Сэм Хелеби
и Денни Мак-Ферсон

Самое полное описание протокола BGP-4

Принципы маршрутизации в Internet

2-е издание

ВИЛЬЯМС

Cisco Systems
Cisco Press

ББК 32.973.26-018.2.75
Х36 УДК 681.3.07

Издательский дом "Вильяме"
Зав, редакцией *С.И. Тригуб* Перевод с английского и редакция *В.В. Ткаченко*
По общим вопросам обращайтесь в Издательский дом "Вильяме" по адресу: info@williamspublishing.com, http://www.williamspublishing.com
Х36 Хелеби, Сэм, Мак-Ферсон, Денни.

Принципы маршрутизации в Internet, 2-е издание. : Пер. с англ. М. : Издательский дом "Вильяме", 2001. — 448 с. : ил. — Парал. тит. англ.
ISBN 5-8459-0188-X (рус.)

Эта книга поможет вам стать экспертом по вопросам интеграции локальной сети вашей компании в глобальную сеть Internet. В ней рассмотрены системы адресации, маршрутизации и способы организации соединений в сети Internet на практических примерах. В книге также освещается широкий круг теоретических понятий и принципов организации маршрутизации в сети Internet с использованием протокола граничного шлюза (Border Gateway Protocol — BGP). После знакомства с этой книгой вы будете способны самостоятельно строить свою сеть и планировать ее развитие. Независимо от того являетесь ли вы потребителем услуг Internet или же вы сами предоставляете услуги этой сети (т.е. вы являетесь сервис-провайдером Internet), в этой книге вы найдете ответы на любые вопросы, касающиеся маршрутизации в сетях на основе протокола TCP/IP.
ББК 32.973.26-018.2.75

ISBN 5-8459-0188-X (рус.)

ISBN 1-57S70-23VX (англ.)

Об авторах

Сэм Хелеби (Sam Halabi) — один из наиболее известных экспертов среди провайдеров Internet. Мистер Хелеби недавно назначен вице-президентом компании Marketing at an IP networking startup, а до этого на протяжении многих лет возглавлял отдел IP-маркетинга в компании Cisco Systems. Он также является крупным специалистом по сложным протоколам маршрутизации и разработке больших IP-сетей.

Сэм Хелеби — активный член технологических форумов Optical Internetworking Forum и MPLS Forum.

Денни Мак-Ферсон (Danny McPherson) в настоящее время является главным инженером компании Amber Networks. До этого он занимал руководящие посты в технических службах компаний — сервис-провайдеров Internet (Qwest, GTE Internetworking, Genuity и internetMCI). В его обязанности входили разработка топологии сетей, организация маршрутизации в них, координация работы различных подразделений и другие вопросы организации сетей. Он — член комитета Internet Engineering Task Force (IETF), т.е. принимает непосредственное участие в разработке стандартов для сети Internet. Денни Мак-Ферсон — признанный эксперт по архитектуре сети Internet и протоколам маршрутизации.

Технические консультанты

Алексей Руднев (Alexei Roudnev) в настоящее время занимает должность системного инженера в консорциуме Genesys Labs/Alcatel в Сан-Франциско, штат Калифорния. На протяжении 10 лет работал инженером по обслуживанию вычислительной сети компании "Релком" - одной из компаний-основательниц сети Internet на территории России. Кроме того, в Москве он около 9 лет занимался разработкой приложений для ОС UNIX.

Абха Ахуджа (Abha Ahuja) -- главный инженер в компании Internap Network Services, занимается проектированием топологии вычислительных сетей и их техническим обслуживанием. Прежде она работала в Merit Network — ведущем исследовательском центре по созданию и учету национальных (для США) точек обмена IP-трафиком -в качестве руководителя проектов разработки сервера маршрутов нового поколения (Route Server Next Generation project) и измерения и учета трафика в сети Internet (Internet Performance Measurement and Analysis project — IPMA). Сегодня она принимает активное участие в жизни сетевого сообщества, продолжая исследовать функционирование протоколов маршрутизации в сетях с различной архитектурой, маршрутизацию между разными доменами, занимаясь сбором сетевой статистики и вопросами сетевой безопасности. Она — сертифицированный специалист по разработке вычислительных сетей.

Посвящения

Денни Мак-Ферсон: *моей супруге Хезер (Heather) и моим дочерям Кортни (Courtney) и Эшли (Ashli). Вы — моя инфраструктура.*

Благодарности

Прежде всего хотелось бы выразить огромную признательность Абхе Ахуджа (Abha Ahuja), Шейну Аманте (Shane Amante), Джонсону Лиу (Johnson Liu), Альваро Ретана (Alvaro Retana) и Алексею Рудневу (Alexei Roudnev) за их ценные консультации, полученные нами при создании этой книги. Следует особо отметить вклад Хенка Смита (Henk Smit), Брюса Коула (Bruce Cole), Энке Чен (Enke Chen), Срихари Рамачандры (Srihari Ramachandra), Рекса Фернандо (Rex Phernando), Сатиндера Сингха (Satinder Singh) и Рави Чандра (Ravi Chandra), а также сотрудников подразделения Cisco "BGP Coders" и всех остальных, внесших посильный вклад во 2-е издание книги. Выражаем благодарность за поддержку и терпение компаниям Amber Networks и Qwest Communications, в которых работал Денни Мак-Ферсон. И наконец, мы хотим поблагодарить Кристофера Кливленда (Christopher Cleveland), Трейси

Хьюджес (Tracy Hughes), Марка Фоулера (Marc Fowler), Гейла Джонсона (Gayle Johnson) и остальных сотрудников издательства Cisco Press за оказанную нам помощь при подготовке книги к печати.

Содержание

Введение

Проект создания вычислительной сети Internet, начинавшийся как академический эксперимент в конце 60-х годов XX века, сегодня можно назвать удачным и констатировать, что он уже давно перерос рамки эксперимента и стал неотъемлемой частью мирового информационного пространства. От сети ARPANET к сети NSFnet, а затем к сети ANYBODYSNET (игра слов: "сеть для каждого" — *Прим. ред.).* Таким образом, сегодня сеть Internet никому не принадлежит, точнее, она принадлежит любому, кто может оплатить адресное пространство в ней. Сегодня десятки миллионов пользователей подключаются к сети Internet и десятки тысяч компаний уже не могут обходиться без ее услуг. Администраторы и разработчики сетей вынуждены учитывать при проектировании сетей новые требования. В настоящее время понимание структуры сети и, в частности, маршрутизации в ней является необходимым условием создания сетей.

Некоторые пользователи удивляются при выходе сети из строя, другие же, наоборот, удивляются, когда сеть работает исправно. Информация о маршрутизации в сетях TCP/IP, доступная проектировщикам и администраторам сетей, по большей части не соответствует их нуждам. Воспользовавшись такой информацией, вы думаете, что обладаете всем необходимым для построения собственной сети. Однако, приступив к реализации своих планов, вы понимаете, что это далеко не так. В первом издании этой книги простым языком, не допускающим двоякого толкования, были описаны принципы маршрутизации, опробованные на практике в реальных сетях.

Кроме новых сведений, во втором издании предлагаются также недавно принятые дополнения к протоколу BGP, обсуждаются вопросы регистрации адресов в сети Internet, а также приводится справочная информация об исследовательских и образовательных вычислительных сетях.

Цели

Цель авторов книги — сделать из вас, дорогие читатели, специалистов по интеграции локальной сети вашей компании в глобальную сеть Internet. Благодаря рассмотрению концепций системы адресации, маршрутизации и структуры соединений в Internet:, а также практических примеров, книга формирует у читателей понятие основных принципов маршрутизации в сети Internet. Ознакомившись с книгой, вы сможете, опираясь на рассмотренные в ней общие принципы, строить свою сеть и планировать ее развитие. Независимо от того, являетесь ли вы потребителем услуг Internet или сами предоставляете сетевые услуги (т.е. являетесь сервис-провайдером Internet), в книге вы найдете ответы на любые вопросы, касающиеся маршрутизации в сетях на основе протокола TCP/IP.

Кому адресована книга

Книга незаменима для организаций, которые планируют подключаться к сети Internet. Будет ли ваша компания сервис-провайдером Internet или вы хотите подключиться к подобной компании, здесь вы найдете все, что нужно для подключения локальной сети к Internet. В книге также рассматриваются перспективы развития сетевой архитектуры и протоколов, которые могут быть интересны сетевым администраторам, сетевым интеграторам и проектировщикам. Хотя книга адресована в основном специалистам по созданию компьютерных сетей, она построена таким образом, что происходит постепенный

переход от простых понятий к более сложным концепциям и проблемам. Для закрепления изложенного материала приводятся практические примеры, которые вы можете использовать в своих целях. Для понимания излагаемого в книге материала не требуется опыт организации маршрутизации в сетях TCP/IP. Следует лишь иметь общее представление о "природе" маршрутизации, чтобы понять предлагаемые концепции.

Как организована эта книга

Книга состоит из четырех частей.

- Часть I. Развитие сети Internet (В главах 1—3 рассматриваются основные аспекты развития и современного состояния сети Internet. В них даются сведения о структуре, компаниях сервис-провайдерах и о системе адресации, принятой в сети Internet. Даже если вы хорошо знакомы со структурой сети Internet, советуем ознакомиться с разделами главы 1, касающимися точек доступа к сети (Network Access Points), проекта арбитража маршрутизации в Internet (Router Arbiter project) и сетевых информационных служб (Network Information Services). Здесь же рассмотрены связанные с маршрутизацией проблемы, возникавшие у администраторов на различных этапах развития сети Internet. В главе 2 дается набор критериев, которыми можно воспользоваться при выборе провайдера Internet. Если же, вы являетесь провайдером или уже подключились к сети Internet через одного из них, то, вероятно, многие вопросы, рассмотренные в этой главе, покажутся вам знакомыми. В главе 3 обсуждаются вопросы бесклассовой междоменной маршрутизации (Classless InterDomain Routing - CIDR.), реализации масок подсетей переменной длины (variable-length subnet masks— VLSM), внедрения протокола IPv6 и другие аспекты адресации в сетях TCP/IP (в частности в сети Internet).
- Часть II. Основы протоколов маршрутизации. В главах 5 и 6 раскрываются основные принципы маршрутизации: свойства и особенности протоколов маршрутизации, основанных на анализе состояния канала связи, и векторных протоколов маршрутизации; рассматриваются вопросы применения междоменных протоколов и принципы их работы. Подробно описывается протокол граничных шлюзов (Border Gateway Protocol - BGP), который на сегодняшний день является стандартным протоколом маршрутизации для сети Internet. В этой части подробно рассматриваются все параметры и свойства протокола BGP.
- Часть III. Эффективные схемы маршрутизации в сетях TCP/IP. В главах 6—10 рассмотрены различные схемы применения протокола BGP. В этой части в действии показаны параметры протокола BGP, описанные в части 2. Все они реализуют различные свойства протокола BGP — резервирование, распределение нагрузки и симметричность. Здесь же рассмотрены проблемы, возникающие при организации внутридоменной и междоменной маршрутизации в крупных и развивающихся сетях, и определяются пути их решения.
- Часть IV. Настройка маршрутизаторов для работы в сетях TCP/IP. Главы 11 и 12 содержат огромное количество листингов реальных конфигураций протокола BOP для разных схем маршрутизации. Эти листинги принесут наибольшую пользу, если вы внимательно изучите предыдущие главы, так как в них описываются определенные концепции и технологии. Таким образом, вы сможете сопоставить обсуждаемые в первых главах идеи с листингами, приведенными в главах 11 и 12. В тексте книги на них ссылаются как на "Примеры настройки". Встретив такую ссылку, вы сможете перейти к указанной странице и просмотреть пример настройки обсуждаемых параметров.

Кроме того, в книге имеется несколько приложений, где представлены дополнительные ссылки на информацию по теме, описание современных команд операционной системы маршрутизаторов Cisco IOS™ для настройки протокола BGP и сведения о модификациях ОС IOS™, с помощью которой обеспечивается интерфейс

командной строки для настройки параметров BGP.

Предисловие

Описывать технические сведения в доступной читателю манере довольно сложно. В рассматриваемой теме кроется так много технических подробностей, опустив которые мы не смогли бы описать те или иные концепции. С другой стороны, перенасыщение книги техническими терминами затрудняет восприятие материала. Мы старались вводить технические термины постепенно и по возможности иллюстрировать их примерами практических настроек оборудования. Самое сложное, на наш взгляд, — описание настроек маршрутизаторов с помощью команд Cisco IOS, поэтому мы оставили этот материал на последние две главы. К тому же все эти настройки опираются на идеи и канонические схемы, представленные в начале книги.

Хотя вашей конечной целью является разработка и внедрение собственных стратегий маршрутизации, все-таки необходимо усвоить основные принципы и концепции. Эта книга сочетает в себе как теоретические концепции, так и практические рекомендации по настройке и выбору стратегии маршрутизации для сети. Читатель при чтении постепенно движется от общих положений к частным случаям и от рассмотрения идей к их реализации. Вопросы, представляющие отдельный интерес, переадресованы с помощью ссылок на примеры настройки, выведены в списки ЧАсто задаваемых ВОпросов (ЧАВО) или проиллюстрированы на примерах.

Подход к созданию книги с использованием сценариев настройки очень важен. Таким образом, рассмотренные топологии, показываются в действии, иллюстрируются все аспекты применения различных протоколов и стратегий маршрутизации. Даже если вы не встретите топологию сети, точно соответствующую вашей, примеры сценариев, приведенных в книге, будут способствовать накоплению опыта, и вы сможете самостоятельно применить полученные знания в определенной ситуации.

Соглашения, принятые в этой книге

В этой книге мы не отказались от описания деталей функционирования протоколов маршрутизации и сведений о разработке сетей на их основе. В то же время основное внимание уделяется построению теоретической базы и разъяснению концепций. Чтобы выделить те или иные идеи, в книге используются:

- указатели на примеры настройки — находятся, как правило, возле текста и указывают на соответствующие листинги в главах 11 и 12;
- сборники часто задаваемых вопросов (ЧАВО) — имеются обычно в конце каждой главы.

Знаки, встречающиеся в командах

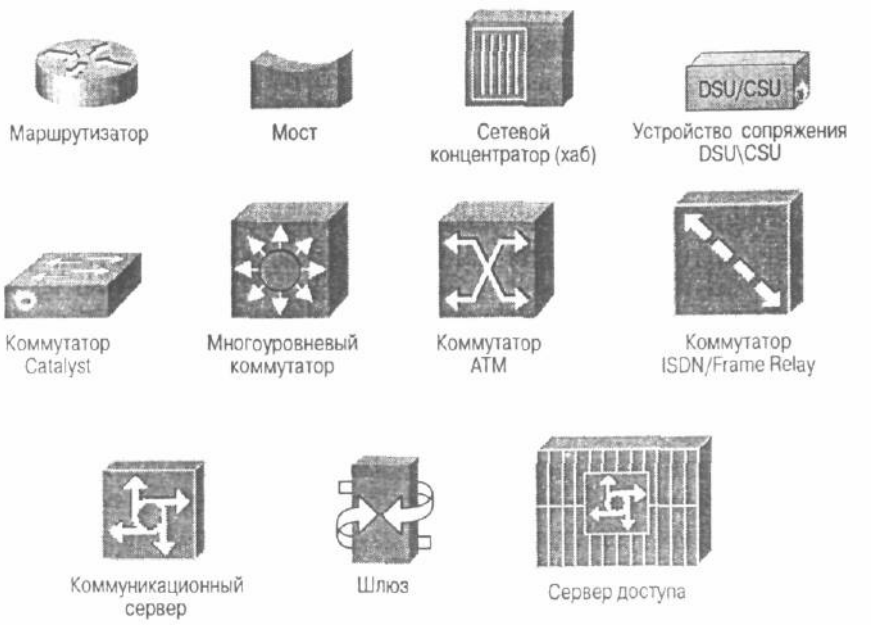
В книге полностью сохранен синтаксис команд Cisco IOS. В описании команд Cisco IOS встречаются следующие обозначения.

- Вертикальные скобки (I) — отдельные взаимоисключающие элементы.
- Квадратные скобки ([ ]) — указывают на необязательные элементы в команде.
- Фигурные скобки ({ }) — указывают на обязательный параметр в команде.
- Фигурные скобки, заключенные в квадратные ([{}]), — обязательный аргумент или

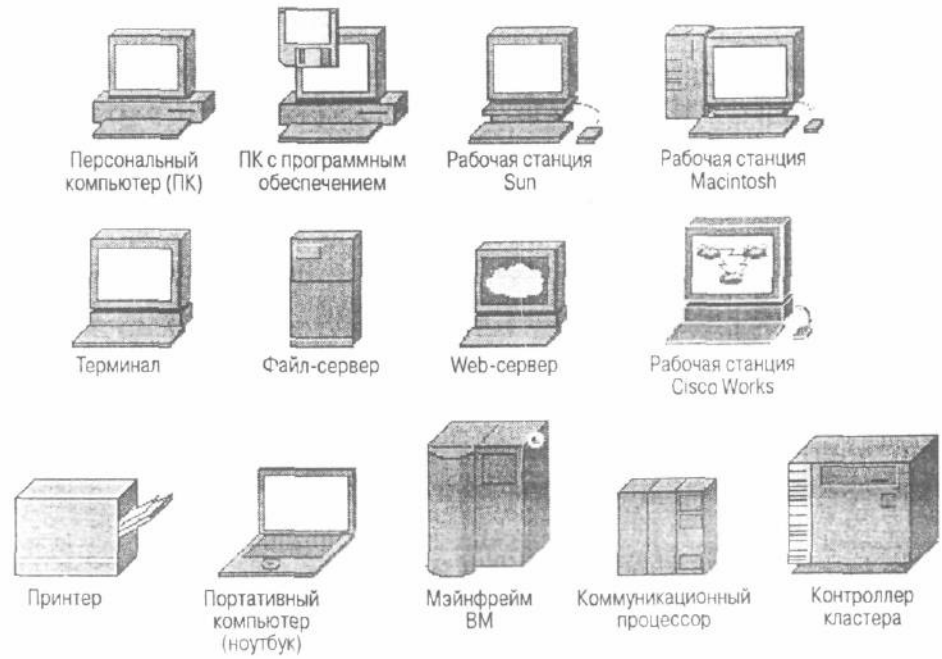
параметр с необязательными дополнительными элементами.

- Полужирный шрифт указывает на то, что команды и ключевые слова должны вводиться именно в представленном виде. В листингах и при перепечатке результатов выполнения команд (что, по сути, не является синтаксисом команд) полужирным шрифтом выделены команды, которые вводятся пользователем вручную (например, команда show).
- *Курсивом* выделяются аргументы, которым назначаются определенные значения.

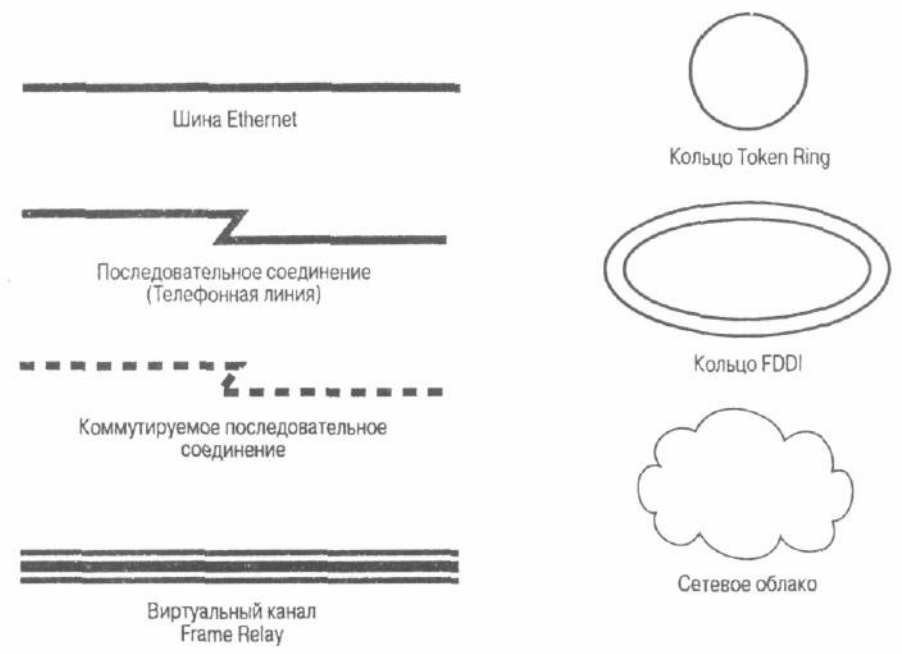
Обозначения, используемые в книге



Для обозначения периферийных и других устройств в книге используются следующие обозначения:



Для обозначения различных типов сетевых соединений в книге будут встречаться следующие обозначения:


Шина Ethernet
Последовательное соединение
(Телефонная линия)
Коммутируемое последовательное
соединение
Виртуальный канал
Frame Relay
Кольцо Token Ring
Кольцо FDDI
Сетевое облако

Часть I.
Развитие сети Internet

В этой части...

Глава 1. Эволюция сети Internet

Глава 2. Услуги, предоставляемые провайдерами Internet и их характеристики

Глава 3. IP-адресация и методы распределения адресов

Сложность проблем при организации маршрутизации в сетях-TCP/IP и их решение тесно связаны с эволюцией и современным состоянием Internet. Таким образом, прежде чем перейти к специфическим вопросам применения протоколов маршрутизации, желательно ознакомиться с эволюцией сети Internet и перспективами ее развития. Поэтому в главе 1 рассмотрены такие исторически важные проекты, как проект арбитража маршрутизации (Router Arbiter project), проект точек доступа к сети (Network Access Point) и сетевых информационных служб (Network Information Services). Эти сведения будут весьма полезны системным администраторам организаций, планирующим в ближайшем будущем подключение к глобальным сетям или развивающим инфраструктуру существующих сетей. В главе 2 представлены основные принципы создания сетей и межсетевых соединений,- осуществляемых провайдерами Internet. В главе 3 рассмотрены основные концепции адресации в сетях TCP/IP, и в частности Internet, а также принципы бесклассовой междоменной маршрутизации, используемой для контроля за распределением пространства IP-адресов.

Ключевые темы этой главы:

- **Истоки и история развития сети Internet.** Рассматривается краткая история создания и развития Сети, первые разработчики и пользователи Internet, приводятся сведения о росте Internet за последние 40 лет. Дается также обзор нескольких докладов Национального научного фонда США (National Science Foundation – NSF)
- **Точки доступа к сети (Network Access Points — NAP).** Способы объединения сетей провайдеров Internet (напрямую или опосредованно, с использованием точек доступа к сети (NAP)). Вам необходимо получить достаточно знаний, чтобы иметь представление, каким образом провайдеры Internet подключаются к NAP,а также где и какие NAP имеются сегодня в мире.
- **Прямые межсетевые соединения** — как альтернатива для точек доступа к сети (NAP) эти соединения за последние несколько лет завоевали огромную популярность среди крупных провайдеров Internet. (В основном по причине того, что такая модель объединения систем устраняет некоторые недостатки модели общих NAP.).
- **Проект арбитража маршрутизации в сети (Routing Arbiter project).** Дается обзор основных концепций, на которых далее строится рассмотрение всего материала книги: принципы работы серверов маршрутов и базы данных арбитража маршрутов (Routing Arbiter Database). Серверы маршрутов являются основными компонентами NAP, сетей провайдеров Internet и других сетей на базе TCP/IP.
- **Региональные провайдеры**. Рассмотрены сложившиеся на сегодняшний день в США схемы соединений провайдерских систем по регионам.
- **Информационные службы**. Приводится обзор информационных служб и агентств, которые сформировались в результате реализации программ Национального научного фонда США (National Science Foundation — NSF) и приватизации сети Internet: сетевая служба InterNlC, службы регистрации, службы каталогов и баз данных, службы сетевой поддержки NIC. Рассмотрена также эволюция других служб регистрации в сети Internet и реестров межсетевых маршрутов (Internetworking Routing Registies).
- **Настоящее и будущее Internet.** Делаются попытки прогнозирования будущего, сети Internet: проекты "Инициатива следующего поколения" (The Next-Generation Initiative), Internet2 и Abilene.

Глава 1.
Эволюция сети Internet

Структура и внешний вид сети Internet изменились вместе с изменениями потребностей общества. Сегодня Internet обслуживает наибольшее разнородное сетевое сообщество в мире. В этой главе делается краткий хронологический обзор развития различных компонентов сети Internet На примере сети Internet проводится анализ этапов развития при построении масштабируемых глобальных сетей.

История развития сети Internet

Сеть Internet началась как эксперимент в конце 60-х годов XX века, проводившийся агентством по передовым исследованиям (Advanced Research Projects Agency — ARPA, позднее переименованное в DARPA), которое находилось в ведении Министерства обороны США (Department of Defense). Агентство DARPA проводило комплексные исследования работы компьютеров, объединенных в сеть, с предоставлением грантов на конкурсной основе нескольким университетам и частным компаниям, вовлекая их таким образом в исследовательский процесс.

В декабре 1969 года была сдана в эксплуатацию экспериментальная сеть, объединявшая четыре узла со скоростью передачи данных 56 Кбит/с. Новая технология, примененная при построении этой сети, доказала свою жизнеспособность и легла в основу создания еще двух военных вычислительных сетей: MILNET — на территории США и MINET — в Европе. Впоследствии к сети ARPANET были подключены тысячи хостов и отдельных сетей (главным образом университетских и государственных), что привело к созданию так называемой "ARPA Internet" — прародительницы современной сети Internet. На рис. 1.1 и 1.2 показана структура сети ARPANET на ее ранних стадиях развития (с момента создания в 1969 году и до 1976 года).

Конгломерат исследовательских, образовательных и государственных сетей, объединенных ядром сети ARPANET, положил начало сети, которая сегодня известна под названием Internet. Однако в сети ARPANET существовал набор правил для пользователей, обязательный для всех (так называемая Acceptable Usage Policy — AUP). Согласно этому своду правил, запрещалось использование сети ARPANET в коммерческих целях. К тому же у пользователей ARPANET стали возникать проблемы при расширении сетей, наиболее очевидными из которых являлись перегрузки на линиях связи. Поэтому Национальный научный фонд (National Science Foundation — NSF) приступил к созданию сети NSFNET2.

Эксплуатация сети ARPANET была полностью прекращена в 1989 году.

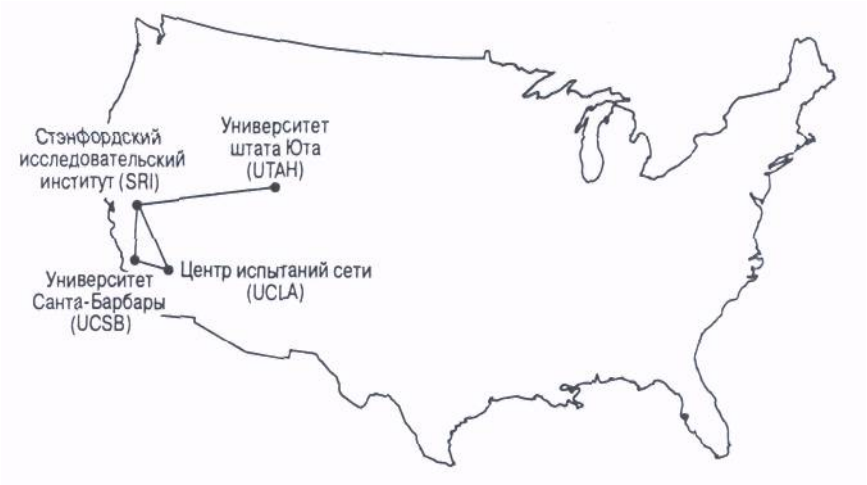


Рис. 1.1. *Структура сети ARPANET (декабрь 1969 г.)*

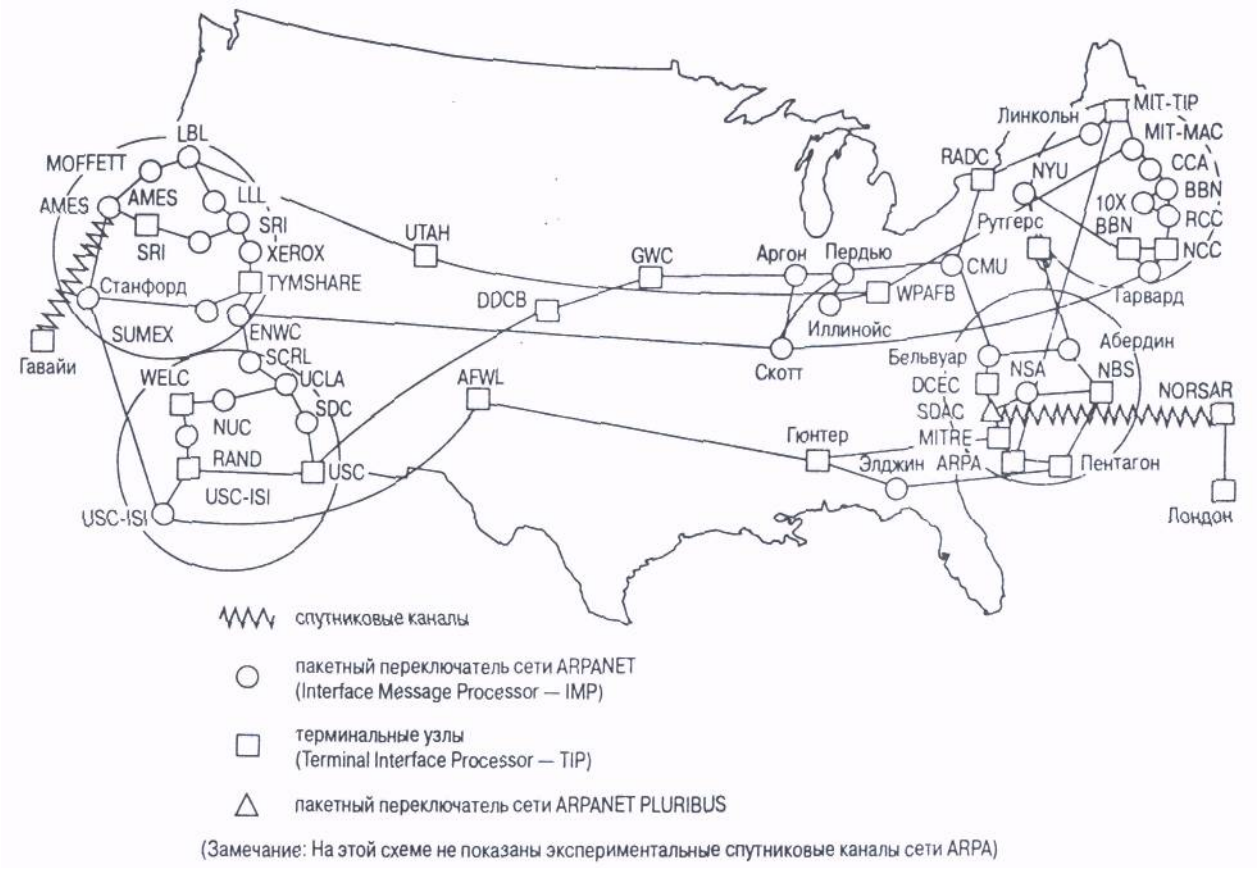


Рис. 1.2*. Структура сети ARPANET (июль 1976г.)*

От ARPANET к NSFNET

Уже к 1985 году сеть ARPANET разрослась до огромных размеров, и администраторы столкнулись с проблемой перегруженности сети. Чтобы исправить ситуацию, NSF инициировал развертывание сети NSFNET. Сеть NSFNET представляла собой объединение нескольких региональных компьютерных сетей и сетей правительственных организаций (таких как NASA Science Network), которые подключались к опорной сети, составляющей ядро всей NSFNET.

Первоначально (1986 год) в NSFNET была создана трехуровневая сетевая

архитектура. Она предполагала подключение университетских городков и исследовательских центров к региональным сетям, которые, в свою очередь, подключались к опорной сети, поддерживаемой шестью общенациональными суперкомпьютерными центрами. В 1988 году существующие каналы передачи данных, работа по которым велась со скоростью 56 Кбит/с, были расширены до более быстрых каналов типа T1 (1,544 Мбит/с). Все эти изменения стали возможными благодаря тендеру на модернизацию и обслуживание сети, объявленному NSF, в котором победила корпорация Merit Network, Inc. И ее партнеры — компании MCI, IBM и администрация штата Мичиган. Таким образом, опорная сеть NSFNET на базе каналов передачи данных T1 объединила 13 узлов по всем Соединенным Штатам Америки — Merit, BARNET, MidNet, Westnet, NorthWestNet, SESQUINET, SURAnet, узел Национального центра исследований атмосферы (National Center for Atmospheric Research — NCAR) и шесть суперкомпъютерных центров самого NSF.

В 1990 году, объединив свои усилия в области обслуживания национальной вычислительной сети, компании Merit3 , IBM и MCI создали организацию по развитию сети и услуг под названием Advanced Network and Services (ANS). Группа инженеров компании Merit разработала базу данных, в которой хранилась информация о маршрутизации, консультировала и осуществляла управление маршрутами в сети NSFNET, а ANS отвечала за работу магистральных маршрутизаторов и управляла сетевым операционным центром (Network Operation Center — NOC).

К 1991 году трафик сети возрос настолько, что понадобилось срочно расширять пропускную способность магистральных каналов опорной сети до каналов T3 (45 Мбит/с). На рис. 1.3 приведена структура сети NSFNET с учетом расположения ее ядра и региональных магистралей.

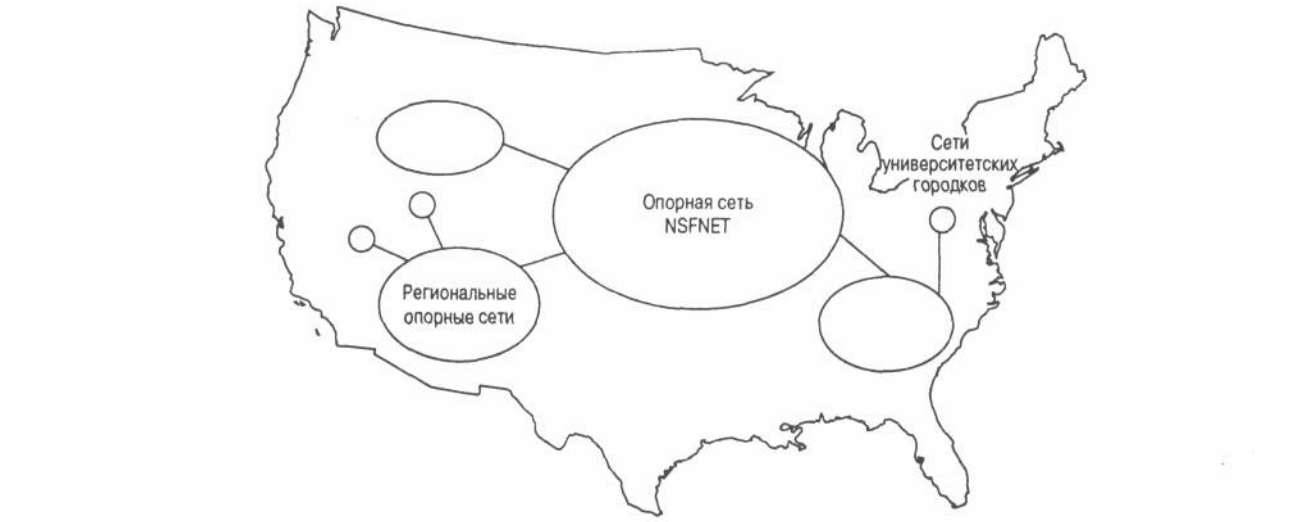


Рис. 1.3. *Структура сети NSFNET*

В начале 90-х годов сеть NSFNET по-прежнему использовалась в исследовательских и образовательных целях. Все магистральные каналы передачи данных были зарезервированы правительственными агентствами для стратегических целей. Вместе с тем ощущалась острая потребность в подключении все большего числа организаций (как правительственных, так и коммерческих). Коммерческие и общенациональные интересы в наращивании сети совпадали, и Internet-сервис провайдеры (Internet Service Provider — ISP) вынуждены были срочно объединить свои интересы, что привело к формированию абсолютно новой индустрии. Компьютерные сети вне США развивались по мере развития международной сети каналов передачи данных. Наравне с создаваемыми и существующими объектами в рамках государств и регионов росла инфраструктура всемирной глобальной сети.

В США сети правительственных агентств объединялись между собой с помощью федеральных точек обмена трафиком (Federal Internet eXchange points — FIX), соединявших Восток и Запад страны. Сети коммерческих организаций сформировали собственные точки обмена трафиком (Commercial Internet eXchange points — CIX), которые объединили весь Запад США. В то же время провайдеры Internet по всему миру, в частности в Европе и Азии,

разработали и подготовили инфраструктуру мировой глобальной сети, объединив магистральные каналы передачи данных между собой.

Чтобы как-то упорядочить рост сети, NSF назначил компанию Sprint менеджером международных соединений (International Connection Manager — ICM), в функции которой входило обеспечение соединений между компьютерными сетями США, Европы и Азии. В апреле 1995 было объявлено о прекращении существования сети NSFNET.

Internet сегодня

Деятельность NSFNET должна была прекращаться постепенно, в несколько этапов, с сохранением существующих соединений между различными государственными институтами и агентствами. Инфраструктура сети Internet сегодня — уход от концентрированного ядра сети (как это было в NSFNET) к более распределенной архитектуре, которая управляется в основном коммерческими провайдерами, такими как UUNET, Qwest, Sprint и тысячи других. Все они соединяются друг с другом через точки обмена трафиком либо напрямую. На рис. 1.4 представлена современная структура сети Internet.

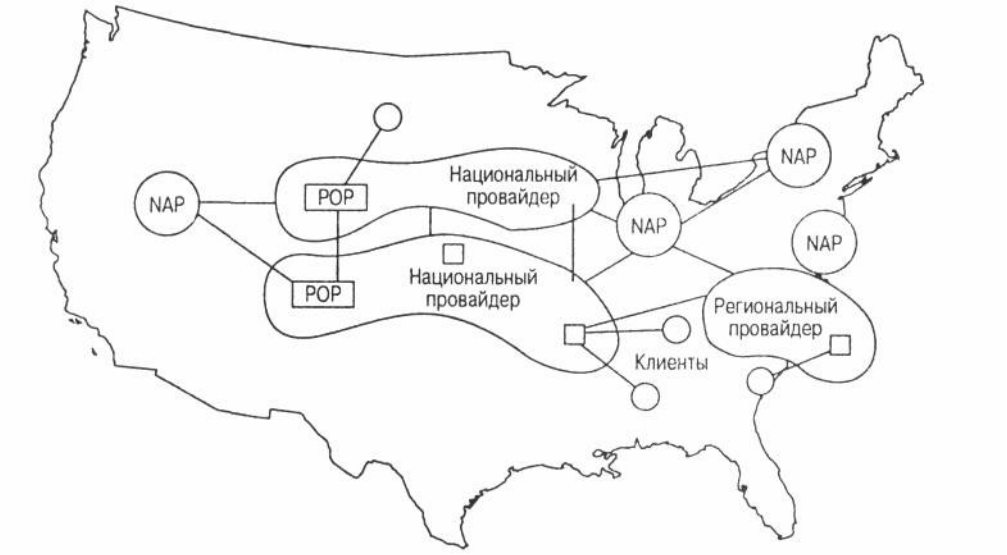


Рис. 1.4. Структура сети Internet

Современная опорная сеть Internet представляет собой объединение провайдеров Internet, у которых в нескольких регионах имеются узлы, называемые *точками присутствия* (Points Of Presence — POP). Это — объединение точек присутствия и совокупность каналов, соединяющих их между собой. Все подключения клиентов к провайдерам Internet осуществляются именно через серверы доступа или другие средства хостинга, расположенные на POP провайдера. Иногда клиенты сами могут быть одновременно провайдерами Internet (их называют субпровайдерами — *Прим. ред.)*. Чаще всего встречающиеся сегодня схемы подключения, используемые провайдерами Internet, обсуждаются в главе 2, "Услуги, предоставляемые провайдерами Internet и их характеристики".

Провайдеры, у которых имеются точки присутствия в большинстве регионов США, считаются *национальными провайдерами (national provider).* Провайдеры, которые обслуживают определенный регион США, называются *региональными провайдерами (regional provider).* Они соединяются друг с другом через одну или несколько точек обмена трафиком. Клиенты, подключенные к одному провайдеру могут общаться по сети с клиентами другого провайдера благодаря точкам доступа к сети (Network Access Points -- NAP) или через непосредственное сетевое соединение между двумя провайдерами. Термин *сервис-провайдер Internet (Internet Service Provider —ISP)* обычно используется применительно к компании, обеспечивающей подключение к сети Internet других провайдеров либо конечных пользователей. Термин *сетевой сервис-провайдер (Network Service Provider—NSP)* традиционно относится лишь к провайдерам, обеспечивающим

подключение к опорной сети передачи данных. Однако сегодня термин NSP все чаще употребляется в связи с провайдером, представленным в NAP и обслуживающем опорную сеть.

Ходатайства NSFNET

Фонд NSF с 1986 года финансировал исследования в области передачи данных к разработки компьютерных сетей. Кроме того, Фондом финансировалась Программа высокопроизводительных вычислений и передачи данных (High Performance Computing and Communications Program — HPCC), которая включала целый ряд передовых научно-исследовательских программ. Одним из проектов в рамках программы развития национальной исследовательской и образовательной сети (National Research and Education Network — NREN) был проект создания к середине 90-х годов сети с пропускной способностью до 1 Гбит/с. Все это, вместе с истечением срока действия в апреле 1995 года Соглашения о сотрудничестве с NSFNET Backbone Network Services, стало причиной появления так называемых ходатайств NSF о сохранении услуг сети NSFNET.

Первое ходатайство от NSF поступило в 1987 году и привело к модернизации опорной сети (магистральные каналы передачи данных *к* концу 1993 г. были заменены на каналы с большей пропускной способностью — T3). В 1992 году NSF собирался подать ходатайство об объединении и усилении роли коммерческих сервис-провайдеров, что создало бы условия для построения универсальной модели сети. В то же время в NSF решили отказаться от обслуживания ядра сети и направили все усилия на разработку новых концепций и продвижение новых технологий. Последнее ходатайство от NSF (NSF 93-52) было издано в мае 1993 года.

В последнее ходатайство вошли четыре отдельных проекта, в которых предполагалось:

- Создать разветвленную сеть точек доступа к сети (NAP), посредством которых провайдеры могли бы осуществлять обмен трафиком.
- Реализовать проект арбитража маршрутизации в сети (Routing Arbiter project — RA), что облегчило бы согласование правил работы в сети и управление адресами между несколькими провайдерами, подключенными к NAP.
- Найти и утвердить единого провайдера для службы обеспечения высокоскоростной магистральной сети (very high-speed Backbone Network Service — vBNS), который бы обеспечивал подключения в интересах государственных и образовательных учреждений.
- Обеспечить транзит трафика по существующим сетям в целях поддержки межрегиональных соединений путем подключения к NSP, которые, в свою очередь, подключены к NAP или напрямую к NAP. Любой избранный для этой цели NSP должен иметь соединения по крайней мере с тремя NAP.

Все ходатайства, поданные правительству США от NSF рассмотрены в этой главе.

Точки доступа к сети

Ходатайство о создании точек доступа к сети (Network Access Points — NAP) было направлено на реализацию предложений, поступивших от некоторых компаний, по учреждению специальных узлов в сети — NAP, где осуществлялись бы межсетевые соединения в рамках vBNS или других сетей. В этих NAP требовалось обеспечить возможность обмена трафиком между региональными сетями, сетями сервис-провайдеров, а также исследовательскими и образовательными центрами по всей Америке. Такого рода точки доступа к сети нужны были для формирования соединений между сетями, которые не

были субъектами «Свода правил применения сети NSF» (NSF Acceptable Usage Policy), ограничившего использование сети Internet только исследовательскими и образовательными целями. Таким образом, использование сети в коммерческих или каких-либо других целях, не соответствующих этому своду правил, автоматически пресекалось на уровне точек доступа к сети.

Что такое NAP?

Согласно терминологии NSF, *точка доступа к сети (NAP)* — это высокоскоростной коммутатор или сеть коммутаторов, к которой подключается определенное число маршрутизаторов для обмена трафиком. Любая NAP должна работать со скоростью не ниже 100 Мбит/с и иметь возможность повышения пропускной способности по запросу или в зависимости от нагрузки. Работа каждой NAP при передаче трафика от одного провайдера другому должна быть так же прозрачна и проста, как работа ATM-коммутатора (45+ Мбит/с) или коммутатора FDDI (100 Мбит/с).

Концепция построения NAP основана на применении точек обмена трафиком FIX и CIX (о них далее, в этом разделе), которые были созданы в свое время вокруг магистральных колец FDDI для подключения сетей на скоростях до 45 Мбит/с.

Трафик через NAP уже не имеет ограничений со стороны Свода правил, т.е. это может быть не только информация в интересах образовательных или исследовательских учреждений. Все сети при подключении к NAP автоматически получают разрешение на обмен трафиком с другими сетями, не нарушая при этом никаких законов и правил.

Фондом NSF были выделены четыре основных NAP:

- Sprint NAP в Пенсаукен (штат Нью-Джерси);
- PacoBell NAP в Сан-Франциско (штат Калифорния);
- Ameritech Advanced Data Services (AADS) NAP в Чикаго (штат Иллинойс);
- MFS Datanet (MAE-East) NAP в Вашингтоне (Федеральный округ Колумбия).

Опорная сеть NSFNET 13 сентября 1994 года была подключена к точке доступа Sprint NAP. В середине октября 1994 года она была подключена к NAP PacoBell, а в начале января 1995 года к NAP Ameritech. И, наконец, 25 марта 1995 года, по предложению MFS (ныне MCI Worldcom), сеть NSFNET была подключена к узлу MAE-East FDDI.

Сети при подключении к NAP должны были иметь пропускную способность, соразмерную с подключенными ранее сетями (т.е. не менее 1,5 Мбит/с), и возможность повышения пропускной способности по запросу, в зависимости от нагрузки или целей применения. Назначенные NSF точки NAP должны были обеспечивать коммутацию не только пакетов протокола IP, но и сетевого протокола без установки соединения CLNP (Connectionless Networking Protocol). Необходимость коммутации CLNP-пакетов и реализации процедур, предусмотренных протоколом междоменной маршрутизации IDRP (Inter-Domain Routing Protocol) и протоколом внешнего шлюза (ISO OSI Exterior Gateway Protocol — EGP), могла быть обусловлена общим уровнем сервиса, предоставляемого на той или иной NAP.

Учреждение должности менеджера NAP

Для каждой NAP назначалось ответственное лицо — менеджер NAP. В его обязанности входило:

- Развертывание и обслуживание NAP при подключении ее к vBNS и другим сетям.
- Установка правил работы и цен на услуги, предоставляемые сервис-провайдерам при подключении к NAP.
- Подание предложений о развертывании NAP в зависимости от географических особенностей местности.
- Подание предложений и утверждение списка стандартных процедур, которые будут

использоваться при взаимодействии с персоналом других NAP, арбитром маршрутов (Routing Arbiter — RA), провайдером vBNS, администраторами региональных и других сетей для решения возникающих проблем и поддержки заданного качества обслуживания (Quality of Service — QoS) для всех сетей и всех пользователей.

- Разработка стандартов по надежности и безопасности работы NAP, а также процедур, с помощью которых можно было бы определить степень соответствия NAP этим стандартам.
- Обеспечение ведения учета трафика в NAP и сбора статистики для последующего анализа состояния NAP.
- Определение соответствующих процедур по ограничению доступа персонала на узлы NAP и контролю за их выполнением.

Федеральные точки обмена трафиком

В переходный период от сети ARPANET к NSFNET были созданы восточная FIX-East (Колледж-Парк, штат Мериленд) и западная FIX-West (NASA AMES, Маунтин Вью, штат Калифорния) федеральные точки обмена трафиком (Federal Internet eXchange — FIX). Вскоре они стали стратегически важными узлами, через которые велся обмен информацией между сетями исследовательских центров, университетов и государственных учреждений. Однако правила работы в них не позволяли вести обмен данными коммерческого характера. Вследствие чего стали появляться коммерческие точки обмена трафиком — Commercial Internet eXchange points (CIX).

В 1996 году была прекращена работа точки FIX-East. Вторая федеральная точка обмена трафиком — FIX-West — до сих пор используется в интересах федеральных служб США.

Коммерческие точки обмена трафиком

Коммерческие точки обмена трафиком (Commercial Internet eXchange — CIX (произносится "кикс" — *Прим. ред.}* — добровольные объединения сервис-провайдеров, способствующие развитию инфраструктуры межсетевых соединений как в масштабах одного государства, так и во всем мире. Создание CIX явилось прямым следствием нежелания операторов федеральных точек обмена трафиком FIX поддерживать сети, которые не являются государственными. Кроме обеспечения межсетевых соединений в интересах коммерческих провайдеров, CIX также стали своего рода форумами, где специалисты могут обмениваться новыми идеями, свежей информацией и проводить эксперименты совместно с другими поставщиками услуг сети. Ниже приведено лишь несколько преимуществ, которыми обладают члены CIX.

- Независимый форум, где можно обсуждать законы и политические вопросы.
- В основу соглашения об использовании CIX положен принцип добровольного объединения сетей. Здесь не существует ограничений по типу трафика, которым члены CIX могут вести обмен друг с другом.
- Доступ ко всем сетям, подключенным к CIX, дает возможность пользоваться всей мощью объединенных сетей: файлами, базами данных и другими информационными сервисами. Таким образом пользователи получают доступ к богатейшему глобальному ресурсу, что повышает ценность сетевого соединения.

Несмотря на то что сегодня CIX в сети Internet играют менее важную роль в объединении сетей, чем NAP, они все же остаются важным звеном в обеспечении физического соединения различных вычислительных сетей, в координации законотворчества для глобальной сети и при разработке правил объединения сетей.

Дополнительную информацию о CIX можно получить на Web-сервере www.cix.org.

Физическая структура NAP

Физическая структура сегодняшних NAP представляет собой совокупность FDDI-, ATM- и Ethernet-коммутаторов (к ним относятся Ethernet, Fast Ethernet и Gigabit Ethernet). В NAP применяется весь спектр методов доступа от FDDI и Gigabit Ethernet до DS3, OC3 и ОСИ ATM. На рис. 1.5 представлена типовая структура современной NAP. Как правило, провайдерами Internet обслуживаются и маршрутизаторы, которые обычно устанавливаются на технических площадках NAP. В обязанности менеджера NAP входит определение конфигурации, правил работы и цен на услуги, предоставляемые NAP.

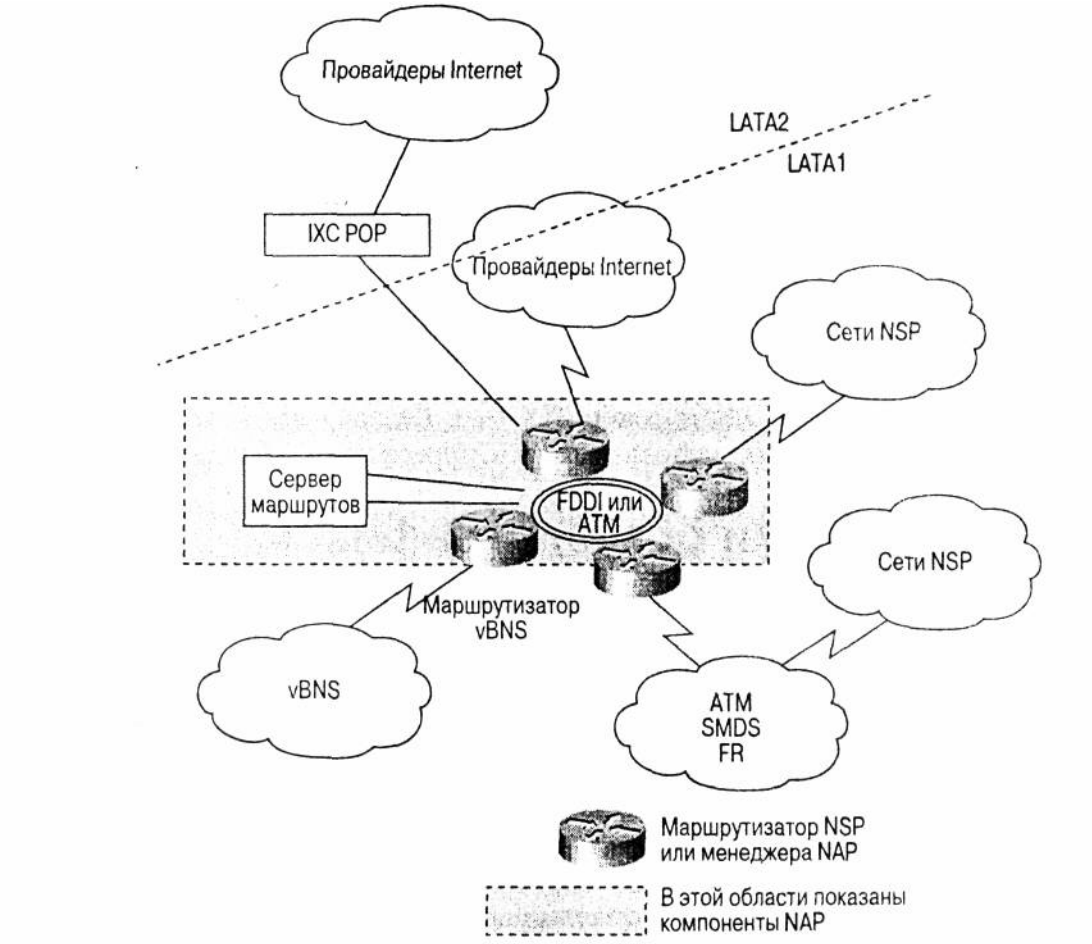


Рис. 1.5. Типовая структура NAP

Альтернатива NAP — прямые межсетевые соединения

С ростом сети Internet стремительно растет и трафик, проходящий через NAP. Размеры трафика сегодня уже таковы, что не все NAP могут обеспечить его нормальное прохождение. Ограниченная пропускная способность некоторых NAP приводит к потерям данных и нестабильной работе сети. Кроме того, крупные частные компьютерные сети и провайдеры Internet иногда неохотно взаимодействуют с третьими лицами, в роли которых выступают менеджеры NAP, при решении вопросов, касающихся обслуживания межсетевых соединений и создании резерва пропускной способности. По этим причинам за последние несколько лет среди провайдеров Internet большую популярность приобрели *прямые межсетевые соединения, с* помощью которых объединяются сети провайдеров. Такие соединения являются альтернативой NAP.

Смысл прямых межсетевых соединений достаточно прозрачен. Обеспечивая соединение сетей напрямую и одновременно избегая NAP, провайдеры Internet могут уменьшить их значимость для сети в целом и увеличить при этом надежность сетевых

соединений со значительным повышением уровня масштабируемости узлов сети. Ширина полосы пропускания и местоположение прямых соединений обсуждаются в двустороннем порядке на взаимовыгодных условиях. К тому же прямые межсетевые соединения, как правило, поддерживаются между двумя сетями пока одна или обе стороны с экономической точки зрения заинтересованы в их работе.

Однако следует помнить о том, что прямые межсетевые соединения обеспечивают не только дополнительную полосу пропускания между двумя сетями. Они также смягчают нагрузку на NAP — освобождают на них полосу пропускания, увеличивая таким образом пропускную способность и производительность последних. Также в связи с тем, что ведущие компании на рынке обладают, как правило, крупными сетями с похожей топологией, это сходство топологий и требований к узлам сети позволяет с помощью прямых межсетевых соединений производить более оптимальное географически распределение потоков данных, чем в NAP. Прямые соединения обеспечивают оптимальный обмен трафиком между региональными сетями, увеличивая вместе с тем пропускную способность сети и снижая время задержки для заданного набора хостов.

Более мелкие провайдеры или провайдеры-новички скорей всего не будут спешить с установкой прямых межсетевых соединений с крупными провайдерами по следующим причинам.

- Высокие затраты со стороны провайдеров на поддержку существующей инфраструктуры прямых соединений.
- Рост затрат, связанных с обслуживанием служебных помещений для локальных и межузловых пунктов обмена трафиком — LEG (local exchange carriers) и IXC (interexchange carriers).

К счастью, большинство крупных провайдеров находят возможность поддерживать работу NAP по обмену трафиком между сетями, в которых пока нет возможности реализовать прямые соединения по экономическим причинам.

Проект арбитража маршрутизации

Еще один проект инициированный Национальным научным фондом NSF — проект арбитража маршрутизации (Routing Arbiter — RA project), который был направлен на обеспечение равноправного обслуживания провайдеров со стороны органов администрирования и управления маршрутизацией. Провайдеры, участвующие в проекте RA, создали общую базу данных маршрутной информации в целях повышения стабильности и управляемости сетей, объединенных в Internet.

Множество провайдеров, подключенных к NAP, повышало масштабируемость сети, так как каждый провайдер должен был обеспечить межсетевые соединения с другими провайдерами для обмена маршрутной информацией и информацией о правилах работы в сети. Проект RA был призван уменьшить требования к организации соединений между провайдерами. Планировалось отойти от схемы соединений типа "каждый с каждым". Вместо этого провайдеры должны были обеспечить лишь соединения с определенной центральной системой, которая называлась *сервером маршрутов (Route Server).* Сервер маршрутов был призван обслуживать базу данных, в которой содержалась необходимая для провайдеров информация о маршрутах и правилах их применения. На рис. 1.6 представлена схема физических соединений и логического взаимодействия между провайдерами и сервером маршрутов. На проект RA возлагались следующие задачи.

- Укрепление стабильности и повышение управляемости маршрутами в сеть Internet. Сервер маршрутов выполняет эти задачи путем уменьшения количеств; узлов BGP, требуемых для соблюдения правил маршрутизации, перед передачей маршрутной информации следующему узлу. Таким образом, путем сокращения объемов информации о маршрутах уменьшается нагрузка на маршрутизаторы.
- Формирование и обслуживание баз данных топологий сети путем обмена маршрутной информацией и ее обновления посредством подключенных к сети автономных систем (Automomous System — AS) с помощью одного из

стандартных протоколов внешнего шлюза Exterior Gateway Protocol (EGP), например таких, как протокол граничного шлюза Border Gateway Protocol (BGP) и IDRP (поддержка IP и CLNP).

- Подача предложений и определение процедур по взаимодействию технического персонала при решении технических проблем, начиная от менеджеров NAP, дежурных операторов провайдеров vBNS, до администраторов региональных и других сетей включительно. Обеспечение единого качества обслуживания (Quality of Service — QoS) во всей сети и устойчивости всех существующих соединений.

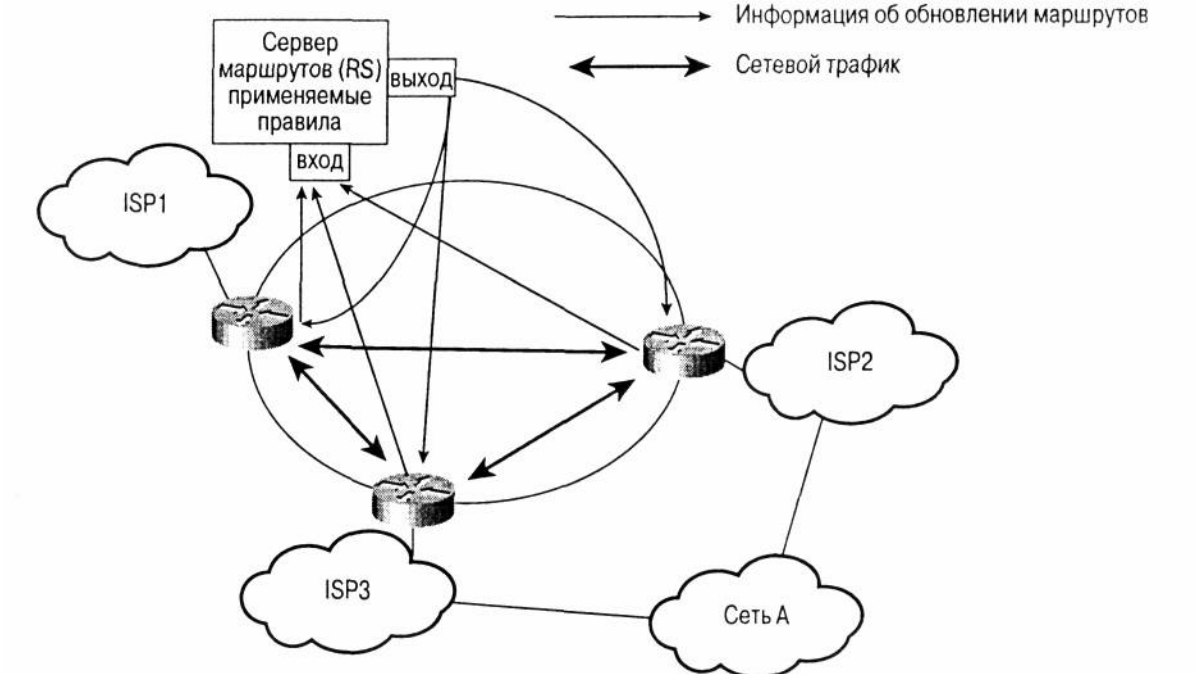


Рис. 1.6. *Обновление маршрутов с помощью сервера маршрутов*

- Разработка новых технологий в обеспечении маршрутизации, т.е. введение таких критериев, как тип сервиса (type of service) и маршрутизация по старшинству, многоадресная передача, выделение полосы пропускания по запросу и формирование службы распределения полосы пропускания совместно со всем сетевым сообществом Internet.
- Внедрение упрощенных стратегий маршрутизации, таких как маршрутизация подключенных сетей по умолчанию.
- Внедрение распределенного управления в сети Internet.

Проект RA осуществлялся объединенными усилиями компании Merit Network, Inc., Института информационных технологий при Южнокалифорнийском университете (University of Southern California Information Sciences Institute — USC ISI), компании Cisco Systems (как субподрядчика ISI) и Университета штата Мичиган (University of Michigan ROC) (с компанией Merit в качестве субподрядчика).

Проект RA подразделялся на четыре дочерних субпроекта.

- Сервер маршрутов (Route Server — RS) — В качестве RS может выступать рабочая станция Sun, установленная на каждой NAP. Сервер маршрутов обменивается с маршрутизаторами провайдеров, установленными в NAP, только информацией о маршрутах. Кроме того, он может отвечать за соблюдение заданных правил маршрутизации для каждого конкретного провайдера (RIPE 181). Сам по себе сервер маршрутов не пересылает пакеты и не выполняет какую-либо коммутацию между сервис-провайдерами.

  Сервер маршрутов всего лишь облегчает взаимодействие провайдеров путем сбора информации о маршрутах от каждого провайдера и ее последующего оптимального распределения среди провайдеров. Таким образом, устраняется необходимость в организации среди провайдеров соединений типа "каждый с каждым", и сокращается общее число узлов с $(n-1)$ до 1, где n — количество необходимых маршрутизаторов.

При такой конфигурации маршрутизаторы провайдеров заняты исключительно коммутацией трафика от одного провайдера к другому и практически не участвуют в фильтрации пакетов и соблюдении правил маршрутизации (эта задача перекладывается на серверы маршрутов).

- Система управления сетью (Network Management System) — Специальное программное обеспечение, с помощью которого осуществляется мониторинг производительности RS. На каждом RS запускаются специальные программы-агенты — *распределенные указатели, с* помощью которых ведется сбор статистики о производительности RS. В центре по мониторингу маршрутов компании Merit (Merit Routing Operations Center) находится центральная станция управления сетью (central network management station — CNMS), которая опрашивает всех агентов и обрабатывает собранную информацию.
- База данных арбитража маршрутизации (Routing Arbiter Database — RADB) — Эта база данных является одной из нескольких баз данных маршрутов, известных как регистр маршрутов сети Internet (Internet Routing Registry — IRR). Правила маршрутизации в RADB выражаются с помощью синтаксиса, описанного в документе RIPE-181, который был разработан координационным центром сети RIPE (RIPE Network Coordination Center — RCC). Совместно с RADB была разработана и база данных правил маршрутизации Policy Routing Database (PRDB). База данных PRDB с 1986 года использовалась для конфигурирования магистральных маршрутизаторов в сети NSFNET. В 1995 году, с вводом в действие определенных в RIPE-181 терминологии и процедур, которые при соблюдении глобальных правил маршрутизации обеспечивали лучшую функциональность, эта база данных прекратила свое существование, а все ее функции были переданы базе RADB.
- Группа инженеров по маршрутизации (Routing Engineering Team) — Эта группа работает совместно с провайдерами при разрешении технических проблем, возникающих в NAP. В обязанности специалистов группы входит предоставление консультаций по стратегиям маршрутизации, разработке планов адресации и другим вопросам, касающимся маршрутизации.

На практике серверы маршрутов играют очень важную роль с точки зрения безопасности, так как при обновлении маршрутной информации они требуют проверки подлинности от участвующих в обмене сторон, предотвращая таким образом распространение искаженной маршрутной информации среди узлов сети.

Итак, основными частями концепции арбитража маршрутов являются серверы маршрутов и RADB. Практическими и административными целями RADB является оказание содействия при подключении провайдеров к NAP. Внесение правильной информации в RADB является существенным при избрании определенного набора правил маршрутизации.

В первом издании этой книги в приложении А был полностью представлен документ RIPE-181. Однако в настоящее время большинство IRR постепенно переходят на новую спецификацию правил — так называемый язык спецификации правил маршрутизации — Routing Policy Specification Language (RPSL). RPSL представляет собой новое поколение языка описания правил маршрутизации в IRR. Он был разработан подразделением Routing Policy Specification (RPS) Working Group, входящим в инженерную группу Internet Engineering Task Force (IETF)8. Позднее он был утвержден в качестве стандарта в RFC 2622 с пояснениями, вышедшими в RFC 2650, под руководством USC ISI. Поскольку RPSL способствовал выходу из употребления RIPE-181, в настоящем издании в приложение А включена информация о RPSL (выдержки из RFC 2650 "Использование RPSL на практике").

Для перехода от RIPE-181 к RPSL имеется несколько объективных причин. В RPSL включены улучшенные механизмы для масштабирования правил маршрутизации и алгоритмы аутентификации, которые позволяют обеспечить целостность адресной и другой информации между операторами сети. Более подробно с языком RPSL можно ознакомиться в приложении А в конце книги.

Как клиенту сервис-провайдера вам, возможно, никогда не придется столкнуться с конфигурированием на языках RIPE-181 или RPSL. Однако в любом случае следует

понимать, каким образом с помощью этих языков формируются наборы правил маршрутизации. Читая книгу, вы увидите, что правила являются основой маршрутизации и функционирования такой сложной структуры, как сеть.

С другой стороны, концепция серверов маршрутов и соединение между собой центральных маршрутизаторов не ограничиваются провайдерами в NAP. То есть предоставляется возможность реализации любой доступной архитектуры сети. В одном из разделов этой книги в качестве практического примера рассматривается реализация средства соединения типа "один со многими" на базе сервера маршрутов. К тому же провайдерам, представленным в NAP, необязательно организовывать соединения с серверами маршрутов, как и не существует причин отказывать провайдерам, которые работают в основном по прямым межсетевым соединениям, в использовании серверов маршрутов.

В марте 1998 года компания Merit успешно завершила реализацию проекта арбитража маршрутизации. Сегодня компания Merit и ее партнер по проекту RA — Институт информационных технологий при Южнокалифорнийском университете продолжают проводить исследования в рамках проекта RA.

Следуя рекомендациям NSF, компания Merit совместно с ISI выполнила комплекс работ по переходу с серверов маршрутов проекта RA на сервер маршрутов следующего поколения — Route Server Next Generation (RSng), согласно которому предполагалось предоставлять услуги сервера маршрутов компании Merit на коммерческой основе путем покупки операторами услуг серверов друг друга. Все операции в NAP также были подвергнуты коммерциализации с учетом того, что они преследуют цель создания на рынке провайдеров Internet равных условий для ведения бизнеса.

В 1997 году компания Merit при финансовой поддержке NSF начала работу над проектом анализа и измерения производительности в Internet — Internet Performance Measurement and Analysis (IPMA). Основное назначение IPMA — изучение производительности сетей и сетевых протоколов путем сбора и анализа статистики о маршрутах и производительности в сетях с целью повышения стабильности работы в сети Internet. Кроме того, в рамках IPMA предполагается разработать средства, которые бы облегчили функционирование сети и некоторые инженерные процедуры.

Североамериканская группа операторов сети North American Network Operators Group (NANOG) была основана NSF с первых дней работы NSFNET и проекта RA. Работа в этой группе началась с региональных технических семинаров, проходивших для персонала, который занимался обслуживанием сети NSFNET. Сегодня она финансируется от регистрационных взносов, собираемых на проведение конференций, которые организует компания Merit. Группа NANOG предоставляет трибуну для обсуждения технических вопросов, связанных с функционированием сетей в Северной Америке.

В настоящее время базы данных и другие инструменты, созданные в процессе работы над проектами RADB широко используются провайдерами Internet и стали сегодня неотъемлемой частью сети Internet.

Для обеспечения стабильности и безопасности глобальной маршрутизации в Internet необходимо проделать еще много работы в области спецификации междоменных правил маршрутизации и порядка контроля за их выполнением. Проекты, подобные проекту арбитража маршрутов, представляют собой яркий пример для понимания архитектуры сети Internet.

Высокоскоростная магистральная сетевая служба

Проект организации высокоскоростной магистральной сетевой службы (very high-speed Backbone Network Service — vBNS) был призван обеспечить специализированную магистральную службу для пользователей с высокопроизводительными компьютерными

системами в государственных суперкомпьютерных центрах (Supercomputer Centers — SCC).

24 апреля 1995 года компании MCI и NSF объявили о начале работы службы vBNS. Компания MCI обязалась обеспечить:

- монтаж и обслуживание транзитной сети с пропускной способностью от 155 Мбит/с и выше, которая бы коммутировала IP и CLNP и объединяла все NAP, финансируемые NSFNET;
- установку и определение метрики сети для мониторинга производительности сети;
- описание наборов правил для менеджеров NAP и RA;
- сервисы мультимедиа;
- разработку новых технологий маршрутизации и предложений по увеличению скорости и качества обслуживания, которые были бы совместимы с требованиями заказчика в лице NSF.

Согласно контракту стоимостью 50 миллионов долларов, заключенному на 5 лет, компания MCI и NSFNET обязались связать воедино пять основных высокопроизводительных коммуникационных центров, принадлежащих NSF:

- Центр теории имени Корнелла (Cornell Theory Center — CTC) в Итаке (штат Нью-Йорк);
- Национальный центр исследований атмосферы (National Center for Atmospheric Research — NCAR) в Боулдере (штат Колорадо);
- Национальный центр по применению суперкомпьютеров (National Center for SuperComputing Applications — NCSA) при университете штата Иллинойс;
- Питгсбургский суперкомпьютерный центр (Pittsburg SuperComputing Center — PSC):
- Суперкомпьютерный центр в Сан-Диего (San Diego Supercomputer Center — SDSC).

Служба vBNS создавалась как своего рода исследовательская лаборатория XXI века. Применение новейших технологий коммутации и оптоволоконных линий связи, режима асинхронной передачи (Asyncronous Transfer Mode — ATM) и создание синхронной оптической сети (Syncronous Optical Network — SONET) позволило обеспечить одновременную передачу с высокой скоростью голосовых и видеосигналов, чувствительных к задержкам.

С разрешения NSF служба vBNS используется для высокоемких приложений, требующих большой полосы пропускания (например, суперкомпьютерное моделирование обледенения самолетов в NCAR). Используется она также и для других задач. Так, в NCSA строятся математические модели, моделирующие работу биологических мембран.

Служба vBNS доступна через четыре NAP, которые расположены в Нью-Йорке, Сан-Франциско, Чикаго и Вашингтоне. На рис. 1.7 показано географическое расположение центров и NAP и связи между ними. Как видите, в основном vBNS собрана на базе каналов OC 12, соединенных посредством высокопроизводительных маршрутизаторов, поставляемых компаниями Juniper Networks и Cisco Systems. Первый магистральный канал OC48c на базе маршрутизаторов Juniper был введен в эксплуатацию в январе 1999 года.

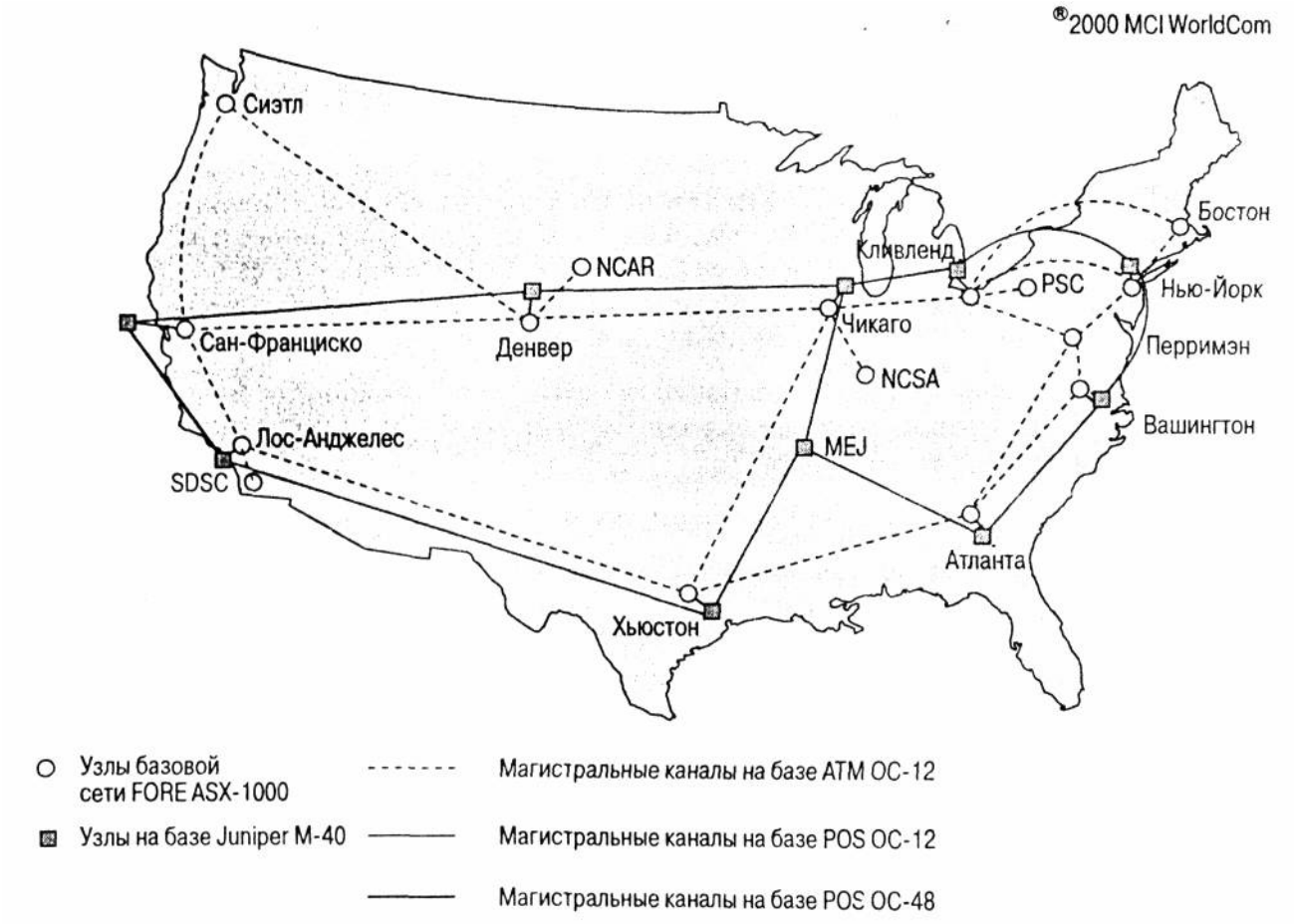


Рис. 1.7. *Карта опорной сети vBNS*

Печатается с разрешения MCIWORLDCOM. Копирайт 2000. Все права защищены. Эти материалы разработаны на основе работ, проведенных при поддержке Национального научного фонда под регистрационным номером NCR 931047. Заключения или рекомендации, содержащиеся в этом материале, являются авторскими и не всегда совпадают с точкой зрения Национального научного фонда США (National Science Foundation).

Служба vBNS представляет собой специализированную сеть, необходимость организации которой была вызвана растущими потребностями в высокоскоростных соединениях для обмена информацией между исследовательскими центрами и научными институтами. Одним из основных потребителей этой службы была на протяжении многих лет сеть NSFNET. Несмотря на то что сама по себе служба vBNS не дает общего представления о глобальной маршрутизации в крупных сетях на базе TCP/IP, ее экспериментальная эксплуатация явилась прекрасным полигоном для "обкатки" новых технологий. Краткий обзор исторических сведений имел своей целью лишь составить у вас общее представление о зарождении и развитии сети NSFNET вплоть до прекращения работы в ней в 1995 году.

Сегодня хосты vBNS в четырех SCC и более 80 университетах предлагают доступ в сеть со скоростями, варьирующимися от DS3 до OC 12. Пользователи постепенно переходят к использованию IPv6 (протокол IP версии 6 — *Прим. ред.),* многоадресным службам и MPLS. С апреля 2000 года между NSF и MCI WorldCom ведутся переговоры по продолжению поддержки vBNS, так как пятилетний контракт истек именно в конце апреля 2000 года. Предполагается привлечь инвесторов для этого проекта путем коммерциализации сервисов vBNS с отказом от Правил использования (Acceptable Use Policy — AUP), принятых ранее. Кроме того, планируется создать специальное целевое сообщество для предоставления этих сервисов.

Ввиду комплексных исследований в области QoS и управления трафиком в настоящее время за службой vBNS закрепился термин *сверхвысокопроизводительная магистральная сетевая служба (very high-performance Backbone Network Service).* Таким

образом, vBNS продолжает традиции, заложенные сетью NSFNET.

Выделение региональных сетей из NSFNET

В ходатайство о переходе на новую архитектуру сети Internet NSF включил раздел, в котором региональные сети (также именуемые *сетями среднего уровня (mid-level networks))* должны были обеспечить соединение магистральных каналов NSFNET с другими провайдерами.

С момента создания NSFNET основную роль в сетях образовательных и научно-исследовательских центров играли региональные сети. Провайдеры региональных сетей (Regional Network Provider — RNP) подключали к сети клиентов лишь на основе их принадлежности к определенным организациям (например, университеты), обеспечивая при этом широкий спектр сетевых услуг для них и поддержку межрегиональных соединений (Inter-Regional Connectivity — IRC).

Ниже приведены оговоренные в NSF 93-52 обязанности RNP.

- Обеспечить межрегиональные соединения путем прямого подключения к NAP с организацией привязки к одному или нескольким NSP.
- Обеспечить снабжение организаций-клиентов новейшими сетевыми информационными службами совместно с InterNIC и менеджером информационных служб NSFNET (NSFNET Information Services Manager).
- Определить процедуры взаимодействия технического персонала всех уровней при решении технических проблем, поддержки работоспособности и .'заданного уровня качества обслуживания для пользователей сети, начиная с менеджеров NAP, RA, провайдеров vBNS и кончая администраторами региональных и других сетей.
- Создать службы, в функции которых входит учет и координация работы поль зователей в сети из числа исследователей и технических специалистов в области телекоммуникаций.
- Обеспечить работу соединений с высокой пропускной способностью (по возможности совместно с NSP) в интересах учреждений и организаций — пользователей сети, где имеются приложения, требующие широкой полосы пропускания.
- Обслуживание сетевых соединений организаций — членов сети.

В процессе выделения региональных сетей из сети NSFNET и формирования новых соединений между ISP, NSF предложил подключать их либо напрямую к NAP, либо через провайдеров, уже подключенных к NAP. На переходный период NSF определил цены на организацию соединений, которые в течение 1 года должны были постепенно снижаться до нуля (по истечении первого срока соглашения о Сотрудничестве между менеджерами NAP и RA (NAP Manager/RA Cooperative Agreement), т.е. четырех лет).

В табл. 1.1 приводится список первых региональных провайдеров NSFNET и их преемников в современной сети Internet. Как видите, большинство из региональных провайдеров перешли на работу в MCInet (теперь Cable & Wireless) либо в Sprintlink. Переход региональных провайдеров к апрелю 1995 года на новую структуру сети Internet был одной из основных задач.

Таблица 1.1. *Эволюция преобразования региональных провайдеров*

Старые региональные сети	Новые провайдеры Internet
Argone	CICnet
BARRnet	MCInet
CA*net	MCInet
CERFnet	CERFnet
CICnet	MCInet
Cornell Theory Center	MCInet

CSUnet	MCInet
DARPA	ANSnet
JvNCnet	MCInet
MOREnet	Sprintlink
NEARnet	MCInet
NevadaNet	Sprintlink
SESQUiNET	MCInet
SURAnet	MCInet
THEnet	MCInet
Westnet	Sprintlink

Создание NIS

Специалисты фонда NSF прекрасно понимали, что наиболее критичным компонентом в разветвленной стихийно развивающейся сети являются информационные сетевые службы. Как результат появился документ об учреждении в сети NSFNET сетевой информационной службы (Network Information Services — NIS). В нем содержались следующие предложения.

- Расширить и координировать действия служб каталогов, баз данных и информационных служб в сети.
- Обеспечить регистрацию в сетях невоенного назначения, входящих в Internet. В военных сетях регистрацию будет продолжать Сетевой информационный центр Агентства оборонных информационных систем (Defense Information Systems Agency Network Information Center — DISA NIC).

Ко времени выхода документа невоенную часть сети Internet составляли NSFNET и другие сети, финансируемые федеральным правительством США, например научная сеть Национального агентства аэрокосмических исследований (NASA Science Internet — NSI) и научная сеть Министерства энергетики США (Energy Sciences Network — ESnet). Поддержание и развитие этих сетей вместе с несколькими другими сетями, входящими в Internet, были включены в бюджет США 1992 года как Национальная исследовательская и образовательная сеть (National Research and Education Network — NREN). Ходатайство о необходимости создания службы баз данных, информационной службы и службы регистрации, поданная NSF правительству США, имела своей целью помочь развитию и NSFNET NREN в целом.

Сетевые информационные службы

Ко времени подачи NSF ходатайства, многие провайдеры уже предлагали в той или иной форме услуги сетевых информационных служб (Network Information Services — NIS). Ниже приведены некоторые из них.

- Информационная служба для конечных пользователей, обеспечиваемая Центром сетевых услуг NSF (NSF Network Services Center — NNSC). Услуги этой службы предоставляла компания Bolt, Beranek and Newman (BBN). Другие пользовательские службы NSFNET, организованные в университетских город ках и организациях, занимающихся обеспечением работоспособности сети.
- Информационные услуги по обслуживанию федеральных опорных сетей различных агентств, как правило, предоставлялись самими агентствами. Так, например, NASA имело свою собственную информационную службу — NSI.
- Услуги регистрации в сети Internet предоставлялись центром DISA NIC, который работал под управлением специально созданной государственной компании

Government Services, Inc. (GSI).

- Информационные услуги на уровне провайдеров в университетских городках предоставлялись сетевыми структурами NSFNET среднего уровня.
- Информационные услуги для организаций среднего уровня NSFNET предоставляла компания Merit, Inc.

Согласно новому ходатайству, менеджеры NIS должны были обеспечить предоставление информационных услуг конечным пользователям, провайдерам университетских городков и провайдерам среднего уровня. Кроме того, они должны взаимодействовать с другими сетевыми организациями, такими как Merit, Inc.

Создание InterNIC

В январе 1993 года в качестве совместного проекта, инициированного компаниями AT&T, General Atomics и Network Solutions, Inc11, был создан Международный сетевой информационный центр (International Network Information Center — InterNIC). Его работа обеспечивалась тремя двусторонними соглашениями сроком на пять лет, заключенными инициаторами проекта с NSF. Однако на втором году действия соглашений NSF прекратил финансирование General Atomics. Компания AT&T стала полностью вести службы каталогов и баз данных (Database and Directory Services), а на долю Network Solutions выпало обслуживание службы регистрации и поддержки сетевого информационного центра (Registration and NIC Support Services).

Служба каталогов и баз данных

Реализация этой службы требует взаимодействия между распределенными базами данных и применения других новейших технологий. Менеджер NIS может координировать работу этой службы для нескольких организаций, которые создали и постоянно поддерживают работу соответствующих каталогов и баз данных. Согласно соглашению с NSF, компания AT&T обеспечивала следующие виды услуг.

- Служба каталогов (Directory Services) (известны также под названием white pages):
  С ее помощью обеспечивается доступ к информации, содержащейся в базе Internet White Pages, посредством X.500, системы WHOIS и netfind.
  Стандартом X.500 предусматривается возможность создания простого глобального каталога с различного рода информацией, разделенной по тематике.
  Система WHOIS обеспечивает унифицированный поиск организаций и отдельных людей в сети Internet на трех основных WHOIS-серверах. Для невоенных доменов и данных, не являющихся персональными, она производит поиск на сервере службы каталогов и баз данных InterNIC. Поиск данных в доменах военных организаций иерархии MIL (military) осуществляется через сервер DISA NIC, а контактная информация распространяется через сервер Регистрационной службы в InterNIC (InterNIC Registration Services server).
  Система Netfind представляет собой поисковую машину Internet по глобальному каталогу "Белые страницы" (white pages). Если задать имя пользователя сети и описать его место работы (организацию), то поисковая система попытается найти информацию о заданном пользователе.
- Служба баз данных (Database Services):
  В нее входят базы данных документов, регламентирующих работу сети Internet, таких как Request For Comments (RFC), Internet Drafts (ID), бюллетени IETF (IETF Meeting minutes), бюллетени Группы управления IETF (IETF Steering Group — IESG) и другие. Служба может также включать в себя базы данных, за определенную плату обслуживаемые другими группами. Компания AT&T

предложила также обслуживать общедоступные базы данных, которые представляют интерес для сетевого сообщества.

- Каталог каталогов (Directory of directories):
  Эта служба ведет учет каталогов и баз данных, подобных приведенным выше. Она представляет собой индекс указателей на ресурсы, продукты и другие службы, доступные в сети Internet. Сюда входят указатели на такие ресурсы, как вычислительные центры, сетевые провайдеры, информационные серверы, каталоги белых и желтых страниц, каталоги библиотек и т.п.

В качестве части этой услуги компания AT&T обеспечивает хранение списков информационных ресурсов, включая тип ресурса, его описание, метод доступа к ресурсу и другие атрибуты. Информационным провайдерам разрешено проводить обновление и добавление информации в базы данных. Доступ к той или иной информации можно получить различными средствами — через Telnet, FTP, электронную почту (E-mail) и "всемирную паутину" World Wide Web.

Службы регистрации

Менеджер NIS был обязан действовать в соответствии с RFC 1174, где говорилось следующее.

> *Для определения и назначения различных цифровых идентификаторов в сети Internet учреждается центральный орган — Организация по распределению цифровых адресов в сети Internet (Internet Assigned Numbers Authority — IANA[12]). Функции IANA выполняются Институтом информационных технологий при Южнокалифорнийском университете (University of California's Information Sciences Institute). Организации IANA принадлежат исключительные права по делегированию ответственности за блоки IP-адресов и номера автономных систем, выделяемых организациям, на правах Реестра сети Internet (Internet Registry — IR).*

Таким образом, менеджер NIS становился либо регистратором IR, либо делегировал это право с разрешения IR кому-либо другому. К вопросам, находящимся в ведении служб регистрации в сети Internet, относились:

- назначение сетевых адресов;
- назначение номеров автономных систем;
- регистрация доменных имен;
- регистрация серверов доменных имен.

С 1993 по 1998 год, согласно Договору о сотрудничестве с правительством США, NSI являлся единственным провайдером для регистрации доменных имен в доменах верхнего уровня .com, .net и .org. В 1998 году в Договор были внесены поправки, и теперь NSI разрабатывает программное обеспечение для поддержки распределенной системы регистрации (Shared Registration System) в доменах верхнего уровня.

Сегодня правительство США начало процесс приватизации системы управления доменными именами в сети в надежде, что это даст новый толчок к развитию конкуренции, а, следовательно, и к развитию всего сетевого сообщества.

Наблюдательные функции за этим процессом были возложены на Корпорацию по назначению адресов и. доменов в сети Internet (Internet Corporation for Assigned Names and Numbers — ICANN). Эта организация отвечает за аккредитацию компаний, регистрирующих доменные имена. Правительством США на нее также были возложены определенные функции по управлению системой доменных имен в сети Internet. Организация ICANN является международной некоммерческой организацией.

Служба поддержки NIC

В первой редакции документа об информационных службах ("Information Services") предполагалось, что их организацию возьмет на себя компания General Atomics. Однако эти функции перешли к службе NSI, которая была переименована в службу поддержки NIC (NIC Suppot Services).

В задачу этой службы входила организация форумов по темам с участием исследователей, студенческих кругов, персонала Сетевых информационных центров (Network Information Center's — NIC) и академического сообщества сети Internet. Кроме того, она должла была выполнять функции, дополнительно возлагаемые на нее InterNIC.

Другие реестры в сети Internet

С приватизацией служб регистрации изменились процедуры распределения пространства IP-адресов и назначения номеров автономным системам (Autonomous System — AS). В настоящее время Региональными реестрами сети Internet (Regional Internet Regisrty — RIR) обеспечивается регистрация в сети Internet по всему земному шару — это Американский реестр адресов сети Internet (American Registry for the Internet Numbers -— ARIN), Европейский сетевой координационный центр (Reseaux IP Europeens Network Coordination Center — RIPE NCC) и Азиатско-Тихоокеанский сетевой информационный центр (APNIC).

Американский реестр адресов сети Internet

В конце 1997 года IANA передала права на администрирование IP-адресов от компании Network Solutions, Inc. Американскому реестру ARIN. Официально реестр ARIN начал свою деятельность 22 октября 1997 года.

В настоящее время ARIN отвечает за распределение IP-адресов в следующих географических регионах:

- Северная Америка.
- Южная Америка.
- Страны Карибского бассейна.
- Центральная и Южная Африка.

В настоящее время ARIN управляет распределением и регистрацией IP-адресов, номеров автономных систем (AS), корневым доменом IN-ADDR.ARPA и экспериментальным доменом IP6.INT. Кроме того, эта организация обеспечивает регистрацию в реестре маршрутизации, т.е. операторы сети могут регистрироваться, получать и обновлять конфигурационную информацию для маршрутизаторов, пользоваться услугами службы WHOIS для просмотра информации по заданным критериям.

ARIN является некоммерческой организацией. Основные источники финансовых поступлений — регистрационные взносы за назначение и обслуживание блоков IP-адресов, поступающие от членов ARIN.

Европейский сетевой координационный центр

Созданный в 1989 году Европейский сетевой координационный центр (Reseaux IP Europeens Network Coordination Center — RIPE NCC или просто RIPE) представлял собой совещательный орган провайдеров сети Internet в Европе. Его основная цель — администрирование и координация работы в Европейском сегменте сети Internet. Для Европы и соседних с ней территорий RIPE выступает в качестве RIR.

RIPE распределяет IP-адреса, координирует работу системы доменных имен (Domain Name System— DNS), а также ведет базу данных с информацией об IP-сетях, серверах DNS и

правилах IP-маршрутизации. Она также обеспечивает функции хранилища программного обеспечения, необходимого для работы с Internet, документов RIPE. С ее помощью поддерживаются службы регистрации в реестре маршрутизации и интерактивные информационные службы.

Подобно ARIN, RIPE является некоммерческой организацией и все финансовые поступления производятся лишь за счет предоставляемых ей услуг.

Азиатско-Тихоокеанский сетевой информационный центр

Созданный в 1993 году Азиатско-Тихоокеанский сетевой информационный центр (Asian Pasific Network Information Center — APNIC) предоставляет услуги, сходные с услугами ARIN. Центр APNIC предоставляет эти услуги во всем Азпатско-Тихоокеанском регионе, включающем 62 страны и региона в Южной и Центральной Азии, в Юго-Восточной Азии, Индокитае и на островах Океании.

В настоящее время APNIC не занимается администрированием DNS, хотя и оказывает посильную помощь в организации этой системы в регионе. Центр APNIC предоставляет комплекс услуг по обучению и подготовке специалистов для сетей, разработке правил работы в сети, а также организовывает региональные сетевые мероприятия. Самая заметная заслуга APNIC — основание Азиатско-Тихоокеанской региональной Internet-конференции по технологиям (Asian Pasific Regional Internet Conference on Operational Technologies — APRICOT), которая стала ведущим форумом для обсуждения насущных проблем среди сетевых операторов и администраторов.

Реестры маршрутизации в Internet

С появлением новых ISP, которые желали развивать инфраструктуру межсетевых соединений, возросли требования к гибкости и управляемости сети. Каждый провайдер руководствуется набором правил или стратегий (так называемая *policy* — линия поведения, политический курс — *Прим. ред.),* который определяет, какую информацию принимать от других сетей и какую отдавать в соседние сети из своей сети. Примеры наборов правил включают в себя определение маршрутов, их фильтрацию от заданного ISP и выбор оптимального пути для доставки информации по ним в пункт назначения. К сожалению, вследствие множества существующих наборов правил между провайдерами часто возникают конфликтные ситуации, что негативно влияет на работу сети в целом.

Реестры маршрутизации в Internet (Internet Routing Registries — IRR) служат также общедоступной базой данных, содержащей информацию об организации маршрутов и об ответственных лицах, с которыми следует контактировать для координации работы системы маршрутизации и при разрешении проблем.

Для того чтобы адресовать соответствующим образом все вызовы для каждого глобального домена, пришлось создать отдельный реестр маршрутизации (routing registry — RR). В каждом RR ведется своя база данных правил маршрутизации, которая обновляется всеми провайдерами, входящими в RR. Совокупность этих баз данных известна под названием Реестра маршрутизации в сети Internet (Internet Routing Registry — IRR).

Главная задача RR — не определение правил маршрутизации, а хранение этих правил и другой информации административного характера. Таким образом, предполагается создать глобальную базу данных правил маршрутизации, используемых всеми провайдерами по всему земному шару. Большинство операторов сети получают информацию о маршрутизации из реестров маршрутизации, что позволяет им оперативно изменять правила маршрутизации.

Автономные системы (Autonomous Systems — AS) для работы друг с другом

используют протоколы внешнего шлюза (Exterior Gateway Protocols — EGP), такие как протокол граничного шлюза (Border Gateway Protocol — BGP). В случае сложных сетей появляется необходимость в стандартизации правил описания и правил объединения различных AS. Поддержка гигантской базы данных, содержащей зарегистрированные по всему миру наборы правил, была бы обременительной и сложной. Поэтому было найдено другое решение — использовать распределенные базы данных. Каждый RR ведет собственную базу данных и обеспечивает ее согласованность с другими базами данных. Ниже приведено несколько существующих сегодня баз данных IRR.

- Реестр маршрутизации RIPE (RIPE Routing Registry), куда входят все европейские провайдеры Internet.
- Реестр маршрутизации компании Cable & Wireless (Cable & Wireless Routing Registry) для клиентов компании.
- Реестр маршрутизации CA*net для пользователей CA*net (CA*net Routine Registry).
- Реестр маршрутизации провайдеров Internet в Японии JPRR (Japanese Internet service providers Routing Registry).
- Общедоступная база данных арбитража маршрутизации (Routing Arbiter Database).
- Общедоступный реестр маршрутизации ARIN (ARIN Routing Registry).

Каждый из приведенных выше реестров служит базой провайдеров Internet для пользователей, за исключением базы данных арбитража маршрутизации (Routing Arbiter Database — RADB) и ARIN, где регистрируются все желающие. Как уже упоминалось, RADB является частью проекта арбитража маршрутизации (Routing Arbiter project).

Ввиду высокой гибкости и других преимуществ локальных реестров несколько других компаний, таких как Qwest, Level(3) и Verio, также учредили собственные RR.

Настоящее и будущее Internet

Коммерциализация сети Internet не только не помешала ее техническому развитию, а, наоборот, способствовала ему. Разработка новых технологий в коммерческом секторе и в исследовательских и образовательных учреждениях велась ускоренными темпами. Сегодня новейшие технологии уже не могут сразу же внедряться в "коммерческую" сеть Internet. Сначала они должны быть тщательно проверены и оптимизированы для реальных условий. Для адаптации и внедрения новейших сетевых технологий были созданы специальные системы отладки.

Инициатива "Internet следующего поколения"

Правительством США финансируется инициатива "Internet следующего поколения" (Next-Generation Internet — NGI Initiative) — федеральная программа, предусматривающая разработку новейших сетевых технологий и приложений, а также систем отладки, которые будут в 1000 раз быстрее, чем существующие сегодня.

В программе NGI, стартовавшей 1 октября 1997 года, принимают участие следующие государственные агентства:

- Оборонное агентство по передовым исследованиям (Defense Advanced Research Projects Agency — DARPA);
- Министерство энергетики США (Department of Energy — DoE);
- Национальное агентство аэрокосмических исследований (National Aeronautics and Space Administration — NASA);
- Национальный институт здоровья (National Institute of Health — NIH);
- Национальный институт стандартов и технологий (National Institute of Standards

and Technology — NIST);

- Национальный научный фонд (National Science Foundation — NSF).

Управление инициативой NGI осуществляется согласно индивидуальным программам агентств и координируется рабочей группой по развитию и исследованию крупномасштабных сетей при подкомитете по компьютеризации, информации и связи (Large-Scale Network Working Group of the Subcommittee on Computing, Information and Communications — CIC R&D), который входит в комитет по технологиям при Национальном совете по науке и технологиям в Белом Доме (White House National Science and Technology Council's Committee on Technology).

Основными целями NGI являются:

- внедрение новейших сетевых технологий;
- установка и обеспечение работы двух систем отладки;
- внедрение новых приложений.

Внедрение новейших сетевых технологий

NGI способствует развертыванию и апробации новых технологий, которые однажды станут частью коммерческой сети Internet. Эти технологии охватывают множество аспектов работы вычислительных сетей:

- надежность;
- универсальность;
- безопасность;
- качество обслуживания (Quality of Service) и дифференциацию обслуживания (включая многоадресную передачу и обмен видеоинформацией);
- управление сетью (включая выделение и распределение полосы пропускания).

Установка и обеспечение работы двух систем отладки

Введение в действие мощных систем отладки — ключевое условие для достижения целей, поставленных перед NGI. Для этого были разработаны две системы отладки — "100x" и "1000x".

Система отладки "100x" объединит около 100 различных организаций — университетов, федеральных научно-исследовательских институтов и др. — и позволит им обмениваться информацией на скорости, в 100 раз превышающей среднюю скорость обмена данными в Internet сегодня.

Система отладки будет строиться на основе следующих федеральных сетей:

- сверхскоростной магистральной сетевой службы NSF (NSF's very high-speed Backbone Network Service — vBNS).
- исследовательской и образовательной сети NASA (NASA's Research and Educational Network — NREN).
- исследовательской и образовательной сети Министерства обороны США (DoD's Defense Research and Education Network — DREN).
- научно-исследовательской сети Министерства энергетики США (DoE's Energy Sciences network — ESnet).

Система отладки "1000x" будет объединять около 10 узлов с пропускной способностью, в 1000 раз превышающей сегодняшнюю. Система отладки "1000x" будет строиться на основе сети DARPA's SuperNet.

Вышеприведенные системы отладки будут использоваться для тестирования новейших технологий и услуг, а также для разработки и тестирования новых приложений.

Внедрение новых приложений

Исследования NGI будут сфокусированы на создании и внедрении приложений и технологий:

- технологии совместного использования;
- цифровые библиотеки;
- распределенные вычисления;

- безопасность при работе в сети;
- удаленная работа и модерирование.

Кроме того, делается упор на разработку специфических приложений:

- естественные науки;
- менеджмент кризисных ситуаций;
- образование;
- охрана окружающей среды;
- федеральные информационные службы;
- охрана здоровья;
- автоматизация производства.

Проект Internet2

Проект Internet2 инициирован университетской корпорацией по развитию сети Internet (University Corporation for Advanced Internet Development — UCAID). Этот проект был анонсирован в октябре 1996 года при участии 34 университетов и имел своей целью сохранить первенство США в развитии, внедрении и эксплуатации сетевых приложений и сетевой архитектуры следующего поколения. Основная роль, отводимая Internet2, — обеспечить благоприятные условия для развития Internet-приложений и сетевых протоколов, которые бы стабилизировали исследовательские и образовательные функции университетов. При экспоненциальном росте сети Internet, который мы наблюдаем сегодня, коммерческие сети, контролируемые провайдерами расширяют свою полосу пропускания и стараются внедрять новые технологии едва ли не быстрее, чем это делается в исследовательских и образовательных сетях. Одна из основных целей создания Internet2 — сохранить лидирующие позиции в области исследования сетей за существующими системами отладки и облегчить перенос этих технологий в глобальную сеть Internet.

Сегодня Internet2 представляет собой плод совместных усилий более чем 160 университетов США и 50 крупнейших корпораций. Финансирование Internet2 осуществляется университетами и корпорациями, которые входят в UCAID. Большинство членов этой корпорации финансируются за счет грантов, предоставляемых NSF и другими федеральными агентствами, принимающими участие в проекте NGI. Финансирование осуществляется также посредством других программ фонда NSF.

Основная задача Internet2 — не заменить существующую сеть Internet, а расширить сферы применения технологий, разработанных на базе Internet2.

Проект Abilene

Проект Abilene — еще одна программа, находящаяся в ведении корпорации UCAID. Он представляет собой дополнение к проекту Internet2, и основной его задачей является обеспечение первичной опорной сети для проекта Internet2. UCAID при содействии Qwest Communications, Nortel Networks и Cisco Systems была создана сеть Abilene Network. Эта сеть обеспечивает региональные узлы Internet2 высокопроизводительными соединениями. В основном сеть Abilene, эксплуатация которой начата в январе 1999 года, на сегодняшний день предоставляет доступ на основе OC3 и OC 12. Главные магистральные каналы этой сети поддерживают доступ OC48c (2,5 Гбит/с) посредством POS (Packet Over SONET).

Подобно vBNS, в сети Abilene проводятся исследования новых Internet-технологий, но в целях обеспечения стабильности работы сети в Abilene используется отдельная высокопроизводительная тестовая сеть для опробования приложений, которые еще не могут применяться в стабильной и передовой сети Abilene. Рабочие группы в Internet2 в настоящее время занимаются проработкой деталей развертывания сети Abilene, концентрируя свои усилия на разработке чистой процедуры многоадресного (циркулярного) обращения, оптимизации правил и структуры маршрутизации, развитии протокола Internet версии 6 (Internet Protocol version 6 — IPv6) и критериев качества обслуживания (Quality of Service —

QoS). В сети Abilene уже сейчас достигнуты определенные успехи по многоадресному обращению и планируется развертывание IPv6 и QoS.

На рис. 1.8 представлена структура сети Abilene.

Рис. 1.8. Карта сети Abilene

Забегая вперед

Прекращение эксплуатации сети NSFNET в 1995 году обозначило начало новой эры. Сегодня сеть Internet — это площадка для игр нескольких тысяч провайдеров, которые борются за свою часть рынка. Исследовательские и экспериментальные сети, такие как Abilene и vBNS, борются за возможность возглавить развивающуюся индустрию, стоимость которой составляет несколько миллиардов долларов. Для большинства компаний и организаций подключение своих сетей к глобальной сети Internet уже не роскошь, а обязательное условие для того, чтобы оставаться конкурентоспособными на рынке.

Структура современной сети Internet для сервис-провайдеров и их клиентов рассматривается только с позиций скорости доступа, надежности и стоимости услуг. Ниже приведены вопросы, которые должны быть продуманы при подключении организации к сети Internet.

- Смогут ли потенциальные провайдеры хорошо спроектировать схемы маршрутизации для вашей сети, предусматривающие изменения маршрутов
- Что и в какой степени клиенты провайдера должны знать и делать для обеспечения нормальной работы маршрутизации?
- Должны ли клиент и провайдер дать четкое определение стабильной сети?
- Является ли предоставляемая полоса пропускания единственным фактором, определяющим скорость соединения с Internet?

В следующей главе делается попытка дать ответы на эти вопросы. В последующих главах подробно рассмотрена организация маршрутизации и различные схемы маршрутизации.

Несмотря на то что междоменная маршрутизация используется уже на протяжении более 10 лет, для многих она все еще является новой технологией. В остальной части этой книги представлена современная структура сети Internet, а также новейшие приемы маршрутизации для сетей на базе протокола TCP/IP.

Часто задаваемые вопросы

В — *Существуют ли другие NAP кроме четырех, основанных NSF?*

О — Да, существуют. По мере роста числа межсетевых соединений, создавались новые NAP. Сегодня их очень много разбросано практически по всему миру. Они есть в Северной Америке, в Европе, Азии и Тихоокеанском регионе, в Африке, Южной Америке и на ближнем Востоке.

В — *Если я уже подключен к Internet через провайдера, то должен ли я подключаться к NAP?*

О — Нет. Точки NAP служат главным образцом для организации соединений между различными провайдерами. Если вы являетесь клиентом какого-либо провайдера, то имеющегося соединения с провайдером вполне достаточно. Однако от того, к какой или к каким NAP подключен ваш провайдер и сколько он имеет прямых межсетевых соединений с другими провайдерами, будет зависеть качество предоставляемых им услуг.

В — *Является ли основной функцией сервера маршрутов в NAP коммутация трафика между провайдерами?*

О — Нет. Сервер маршрутов содержит базу данных правил маршрутизации, используемых различными провайдерами. Для обмена трафиком провайдеры используют другие физические устройства в NAP.

В — *Все ли провайдеры, подключающиеся к NAP, должны подключаться и к серверу маршрутов?*

О — Хотя это настоятельно рекомендуется делать, такая процедура не является обязательной и провайдеры могут не следовать ей.

В — *Какая разница между IR и IRR?*

О — Реестры сети Internet (Internet Registry— IR), такие как Network Solutions, Inc, отвечают за регистрацию в сети Internet (т.е. за закрепление доменных имен и т.п.). Реестры маршрутизации в сети Internet (Internet Routing Registry— IRR), такие как RADB, отвечают за поддержку базы данных правил маршрутизации для сервис-провайдеров.

В — *Чем отличаются службы баз данных от баз данных арбитража маршрутизации (Routing Arbiter Databases) ?*

О — Службы баз данных являются частью сетевых информационных служб. Они обеспечивают сохранность важных технических документов (таких как RFC и др.). База RADB является базой данных правил маршрутизации.

Ссылки

1. www.darpa.mil
2. www.nsf.gov
3. www.merit.edu
4. www.ra.net
5. www.isi.edu
6. http://www.ietf.org/rfc/rfcl786.txt
7. www.merit.edu
8. www.ietf.org
9. www.nanog.org
10. www.vbns.net
11. www.internic.net

12. www.iana.org
13. www.icann.org
14. www.arin.net
15. www.ripe.net
16. www.apnic.net
17. www.ngi.gov
18. www.internet2.edu
19. www.internet2.edu/abilene

Ключевые темы этой главы:

- **Услуги, предоставляемые провайдерами Internet (Internet Service Provider — ISP).**Проводится классификация ISP с точки зрения методов доступа в Internet, набора базовых услуг и обеспечения безопасности.
- **Оплата услуг ISP.** Дается обзор факторов, влияющих на формирование цен на услуги ISP.
- **Критерии выбора ISP.** Рассматриваются критерии оценки ISP с позиций топологии сети и соглашений об обмене трафиком.
- **Точка демаркации.** Разделение сети, оборудования и зоны ответственности провайдера и клиента.

Глава 2. Услуги, предоставляемые провайдерами Internet и их характеристики

Прежде чем углубиться в рассмотрение междоменной маршрутизации, очень важно, чтобы вы получили базовые знания об основных услугах, предоставляемых провайдерами Internet, и их характеристиках, которые влияют на качество соединения с Internet. В принципе любое физическое и юридическое лицо, предлагающее подключение к сети Internet, может считаться сервис-провайдером, т.е. термин "провайдер" охватывает все компании и организации — от провайдеров с крупномасштабной инфраструктурой по всему миру и магистральными каналами стоимостью несколько миллионов долларов до провайдера с одним маршрутизатором и сервером доступа где-нибудь в гараже за домом. При выборе ISP не следует руководствоваться ценой предоставляемых услуг. Прежде всего выясните такие вопросы,,как набор услуг, предоставляемых провайдером, структу-/i ра его магистральных каналов, отказоустойчивость, резервирование оборудования, стабильность работы, порядок размещения клиентского оборудования у провайдера и т.д.

Формирование маршрутов в сети Internet и их динамика зависит от работы протоколов маршрутизации и потоков Данных, проходящих по уже существующей инфраструктуре сети. Правильно построенная сеть и постоянное ее обслуживание — залог нормальной "здоровой" маршрутизации в сети Internet.

Услуги, предлагаемые провайдерами Internet

Различные'провайдеры предлагают разные наборы услуг, в зависимости от инфра-щспгруктурьг"их cefeu и от того; ^насколько крупными они являются. В основном провайдеров можно" классифицировать по методам доступа к сети Internet, которые они предлагают, по приложениям й%о услугам безопасности работы в сети.

В последующих раздёлах-мы рассмотрим способы подключения, которые сегодня наиболее широко применяются провайдерами в сети Internet. Вы увидите, какое множество методов доступа предлагается сегодня, начиная от обычного дозвона через модем по телефонной линии из дому (так называемый dial-up) до создания центров обработки данных, когда вы размещаете свое оборудование на узле провайдера и получаете подключение'по локальной сети.

Доступ в Internet по выделенной линии

Как правило, доступ в Internet по выделенной линии предоставляется на скоростях от 56 Кбит/с или 64 Кбит/с до 1,5 Мбит/с и 2 Мбит/с (каналы E1 и T1, соответственно) для конечных пользователей и от 45 и 34 Мбит/с (T3/E3) до 155 Мбит/с (OC3) для субпровайдеров. Сегодня провайдеры, специализирующиеся на предоставлении

доступа по выделенным линиям, предлагают высокоскоростной доступ на скорости 622 Мбит/с (OCИ) и даже 2,5 Гбит/с (OC48). Доступ по выделенной линии предпочтителен, когда имеется прогнозируемая нагрузка на полосу пропускания и частота доступа к сети достаточно интенсивна, чтобы обеспечить загрузку канала 24 часа в сутки. Естественно, стоимость такого подключения довольно высока и, как правило, намного превышает стоимость других методов подключения.

Обычно доступ в сеть Internet по выделенной линии требует завершения физической цепи на оборудовании, принадлежащем клиенту (customer premises equipment — CPE) с одной стороны и замыкания этой цепи на IP-маршрутизаторе провайдера с другой. Протоколы канального уровня, такие как PPP или Cisco HDLC (производный от PPP), используются для передачи кадров и обеспечения сигнализации через существующее соединение. На рис. 2.1 представлена схема типового подключения к сети Internet по выделенной линии.

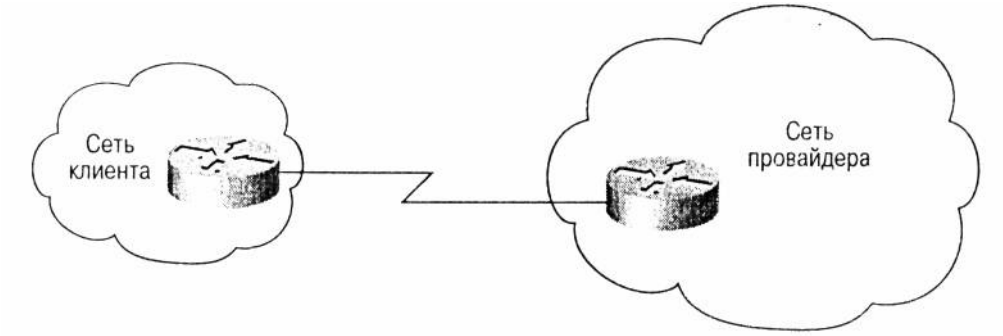


Рис. 2.1. *Доступ в Internet по выделенной линии*

Доступ в Internet с помощью технологий Frame Relay и ATM

Технологии Frame Relay и ATM (Asynchronous Transfer Mode — режим асинхронной передачи) уже на протяжении нескольких лет успешно используются для подключения корпоративных пользователей к Internet. Оплата выделенных линий с подходящей пропускной способностью может оказаться слишком дорогостоящей для большинства компаний. В таком случае наиболее оптимальным решением для них будет доступ по Frame Relay или ATM. С использованием этих технологий компании могут за приемлемую цену получить полосу пропускания, которая будет удовлетворять существующую потребность и даже обеспечит некоторый запас при росте требований к полосе пропускания.

Ввиду того что при организации доступа по Frame Relay и ATM провайдеры могут статически мультиплексировать данные от нескольких абонентов в один поток и затем распределять их по различным IP-сетям, цены на эти услуги значительно ниже, чем оплата выделенного соединения с провайдером. На рис. 2.2 представлена типовая схема подключения к сети Internet посредством Frame Relay.

Доступ в Internet no Frame Relay и ATM в особенности востребован компаниями, в которых уже существуют сети на базе этих технологий. Это также вызвано тем, что довольно часто провайдеры Internet обеспечивают для подобных сетей доступ в свои IP-сети через шлюзы. Таким образом, для поддержания работы таких соединений не требуется дополнительных затрат со стороны клиента.

Несмотря на то что на физическом уровне Frame Relay, ATM и доступ по выделенной линии реализуются по одной и той же технологии, очень важно уловить разницу между Frame Relay, ATM и доступом по выделенной линии. При подключении по Frame Relay и ATM статическое мультиплексирование осуществляется перед подключением к IP-сети. Таким образом, статическое мультиплексирование позволяет сервис-провайдерам объединять потоки данных, что снижает общую стоимость подключения.

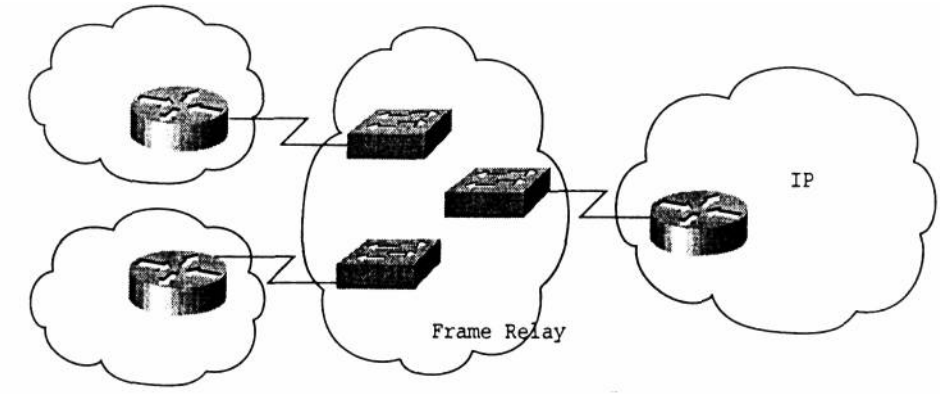


Рис. 2.2. Доступ в Internet по Frame Relay

Очень важно представлять себе размеры трафика в сетях Frame Relay и ATM, который является дополнительной нагрузкой на шлюзы сети Internet. Например, на перегруженном шлюзе (где имеется большое число абонентов) это может привести к значительному снижению производительности на существующих соединениях для доступа в Internet.

Службы доступа по коммутируемым линиям

К службам доступа по коммутируемым линиям относится традиционный доступ с набором номера по телефонной линии посредством модема на скоростях до 56 Кбит/с. К этим службам также можно отнести доступ с помощью цифровой сети с интеграцией услуг ISDN (Integrated Services Digital Network) через базовый интерфейс обмена BRI (Basic Rate Interface) на скорости до 128 Кбит/с и первичный интерфейс обмена PRI (Primary Rate Interface) на скоростях до 1,5 Мбит/с. Службы доступа по коммутируемым линиям используются как отдельными пользователями, так и корпоративными клиентами для получения доступа к сети Internet. За последние несколько лет службы ISDN BRI и PRI получили широкое распространение в основном благодаря возможности предоставления полосы по запросу (т.е. когда в ней возникает потребность), а также возможности передавать цифровые сигналы, применяемые мультимедиа-приложениями, такими как видео в реальном времени и телеконференции.

Цифровые абонентские линии

Услуги цифровой абонентской линии (Digital Subscriber Line — DSL) обеспечивают высокоскоростной доступ в сеть Internet при низкой стоимости. С точки зрения обеспечиваемой скорости и стоимости подключения они находятся между доступом по коммутируемой линии и доступом по выделенной линии. Типы услуг DSL зависят от технологии, на которой они основаны. Для обозначения базовых услуг DSL используется термин *xDSL,* где *x* может представлять собой различные технологии кодирования при передаче цифровых сигналов по физической линии. В табл. 2.1 представлены наиболее распространенные типы технологии DSL и их характеристики.

Таблица 2.1. *Типы DSL*

Тип DSL	Скорость входного потока	Скорость выходного потока	Симметрия	Совместимость с обычной телефонной сетью (POTS)	Стандартизация
ADSL (Asymetrical Digital Subscriber Line)	16 – 640 Kbps	1,5 – 8Mbps	Нет	Да	Да
HDSL (High-bit-rate Digital Subscriber	Фиксирована на 1,544 Mbps	Фиксирована на 1,544 Mbps	Есть	Нет	Да

Line)	и 2,048 Mbps	и 2,048 Mbps			
SDSL (Symmetric Digital Subscriber Line)	1,5 или 2.048 Mbps	1,5 или 2.048 Mbps	Есть	Да	Нет
VDSL (Very high-bit-rate Digital Subscriber Line)	1,6 – 19,2 Mbps (в завис. от протяженности)	12,96 Mbps (до 1,4км) 55,2 Mbps (до 300м)	Есть (может и не быть)	Да	Разрабатывается

Главное преимущество технологии DSL состоит в том, что она может работать на существующей сети кабелей для обычной старой телефонной системы (Plain Old Telephone System — POTS), что делает ее привлекательной для малого и среднего бизнеса. Как правило, услуги DSL, в зависимости от возможностей провайдеров и региона могут значительно отличаться по скорости (от 64 Кбит/с до 52 Мбит/с). Производительность и пропускную способность DSL-соединения в значительной мере определяют качество кабелей и дальность центрального офиса от узла провайдера. В 1999 году в США было введено в действие полмиллиона линий DSL.

Кабельные модемы

Помимо технологии DSL, в настоящее время так же активно развивается новое средство доступа в сеть Internet — кабельные модемы. Кабельные модемы являются прекрасным средством для использования потенциала телевизионных кабельных линий как среды доступа к сети, без дополнительных капиталовложений.

Ввиду того что кабельные модемы были разработаны для применения в существующих сетях кабельного телевидения на базе коаксиального и оптоволоконного кабелей, которые оптимизированы для передачи односторонних широковещательных сигналов (от станции к телевизионным приемникам), то доступная для работы полоса пропускания по своей природе имеет значительную асимметрию. Например, типовые услуги доступа предполагают пропускную способность до 2 Мбит/с (от узла провайдера до абонента) и до 64 Кбит/с в обратном направлении (от абонента до узла провайдера).

В отличие от DSL, которая является по сути технологией по обеспечению соединений типа "точка-точка", выходной поток может совместно использоваться несколькими пользователями, что создает определенную угрозу безопасности как для производителей и сервис-провайдеров, так и для клиентов.

Несмотря на все эти трудности доступ с помощью кабельных модемов предоставляется уже на протяжении нескольких лет. Количество абонентов и провайдеров, предоставляющих доступ по кабельным модемам, стремительно растет. На сегодняшний день *в США* насчитывается около 2 миллионов абонентов, подключенных к Internet посредством кабельных модемов, а к концу 2003 года ожидается прирост подобных подключений до 16 миллионов.

Услуги выделенного хостинга

Хотя хостинг (хранение данных на узле провайдера — *Прим. ред.)* как услуга предоставляется уже достаточно долгое время, за последние годы он приобрел огромную популярность Появились даже провайдеры, которые работают именно на этом сегменте рынка (предоставление услуг по размещению информации на Web-узлах). Крупные провайдеры, которые специализируются на предоставлении услуг выделенного хостинга, обычно именуются *контент-провайдерами (content-provider).* Эти провайдеры, как правило, обеспечивают рачвертывание и поддержку высоконадежных отказоустойчивых информационных центров (data center), в которых находится оборудование для предоставления услуг Web-хостинга. В этих помещениях клиенты могут либо размещать свое оборудование, либо арендовать его у провайдера. Затем провайдеры продают услуги

доступа в Internet, подключая это оборудование локально к своей сети на базе Fast Ethernet (100 Мбит/с) или Gigabit Ethernet (1 Гбит/с). При такой организации доступа к Internet реализуются различные схемы оплапы услуг провайдера — как на основе фиксированных тарифов, так и в зависимости от частоты использования той или иной услуги.

Для объединения трафика от сотен, а иногда и тысяч серверов на узлах хостинг-провайдеров часто применяются высокопроизводительные коммутаторы Ethernet. Клиенты должны придерживаться заявленной полосы пропускания и механизмов защиты от отказов, которые применяются на узле провайдера. Кроме того, в целях обеспечения безопасности в крупных коммутируемых сетях клиенты должны представлять себе, каким образом на узле провайдера осуществляется разделение широковещательных доменов в роли которых обычно выступают виртуальные локальные сети. В коммутируемой сети с совместным доступом, общей для модели контент-хостинга, понимание этих механизмов является ключевым условием для предотвращения возможных атак типа "отказ в обслуживании" (Denial of Service — DoS), несанкционированного доступа к данным и лругих проблем, связанных с работой системы.

Услуги хостинга становятся все более привлекательными и уже сегодня являются сегментом рынка с многомиллиардным оборотом. Однако клиенты должны быть весьма осторожны и представлять себе, какие, кто и каким образом обеспечивает услуги. Более детально о коммутаторах, виртуальных локальных вычислительных сетях (Virtual Local Area Networks —VLAN) и широковещательных доменах читайте в книгах *"Межсетевое оборудование: мосты, маршрутизаторы, коммутаторы и межсетевые протоколы" 2-е изд., ("Intcnvnturfions: Bridges, Routers, Switches and Internetworking Protocols", 2nd edition., Addison-Wesley, 1999)* Рэдии Перлман (Radia Perlman) или *"Коммутация ЛВС с помощью оборудования компании Cisco" ("Cisco LAN Switching", Cisco Press, 1999)* Кеннеди Кларка (Kennedy (Lark) и Кевина Гамильтона (Kevin Hamilton).

Другие услуги, предоставляемые провайдерами Internet

К другим услугам, которые также широко используются клиентами, относятся электронная почта, службы новостей, виртуальные частные сети (Virtual Private Networks — VPN) и организация многоадресных IP-запросов. Желая получить эти услуги, клиенты должны взвесить все “за" и "против" их использования. В особенности мы хотели бы обратить ваше внимание на то, как эти услуги предоставляются, а также на уровень знаний технического персонала, необходимых для поддержания работы подобных служб.

Большинство ISP предоставляют услуги консультативного характера и другие услуги, такие как, например, обеспечение безопасности сети. Простейшие мероприятия по обеспечению безопасности включают в себя фильтрацию пакетов, проходящих через устройство доступа к сети. Сейчас также распространены услуги по шифрованию данных и защите от компьютерных вирусов.

Цены на предоставляемые услуги значительно разнятся в зависимости от надежности метода доступа к сети на узле провайдера (об этом более подробно в следующем разделе). Кроме того, цены сильно зависят от объемов капиталовложений, сделанных провайдером в инфраструктуру, оборудование и персонал.

Цены на услуги ISP, соглашения об уровне обслуживания и технические характеристики

Оценив доступные услуги, клиенты при выборе провайдера должны также проанализировать цены и технические характеристики, предлагаемых услуг. Хотя технические характеристики кажутся немного пугающими, тем не менее их необходимо знать, хотя бы для общения с провайдером, которого вы уже, вероятно, выбрали. Технические вопросы, рассматриваемые в этом разделе, включают в себя характеристики магистральных каналов, разграничение сетей и хостинг.

Цены на услуги, предоставляемые ISP

Цены на одни и те же услуги могут значительно отличаться у разных провайдеров, даже если они находятся в одном регионе. Довольно часто авторитет провайдера или объем сделанных им капиталовложений являются главными факторами, определяющими конечную цену услуги. Так, например, провайдер, который продает доступ по Frame Relay, скорее всего предложит вам подключение по более привлекательной цене, чем начинающая компания-провайдер. С другой стороны, новый провайдер будет более конкурентоспособным, так как он не вкладывал деньги в устаревшую инфраструктуру, а предоставшет свои услуги на базе нового оборудования с более широкими возможностями.

Таким образом, выбор провайдера лишь на основании цены не совсем оправдан, так как за одну и ту же цену у разных провайдеров вы получите различные услуги, как по качеству, так и по содержанию. Так, например, доступ по выделенной линии некоторыми провайдерами организуется с использованием оборудования на стороне клиента (Customer Pertinent Equipment — CPE) (более подробно об этом далее в этой главе), такого как маршрутизатор и устройства сопряжения с цифровыми каналами CSU/DSU (Channel Service Unit/Data Service Unit). Другие провайдеры могут потребовать отдельной платы за установку CPE или адаптировать его для работы со своим оборудованием, что может стоить намного дороже самой услуги по подключению. В этом случае, возможно, дешевле оплатить услуги провайдера по закупке и/или обслуживанию CPE.

Крупные компании часто закупают подключение к сети Internet и другие телекоммуникационные услуги у одного провайдера. Такого рода интегрированные решения от одного провайдера обычно обеспечивают лучшее управление и координацию действий по предоставлению различных услуг в одной сети по регионам. Некоторые провайдеры предлагают объединенные пакеты услуг, предоставляемых как в пределах одной страны, так и на мировом рынке, и довольно часто предоставляют клиентам, закупающим пакеты услуг, значительные скидки. Например, если организация закупает междугородные переговоры и доступ в Internet. Такое объединение счетов упрощает бухгалтерский учет (выписывается один счет и оплачивается один счет), что для большинства компаний является очевидным преимуществом. Конечно, если удобство объединенного счета за услуги не является важным фактором для вашей компании, то вполне можно найти и другие предложения от провайдеров Internet.

Соглашения об уровне обслуживания

Сегодня большинство провайдеров предоставляют свои услуги на основе

соглашений о гарантиях и уровне обслуживания (Service-Level Agreements/Service-Level Guarantees — SLA/SLG), в которых определяются критерии для гарантированной производительности и доступности использования различных услуг. При выборе провайдера следует убедиться, что в этих соглашениях четко описана ответственность сторон в случае сбоев и простоев системы и др. Осведомитесь у провайдера, каким образом реализуются гарантии и генерируются ли автоматически сообщения при отказах оборудования со стороны клиента или со стороны провайдера, или уведомление об отказах в работе оборудования должно исходить непосредственно от заказчика.

В соглашениях обычно указывается допустимый уровень потерь пакетов и обязательное время задержки при доступе к сети провайдера, а также вопросы обслуживания и/или уведомления о простоях каналов связи.

Финансовая ответственность, закладываемая в соглашениях с провайдером, может стать реальным рычагом обеспечения заявленного качества услуг, однако идентификация нарушений соглашения и сбор штрафов могут оказаться для неподготовленного клиента непосильной задачей.

Критерии выбора магистральных каналов ISP

Сеть магистральных каналов ISP имеет несколько довольно важных технических характеристик.

- Физическая топология сети.
- "Бутылочные горлышки" в сети и их описание.
- Уровень отказоустойчивости элементов сети и сети в целом.
- Межсетевые соединения с другими сетями, включая расстояние между узлами и соглашения об обмене трафиком.

Этот раздел предназначен как для пользователей, так и для разработчиков сетей ISP. Клиенты должны при выборе провайдера оценить все нижеприведенные характеристики. Это даже более важно, чем цены на услуги. При расширении существующих или построении новых сетей инженерам следует проанализировать все преимущества и недостатки, связанные с этими характеристиками.

Физические соединения

Клиентам следует ознакомиться с топологией сети провайдера, а провайдер должен предоставить для этого карту своей сети со всеми соединениями и последними изменениями. Что касается соединений, правильная физическая топология сети — один из важнейших факторов, обеспечивающих адекватное разделение полосы пропускания в зависимости от направления трафика, даже в том случае, если одно или несколько соединений выйдут из строя. Существование у провайдера магистральных подключений на базе каналов OCИ и OC48 само по себе не гарантирует высокоскоростной доступ для конечных пользователей. Ваш трафик может поступать в сеть провайдера через низкоскоростное магистральное соединение или через высокоскоростное, но перегруженное абонентами соединение. Все эти факторы влияют на качество соединения.

Потенциальные "бутылочные горлышки" на узлах ISP и коэффициенты абонентской нагрузки

Мощность сети провайдера определяется ее самым слабым соединением. Потенциально на узлах провайдеров существуют два "бутылочных горлышка" (bottlenecks), или, другими словами, "узких места", которые создают задержки при прохождении трафика, — это превышение допустимого числа абонентов на магистральных каналах (так называемый oversubscription) и образование шлейфов при работе в точках присутствия (Points of Presence — POP) или от узла провайдера в сторону клиента. Провайдерам следует быть осторожными и не перегружать свои соединения. Провайдеры, пытающиеся сэкономить деньги путем повышения нагрузки на свои маршрутизаторы и каналы, закончат

тем, что потеряют доверие клиентов на долгое время.

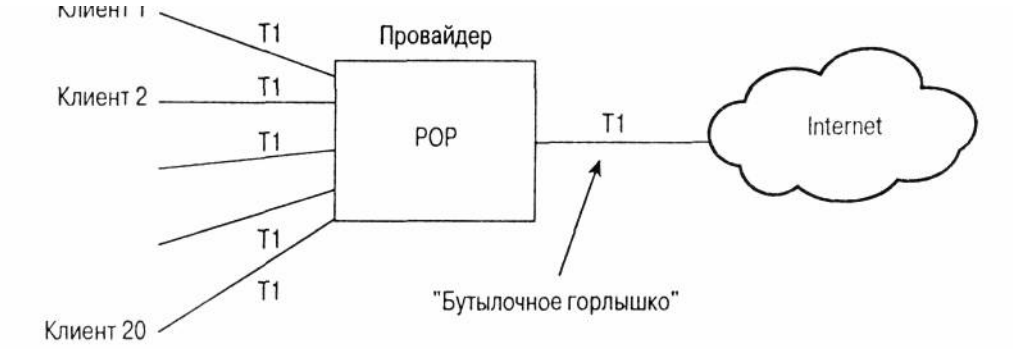


Рис. 2.3. Производительность каналов ISP определяется самым слабым соединением

Превышение допустимого числа абонентов происходит в том случае, когда совокупная нагрузка соединений превышает имеющуюся полосу пропускания. Так, провайдер, продавая 20 каналов типа TI в POP и подключаясь, в свою очередь, к NAP тоже по каналу T1, создает так называемое "бутылочное горлышко". Как видно из рис. 2.3 при примерном подсчете получим соотношение 5:1, т.е. в этом случае на одно магистральное соединение типа T1 нельзя подключать более пяти каналов T1. Коэффициенты абонентской нагрузки (отношение совокупной полосы пропускания подключаемых каналов к полосе пропускания выходного канала) могут значительно варьироваться в зависимости от предлагаемых услуг. Как правило, провайдеры, предоставляющие услуги выделенного хостинга, пользуются соотношениями 8:1 и даже 10:1. Эти значения обычно основаны на практическом опыте и проектной нагрузке, однако если они небрежно выбраны или не соблюдаются, то это может привести к перегрузке канала.

Рис. 2.4. Скорость доступа ограничивается наименьшей полосой пропускания

В качестве еще одного примера создания эффекта "бутылочного горлышка" можно привести ситуацию, когда системы с высокоскоростными каналами пытаются получить доступ к системам, которые подключены через низкоскоростные соединения.

Так, доступ к Web-серверу расположенному на узле, подключенном к Internet со скоростью 56 Кбит/с, может осуществляться лишь на этой скорости, независимо от того, с какой скоростью работает с Internet пользователь. На рис. 2.4 показано, как клиент, имеющий доступ в Internet по каналу типа T3, при подключении к Web-серверу будет работать лишь со скоростью 56 Кбит/с, т.е. со скоростью, на которой этот сервер подключен к Internet. Обратите внимание, что и другие пользователи, получающие одновременно доступ к этому серверу также будут ограничены скоростью 56 Кбит/с, причем все они должны будут совместно использовать эту полосу пропускания.

Очень важно, чтобы провайдеры вели постоянный мониторинг соединений и управляли нагрузкой в своих сетях. Перед приобретением услуг клиентам следует задать потенциальному провайдеру следующие вопросы.

- Каким образом осуществляется управление нагрузкой на каналы?
- При каких граничных условиях подключаются резервные системы?

- Какие типовые коэффициенты абонентской нагрузки (отношение доступной производительности к используемой производительности) установлены для той или иной услуги?
- Какие типовые коэффициенты абонентской нагрузки установлены для магист ральных каналов и узлов?
- Теоретическое ограничение пропускной способности для данного вида услуги?

Уровень резервирования оборудования ISP

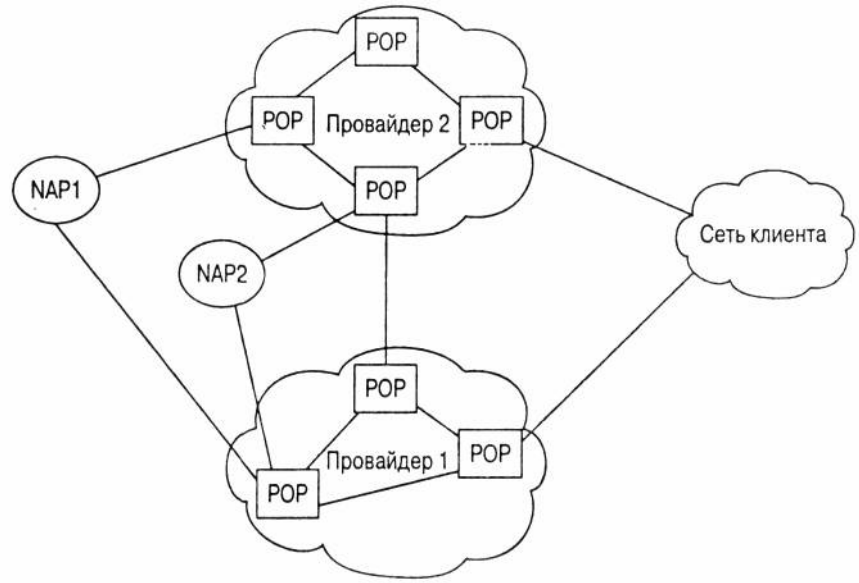


Рис. 2.5. *В сети с резервированием обеспечивается более высокая надежность соединений*

Итак, Мэрфи уже здесь и готов сделать вашу жизнь ужасной. В силу плохой погоды либо проблем с карьерой, или просто из-за неудачного дня соединение вашего провайдера с NAP, с другим провайдером или с другой POP в какой-то момент времени пропадает, что приводит к невозможности достичь всех или определенного списка узлов. Сеть с дублирующим оборудованием позволяет в аварийных ситуациях направлять трафик по альтернативным маршрутам до тех пор, пока неполадка не будет устранена. Правильно сконструировать сеть ISP — весьма ответственная задача. Наиболее оптимальной считается схема, при которой POP провайдера подключаются к нескольким NAP и сетям других провайдеров, а также к POP других провайдеров, как показано на рис. 2.5.

Важно понимать, что резервирование одноранговых соединений и узлового оборудования обычно производится глобально, т.е. во всех сетях сразу. Другими словами, если соединение с провайдером становится недоступным через основную точку обмена трафиком, то выбирается следующая ближайшая точка обмена трафиком. Таким образом, основная идея заключается не в обеспечении резервирования оборудования на одном узле, а в создании резервных соединений и магистральных каналов для быстрого реагирования на один или несколько отказов на различных узлах сети. При таком подходе обеспечение избыточных межсетевых соединений и соединений с ближайшими географически NAP вполне может возместить затраты на резервирование. В результате вы получите более надежную сеть, чем при реализации резервирования лишь между провайдерскими POP. На рис. 2.6 представлена схема неоптимального соединения между двумя провайдерами, а на рис. 2.7 — схема соединения с полным резервированием.

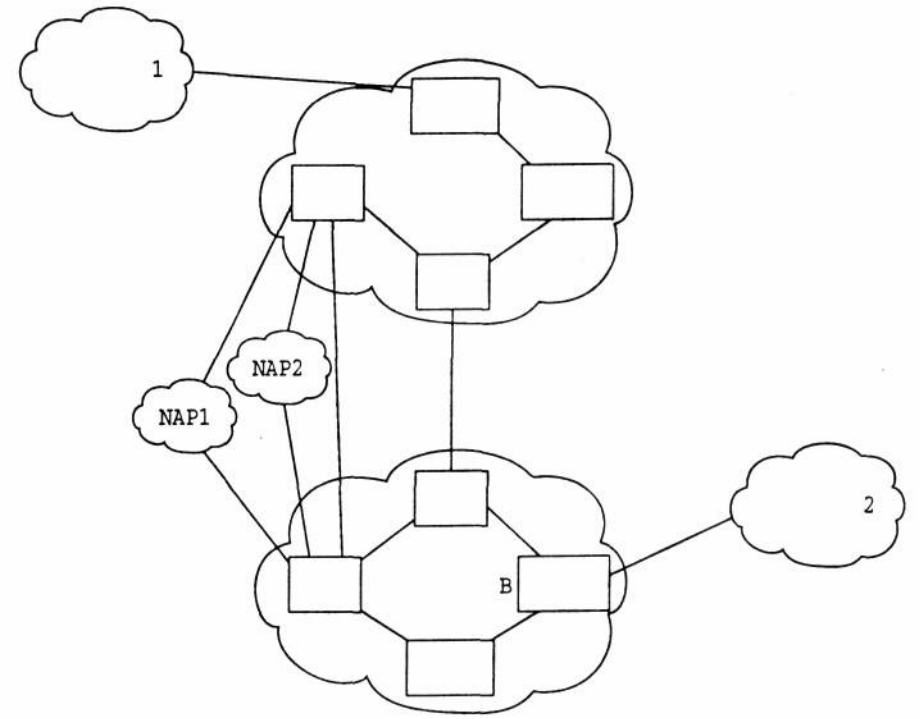


Рис. 2.6. Неоптимальное соединение провайдеров

Следует обсудить с провайдером план резервирования. Большинство провайдеров имеет на своих рабочих площадках определенный запас критичного к сбоям оборудования. Количество запасных компонентов обычно зависит от критичности того или иного оборудования, а также от времени наработки на отказ (Mean Time Between Failures — MTBF).

Некоторые провайдеры выбирают резервирование отдельных услуг для географически близких дилеров, которые имеют сети хранилищ и предоставляют свои площади клиентам. Хотя такой подход увеличивает гарантированное время до первого ремонта (Mean Time To Repair — MTTR), при возникновении проблем он все-таки лучше, чем отсутствие какого-либо резервирования вообще. Иногда возникает дефицит наиболее востребованных компонентов оборудования, что также при отсутствии плана резервирования скажется на сроках восстановления нормальной работоспособности.

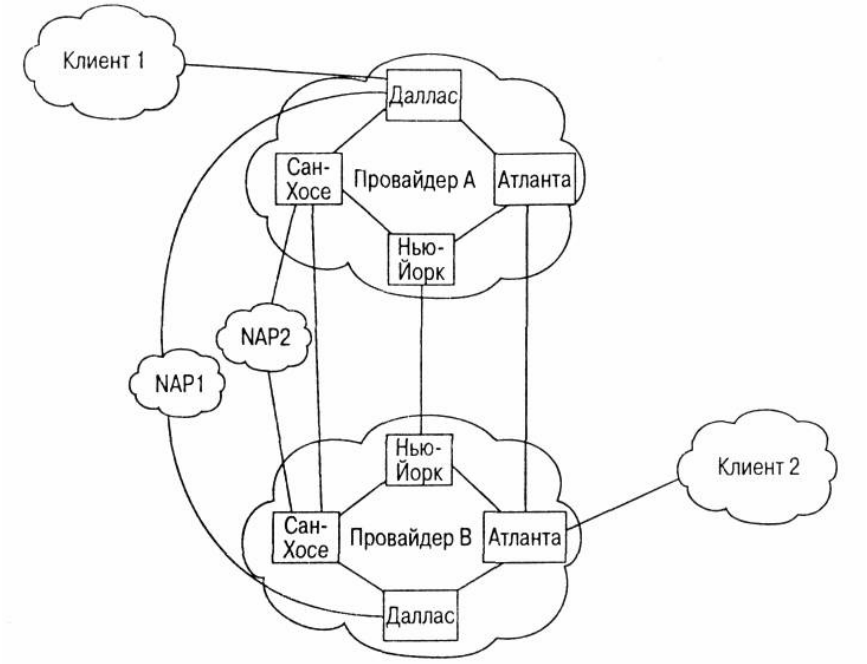


Рис. 2.7. *Соединение провайдеров с полным резервированием*

Расстояние до удаленных узлов

Типичное заблуждение заключается в том, что единственный вопрос, который должен волновать клиентов — это количество переприемов IP-пакетов *(IP hops),* т.е. количество маршрутизаторов, через которые должен пройти пакет, чтобы достичь нужной сети. В прошлом это обстоятельство в некоторой степени действительно оказывало влияние на прохождение IP-пакетов, так как чем больше переприемов, тем больше была задержка в каждом из них, возрастала также вероятность неправильной маршрутизации пакетов и даже их потери. Однако сегодня большинство магистральных сетей провайдеров строятся на основе технологий мультипротокольной коммутации (Multiprotocol Label Switching — MPLS), ATM или Frame Relay. Благодаря этому большинство узловых устройств прозрачны для средств определения IP-маршрутов, таких как *traceroufe.*

Меньшее количество переприемов IP-пакетов для заданной сети может свидетельствовать о более коротком маршруте к пункту назначения, чем путь через сеть с большим количеством переприемов. Однако знание технологий, на которых строятся межсетевые соединения, поможет вам принять правильное решение. Например, возможно более эффективным будет воспользоваться несколькими сквозными высокоскоростными каналами, чем одним низкоскоростным.

Как вам уже известно, сеть Internet представляет собой конгломерат наложенных друг на друга магистральных сетей, соединенных точками обмена трафиком и посредством прямых межсетевых соединений. Справедливо оценивать количество *сетей* или число *переприемов между AS* (количество пересекающихся доменов маршрутизации) для заданного набора пунктов назначения. Расстояние до удаленных узлов будет зависеть от того, сколько каналов закупают у провайдера администраторы удаленных сетей и насколько хорошо провайдер соединен с другими сетями, т.е. насколько развита инфраструктура его сети. Мелкие компании-провайдеры могут быть подключены только к одной NAP или вообще не иметь соединения ни с одной из них. Более крупные провайдеры часто подключаются к другим сетям как посредством NAP, так и с помощью прямых межсетевых соединений.

Соглашения об обмене трафиком

При выборе провайдера непременным условием является наличие у него двусторонних соглашений об обмене трафиком с другими провайдерами на равноправной основе. Существующая сегодня архитектура сети Internet и незначительное вмешательство в регулирование работы сети по вопросам, кто и каким образом должен соединяться (напрямую или через NAP), вынуждает провайдеров решать эти вопросы самостоятельно.

Многие годы провайдеры вынашивали идеи относительно урегулирования межсетевого взаимодействия, но в вопросах о том, кто, кому и за что должен платить, не достигали консенсуса. Как отмечалось в главе 1, "Эволюция сети Internet", наиболее крупные провайдеры начали постепенно переходить к модели распределенных межсетевых соединений, используя NAP для подключения только мелких провайдеров. Более крупные провайдеры ужесточили также режим обмена трафиком в NAP, т.е. они соглашались организовывать межсетевые соединения в NAP только с равноценными компаниям и-провайдерами. Эти положения фиксируются в так называемом совместном соглашении о неразглашении (Nondisclosure Agreement — NDA), которое заключается обеими сторонами.

Хотя потенциальные провайдеры скорее всего не будут столь строги насчет заключения специфических соглашений об обмене трафиком со всеми сетями, они лишь потребуют сведения о количестве задействованных адресов и техническую информацию о межсетевых соединениях и правилах их взаимодействия. То, каким образом провайдер соединяется с другими сетями, может быть наиболее важным с точки зрения производительности соединения, которое вы будете оплачивать.

Точка демаркации

Кроме ценовой политики, магистральных каналов и способов межсетевых соединений, следует также обсудить с провайдером вопрос о точке демаркации (demarcation point — DP). *Тонка демаркации* — это точка, в которой происходит разделение сети на зоны ответственности провайдера и клиента (или клиентов). Частично это определение верно и для провайдера, предоставляющего услуги выделенного хостинга. Очень важно определить и правильно понимать различия между зоной ответственности провайдера и клиента. Точки демаркации определяются вплоть до конкретных кабелей и разъемов для того, чтобы избежать конфликтных ситуаций при возникновении каких-либо проблем с оборудованием или в работе сети. На рис. 2.8 представлена типовая точка демаркации между сетью провайдера и сетью клиента.

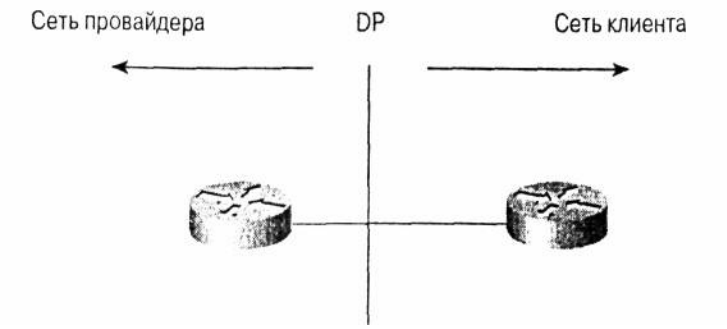


Рис. 2.8. Точка демаркации

Провайдеры по-разному определяют точку демаркации. Как правило, выбор точки демаркации зависит от того, кто платит за оборудование и за канал доступа, где находится оборудование и кто проводит его обслуживание.

Оборудование принадлежащее клиенту

Оборудование, принадлежащее клиенту (Customer Premises Equipment — CPE), обычно включает в себя маршрутизатор, устройство сопряжения с цифровым каналом CSU/DSU, систему кабелей и иногда аналоговый модем для мониторинга и управления вышеуказанным оборудованием вне основной полосы (out-of-bandwidth — OOB). Обычно провайдер предлагает клиенту выбор в приобретении CPE и канала для доступа к своему узлу. Так, вы можете оплатить только канал доступа или же выплачивать ежемесячно набор услуг, в который входит аренда и обслуживание всего оборудования и канала доступа, но при этом все эти задачи выполняются персоналом провайдера. Практически всегда удается достичь с провайдером хорошего соотношения цена/качество для предоставляемых услуг. Обычно ISP несет ответственность за обслуживание оборудования или за качество предоставляемых с его помощью услуг. Как правило, провайдеры заранее формируют различные пакеты услуг, в которые входит и обслуживание CPE. Если же клиент не желает приобрести стандартный пакет услуг, то ему придется выбрать оборудование, заранее одобренное провайдером. После этого покупатель будет самостоятельно нести ответственность за обслуживание и решение технических проблем, возникающих при эксплуатации этого оборудования. Провайдер всегда окажет помощь в решении технических проблем, но уже за отдельную плату.

На рис. 2.9—2.11 представлены схемы предоставления стандартных пакетов услуг.

На схеме, представленной на рис. 2.9, провайдер отвечает за соединительную линию и за CSU/DSU вплоть до разъема CSU на технической площадке клиента. В этом случае провайдером могут также предъявляться определенные требования к принадлежащему клиенту маршрутизатору, в частности к объему памяти и версии операционной системы.

На схеме, представленной на рис. 2.10 провайдер предоставляет все оборудование, и зона его ответственности заканчивается портом локальной вычислительной сети на маршрутизаторе, который расположен у клиента.

На рис. 2.11 представлена схема, когда клиент предоставляет и CPE, и соединительную линию. В этом случае ответственность провайдера заканчивается на коммутационном шкафу, расположенном на одной из POP провайдера, где осуществляется

коммутация технической площадки провайдера с центральным офисом.

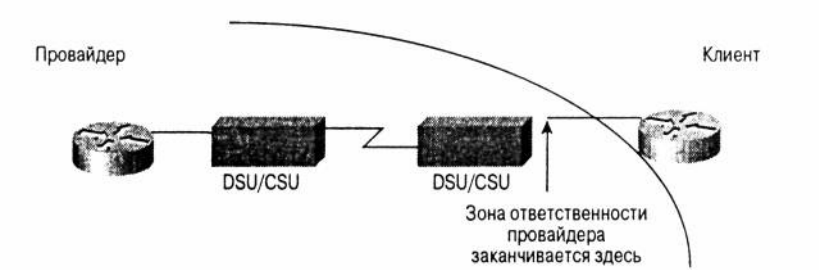


Рис. 2.9. *Провайдер обеспечивает доступ до CSU/DSU; клиент предоставляет маршрутизатор (пример)*

Расположение маршрутизаторов

Расположением (или монтажом) (collocation) называют действия по размещению оборудования одной стороны на технической площадке другой стороны. В качестве примера можно привести установку маршрутизатора клиента на технической площадке провайдера или в хостинг-центре, как показано на рис. 2.12. Стимулирующими факторами для расположения клиентом своего оборудования на узле провайдера является более быстрый доступ и упрощение мониторинга оборудования со стороны провайдера или предоставление клиенту возможности более гибкого управления нагрузкой на полосу пропускания.

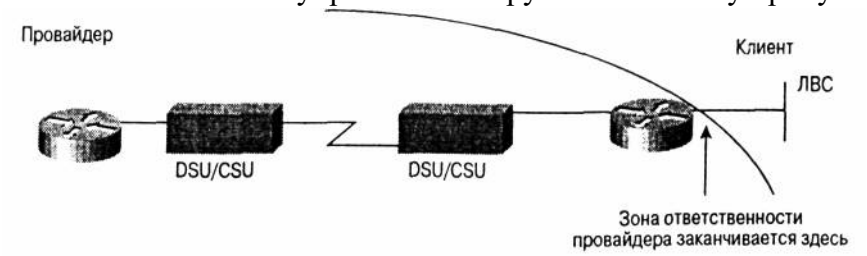


Рис. 2.10. Провайдер предоставляет доступ, маршрутизатор и CSU/DSU (пример)

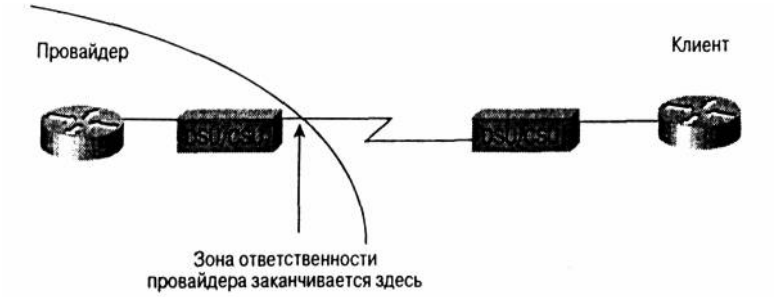


Рис. 2.11. Все оборудование предоставляется клиентом (пример

Рис. 2.12. Размещение маршрутизатора на технической площадке провайдера (пример)

Ситуация, обратная показанной на рис. 2.12, — размещение провайдером своего маршрутизатора на площадях клиента (рис. 2.13). В таком случае ISP оплачивает и маршрутизатор, и соединительную линию, а клиент выплачивает полную стоимость услуги по подключению.

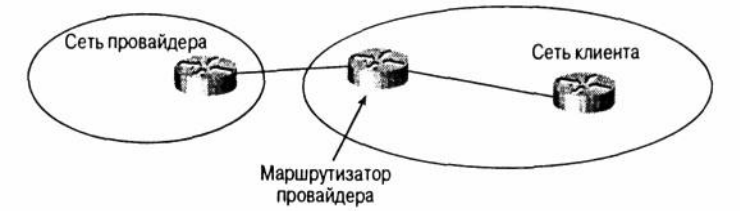


Рис. 2.13. *Размещение провайдером своего маршрутизатора у клиента (пример)*

Забегая вперед

Технические характеристики сети провайдера определяют в значительной степени качество услуг, предоставляемых клиентам, и оптимальность маршрутизации. Поскольку клиент не может в полной мере контролировать все технические параметры, он должен уметь оценивать самые важные из них и определять, соответствуют ли они заданным требованиям к структуре и качеству обслуживания.

Если вы являетесь клиентом какого-либо провайдера и оговоренная точка демаркации вменяет вам в обязанности обслуживание собственного оборудования, даже если вы только арендуете его, то вы должны взять на себя вопросы развития сети, а также разработки и реализации правил маршрутизации для нее. Даже если не вы устанавливаете и не обслу живаете сетевое оборудование, все равно вам придется принимать определенные решения, для чего необходимо иметь представление об организации маршрутизации в сети.

В следующей главе обсуждением фундаментальных положений IP-адресации и рс зервировании адресного пространства завершается формирование теоретической базы После этого во второй части книги мы перейдем к рассмотрению различных протоко лов маршрутизации.

Часто задаваемые вопросы

В —Говорит ли высокая цена на услуги провайдера о том, что я получу более быстрое и качественное соединение с провайдером?

О - Нет, это не совсем так. Иногда более высокая цена означает, что провайдер обладает собственными высокоскоростными магистральными каналами типа ОС12 или ОС48, в которые он вложил свои деньги. Однако, наличие этих каналов отнюдь не означает, что вы получмте более быстрое соединение. При неправильной комбинации высокоскоростных соединений с низкоскоростными может произойти общее падение производительности в сети провайдера и даже в связанных с ней сетями. В конечном итоге цена – лишь один фактор из многих, на которые следует опираться при выборе провайдера. Более важно узнать побольше о топологии сети провайдера, о том в достаточной ли степени зарезервировано в ней оборудование и соответствует ли предлагаемая полоса вашим нуждам.

В — *Что может привести к появлению "бутылочных горлышек" на магистральных каналах провайдера?*

О – Как правило, появление "бутылочных горлышек” обусловлено превышением допустимого предела количества абонентских линий или перегрузкой самой полосы пропускания

В — *Следует ли мне при подключении к провайдеру приобретать собственное оборудование?*

О — Существуют несколько "за" и "против” такого решения, и принимать его следует лишь исходя из потребностей вашей организации. Выясните, во-первых, будет ли провайдер настаивать, чтобы вы работали исключительно с его оборудованием (некоторые

требуют этого). Даже если провайдер позволит вам закупить собственное CPE, то, вероятнее всего, он потребует определенной аппаратной и программной конфигурации этого оборудования, которая бы согласовалась с его системой. Оплата всех этих мероприятий будет целиком возлагаться на вас. Сможет ли ваша организация позволить себе подобные капиталовложения, включая модернизацию и расширение парка оборудования? Приобретая подобные устройства, вы также обрекаете себя на проведение их технического обслуживания, хотя некоторые провайдеры и берут на себя эти функции (за отдельную;плату).

В большинстве стран принимаются специальные законы, которые ограничивают спектр используемого провайдерами оборудования. Принимая решение вы должны учитывать и этот аспект.

В - *Если соединение с провайдером разрывается по причине аппаратного сбоя, кто несет за это ответственность ?*

О — Все зависит от комплекса услуг, который вам предоставляет провайдер. Границы ответственности между вами и провайдером определяются заранее при установлении точки демаркации

Ключевые темы этой главы:

- **Обзор системы адресации** в Internet. Дается обзор системы адресации IPv4, адресов классов A, B и C, а тарже рассматриваются основные концепции разбиения сетей на подсети.
- **Маски подсети переменной длины (Variable-length subnet masks—VLSM).** Описаны маски подсети и их применение при назначении адресного пространства IP.
- **Исчерпывание пространства IP-адресов. Обс**уждается проблема нехватки IP-адресов. Освещаются также вопросы распределения адресов на современном этапе, выделение адресов для крупных сетей, частных лиц, а также обсуждаются протоколы следуюшего поколения.
- **Адресация в частных сетях и преобразование сетевых адресов (Network Address Translations** - NAT). Рассматриваются вопросы использования программного обспечения с функциями NAT для преобразования адресов при выходе из частных сетей в глобальные IP-сети.
- **IP версии 6 (IPv6).** Дается обзор следуюшего поколения системы адресации (IPng), а также рассматривается ее соответствие иерархической модели, сформировавшейся на базе бесклассовой междоменной маршрутизации (classless interdomain routiftg.-CIDR) и IPv4.

Глава 3. IP адресация и методы распределения адресов

Эту главу мы начнем с краткого экскурса в историю возникновения системы адресации в сети Internet, подробно рассмотрим традиционную 4-ю версию адресации IP (Ipv4) и схемы разделения сетей на подсети. В этой главе вы узнаете о проблеме истощения адресного пространства в сети Internet. Затем мы рассмотрим несколько методов назначения и распределения IP-адресов, а также технические приемы, такие как маскирование сетей с помощью масок пёременной длины (Variable-Length Subnet Mask — VLSM), бесклассовая междоменная маршрутизация (Classless Interdomain Routing — CIDR) и преобразование сетевых адресов (Network Address Translation — NAT). И в заключение мы приводим в этой главе общие сведения об IP версии 6 (IPv6).

Для любой сети правила распределещя адресов являются фундаментальным вопросом при организации маршрутизации. Одна до основных функций системы маршрутизации и маршрутизаторов — обеспечить адресами все точки прохождения трафика. При стремительном росте сети Internet нехватка адресного пространства и появление новых систем адресации испытывают существующую структуру маршрутизации на прочность. Знание истории развития и основ адресации в IP-сетях, без сомнения, будет играть ключевую роль в усвоении новых концепций, закладываемых сегодня в протоколы маршрутизации.

История развития системы адресации в Internet

Система адресации, применяемая сегодня в сети Internet, основана на протоколе IP версии 4 (IPv4)1, который, как правило, называют просто *IP*. В этом разделе обсуждаются:

Основы IP-адресации.

Основы формирования подсетей на базе IP2.

Маски подсети с переменной длиной (VLSM)3.

Основы IP-адресации

Итак, IP-адрес представляет собой уникальную четырехоктетную (32-битовую) величину, выраженную в десятичных числах, разделенных точками, в форме W.X.Y.Z, где точки используются для разделения октетов (например, 10.0.0.1). Поле адреса размером 32 бита состоит из двух частей: *адрес сети* или *связи* (который представляет собой сетевую часть адреса) и *адрес хоста* (идентифицирующий хост в сетевом сегменте).

Разграничение сетей по количеству хостов в них традиционно осуществляется на

основе так называемых классов IP-адресов. Сегодня существует 5 классов IP-адресов (три из которых используются для уникальной адресации сетей и хостов): A, B, C, D и E. Все они представлены в табл. 3.1.

Таблица 3.1. *Классы IP-адресов и их функции*

Класс	Диапазон адресов	Старшие биты	Биты для обозначения сети	Биты для обозначения хоста	Функция
A	0.0.0.0 – 127.255.255.255	0	7	24	Уникальные адреса
B	128.0.0.0 – 191.255.255.255	10	14	16	Уникальные адреса
C	192.0.0.0 – 223.255.255.255	110	21	8	Уникальные адреса
D	224.0.0.0 – 239.255.255.255	1110			Многоадресное обращение
E	240.0.0.0 – 255.255.255.255	1111			Зарезервировано

Обратите внимание, что только адреса классов A, B и C могут использоваться как уникальные. Адреса класса D применяются для обращения к набору узлов, а адреса класса E зарезервированы для исследовательских целей и в настоящее время не используются. Несколько адресов во всех классах зарезервированы для специальных целей. Некоторые из них представлены в табл. 3.2.

Таблица 3.2. *IP-адреса, зарезервированные для специальных целей*

Диапазон адресов	Назначение
0.0.0.0	Неизвестная сеть (обычно представляет сеть по умолчанию)
10.0.0.0 - 10.255.255.255	Зарезервировано для частных сетей (RFC 1918)
127.0.0.1 - 127.255.255.255	Зарезервировано для локальных адресов типа "петля"
172.16.0.0-172.31.255.255	Зарезервировано для частных сетей (RFC 1918)
192.168.0.0 - 192.168.255.255	Зарезервировано для частных сетей (RFC 1918)
192.168.255.255 – 255.255.255.255	Широковещательный запрос

Такая система адресации, основанная на классах, часто именуется *классовой моделью (classful model).* Различные классы определяются также различными конфигурациями сетей, в зависимости от желаемого количества подсетей в сети и числа хостов в них. По мере рассмотрения материала в этой главе станут более четко видны различия между классами IP-адресов. В последующих разделах мы подробно остановимся на характеристиках каждого класса.

Адреса класса A

Сети класса A определяются значением 0 самого старшего (левого) бита в адресе. Первый октет (биты с 0 по 7), начинаются с левого бита в адресе. Этот октет определяет количество подсетей в сети, в то время как оставшиеся три октета (биты с 8 по 31) представляют количество хостов в сети. Возьмем для примера адрес в сети класса A 124.0.0.1. Здесь 124.0.0.0 представляет собой адрес сети, а единица в конце адреса обозначает первый хост в этой сети. В результате такого представления (рис. 3.1) в сетях класса A можно адресовать 128 (27) подсетей. Однако ввиду того, что адрес 0.0.0.0 не является нормальным адресом сети, то реально в сетях класса A доступно только 127 (27-1) адресов.

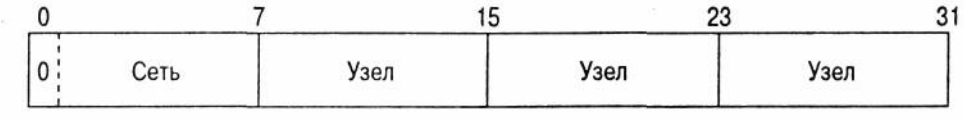


Рис. 3.1*. Общий вид IP-адреса класса A*

После определения сети, первый и последний адреса хостов в ней выполняют

специальные функции. Так, первый адрес 124.0.0.0 (из приведенного выше примера) используется в качестве адреса сети, а последний адрес (124.255.255.255) представляет собой широковещательный адрес для этой сети. Таким образом, с помощью адресов класса A можно представить только 16777214 (216—1) хостов в каждой сети, а не 16777216 (224).

Адреса класса B

Сети класса B определяются значениями 1 и 0 в старших битах адреса. Первые два октета в адресе (биты с 0 по 15) служат для представления адресов сетей, а оставшиеся два октета (биты с 16 по 31) представляют номера хостов в этих сетях. В результате мы получим 16384 (214) адреса сетей с 65534 (216—2) хостов в каждой (рис. 3.2). Так, например, в адресе класса B вида 172.16.0.1, адрес 172.16.0.0 представляет собой адрес сети класса B, а 1 — номер хоста в этой сети.

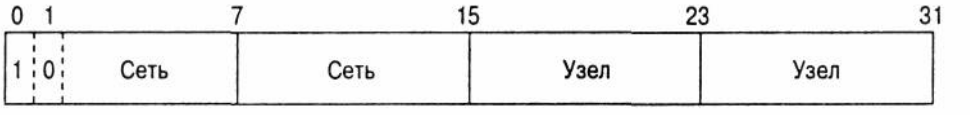


Рис. 3.2. *Общий вид адреса класса B*

Адреса класса C

Сети класса C определяются значениями 1, 1 и 0 старших битов в адресе. Первые три октета (биты с 0 по 23) используются для представления номеров сетей, а последний октет (биты с 24 по 31) представляют собой номера хостов в сети. Таким образом, получаем 2097152 (221) сетей, в каждой из которых находится 254 (28—2) хоста (рис. 3.3). Для примера возьмем адрес в сети класса C 192.11.1.1, где 192.11.1.0 представляет собой адрес сети, а номер хоста в сети — 1.

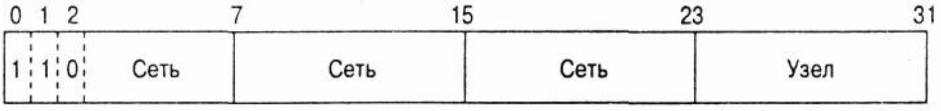


Рис. 3.3. *Общий вид адреса класса C*

Адреса класса D

Сети класса D определяются значениями 1, 1, 1 и 0 в первых четырех битах IP-адреса. Адресное пространство класса D зарезервировано для представления групповых IP-адресов, которые используются для адресации набора узлов. Это означает, что данный пакет должен быть доставлен сразу нескольким узлам, которые образуют группу с номером, указанным в поле адреса.

Адреса класса E

Сети класса E определяются значениями 1, 1, 1 и 1 в старших четырех битах IP-адреса. В настоящее время адреса этого диапазона не используются. Они зарезервированы для экспериментальных целей.

Основы формирования подсетей на базе протокола IP

Довольно часто основные вопросы формирования подсетей и понятие о подсетях переменной длины трудно воспринимаются даже подготовленными специалистами. В этом разделе дается краткое введение в основы построения подсетей, а в следующем разделе обсуждаются маски подсети переменной длины (variable-length subnet masks — VLSM).

Подсетью (subnet или *subnetwork)* называют отдельную IP-сеть класса A, B или C. Чтобы лучше представить себе разбиение на подсети, давайте подробно рассмотрим IP-

адреса, которые не отнесены к подсетям. Как уже отмечалось, в IP-адресах заключена часть, представляющая номер сети, и часть, адресующая номер хоста в сети. Таким образом мы получаем статическую двухуровневую иерархическую модель адреса (сети и хосты). Формирование подсетей в IP-сетях2 представляет собой третий уровень в этой иерархии, и оно производится с использованием сетевых масок (netmask). Маской сети служит битовая маска, в которой набор битов соответствует битам, используемым для нумерации IP-сети, а дополнительные биты соответствуют номеру подсети. Другими словами маска — это число, двоичная запись которого содержит единицы в тех разрядах, которые должны интерпретироваться как номер сети.

Так, на рис. 3.4 маска 255.0.0.0 определяет сеть 10.0.0.0. В двоичной записи она представляет собой непрерывную последовательность из единиц и нулей. Группа единиц представляет здесь сетевую часть IP-адреса, а нули — часть адреса, отвечающую за нумерацию хостов. Таким образом обеспечивается механизм разбиения IP-адреса 10.0.0.1 на сетевую часть (номер сети 10) и узловую часть (номер хоста 1).

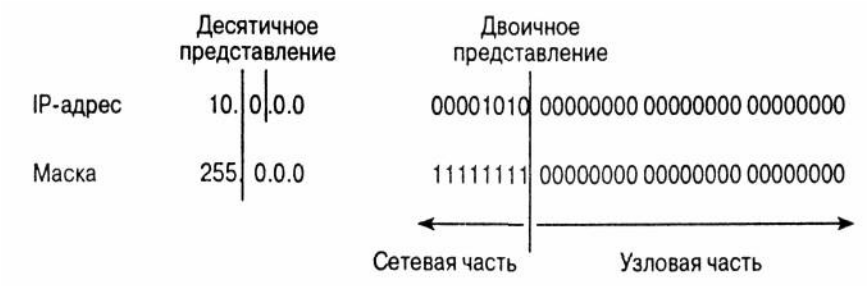


Рис. 3.4. Формирование сетевых масок

Адреса класса A, B и C имеют стандартные маски, которые задаются с учетом максимального количества сетей и узлов в них для каждого класса. Для стандартных классов сетей маски имеют следующие значения:

- 255.0.0.0 — маска для сети класса A.
- 255.255.0.0 — маска для сети класса B.
- 255.255.255.0 — маска для сети класса C.

Разделение сетевой и узловой частей IP-адреса облегчает процесс создания подсетей на основе масок. Без подсетей номера сетей быстро были бы исчерпаны. Обычно каждый физический сегмент, такой как Ethernet, Token Ring или FDDI, связан с одним или несколькими номерами сетей. Если сеть не разбивается на подсети, то, например, в сети класса A с адресом 10.0.0.0 будет находиться только один физический сегмент, в котором можно адресовать около 16 миллионов хостов, как показано на рис. 3.5.

При применении масок сети можно делить на более мелкие подсети путем расширения сетевой части адреса и уменьшения узловой части. Технология разбиения сетей на подсети позволяет создавать большее число сетей с меньшим количеством хостов в них.

На рис 3.6 рассмотрена маска 255.255.0.0 для сети 10.0.0.0. Тогда IP-адрес 10.0.0.1 рассматривается следующим образом: номер сети — 10; номер подсети — 0; и номер хоста — 1. С помощью маски 255.255.0.0 часть адреса, отведенная под нумерацию узлов, отбирается и применяется для нумерации сетей (а точнее подсетей в сети). В результате такой операции адресное пространство сети с номером 10 расширяется с одной сети до 256 подсетей в диапазоне от 10.0.0.0 до 10.255.0.0. Одновременно этим достигается снижение числа хостов в каждой подсети с 16777214 до 65534.

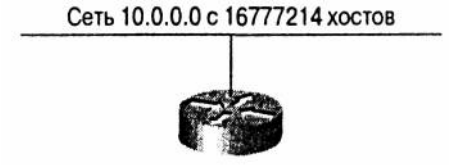


Рис. 3.5. *Адресное пространство класса A без разбиения на подсети*

Примечание

Обратите внимание, что в приведенном нами примере сеть 10.0.0.0 представляет собой сеть без подсетей. Некоторое устаревшее программное обеспечение для маршрутизаторов не позволяет работать с адресным пространством, не разбитым на подсети, и его нельзя использовать по умолчанию в маршрутизаторах типа Cisco. В порядке разрешения использования "нулевых" подсетей в ОС IOS нужно сконфигурировать ip subnet-zero.

Рис. 3.6. *Формирование подсетей*

Маски подсети переменной длины

Термин *маска подсети переменной длины (variable-length subnet mask — VLSM)* означает, что одна сеть может быть сконфигурирована с различными масками. Основная идея применения VLSM3 заключается в предоставлении большей гибкости при разбиении сети на несколько подсетей, т.е. для оптимального распределения допустимого количества хостов в различных подсетях. Без VLSM для всей сети может использоваться только одна маска подсети. Тогда количество хостов в подсетях будет строго ограничено. Если же вы выберете маску, которая предоставит нужное количество подсетей, то, возможно, вам будет недостаточно допустимого количества хостов для каждой подсети. Та же ситуация справедлива и для хостов, т.е. маска, обеспечивающая достаточное количество хостов, ограничивает вас в числе подсетей. Маски переменной длины предоставляют возможность выделять подсети с различным количеством хостов в них, что позволяет сетевому администратору более эффективно использовать доступное адресное пространство.

Допустим для примера, что вам выделена сеть класса С с адресом 192.214.11.0, и требуется разделить ее на три подсети. В одной подсети должно быть около 100 хостов, а в двух других — около 50 хостов в каждой. Исключая два адреса, 0 (номер сети) и 255 (широковещательный адрес для сети) вам теоретически доступно 256 адресов хостов для сети класса С, т.е. с 192.214.11.0 до 192.214.11.255. Как видите разбить такую сеть на подсети с требуемым количеством хостов без использования VLSM невозможно.

Чтобы определить параметры подсети в сети 192.214.11.0, сначала необходимо определить маску сети, которая для обычной сети класса С будет представлена в виде 255.255.255.0 (все биты равны 1 в первых трех октетах). Для разделения сети класса С с адресом 192.214.11.0 на подсети можно использовать несколько масок вида 255.255.255.X. Маска, начиная со старшего (самого левого) бита, должна иметь непрерывный ряд единиц и оканчиваться нулями.

Примечание

Изначально маски не обязательно должны были состоять из непрерывных групп 1 и оканчиваться 0. Иногда, например, практиковалось использование "средних битов" в маске для определения адресной части, отвечающей за идентификацию хоста, при этом младшие биты определяли адрес подсети. Хотя подобная гибкость в работе с масками и помогает сетевым администраторам при распределении адресов, все же эта методика значительно затрудняет маршрутизацию в сетях. Вследствие этого, согласно новым спецификациям, требуется, чтобы маски состояли из групп непрерывных единиц.

В табл. 3.3 приведены потенциальные маски, которые могут применяться для сегментирования адресного пространства из 256 адресов на подсети.

Таблица 3.3 *Разделение сети класса С на подсети*

Последний октет	Двоичное представление	Количество подсетей	Число хостов*
128	1000 0000	2	128
192	1100 0000	4	64
224	1110 0000	8	32
240	1111 0000	16	16
248	1111 1000	32	8
252	1111 1100	64	4

*Обратите внимание на то, что в поле таблицы "Число хостов" включены и адрес подсети и широковещательный адрес.

До появления VLSM сети обычно делились лишь простыми масками, как указано в табл. 3.3. В этом случае у вас был выбор применять маску 255.255.255.128 и разбить адресное пространство на две подсети по 128 хостов в каждой или разбить его маской 255.255.255.192 на четыре подсети по 64 хоста в каждой. Однако ни одна из этих процедур не соответствует вашим требованиям получить сегмент сети размером 100 хостов и еще два сегмента по 50 хостов в каждом.

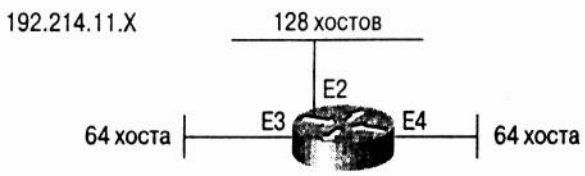


Рис. 3.7. *Пример сети класса С, разделенной на три подсети*

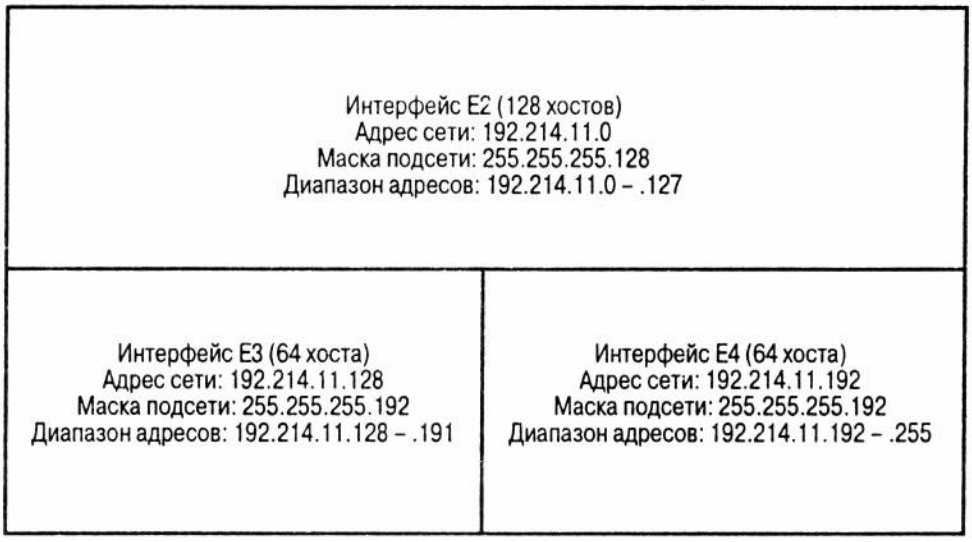


Рис. 3.8. *Применение VLSM для неравного деления адресного пространства на подсети*

Прибегнув к использованию масок переменной длины, вы можете выполнить поставленную задачу. Представим, что вы получили сеть 192.214.11.0. Во-первых, разделите эту сеть на две подсети маской 255.255.255.128. Вы получите две подсети по 128 хостов в каждой. Эти две подсети будут представлены адресами 192.214.11.0 (от .0 до .127) и 192.214.11.128 (от .128 до .255). Затем вторую подсеть с адресом 192.214.11.128 разбейте еще на две подсети с помощью маски 255.255.255.192 — вы получите две подсети по 64 адреса в

каждой: подсети 192.214.11.128 (адреса от .128 до 191) и 192.214.11.192 (адреса от .192 до 255). На рис. 3.7 и 3.8 представлен механизм деления адресного пространства на подсети. Обратите внимание, что адрес подсети и ее широковещательный адрес также включены в число адресов хостов.

Конечно, далеко не все протоколы маршрутизации поддерживают VLSM. Так, протокол информации о маршрутах версии 1 Routing Information Protocol (RIP-1) и протокол маршрутизации внутреннего шлюза Interior Gateway Routing Protocol (IGRP) не передают информацию о сетевых масках при обновлениях маршрутной информации и, следовательно, не могут корректно маршрутизировать сети с подсетями переменной длины. Сегодня, несмотря на то, что протоколы маршрутизации, такие как протокол кратчайшего свободного пути Open Shortest Path First (OSPF), расширенный IGRP (Enchanced IGRP или EIGRP), протокол информации о маршрутах версии 2 Routing Information Protocol (RIP-2) и протокол связи промежуточных систем Intermediate System-to-Intermediate System (IS-IS), поддерживают работу с VLSM, администраторы до сих пор испытывают трудности при реализации этой методики разделения сетей. Построенные ранее на базе протоколов RIP-1 и IGRP сети имеют структуру IP-адресов, распределенных таким образом, что невозможно более оптимально сгруппировать их в блоки различной длины. Таким образом, ввиду разброса IP-адресов администраторам пришлось бы перенумеровать все хосты в сети для того, чтобы привести их в соответствие с новой системой адресации. Такая перенумерация является довольно сложной процедурой, и администраторы чаще всего сразу же отвергают подобную перспективу. Однако одновременное сосуществование двух систем осложняет ситуацию и вынуждает администраторов всячески маневрировать и применять статическую маршрутизацию для обеспечения нормальной работы в сети.

Исчерпание адресного пространства IP

Растущая потребность в IP-адресах стала суровым испытанием для классовой модели. Большинство компаний стремилось получить адреса класса B, так как они наилучшим образом соответствовали их нуждам вследствие оптимального соотношения между количеством сетей и числом хостов в них. Дело в том, что для корпоративных сетей сеть класса A с 16 миллионами хостов предоставляла большие возможности, чем требовалось, а сеть класса C не могла удовлетворить потребности крупной компании из-за небольшого числа хостов, которые можно было адресовать. К 1991 году стало очевидным, что расход адресного пространства класса B приобрел угрожающие масштабы и следует принимать срочные меры для того, чтобы предотвратить исчерпание адресного пространства.

Для решения этой проблемы были приняты определенные меры: совещательный подход к распределению IP-адресов и побуждение организаций, не подключенных к сети Internet, к использованию диапазонов частных IP-адресов. Кроме того, были созданы специальные рабочие группы и управления, такие как рабочая группа по маршрутизации и адресации Routing and Addressing (ROAD) и управление по созданию протокола IP следующего поколения (IP next generation directorate — IPng). В 1992 году рабочая группа ROAD предложила в качестве мероприятий по отказу от классовой модели IP адресации использовать технологию бесклассовой междоменной маршрутизации (classless interdo-main routing — CIDR). В то же время управление IPng разрабатывав новую улучшенную систему IP-адресации с применением протокола IP версии 6 (IPv6), с помощью которого проблема истощения адресного пространства была бы полностью решена.

Меры по предотвращению исчерпания адресного пространства можно разделить на четыре категории.

- Распределение IP-адресов на совещательной основе.
- Бесклассовая междоменная маршрутизация (classless interdomain routing —

CIDR)4.

- Использование общедоступных (частных) IP-адресов и механизма преобразования сетевых адресов Network Address Translation (NAT)5,7.
- IP версии 6 (IPv6)8 .

Растущая потребность в IP-адресах совместно с проблемами по их исчерпанию вызвали необходимость в пересмотре процедур выделения IP-адресов центральными органами. Изначально полный контроль за назначением и распределением IP-адресов осуществлялся организацией IANA и реестром Internet — Internet Registry (IR). Все IP-адреса распределялись последовательно среди всех организаций, независимо от их географического местоположения и способа подключения к сети Internet. Метод распределения адресов существовавший в то время, напоминал скорее затыкание дыр в пространстве IP-адресов — выделялись лишь индивидуальные или небольшие группы IP-адресов, а большие, идущие подряд блоки адресов не выдавались.

Для выделения больших непрерывных блоков IP-адресов необходимо было найти другой подход. Было решено поручить эту задачу нескольким администрациям (таким как сервис-провайдеры), которые, в свою очередь, распределяли бы адреса среди клиентов, но уже из предоставленных им диапазонов. В общем такой метод распределения адресов показал себя лучше, чем метод централизованного распределения IP-адресов. Он в некоторой степени имеет сходство с подходом, используемым при распределении номеров в телефонной сети, где коды соответствуют географическим областям (сетям провайдеров), префиксы при наборе номера — регионам или районам города (клиентам провайдеров), а остальная часть — индивидуальным абонентам (хостам).

Распределение IP-адресов

Количество сетевых адресов класса A довольно ограничено, и распределение адресов из этого диапазона строго контролируется. Хотя адресное пространство класса A будет и дальше распределяться между субъектами сети, это будет происходить на основе распределения подсетей в сетях класса A, а не полных соответствующих классу блоков IP-адресов. Распределение адресов класса B также контролируется и проводится путем выделения подсетей из сетей класса B. Адреса класса C, как правило, выделяются из адресного пространства сервис-провайдера. В табл. 3.4 приведены критерии распределения адресного пространства класса C.

Таблица 3.4. *Назначение адресов класса C*

Потребности организации в адресном пространстве	**Назначение адресов**
Менее 256 адресов	1 сеть класса C
Менее 512, но более 256 адресов	2 непрерывных сети класса C
Менее 1024, но более 512 адресов	4 непрерывных сети класса C
Менее 2048, но более 1024 адресов	8 непрерывных сети класса C
Менее 4096, но более 2048 адресов	16 непрерывных сети класса C
Менее 8192, но более 4096 адресов	32 непрерывных сети класса C
Менее 16384, но более 4096 адресов	64 непрерывных сети класса C

Региональные реестры сети Internet, такие как Американский реестр адресов Internet — American Registry for Internet Numbers (ARIN), в последнее время довольно неохотно выделяет блоки адресов напрямую конечным пользователям. Для получения адресного пространства напрямую от ARIN необходимо подтвердить готовность к освоению по крайней мере 16 блоков адресов класса C или 4096 хостов. Даже при обосновании необходимости такого количества адресов администраторам рекомендуется получать их от своих провайдеров.

Вся информация по получению и распределению адресов, включая руководящие документы, шаблоны запросов, принципы распределения адресов и др., находится на Web-

сервере ARIN (www.arin.net).

Географически адресное пространство распределено между четырьмя основными регионами: Европой, Северной Америкой и Северной Африкой, Тихоокеанским регионом, а также Южной и Центральной Америкой. В табл. 3.5 показано распределение адресов между этими регионами. Некоторые сетевые адреса, согласно плану распределения, выдавались нескольким регионам.

Таблица 3.5. *Распределение адресного пространства по регионам*

Адресное пространство	Кому выдано	Когда выдано
61.0.0.0 — 61.255.255.255	APNIC — Тихоокеанский регион	апрель 1997
62.0.0.0 — 62.255.255.255	RIPE NCC — Европа	апрель 1997
63.0.0.0 — 63.255.255.255	ARIN	апрель 1997
64.0.0.0 — 64.255.255.255	ARIN	июль 1999
128.0.0.0 — 191.255.255.255	Различные реестры	май 1993
192.0.0.0 — 192.255.255.255	Различные реестры	май 1993
193.0.0.0—195.255.255.255	RIPE NCC — Европа	май 1993
196.0.0.0 — 198.255.255.255	Различные реестры	май 1993
199.0.0.0 — 199.255.255.255	ARIN — Северная Америка	май 1993
200.0.0.0 — 200.255.255.255	ARIN — Центральная и Южная Америка	май 1993
201.0.0.0 — 201.255.255.255	Зарезервировано - Центральная и Южная Африка	май 1993
202.0.0.0 — 203.255.255.255	APNIC — Тихоокеанский регион	май 1993
204.0.0.0 — 205.255.255.255	ARIN — Северная Америка	май 1993
206.0.0.0 — 206.255.255.255	ARIN — Северная Америка	март 1994
207.0.0.0 — 207.255.255.255	ARIN — Северная Америка	апрель 1995
208.0.0.0 — 208.255.255.255	ARIN — Северная Америка	ноябрь 1995
209.0.0.0 — 209.255.255.255	ARIN — Северная Америка	апрель 1996
210.0.0.0 — 210.255.255.255	APNIC — Тихоокеанский регион	июнь 1996
211.0.0.0 — 211.255.255.255	APNIC — Тихоокеанский регион	июнь 1996
212.0.0.0 — 212.255.255.255	RIPE NCC — Европа	октябрь 1997
213.0.0.0 — 213.255.255.255	RIPE NCC — Европа	март 1999
216.0.0.0 — 217.255.255.255	ARIN — Северная Америка	апрель 1998

Бесклассовая междоменная маршрутизация

Всего несколько лет назад глобальные таблицы маршрутов выросли настолько, что маршрутизаторам стало не хватать мощностей и памяти для их обработки. Статистические данные показывают, что с 1991 по 1995 год удвоение размера таблиц маршрутов происходило каждые 10 месяцев; наиболее заметный рост наблюдался с 1998 года. На рис. 3.9 представлена диаграмма роста таблиц маршрутов в сети Internet.

Если бы не предпринималось никаких действий по урегулированию роста таблиц маршрутов, то они выросли бы до немыслимых размеров — около 80000 маршрутов в 1995 году. Однако, как видим, в начале 2000 года в таблицах маршрутов было около 76000 маршрутов. Такое снижение роста маршрутных таблиц было достигнуто благодаря схеме распределения IP-адресов, рассмотренной нами в предыдущем разделе, а также за счет реализации бесклассовой междоменной маршрутизации (Classless Interdomain Routing — CIDR).

Бесклассовая междоменная маршрутизация была новым шагом на пути эволюции IP-адресов классов A, B и C. При работе с CIDR любая IP-сеть представляется префиксом, в котором в IP-адрес сети через косую черту вносится число, показывающее количество бит соответствующих маске подсети, связанной с данным сетевым адресом. Например, рассмотрим сеть с адресом 198.32.0.0 и префиксом /16, которая записывается как 198.32.0.0/16. Здесь префикс /16 указывает на то, что в маске используются старшие 16 битов. Таким образом, это соответствует IP-сети 198.32.0.0 с маской 255.255.0.0.

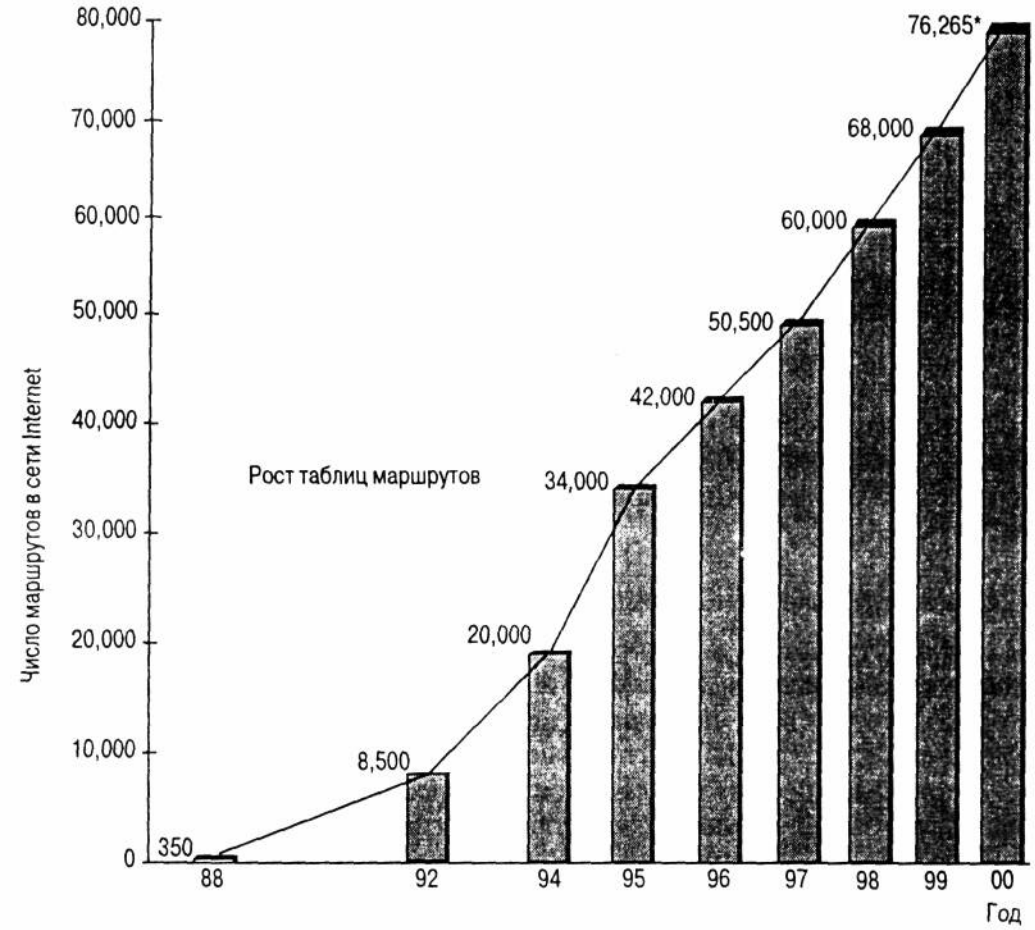


Рис. 3.9. Рост таблиц маршрутов в сети Internet

*Информация получена из отчета CIDR от 25 мая 2000 года

Сеть можно называть *суперсетью (supernet),* если префикс маски содержит меньше битов, чем обычная маска сети. Например, сеть класса А с адресом 198.32.1.0 имеет нормальную маску 255.255.255.0, которая соответствует префиксу /24 в представлении CIDR. Сеть 198.32.0.0 с маской 255.255.0.0 можно представить в записи 198.32.0.0/16, при этом обе записи предполагают, что маска будет меньше нормальной маски для сети класса С (16 меньше 24), следовательно, сеть можно отнести к суперсетям. На рис. 3.10 представлены эти схемы адресации.

При такой системе записи обеспечивается механизм для оптимального распределения маршрутов во всей сети 198.32.0.0/16 (таких как 198.32.0.0, 198.32.1.0, 198.32.2.0 и т.д.) с помощью однократного их объявления, которое также называется *совокупным или объединенным (aggregate).*

Вся эта терминология может ввести вас в заблуждение, так как термины "совокупное объявление маршрутов", "блок CIDR" и "суперсеть" часто взаимно заменяют друг друга и обозначают одно и то же. В принципе, все эти термины указывают на то, что группа непрерывных IP-сетей объединяется, и при объявлении маршрутов все сети объявляются одновременно. Выражаясь точнее, при записи маршрутов в представлении CIDR (префикс/длина), у суперсетей длина префикса меньше, чем нормальная маска, и совокупность сетей может быть маршрутизирована по одному любому маршруту.

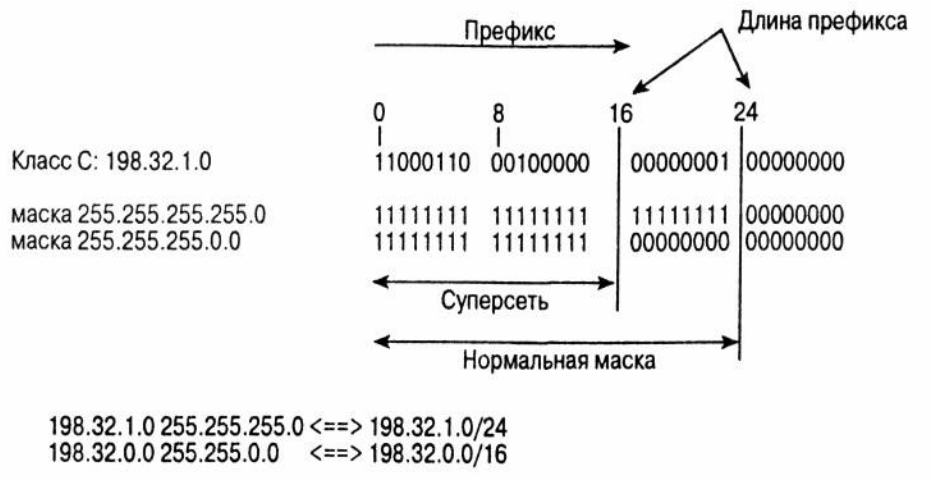


Рис. 3.10. *Адресация на основе CIDR*

Все сети, которые входят в набор совокупных сетей, или, по-другому, блоки CIDR, обычно относят к специфичным, так как они предоставляют больше информации о расположении сети. Однозначно определенные префиксы длиннее совокупных:

- 198.213.0.0/16 имеет совокупную длину 16 бит.
- 198.213.1.0/20 имеет однозначно определенный префикс длиной 20 бит.

Домены маршрутизации, в которых возможно производить разбиение на блоки CIDR, относят к *бесклассовым (classless),* в отличие от традиционных классовых доменов, в которых невозможно провести такую процедуру. С помощью CIDR можно описать новую структуру сети Internet, где каждый домен получает свой IP-адрес от высшего по иерархии уровня. Таким образом предполагается достичь существенного сокращения рассеивания маршрутов, особенно при суммировании в концевых или, как их еще называют, *сетях-заглушках (stub network).* Концевые сети, или сети-заглушки, представляют собой конечные точки глобальных сетей; они, в свою очередь, обеспечивают соединение с другими сетями в Internet. Провайдеры, поддерживающие несколько листовидных сетей делят подсети в них на небольшие блоки адресов, которые предоставляются конечным пользователям. Объединение сетей позволяет провайдерам объявлять одну IP-сеть, которая представляется как суперсеть, вместо того, чтобы проводить несколько отдельных объявлений. В результате, наряду с повышением стабильности объявления маршрутов, создаются более эффективные схемы маршрутизации и процедуры распространения маршрутов. На рис. 3.11 показана сравнительная характеристика разных систем маршрутизации.

Итак, в приведенном примере ISP3 был выделен блок IP-адресов, начиная с 198.0.0.0 до 198.1.255.255 (198.0.0.0/15). Затем провайдер разбивает этот блок адресов на два более мелких блока и выдает их провайдерам ISP1 и ISP2. Так, ISP1 получает диапазон адресов от 198.1.0.0 до 198.1.127.255 (198.1.0.0/17), а ISP2 получает диапазон от 198.1.128.0 до 198.1.255.255 (198.1.128.0/17). Точно так же провайдеры ISP1 и ISP2 распределяют эти адреса между своими клиентами. Частный случай, представленный на рис. 3.11 слева, показывает, что происходит, если не использовать CIDR: провайдеры ISP1 и ISP2 объявляют маршруты ко всем подсетям, выделенным своим клиентам, и провайдер 1SP3, в свою очередь, передает всю эту информацию во внешний мир. Результатом таких действий становится увеличение глобальных таблиц маршрутов в сети на основе TCP/IP.

В правой части рис. 3.11 представлен тот же сценарий распределения адресов, но уже с использованием CIDR. Как видите, в этом случае провайдеры ISP1 и ISP2 объединяют подсети своих клиентов. Провайдер ISP1 при этом объявляет лишь о маршрутах в объединенную сеть с адресом 198.1.0.0/17. В то же время провайдер ISP2 объявляет о маршрутах в объединенную сеть 198.1.128.0/17. Далее провайдер ISP3 точно так же объединяет подсети своих клиентов — провайдеров ISP1 и ISP2 — и посылает затем информацию лишь о маршруте в одну сеть (198.0.0.0/15). Таким образом, значительно сокращаются глобальные таблицы маршрутов.

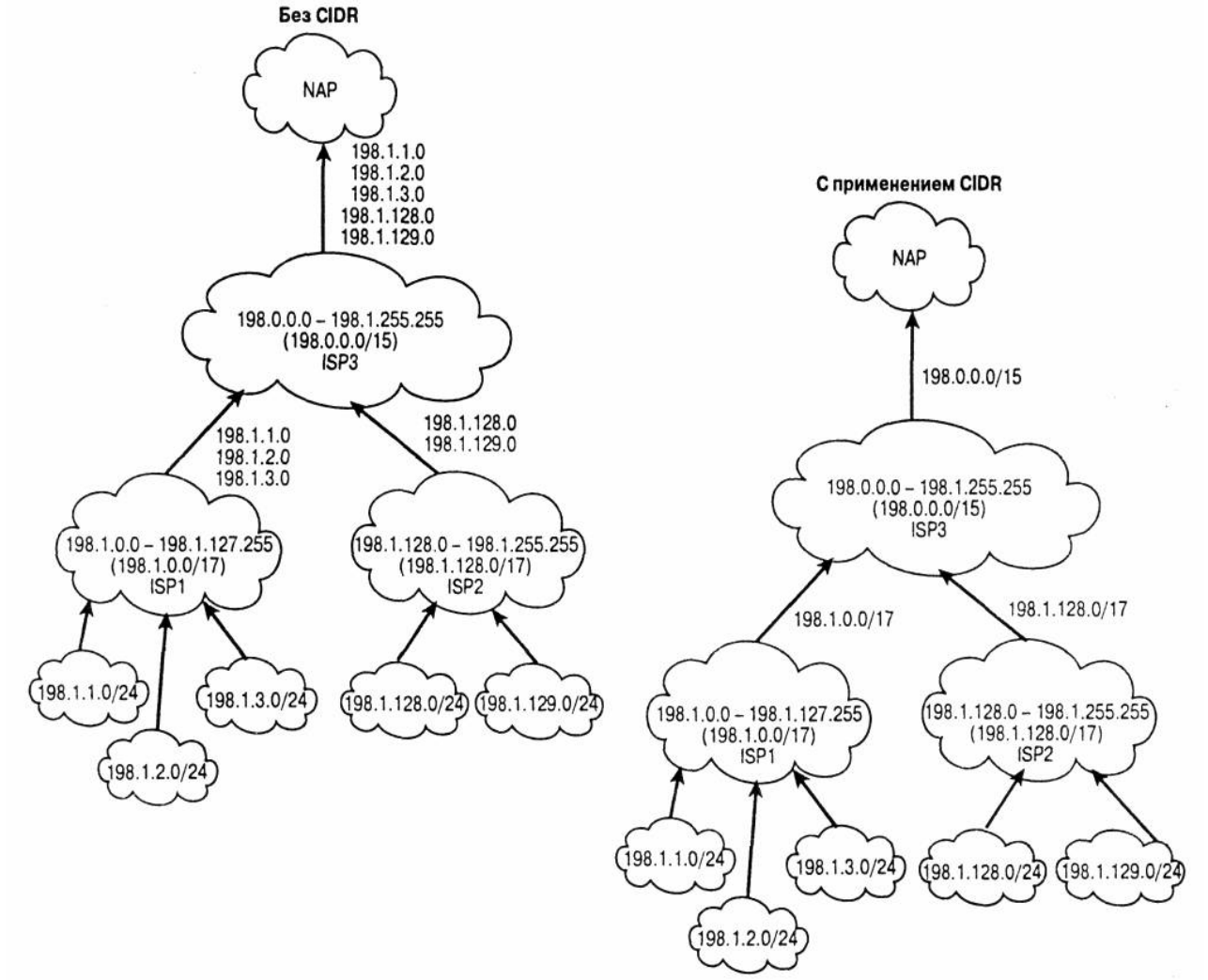


Рис. 3.11. Классовая адресация и адресация на основе CIDR

Итак, можно сделать вывод о том, что большая эффективность достигается при объединении сетей вблизи концевых узлов (так называемых *leaf nodes),* так как большинство объединяющихся подсетей являются сетями конечных пользователей. Процесс объединения, или, как его еще иногда называют, агрегирования, сетей (от англ. *aggregation — Прим. ред.}* на более высоких уровнях, таких как, например, на уровне ISP3, дает менее впечатляющие результаты, так как он по сути является объединением небольшого числа сетей, которые состоят из множества клиентских сетей.

Наиболее оптимально процесс агрегирования сетей работает при условии, что каждый клиент подключен к провайдеру только через одно соединение, которое называют *одноканальным (single-homing],* и получает блоки IP-адресов только из блоков CIDR провайдера. К сожалению, в реальной жизни вы редко встретите подобную ситуацию. Чаще всего возникают другие варианты развития событий. Например, клиент уже получил IP-адреса, но не из диапазона, предлагаемого его провайдером. Или несколько клиентов (которые также могут быть провайдерами) нуждаются в соединении не с одним, а с несколькими провайдерами одновременно. Эта схема носит название *многоканальной (multihoming).* В этих случаях возникают определенные трудности при объединении сетей, и теряется гибкость маршрутизации, достигаемая агрегированием.

Правило длиннейшего подходящего маршрута

Маршрутизация в сетях TCP/IP от одного узла к другому всегда выполняется на основе длиннейшего подходящего маршрута. Суть этого метода маршрутизации заключается в том, что маршрутизатор, принимающий для каждой сети решение на основе префиксов различной длины, всегда будет выбирать маршрут с более длинной маской. Представим, например, что в таблице маршрутов на маршрутизаторе имеются две записи:

198.32.1.0/24 по маршруту 1 198.32.0.0/16 по маршруту 2

При попытке доставить трафик на хост 198.32.1.1 маршрутизатор пытается найти маршрут с пунктом назначения, имеющим самый длинный префикс, и, найдя его, отправит все пакеты по маршруту 1.

На рис. 3.12 показана работа правила длиннейшего подходящего маршрута. Здесь домен В получает сведения о двух маршрутах: 198.32.1.0/24 и 198.32.0.0/16, и маршрутизатор в этом домене направляет трафик на хост с адресом 198.32.1.1 по маршруту 1.

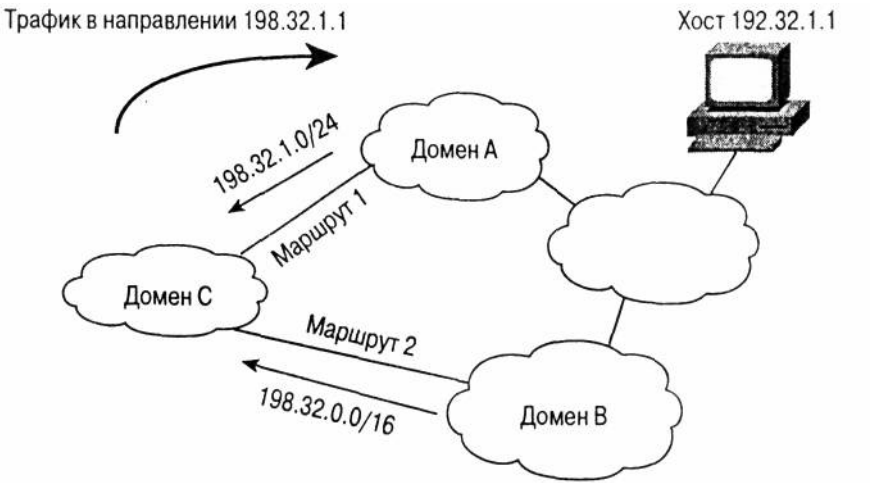


Рис. 3.12. Правило длиннейшего подходящего маршрута

Если по какой-либо причине маршрут 1 станет недоступным, то трафик будет передаваться через следующий ближайший маршрут в таблице маршрутов, который ведет к нужному хосту (в нашем примере это маршрут 2). Если домен В получил сведения о маршрутах с масками одинаковой длины от доменов А и Б, то маршрутизатор выбирает любой из маршрутов или, если в домене используется один из методов балансировки нагрузки, используются оба маршрута.

Правило длиннейшего подходящего маршрута предполагает, что маршруты к пункту назначения, подключенному к нескольким доменам, обязательно должны быть объявлены, причем в специфичной для этих доменов форме, т.е. до проведения объединения сетей. На рис. 3.12, ввиду того, что домен В не объявил явным образом маршрут 198.32.1.0/24, весь трафик от клиента на заданный хост всегда будет пересылаться по маршруту, выбранному с помощью правила длиннейшего подходящего префикса, т.е. через домен А. Однако такая конфигурация маршрутов может вызвать нежелательную перегрузку трафиком домена А.

Менее специфичные (неопределенные) маршруты при объединении сети

Существует специальное правило маршрутизации, которое гласит: для предотвращения образования петель маршрутизации (routing loops) в сети не должны возникать маршруты к пунктам назначения, которые совпадают с маршрутами к сетям, составляющим объединенную (агрегированную) сеть. *Петля маршрутизации (routing loop}* возникает, когда трафик образует кольцо между элементами сети и следует по нему взад и вперед, не имея возможности достичь конечного пункта назначения. Маршруты вида 0.0.0.0/0, используемые по умолчанию, являются частным случаем этого правила. В сети не следует обозначать маршруты по умолчанию для пунктов назначения, которые являются частью объединенных сетей. Вот почему протоколы маршрутизации, которые обрабатывают маршруты к объединенным сетям, всегда содержат специальную битовую корзину (bit bucket) (в терминах Cisco маршрут NullO) для выделения маршрута самого из себя. Трафик, направленный в битовую корзину, будет уничтожен, что предотвращает образование потенциальных петель в маршрутизации.

Совет

При организации маршрутов по умолчанию избегайте создания петель маршрутизации. Для этого используйте битовую корзину.

На рис. 3.13 приведен пример объединения всего домена провайдера ISP1 в один маршрут 198.32.0.0/13.

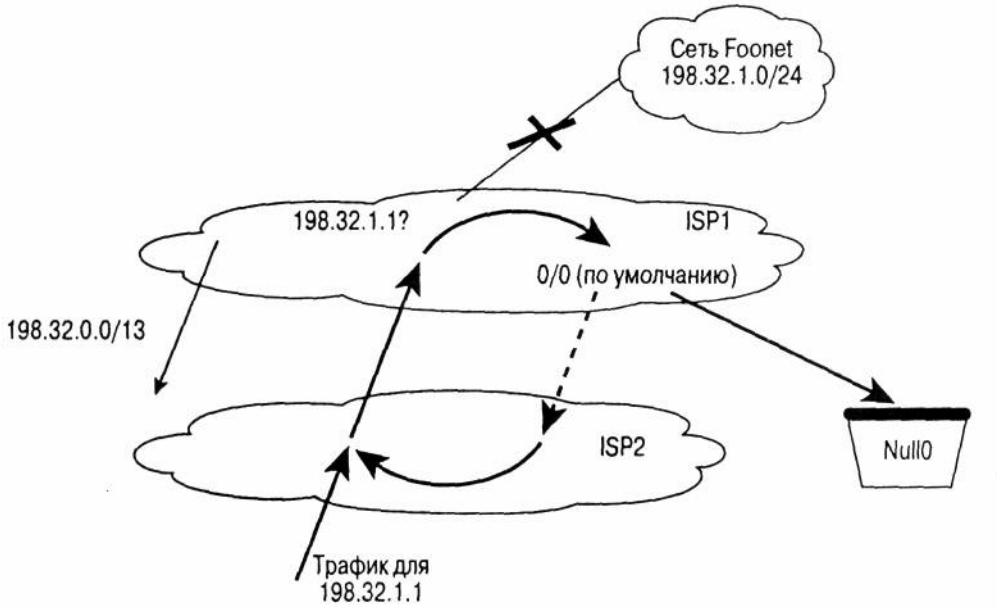


Рис. 3.13. *Следование менее специфичным маршрутам при объединении сети приводит к появлению петель маршрутизации*

Предположим, что соединение между провайдером ISP1 и его клиентом — компанией Foonet (где расположена сеть 198.32.1.0/24) пропадает по каким-либо причинам. Представим также, что у провайдера ISP1 имеется маршрут по умолчанию 0.0.0.0/0, указывающий на сеть провайдера ISP2. По этому маршруту отсылается весь трафик, адреса пунктов назначения которого неизвестны провайдеру ISP1. Тогда трафик, направленный на адрес 198.32.1.1, следуя маршруту для объединенной сети, попадет в сеть провайдера ISP1 и, не найдя пункта назначения, возвращается по маршруту, заданному по умолчанию, в сеть провайдера ISP2. Как видите, трафик передвигается вперед и назад между сетями провайдеров ISP1 и ISP2, образуя петлю. Чтобы избежать появления такого рода петли, на маршрутизаторах провайдера ISP1 создается нулевая запись для объединенного маршрута 198.32.0.0/13. Запись о нулевом маршруте позволяет уничтожать все пакеты, которые следуют в пункт назначения, недоступный в настоящее время, к тому же маршрут к нему определен менее жестко, чем агрегированный маршрут.

Итак, агрегирование (или объединение) при неправильном применении может привести к появлению петель маршрутизации, или, как их еще называют, *"черных дыр" (black holes).* "Черная дыра" возникает, когда трафик, несмотря на то, что он предназначен другому узлу, достигает определенного узла и не передается дальше, так как отсутствует маршрут к узлу назначения. Все эти сложности станут для вас менее пугающими, как только вы изучите схемы составного распределения адресов и их взаимодействие с процессом агрегации.

Одноканальное соединение — адреса получены из адресного пространства другого провайдера

Обсуждаемые нами выше правила маршрутизации, наряду с системой сетевого адресного пространства, независимо от подключения (одно- или многоканального), так или иначе участвуют в реализации объединения адресов в сети. В этом и в последующих трех разделах мы рассмотрим несколько схем межсетевого взаимодействия.

При одноканальном соединении (single-homing scenario) клиент подключается к одному провайдеру, но имеет свое адресное пространство IP, которое совершенно отлично от адресного пространства провайдера. Такая ситуация может возникнуть, если клиент сменил провайдера, но сохранил право на адреса, выданные ему предыдущим провайдером. Обычно в этом случае клиентам настоятельно рекомендуется или даже требуется сменить свои адреса на новые. Если не провести перенумерацию адресов, то новый провайдер не сможет агрегировать адреса клиента. Кроме того, прежний провайдер не сможет выполнять агрегирование так же эффективно, как раньше, ввиду того, что в его адресном пространстве образовалась дыра. В итоге от использования старого адресного пространства мы получим лишь увеличение глобальных таблиц маршрутизации, так как для такой сети потребуется введение дополнительных маршрутов.

Такая схема соединения будет обсуждаться в последующих главах, но здесь следует отметить, что сетевые администраторы, сети которых имеют одноканальное соединение, должны руководствоваться принципом: лучшие параметры маршрутизации — самые простые. В случае одноканального соединения самым беспроблемным является подход, когда провайдер статически маршрутизирует адресное пространство для вашей сети, а вы указываете маршрут по умолчанию на сеть провайдера. Более сложные решения при построении системы маршрутизации следует применять лишь в случаях, когда имеется несколько соединений с различными провайдерами или предъявляются повышенные требования к резервированию соединений.

Принцип упрощения конструкции "Keep It Simple, Stupid (KISS)", которым должен руководствоваться каждый конструктор, архитектор, инженер и администратор, предполагает, что самое простое доступное решение проблемы, чаще всего является самым лучшим.

Многоканальное соединение — адреса получены от одного провайдера

При этой схеме подключения (фрагмент ее вы можете увидеть на рис. 3.14) клиенты имеют несколько каналов, подключенных к различным провайдерам. При этом сеть клиента достаточно мала, чтобы адресного пространства, выделяемого одним из провайдеров, хватило на удовлетворение всех нужд, или пространство было выделено, когда клиент имел одноканальное соединение с одним провайдером. Мы рассмотрим следующую схему: два провайдера (1SP1 и ISP2) и их клиенты — компании Onenet, Twonet и Stubnet взаимодействуют друг с другом. Для каждого домена в табл. 3.6 приведены диапазоны IP-адресов, соответствующих им агрегированных (объединенных) сетей и провайдеров.

Обратите внимание на то, что сети Onenet и Twonet имеют многоканальные соединения с провайдерами ISP1 и ISP2, при этом диапазоны IP-адресов получены ими от провайдера ISP1 (рис. 3.14).

Таблица 3.6. *Список клиентов и их провайдеров*

Домен	Диапазон адресов	Агрегированная сеть	Провайдер	От кого получены адреса
ISP1	198.24.0 — 198.31.255.255	198.24.0.0/13		
Onenet	198.24.0.0 — 198.24.15.0	198.24.0.0/20	ISP1, ISP2	ISP1
Stubnet	198.24.16.0 — 198.24.23.0	198.24.16.0/21	ISP1	ISP1
Twonet	198.24.56.0 — 198.24.63.0	198.24.56.0/21	ISP1, ISP2	ISP1
ISP2	198.32.0.0 — 198.39.255.255	198.32.0.0/13		

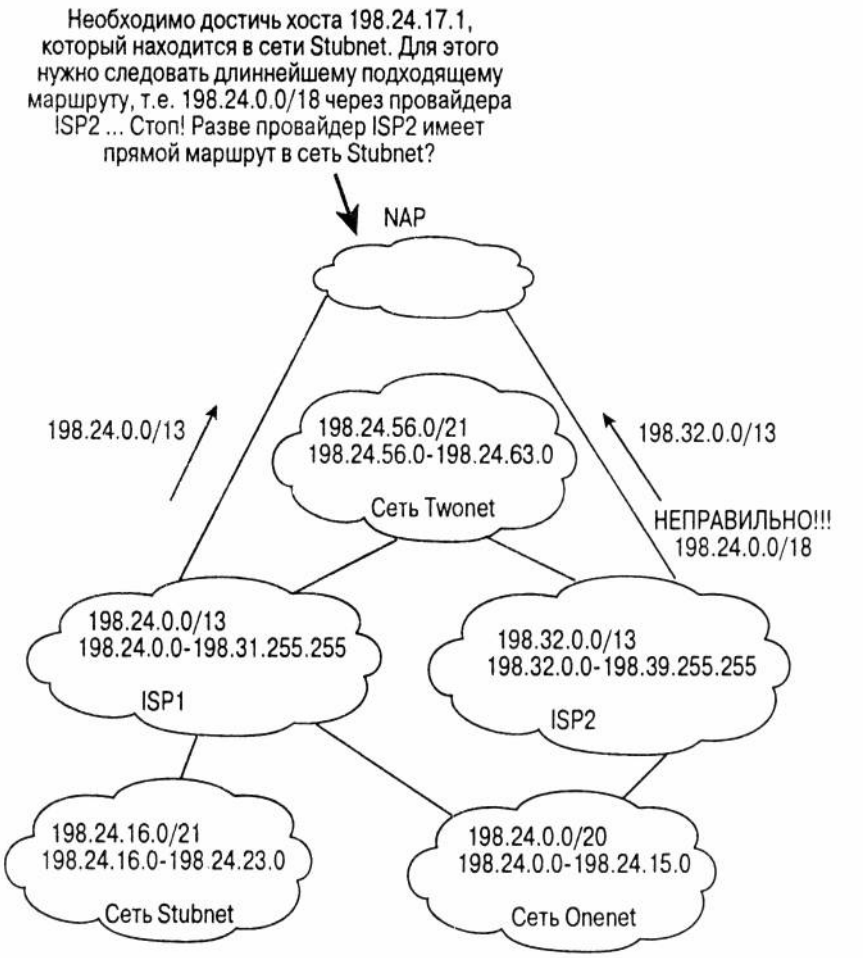


Рис. 3.14. *При неправильном объявлении агрегированных маршрутов могут возникать "черные дыры "*

Объявление агрегированных маршрутов — весьма сложное и тонкое дело. И клиенты, и их провайдеры должны очень осторожно подходить к вопросу выделения IP-адресов, которые входят затем в агрегированную сеть. При этом абсолютно исключается возможность агрегировать еще какие-либо маршруты (так называемая прокси-агрегация (proxy aggregation)), если одна из сторон не является суперсетью для другой стороны или между ними не заключено соответствующее соглашение об этой процедуре. В следующем примере вы увидите, каким образом провайдер ISP2 может создать "черную дыру" в маршрутизации. Это осуществляется объединением диапазонов сетей Onenet и Twonet.

Примечание

"Черные дыры" чаще всего являются результатом неправильного объединения маршрутов.

Если ISP2 при обновлении маршрута посылает информацию об объединенном маршруте, в который включены сети Onenet и Twonet (198.24.0.0/18), как показано на рис. 3.14, то в маршрутизации возникает "черная дыра". Так, например, сеть компании Stubnet, которая является клиентом ISP1, имеет адресное пространство, которое входит в объединенную сеть 198.24.0.0/18. Если провайдер ISP2 объявит такой объединенный маршрут, то весь трафик в сеть компании Stubnet, следуя правилу длиннейшего подходящего префикса в IP-адресе, будет заканчиваться в сети самого провайдера ISP2, что вызовет появление "черной дыры". Именно поэтому провайдер ISP2 должен однозначно анонсировать маршруты к каждому диапазону адресов своих клиентов, адресное пространство которых не входит в его диапазон IP-адресов (198.24.0.0/20 для Onenet и 198.24.0.0/21 для Twonet), который он также получил от ISP1. Кроме того, провайдер ISP2

должен вместе с маршрутами к сетям клиентов анонсировать и свое собственное адресное пространство 198.32.0.0/13.

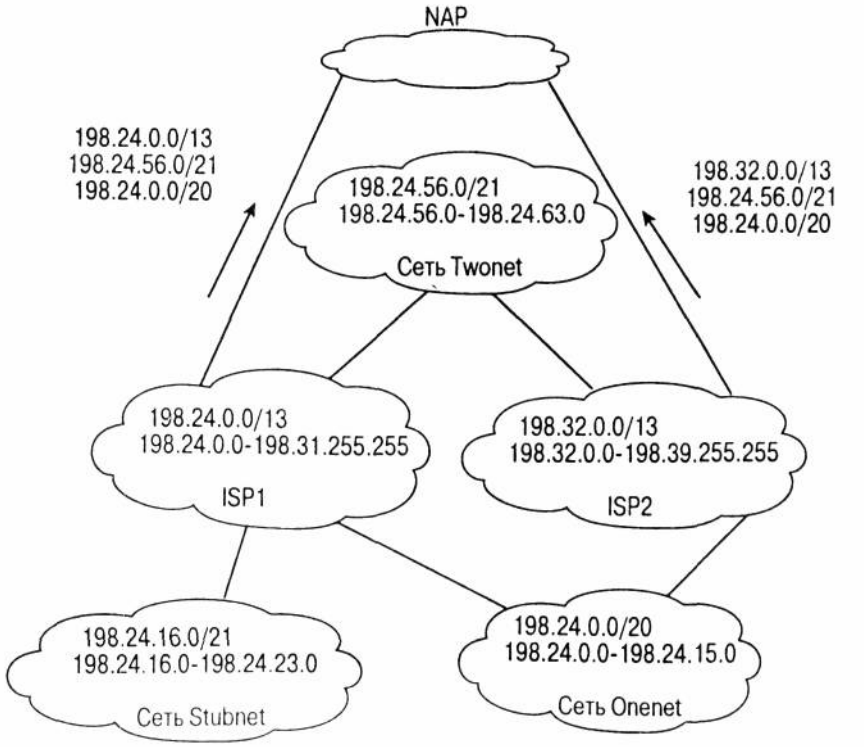


Рис. 3.15. Правильно сформированные объединенные маршруты предотвращают возникновение "черных дыр"

На рис. 3.15 представлена схема правильного объявления маршрутов к объединенным сетям. Здесь, провайдер ISP2 явным образом объявляет маршруты к сетям компаний Onenet и Twonet. В таком случае трафик, предназначенный сети Stubnet, никогда не попадет в сеть провайдера ISP2.

Заметим, что провайдер ISP1 также объявляет явным образом объединенные сети Onenet и Twonet (рис. 3.15). Если бы провайдер ISP1 анонсировал только неопределенный объединенный маршрут 198.24.0.0/13, то весь трафик в направлении сетей Onenet и Twonet пошел бы по более определенному длиннейшему подходящему пути, т.е. через сеть провайдера ISP2.

Многоканальное соединение —адреса получены от разных провайдеров

Как правило, в больших доменах довольно высока вероятность получения IP-адресов от различных провайдеров, в зависимости от их географического положения. Рассмотрим рис. 3.16. Сеть Largenet получила диапазон IP-адресов от двух различных провайдеров — ISP1 и ISP2. При этом каждый из провайдеров способен объединить собственное адресное пространство без необходимости указывать явным образом диапазоны адресов другого провайдера. Таким образом, ISP1 объявляет о маршруте к объединенной сети 198.24.0.0/13, а ISP2 — о маршруте к сети 198.32.0.0/13. При этом обе объединенные сети являются суперсетями для блоков IP-адресов, составляющих сеть Largenet.

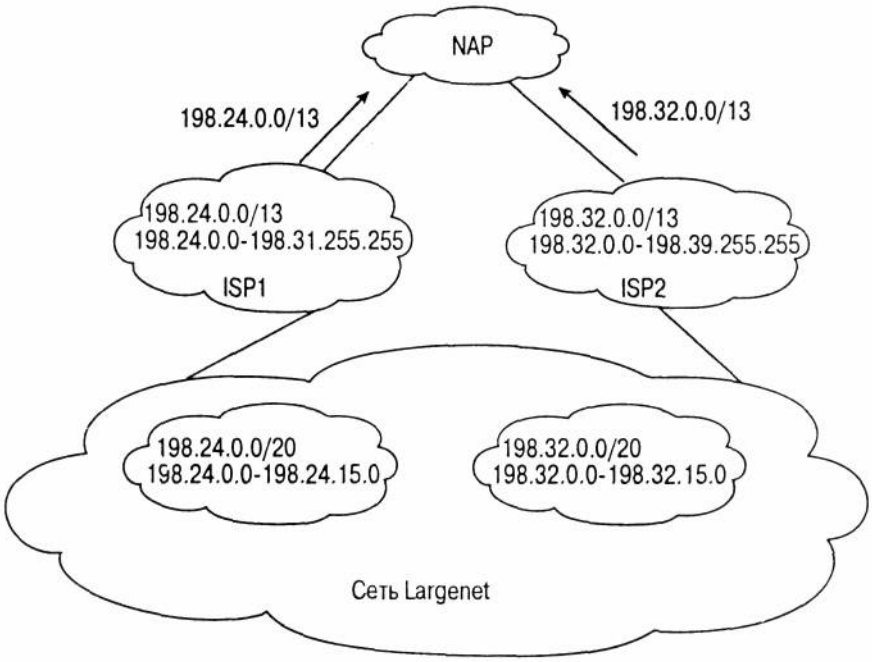


Рис. 3.16. Многоканальное соединение с IP-адресами, полученными от разных провайдеров

Основной недостаток схемы, представленной на рис. 3.16, — невозможность обслуживания резервных маршрутов для организаций с многоканальными подключениями. Дело в том, что провайдер ISP2 анонсирует только блок своих адресов и не дает никаких сведений об адресах из диапазона провайдера ISP1. Если на узле провайдера ISP1 возникнет непредвиденная ситуация и маршруты к сети 198.24.0.0/13 будут потеряны, то весь трафик в сети Largenet, адресованный в сеть 198.24.0.0/20, также не достигнет пунктов назначения, так как этот маршрут нигде больше не объявлен. Та же участь ожидает и адреса Largenet, полученные от провайдера ISP2. При пропадании канала с 1SP2 доступ к диапазону адресов 198.32.0.0/20 будет затруднен. Для того чтобы исправить эту ситуацию, провайдер ISP1 должен объявить у себя маршрут к сети 198.32.0.0/20, а ISP2 в свою очередь объявить маршрут в сеть 198.24.0.0/20.

Многоканальное соединение — адреса получены от третьей стороны

На рис. 3.17 показана ситуация, когда адреса, полученные клиентом, не принадлежат ни к одному из диапазонов адресов провайдеров ISP1 и ISP2, т.е. получены у третьей стороны. В этом случае и ISP1, и ISP2 объявляют специальный агрегированный маршрут (202.24.0.0/20) в дополнение к своим собственным диапазонам адресов (198.24.0.0/13 и 198.32.0.0/13). Недостаток этого метода заключается в том, что все маршрутизаторы в сети Internet должны иметь специальный маршрут к новому диапазону адресов. Большое количество подобных схем подключения может привести к значительному росту глобальных таблиц маршрутов в сети Internet.

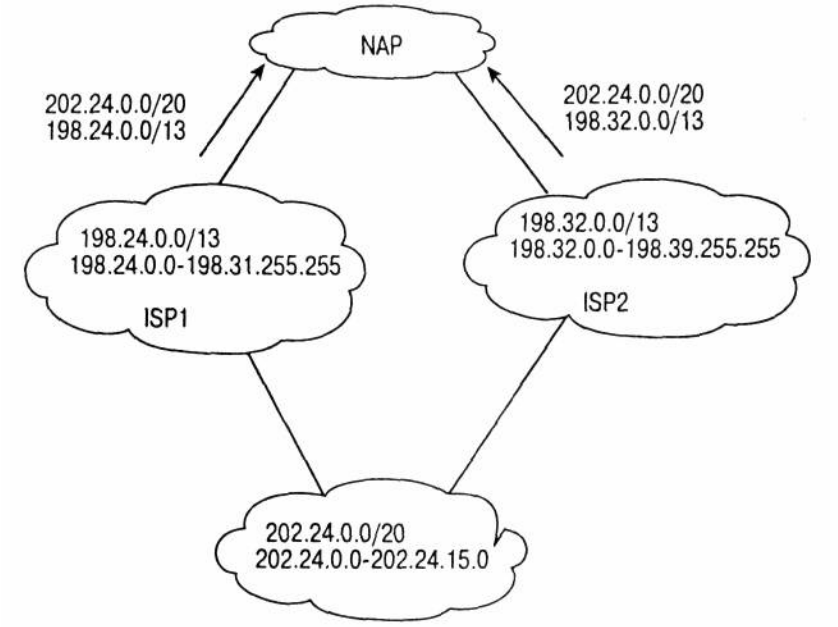


Рис. 3.17. *Получение адресов от третьей стороны (вне диапазона провайдеров)*

Рекомендации по проведению объединения (агрегирования)

Домен, для которого был выделен диапазон IP-адресов, должен обеспечить возможность их объединения (агрегирования). При выполнении объединения адресов в домене следует подвергать этой процедуре максимально возможное количество адресов, исключая всякую неоднозначность, присущую многоканальным сетям.

Однако следует помнить, что каждый конкретный случай требует индивидуального подхода. Не существует единого простого решения на все случаи жизни. Для одноканальных схем подключения клиентам рекомендуется получать простые непрерывные блоки IP-адресов от своего провайдера и выполнять совместно статическую маршрутизацию, если есть возможность не прибегать к более сложным схемам подключения. При смене провайдера клиентам с одноканальным соединением желательно перейти на использование адресного диапазона нового провайдера. Для клиентов с многоканальными соединениями назначение адресов должно проводиться с учетом максимально возможного объединения адресного пространства. В тех случаях, когда объединение влияет на резервирование, вопросы резервирования должны превалировать, для чего следует определять дополнительные сети с соответствующими маршрутами к ним.

Введение CIDR в течение последних нескольких лет помогло остановить лавинообразный рост глобальных таблиц маршрутов в сети Internet. Для сети Internet де-факто стад стандартом междоменный протокол маршрутизации граничного шлюза версии 4 Border Gateway Protocol (BGP-4), частично благодаря эффективной обработке объединенных маршрутов и их распространению между доменами маршрутизации. По мере чтения книги вы познакомитесь с примерами, поясняющими важность применения CIDR в управлении трафиком и сохранении стабильной работы сети.

Дополнительную информацию о CIDR, текущих и исторических сведениях о размерах таблиц маршрутов в сети Internet и другую интересную информацию вы можете получить, обратившись к приложению А в конце книги.

Частные адреса и преобразование сетевых адресов

Для снижения темпов распределения IP-адресов очень важно было определить требования к создаваемым сетям и распределению адресов в них. Сети организаций могут иметь:

- глобальную связность (Global connectivity);
- внутреннюю связность (Private connectivity).

Глобальная связность

Глобальная связность означает, что все хосты внутри организации должны иметь доступ как к хостам внутренней корпоративной сети, так и к хостам сети Internet. В этом случае хостам нужно назначать уникальные IP-адреса, которые опознавались бы как во внутренней сети организации, так и во внешней глобальной сети. Организации, которым необходима глобальная связность, должны запрашивать IP-адреса у своих сервис-провайдеров.

Внутренняя связность

Внутренняя связность означает, что доступ друг к другу должны иметь только хосты внутри корпоративной сети организации, при этом не требуется, чтобы они могли соединяться с хостами в сети Internet. Примерами хостов, нуждающихся лишь во внутренней связности, могут выступать банкоматы, электронные кассовые аппараты в магазинах розничной торговли и другое оборудование, для которого не требуется соединение с хостами вне сети компании. Внутренние хосты могут иметь уникальные IP-адреса лишь внутри корпоративной сети организации. Для этих целей организацией IANA зарезервированы три блока IP-адресов, которые предназначены для использования во внутренних сетях организаций (так называемых "частных сетях" (private internet)):

- 10.0.0.0 — 10.255.255.255 (одна сеть класса A)
- 172.16.0.0 — 172.31.255.255 (16 непрерывных блоков класса B)
- 192.168.0.0 — 192.168.255.255 (256 непрерывных сетей класса C)

Исчерпывающую информацию об их использовании, а также сведения о других зарезервированных адресах можно найти в документе RFC 19185.

Компания, выбирающая адреса для внутренней сети из вышеприведенных диапазонов, не нуждается в специальном разрешении на их использование от IANA или реестра сети Internet. Все хосты, которые получают внутренние IP-адреса, могут соединяться с хостами внутри компании, но не могут соединяться с хостами вне сети организации без специальных устройств-посредников (proxy) или шлюза (gateway). Хосты внутренней сети не смогут общаться с хостами сети Internet, потому что пакеты, исходящие из внутренней сети, имеют IP-адрес отправителя, который не определен в сети Internet, и, следовательно, они не могут пересылаться через глобальные сети общего пользования. Дело в том, что несколько компаний могут строить свои частные сети с использованием одних и тех же внутренних IP-адресов, что недопустимо. Для этих целей можно использовать только уникальные глобальные IP-адреса.

Однако следует отметить, что хосты с внутренними IP-адресами могут нормально сосуществовать с хостами, имеющими глобальные IP-адреса. На рис. 3.18 представлен пример такой среды.

Часть хостов компании может находиться во внутренней сети, а отдельные сегменты иметь выход в глобальные сети. При такой схеме половина хостов в корпоративной сети может достичь сети Internet без каких-либо дополнительных ухищрений. Компании, использующие лишь внутренние адреса, также могут организовывать доступ своим хостам в сеть Internet, но при этом требуется применение специальных сетевых фильтров, которые не пропускают пакеты с внутренними адресами в сеть Internet. Большинство провайдеров Internet все же пользуются определенными IP-адресами для организации маршрутизации с клиентами, которые используют обычные IP-адреса.

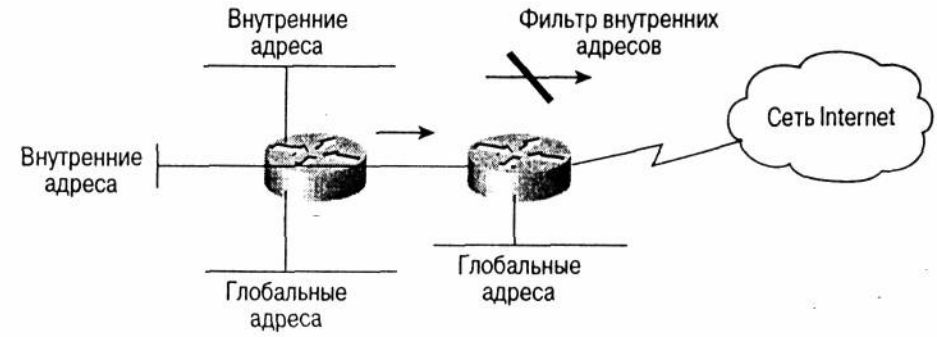


Рис. 3.18. *Общий вид сети со смешанной связностью*

Недостаток такого подхода заключается в следующем: если организация позднее решит разрешить доступ своим хостам в сеть Internet, то придется провести перенумерование всех хостов в сети с внутренних IP-адресов на глобальные IP-адреса. С появлением и внедрением новых протоколов, таких как протокол динамической конфигурации хоста Dynamic Host Configuration Protocol (DHCP)6, эта задача намного упростилась. Протокол DHCP обеспечивает механизм для передачи параметров конфигурации (включая IP-адреса) для хостов, в которых используется стек протоколов TCP/IP. Если на хостах разрешено использование протокола DHCP, то они могут получать IP-адреса динамически от центрального сервера.

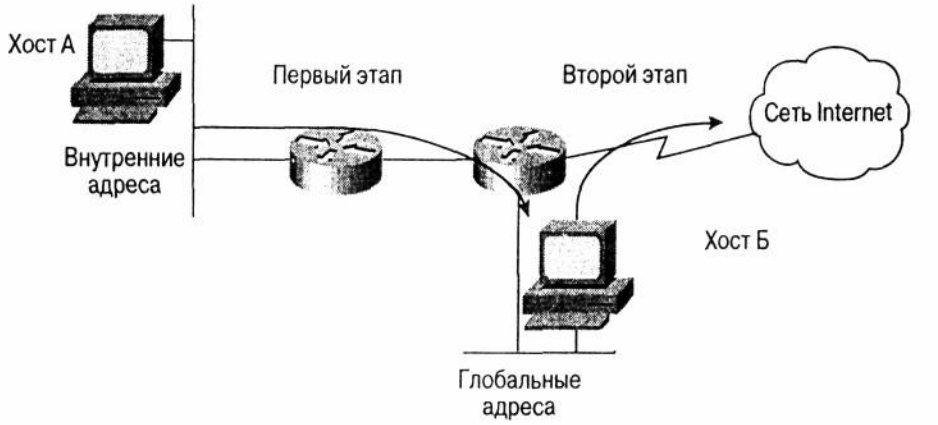


Рис. 3.19. *Доступ к ресурсам Internet из внутренней корпоративной сети*

Вторая возможность — установка так называемого бастионного хоста (bastion host), который действует как шлюз между внутренней сетью и глобальной сетью Internet. Итак, хост А (рис. 3.19) имеет внутренний IP-адрес. Если с этого хоста необходимо установить сеанс Telnet с хостом, который находится в Internet, то это можно проделать, связавшись по Telnet с хостом Б, который затем выходит в наружную сеть. Теперь все пакеты, исходящие из корпоративной сети компании, будут иметь в качестве адреса отправителя IP-адрес хоста Б, который "виден" из сети Internet. И третье — использовать транслятор сетевого адреса (Network Address Translator).

Транслятор сетевого адреса

Некоторые компании при переходе от использования внутренних IP-адресов к глобальным IP-адресам могут воспользоваться технологией трансляции (или преобразования) сетевых адресов Network Address Translator (NAT)7. Технология преобразования адресов NAT позволяет подключать корпоративные сети с внутренними адресами к сети Internet без перенумерования IP-адресов хостов во внутренней сети. Маршрутизатор с NAT располагается обычно на границе домена и преобразует внутренние IP-адреса в обычные для сети Internet глобальные адреса, и наоборот, для обеспечения нормальной работы хостов из внутренней сети с хостами в сети Internet.

На рис. 3.20 представлены хосты А и Б, которые имеют IP-адреса 10.1.1.1 и 10.1.1.2, соответственно.

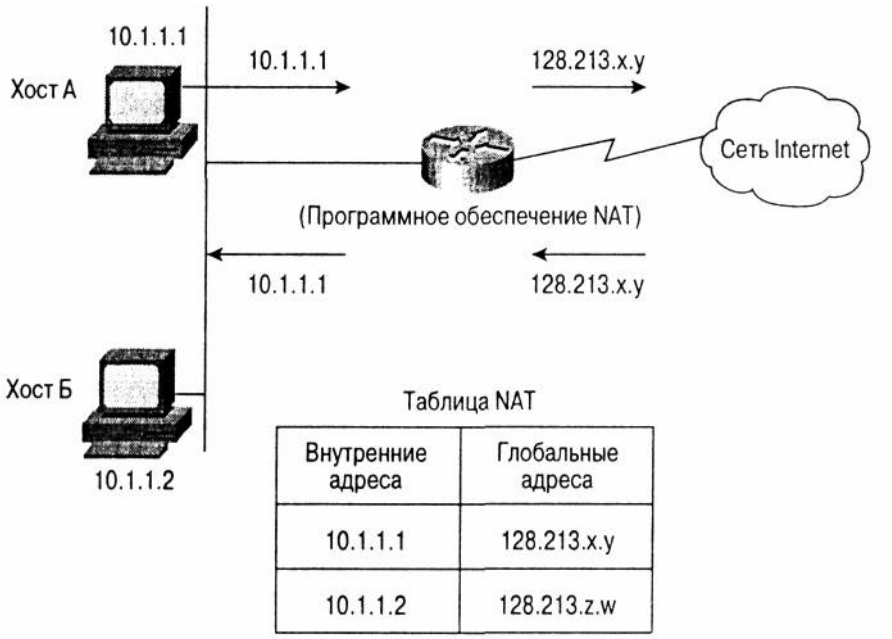


Рис. 3.20. Пример преобразования сетевых адресов

Если хостам А и Б требуется связаться с узлом вне внутренней сети компании, то NAT-устройство преобразует IP-адрес источника в пакетах согласно установленному (или динамически выделяемому) IP-адресу из таблицы NAT. Таким образом, пакеты от хоста А достигнут удаленного узла с IP-адресом источника 128.213.x.y. Хостам в глобальной сети даже не будет известно о преобразовании IP-адресов, и они будут отвечать на глобальный IP-адрес. Ответные пакеты, приходящие из глобальной сети, будут иметь IP-адрес пункта назначения узла, где проводилось преобразование сетевых адресов. Далее этот глобальный адрес согласно таблице NAT, будет преобразован во внутренний IP-адрес хоста, который начал сеанс связи.

Подробное рассмотрение работы NAT-устройств не входит в круг вопросов, рассматриваемых в книге, так как эта тема имеет много "острых углов" и достойна отдельной книги. В число частных случаев использования NAT входит ситуация, когда в сети предприятия применялись адреса, которые не предназначены IANA для корпоративного использования. В этом случае может возникнуть ситуация, когда реестром Internet эти адреса уже выделены какой-либо другой организации. Случается также, что предприятию или организации выделено меньше адресов, чем им необходимо. В этом случае с помощью NAT можно динамически транслировать внутренние IP-адреса в меньшие пулы глобальных адресов.

В принципе функционирование NAT не всегда требует выделения отдельного устройства, и часто его можно организовать на базе программного обеспечения маршрутизатора, который уже развернут в сети. Так, например, функции NAT являются частью операционной системы маршрутизаторов, выпускаемых компанией Cisco Systems — Cisco Internetwork Operating System (IOS).

IP версии 6

Протокол IP версии 6 (IPv6)8 также известен под названием IP следующего поколения (IP next generation — IPng) и является следующим шагом по улучшению существующего протокола IPv4.

Первые предложения по использованию IPng были выдвинуты в июле 1992 года на встрече группы инженеров Internet Engineering Task Force (IETF) в Бостоне (США). Было сформировано несколько рабочих групп по разработке нового протокола. Протокол IPv6 решает проблему исчерпания адресного пространства IP, закрепляет критерии качества обслуживания (Quality of Service), имеет улучшенные механизмы по автоматической

настройке узлов и аутентификации пользователей, обладает повышенной безопасностью.

В настоящее время IPv6 находится все еще в экспериментальной стадии. Для компаний и администраторов, которые вложили большие средства в развитие инфраструктуры под IPv4 нелегко перестроиться на IPv6. Пока реализации протокола IPv4 обеспечивают различные методы и технологии (какими бы громоздкими они ни казались), позволяющие выполнять основные задачи, для решения которых был разработан IPv6, переход к этой версии протокола IP не кажется компаниям необходимым. Пока невозможно точно сказать, когда начнется повальный переход на IPv6. В этой книге мы затронем лишь часть системы адресации протокола IPv6 и сравним ее с тем, что вам известно как IPv4.

Адреса IPv6 имеют 128 битов, в отличие от 32-битовых адресов IPv4. Это должно обеспечить достаточное адресное пространство, чтобы избежать проблемы его исчерпания и масштабируемости в сети Internet. При адресации с помощью 128 бит можно адресовать 2128 хостов, а это огромное количество!

Типы адресов IPv6 также определяются старшими битами в адресе в поле переменной длины, которое называется префикс формата Format Prefix (FP) (рис. 3.21).

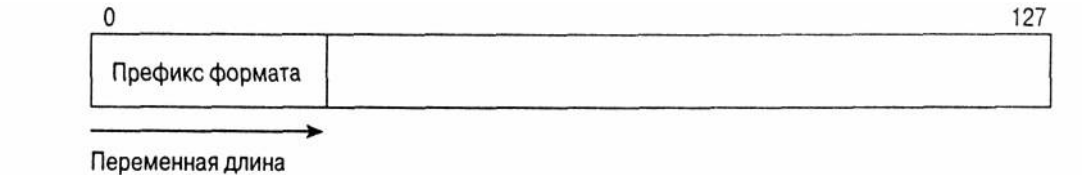


Рис. 3.21. *Формат адреса и префикса в IPv6*

В табл. 3.7 выделены начальные префиксы. Как уже отмечалось, в IPv6 определено несколько типов адресов. Нас в этой книге интересуют лишь уникальные адреса, используемые провайдерами, и локальное применение адресов IPv4 различными компаниями.

Таблица 3.7. *Распределение префиксов в IPv6*

Описание	**Префикс формата**
Зарезервирован	0000 0000
Не определен	0000 0001
Зарезервирован для NSAP	0000 001
Зарезервирован для IPX	0000 010
Не определен	0000 011
Не определен	0000 1
Не определен	0001
Не определен	001
Уникальные адреса для провайдеров	**010**
Неопределен	011
Зарезервирован для уникальных адресов географических зон	100
Неопределен	101
Неопределен	110
Неопределен	1110
Неопределен	1111 0
Неопределен	1111 10
Неопределен	1111 110
Неопределен	1111 1110
Адреса для локального пользования при организации соединений	**1111 1110 10**
Адреса для локального пользования при организации узлов	**1111 1110 11**
Групповые адреса	1111 1111

Уникальные адреса для провайдеров

Уникальные адреса для провайдеров во многом имеют сходство с глобальными IP-адресами IPv4. Их общий вид представлен на рис. 3.22.

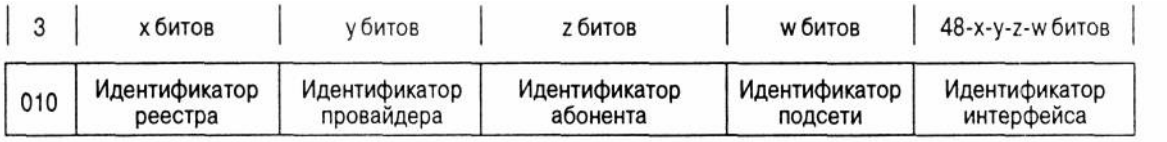


Рис. 3.22. *Формат уникального адреса IPv6 для провайдеров*

Ниже описаны поля в уникальных адресах для провайдеров.

- Префикс формата (Format prefix) — первые три бита 010, указывающие на то, что данный адрес является уникальным адресом провайдера.
- Идентификатор реестра (REGISTRY ID) — указывает, какой реестр Internet выдал идентификатор провайдера (PROVIDER ID).
- Идентификатор провайдера (PROVIDER ID) — идентифицирует провайдера, которому присвоен данный адрес.
- Идентификатор абонента (SUBSCRIBER ID) — идентифицирует абонента, подключенного к провайдеру.
- Идентификатор подсети (SUBNET ID) — идентифицирует физический канал, к которому принадлежит данный адрес.
- Идентификатор интерфейса (INTERFACE ID) — определяет интерфейс среди множества интерфейсов, принадлежащих подсети с заданным SUBNET ID. Например, в этом качестве может выступать 48-битовый адрес управления доступом к среде передачи Media Access Control (MAC), описанный стандартом IEEE-802.

Глобальные адреса IPv6 включают в себя все функции бесклассовой междоменной маршрутизации CIDR, реализованные для IPv4. Все адреса определяются иерархически, т.е. каждый элемент адреса состоит из части вышестоящего адреса (рис. 3.23).

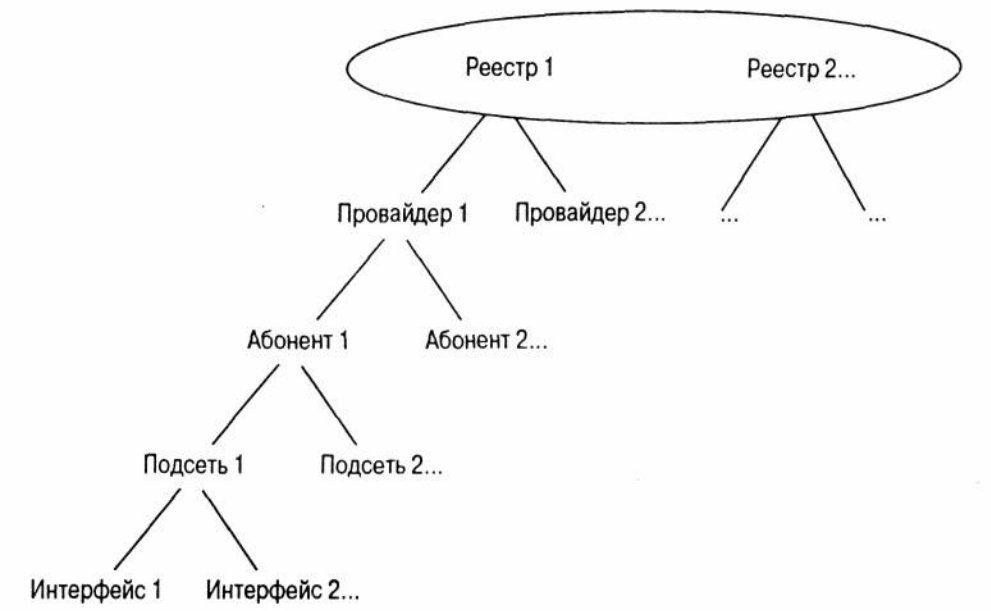


Рис. 3.23. *Иерархия в назначении адресов IPv6*

Адреса для локального пользования

Адреса для локального пользования имеют сходство с внутренними адресами для общего пользования IPv4, описанными в RFC 1918. Адреса для локального пользования делятся на две категории:

- Адреса для локального использования при организации соединений (префикс

1111111010), которые являются внутренними для определенного физического сегмента.

- Адреса для локального использования при организации узлов (префикс 1111111011), которые являются внутренними для определенного узла.

На рис. 3.24 приведен формат адресов для локального использования.

Адреса для локального использования имеют значение только для локального сегмента сети (адреса для организации соединений) и только для определенного узла (адреса для организации узлов).

Компании, которые не имеют соединения с сетью Internet, могут легко назначать собственные адреса, при этом не требуется запрашивать соответствующие префиксы из глобального адресного пространства. Если позднее компания решит подключиться к глобальной сети, то к уже имеющимся локальным адресам добавятся лишь значения REGISTRY ID, PROVIDER ID и SUBSCRIBER ID. Это одно из основных преимуществ IPv6 перед IPv4, так как в этом случае не требуется замена всех внутренних адресов глобальными или применение таблиц NAT для нормальной работы с внутренними адресами в сети Internet.

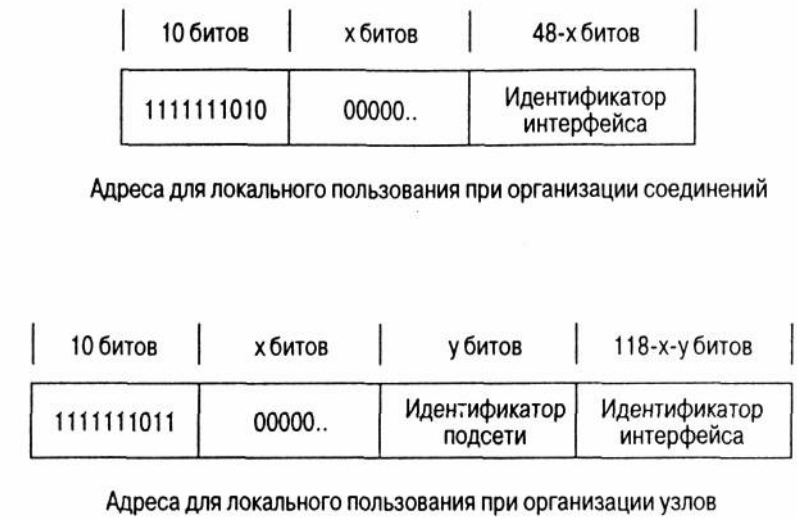


Рис. 3.24. *Форматы адресов для локального использования*

Забегая вперед

Итак, IP-адреса и система адресации являются основными элементами междоменной маршрутизации. С помощью IP-адресации определяется местонахождение нужной информации, но при этом не указывается путь, по которому можно получить доступ к этой информации. Хосты нуждаются в специальном механизме для обмена информацией о пунктах назначения и для вычисления оптимального маршрута к заданному пункту назначения. И, естественно, этим механизмом в сети является маршрутизация.

Этой главой завершается изложение фундаментального материала, который необходим для дальнейшего изучения структуры системы маршрутизации сети Internet. В следующей главе рассмотрены основы междоменной маршрутизации, концепции адресации, глобальные сети и домены, которых мы уже коснулись ранее. В следующей части книги будут рассмотрены протоколы маршрутизации и в частности протокол BGP, реализации которого подробно рассмотрены в главе 5 и далее.

Часто задаваемые вопросы:

В – Чем отличается применение VLSM от простого разбиения на подсети?

О — Применение VLSM ничем не отличается от разбиения на подсети, а является лишь расширением для проведения такого разбиения, посредством чего адреса классов A, B и C могут образовывать подсети с помощью масок различной длины.

В — Для чего может понадобиться VLSM?

О —VLSM предоставляют возможность более эффективного назначения IP-адресов. С их : помощью обеспечивается большая гибкость при назначении соответствующих адресов хостов и подсетей при ограниченном количестве IP-адресов.

В— Какая разница между CIDR и созданием суперсетей?

О - Бесклассовая междоменная маршрутизация представляет собой механизм, позволяющий сетям объявлять и суперсети, и подсети вне диапазона адресов классовой се:ти. При создании суперсетей разрешается использование сетевых масок, которые короче нормальных масок, что позволяет выделять суперсети.

В — Влияет ли классовая модель на рост глобальных таблиц маршрутов?

О — Нет. Рост глобальных таблиц маршрутов обусловлен ростом числа организаций, подключаемых к сети Internet. Дело в том, что классовая модель не предусматривает масштабируемого решения для того, чтобы справиться с подобным ростом.

В — В моей сети используются устаревшие протоколы, такие как RIP-1 и IGRP. Что мне может понадобиться при переходе на новые протоколы маршрутизации, которые поддерживают CIDR и VLSM?

О — Если вы считаете, что реализация в сети VLSM и CIDR поможет более эффективно использовать ваше адресное пространство и улучшит общую маршрутизацию в сети, то вам следует провести модернизацию. Одним из критериев при принятии такого решения является способность вашего аппаратного обеспечения работать с новыми протоколами, которые предъявляют повышенные требования к производительности и размерам оперативной памяти. Конечно, это зависит и от протокола, на который вы переходите. Нужно также не забывать о необходимости обеспечения совместимости при совместной работе нового и старого протоколов. Ввиду того что модернизация сетей проводится, как правило, поэтапно, рано или поздно вы столкнетесь с ситуацией, когда новый и старый протоколы работают в сети параллельно. Так как старые протоколы не "понимают" VLSM и CIDR, то может понадобиться широкое применение статической маршрутизации для обеспечения нормальной работы домена в переходный период.

В — Могу ли я объединять любые маршруты в таблице маршрутов?

О – Вы можете объединять лишь те маршруты, за администрирование которых несете ответственность. При объединении маршрутов вне своего домена вы можете спровоцировать появление "черных дыр" в маршрутизации.

В – Если я меняю провайдеров, могу ли я не менять свои IP-адреса?

О – На сегодняшний день, с точки зрения обеспечения наилучшего объединения маршрутов рекомендуется (а иногда и требуется), чтобы вы вернули старые адреса бывшему провайдёру и получили новые от вашего нового провайдера. Проконсультируйтесь у своего провайдера о правилах подключения клиентов к сети, которыми он руководствуется.

В — У меня есть хосты, которым необходимо обеспечить выход в Internet, остальным же доступ в Internet требуется закрыть. Могу ли я использовать внутренние IP-адреса на одних хостах в сети, и глобальные на других?

О — Да, совместное применение внутренних и глобальных IP-адресов в одной сети полне допустимой При этом при объявлении маршрутов в сеть провайдера, вам следует указать лишь легальные IP-сети (не внутренние).

В — Мне необходимо подключиться к сети Internet, но не все адреса в моей сети зарегистрированы в соответствующих инстанциях. Я не могу провести перенумерование всех хостов в сети. Что мне следует делать в этом случае?

О — В таком случае вы можете воспользо&аться процедурой преобразования сетевых адресрв Network Address Translation (NAT), которая позволяет преобразовать незарегистрированные адреса в диапазон легальных IP-адресов, полученных от вашего провайдера.

Ссылки

1 RFC 791, "Internet Protocol (IP)", www.isi.edu/in-notes/rfc791.txt
2 RFC 917, "Internet Subnets", www.isi.edu/in-notes/rfc917.txt
3 RFC 1878, "Variable Length Subnet Table for IPv4", www.isi.edu/in-notes/rfcl878.txt
4 RFC 1519, "Classless Inter-Domain Routing (CIDR)", www.isi.edu/in-notes/rfcl519.txt
5 RFC 1918, "Address Allocation for Private Internets", www.isi.edu/in-notes/rfcl918.txt
6 RFC 1541, "Dynamic Host Configuration Protocol", www.isi.edu/in-notes/rfcl541.txt
7 RFC 1631, "The IP Network Address Translator", www.isi.edu/in-notes/rfcl631.txt
8 RFC 1884, "IP version 6 Addressing Architecture", www.isi.edu/in-notes/rfcl884.txt

Часть II. Основы протоколов маршрутизации

В этой части...

Глава 4. Основы междоменной маршрутизации
Глава 5. Протокол граничного шлюза Border Gateway Protocol версии 4

Несмотря на то что эта книга посвящена в основном протоколам внешнего : шлюза, т.е. маршрутизации между различными автономными системами, мы все-таки решили сначала рассмотреть некоторые вопросы функционирования протоколов внутреннего шлюза, так как и концептуально, и на практике они влияют друг на друга. Так, глава 4 начинается с описания протоколов, предназначенных для организации маршрутизации между автономными системами. Затем мы рассмотрим протоколы внешнего шлюза, в частности протокол граничного шлюза Border Gateway Protocol (BGP). В главе 5 описаны протоколы BGP версии 4 (BGP-4) и процедуры ведения переговоров с соседними маршрутизаторами. Кроме того, в главе 5 приведены сведения о мультипротокольных расширениях для BGP-4, правилах ведения переговоров (Capabilities Negotiation) в BGP-4 и о параметре шифрованной подписи MD5 (TCP MD5 Signature Option) для BGP. Понимание основ работы протокола BGP, описанных в части II, необходимо для перехода к практическому использованию его возможностей при рещёнии проблем маршрутизации и для восприятия дальнейшего материала книги.

Ключевые темы этой главы:

- **Обзор маршрутизаторов и схем маршрутизации.** Приводится краткий обзор основных вопросов и задач маршрутизации, а также протоколов внутреннегошлюза (Interior Gateway Protocols — IGP). Подчеркиваются основные отличияэтих протоколов от протоколов внешнего шлюза, рассматриваемых в следующей главе.
- **Концепции протоколов маршрутизации.** В этом разделе делается обзор маршрутизации дистанционно-векторных алгоритмов и алгоритмов, основанных на анализе состояния канала.
- **Разделение мира на автономные системы.** Автономные системы представляют собой наборы маршрутизаторов, которые совместно используют один и тот же набор правил маршрутизации (routing policy). Существует множество конфигураций автономных системой их типы варьируются в зависимости от требуемого количества точек выхода во внешние сети или с учетом того, будет ли разрешено прохождение транзитного трафика через автономную систему.

Глава 4. Основы междоменной маршрутизации

Сеть Internet представляет собой конгломерат автономных систем (autonomous systems), которые делят между собой зоны административной ответственности и определяют для различных организаций правила маршрутизации. Автономные системы создаются на основе маршрутизаторов, которые могут работать с протоколами внутреннего шлюза Interior Gateway Protocols (IGP), такими как: протокол информации о маршрутах Routing Information Protocol (RIP), расширенный протокол внутреннего шлюза Enhanced Interior Gateway Protocol (EIGRP), протокол предпочтительного выбора кратчайшего пути Open Shortest Path First (OSPF) и протокол обмена маршрутной информацией между промежуточными системами Intermediate System-to-Intermediate System (IS-IS). Все эти протоколы и граничные с ними взаимодействуют посредством протокола внешнего шлюза Exterior Gateway Protocol (EGP). В настоящее время стандартом де-факто для сети Internet является протокол граничного шлюза версии 4 (Border Gateway Protocol Version 4 — BGP-4), описанный в RFC 17711.

Обзор маршрутизаторов и схем маршрутизации

Маршрутизаторы — это устройства, которые регулируют направление трафика между хостами. Они создают маршрутные таблицы, которые содержат информацию обо всех возможных маршрутах ко всем известным гол узлам. Ниже описаны этапы маршрутизации.

1. На маршрутизаторах запускаются специальные программы, которые называют*протоколами маршрутизации (routing protocols).* Эти программы служат для приема и передачи маршрутной информации другим маршрутизаторам в сети.
2. Маршрутизаторы используют информацию о маршрутах для заполнения своих таблиц маршрутов, которые связаны с соответствующим протоколом маршрутизации.
3. Маршрутизаторы сканируют таблицы маршрутов различных протоколов маршрутизации (если используется несколько протоколов маршрутизации) и выбирают одинили несколько наилучших путей для доставки трафика в пункт назначения.
4. Маршрутизаторы взаимодействуют при передаче трафика со следующими ближайшими устройствами (next-hop devices), обладающими адресом канальногоуровня, и с локальными интерфейсами, которые используются для пересылкипакетов в пункты назначения. Обратите внимание, что в качестве следующего ближайшего устройства может выступать еще один маршрутизатор и даже просто удаленный узел, которому предназначен пакет.
5. Информация о пересылке для следующих ближайших устройств (адрес канального уровня и исходящий интерфейс) помещается на маршрутизаторе в таблицумаршрутов пересылки (forwarding table).

6. Когда маршрутизатор принимает пакет, он анализирует заголовок пакета и выделяет адрес получателя.
7. Далее маршрутизатор консультируется с таблицей маршрутов пересылки и получает оттуда сведения об исходящем интерфейсе (через что передавать) и обадресе следующего ближайшего устройства, откуда можно попасть в пункт назначения (куда передавать).
8. Кроме того, маршрутизатор выполняет все необходимые дополнительные функции(такие как уменьшение значения "времени жизни" IP TTL и управление параметрамитипа сервиса IP TOS) и затем пересылает пакет на соответствующее устройство.
9. Все эти действия продолжаются, пока пакет не достигнет хоста получателя. Такая схема отражает парадигму о маршрутизации через промежуточные узлы(hop-by-hop), которая обычно применяется в сетях с коммутацией пакетов.

Причиной разработки протоколов внешнего шлюза, таких как BOP, было то, что протоколы внутреннего шлюза не очень хорошо масштабировались в сетях, крупнее уровня предприятия, с тысячами узлов и сотнями тысяч маршрутов. Протоколы внутреннего шлюза не были предусмотрены для этих целей. В этой главе мы рассмотрим только основы функционирования протоколов IGP.

Пример простейшей маршрутизации

На рис. 4.1 представлены три маршрутизатора — RTA, RTB и RTC, объединяющие три локальных вычислительных сети — 192.10.1.0, 192.10.5.0 и 192.10.6.0 посредством последовательных соединений. Каждое последовательное соединение имеет собственный сетевой адрес, т. е. в результате получаются три дополнительные сети — 192.10.2.0, 192.10.3.0 и 192.10.4.0. Каждая сеть имеет свою метрику, которая показывает уровень сложности служебных операций (вес) при передаче трафика по определенному каналу. Так, например, соединение между RTA и RTB имеет вес 2000, что намного выше веса соединения между RTA и RTC, равного 60. На самом деле соединение между RTA и RTB может быть организовано на скорости 56 Кбит/с с более существенными задержками, чем цифровой канал типа T1 между RTA и RTC и между RTC и RTB.

Маршрутизаторы RTA, RTB и RTC посредством одного из протоколов IGP обмениваются сетевой информацией и строят в соответствии с ней свои таблицы маршрутов. На рис. 4.1 представлены примеры таблицы маршрутов на маршрутизаторе RTA для двух различных ситуаций. В одном случае маршрутизаторы обмениваются информацией о маршрутах по протоколу RIP, а в другом — с помощью протокола OSPF.

В качестве примера того, как трафик перемещается между конечными станциями, рассмотрим такой случай. Допустим, что хост 192.10.1.2 попытался передать пакет на хост 192.10.6.2. При этом ему придется использовать маршрут, вручную установленный по умолчанию, и переслать пакет сначала на маршрутизатор RTA. Далее RTA просматривает свою таблицу маршрутов в поиске сети, к которой принадлежит узел получателя, и выясняет, что сеть 192.10.6.0 доступна через следующий ближайший узел 192.10.3.2 (RTC), соединенный последовательным каналом под номером 2 (S2). Маршрутизатор RTC, получив пакет, попытается найти получателя в своей таблице маршрутов (она не показана на рисунке). Затем маршрутизатор RTC обнаружит, что хост получателя напрямую подключен к его интерфейсу Ethernet O (EO) и передаст пакет на хост 192.10.6.2.

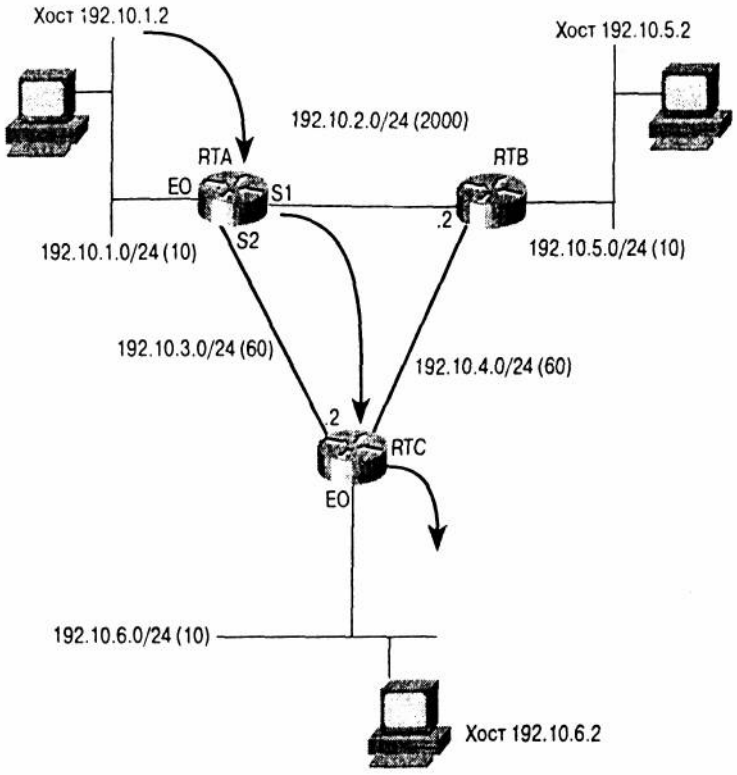


Таблица IP-маршрутов для маршрутизатора RTA (RIP)

Пункт назначения	Следующий узел	Кол-во переприемов
192.10.1.0	Подключен (E0)	-
192.10.2.0	Подключен (S1)	-
192.10.3.0	Подключен (S2)	-
192.10.4.0	192.10.2.2 (S1) 192.10.3.2 (S2)	1 1
192.10.5.0	192.10.2.2 (S1)	1
192.10.6.0	192.10.3.2 (S2)	1

Таблица IP-маршрутов для маршрутизатора RTA (OSPF)

Пункт назначения	Следующий узел	Кол-во переприемов
192.10.1.0	Подключен (E0)	-
192.10.2.0	Подключен (S1)	-
192.10.3.0	Подключен (S2)	-
192.10.4.0	192.10.3.2 (S2)	120
192.10.5.0	192.10.3.2 (S2)	130
192.10.6.0	192.10.3.2 (S2)	70

***Рис. 4.1**. Механизм простейшей маршрутизации*

В этом примере маршрутизация происходит одинаково и при использовании протокола RIP, и при работе по протоколу OSPF. Однако следует помнить, что RIP относится к дистанционно-векторным протоколам, а OSPF — к протоколам маршрутизации на основе анализа состояния канала связи. Для различных примеров маршрутизации на базе схемы, представленной на рис. 4.1, при применении RIP и OSPF могут быть получены различные результаты. Теперь вполне закономерно перейти к рассмотрению характеристик обеих категорий протоколов IGP, истории их развития и тенденцией к общему усложнению системы маршрутизации.

Концепции протоколов маршрутизации

Большинство протоколов маршрутизации, используемых сегодня, основано на одном из двух алгоритмов распределенной маршрутизации: анализ состояния канала и дистанционный вектор. В последующих разделах мы обсудим различные свойства, присущие алгоритмам дистанционного вектора и анализа состояния канала.

Дистанционно-векторные протоколы маршрутизации

Дистанционно-векторные протоколы маршрутизации иногда именуются протоколами Беллмана-Форда (Bellman-Ford) в честь изобретателей алгоритма вычислений кратчайших маршрутов2, которые впервые описали механизм распределенного применения этого алгоритма3. Термин *дистанционный вектор (distance vector)* возник ввиду того, что в протоколе имеется вектор (список) расстояний (счетчик переприемов или другие параметры), который связан с каждым префиксом получателя, содержащимся в сообщении о маршруте.

Дистанционно-векторные протоколы маршрутизации, такие как протокол маршрутной информации Routing Information Protocol (RIP), при расчете маршрута используют механизм распределенных вычислений для каждого префикса пункта назначения. Другими словами для работы дистанционно-векторных протоколов необходимо, чтобы каждый узел отдельно занимался вычислением наилучшего маршрута (исходящего соединения) для каждого префикса пункта назначения.

Выбрав наилучший маршрут, маршрутизатор посылает дистанционные векторы своим соседям, уведомляя их таким образом о доступности каждого из пунктов назначения и о метриках маршрутов, которые выбраны для доставки данных в соответствующий пункт назначения. Параллельно соседние с маршрутизатором узлы также вычисляют наилучший маршрут к каждому пункту назначения и уведомляют своих соседей о доступных маршрутах (и связанных с ними метриках), с помощью которых можно достичь заданного пункта назначения. На основании квитанций (отчетных сообщений) от соседей, где детально описывается маршрут к пункту назначения и его метрики, маршрутизатор может "решить", что существует лучший маршрут через другого соседа. Затем он повторно рассылает уведомления об имеющихся маршрутах и их метриках своим соседям. Эти процедуры повторяются до тех пор, пока все маршрутизаторы не определят наилучшие маршруты для каждого пункта назначения.

Начальные спецификации дистанционно-векторных протоколов, таких как RIP версии 1 (RIP-1), имели серьезные недостатки. Например, подсчет количества переприемов был единственной метрикой в RIP-1, которая использовалась при выборе маршрута. Кроме того, этот протокол имел несколько ограничений. Рассмотрим, например, маршрутные таблицы маршрутизатора RTA (рис. 4.1). В одной из них представлена информация о маршрутах, собранная протоколом RIP, а в другой — протоколом OSPF (этот протокол маршрутизации на основе анализа состояния канала будет обсуждаться в последующих разделах).

При использовании RIP-1 маршрутизатор RTA выберет прямое соединение между RTA и RTB, чтобы достичь сети 192.10.5.0. Маршрутизатор RTA выбирает это соединение потому, что при непосредственном соединении для того, чтобы достичь заданной сети, используется лишь один переприем через узел RTB, против двух переприемов при выборе маршрута через узлы RTC и RTB. Однако маршрутизатор RTA "знает" о том, что канал RTA-RTB имеет меньшую производительность и большое время задержки, а канал RTC-RTB обеспечит более высокое качество обслуживания.

С другой стороны, при использовании протокола OSPF и метрик при выборе маршрута, помимо подсчета количества переприемов, маршрутизатор RTA обнаружит, что путь к маршрутизатору RTB через RTC (вес: 60 + 60 = 120; 2 переприема) является более оптимальным, чем прямой путь (вес: 2000; 1 переприем).

Еще при подсчете переприемов следует учитывать ограничения, налагаемые на количество переприемов, т. е. их не может быть бесконечное множество. В дистанционно-векторных протоколах (например, в RIP-1) количество переприемов ограничено, как правило, числом 15. При превышении этого предела узел считается недоступным по заданному маршруту. Таким образом, распространение информации о маршрутах в больших сетях также вызывало определенные проблемы (в тех из них, где насчитывалось более 15 узлов на маршрут). Зависимость от количества переприемов — одна из определяющих

характеристик дистанционно-векторных протоколов, хотя более новые протоколы этой категории (RIP-2 и EIGRP) не столь строги.

Еще один недостаток — способ обмена маршрутной информацией. Для традиционных дистанционно-векторных протоколов в настоящее время применяется следующая концепция: маршрутизаторы ведут обмен всеми IP-адресами, которые могут быть достигнуты при периодическом обмене данными посредством широковещательных анонсов дистанционных векторов. Эти широковещательные сообщения рассылаются согласно "таймеру обновлений" (refresh timer), установленному для каждого сообщения. Таким образом, если истекает срок работы "таймера обновлений" и при этом поступает новая маршрутная информация, требующая пересылки соседям, этот таймер сбрасывается, и маршрутная информация не пересылается до тех пор, пока срок работы таймера снова не истечет. Теперь рассмотрим, что бы произошло, если бы соединение или определенный маршрут вдруг стали недоступны по каким-либо причинам сразу после обновления маршрутов. Распространение маршрутной информации со сведениями о нерабочем маршруте было бы задержано на время до окончания срока работы "таймера обновления", следовательно, возникло бы значительное замедление при обновлении маршрутной информации.

К счастью, в новые модификации дистанционно-векторных протоколов, таких как EIGRP и RIP-2, введена концепция *триггерных обновлений (triggerred updates).* Триг-герные обновления распространяют сообщения об отказах по мере их появления, что значительно ускоряет обмен маршрутной информацией.

Итак, можно сделать вывод о том, что в крупных и даже небольших сетях с большим количеством узлов периодический обмен таблицами маршрутов с соседними узлами может быть очень большим по объему, что затрудняет обслуживание и замедляет обмен маршрутной информацией. Нагрузка на процессоры и каналы связи, вызванная периодическим обменом маршрутной информацией, также может негативно влиять на общую производительность сети. Еще одно свойство, которым обладают новые дистанционно-векторные протоколы, — повышенная надежность при передаче дистанционных векторов между соседями, что исключает необходимость периодически повторять полные таблицы маршрутов.

Конвергенция (convergence) — это интервал времени, за который обновляются все маршруты в сети, т.е. устанавливается факт существования, отсутствия или изменения того или иного маршрута. Старые дистанционно-векторные протоколы работали по принципу периодического обновления маршрутов с использованием таймеров удержания: если в течение определенного времени информация о маршруте не поступала, то этот маршрут "замораживался" (удерживался) и исключался из таблицы маршрутов. Процесс удержания и исключения из таблицы маршрутов в больших сетях мог длиться несколько минут, пока не проходила полная конвергенция, т. е. пока всем узлам сети сообщалась информация об исчезновении маршрута. Задержка между моментом, когда маршрут становился недоступным, и его исключением из таблицы маршрутов могла привести к образованию временных петель или даже "черных дыр".

В некоторых дистанционно-векторных протоколах (например, в RIP) при пропадании активного маршрута и его появлении, но уже с более высокой метрикой (предположительно сгенерированной другим маршрутизатором, который сообщил о возможном альтернативном маршруте) маршрут по-прежнему остается в "замороженном" состоянии. Таким образом, время конвергенции для всей сети остается достаточно большим.

Еще один серьезный недостаток дистанционно-векторных протоколов первого поколения — их классовая природа и отсутствие полноценной поддержки VLSM и CIDR. При обновлении маршрутной информации эти дистанционно-векторные протоколы не передают сведения о сетевых масках и, следовательно, не могут поддерживать эти технологии. В протоколе RIP-1 маршрутизатор, принимающий обновление маршрутов через определенный интерфейс, будет подставлять в эту посылку свою локальную маску подсети. Протокол IGRP делает то же самое, что и RIP-1, но он, кроме того, привязывается к сетевым маскам сетей класса A, B и C, если часть переданного сетевого адреса не соответствует локальному сетевому адресу. Все это приводит к определенным затруднениям (в том случае,

если интерфейс принадлежит сети, которая разбита на подсети с помощью масок переменной длины) и неправильной интерпретации принимаемых обновлений маршрутов. В новейших дистанционно-векторных протоколах, таких как RIP-2 и EIGRP, указанные недостатки устранены.

С целью исправления недостатков старых дистанционно-векторных протоколов маршрутизации было разработано несколько их модификаций. Так, например, протоколы RIP-2 и EIGRP уже поддерживают работу с VLSM и CIDR. К тому же протоколы IGRP и EIGRP способны воспринимать сложные метрики, которые используются для представления характеристик, соединений составляющих маршрут (таких как полоса пропускания, текущая нагрузка, задержки, размер передаваемого блока (MTU) и т.д.), с помощью которых можно вычислить более оптимальный маршрут, чем при простом подсчете числа переприемов.

Простота и завершенность дистанционно-векторных протоколов стала причиной их широкой популярности. Основной недостаток протоколов этого класса — медленная конвергенция, что может стать катализатором образования петель и "черных дыр" при изменении топологии сети. Однако в последних модификациях дистанционно-векторных протоколов, в частности в EIGRP, достигается довольно хорошая конвергенция.

Этот раздел мы не могли бы завершить, не упомянув, что протокол BGP также относится к семейству дистанционно-векторных протоколов. Кроме обычных параметров, свойственных этим протоколам, в BGP используется дополнительный механизм, именуемый *вектором маршрута (path vector),* благодаря которому устраняется проблема ограничения числа переприемов. По сути, вектор маршрута содержит список доменов маршрутизации (номера автономных систем), по которому пролегает тот или иной маршрут. Если домен получает информацию о маршруте, который уже имеет идентификатор домена, то такой маршрут игнорируется. Эта маршрутная информация позволяет избежать образования петель маршрутизации. Кроме того, ее можно использовать как основу для создания правил маршрутизации в домене. Этот атрибут маршрута более подробно обсуждается в последующих главах.

Протоколы маршрутизации на основе анализа состояния канала

Протоколы маршрутизации на основе анализа состояния канала (link-state routing protocols), такие как протокол первого кратчайшего открытого маршрута Open Shortest Path First (OSPF)4 и протокол обмена данными между промежуточными системами Intermediate Sys-tem-to-Intermediate System (IS-IS)5, используют в работе модель распределенных баз данных и считаются более сложными протоколами маршрутизации. Протоколы с анализом состояния канала работают на основе обмена между маршрутизаторами специальными сообщениями, которые называются *отчетами о состоянии канала (link states).* В этих отчетах содержится информация о соединениях и узлах домена маршрутизации. Это означает, что на маршрутизаторах, где запущены протоколы анализа состояния канала, не проводится обмен маршрутными таблицами, как это делается в дистанционно-векторных протоколах. Вместо этого маршрутизаторы обмениваются информацией о ближайших соседях и о сетях, а также сведениями о метрике для каждого своего соединения.

Один из способов рассмотрения работы протоколов маршрутизации на основе анализа состояния канала очень похож на составление картинки-головоломки (паззла). Каждый маршрутизатор в сети генерирует один из элементов головоломки (отчет о состоянии канала), где описывается его состояние и способ соединения с другими элементами головоломки. Кроме того, для соединения каждого элемента головоломки он предоставляет список соответствующих им метрик. Элемент головоломки, который представляет сам маршрутизатор, затем рассылается по всей сети от одного маршрутизатора к другому до тех пор, пока все узлы в домене не получат его копию. После завершения этой процедуры все маршрутизаторы в сети будут иметь копии каждого элемента головоломки и хранить их в так называемой *базе данных состояния каналов (link-state database).* Далее маршрутизаторы автономно собирают головоломку, в результате чего на каждом из них

хранятся идентичные копии всех маршрутизаторов в сети.

Затем, применяя алгоритм кратчайшего пути (shortest path first — SPF), который более известен как алгоритм Дейкстра, маршрутизаторы вычисляют дерево кратчайших маршрутов к каждому удаленному узлу, помещая себя в корень этого дерева.

Ниже перечислены преимущества протоколов анализа состояния канала.

- Не ведется подсчет количества переприемов. Отсутствуют какие-либо ограничения на количество переприемов, составляющих маршрут. Протоколы анализасостояния канала работают на основе анализа метрики канала, а не подсчетаколичества переприемов.

  Для того чтобы убедиться в том, что эти протоколы не зависят от подсчета числа переприемов, обратимся к таблицам маршрутов для маршрутизатора RTA, представленным на рис. 4.1. В случае применения протокола OSPF, маршрутизатор RTA для того, чтобы достичь маршрутизатора RTB, подбирает оптимальный маршрут на основе анализа весовых коэффициентов различных соединений. В его таблице маршрутов следующим узлом на пути в сеть 192.10.5.0 (RTB) значится узел 192.10.3.2 (RTC). Это коренным образом отличается от работы по протоколу RIP, при котором выбирается неоптимальный маршрут.
- Предоставление сведений о полосе пропускания. Полоса пропускания канала изадержки в нем могут быть (вручную или динамически) учтены при расчетекратчайшего маршрута к заданному узлу. Это позволяет сбалансировать нагрузку на канал лучше, чем при подсчете переприемов.
- Лучшая конвергенция. Изменения в канале и на узле моментально распространяются по всему домену посредством отчетов о состоянии канала. Все маршрутизаторы в домене будут постоянно обновлять свои таблицы маршрутов (этипроцедуры немного походят на триггерные обновления).
- Поддержка VLSM и CIDR. Протоколы анализа состояния канала позволяютвести обмен информацией о сетевых масках как части информационных элементов, пересылаемых в домене. В результате сети с масками переменной длины могут легко распознаваться и маршрутизироваться.
- Улучшенная иерархическая структура. В то время как сети с дистанционно-векторными протоколами являются плоскими, протоколы анализа состоянияканала обеспечивают механизмы для разбиения домена на уровни или области.Иерархическая структура сети позволяет лучше выявлять нестабильные участки.

Несмотря на то что алгоритмы анализа состояния связей повышают масштабируемость при маршрутизации, что позволяло применять их в более крупных и сложных сетях, они подходят лишь для внутренней маршрутизации. Сами по себе эти протоколы не предлагают глобального решения для многосвязной сети с междоменной маршрутизацией, которой является сеть Internet. В очень больших сетях и при колебаниях маршрутов, что может быть вызвано нестабильностью канала связи, повторная передача информации о состоянии канала и связанный с этим пересчет метрик создает нежелательную нагрузку на отдельный маршрутизатор.

Хотя более подробное обсуждение протоколов IGP не входит в число вопросов, затрагиваемых в этой книге, мы приведем две прекрасные ссылки на книга, где рассмотрены дистанционно-векторные протоколы маршрутизации и протоколы анализа состояния канала: *"Межсетевые соединения, 2-е изд: мосты, маршрутизаторы, коммутаторы и межсетевые протоколы" Радии Перлман ("Interconnections, Second Edition: Bridges, Routers, Switches and Internetworking Protocols" Radio Perlmanf* и *"OSPF: Анатомия протокола маршрутизации" Джона Т. Мойя ("OSPF: Anatomy of an Internet Routing Protocol" John T. May)1.*

Организуя внутреннюю маршрутизацию между автономными системами (AS), большинство крупных сервис-провайдеров используют протоколы анализа состояния канала в силу их способности к быстрой конвергенции. Чаще всего из протоколов анализа состояния канала применяются протоколы OSPF и IS-IS.

Многие провайдеры, которые уже давно присутствуют на рынке, в качестве протокола IGP выбрали для себя протокол IS-IS, а более молодые провайдеры используют OSPF либо также IS-IS. Изначально в старых сетях чаше применялся протокол IS-IS, так как

правительство США требовало поддержки стандартов ISP CLNP в тех сетях, с которыми заключались федеральные контракты. (Следует обратить ваше внимание на то, что протокол IS-IS способен передавать информацию как уровня CLNP, так и сетевого уровня IP, в то время как протокол OSPF рассчитан для работы только по протоколу IP). Однако если обратиться к истории Internet, основным "руководящим фактором" при выборе протоколов маршрутизации первые провайдеры избрали более стабильную реализацию протокола IS-IS, в отличие от только зарождавшегося OSPF. Очевидно, что стабильность протокола при выборе ЮР имела для провайдеров решающее значение.

Сегодня в сетях провайдеров широко используются оба протокола — и IS-IS, и OSPF. Завершенность и стабильность IS-IS явилась результатом того, что он успешно используется в крупных сетях и продолжает применяться в развертываемых сегодня сетях.

Разделение мира на автономные системы

Внешние протоколы маршрутизации были разработаны для управления разросшимися таблицами маршрутов и для повышения структурированности сети Internet путем разделения доменов маршрутизации на различные административные единицы, которые называются *автономными системами (autonomous systems — AS),* имеющими собственные независимые правила маршрутизации и уникальные внутренние протоколы IGP.

На начальном этапе развития в сети Internet использовался протокол внешнего шлюза EGP8 (Exterior Gateway Protocol) (не путайте с обобщающим названием протоколов внешнего шлюза!). Так, в сети NSFNET этот протокол использовался для обмена информацией о взаимной достижимости между магистральной и региональными сетями. Хотя протокол EGP применялся очень широко, его ограничения по топологии, неэффективность в распознавании петель маршрутизации и при задании правил маршрутизации породили потребность в новом универсальном протоколе, лишенном этих недостатков. В настоящее время стандартом де-факто для организации междоменной маршрутизации в сети Internet является протокол BGP-4.

Примечание

Обратите внимание, что основное отличие внутренней и внешней маршрутизации в AS заключается в том, что маршрутизация внутри AS обычно оптимизирована под технические требования, в то время как внешняя маршрутизация AS отражает политические и деловые отношения между сетями и компаниями, которым они принадлежат.

Статическая маршрутизация, маршрутизация по умолчанию и динамическая маршрутизация

Прежде чем представить вам основные способы подключения автономных систем к провайдерам Internet, приведем несколько определений и положений.

- *Статическая маршрутизация* означает, что маршруты к узлам задаются вручную,или статически, по мере их поступления на маршрутизатор. В этом случае достижимость той или иной сети не зависит от наличия и состояния самой сети. Длястатических маршрутов не имеет значения, можно ли в данный момент доставитьтрафик в пункт назначения или нет. В любом случае эти маршруты будут находиться в таблице маршрутов, а трафик пересылаться в заданном направлении.
- *Маршрутизация по умолчанию* — это, как говорят, "последнее средство" накрайний случай. Трафик, который необходимо переслать на узел, неизвестныймаршрутизатору,

отправляется по маршруту, заданному по умолчанию. Маршрутизация по умолчанию представляет собой простейшую форму маршрутизации в домене с одной точкой выхода.

- *Динамическая маршрутизация* означает, что все маршруты изучаются маршрутизатором с помощью внутреннего или внешнего протокола маршрутизации.Здесь достижимость той или иной сети напрямую зависит от существования исостояния сети. Если узел получателя выключен, то маршрут к нему исчезает изтаблицы маршрутов, и трафик не отправляется по указанному адресу.

Эти три подхода к организации системы маршрутизации и представляют собой возможные конфигурации AS, которые мы рассмотрим в последующих разделах, но существует еще один подход — оптимальный (т.е. комбинация нескольких подходов). Таким образом, говоря в этой главе о различных автономных системах, мы рассмотрим динамический, статический подходы, маршрутизацию по умолчанию и оптимальный подход, который является комбинацией нескольких подходов. В этой главе также рассматривается вопрос о том, какие протоколы лучше применять в автономных системах -— внешние или внутренние. Подробно схемы маршрутизации для автономных систем различной топологии рассматриваются в главе 6, "Настройка параметров BGP".

Всегда помните, что статическая маршрутизация и маршрутизация по умолчанию не являются вашими врагами. Наиболее стабильная (но иногда менее гибкая) конфигурация всегда основана на статической маршрутизации. Многие ошибочно полагают, что они безнадежно отстали, так как не используют в сети динамическую маршрутизацию. Попытки насильственного внедрения динамической маршрутизации, когда в этом нет необходимости, являются растратой полосы пропускания, усилий и денег. Помните о принципе KISS для информационно-вычислительных систем, который мы приводили в предыдущей главе!

Автономные системы

Автономная система (autonomous system — AS) — это набор маршрутизаторов, имеющих единые правила маршрутизации, управляемых одной технической администрацией и работающих на одном из протоколов IGP (для внутренней маршрутизации AS может использовать и несколько ЮР). Для всей остальной сети AS является конечным простейшим элементом и воспринимается как единое целое. Реестром сети Internet либо провайдером каждой AS назначается уникальный идентификационный номер. Маршрутизация между различными AS осуществляется с помощью протокола внешнего шлюза, такого как BGP-4 (рис. 4.2).

Давайте разберемся, какие преимущества имеет разбиение крупной сети на административные участки (с учетом iv.ro, что сеть Internet благодаря использованию протоколов OSPF или IS-IS могла бы быть одной сложной сетью). Более мелкие сети, представляемые в качестве AS, способны реализовывать собственные правила маршрутизации, которые бы уникально их характеризовали и описывали все услуги, предоставляемые другими сетями. Теперь в каждой AS можно запускать свой пакет протоколов ЮР, независимо от того, какие наборы ЮР запущены на других AS.

В последующих разделах мы рассмотрим конфигурации сетей с заглушками или одноканальных сетей (single-homed), многоканальных нетранзитных сетей (multi-homed nontranzit) и многоканальных сетей с транзитом (multihomed tranzit).

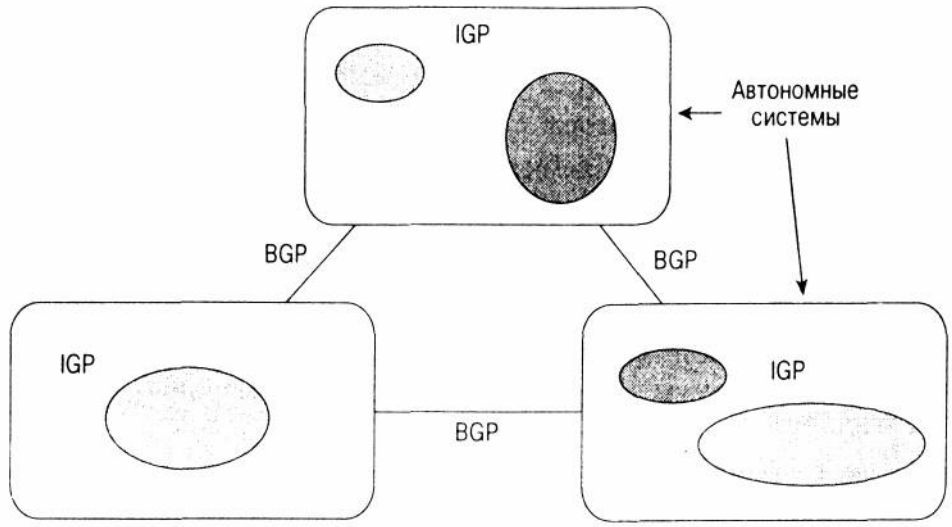


Рис. 4.2. *Обмен маршрутной информацией между автономными системами*

AS с заглушками

Автономная система считается системой с заглушкой (stub) при условии, что все маршруты из нее к другим сетям проходят через одну точку. Эти AS также называют *одноканальными (single-homed)* по отношению к провайдерам. На рис. 4.3 приведен пример одноканальной AS, или AS с заглушкой.

Рис. 4.3. Одноканальная AS (AS с засушкой)

Одноканальным AS не нужно получать информацию обо всех маршрутах в Internet от провайдера. Ввиду того что для таких AS имеется всего один выход во внешний мир, весь трафик может по умолчанию отправляться провайдеру. Имея AS такой конфигурации, провайдер может использовать различные способы для объявления другим провайдерам маршрутов в клиентскую сеть.

Один из таких способов — хранение записей о статических маршрутах в подсети клиентов на маршрутизаторе провайдера. Затем провайдер может с помощью BOP объявить все эти статические маршруты по сети Internet. Этот метод обеспечивает хорошую масштабируемость, если маршруты в клиентские сети представлены небольшими наборами объединенных маршрутов. Когда клиентам выделяется слишком большое число подсетей с адресами, которые следуют вразнобой, то их содержание в списке на маршрутизаторе становится неэффективным. Лучше всего, если клиентам выделены непрерывные блоки IP-адресов. Этим достигается наилучшее объединение маршрутов.

В качестве альтернативы для объявления маршрутов к сетям своих клиентов провайдер может использовать протоколы IGP. Эти протоколы могут использоваться между клиентом и провайдером для рассылки маршрутной информации. Такая схема предоставляет преимущества динамической маршрутизации, когда любые изменения и другая сетевая информация динамически посылаются на узел провайдера. Однако этот метод не нашел широкого применения из-за низкой масштабируемости и ввиду того, что нестабильность канала делает нестабильной работу самих протоколов IGP.

И третий способ, с помощью которого провайдер может получать информацию о

маршрутах и объявлять о маршрутах к клиентским сетям, — применение протокола BGP между клиентом и провайдером. При работе в AS с заглушкой довольно сложно зарегистрировать номер AS в реестре Internet, так как правила маршрутизации в клиентских сетях являются по сути дополнением к правилам маршрутизации провайдера.

Примечание

В RFC 1930[9] приведен набор основных указаний по подбору, созданию и регистрации номеров автономных систем.

Провайдер также может присвоить клиентской AS номер из диапазона номеров для частных AS (65412—65535), допуская, что правила маршрутизации провайдера обеспечивают поддержку работы частного пространства AS у своих клиентов, как это описано в RFC 227010.

Для обеспечения работы между провайдером и клиентом можно использовать несколько комбинаций протоколов. На рис. 4.4 представлены возможные конфигурации протоколов между клиентом и провайдером — в качестве примера использована AS с заглушкой. (Разница между протоколами EBGP и IBGP будет рассмотрена в последующих разделах). Как видите, провайдеры могут переносить маршрутизаторы клиента в свои точки присутствия или свои маршрутизаторы — на технические площадки клиентов. Обратите внимание, что не во всех случаях требуется, чтобы клиент с провайдером работал по протоколу BGP.

Многоканальные AS без транзита

Автономная система считается многоканальной (multihomed), если в ней имеется более одной точки выхода во внешний мир. Автономная система может быть многоканальной по отношению к одному или к нескольким провайдерам. В нетранзитных (nontranzit) AS не разрешается транзит трафика через автономную систему. *Транзитным* считается трафик, отправитель и получатель которого находится вне данной AS. На рис. 4.5 показана многоканальная AS (AS1), которая не является транзитной и подключена к двум провайдерам (ISP1 и ISP2).

В AS без транзита объявляются только ее собственные маршруты, а информация о маршрутах от других AS не распространяется. Таким образом, трафик для пункта назначения, который не принадлежит данной AS, не будет направляться на нее. На рис. 4.5 автономная система AS1 получает сведения о маршрутах п3 и п2 через ISP1 и сведения о маршрутах п5 и пб — через ISP2. При этом AS1 сама объявляет только свои локальные маршруты (nl, п2). Она не передает на узел провайдера ISP2 информацию о маршрутах, полученную от ISP1, а провайдеру ISP1 — сведения о маршрутах, полученные от ISP2. Таким образом AS1 не открыта для прохождения внешнего трафика, например, в ситуации когда ISP1 попытается достичь п5 и пб или ISP2 попытается через AS1 выйти на п3 или п4. Конечно, провайдеры ISP1 и 1SP2 могут заставить свой трафик следовать через AS1 с помощью статической маршрутизации или прописав маршрут по умолчанию. Чтобы избежать такой ситуации, AS1 фильтрует входяший трафик по адресу получателя и на основе принадлежности его к данной AS пропускает или отвергает его.

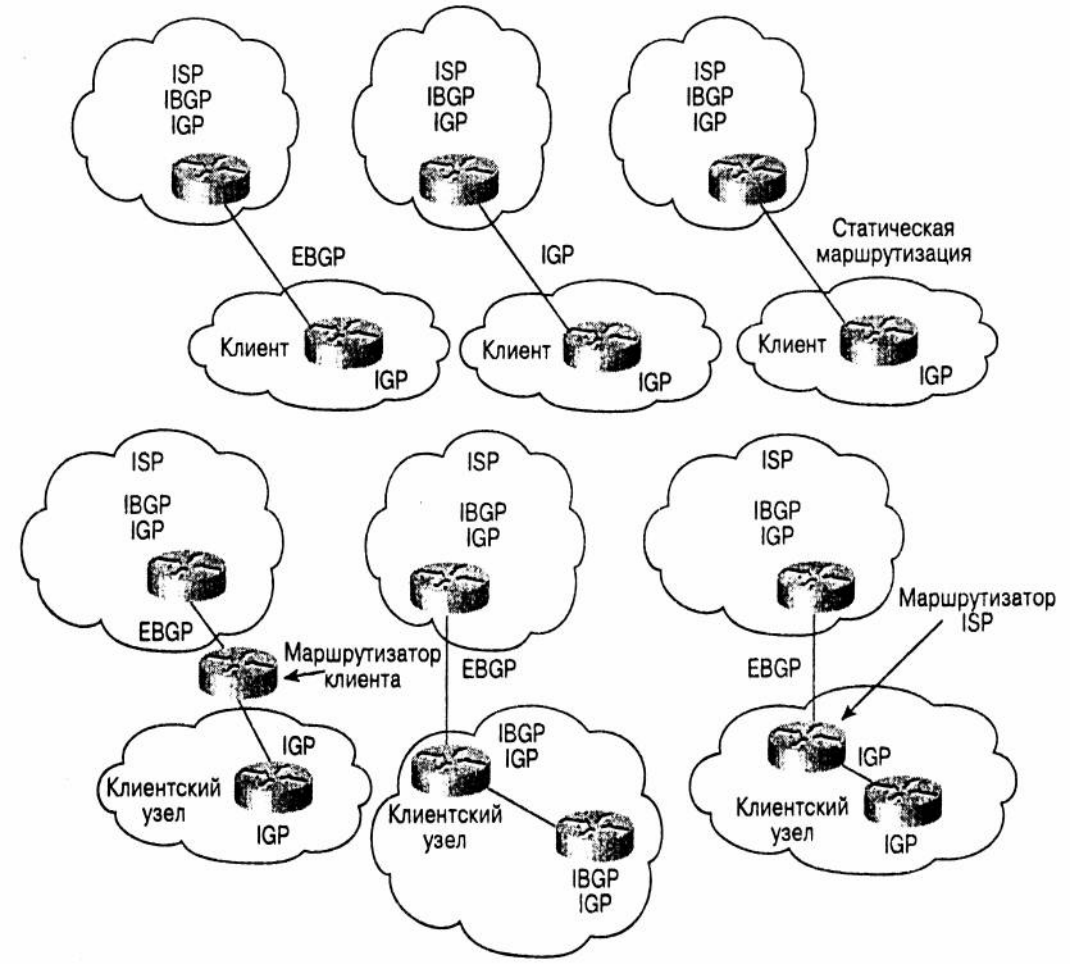


Рис. 4.4. *AS с заглушкой: варианты реализации различных протоколов*

В принципе, многоканальные AS без транзита не нуждаются в организации работы по протоколу BOP со своими провайдерами, хотя это и рекомендуется, а в большинстве случаев и требуется провайдером. Как вы увидите далее из материалов книги, работа с провайдерами по протоколу BGP-4 имеет множество преимуществ и для контроля за распространением маршрутов, и для фильтрации.

Многоканальные транзитные AS

Многоканальные транзитные AS также имеют несколько соединений с внешним миром и могут использоваться для транзита трафика в интересах других AS (рис. 4.6). Транзитным (по отношению к многоканальной AS) является любой трафик, отправитель и получатель которого не принадлежат к локальной AS.

Хотя протокол BGP-4 является протоколом внешнего шлюза, он может использоваться и внутри AS для обмена обновлениями маршрутов на основе протокола BOP. Соединения между маршрутизаторами на базе протокола BOP внутри автономных систем относятся к *внутреннему BGP (Internal BGP — IBGP),* в то время как соединения между маршрутизаторами различных автономных систем относят к *внешнему BGP (External BGP— EBGP).* Маршрутизаторы, работающие на базе IBGP, называют также *транзитными маршрутизаторами.* Они занимаются пересылкой транзитного трафика, поступающего в AS.

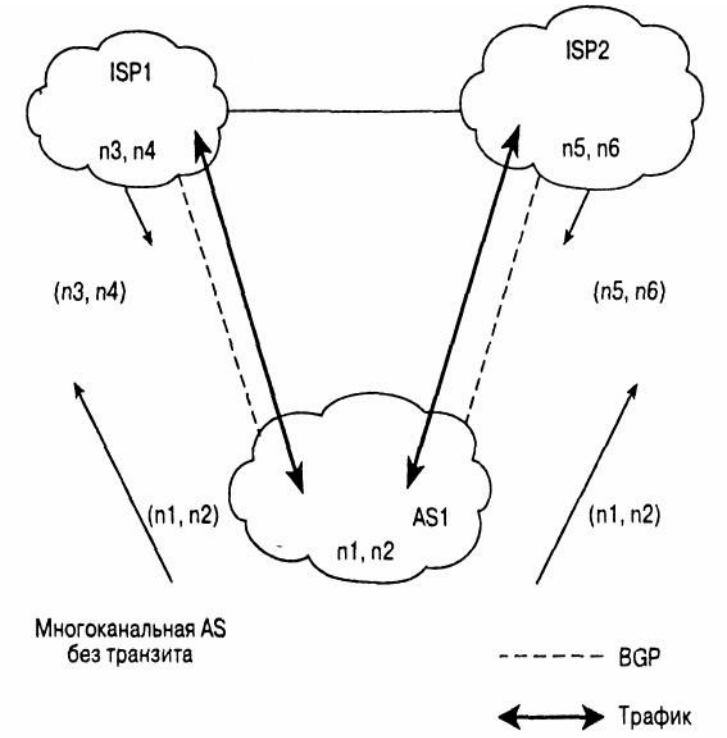


Рис. 4.5. Пример многоканальной AS без транзита

Транзитные AS объявляют и маршруты, полученные ими от других AS. Таким образом, транзитная AS будет открыта для трафика, который адресован в другую AS. В многоканальных транзитных AS рекомендуется использовать протокол BGP-4 для обеспечения соединений с другими AS и защищать внутренние не транзитные маршрутизаторы от получения маршрутов из сети Internet. Необязательно, чтобы все маршрутизаторы в домене работали по протоколу BGP. Внутренние маршрутизаторы без транзита вполне могут обойтись маршрутами по умолчанию на маршрутизаторы с BGP, которые уменьшают количество маршрутов, приходящихся на маршрутизаторы без транзита. В сетях наиболее крупных сервис-провайдеров все маршрутизаторы обычно обрабатывают весь набор BGP маршрутов самостоятельно.

На рис. 4.6 представлена многоканальная транзитная автономная система AS1, подключенная к двум различным провайдерам, ISP1 и ISP2. Автономная система AS1 получает сведения о маршрутах п3, п4, п5 и пб и от ISP1, и от ISP2, а затем, добавив свои локальные маршруты, делится этой информацией с провайдерами ISP1 и ISP2. В этом случае провайдер ISP1 может использовать AS1 в качестве транзитной AS для того, чтобы отправить трафик в сети п5 и пб, а провайдер ISP2 может использовать AS 1 для того, чтобы достичь сетей п3 и п4.

Забегая вперед

Протокол граничного шлюза Border Gateway Protocol определил архитектуру системы маршрутизации в сети Internet. Разделение сетей на автономные системы позволило логически провести административные и политические границы между различными организациями. Протоколы внутреннего шлюза Interior Gateway Protocols могут использоваться сегодня независимо друг от друга, но вместе с тем сети по-прежнему могут для обеспечения глобальной маршрутизации взаимодействовать друг с другом по протоколу BGP.

В главе 5, "Протокол граничного шлюза Border Gateway Protocol версии 4" Дается детальный обзор работы протокола BGP-4, включая детальное описание форматов заголовков сообщений этого протокола.

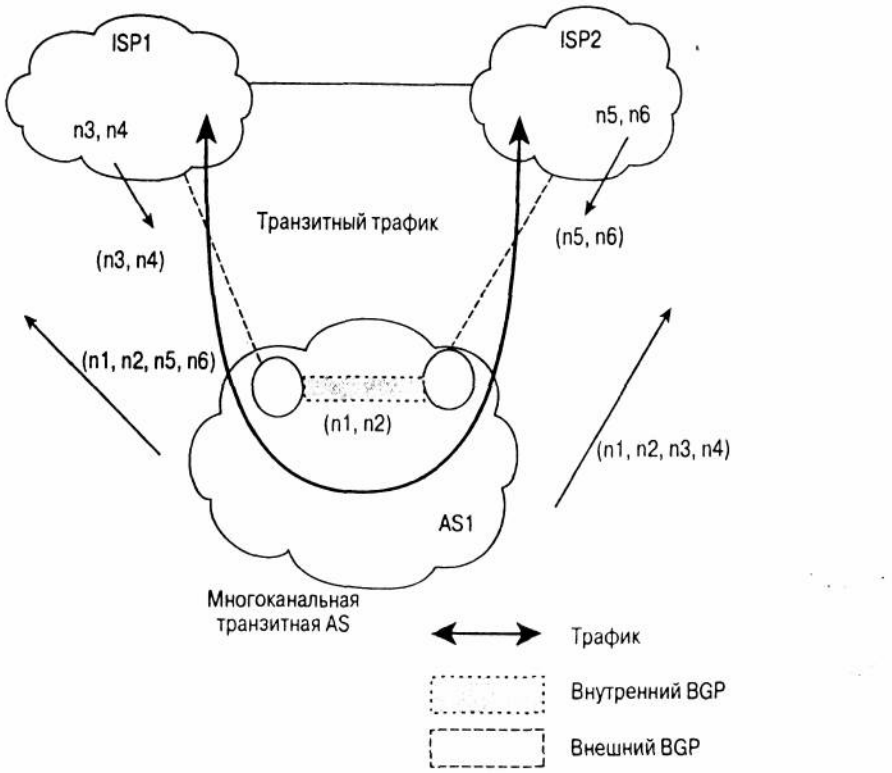


Рис. 4.6. *Многоканальная транзитная AS с использованием внешнего и внутреннего протоколов BGP*

Часто задаваемые вопросы

В — Какая разница между доменом и автономной системой?

О:— Оба термина применяются для обозначения совокупности маршрутизаторов. Термин "домен" обычно применяется к набору маршрутизаторов, в котором используется одинаковый протокол маршрутизации, например RIP-домен или домен OSPF. Автономная система представляет собой по сути один или несколько доменов, которые находятся в ведении одной администрации и имеют унифицированные правила маршрутизации с другими AS.

В — Моя компания подключена к провайдеру Internet по протоколу RIP. Нужно ли мне перейти к использованию BGP?

О —Если в ближайшем будущем предполагается подключение к нескольким прбвайдерам, вам следует обсудить возможность перехода к использованию BGP. Если же ваш трафик не нуждается в многоканальной системе, подключенной к нескольким провайдерам, то достаточно и имеющихся возможностей, предоставляемых протоколом RIP.

В – У меня уже есть одно соединение с провайдером по одному из протоколов IGP. Я хотел бы подключиться к этому же провайдеру, но в другом месте. Могу я это сделать только с помощью IGP или мне нужно перейти на BGP?

О — В этом случае все зависит от провайдера. Некоторые провайдеры могут предоставить вам возможность подключения нескольких узлов по протоколам IGP, другие предпрчитают использовать BGP. При использовании BGP вы сможете лучше управлять своим трафиком, но об этом в следующей главе.

В — Я думал, что BGP может использоваться лишь между AS. Я не очень понимаю, как можно использовать BGP внутри AS?

О – При рассмотрении BGP внутри AS (IBGP) представте себе своего рода туннель по которому пересылается информация о маршрутах, если ваша AS является транзитной, то IBGP защитит все ваши внутренние нетранзитные маршрутизаторы от переполнения информацией о внешних маршрутах. С другой стороны, даже если у вас AS без транзита, то

все равно по мере чтения книги вы поймете преимущества в управлении точками входа и выхода трафика, которые дает вам протокол IBGP.

В — Вы все время упоминаете только о BGP-4, а использует ли кто-нибудь BGP-1,-2 или -3? И что такое протокол EGP?

О— Протокол BGP-4 является де-факто стандартным протоколом междоменной маршрутизации, который используется в сети Internet. Протоколы EGP, BGP-1, 2 и 3 в настоящее время вышли из употребления. Поддержка CIDR в BGP-4, дополнительные обновления сведений о маршрутах и улучшенные механизмы .фильтрации и задания правил маршрутизации подталкивают всех к перёходу на этот новый протокол.

В —Я планирую организовать второе соединение с моим лровайдером Internet. Нужно ли мне получить номер AS в региональном реестре сети lntemet?

О — Получение номера для AS на самом деле необязательно. Возможно, ваш провайдер выделит вам номера из диапазона для частных AS для клиентов, которые подключаются к одному провайдеру по нескольким каналам. Вы можете проконсультироваться со специалистами регионального реестра Internet по вопросу выделения номеров автономным системам в сетях, подключенных только к одному провайдеру.

Ссылки

1 RFC 1771, "A Border Gateway Protocol 4 (BGP-4)," www.isi.edu/in-notes/rfcl771.txt

2 Bellman, R. *Dynamic Programming* (Princeton University Press, 1957)

3 Ford, L. R., Jr. And D. R. Fulkerson. *Flows in Networks* (Princeton University Press, 1962)

4 RFC 1583, "OSPF version 2," www.isi.edu/in-notes/rfcl583.txt

5 ISO 10589, "Intermediate System to Intermediate System"; RFC 1195, "Use of OSIIS-IS for Routing in TCP/IP and Dual Environments," www.isi.edu/in-notes/rfcll95.txt

6 Perlman, Radia. *Interconnections, Second Edition: Bridges, Routers, Switches and Internetworking Protocols* (Boston, Mass.: Addison-Wesley Longman, Inc., 1998)

7 Moy, John. *OSPF: Anatomy of an Internet Routing Protocol* (Boston, Mass.: Addison-Wesley Longman, Inc., 1998)

8 RFC 904, "Exterior Gateway Protocol Formal Specification," www.isi.edu/in-notes/rfc904.txt

9 RFC 1930, "Guidelines for creation, selection, and registration of an Autonomous System (AS)," www.isi.edu/in-notes/rfcl930.txt

10 RFC 2270, "Using a dedicated AS for Sites Homed to a Single Provider, "www.isi.edu/in-notes/rfc2270.txt”

Ключевые темы этой:

- **Как работает BGP.** Рассмотрены вопросы функционирования протокола BGP версии 4, включая:ф6рматьт заголовков сообщений и фазы ведения переговоров с соседними маршрутизаторами. Мы рассмотрим назначение и форматы четырех основных типов сообщений BGP — PEN, NOTIFICATION, KEEPALIVE и UPDATE.
- **Мультипротокольные расширения** для BGP-4. Здесь мы обсудим мультипротокольные расширения для протокола BGP, специально разработанные для работы в междоменной среде с групповыми адресами, но способные взаимодействовать и с другими протоколами.
- **Возможности ведения переговоров в BGP-4.** Возможности ведения переговоров в протоколе BGP обеспечивают механизм четкой реализации новых возможностей BGP. Их мы обсудим более подробно следующих главах.
- **Параметр зашифрованной подписи TCR-MD5 для BGP.** Параметр зашифрованной подписи для BGP (TCP MD5 Signature Option) был добавлен в протокол для зашита от ложных TCP-сегментов, вызывающих сброс сеансов TCP. Мы рассмотрим использование зашифрованной подписи и дадим несколько рекомендаций по обеспечению безопасности при работе BGP.

Глава 5. Протокол граничного шлюза Border Gateway Protocol версии 4

Протокол граничного шлюза Border Gateway Protocol (BGP) претерпел несколько изменений с момента выхода его первой версии BGP-1 в 1989 году. Повсеместное внедрение BGP-4 началось в 1993 году. Это первая из версий BGP, в которой появились возможности агрегации (объединения), что позволило реализовать бесклассовую междоменну маршрутизацию (classless interdomain routing — CIDR), и обеспечить поддержку суперсетей.

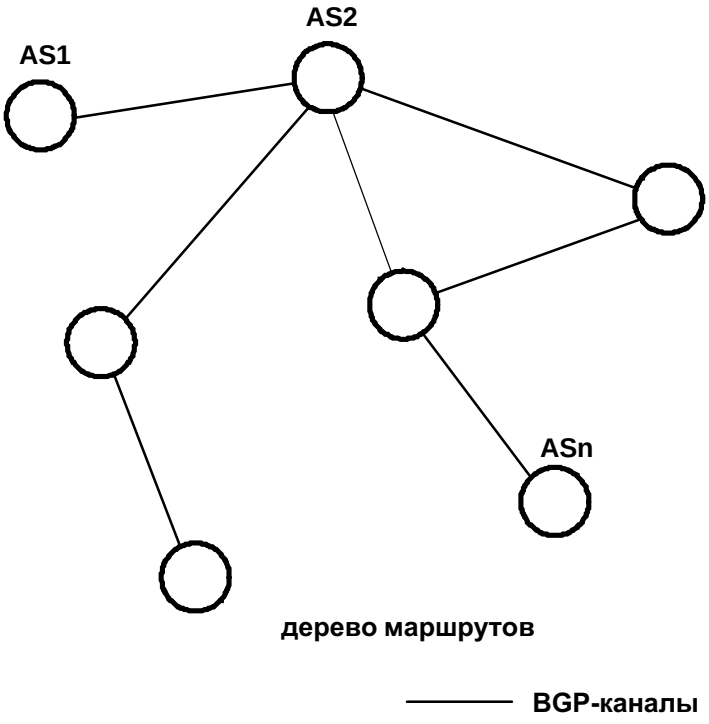


Рис.5. 1. *Пример дерева маршрутов между AS*

Протокол BGP не предъявляет никаких требований к топологии сети. Принцип его действия предполагает, что маршрутизация внутри автономной системы выполняется с помощью внутренних протоколов маршрутизации, или, как их еще называют, интра-протоколов (например, Interior Gateway Protocol — IGP). В этой книге термин *"внутренний" (intra)* обозначает все, что относится к действиям внутри субъекта, а термин *"внешний " (inter)* означает события или действия, которые имеют место между субъектами. Протоколом BOP на основе информации, полученной от различных маршрутизаторов, выстраивается граф автономных систем со всеми связями между узлами. Такой граф иногда называют *деревом.* Если рассматривать сеть Internet "глазами" протокола BGP, то это будет граф, состоящий из автономных систем (AS), где каждой AS соответствует уникальный номер. Соединение между двумя AS формирует путь, а информация о совокупности путей от одного узла в AS к узлу в другой AS составляет маршрут. Протокол BGP активно использует информацию о маршрутах к заданному пункту назначения, что позволяет избежать образования петель маршрутизации между доменами. На рис. 5.1 представлена концепция дерева маршрутов, положенная в основу протокола BGP.

Как работает BGP

Протокол BGP является протоколом вектора маршрута и используется для обмена маршрутной информацией между автономными системами. Термин *вектор маршрута (path vector)* происходит из самого принципа действия BGP: маршрутная информация содержит последовательности номеров AS, через которые прошел пакет с заданным префиксом сети. Маршрутная информация, связанная с префиксом, используется для профилактики образования петель в маршрутах.

В качестве транспортного протокола в BGP используется протокол TCP (порт 179). Таким образом вся надежность доставки (включая повторную передачу) возлагается на протокол TCP и не требует отдельной реализации в самом BGP, что естественно упрощает механизмы надежности в BGP.

Маршрутизаторы, которые работают с протоколом BGP, часто называют *спикерами BGP (BGP speakers).* Два спикера BGP, образующих TCP-соединение друг с другом для обмена маршрутной информацией, называют *соседними (neighbors)* или *взаимодействующими (peers).* На рис. 5.2 показана схема такого взаимодействия. Взаимодействующие маршрутизаторы сначала обмениваются открытыми сообщениями для того, чтобы определить параметры соединения. Эти сообщения используются для согласования параметров, таких как номер версии BGP и др.

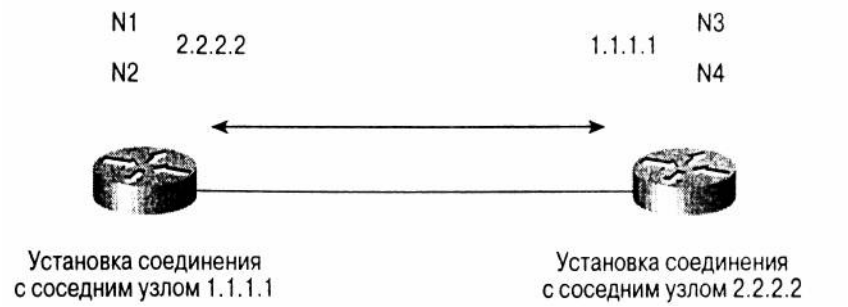


Рис. 5.2. *Маршрутизаторы BGP становятся соседями*

Протокол BGP также обеспечивает очень изящный механизм закрытия соединения с соседним маршрутизатором. Другими словами, в случае негативного исхода переговоров между взаимодействующими маршрутизаторами, что может быть результатом несовместимости их конфигураций, вмешательства оператора или вызвано другими причинами, генерируется и посылается сообщение об ошибке NOTIFICATION. Получив это сообщение, вторая сторона должна прекратить попытки установить соединение или разорвать его, если оно было установлено ранее. Преимущество такого механизма заключается в том, что обе стороны уведомляются о невозможности установки соединения и не тратят свои мощности на обслуживание этого соединения или попытки повторно установить связь. Процедура закрытия также гарантирует, что обе стороны до закрытия сеанса TCP, получат все сообщения об ошибках, в частности сообщение NOTIFICATION.

В начале сеанса BOP между несколькими спикерами BOP ведется обмен всеми маршрутами, которые могут далее использоваться в работе по протоколу BOP (рис. 5.3). После того как соединение установлено и проведен начальный обмен маршрутами, по сети рассылается лишь информация о новых маршрутах — так называемые *инкремент-ные обновления (incremental updates).* Применение инкрементных обновлений, по сравнению с периодическим обновлением маршрутов, которое использовалось в других протоколах, таких как EGP, позволило многократно увеличить производительность центральных процессоров на маршрутизаторах и разгрузить полосу пропускания.

Согласно протоколу BOP, пара маршрутизаторов уведомляется о маршрутах и изменениях в них с помощью сообщения UPDATE. Сообщение UPDATE, помимо другой полезной информации, содержит список записей типа <длина, префикс> (<Iength, prefix>),

указывающих на список узлов, на которые можно доставить трафик через спикер BOP. В сообщение UPDATE также включены атрибуты маршрута. К ним относятся: степень предпочтения определенного маршрута и список AS, через которые пролегает маршрут.

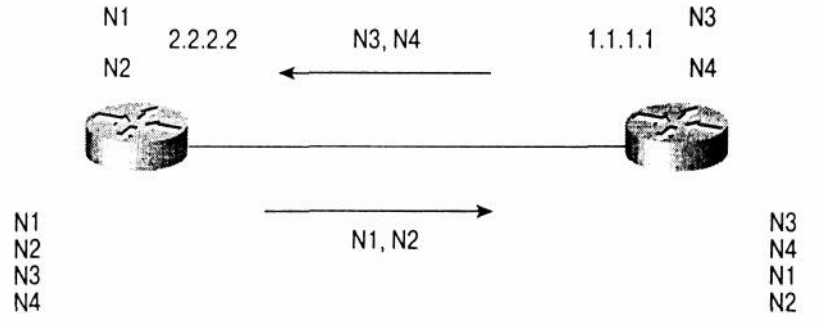


Рис. 5.3. *Обмен обновлениями маршрутной информации*

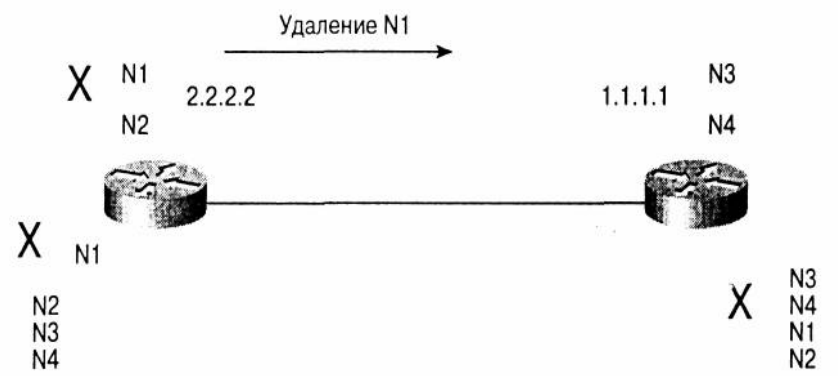


Рис. 5.4. *Маршрут NJ выходит из строя. Посылается частичное обновление*

В случае если маршрут становится недействительным, т.е. по нему невозможно достичь пункта назначения, спикер BOP информирует об этом своих соседей и удаляет недействительный маршрут. Как показано на рис. 5.4, удаляемые маршруты также включаются в сообщение UPDATE. Таким образом, эти маршруты уже нельзя использовать. Если же информация о маршруте изменилась или для того же префикса выбран новый маршрут, то процедура удаления не выполняется; в этом случае достаточно лишь объявить о замене маршрута.

На рис. 5.4 показана система в *уравновешенном состоянии (steady state):* если нет никаких изменений в структуре маршрутов, то маршрутизаторы обмениваются только пакетами KEEPALIVE.

Сообщения KEEPALIVE периодически посылаются между соседними маршрутизаторами BOP, чтобы убедиться, что соединение находится в нормальном состоянии. Пакеты KEEPALIVE (длиной 19 байт каждый) не создают практически никакой нагрузки на процессор маршрутизатора и полосу пропускания, так как им требуется очень незначительная полоса пропускания (один 152-битовый пакет каждые 60 секунд, т.е. около 2,5 байт/с).

В протоколе BGP учитывается номер версии таблицы маршрутов, чтобы отслеживать изменения маршрутов. Если в таблицу маршрутов вносятся какие-либо изменения, то BGP автоматически увеличивает номер версии таблицы. Быстро растущие номера версии таблицы обычно указывают на то, что в сети имеется нестабильно работающий участок (хотя это довольно обычная ситуация для сетей крупных провайдеров Internet). Нестабильность сетей, подключенных к Internet по всему миру, приводит к росту номеров версий таблиц маршрутов на каждом спикере BGP, где имеются сведения обо всех маршрутных таблицах сети Internet. Для снижения воздействия этих неоднородностей в Internet были разработаны механизмы коммутации маршрутов и другие мероприятия (более подробно о них в главе 10, "Проектирование стабильных сетей на базе TCP/IP").

Формат заголовка сообщения протокола BGP

Формат заголовка сообщения в BGP представляет собой поле маркера длиной 16 байт, за которым следует поле длины (2 байта) и поле типа (1 байт). На рис. 5.6 представлен формат заголовка сообщения протокола BGP.

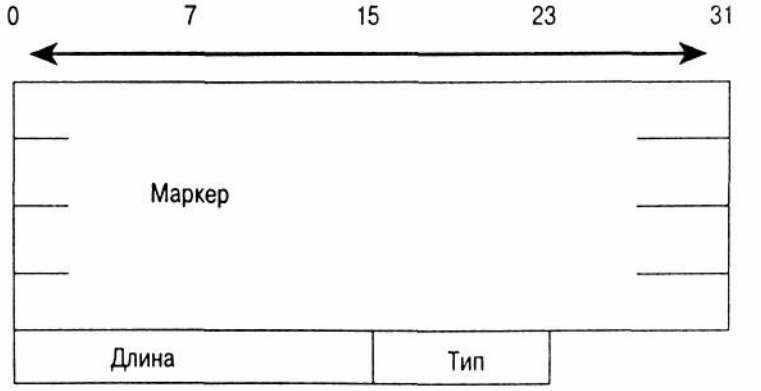


Рис. 5.6. *Формат заголовка сообщения BGP*

В зависимости от типа сообщения в сообщении протокола BGP за заголовком может следовать или не следовать блок данных. Так, например, сообщения KEEPALIVE состоят только из заголовка и никаких данных не передают.

Поле маркера длиной 16 байт используется для аутентификации входящих сообщений BGP либо для детектирования потери синхронизации между двумя взаимодействующими по BGP маршрутизаторами. Поле маркера бывает двух форматов.

- Если послано сообщение типа OPEN или в нем отсутствует информация об аутентификации, то в поле маркера все позиции выставляются в 1.
- В другом случае значение поля маркера вычисляется в соответствии с используемым механизмом аутентификации. Далее в этой главе мы рассмотрим, каким образом на формирование маркера влияет применение параметра зашифрованной подписи TCP MD5.

Поле длины размером 2 байта используется для отображения полной длины сообщения BOP, включая заголовок. Наименьшая длина сообщения BOP составляет 19 байт (16+2+1), а наибольшая — 4096 байт.

Поле типа размером 1 байт определяет тип сообщения. Возможны следующие значения:

- OPEN (Открытие соединения)
- UPDATE (Обновление маршрутной информации)
- NOTIFICATION (Уведомление об ошибке)
- KEEPALIVE (Проверка состояния соединения)

В последующих разделах мы рассмотрим назначение и формат каждого типа сообщений более детально.

Переговоры с соседними BGP-узлами

Одно из основных положений протокола BOP состоит в том, что взаимодействующие узлы устанавливают между собой сеансы BOP. Если этот этап по каким-либо причинам не был выполнен, то обмен маршрутной информацией или ее обновление не проводится. Переговоры с соседними узлами основаны на успешном установлении соединения по протоколу TCP, успешной обработке сообщения OPEN и периодическом обмене сообщениями UPDATE и KEEPALIVE.

Формат сообщения OPEN

На рис. 5.7 представлен формат сообщения OPEN.

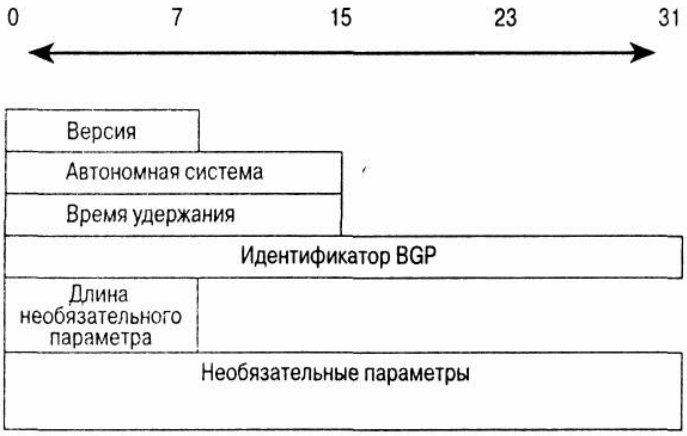


Рис. 5.7. *Формат сообщения OPEN*

Рассмотрим назначение каждого из полей сообщения OPEN.

- **Версия** (Version) — целое число длиной 1 байт, которое отражает номер версии протокола BOP, такой как BGP-3 или BGP-4. В течение фазы переговоров с соседями стороны, участвующие в BGP-сеансе, должны согласовать номер версии протокола BOP. Вначале стороны пытаются "договориться" о наивысшей версии, которую они могут поддерживать. На этом этапе стороны могут сбрасывать сеанс BOP и проводить повторные переговоры до тех пор, пока не согласуют, по какой версии BOP будет проводиться сеанс. Для ускорения процесса переговоров компания Cisco Systems ввела специальный параметр, в котором определяется версия протокола. Как правило, номер версии устанавливается статически, когда версии BOP сторон уже известны, хотя большинство реализаций по умолчанию начинают переговоры с BGP-4.
- **Автономная система (My autonomous system)** — поле размером 2 байта, где указывается номер AS спикера BOP.
- **Таймер удержания (Hold timer).** В поле "Таймер удержания", имеющее в длину 2байта, включаются целые числа, указывающие максимальный интервал времени между приемом сообщений KEEPALIVE и UPDATE. По сути таймер удержания представляет собой счетчик, величина которого увеличиваются от 0 дозначения времени удержания. Прием сообщений типа KEEPALIVE илиUPDATE сбрасывает таймер в 0. Если время удержания для заданного соседнего узла превышено, делается вывод о недоступности такого узла.

  Маршрутизатор, поддерживающий работу по BOP, в фазе переговоров со своим соседом подбирает для него время удержания. Выбор времени удержания между соседними маршрутизаторами производится на основе наименьшего времени удержания. Таймер удержания может быть равным 0, но тогда ни он, ни таймер состояния соединения (KEEPALIVE timer) никогда не будут сбрасываться. Другими словами, оба таймера всегда будут иметь значение 0, следовательно, соединение будет считаться активным. Если таймер не установлен в 0, то по умолчанию минимальное значение времени ожидания для таймера удержания 3 секунды.

  Обращаем ваше внимание на то, что переговоры с целью определения номера версии (сводящиеся в действительности к повторному установлению сеанса, пока узлы не согласуют номер версии протокола) и для определения начального значения таймера удержания (использование минимального значения одного из двух спикеров BOP) отличаются коренным образом. В обоих случаях каждому маршрутизатору посылается только сообщение OPEN. Однако при несовпадении значений (в случае таймера удержания) сеанс не прерывается.
- **Идентификатор BGP (BGP Identifier)** представляет собой четырехбайтовое целоечисло, которое отображает значение идентификатора BOP узла отправителя. Вмаршрутизаторах компании Cisco это значение обычно соответствует идентификатору маршрутизатора (Router ID — RID), который вычисляется из наивысшего IP-адреса на маршрутизаторе или из наивысшего адреса обратной петли в

начале сеанса BOP. Адрес обратной петли представляет собой IP-адреспрограммного виртуального интерфейса, который считается всегда активным,независимо от состояния физического интерфейса на маршрутизаторе.

- **Длина поля необязательных параметров (Optional Parameter Length — Opt ParmLen).** Это однобайтовый целочисленный параметр, который отражает полнуюдлину в байтах поля "Необязательные параметры". Если длина равна 0, то необязательные параметры отсутствуют.
- **Необязательные параметры (Optional Parameters).** Это поле переменной длины, вкотором отображается список необязательных параметров, используемых протоколом BOP при ведении переговоров между соседними узлами. В этом полемогут отображаться параметры <Параметр типа, параметр длины, параметр значения> (<Parameter Type, Parameter Length, Parameter Value>) длиной по одному байту и переменной длины, соответственно. Примером необязательных параметров может служить параметр информации об аутентификации (тип 1), который применяется для аутентификации сторон в сеансе BOP.

Модель конечных состояний

Процесс переговоров между BGP-соседями до полной установки соединения происходит в несколько этапов. На рис. 5.8 приведена упрощенная модель конечных состояний (finite state machine — FSM), с помощью которой можно рассматривать основные события, в результате которых генерируются сообщения от одного узла другому и реакцию на них другой стороны.

Ниже приведен список основных состояний модели FSM.

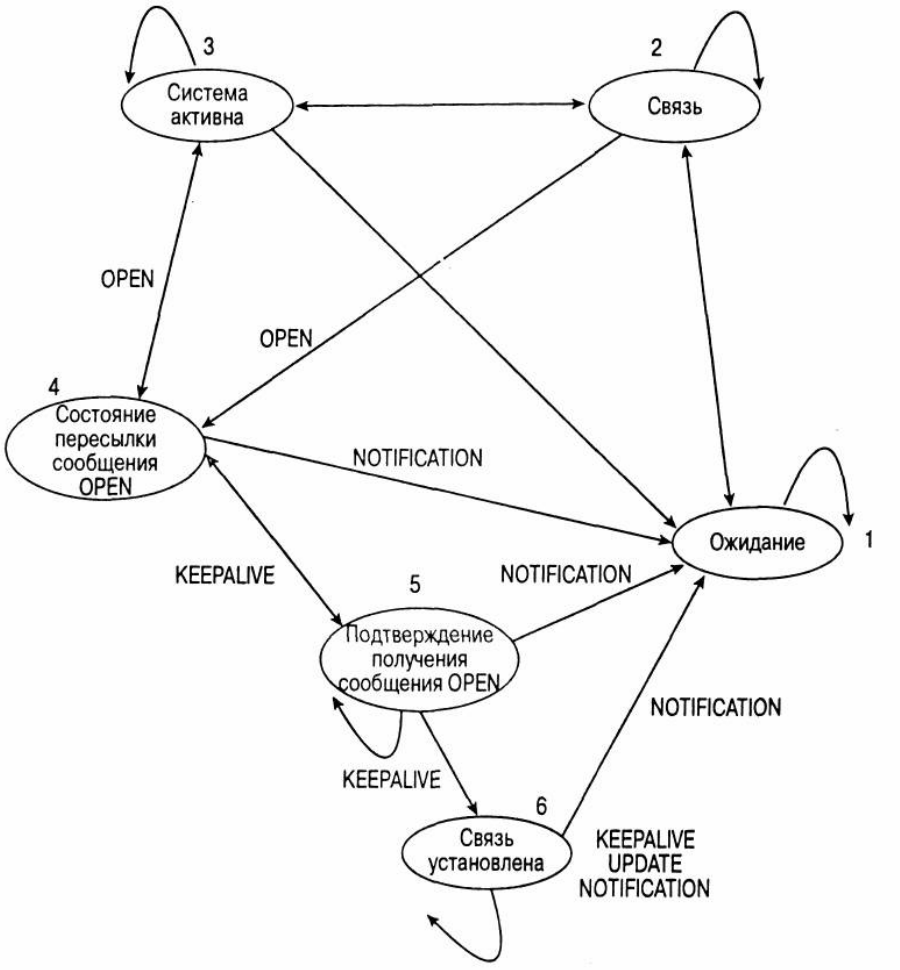


Рис. 5.8. *Модель конечных состояний при переговорах по протоколу BOP между соседними узлами*

1. **Ожидание (Idle).** Это первое состояние, в котором находятся системы перед установлением соединения. В протоколе BOP ожидается наступление события "Пуск" (Start), инициированного оператором или самой BGP-системой. Администратор вызывает наступление события "Пуск" путем установления BGP-сеанса маршрутизатором или посредством сброса существующего BGP-сеанса. После наступления события "Пуск" BGP-система инициализирует свои ресурсы, сбрасывает таймер повторных попыток установки соединения (ConnectRetry timer), устанавливает транспортное соединение по протоколу TCP и находится в режиме ожидания соединения с удаленной стороной. Затем BGP-система переходит в состояние ведения связи. В случае появления каких-либо ошибок BGP-система возвращается в состояние ожидания.
2. **Связь (Connect).** В этом состоянии BGP-система ожидает полного установлениясоединения транспортным протоколом. Если TCP-соединение установлено успешно, то система переходит в состояние пересылки сообщения OPEN(OpenSent) (т.е. на этом этапе удаленной системе посылается сообщение об открытии соединения OPEN). Если истекает время, заданное таймером повторныхпопыток ConnectRetry timer, то система остается в состоянии "Связь", таймерсбрасывается, и повторно начинается установка соединения транспортным протоколом. При наступлении каких-либо других событий, инициированных оператором или самой системой, BGP-система возвращается в состояние ожидания.
3. **Система активна (Active).** На этом этапе BGP-система пытается достичь удаленной системы путем открытия соединения транспортного протокола. Если установлено транспортное соединение, то система переходит в состояние пересылки сообщения OPEN (OpenSent), при котором генерируется и посылается сообщение OPEN. Если истекает интервал времени, заданный таймеромповторных попыток, то система перезапускает таймер и возвращается в состояние "Связь". При этом BGP-система продолжает ожидать появления входящегосоединения от удаленного узла. При наступлении еще каких-либо событий система может вернуться и в состояние ожидания. Таким событием может бытьсобытие "Останов" (Stop), инициированное самой системой или оператором.

   Если система находится в состоянии, колеблющимся между состояниями "Связь" и "Система активна", — это признак того, что при установке транспортного TCP-соединения что-то происходит неправильно. Причинами такого состояния может быть большое количество повторных передач пакетов по протоколу TCP или невозможность соседнего узла распознать IP-адрес удаленной стороны.
4. **Состояние пересылки сообщения OPEN (OpenSent).** В этом состоянии BGP-система ожидает получения сообщения OPEN от удаленной стороны. Полученноесообщение проверяется на целостность. Если в нем содержатся ошибки, такиекак искаженный номер версии протокола или недопустимый номер AS, системаотправляет удаленной стороне сообщение об ошибке NOTIFICATION и возвращается в состояние ожидания. Если ошибок не обнаружено, BGP-система начинает посылать сообщения KEEPALIVE и сбрасывает свой таймер проверки состояния канала (KEEPALIVE timer) в 0. С этого момента оговаривается такжевремя удержания и устанавливается наименьшее его значение из связанных систем. Если согласованное время удержания равно 0, то таймер удержания (Holdtimer) и таймер проверки состояния (KEEPALIVE timer) не перезапускаются.

   В состоянии пересылки сообщения OPEN BGP-система путем сравнения собственного номера AS с номером AS удаленной системы выясняет, принадлежит ли маршрутизатор, с которым установлена связь, к той же автономной системе (внутренний Internal BGP) или это различные AS (внешний External BGP).

   При разрыве TCP-соединения система возвращается в состояние "Система активна". При возникновении других событий, таких как истечение времени, заданного таймером удержания, BGP-система посылает сообщение NOTIFICATION, в котором содержится код ошибки, и возвращается в состояние ожидания. Кроме того, в ответ

на событие "Останов", инициированное системой или оператором, BGP-система также переходит в состояние ожидания.

5. **Подтверждение получения сообщения OPEN (OpenConfirm).** В этом состоянииBGP-система ожидает поступления сообщения KEEPALIVE. Приняв такое сообщение, система переходит в следующее состояние "Связь установлена"(Established) и переговоры с соседним узлом завершаются. Приняв сообщениеKEEPALIVE, система перезапускает свой таймер удержания (при условии, что оговоренное значение времени ожидания не равно 0). Если же система получает сообщение NOTIFICATION, то она возвращается в состояние ожидания. Система периодически посылает другой стороне сообщения KEEPALIVE с частотой, установленной таймером проверки состояния канала. В случае любого разрыва транспортного соединения или в ответ на событие "Останов", инициированное самой системой или оператором, система также возвращается в состояние ожидания. При наступлении какого-либо другого события система посылает сообщение NOTIFICATION, содержащее код ошибки модели конечных состояний FSM, и возвращается в состояние ожидания.
6. **Связь установлена (Established).** Это последнее состояние, в котором находятся соседние узлы при ведении переговоров. В этом состоянии BGP-система начинает обмен пакетами UPDATE со своими соседями. Предположим, что таймер удержания не равен 0. Тогда он будет перезапускаться каждый раз при приеме сообщения UPDATE или KEEPALIVE. Если же система получает сообщение NOTIFICATION (в случае возникновения какой-либо ошибки), то она возвращается в состояние ожидания.

   Сообщения UPDATE также проверяются на наличие ошибок, таких как недостающие атрибуты, дублированные атрибуты и другие. При обнаружении ошибки взаимодействующей стороне высылается сообщение NOTIFICATION, и система переводится в состояние ожидания. В состояние ожидания система возвращается также по истечении времени, заданного таймером удержания, при получении уведомления о разрыве транспортного соединения или при наступлении события "Останов", принятого от другого узла или наступившего в результате какого-либо другого события.

Сообщение NOTIFICATION

После рассмотрения модели конечных состояний вам, должно быть, понятно, сколько существует возможных вариантов развития событий при обнаружении ошибок. В любом случае при обнаружении ошибки другой стороне, участвующей в соединении, посылается уведомление об ошибке — сообщение NOTIFICATION. После этого, узел, пославший сообщение, разрывает соединение. Сетевые администраторы должны уметь анализировать содержимое сообщения NOTIFICATION, чтобы определить причины, вызвавшие ошибку протокола маршрутизации. На рис. 5.9 приведен формат сообщения NOTIFICATION.

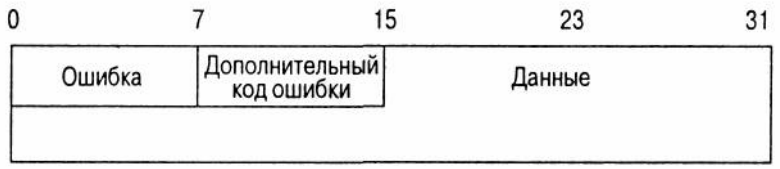


Рис. 5.9. *Формат сообщения NOTIFICATION*

Сообщение NOTIFICATION состоит из кода ошибки (1 байт), дополнительного кода ошибки (1 байт) и поля данных переменной длины.

Код ошибки (Error code) определяет тип уведомления об ошибке, а дополнительный код ошибки (Error subcode) предоставляет более детальную информацию о природе ошибки.

В поле данных (Data field) содержатся сведения об ошибке, такой как неправильный заголовок, запрещенный номер AS и т.д. В табл. 5.1 приведен список возможных ошибок и их дополнительные коды.

Таблица 5.1. *Коды ошибок протокола BGP*

Код ошибки	**Дополнительный код ошибки**
1 — Ошибка в заголовке	1 — Соединение не синхронизировано сообщения
	2 — Неправильная длина сообщения
	3 — Неправильный тип сообщения
2 — Ошибка в сообщении OPEN	1 — Номер версии не поддерживается
	2 — Неправильный номер AS взаимодействующего узла
	3 — Неправильный идентификатор BGP
	4 — Необязательный параметр не поддерживается
	5 — Ошибка аутентификации
	6 — Неприемлемое значение таймера удержания
	7 — Параметр не поддерживается
3 — Ошибка в сообщении UPDATE	1 — Список атрибутов сформирован неправильно
	2 — Общеизвестный атрибут не опознан
	3 — Общеизвестный атрибут не найден
	4 — Ошибка в атрибуте Flags
	5 — Ошибка в атрибуте Length
	6 — Неправильный атрибут Origin
	7 — Петля маршрутизации между AS
	8 — Неправильный атрибут NEXTJHOP
	9 — Ошибка в необязательном атрибуте
	10 — Неправильное поле "Сеть”
	11 — Ошибка в атрибуте AS_PATH
4 — Окончание работы таймера удержания	Нет
5 — Ошибка модели конечных состояний (для ошибок обнаруженных FSM)	Нет
6 — Останов (при других серьезных ошибках, кроме вышеуказанных)	Нет

Сообщение KEEPALIVE

Стороны, участвующие в сеансе связи, периодически обмениваются сообщениями типа KEEPALIVE для того, чтобы определить наличие канала связи и возможность достижения по нему удаленного узла. Как уже отмечалось, время удержания определяет максимальный интервал времени между успешным приемом двух сообщений типа KEEPALIVE или UPDATE. Сообщения типа KEEPALIVE посылаются обычно с частотой, меньшей времени, установленного таймером удержания, на основании чего делается вывод о нормальном течении сеанса. Рекомендуемый интервал времени для посылки сообщений KEEPALIVE — 1/3 от значения таймера удержания. Если же таймер удержания установлен в 0, то обмен сообщениями KEEPALIVE не ведется. Как мы уже говорили ранее, сообщение типа KEEPALIVE представляет собой 19-байтовый заголовок протокола BGP, без каких-либо значений в поле данных. Сообщения этого типа могут подавляться в течение передачи сообщения UPDATE.

Сообщение UPDATE и маршрутная информация

Основу протокола BGP составляет концепция обновлений маршрутной информации.

Обновления маршрутов несут в себе всю необходимую информацию, которая используется в протоколе BGP для построения сети без петель маршрутизации. В сообщение UPDATE, как правило, входят три основных блока:

- информация сетевого уровня о доступности сети (Network Layer ReachabilityInformation — NLRI);
- атрибуты маршрута;
- недостижимые маршруты.

Ка рис. 5.10 показаны все компоненты сообщения UPDATE.

Блок NLRI отображает форму записи IP-префикса маршрута к объявляемой сети. Список атрибутов маршрута позволяет протоколу BGP обнаруживать петли в маршрутизации и придает ему дополнительную гибкость при определении локальных и глобальных правил маршрутизации. В качестве примера атрибутов маршрута можно привести атрибут AS_PATH, с помощью которого определяется последовательность номеров AS, составляющих маршрут до маршрутизатора BGP.

Например, на рис. 5.11 автономная система AS3 получает сообщения UPDATE от AS2, где указывается, что в сеть 10.10.1.0/24 (NLRI) можно попасть через два промежуточных узла — AS2 и AS1. На основе этой информации система AS3 может направлять трафик в сеть 10.10.1.0/24 через транзитный узел AS2 в пункт назначения, подключенный к AS1.

Рис. 5.10. *Формат сообщения UPDATE протокола BGP*

Третья часть сообщения UPDATE представляет собой список маршрутов, которые стали недействительными, или, согласно терминологии, принятой в BGP, удаленными. Из рис. 5.11 видно, что если сеть 10.10.1.0/24 становится недостижимой или в информацию об атрибутах маршрута внесены какие-либо изменения, протокол BGP на всех трех AS может удалить маршрут и разослать сообщение UPDATE со списком новой информации об атрибутах или о невозможности попасть по заявленному ранее маршруту в заданную сеть.

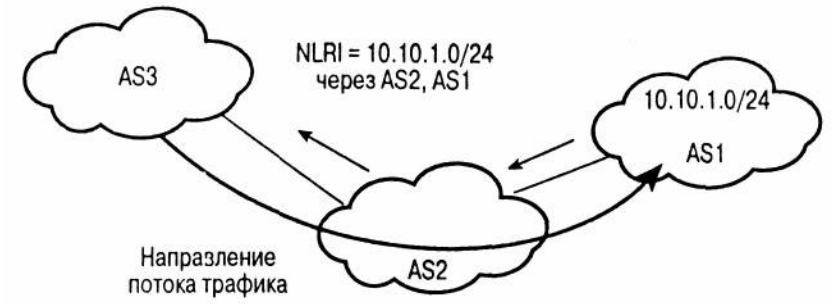


Рис. 5.11. *Пример обновления маршрутной информации в BOP*

Информация сетевого уровня о доступности сети

Одно из основных улучшений, введенных в протокол BGP-4, по сравнению со старыми версиями, — это набор процедур, реализующих механизм бесклассовой междоменной маршрутизации (CIDR). В главе 3, "IP-адресация и методы распределения адресов", уже говорилось о том, что концепция CIDR является отступлением от традиционной классовой модели протокола IP (с разбиением адресов на классы A, B и C) и основана на использовании IP-префиксов и бесклассовой модели.

IP-префикс представляет собой IP-адрес сети с отображением количества битов (слева направо), которые составляют сетевой адрес. Информация сетевого уровня о доступности сети (Network Layer Reachability Information — NLRI) является механизмом, с помощью которого в протоколе BOP осуществляется поддержка бесклассовой маршрутизации. Блок NLRI является частью BGP-сообщения UPDATE. В этом блоке указывается список узлов, о маршрутах к которым BGP-система пытается проинформировать своих соседей. В блок NLRI входят одна или несколько записей формата *<длина, пре-фикс>,* где *длина* — количество маскируемых битов, входящих в заданный префикс.

На рис. 5.12 представлен пример NLRI— <19, 198.24.160.0>. Здесь префикс — 198.24.160.0, а длина маски — 19 бит (считая с левой позиции префикса).

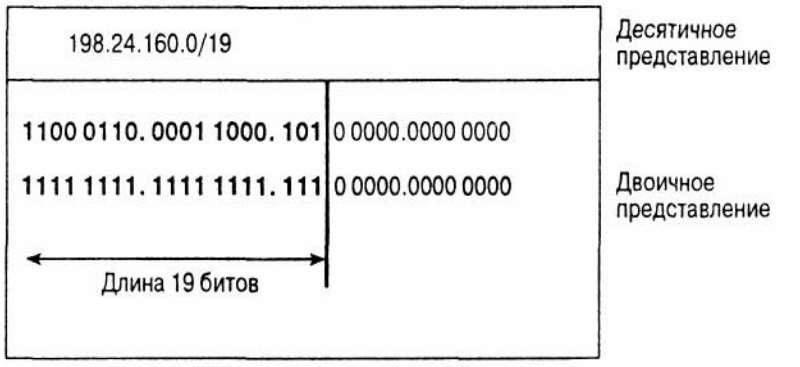


Рис. 5.12. *Пример NLRI*

Удаляемые маршруты

Удаление маршрутов позволяет создать список маршрутов, которые больше не обслуживаются или по каким-либо причинам недоступны в данный момент. Эти маршруты следует изъять (удалить) из маршрутных таблиц BOP. Удаляемые маршруты записываются в том же формате, что и NLRI: IP-адрес и число бит, используемых в нем, считая слева (рис. 5.13).

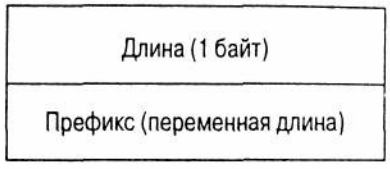


Рис. 5.13. *Общий вид поля удаляемых маршрутов*

Удаляемые маршруты также представляются в формате <длина, префиксX Запись в форме <18, 192.213.128.0> указывает на то, что будет удален маршрут к сети 192.213.128.0 255.255.192.0 или в формате CIDR - 192.213.128.0/18.

Поле длины недостижимых маршрутов (Unfeasible Routes Length) в сообщении UPDATE показывает длину в байтах всех удаляемых маршрутов. В одном сообщении UPDATE может содержаться список из нескольких маршрутов, которые будут удалены, или ни одного удаляемого маршрута.

Так, если поле длины недостижимых маршрутов заполнено нулями, то это означает, что нет ни одного удаляемого маршрута. С другой стороны, в сообщении UPDATE можно

объявлять более одного маршрута, каждый из которых можно описать с помощью нескольких атрибутов маршрута. Сообщения UPDATE, не содержащие информации о NLRI или об атрибутах маршрутов, используются только для уведомления о маршрутах, которые выводятся из эксплуатации (об удаляемых маршрутах).

Атрибуты маршрута

Атрибуты маршрута в BOP — это набор параметров, используемых для характеризации маршрутной информации, такой как информация о пути следования, степень предпочтения маршрута, значение переменной NEXT_HOP для маршрута и информация о возможной агрегации. Все эти параметры используются при фильтрации и выборе маршрутов на базе протокола BOP. Каждое сообщение типа UPDATE включает в себя последовательность атрибутов маршрута переменной длины. Атрибут маршрута имеет три составляющие и записывается в форме <тип атрибута, длина атрибута, значение атрибутах Тип атрибута — это двухбайтовое поле, состоящее из однобайтового флага атрибута и однобайтового кода типа атрибута. На рис. 5.14 представлен общий вид поля типа атрибутов маршрута.

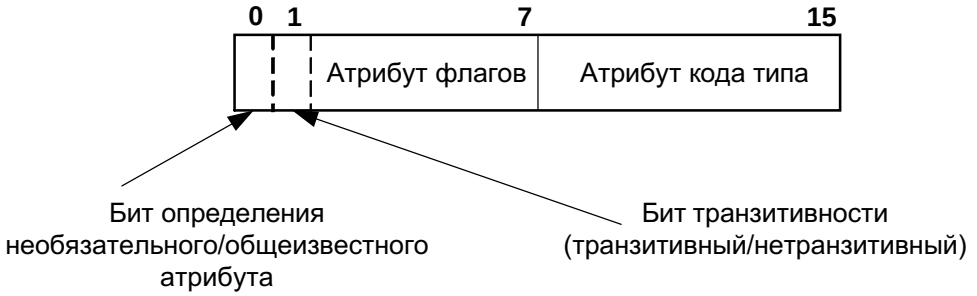


Рис. 5.14. *Формат поля типа атрибутов маршрута*

Атрибуты маршрута подразделяются на четыре категории: обязательные общеизвестные; общеизвестные, предоставленные на собственное усмотрение; необязательные транзитивные и необязательные нетранзитивные. Принадлежность к одной из этих категорий определяется первыми двумя битами в поле флагов атрибута (Atribute Flags).

- Первый бит в поле флагов атрибута (бит 0) указывает на то, является ли атрибут общеизвестным (0) или необязательным (1).
- Второй бит (бит 1) указывает, является ли необязательный атрибут нетранзитивным (0) или транзитивным (1). Общеизвестные атрибуты всегда транзитив-ны, так что второй бит всегда установлен в 1.
- Третий бит (бит 2) показывает, является ли информация, содержащаяся в необязательном транзитивном атрибуте, полной (0) или частичной (1).
- В четвертом бите (бит 3) определяется длина атрибута — 1 байт (0) или 2 байта (1).
- Младшие биты (с 4 по 7) в поле атрибута флага в настоящее время не используются и всегда установлены в 0.

Ниже подробно описана каждая категория атрибутов.

- **Обязательные общеизвестные (Well-known mandatory).** Атрибуты, которые обязательнодолжны присутствовать в пакете UPDATE протокола BOP. Если отсутствует общеизвестный атрибут, то автоматически генерируется сообщение об ошибке NOTIFICATION и сеанс прекращается. Это служит подтверждением того, что во всехреализациях BGP может использоваться стандартный набор атрибутов. В качествепримера общеизвестного обязательного атрибута можно привести атрибут AS_PATH.
- **Общеизвестные, предоставленные на собственное усмотрение (Well-known discretionary).** Эти атрибуты опознаются всеми реализациями BGP, но при этом могут и не посылаться в сообщении UPDATE протокола BGP. Атрибутом этой категории является атрибут LOCAL_PREF.

Кроме общеизвестных атрибутов, маршрут может обладать одним и более необязательными атрибутами. Эти атрибуты необязательно будут поддерживаться всеми реализациями протокола BGP. Необязательные атрибуты могут быть транзитивными и нетранзитивными.

- **Необязательные транзитивные (Optional transitive).** Если необязательный атрибут неопознан протоколом BGP, то обращается внимание на флаг транзитивности. Если он установлен для этого атрибута, что говорит о транзитивности последнего,то используемый BGP принимает атрибут и передает его другим спикерам BGP.
- **Необязательные нетранзитивные (Optional nontransitive).** Когда необязательные атрибуты не опознаны и не установлен флаг транзитивности, что указывает нанетранзитивность атрибута, то такой атрибут просто игнорируется и не передается другим BGP-системам.

Байт кода типа атрибута содержит код атрибута. В настоящее время определены следующие атрибуты (их коды приведены в табл. 5.2).

Таблица 5.2. *Коды типов атрибутов*

Номер атрибута	**Имя атрибута**	**Код категории/типа**	**Соответствующий RFC/проект документа**
1	ORIGIN	Общеизвестный обязательный код типа 1	RFC 1771
2	AS_PATH	Общеизвестный обязательный код типа 2	RFC 1771
3	NEXT_HOP	Общеизвестный обязательный код типа 3	RFC 1771
4	MULTI_EXIT_DISC	Необязательный нетранзитивный, код типа 4	RFC 1771
5	I_OCAL_PREF	Общеизвестный предоставленный на собственное усмотрение, код типа 5	RFC 1771
6	ATOMIC_AGGREGATE	Общеизвестный предоставленный на собственное усмотрение, код типа 6	RFC 1771
7	AGGREGATOR	Необязательный транзитивный, код типа 7	RFC 1771
8	COMMUNITY	Необязательный транзитивный код типа 8	RFC1997[1]
9	ORIGINATORJD	Необязательный нетранзитивный, код типа 9	PFC1966[2]
10	Список кластеров (Cluster List)	Необязательный нетранзитивный, код типа 10	RFC 1966
11	DPA (Destination Point Atribute)	Атрибут точки назначения для BGP	Не использующийся документ
12	Объявитель маршрутов (Advertiser)	Сервер маршрутов BGP/IDRP	RFC1863[3]
13	FaD_PATHO-USTTJRJD	Сервер маршрутов BGP/IDRP	RFC 1863
14	NLRI, допускающий работу в мультипротокольном режиме (Multiprotocol Reachable NLRI)	Необязательный нетранзитивный, код типа 14	RFC 2283[4]
15	NLRI, запрещающий работу в мультипротокольном режиме (Multiprotocol Unreachable NLRI)	Необязательный нетранзитивный, код типа 15	RFC 2283

16	Дополнительные сообщества (Extended Communities)	"в стадии разработки" draft-ramachandra-bgp-ext-communities-00.txt
256		Зарезервировано на будущее

Возможности ведения переговоровв BGP

Поскольку правила ведения переговоров в протоколе BGP (BGP Capabilities Negotiation) еще находятся в стадии разработки (ее осуществляет рабочая группа междоменной маршрутизации IDR (Inter-Domain Routing) Working Group[5], входящая в состав IETF), мы обсудим лишь новые возможности протокола BGP.

Спецификация правил ведения переговоров в BGP-4 существует пока только в черновом варианте, и работа над ней еще не закончена. Основная цель спецификации — ввести в BGP-4 дополнительный параметр, который называется параметром возможностей. Этот параметр помогает реализовать ведение переговоров без разрыва соединения с удаленной стороной.

Если спикер BGP поддерживает возможность ведения переговоров, то при отправке сообщения OPEN взаимодействующей стороне, он включает в него необязательный параметр возможностей (Optional Capabilities parameter). Затем спикер BGP анализирует принятую в сообщении OPEN информацию и в частности параметр необязательных возможностей, чтобы определить, какие возможности поддерживаются удаленным узлом. Если спикер определил, что взаимодействующий узел поддерживает заданные возможности, то он может использовать их при работе с другой стороной.

Если спикер BOP определяет, что другая сторона не поддерживает расширенные возможности ведения переговоров, то в ответ на сообщение OPEN с параметром необязательных возможностей он получает уведомление об ошибке — сообщение NOTIFICATION, в котором содержится дополнительный код ошибки "Необязательный параметр не поддерживается". Тогда спикер BOP должен повторно установить соединение уже без посылки параметра дополнительных возможностей на взаимодействующий узел.

Дополнительные возможности BOP характеризуются параметром типа 24 и содержат одну или несколько записей вида <Код возможности, длина возможности, значение возможности>, где каждое поле имеет вид, как на рис. 5.15.

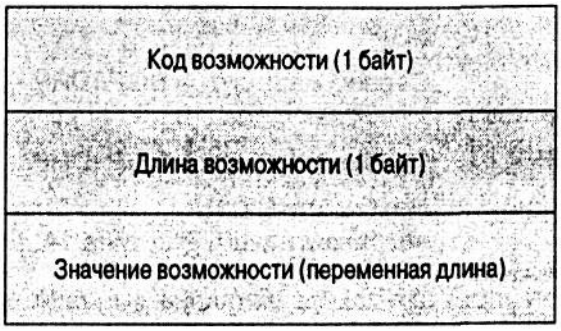


Рис. 5.15. *Формат параметра необязательных возможностей BOP*

Значения полей приведены ниже.

- **Код возможности (Capability Code)** — однобайтовое поле, которое однозначноидентифицирует индивидуальные возможности BOP-системы.

- **Длина возможности (Capability length)** — однобайтовое поле, которое содержитсведения о длине поля значения возможности в байтах.
- **Значение возможности (Capability Value)** — поле переменной длины, интерпретируемое согласно значению поля "Код возможности".

В настоящее время код возможности 0 зарезервирован. Коды 128—255 также зарезервированы для специальных приложений различных производителей.

В операционной системе маршрутизаторов Cisco IOS используется процедура ведения переговоров в BOP для введения двух новых характеристик BOP — регенерации BGP-маршрутов (BOP Route Refresh) и фильтра исходящих маршрутов (Outbound Route Filter — ORF), которые были успешно реализованы. Более подробно об этих характеристиках читайте в последующих главах.

Итак, дополнительные возможности BOP имеют огромный потенциал для внедрения новых характеристик протокола BOP. Однако следует отметить, что многие из этих функций работают лишь с оборудованием от определенного производителя.

Мультипротокольные расширения для BGP

Протокол BOP с мультипротокольными расширениями (Multiprotocol BOP — MBGP) иногда также называют BGP-4+ и ошибочно причисляют его к групповому протоколу BOP (Multicast BOP). На самом деле протокол BOP с мультипротокольными расширениями описан в RFC 2283, и его работа организуется на основе возможностей по ведению переговоров, принятых в протоколе BGP. Протокол MBGP обеспечивает обратную совместимость с протоколом BGP-4, что позволяет ему переносить в себе информацию для протоколов сетевого уровня (кроме IPv4), таких как IPv6 и IPX. Хотя мы не можем здесь рассмотреть все аспекты работы протокола MBGP (у нас просто нет на это времени), мы все же остановимся на типах новых атрибутов, а также сферах применения этого протокола.

В порядке поддержки мультипротокольных расширений в протокол BGP-4 было введено два дополнительных атрибута: NLRI, допускающий работу в мультипротокольном режиме (Multiprotocol Reachable NLRI — MP_REACH_NLRI), и NLRI, запрещающий работу в мультипротокольном режиме (Multiprotocol Unreachable NLRI — MP_UNRICH_NLRI).

NLRI, допускающий работу в мультипротокольном режиме (MP_REACH_NLRI), является необязательным нетранзитивным атрибутом, который может использоваться для следующих целей.

- Объявление действующих маршрутов соседнему узлу.
- Разрешение маршрутизатору объявлять адреса сетевого уровня следующемуближайшему маршрутизатору, который будет использоваться в качестве промежуточного узла, чтобы достичь пункта назначения, указанного в поле NLRI атрибута MP_ NLRI.
- Разрешение заданному маршрутизатору выдавать отчет о некоторых или обовсех точках подключения подсетей (Subnetwork Points of Attachment — SNPA),которые имеются на локальной системе.

NLRI, запрещающий работу в мультипротокольном режиме (Multiprotocol Unreachable NLRI — MP_UNRICH_NLRI), также представляет собой необязательный нетранзитивный атрибут, который может использоваться для удаления одного или нескольких недействующих маршрутов.

Эти новые атрибуты были введены в действие протоколом MBGP для обеспечения возможности связывания определенного протокола сетевого уровня с информацией о соседнем узле и с NLRI. Адресная информация, описанная в RFC 1700[6], используется для

идентификации протоколов сетевого уровня.

Междоменная групповая маршрутизация — наиболее общий случай применения мультипротокольных расширений протокола BGP. Вероятно, в этом и кроется причина того, что некоторые под MBGP подразумевают Multicast (т.е. групповой) BGP, а не Multiprotocol (мультипротокольный) BGP. При использовании MBGP для работы с группами адресов протокол BGP одновременно передает информацию о двух наборах маршрутов — об уникальных маршрутах и о маршрутах, используемых в групповой маршрутизации. Маршруты для групповой маршрутизации используются затем групповым независимым протоколом (Protocol-Independent Multicast — PIM) для процедур пересылки обратных маршрутов (Reverse Path Forwarding — RPF), с помощью которых осуществлялось построение деревьев распределения данных.

До появления MBGP в групповой междоменной маршрутизации использовались обычные системы с однозначными маршрутами, что требовало взаимного соответствия топологий сети для одно- и многоадресной маршрутизации. С появлением протокола MBGP работа с многоадресной междоменной маршрутизацией стала более гибкой, и у сетевого администратора появились новые возможности для управления сетевыми ресурсами.

Более подробно с элементами групповой (или, как ее еще называют, многоадресной) маршрутизации, протоколами для обеспечения ее работы и их функционированием вы можете ознакомиться в книге Бью Вильямсона (Beau Williamson) *"Разработка групповых IP-сетей " (Developing IP Multicast Networks[7]).*

Зашифрованная подпись TCP MD5

Параметр зашифрованной алгоритмом MD5 (Message Digest Algorithm) TCP-подписи описан в RFC 2385[8] и применяется исключительно для зашиты протокола BGP от ложных TCP-сегментов, в частности от несанкционированного сброса TCP-соединений. Зашифрованная TCP-подпись использует алгоритм MD5, описанный в RFC 1321[9]. Более подробные сведения об эффективности применения параметра зашифрованной подписи TCP MD5 вы найдете в спецификации на него.

Итак, данное расширение протокола BOP обеспечивает механизм переноса составного сообщения посредством протокола TCP в каждый TCP-сегмент, где составное сообщение (дайджест) заключает в себе информацию, известную только конечной точке маршрута, и выполняет для сегмента функции цифровой подписи.

Применение алгоритма MD5 к заданным элементам в порядке их приведения позволяет создать дайджест для данного сегмента.

1. Псевдозаголовок TCP в следующем порядке: IP-адрес отправителя, IP-адресполучателя, сокращенный номер протокола и длина сегмента.
2. Заголовок TCP, исключая параметры и полагая, что его контрольная сумма равна 0.
3. Данные TCP-сегмента.
4. Заданный ключ или пароль, известный отправителю и получателю TCP-сегмента.

Сторона, принявшая по протоколу TCP-сегмент с подписью, должна подтвердить действительность подписи своим локальным ключом. Это делается путем вычисления собственного дайджеста и сравнения его значения с принятым. Если в результате сравнения были получены разные величины, то принятый сегмент уничтожается, и отправитель об этом не уведомляется.

Если принимающая сторона настроена для работы с зашифрованной подписью, то отсутствие подписи в получаемом пакете не приводит к запрещению работы этого параметра.

Параметр MD5 формируется для каждого сегмента и всегда имеет длину 16 байт. (Помните, 16 байт в заголовке сообщения BOP, которые были зарезервированы именно для этой цели?) Формат записи этого параметра представлен на рис. 5.16.

В результате применения параметра MD5 все потенциально враждебные попытки сброса TCP-соединений будут игнорироваться принимающим узлом, если отправители не знают значения ключа. Обратите внимание, что обмен ключами во время сеанса по каналу связи не ведется, что исключает возможность их перехвата; они должны быть известны только конечным точкам.

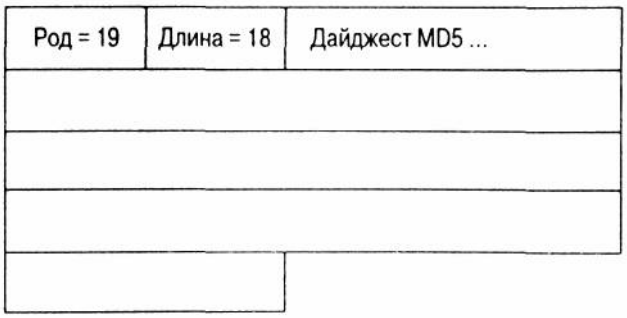


Рис. 5.16. *Формат параметра MD5*

С использованием зашифрованной цифровой подписи связано несколько проблем. Было обнаружено, что алгоритм MD5 уязвим для атак, основанных на коллизиях поиска, поэтому он был признан непригодным для подобных реализаций. К счастью, текущая спецификация не запрещает применение альтернативного алгоритма с функцией хеширования.

При вычислении значения дайджеста наблюдаются потери производительности. Это связано с тем, что в каждом TCP-сегменте для генерации ключа требуется, чтобы и отправитель, и получатель выполнили функцию хеширования, после чего получатель должен сравнить полученные значения и лишь затем принять сообщение. В результате этих операций возникает заметная задержка при обработке и генерации сообщений протокола BGP.

Несмотря на эти трудности цифровая подпись широко используется и в междоменной маршрутизации и при организации маршрутизации внутри доменов.

Забегая вперед

Протокол BGP предоставляет основные элементы маршрутизации, которые обеспечивают администратору достаточную гибкость при управлении. Вся сила BGP заключается в атрибутах и технологиях фильтрации маршрутов. Атрибуты представляют собой параметры, которые можно изменять в процессе выбора маршрутов в BGP. Фильтрация маршрутов может выполняться как на уровне префиксов, так и над самими маршрутами. С помощью комбинирования фильтрации и манипулирования атрибутами можно добиться оптимальной работы системы маршрутизации. Ввиду того что трафик следует по карте маршрутов, которая строится на основе обновлений маршрутов, любое изменение правил маршрутизации неизбежно повлечет за собой изменение траекторий трафика. В следующей главе мы рассмотрим основные вопросы, связанные с установкой правил маршрутизации в протоколе BGP.

Часто задаваемые вопросы

В — *Посылаются ли* в *протоколе BGP периодические обновления маршрутной информации, как это делает RIP?*

О – Нет. В протоколе BGP обмен маршрутной информацией проводится только один раз в начале сеанса связи. После этого узлы обмениваются лишь информацией об изменении маршрутов.

В — *Переходит ли сеанс BGP в состояние «Связь установлена» после обмена маршрутной информацией между соседними BGP-системами?*

О — Нет. Все происходит иначе. Обмен маршрутной информацией не играет никакой роли пока обе стороны, участвующие в сеансе BGP, не согласовали все параметры соединения.

В – Влияет ли информация сетевого уровня о доступности сети (NLRI) на формирование обновлений маршрутов в BGP?

О – Нет. NLRI является просто одним из элементов, который содержится в сообщении UPDATE. Другие элементы этого сообщения – атрибуты и недоступные сети.

В — *Вы говорите об аутентификации как о необязательном параметре в BGP. Насколько важно использование аутентификации?*

О – Аутентификация дает возможность убедиться, что взаимодействующая сторона является именно тем за кого себя выдает. Это позволяет предотвратить получение неправильной маршрутной инсрормации, если хакеры представят себя в качестве одной из соседних систем и будут снабжать вас неверными сведениями о маршрутах. При аутентификации обе стороны принимают решение о надежности другой системы на основе паролей.

В — Где в BGP переносится информация о номерах AS?

О — Номера AS хранятся в атрибуте AS_PATH, который включается в сообщение UPDATE.

В – *Является ли соединение по BGP симметричным или в нем используются отношения типа ведущий-ведомый?*

О – В протоколе BGP не предусмотрено использование ролей ведущего и ведомого. На траспортном уровне соединение всегда инициируется одной из сторон, которая считается клиентом (с номером порта выше 2048). Клиент, в свою очередь, подключается к серверу (порт 179), но эти отношения никак не сказываются на уровне протокола.

В – В используемом соединении с провайдером имеется брандмауэр. Что необходимо сделать для обеспечения нормальной работы BGP?

О – Вам нужно лишь сконфигурировать брэндмауэр таким образом, чтобы он разрешал установку TCP-соединений с портом 179 хотя бы в одном направлении (от провайдера к вам или от вас к провайдеру). Будьте внимательны, т.к. некоторые провайдеры используют BGP в пассивном режиме (т.е. их маршрутизаторы не пытаются самостоятельно устанавливать соединение по BGP.

Ссылки

1 RFC 1997, "BGP Communities Attribute," www.isi.edu/in-notes/rfcl997.txt

2 RFC 1966 "BGP Route Reflection: An alternative to full mesh IBGP," www.isi.edu/in-notes/rfc 1966.txt

3 RFC 1863, "XBGP/IDRP Route Server alternative to a full mesh routing, "www.isi.edu/in-notes/rfcl863.txt

4 RFC 2283, "Multiprotocol Extensions for BGP-4," www.isi.edu/in-notes/rfc2283.txt

5 IETF Inter-Domain Routing Working Group, www.ietf.org/html.chartrers/idr-charter.html

6 RFC 1700, "Assigned Numbers,"www.isi.edu/in-notes/rfc 1700.txt

7 Williamson, Beau. Developing IP Multicast Networks (Indianapolis, Ind.: Cisco Press, 1999)

8 RFC 2385, "Protection of BGP Sessions via the TCP MD5 Signature Option, "www.isi.edu/in-notes/rfc2385.txt

9 RFC 1321, "The MD5 Message-Digest Algorithm, "www.isi.edu/in-notes/rfcl321.txt

Часть III. Зффективные схемы маршрутизации в сетях TCP/IP

В этой части...

Итак, вы уже готовы приступить к использованию атрибутов и функциональности протокола BGP для решеняя практических проблем маршрутизации. В главе 6 мы рассмотрим технологии манипулирования атрибутами BGP и применение фильтрации маршрутов. Определим влияние фильтрации на процесс принятия решений в протоколе BGP. Глава 7 посвящена трем фундаментальным критериям создания сетей на базе TCP/IP - избыточности, симметрии и распределению нагрузки, которыми должны руководствоваться разработчики сетей и отражать их в правилах маршрутизации. В главе 8 рассмотрена интеграция BGP с внутренними протоколами, а в главе 9 раскрывается, как использовать потенциал BGP для управления крупными быстрорастущими сетями. В главе 10 поднимается проблема стабильности сети в связи с продолжающимся ростом Internet. Протокол BGP обладает специально разработанными функциями для поддержания стабильной работы сети. В целом в части III используется подход разъяснения на примерах с использованием специфических топологий и сценариев, что способствует восприятию материала и позволяет наглядно проиллюстрировать те или иные концепции маршрутизации и их реализацию.

Ключевые темы этой главы:

- **Структура** сеанса **связи между взаимодействующми маршрутизаторамй.** Дается полное описание процесса переговоров между BGP-системой и соседними BGP-узлами.
- **Источники обновления маршрутов.** Рассматриваются источники и методы включения маршрутной информации в протокол BGP, а также, их влияние на точность и правильность маршрутной информации.
- **Наложение протоколов: "черные ходы".** Когда накладывающимися протоколами предлагаются альтернативные маршруты в/из сети, то вступает в действие механизм ранжирования их по предпочтительности.
- **Упрощение маршрутизации.** Рассматривается модель принятия решений о пересылке трафика, как непрерывный процесс, согласно которой протокол BGP принимает, фильтрует, выбирает и объявляет маршруты.
- **Управление; маршрутами в BGP.** В ядре протокола BGP реализован набор атрибутов, которые могут быть использованы администраторами для управления .маршрутизацией.
- **Фильтрация маршрутов и управление атрибутами.** На примерах представлен систематизированный обзор процедур, выполняемых BGP, при разрешении или запрещении использования маршрутов, применении фильтров и управлении атрибутами. Таким образом, определяется набор обновлений маршрутов, которые поступают в автономную систему и затем покидают ее.
- **Агрегация в BGP-4.** Приводятся варианты агрегации (объединения) маршрутов и механизмы их реализации с помощью BGP-4.

Глава 6. Настройка параметров BGP

До сих пор мы рассматривали в основном общие положения и определения протоколов внутреннего и внешнего шлюза, а также выполняемые ими задачи. Обсуждались перспективы развития функциональных элементов протокола граничного шлюза Border Gateway-Protocol **(BGP)** В этой главе мы будем подробно рассматривать практические задачи, решаемые протоколом BGP. По сути все они являются частью проблемы надежного взаимрдействия сетей в глобальной сети Internet. В этой главе описаны специфические атрибуты протокола BGP и способ их применения (отдельно или в группе) с целью создания надежных сетей. Хотя терминология, атрибуты и другие детали в этой главе относятся к протоколу BGP, поднимаемые проблемы и рассматриваемые концепции являются общими для разработки структуры маршрутизации, независимо от применяемого протокола маршрутизации.

Структура сеанса связи между взаимодействующими маршрутизаторами

В предыдущей главе мы рассмотрели процесс переговоров между соседними BGP-системами с чисто технической точки зрения. В ней внимание акцентировалось на форматах сообщений, обмен которыми ведется в процессе переговоров. В этой главе мы рассмотрим этот процесс подробно. Кроме того, будут описаны характерные признаки и отличия внутреннего и внешнего протоколов BGP, знание которых необходимо при организации сеанса связи между двумя взаимодействующими маршрутизаторами.

Хотя наиболее широко протокол BGP применяется при построении топологии без петель маршрутизации, он также используется и внутри автономных систем (autonomous system — AS) для обеспечения внутренних маршрутизаторов информацией о доступности различных внешних узлов. *Соединение с соседним узлом (neighbor connection)* (или как его еще называют, *соединение с взаимодействующим узлом (peer connection))* может быть установлено между двумя маршрутизаторами одной и той же AS. В этом случае, протокол BGP называется внутренним BGP (Internal BGP -IBGP). Точно так же соединение между взаимодействующими маршрутизаторами, принадлежащими различным AS, организуется посредством внешнего протокола BGP (External BGP —EBGP). Эти положения представлены на рис.6.1.

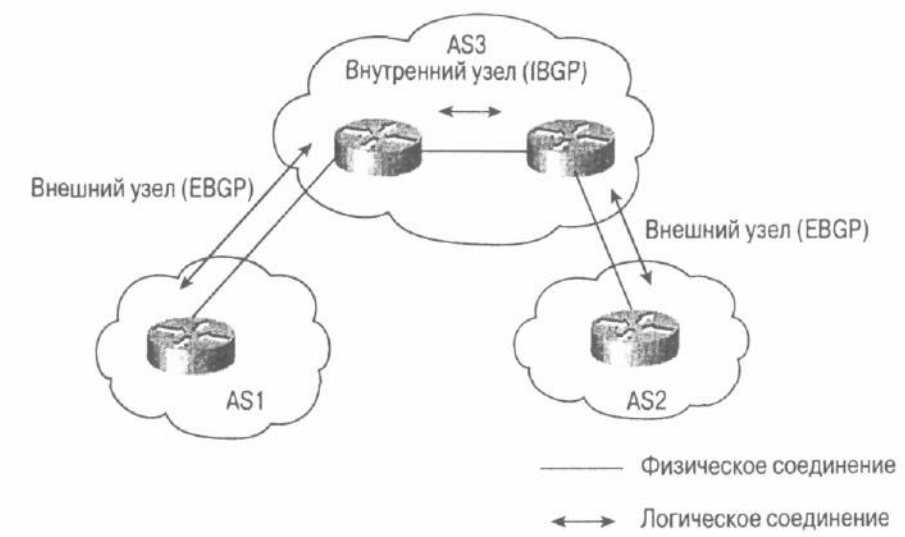


Рис.6.1. *Внутренняя и внешняя реализации протокола BGP*

Во время установления сеанса между соседними BGP-системами и в течение фазы обмена сообщением OPEN маршрутизаторы сравнивают номера AS и определяют, принадлежит ли другая сторона к той же самой AS или другой AS. Разница между EBGP и IBGP проявляется в способе обработки удаленной стороной обновлений маршрутной информации, поступающих от соседней BGP-системы, и в способе передачи различных атрибутов BGP, которые для внутренних и внешних соединений различны.

Процесс переговоров с соседней BGP-системой на этапе установки транспортного соединения на базе протокола TCP/IP для внутренних и для внешних соединений практически ничем не отличается. Обязательное условие для установления транспортного соединения — наличие у взаимодействующих BGP-систем IP-адресов, т.е. соединения по протоколу IP. Такое соединение должно быть установлено по протоколу, отличному от BGP, в противном случае сеанс связи будет развиваться по приведенному ниже сценарию.

Соседние маршрутизаторы могут связываться друг с другом посредством одного из протоколов внутреннего шлюза (Interior Gateway Protocol -- IGP), сеанс BGP установлен и ведется обмен сообщениями BGP. Затем по какой-то причине IGP-соединение разрывается, но при этом продолжается работа в сеансе TCP, так как соседние маршрутизаторы все могут связаться друг с другом по протоколу BGP. В итоге сеанс все равно будет прерван, так как он не зависит от соединения между двумя маршрутизаторами — для этого имеется специальный параметр, определяющий доступен ли ближайший соседний хост (NEXT_HOP). В другом случае маршрут является более специфичным (однозначно определенным), чем используемый в данном сеансе связи, и сведения о нем доводятся посредством BGP.

Чаще всего для обеспечения работы по IBGP используется протокол внутреннего шлюза (Interior Gateway Protocol — IGP) или статический маршрут. В сущности, посылаемый программой PING пакет, содержащий IP-адрес отправителя (IP-адрес одной из BGP-систем) и IP-адрес получателя (другой BGP-системы), вполне может подойти для инициирования транспортного сеанса связи. Для внешних BGP-сеансов маршрут через непосредственно подключенный интерфейс позволяет установить IP-соединение.

Физические и логические соединения

Все внешние взаимодействующие BGP-системы должны иметь физическое соединение друг с другом. BGP-система отказывается принимать сообщения UPDATE от внешних соседей, если у нее нет с ними физического соединения. Однако ее можно аставить делать это с помощью определенных технических приемов. В некоторых случаях внешние соседние маршрутизаторы не могут находиться в одном физическом сегменте. Такие соседние системы имеют логическое соединение (т.е. через несколько промежуточных узлов IP), а не физическое. В качестве примера можно привести работу по протоколу BGP между

двумя внешними маршрутизаторами через несколько других маршрутизаторов, на которых не поддерживается BGP. Для подобных случаев компания Cisco (и некоторые другие производители оборудования) предлагает специальный прием, помогающий обойти это ограничение. При этом в конфигурации BGP-системы требуется задать дополнительные параметры, которые указывали бы на то, что внешняя взаимодействующая сторона не подключена к ней физически.

Примечание

Для внешних взаимодействующих систем, которые не подключены напрямую друг к другу, требуется дополнительная конфигурация.

BGP-сеанс между двумя внешними BGP-системами, не имеющими физического соединения между собой, называется *мультиузловым EBGP-сеансом (multihop EBGP).* Как показано на рис. 6.2, на маршрутизаторе RT2 не поддерживается протокол BGP, а на RT1 и RT3 поддерживается. Таким образом, внешние маршрутизаторы RT1 и RT3 являются логически соединенными и взаимодействуют друг с другом посредством мультиузлового EBGP. (Однако следует помнить, что маршрутизатор RT2 должен каким-то образом получить соответствующую маршрутную информацию, чтобы избежать образования петель или "черных дыр" в маршрутизации).

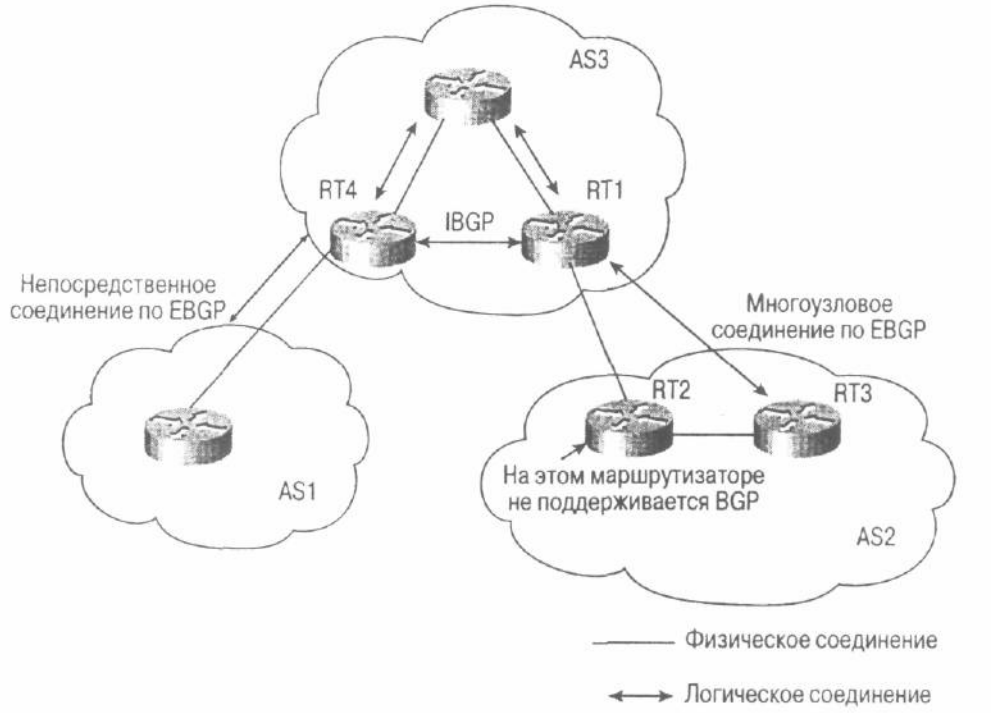


Рис. 6.2. *Внешний BGP в многоузловой среде*

С другой стороны, на соседние системы, принадлежащие к одной и той же автономной системе (внутренние соседи), не налагаются подобные ограничения, т.е. не имеет значения, подключены ли они друг к другу физически или разделены несколькими узлами. В любом случае, если между двумя взаимодействующими узлами имеется IP-соединение, то никаких дополнений в конфигурацию BGP вносить не требуется. Из рис. 6.2 видно, что между маршрутизаторами RT1 и RT4 имеется логическое соединение. Благодаря тому что оба эти маршрутизатора принадлежат одной и той же AS, для работы по протоколу IBGP не требуется менять их конфигурацию.

Назначение IP-адреса

В качестве IP-адреса узла может выступать адрес любого из интерфейсов маршрутизатора: Ethernet, Token Ring или последовательный порт. Помните, что стабильность соединения с удаленной стороной зависит от выбранного вами IP-адреса.

Примечание
Стабильность сеанса связи зависит от выбранных IP-адресов

Если IP-адрес принадлежит неисправному сетевому адаптеру (или порту) Ethernet, который каждые несколько минут выключается, то соединение с взаимодействующим узлом и стабильность всей системы маршрутизации под угрозой. Компания Cisco предоставляет возможность конфигурирования виртуального интерфейса, который носит название *петельного интерфейса (loopback interface)* и всегда находится в активном состоянии. При соединении с удаленной BGP-системой такой петельный интерфейс поможет локализовать неисправность, что при использовании обычного аппаратного интерфейса довольно проблематично.

Добавление петельных интерфейсов не является необходимой процедурой (к тому же это требует дополнительной конфигурации). Если внешние BGP-системы связаны друг с другом напрямую и их IP-адреса, используемые в процессе переговоров, принадлежат к одному сегменту, то адрес петли не добавляется. Если физическое соединение между двумя системами нестабильно, то сеанс может быть прерван с помощью или без помощи петли.

См. в главе 11 раздел "Сеанс связи между взаимодействующими маршрутизаторами"

Аутентификация сеанса BGP

Как уже отмечалось в главе 4, "Основы междоменной маршрутизации", с помощью заголовка сообщения BGP можно проводить аутентификацию удаленной стороны. Аутентификация дает возможность противодействовать хакерам, которые, выдавая себя за одну из взаимодействующих BGP-систем, передают вашей AS некорректную маршрутную информацию. Аутентификация между двумя BGP-системами позволяет удостовериться в полномочиях сторон, организующих сеанс, путем использования секретных ключей на обоих сторонах. Тогда любая система, попытавшаяся установить соединение с вашей системой без надлежащего ключа, будет игнорироваться. В настоящее время в BGP-4 имеется возможность применения алгоритма шифрования с помощью сообщений-дайджестов версии 5 (Message Digest algorithm version 5 — MD5). Детальное рассмотрение алгоритма MD5 не входит в круг вопросов, освещаемых в этой книге, но, как уже отмечалось, он предоставляет дополнительные возможности по обеспечению безопасности транспортного TCP-соединения.

Целостность BGP внутри AS

Кроме случая отражения маршрута, во избежание образования петель маршрутов внутри AS протокол BGP не проводит повторное объявление маршрутов для внутренних узлов, получающих информацию еще и от других узлов IBGP. Таким образом, очень важно поддерживать работоспособность всей совокупности узлов IBGP в пределах AS. Другими словами, каждый маршрутизатор с BGP должен установить сеансы вязи с остальными BGP-маршрутизаторами внутри AS. На рис. 6.3 приводится одна из наиболее распространенных ошибок, допускаемых администраторами при организации маршрутизации на базе протокола BGP внутри **AS.**

В ситуации, показанной на рис. 6.3, провайдер имеет несколько точек присутствия (POP) — в Сан-Хосе, Сан-Франциско и в Лос-Анджелесе. В каждой точке присутствия имеется несколько маршрутизаторов без поддержки протокола BGP и по одному граничному маршрутизатору BGP, соединенному с другими AS посредством протокола EBGP. Администратор настраивает соединение между граничными маршрутизаторами в Сан-Хосе и Сан-Франциско и использует при этом протокол IBGP. Затем он настраивает еще одно соединение на базе IBGP между маршрутизаторами в Сан-Франциско и Лос-Анджелесе. При такой конфигурации EBGP-маршруты, полученные маршрутизатором в Сан-Хосе, будут

передаваться в Сан-Франциско, EBGP-маршруты от маршрутизатора в Сан-Франциско будут передаваться в Сан-Хосе и Лос-Анджелес, а EBGP-маршруты, полученные маршрутизатором в Лос-Анджелесе, попадут в Сан-Франциско. Как видите, маршрутизация, представленная на этом рисунке, не является завершенной — EBGP-маршруты через маршрутизатор в Сан-Хосе не будут переданы в Лос-Анджелес, а EBGP-маршруты через маршрутизатор в Лос-Анджелесе не будут доступны в Сан-Хосе. Причина этого заключается в том, что маршрутизатор в Сан-Франциско не пропускает IBGP-маршруты между Сан-Хосе и Лос-Анджелесом. Все, что необходимо в этом случае, — организовать дополнительное соединение между Сан-Хосе и Лос-Анджелесом (показано пунктирной линией). В главе 9, "Управление крупномасштабными автономными системами", вы увидите, как подобная ситуация может быть исправлена с помощью отражателей маршрута — специального параметра, который обеспечивает лучшее масштабирование в AS с большим числом маршрутизаторов IBGP.

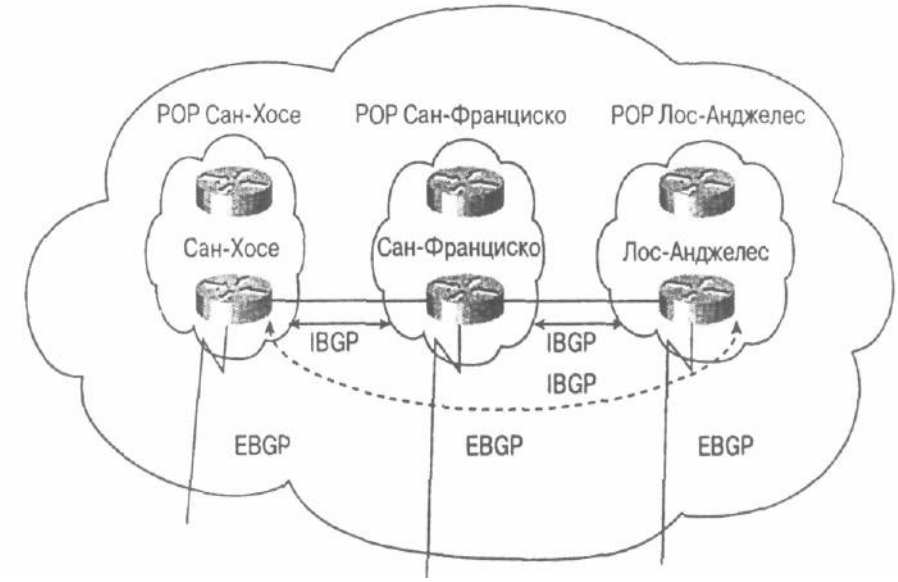


Рис. 6.3. *Нарушение целостности BGP внутри AS*

Синхронизация внутри AS

Согласно определению, для нормальной работы протокола BGP по умолчанию он должен быть синхронизирован с IGP, после чего BGP может уведомлять другие AS о транзитных маршрутах. Очень важно, чтобы все маршруты, о которых объявляет ваша AS, были согласованы, в противном случае в маршрутизации могут появиться "черные дыры", что чревато потерями трафика. Например, если спикер IBGP объявит маршруты внешнему маршрутизатору во время обмена информацией о IGP-маршрутах внутри самой AS, то ваша AS может принять трафик, направленный узлам, сведения о которых еще не получены маршрутизаторами. В результате маршрутизаторы просто не смогут доставить трафик в пункты назначения.

В любом случае, когда маршрутизатор принимает информацию об обновлении маршрута к определенному узлу от взаимодействующей с ним по IBGP стороны, он пытается самостоятельно проверить возможность достижения нужного узла, прежде чем передать эту информацию внешнему маршрутизатору с EBGP. Эту проверку маршрутизатор выполняет путем исследования префикса пункта назначения, чтобы убедиться в существовании маршрута к ближайшему соседнему маршрутизатору и уточнить, существует ли префикс для данного пункта назначения в протоколе JGP. Подобная проверка выявляет, могут ли маршрутизаторы без помощи BGP доставить трафик в заданный пункт назначения. Допуская, что протокол IGP распознает заданный пункт назначения, маршрутизатор анонсирует маршрут к нему другим внешним маршрутизаторам с EBGP. В противном случае маршрутизатор воспринимает префикс пункта назначения как несинхронизированный с

протоколом 1GP и не объявляет об этом маршруте.

Рассмотрим ситуацию, показанную на рис. 6.4. Провайдеры ISP1 и ISP2 используют узел провайдера ISP3 как транзитную AS. В состав AS провайдера ISP3 входит несколько маршрутизаторов, и протокол BGP используется только на граничных маршрутизаторах этого провайдера. (Хотя маршрутизаторы RTB и RTD тоже участвуют в транспортировке транзитного трафика, провайдер ISP3 не сконфигурировал их для работы с протоколом BGP). Для обеспечения работы внутри AS провайдер ISP3 использует протокол внутреннего шлюза Interior Gateway Protocol.

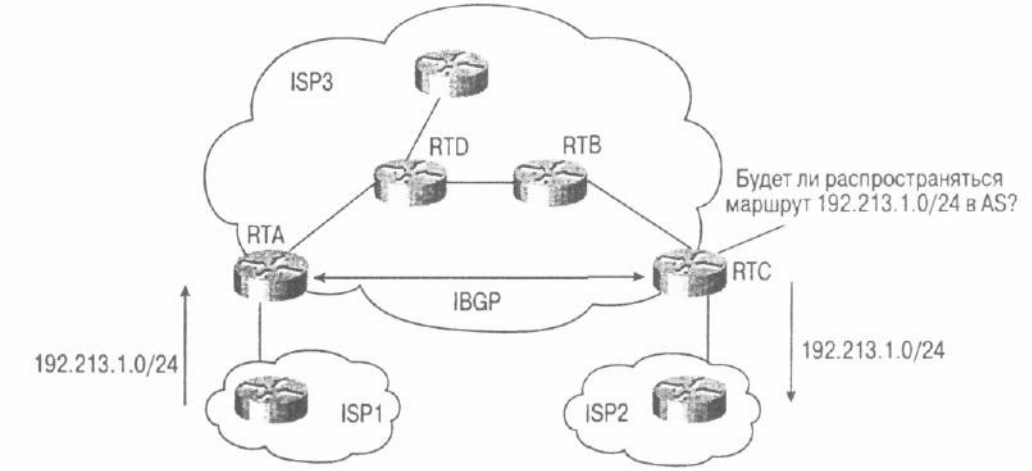


Рис. 6.4. *Синхронизация маршрутов в BGP*

Допустим, что провайдер ISP1 объявляет маршрут 192.213.1.0/24 к провайдеру ISP3. Так как маршрутизаторы RTA и RTC работают по протоколу IBGP, то маршрутизатор RTA уведомляет об этом маршруте маршрутизатор RTC. Обратите внимание на то, что на других маршрутизаторах, кроме RTA и RTC, протокол BGP не поддерживается, и им неизвестно о существовании маршрута 192.213.1.0/24.

В ситуации, представленной на рис. 6.4, если маршрутизатор RTC объявит маршрут к провайдеру ISP2, то трафик, направленный в сеть 192.213.1.0/24, начнет поступать на маршрутизатор RTC. Затем на маршрутизаторе RTC выполняется анализ таблицы маршрутов, и трафик перенаправляется на маршрутизатор RTB. Маршрутизатору RTB не известно о существовании маршрутов BGP, и он игнорирует весь трафик, так как не имеет сведений о пункте назначения, в который направляется трафик. Здесь мы видим наглядный пример потери трафика ввиду отсутствия синхронизации между BGP и IGP.

Основное правило для протокола BGP гласит: "Маршрутизатор, работающий на базе протокола BGP, не должен уведомлять внешние маршрутизаторы о маршрутах к узлам, полученным от своих соседей по протоколу IBGP, если только эти сведения не были подверждены протоколом IGP". Это и есть правило синхронизации. Если маршрутизатор получает сведения о маршрутах в пункты назначения по IGP, то это означает, что сведения о маршрутах уже распространились внутри AS и доступность всех узлов подтверждена.

Однако последствия ввода BGP-маршрутов в IGP не проходят бесследно. Преобразование маршрутов из BGP в IGP приводит к перегрузке внутренних маршрутизаторов -как из-за масштабируемости самого протокола IGP, так и вследствие того, что протоколы типа IGP не очень хорошо справляются с большим количеством маршрутов. Кроме того, нет необходимости в переносе абсолютно всех внешних маршрутов внутри AS. Вся маршрутизация может выполняться путем направления всех маршрутизаторов без BGP по умолчанию на один из маршрутизаторов с поддержкой BGP. Конечно, эта маршрутизация будет неоптимальной, так как не гарантируется, что в каждом маршруте будет использоваться кратчайший путь к заданному пункту назначения, но это минимальная цена, которую лучше заплатить, в противном случае вам придется иметь дело с обслуживанием нескольких тысяч маршрутов внутри AS. Естественно, управление маршрутами, заданными по умолчанию, в ситуации подобной этой может показаться довольно сложной задачей и иногда приводит к образованию петель маршрутизации.

Однако в большинстве реализаций BGP предлагается программное решение,

позволяющее сетевому оператору запретить синхронизацию. Как вы уже догадались, речь идет о вспомогательной команде в реализации BGP для оборудования Cisco — **no syncronization,** с помощью которой запрещается использовать синхронизацию в BGP и объявлять маршруты, полученные посредством IBGP, несмотря на наличие маршрутов протокола IGP. На практике с целью обеспечения полной изоляции граничных маршрутизаторов синхронизацию отключают, допуская, что транзитные маршрутизаторы в AS работают в связке по протоколу IBGP. В этом случае доступность узлов внутри AS гарантирована, так как маршрут, полученный посредством EBGP, будет автоматически передан по BGP всем транзитным маршрутизаторам.

Все это говорит о том, что в общем случае для сетей, подключенных к Internet, запрещается применение синхронизации BGP и делается упор на сеть маршрутизаторов, взаимосвязанных по протоколу IBGP. Сама мысль о подстановке в IGP десятков тысяч маршрутов BGP выглядит пугающе.

Источники обновления маршрутов

В таких сложных сетях, какой является сегодня Internet, стабильность маршрутов – важнейшая характеристика. С учетом флуктуации маршрутов в сети Internet существует тесная взаимосвязь между стабильностью каналов и способом распространения маршрутной информации посредством протокола BGP. Информация о маршрутах может включаться в протокол BGP динамически или статически. Динамически вложенные маршруты поступают прямо в таблицу BGP-маршрутов, в зависимости от статуса сетей, которые они идентифицируют. Маршруты, вложенные статически, постоянно обслуживаются таблицами BGP-маршрутов, независимо от статуса сетей, которые они идентифицируют. Таким образом, при недоступности объявляемой сети динамическое объявление маршрутов прекращается, а статическое продолжается. Каждый метод имеет свои преимущества и недостатки, но об этом позднее.

См. в главе 11 раздел "Источники обновления маршрутов"

Динамическое вложение информации в BGP

Информация, поступающая в BGP динамически, далее может подразделяться на чисто динамическую — при этом все маршруты протокола IGP преобразуются в BGP (с помощью дополнительной команды протокола BGP redistribute) — и полудинамическую, когда в BGP преобразуются только определенные IGP-маршруты (с помошью дополнительной BGP-команды network). Такое различие отражает и уровень вмешательства пользователя, и уровень управления при выборе объявляемых маршрутов.

Информация динамически подставляется в BGP посредством разрешения автоматического преобразования всех маршрутов из IGP в BGP. В настоящее время в автономных системах используется несколько разновидностей протоколов IGP, включая протокол обмена маршрутной информацией Routing Information Protocol (RIP), протокол маршрутизации внутреннего шлюза Interior Gateway Routing Protocol (IGRP), расширенный протокол IGRP Enchanced IGRP (EIGRP), протокол маршрутизации по кратчайшему открытому пути Open Shortest Path First (OSPF) и протокол взаимодействия промежуточных систем Intermediate System-to-Intermediate System (IS-IS). Динамическое преобразование адресов позволяет упростить конфигурацию маршрутизаторов — все внутренние IGP-маршруты будут динамически переноситься в BGP, независимо от используемого протокола.

См. в главе 11 раздел "Динамическое вложение информации в BGP"

Еще один, не совсем динамический метод подстановки информации о маршрутах в протокол BGP заключается в определении подлежащих объявлению подсетей на базе протокола IGP вручную, т.е. самим сетевым администратором. Вложение этих маршрутов в BGP вручную проводится с помощью команды network. Этот метод в принципе не является чисто динамическим, так как список объявляемых префиксов, необходимо задавать на маршрутизаторе вручную. Маршрутизатор в этом случае не преобразует все маршруты из протокола IGP в BGP автоматически. При этом, если список префиксов увеличивается, его обслуживание становится практически невозможным.

В протоколе BGP предполагается, что сети, заданные командой network, существуют, и делается попытка проверить их доступность через таблицы IP-маршрутов. Если протокол BGP не находит точного соответствия для заявленных сетей, то маршруты к ним не объявляются. Если сопутствующая сети маска (дополнительная команда mask) не указана в команде network (например, network 10.10.0.0 mask 255.255.0.0) и разрешено автоматическое суммирование (по умолчанию), то существование любого набора классовых префиксов, указанных в команде network (например, network 10.0.0.0 с существующей сетью 10.10.10.0/24), приведет к объявлению набора классовых префиксов. (В нашем случае будет объявлен маршрут к сети 10.0.0.0/8). Такая интеллектуальная проверка многократно увеличивает устойчивость BGP к появлению ошибок, что не позволяет внешним сетям получать неверную информацию о неподключенной или неизвестной по каким-либо другим причинам сети.

Подстановка маршрутов в BGP с помощью команды network позволяет эффективнее управлять объявлением маршрутов. Подстановка маршрутов из IGP в BGP путем их преобразования может привести к появлению в BGP нежелательной сторонней или некорректной информации, о чем мы будем говорить позже.

Вложение нежелательной или некорректной информации

Полное преобразование IGP в BGP может привести к поступлению в BGP нежелательной информации. К этой информации можно отнести внутренние незарегистрированные IP-адреса (известные также под названием общедоступных адресов, разрешенных для применения в корпоративных IP-сетях — *Прим. ред.),* которые действительны только внутри данной AS. В число этих сведений может входить информация о маршрутах, имеющих длину префикса, не согласующуюся с правилами агрегации, принятыми провайдером, например маршрут на хост с длиной префикса /32. Для того чтобы не допустить формирования условий для вложения нежелательных сведений требуется проводить жесткую фильтрацию. Кроме того, в более новых версиях IOS автосуммирование маршрутов (в классовых пределах) разрешено по умолчанию.

См. в главе 11 раздел "Динамическое вложение информации в BGP"

Некорректная информация может поступать в BGP при взаимном обмене маршрутами между BGP и IGP. Как IGP-маршруты могут быть преобразованы в BGP, так и маршруты протокола BGP могут быть вложены в пределах AS путем обратного преобразования в протокол IGP. Когда преобразование маршрутной информации происходит в двух направлениях, такая ситуация называется *взаимным преобразованием маршрутов (mutual redistribution).* При взаимном преобразовании маршрутов информация о маршрутах, поступающая в AS извне, может посылаться обратно в сеть Internet, при этом она уже будет считаться исходящей из AS. На рис. 6.5 показана опасность взаимного преобразования маршрутов между протоколами.

На рис. 6.5 показано, как AS100, которая принадлежит к сети A, передает информацию о маршрутах по протоколу BGP на AS200. Граничный маршрутизатор RTC помешает информацию протокола BGP в один из протоколов IGP. Маршрутизатор RTB получает информацию о маршрутах в AS 100 посредством протокола IGP. Он

сконфигурирован для выполнения преобразования сведений о маршрутах из IGP в BGP. В большинстве протоколов IGP отсутствуют средства для различения префиксов AS100 от префиксов собственной AS, так как сведения такого рода не передаются протоколами IGP. Следовательно, сеть A будет объявлена по BGP в сети Internet, но уже как относящаяся к AS200. Такая ситуация, естественно, приводит к путанице среди AS, подключенных к Internet, так как в этом случае у сети A имеется два источника вместо одного — AS 100.

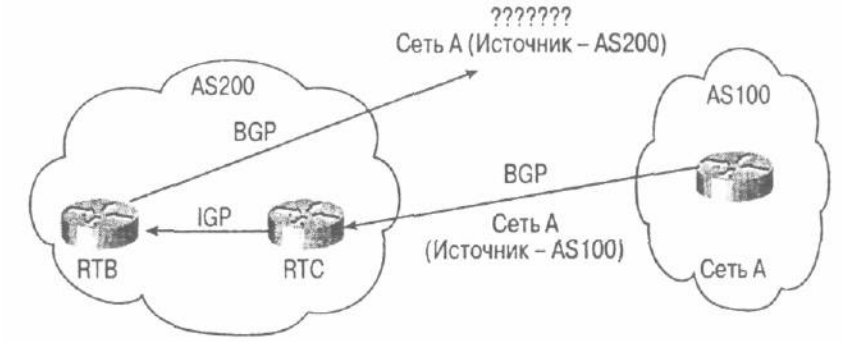


Рис. 6.5. *Пример взаимного преобразования маршрутной информации*

Для того чтобы в подобном случае восстановить нормальную работу на граничных маршрутизаторах, следует прибегнуть к фильтрации, с помощью которой можно задавать сети, включаемые из IGP в BGP. Это предотвратит обратное преобразование маршрута к сети A через BGP на маршрутизаторе RTB. Администратор может конфигурировать протоколы, которые различают внешние и внутренние маршруты, такие как OSPF, таким образом, чтобы преобразование маршрутов в BGP касалось только внутренних маршрутов. (В реализации Cisco внешние маршруты OSPF автоматически блокируются и преобразованию в BGP не подвергаются, однако администратор может изменить этот параметр). Для протоколов, не отличающих внешние маршруты от внутренних, таких как RIP или 1GRP, имеются специальные дополнительные процедуры, с помощью которых эти протоколы могут различать внешние и внутренние маршруты.

Нестабильные маршруты

Процедура динамического или полудинамического вложения IGP-маршрутов в BGP основывается на зависимости маршрутов в BGP от IGP-маршрутов. Хотя это можно не считать негативным, поскольку отражается состояние всех сетей, все же эта зависимость имеет определенные недостатки. Помните о том, что при сегодняшнем уровне глобальной интеграции сетей любые флуктуации маршрутов внутри вашей AS приведут к нестабильной работе в других подключенных к Internet сетях, которые работают друг с другом по протоколу BGP. Маршрут, объявляемый вами в IGP, будет преобразован в BGP. Если такой маршрут по каким-либо причинам становится непригодным для работы, то посредством BGP посылается сообщение WITHDRAWN, согласно которому взаимодействующие узлы должны будут удалить данный маршрут из своих маршрутных BGP-таблиц. Маршрут, который на вашей AS постоянно переходит из рабочего в нерабочее состояние, на других AS непрерывно удаляется и возобновляется в таблицах маршрутов. Пример флуктуации одного маршрута не так впечатляет. Представьте себе флуктуацию сотен маршрутов на сотнях AS. Такая ситуация может очень негативно повлиять на стабильность работы всей сети Internet.

Для снижения влияния флуктуации маршрутов в сети Internet были приняты все возможные меры. В главе 10, "Проектирование стабильных сетей на базе TCP/IP", вы увидите, как с помощью процесса, называемого разгрузкой маршрутов (route dampening), прекращается объявление нестабильных маршрутов, в зависимости от степени их флуктуации. Объявление таких маршрутов может подавляться на несколько минут и даже на несколько часов (до тех пор пока они не стабилизируются).

Управление стабильностью маршрутов — довольно непростая задача, поскольку, как правило, факторы, дестабилизирующие работу, не контролируются администратором и имеют внешнюю природу. Такими факторами могут быть нестабильные каналы связи или сбои в аппаратном обеспечении. Один из путей минимизации нестабильности маршрутов —

их агрегация (объединение). Когда в объединенный маршрут входит два и более маршрутов, то флуктуации в одном из них не сказываются на всем (объединенном) маршруте. Агрегация маршрутов может выполняться как на узле провайдера, так и на узле клиента, в зависимости от уровня информации, на котором работают между собой клиент и провайдер. Если эта процедура выполняется на стороне клиента, то это облегчает провайдеру задачу выявления флуктуации отдельных маршрутов в сети клиента. Если же агрегация выполняется на узле провайдера, то флуктуации маршрутов клиента будут попадать на узел провайдера, но не в сеть Internet. Вопросы агрегации в BGP-4 будут обсуждаться в конце этой главы, после обсуждения нескольких приемов настройки BGP.

Еще один путь управления стабильностью маршрутов заключается в "развязывании" объединенных маршрутов на отдельные маршруты. Этот процесс называется статическим вложением маршрутов в BGP и описан в следующем разделе.

Статическое вложение информации о маршрутах в BGP

На сегодняшний день статическое вложение информации о маршрутах в BGP является наиболее эффективным методом обеспечения стабильности маршрутов. Конечно же, ничто не совершенно в этом мире, и этот метод тоже имеет свои недостатки.

См. в главе 11 раздел "Статическое вложение информации в BGP"

Для статического вложения маршрутной информации в BGP маршруты протокола IGP (или объединенные маршруты), о которых требуется уведомить другие узлы, определяются вручную, т.е. прописываются статически. Таким образом, эти маршруты никогда не будут удалены из маршрутной таблицы IP и, следовательно, всегда будут доступны для объявления. Так как администраторы в большинстве случаев неохотно объявляют маршруты к недоступным или недействующим сетям, то статическое вложение информации о маршрутах зависит от конкретной ситуации.

Например, если объявление маршрута в Internet проводится из одной точки, то объявление маршрута, который на самом деле недействителен, — не большая проблема. Попытки получить соединение с удаленным узлом по данному маршруту для хостов будут неудачными, независимо от того, будет объявлен маршрут или нет.

С другой стороны, если объявление маршрута в Internet проводится статически из нескольких точек, то это может привести к появлению "черной дыры", в которую устремится весь трафик. Если по каким-либо причинам, вызванным проблемами внутри AS, граничный маршрутизатор не сможет опросить сеть, информацию о которой он рассылает, то трафик, направленный в эту сеть, будет уничтожаться, даже если в эту сеть можно было бы попасть через какой-либо другой промежуточный узел.

Объявление статических маршрутов может выполняться с помощью любого способа из представленных в разделе "Динамическое вложение информации о маршрутах в BGP". Объявление маршрутов может быть выполнено путем перераспределения всех статических маршрутов с помощью команды redistribute или набора статических маршрутов с помощью команды network. Последний из методов обеспечивает лучшее управление маршрутами, так как перераспределение может привести к попаданию нежелательной информации о статических маршрутах в BGP, хотя это может случиться и при фильтрации маршрутов.

Атрибут маршрута ORIGIN

В протоколе BGP маршруты к сети, объявленные с помощью команды network или через агрегацию внутренних маршрутов в AS, включают для каждого маршрута IGP(i) атрибут ORIGIN. С другой стороны, при вложении маршрута в BGP путем преобразования (статического или динамического) атрибут маршрута ORIGIN будет неполным (INCOMPLETE), так как преобразованные маршруты могут поступать из любой точки сети,

И наконец, если информация о маршрутах была получена посредством протокола EGP, то всегда определяется атрибут ORIGIN. Обращаем ваше внимание на то, что объединенные маршруты имеют наихудшее значение атрибута ORIGIN из всех составных

маршрутов.

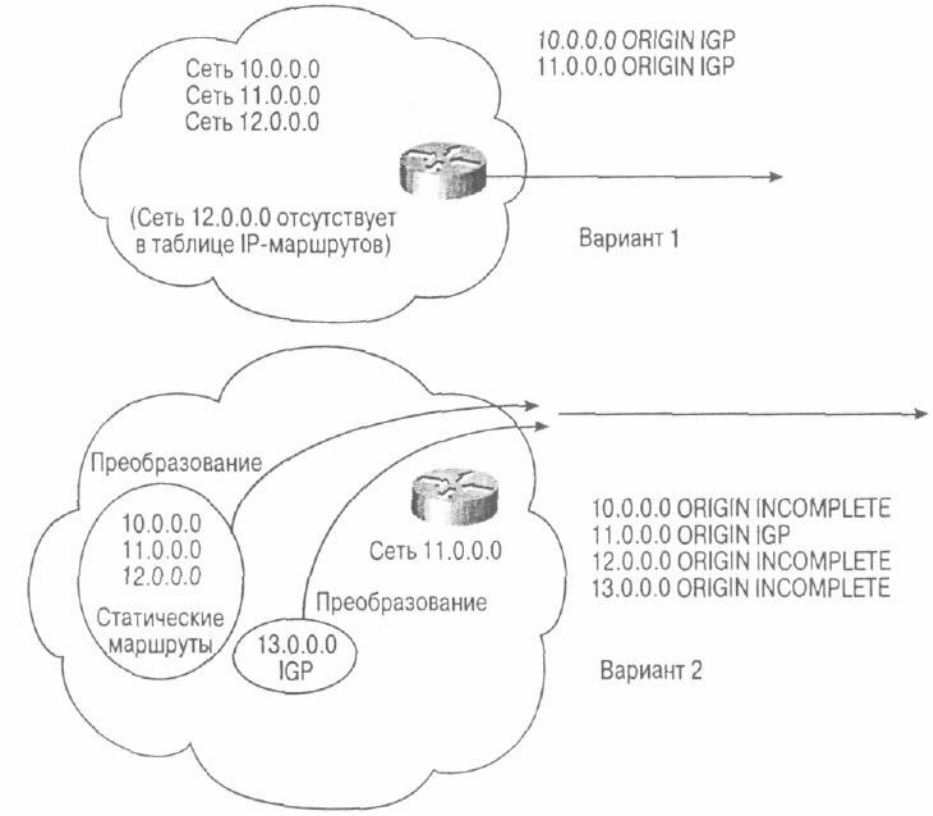


Рис. 6.6. *Сравнение процедур формирования атрибута ORIGIN*

На рис. 6.6 отображен принцип работы атрибута ORIGIN. В варианте 1 сведения обо всех сетях в BGP распространяются с помощью команды network. Отметим, что BGP предполагает для сетей 10.0.0.0 и 11.0.0.0 известные атрибуты IGP. Лишь сеть с адресом 12.0.0.0 не известна маршрутизатору (так как о ней нет записи в таблице маршрутов протокола IP). Как видите, сеть 12.0.0.0 не объявляется посредством BGP, даже если она была задана в команде network.

В варианте 2 маршруты к сетям 10.0.0.0, П.0.0.0 и 12.0.0.0 определены статически. Кроме того, сеть 11.0.0.0 была также определена командой network. Сведения о сети 13.0.0.0 были получены маршрутизатором динамически посредством протокола IGP. Маршруты ко всем этим сетям были путем преобразования вложены в BGP. В результате этого маршруты к сетям 10.0.0.0, 12.0.0.0 и 13.0.0.0 были объявлены с неполным атрибутом ORIGIN INCOMPLETE.

Несмотря на то что маршрут в сеть 11.0.0.0 был вложен в BGP путем явного преобразования статических маршрутов, он также был объявлен с помощью команды network и поэтому послан с атрибутом ORIGIN IGP(i). Если маршрут в сеть 11.0.0.0 не был определен статически командой network, то он получит атрибут ORIGIN INCOMPLETE. Следует заметить, что маршрут в сеть 11.0.0.0 не нужно подвергать преобразованию, так как для его вложения в BGP достаточно задать его статически и указать в команде network.

Хотя атрибут ORIGIN не является важным для работы по протоколу BGP, он используется в нем в процессе выбора маршрута при предпочтении одного маршрута другому.

Статическая маршрутизация против динамической: мобильные сети

В армии принято, чтобы воинские части находились в постоянной готовности и были мобильными, но вместе с этим возникает проблема назначения IP-адресов. Обычно эти мобильные части должны разBGPачивать свои вычислительные сети в любой точке земного шара и работать с IP-адресами так, будто они никуда не перемещались. Если эти сети

являются частью глобальной сети и объявляются с помощью протокола BGP, то задание к ним статических маршрутов не обеспечит нормальной работы. Статические маршруты в этом случае следует удалить из граничных маршрутизаторов в одной AS и каждый раз по прибытии части на новое место устанавливать новые статические маршруты на граничном маршрутизаторе другой AS.

Чтобы избежать подобных сложностей, динамическая маршрутизация сетей в протоколе BGP становится просто необходимой. Одно из решений этой проблемы - обеспечить вложение IGP в BGP на всех узлах глобальной сети. При этом при переносе информации об IP-адресах с одного места на другое, объявление маршрутов из одной точки глобальной сети прекращается и возобновляется в другой точке сети. В некоторых случаях сетевые администраторы не любят прибегать к подобной практике по причинам, указанным нами ранее, таким как взаимное преобразование маршрутов и необходимость в тщательной фильтрации маршрутов.

Еще одна возможность для объявления мобильных сетей заключается в описании их на всех граничных маршрутизаторах во всех предполагаемых точках дислокации с помощью команды network. Так как протокол BGP перед анонсированием маршрутов проверяет их существование в маршрутных таблицах протокола IP, то посредством BGP будут анонсированы только маршруты к мобильным сетям, развернутым в определенных местах. Все остальные точки будут автоматически исключены из анонсирования маршрутов, так как они не являются частью IGP-маршрутов в данной AS.

Наложение протоколов: "черные ходы"

При наличии множества различных протоколов IGP и EGP, совместно выполняющих задачи маршрутизации, распространение маршрутной информации может проводиться с помощью различных протоколов. Выбор того или иного протокола влияет на прохождение трафика. Например, если трафик следует по маршруту согласно протоколу RIP, то он может проходить по одному каналу, если же он следует по внешнему маршруту BGP, то он может передаваться по другому каналу. Запасные соединения, или, как их еще называют, "черные ходы", позволяют использовать альтернативный IGP-маршрут вместо внешнего BGP-маршрута. IGP-маршруты, которые могут быть организованы по запасным каналам, называют *обходными маршрутами (backdoor routes). С* учетом существования подобных альтернативных маршрутов требуется механизм, который бы отдавал предпочтение одному протоколу перед другим. Компанией Cisco Systems введен параметр предпочтения, который называется *административной дистанцией (administrative distance)* протокола. Чем меньше административная дистанция протокола, тем выше степень предпочтения этого протокола.

Нужно отметить, что административная дистанция представляет собой параметр, относящийся только к локальному маршрутизатору. Этот параметр не известен и никак не передается другим маршрутизаторам внутри AS. Таким образом, если вы собираетесь изменить административную дистанцию на одном маршрутизаторе в AS, настоятельно рекомендуется аналогично внести изменения во все маршрутизаторы в AS для того, чтобы обеспечить согласованное принятие решений об использовании того или иного маршрута. В табл. 6.1 приведены все дистанции, реализованные в оборудовании Cisco.

Совет

Более подробно об этой теме см. раздел "Наложение протоколов: "черные ходы" в главе 11.

Таблица 6.1. Значения дистанций протоколов по умолчанию	
Протокол	**Дистанция**
Непосредственное подключение	**0**
Статический маршрут	**1**
EBGP	**20**
EIGRP (внутренний)	**90**
IGRP	**100**
OSPF	**110**
ISIS	**115**
RIP	**120**
EGP	**140**
EIGRP (внешний)	**170**
IBGP	**200**
Локальный BGP	**200**
Неизвестный	**255**

Как видно из табл. 6.1, непосредственное соединение узлов более предпочтительно, чем статический маршрут, который, в свою очередь, более предпочтителен, чем маршрут EBGP, и так далее. Обратите внимание, что маршруты EBGP со значением дистанции 20 предпочтительнее любых других маршрутов IGP.

На рис. 6.7 показан механизм использования обходных маршрутов. На этом рисунке AS1 принимает обновление маршрутной информации о сети A из двух различных источников. Так, AS1 по протоколу EBGP получает маршруты от AS3 и через запасное соединение между AS1 и AS2 по протоколу RIP. Согласно табл. 6.1, маршрутизатор автоматически присвоит дистанцию 20 маршруту EBGP и 120 — маршруту RIP. В AS1 маршрутизаторам, которые получили сведения о маршрутах по протоколу EBGP (граничные маршрутизаторы AS), будет присвоено меньшее значение дистанции в таблице маршрутов. Следовательно, трафик в направлении сети A будет направлен по непрямому маршруту BGP через AS3 и затем AS2, вместо того, чтобы воспользоваться прямым маршрутом RIP через AS2.

Компания Cisco предоставляет способ заставить маршрутизаторы отдавать предпочтение маршрутам IGP перед EBGP. Определенные маршруты EBGP могут быть отмечены как обходные, что влечет за собой изменение значения дистанции для этих маршрутов, и оно снижается до значения локального BGP-маршрута (т.е. по умолчанию 200). Согласно табл. 6.1, эта дистанция намного превышает любой известный маршрут протокола IGP, таким образом, предпочтение будет отдано обходному IGP-маршруту.

Для изменения административной дистанции у всех префиксов BGP, известных маршрутизатору, можно также использовать BGP-команду distance.

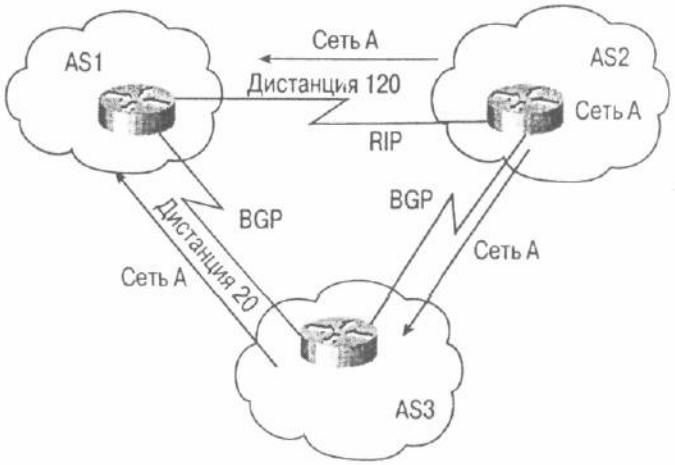


Рис. 6.7. *Конфликты при маршрутизации по обходным маршрутам*

Упрощенная схема процесса маршрутизации

До сих пор в этой главе мы рассматривали отдельные аспекты маршрутизации, в частности ведение переговоров между системами, а также статическую и динамическую маршрутизацию. Прежде чем перейти к рассмотрению деталей в настройке конфигурации маршрутов, имеет смысл остановиться и сделать краткий обзор всего процесса маршрутизации по протоколу BGP.

Протокол BGP сам по себе довольно простой и поэтому достаточно гибкий. С помощью сообщений UPDATE происходит обмен маршрутами между взаимодействующими BGP-системами. BGP-маршрутизаторы принимают сообщения UPDATE, обрабатывают и фильтруют их согласно принятым правилам маршрутизации, а затем передают маршруты другим BGP-узлам. Для того чтобы маршрутные таблицы протоколов BGP и IP хранились отдельно, требуются специальные средства. В случае существования нескольких маршрутов к одному и тому же узлу, BGP-узел не передает сведения обо всех этих маршрутах своим соседям. Наоборот, из имеющихся маршрутов он подбирает наилучший и посылает сведения о нем другим узлам. Кроме передачи всех EBGP-маршрутов от различных узлов или IBGP-маршрутов, поступающих от клиентов-отражателей и маршрутов, BGP-маршрутизатор может самостоятельно генерировать обновления маршрутов для объявления внутренних сетей, принадлежащих к его автономной системе. Действующие локальные маршруты, сгенерированные самой системой, и наилучшие маршруты, полученные от других BGP-узлов, затем помещаются в маршрутную таблицу IP. Таблица IP-маршрутов является конечным решением о маршрутизации и применяется для заполнения таблицы пересылки.

Для моделирования работы протокола BGP представим, что у каждого спикера BGP имеется свой набор маршрутов и собственный набор правил маршрутизации, согласно которому обслуживаются маршруты (хотя в реальной жизни встречается только набор маршрутов). Итак, в модель включены следующие компоненты.

- Набор маршрутов, получаемый маршрутизатором от других узлов.
- Набор входных правил маршрутизации (Input Policy Engine), с помощью кото рых проводится фильтрация маршрутов и изменение их атрибутов.
- Процесс принятия решения, во время которого выбирается маршрут, исполь зуемый самим маршрутизатором.
- Набор маршрутов, которые маршрутизатор формирует и использует сам.
- Набор выходных правил маршрутизации (Output Policy Engine), согласно кото рым проводится фильтрация маршрутов и управление их атрибутами.
- Набор маршрутов, которые маршрутизатор объявляет другим узлам сети.

На рис. 6.8 представлена описанная модель процесса маршрутизации. Дальнейшие разъяснения позволят вам более подробно ознакомиться с функционированием каждого компонента.

Рис. 6.8. *Схема процесса маршрутизации*

BGP-маршруты: объявление и хранение

Как сказано в RFC 1771:

"Маршрутом следует считать единицу информации, включающую сведения о пункте назначения и атрибутах маршрута в этот пункт назначения.

Маршруты объявляются между парами спикеров BGP в сообщениях UPDATE. При этом под пунктом назначения следует понимать систему, IP-адреса которой указаны в поле информации о доступности сети сетевого уровня {Network Layer Reachability Information — NLRI), а под маршрутом следования подразумевается информация, содержащаяся в поле атрибутов маршрута того же сообщения UPDATE.

Маршруты хранятся в информационных базах маршрутизации (Routing Information Bases - RIBs), которые называются: Adj-RIBs-In, Loc-RIB и Adj-RIBs-Out. Маршруты, которые будут объявляться другим спикерам BGP, должны помещаться в базу и Adj-RIBs-Out. Маршруты, используемые спикером BGP локально, должны присутствовать в базе Loc-RIB. Для каждого маршрута в информационной базе пересылки (Forwarding Information Base - FIB) на локальном спикере BGP должны быть сведения о ближайшем следующем узле. Маршруты, получаемые от других спикеров BGP, заносятся в базу Adj~RIBs-In."

Если спикер BGP принимает решение объявить маршрут, то перед объявлением он может добавить или модифицировать его атрибуты.

Отметим, что с этого момента и далее термин *маршрут (route)* в контексте протокола BGP будет представлять собой единицу информации, состоящую из адреса пункта назначения и атрибутов пути следования к этому пункту назначения.

Информационные базы маршрутизации в BGP

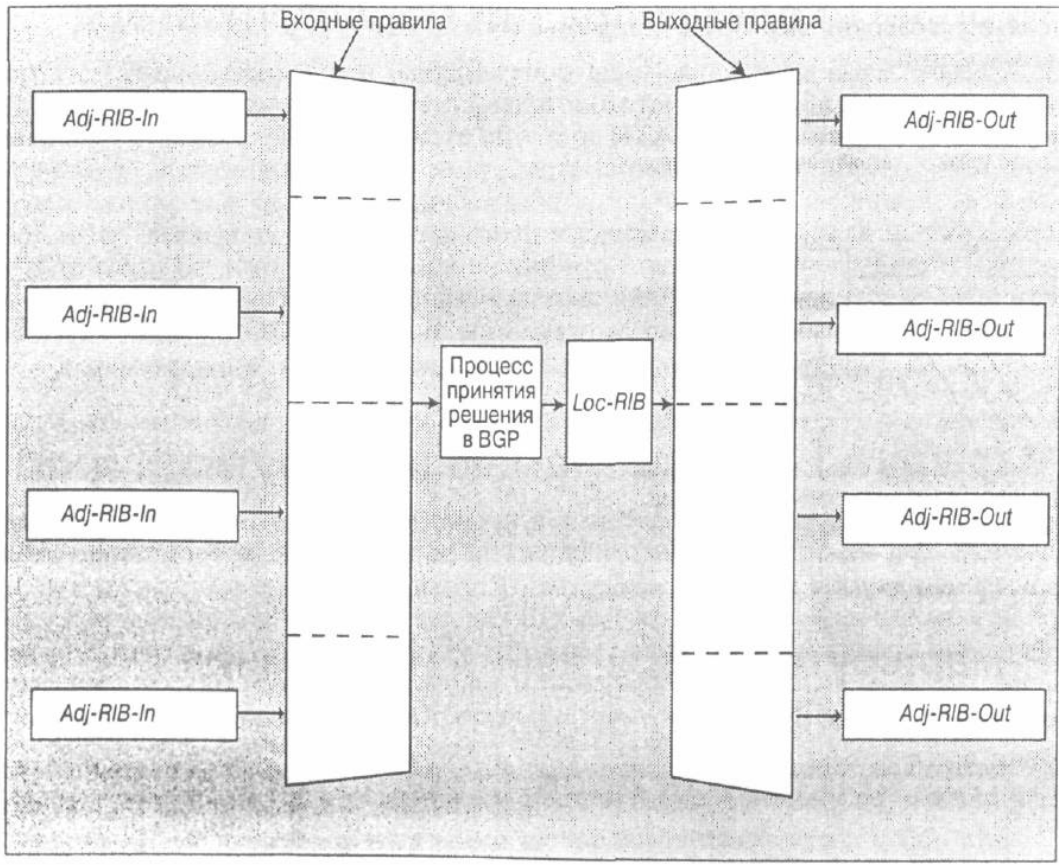


Рис. 6.9. *Логическое представление таблицы маршрутов BGP*

Как показано на рис. 6.9, таблица BGP-маршрутов состоит из трех отдельных частей: информационной базы маршрутизации (Routing Information Base - RIB) Adj-RIB-In, базы Loc-RIB и базы Adj-RIB-Out.

База Adj-RIB-In логически связана с каждым узлом, с которым взаимодействует спикер BGP. В ней хранится маршрутная информация, поступившая от других узлов посредством сообщений UPDATE. Содержимое всех баз Adj-RIB-In доступно в качестве исходных данных для процесса принятия решения после обработки или фильтрации набором входных правил маршрутизации, принятых на данном узле.

База Loc-RIB содержит только предпочтительные маршруты, выбранные в качестве наилучших путей в заданные пункты назначения. Содержимое базы Loc-RIB является результатом процесса принятия решения после применения соответствующих локальных правил маршрутизации, входящих в набор входных правил маршрутизации.

База Adj-RIB-Out логически связана с каждым узлом, взаимодействующим с данным спикером BGP. В ней хранится маршрутная информация, собранная спикером BGP, для объявления другим узлам после применения набора выходных правил маршрутизации.

Несмотря на то что в представленной концептуальной модели рассмотрены три базы — Adj-RIB-In, Loc-RIB и базы Adj-RIB-Out, — хранение трех копий маршрутной информации не является обязательным. На практике для экономии памяти обычно хранится одна копия маршрутной информации с указателями.

Маршруты, полученные от других узлов

Спикер BGP получает маршруты (с соответствующими атрибутами) от внешних и/или внутренних узлов в сообщениях UPDATE. В зависимости от наборов входных правил маршрутизации, все или некоторые из этих маршрутов далее поступают в базу Loc-RIB.

Наборы входных правил маршрутизатора

Наборы входных правил (Input policy Engine) позволяют фильтровать маршруты и изменять их атрибуты. Фильтрация проводится на основе различных параметров, таких как префиксы IP, AS_PATH и другие атрибуты BGP. В BGP набор входных правил используется для управления атрибутами маршрута для того, чтобы повлиять на процесс принятия решения и, следовательно, на выбор маршрута, который будет применяться для доставки трафика в пункт назначения. Например, если для BGP определено фильтровать заданный префикс с какого-либо узла, то через этот узел нежелательно пересылать трафик для доставки получателю. Точно так же, если в BGP определенному префиксу задано оптимальное значение LOCAL_PREF, то в BGP среди нескольких узлов предпочтение отдается префиксу от определенного узла. Набор входных правил конфигурируется оператором.

Маршруты, используемые маршрутизатором

Наилучшие маршруты выявленные в процессе принятия решения, помещаются в базу Loc-RIB. Эти маршруты становятся кандидатами, которые могут быть объявлены другим узлам или помещены в таблицу маршрутов IP. Если маршрут не поступает в базу Loc-RIB, то он не может быть помещен в базу Adj-RIB-Out для дальнейшего объявления другим узлам.

Кроме приема маршрутов от сторонних узлов, маршрутизатор (если это задано в его конфигурации) генерирует сообщения о маршрутах к сетям внутри автономной системы. Таким образом, внешним узлам объявляются маршруты ко внутренним сетям AS.

Набор выходных правил маршрутизации

Этот набор представляет собой тот же набор входных правил, но применяемый на выходе. Маршруты, используемые маршрутизатором (т.е. наилучшие маршруты) в дополнение к маршрутам, сгенерированным маршрутизатором локально, обрабатываются именно этим набором правил. В наборе выходных правил маршрутизации (Output Policy Engine) могут применяться фильтры и вноситься изменения в некоторые атрибуты BGP (такие как AS_PATH) перед отправкой сообщения об обновлении маршрутной информации.

Набор выходных правил маршрутизации налагает ограничения на распространение маршрутов между внутренними и внешними узлами; например, маршруты, полученные от одного внутреннего узла, не могут быть переданы другому внутреннему узлу.

Маршруты, объявляемые другим узлам

Набор маршрутов, объявляемых другим узлам, включает в себя маршруты, которые успешно прошли обработку набором выходных правил и готовы быть объявленными другим узлам BGP — внутренним или внешним.

Пример организации маршрутизации

На рис. 6.10 представлен процесс, которому подвергаются BGP-маршруты. На этом рисунке AS5 получает сведения о маршрутах от AS1 и AS2 и, кроме того, генерирует собственные маршруты (172.16.10.0/24). С целью упрощения предположим, что поток обновлений маршрутов происходит в одном направлении — слева направо.

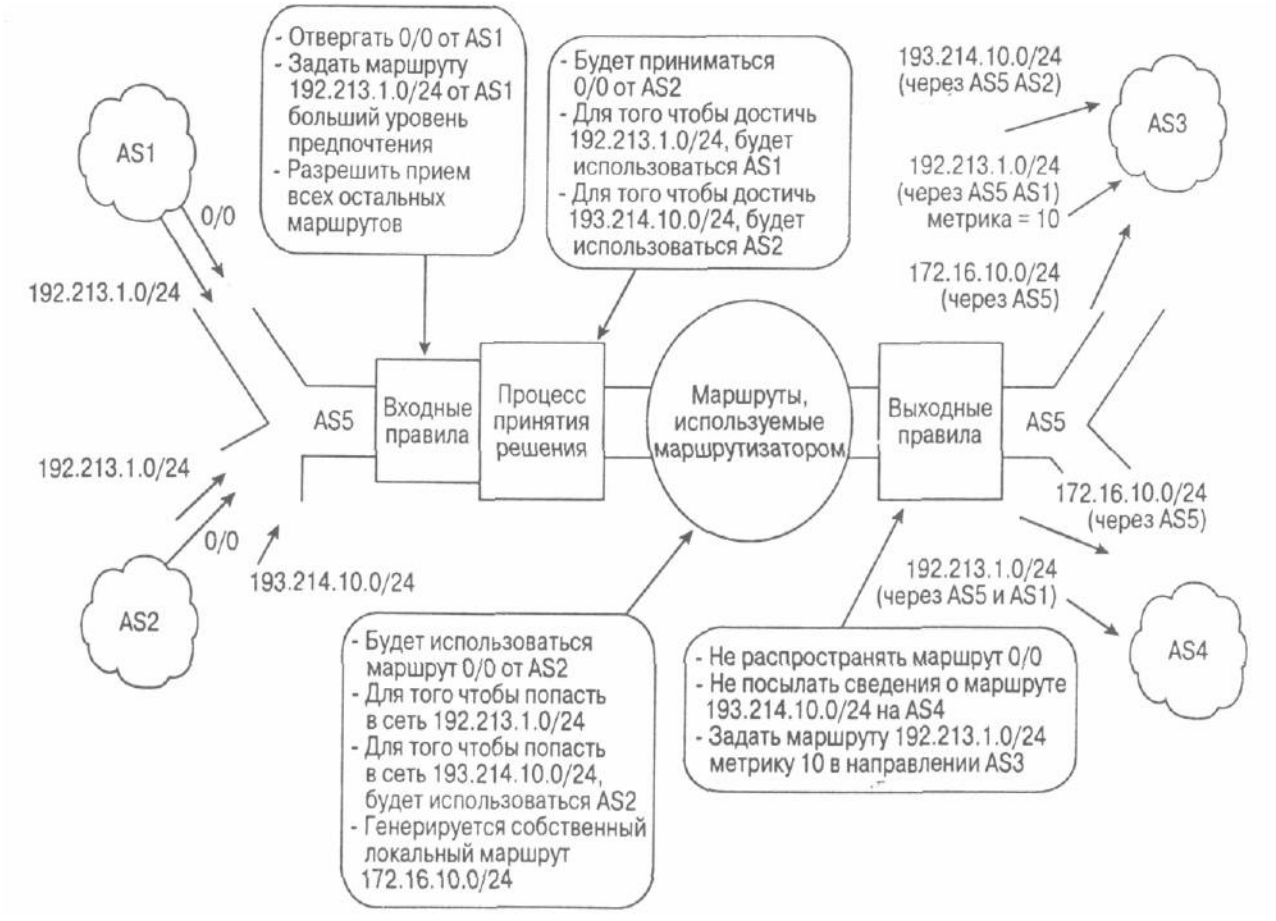


Рис. 6.10. *Пример организации маршрутизации*

Применение модели работы протокола BGP к AS5 дает следующие результаты. Маршруты, получаемые от взаимодействующих узлов (маршруты от AS1 и AS2), включают в себя:

- 192.213.1.0/24 через AS1.
- 0/0 через AS1 (это маршрут по умолчанию).
- 193.214.10.0/24 через AS2.

- 0/0 через AS2 (это тоже маршрут по умолчанию).
- 192.213.1.0/24 через AS2.

Согласно критериям, заданным набором входных правил, получим:
- не принимать маршрут по умолчанию 0/0 от AS1;
- отдавать предпочтение маршруту 192.213.1.0/24 через AS1 перед маршрутом 192.213.1.0/24 через AS2;
- принимать все остальные маршруты (за исключением 193.214.10.0/24).

В процессе принятия решения делаются следующие выводы.
- так как маршрут в сеть 192.213.1.0/24 через AS1 более предпочтителен, то его и будем использовать для того, чтобы попасть в сеть 192.213.1.0/24.
- в сеть 193.214.10.0/24 попадаем через AS2.
- принимаем маршрут 0/0 через AS2.

Маршруты, используемые маршрутизатором:
- в качестве маршрута по умолчанию будем использовать 0/0 через AS2;
- в сеть 192.213.1.0/24 можем попасть через AS1;
- в сеть 193.214.10.0/24 можем попасть через AS2;
- сеть 172.16.10.0/24 является одной из локальных сетей, маршруты к которым мы хотим объявить.

Согласно критериям, выдвигаемым набором выходных правил маршрутизации:
- не сообщать сведения о маршруте по умолчанию 0/0;
- не объявлять маршрут в сеть 193.214.1.0/24 через AS4;
- при пересылке маршрута к сети 192.213.1.0/24 на AS3 присвоить ему метрику 10.

Маршруты, объявляемые другим узлам в направлении AS3, включают в себя:
- 192.213.1.0/24 (через AS5 и AS1) (т.е. сначала на AS5, а затем на AS1) с метрикой 10;
- 172.16.10.0/24 (через AS5);
- 193.214.10.0/24 (через AS5 и AS2).

Маршруты, объявляемые другим узлам в направлении AS4, включают в себя:
- 192.213.1.0/24 (через AS5 и ASI);
- 172.16.10.0/24 (через AS5).

Процесс принятия решения в BGP

Процесс принятия решения об использовании маршрута в протоколе BGP основан на обработке атрибутов. При наличии нескольких маршрутов к одному узлу с префиксами одинаковой длины в BGP выбирается наилучший из них, по которому и направляется весь трафик. Ниже описывается процесс выбора наилучшего маршрута в BGP:

1. Если ближайший следующий узел недоступен, то маршрут игнорируется. (Поэтому важно всегда иметь IGP-маршрут на ближайший соседний узел).
2. Предпочитается маршрут с наибольшим весом. (Вес маршрута — специальный параметр, используемый в маршрутизаторах компании Cisco локально).
3. Если веса маршрутов оказались одинаковыми, то следует выбрать из них маршрут с наибольшим значением коэффициента предпочтения.
4. Если нет локально сгенерированных маршрутов и коэффициент предпочтения оказался одинаковым, то следует предпочесть маршрут с наименьшим значением атрибута AS_PATH (т.е. самый короткий путь).

5. Если длина AS_PATH у маршрутов совпадает, то следует выбрать маршрут с наименьшим значением атрибута типа протокола ORIGIN (где IGP стоит ниже EGP, а EGP - ниже, чем INCOMPLETE).
6. Если атрибут типа протокола также совпадает, то следует выбрать маршрут с наименьшим значением атрибута MED, если маршруты были приняты от одной и той же AS (или если была задана команда bgp always-compare-med).
7. Если у маршрутов равные значения MED, то IBGP-марщрутам следует предпочесть EBGP-маршруты.
8. Если во всех предыдущих случаях получены совпадения, то следует предпочесть маршрут, который пролегает через ближайшего соседа по IGP, т.е. предлагается избрать кратчайший путь к удаленному узлу внутри AS. (Следовать кратчайше му пути до узла, указанного в NEXT_HOP).
9. Если и внутренние маршруты окажутся одинаковыми, то для решения этой за дачи следует использовать атрибут ROUTER_ID. В этом случае следует предпочесть маршрут, полученный от маршрутизатора BGP, с наименьшим значением RID. В Cisco IOS в качестве RID выступает адрес петли, если такой сконфигурирован; в противном случае -- наибольший IP-адрес маршрутизатора. Установление RID зависит от изготовителя конкретного оборудования.

Если разрешено применение BGP Multipath (см. главу 11), то шаги с седьмого по девятый можно пропустить. Тогда все маршруты с одинаковой длиной атрибута AS_PATH и одинаковыми значениями MED могут помещаться в таблицу маршрутов. Некоторые реализации позволяют пропускать девятый шаг и использовать в качестве активных маршруты, "установленные изначально".

Управление маршрутами в BGP

В предыдущем разделе мы обсудили работу наборов правил, с помощью которых можно управлять атрибутами и проводить фильтрацию маршрутов. В этом разделе детально рассматриваются фильтрация и управление атрибутами маршрутов, которые являются ключевыми для управления маршрутами. Подробно рассматривается каждый атрибут BGP, чтобы показать, на что и как именно он влияет.

Трафик внутри и вне AS всегда перемещается согласно "атласу автомобильных дорог", сформированному на основе маршрутов. Внесение изменений в маршруты приводит к изменению направления потоков трафика. Ниже приведены наиболее часто задаваемые вопросы, касающиеся управления маршрутами.

- Как запретить объявление внутренних сетей во внешнюю сеть?
- Как организовать фильтрацию обновлений маршрутов, поступающих от опре деленного узла?
- Чем можно доказать, что используется именно нужное соединение и именно от нашего провайдера, а не какое-то другое?

Как вы увидите далее, в BGP имеются все необходимые средства и атрибуты, позволяющие получить ответы на эти и многие другие вопросы.

Атрибуты маршрутов в BGP

Атрибуты BGP представляют собой набор параметров, которые описывают различные характеристики префиксов (маршрутов). В процессе принятия решения в BGP все атрибуты связываются с префиксами, которые они описывают, сравниваются все доступные маршруты к заданному пункту назначения и затем выбираются лучшие из них. Помните о том, что атрибуты являются неотъемлемой частью пакета UPDATE протокола BGP и заключают в себе информацию описательного характера для определенного префикса. В последующих разделах будут рассмотрены различные атрибуты и их воздействие на работу системы маршрутизации.

Прежде чем перейти к описанию атрибутов маршрутов, давайте рассмотрим их основные категории.

- Обязательные общеизвестные (Well-known mandatory).
- Общеизвестные, используемые по собственному усмотрению (Well-known discretionary).
- Необязательные транзитивные (Optional transitive).
- Необязательные нетранзитивные (Optional nontransitive).

Вот что сказано о категориях атрибутов в RFC 1771.

"Общеизвестные атрибуты должны распознаваться всеми реализациями протокола BGP. Часть этих атрибутов обязательна для применения и должна всегда включаться в состав сообщения UPDATE. Другие - - оставлены на ваше усмотрение и могут или включаться не включаться в сообщение UPDATE.

Общеизвестные атрибуты должны передаваться другим узлам (после соответствующего обновления, если это необходимо).

Кроме общеизвестных атрибутов, каждому маршруту можно присваивать один или несколько необязательных атрибутов. Поддержка этих атрибутов всеми реализациями BGP не требуется. Возможность обработки неопознанного необязательного атрибута определяется установкой бита транзитивности из октета флагов атрибутов в определенное значение. Маршруты с неопознанными необязательными транзитивными атрибутами должны быть обработаны. Если маршрут с неопознанным необязательным транзитивным атрибутом обрабатывается и пересылается другим узлам BGP, то и этот атрибут также передается на другие узлы с битом частичности (Partial bit) и октетом флагов атрибутов, установленными в 1. Если принят и передан на другие узлы BGP-маршрут с опознанным необязательным транзитивным атрибутом, то бит частичности и октет флагов атрибутов уже установлены в 1 предыдущей AS и не сбрасываются в 0 текущей AS. Неопознанные нетранзитивные необязательные атрибуты должны игнорироваться, и передавать сведения о них на другие узлы BGP не нужно.

На маршруте следования можно добавлять новые необязательные транзитивные атрибуты. Это может выполнять узел, генерирующий маршрут , либо другая AS, через которую проходит маршрут. Если маршруты были добавлены другой AS, то бит частичности и октет флагов атрибутов устанавливаются в 1. Правила добавления новых необязательных нетранзитивных атрибутов зависят от свойств добавляемых атрибутов. Документация на каждый новый необязательный нетранзитивный атрибут, вероятнее всего, будет содержать эти правила. (В качестве примера можно привести атрибут MULTI_EXIT_DISC) . Все необязательные атрибуты (и транзитивные, и нетранзитивные) могут быть внесены любой AS на маршруте (если это разрешено)."

Определенные для протокола BGP атрибуты приведены в списке. Подробную информацию о действии каждого из них мы представим в последующих разделах.

- **ORIGIN (код типа 1)** — обязательный общеизвестный атрибут, который определяет источник, сгенерировавший информацию о маршруте. В его октете данных разрешены следующие значения:
  - **0: IGP** — информация сетевого уровня о доступности сети, которая является внутренней для данной AS;
  - **1: EGP** — информация сетевого уровня о доступности сети, полученная через протокол EGP;
  - **2: INCOMPLETE** — информация сетевого уровня о доступности сети, полученная другими средствами.
- **AS_PATH (код типа 2)** -- обязательный общеизвестный атрибут, который состоит из последовательности сегментов AS, составляющих маршрут. Каждый сегмент AS, составляющий маршрут, представлен параметрами <тип сегмента маршрута,

длина сегмента, значение сегментах

- **NEXTJttOP (код типа 3)** — обязательный общеизвестный атрибут, который определяет IP-адрес граничного маршрутизатора. Этот адрес следует использовать в качестве следующего ближайшего к пункту назначения узла. Он помещается в поле информации сетевого уровня о доступности сети в сообщении UPDATE.
- **MULTI_EXIT_DISC (код типа 4)** — необязательный нетранзитивный атрибут, который представляет собой целое неотрицательное число, занимающее четыре октета. Значение этого атрибута может использоваться в процессе принятия решения спикером BGP для выделения нескольких точек выхода от соседних автономных систем.
- **LOCAL_PREF (код типа 5)** — общеизвестный атрибут, используемый по усмотрению, который состоит из целого неотрицательного числа размером четыре октета. Он ис пользуется спикером BGP для информирования других спикеров BGP в автономной системе о степени предпочтительности того или иного объявляемого маршрута.
- **ATOMIC_AGGREGATE (код типа 6)** — общеизвестный атрибут, используемый по усмотрению, длиной 0 октетов. Он используется спикером BGP для информирования других спикеров BGP о том, что в локальной системе выбран менее специфичный (неопределенный) маршрут вместо более специфичного (однозначно определенного), который уже включен в него.
- **AGGREGATOR (код типа 7)** — необязательный транзитивный атрибут длиной 6 октетов. Этот атрибут содержит номер последней AS, где был сформирован объединенный маршрут (закодированный двумя октетами), за которым следует IP-адрес спикера BGP, сформировавшего объединенный маршрут (заключенный в четыре октета).
- **COMMUNITY (код типа 8)** — необязательный транзитивный атрибут переменной длины. Этот атрибут состоит из набора четырех октетов, каждый из которых определяет сообщество. Все маршруты, содержащие этот атрибут, принадлежат к сообществу, заданному в атрибуте.

См. в главе 11 на с. 299 раздел "Атрибуты BGP"

Атрибут ORIGIN

ORIGIN — обязательный общеизвестный атрибут (код типа 1), который указывает на источник обновления маршрута с учетом автономной системы. В протоколе BGP допускаются следующие типы источников.

- IGP — Информация сетевого уровня о доступности сети (Network Layer Reachability Information -- NLRI), являющаяся внутренней по отношению к AS, где был сформирован маршрут.
- EGP – Информация сетевого уровня о доступности сети, полученная через протокол внешнего шлюза Exterior Gateway Protocol (EGP).
- INCOMPLETE — Информация сетевого уровня о доступности сети полученная другими средствами.

В протоколе BGP атрибут ORIGIN используется в процессе принятия решения для установления предпочтительности маршрутов. Как правило, в BGP предпочитается маршрут с самым младшим типом источника (IGP младше EGP, а EGP младше INCOMPLETE). Более детально о механизме формирования атрибута ORIGIN читайте в разделе "Атрибут маршрута ORIGIN".

Атрибут AS_PATH

AS_PATH является обязательным общедоступным атрибутом (код типа 2), в котором содержится последовательность номеров автономных систем, пересекаемых на маршруте. При прохождении маршрутов через спикеры BGP внутри AS, изменения в атрибут AS_PATH не вносятся. Однако при пересылке маршрутов внешним BGP-узлам в

этот атрибут подставляется номер AS, где был сформирован маршрут. Впоследствии каждая AS, которая принимает маршрут и пересылает его другим EBGP-узлам, будет помещать свой номер в список AS. *Размещение (prepending)* — это добавление номера AS в начало списка номеров. В последнем списке должны быть представлены номера всех AS, через которые пролегает маршрут. Номер AS, где был сгенерирован маршрут, помещается в конце списка номеров (перед кодом атрибута ORIGIN). Этот тип списка атрибута AS_PATH называют еще последовательностью AS *(AS_SEQUENCE),* так как все номера AS расположены последовательно.

См. в главе 11 раздел "Атрибут AS_PATH"

Атрибут AS_PATH используется в BGP как часть обновления маршрутов (в пакете UPDATE) для того, чтобы предотвратить образование петель маршрутов в сети Internet. Каждый маршрут, передаваемый между BGP-системами, переносит и список всех номеров AS, через которые он прошел. Если маршрут объявлен AS, номер которой уже имеется в AS_SEQUENCE, то сообщение UPDATE игнорируется. Спикеры BGP помещают номера своих AS при объявлении маршрутов другим AS. Когда маршрут передается спикеру BGP внутри одной AS, то изменения в AS_PATH не вносятся.

На рис. 6.11 представлено формирование атрибута AS_PATH в каждой точке маршрута к сети 172.16.10.0/24. Этот маршрут образован в AS1, передан на AS2, затем — на AS3 и AS4 и возвращен на AS 1. Обратите внимание на изменения, вносимые каждой AS на пути следования. Каждая AS добавляет свой номер в начало списка номеров. При возвращении на AS1 граничный маршрутизатор этой AS обнаруживает, что этот маршрут уже проходил через ASI (так как номер AS1 имеется в списке), поэтому он отвергает маршрут.

Информация, заключенная в атрибуте AS_PATH, используется протоколом BGP при определении наилучшего маршрута к пункту назначения. При сравнении нескольких маршрутов, если остальные, более приоритетные атрибуты одинаковы, то предпочтение отдается маршруту с более коротким списком в AS_PATH. В случае возникновения конфликтной ситуации для определения наилучшего маршрута используются другие атрибуты (см. раздел "Процесс принятия решения в BGP").

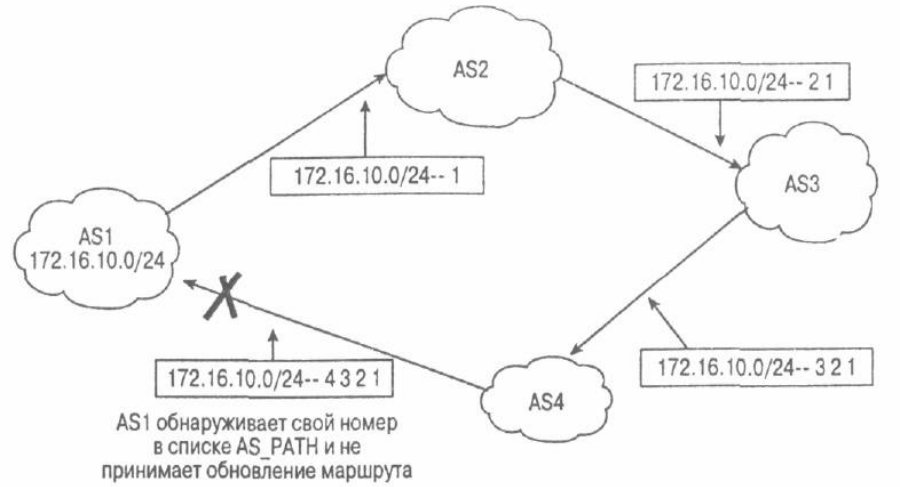


Рис. 6.11. *Пример образования петли и работы атрибута AS_PATH*

Атрибут NEXT_HOP

NEXT_HOP — обязательный общеизвестный атрибут (код типа 3). Он может немного видоизменяться при использовании в контексте протокола IGP, где в качестве ближайшего следующего узла используется IP-адрес интерфейса маршрутизатора, объявленного при анонсировании маршрута.

См. в главе 11 на с. 302 раздел "Атрибут NEXTJHOP"

Концепция работы с ближайшим следующим узлом в протоколе BGP сложнее, чем в

IGP.

- Для EBGP-сеансов следующий узел — это IP-адрес соседнего узла, анонсированного в маршруте.
- Для IBGP-сеансов, т.е. для маршрутов, сгенерированных внутри AS, следующий узел — это также IP-адрес соседнего узла, анонсированного в маршруте.
- Для маршрутов, поступивших в AS по EBGP, сведения о следующем узле из EBGP переносятся без изменений в IBGP. Следующий узел в этом случае - IP-адрес соседнего E
- BGP-узла, от которого был получен маршрут.
- Когда информация о маршрутах распространяется в среде с множественным доступом (такой как Ethernet, Frame Relay и т.д.), под следующим узлом обычно подразумевают IP-адрес интерфейса маршрутизатора, который подключен к среде, сгенерировавшей маршрут.

Маршрутизатор SF участвует в сеансе EBGP с маршрутизатором LA и в сеансе IBGP с маршрутизатором SJ. Маршрутизатор SF получает сведения о маршруте в сеть 128.213.1.0/24 от маршрутизатора LA. В свою очередь, маршрутизатор SF посылает сведения о своей локальной сети 192.212.1.0/24 в BGP.

Маршрутизатор SJ получает сведения о маршруте в сеть 192.212.1.0/24 через интерфейс с IP-адресом 2.2.2.2, который является адресом узла IBGP, анонсировавшего маршрут. Таким образом, 2.2.2.2, согласно определению, является следующим узлом для SJ для того, чтобы попасть в сеть 192.212.1.0/24. Точно так же маршрутизатор SF рассматривает и маршрут к сети 128.213.1.0/24, поступающий от маршрутизатора LA через ближайший узел 1.1.1.1. Когда он сообщает об обновлении маршрута маршрутизатору SJ по IBGP, то маршрутизатор SF не вносит никаких изменений в информацию о следующем узле. Таким образом, маршрутизатор SJ получит информацию о маршруте в сеть 128.213.1.0/24 по BGP через следующий узел с адресом 1.1.1.1. Это яркий пример переноса информации о ближайшем узле EBGP в IBGP.

Как видно из примера, следующим узлом не обязательно должен быть узел, с которым имеется прямое соединение. Например, следующий узел на маршрутизаторе SJ для сети 128.213.1.0/24 — 1.1.1.1, но чтобы достичь его, нужно пройти через узел 3.3.3.3. Таким образом, работа со следующими узлами требует рекурсивных запросов по IP, чтобы маршрутизатор "знал" куда ему отправлять пакеты. Чтобы попасть на узел 1.1.1.1, маршрутизатор SJ будет рекурсивно опрашивать свою таблицу маршрутов IGP, чтобы определить, каким образом можно достичь узла 1.1.1.1. Подобный рекурсивный поиск продолжается до тех пор, пока маршрутизатор не свяжет пункт назначения 1.1.1.1 с исходящим интерфейсом. Точно так же выполняются процедуры для узла 2.2.2.2. Если невозможно достичь определенного узла, то BGP принимается решение о недоступности маршрута.

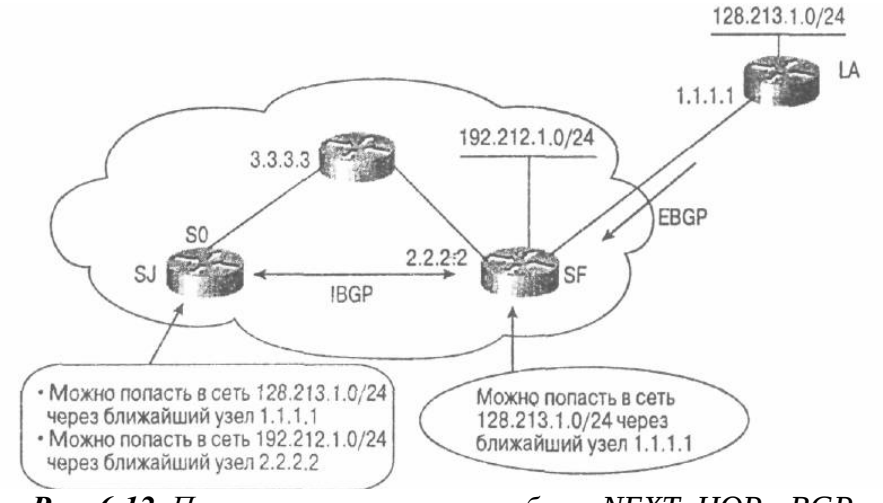


***Рис. 6.12**. Пример применения атрибута NEXT_HOP в BGP*

Ниже мы приводим пример использования рекурсивного запроса IP для направления трафика в пункт конечного назначения. В табл. 6.2 и 6.3 приведены таблицы BGP- и IP-

маршрутов для маршрутизатора SJ из рис. 6.12.

Таблица 6.2. *Таблица BGP-маршруов для маршрутизатора SJ*

Пункт назначения	Следующий узел
192.212.1.0/24	2.2.2.2
128.213.1.0/24	1.1.1.1

Таблица 6.3. *Таблица IP-маршрутов для маршрутизатора SJ*

Пункт назначения	Следующий узел
192.212.1.0/24	2.2.2.2
2.2.2.0/24	3.3.3.3
3.3.3.0/24	Подключено; Порт 0
128.213.1.0/24	1.1.1.1
1.1.1.0/24	3.3.3.3

В табл. 6.2 показано, что в сеть 128.213.1.0/24 можно попасть через узел 1.1.1.1. Из таблицы IP-маршрутов видно, что в сеть 1.1.1.0/24 можно попасть через узел с адресом 3.3.3.3. Еще один рекурсивный запрос таблицы IP-маршрутов свидетельствует о том, что сеть 3.3.3.0/24 напрямую подключена к порту 0 маршрутизатора. Таким образом, трафик в направлении 1.1.1.1 следует пересылать по порту 0. Руководствуясь теми же принципами, трафик пересылается на узел с адресом 2.2.2.2.

Чтобы убедиться в доступности узлов, указанных в атрибуте NEXT_HOP, следует обеспечить маршрутизацию к ним посредством какого-либо протокола IGP или с помощью статических маршрутов. Если узел, указанный в атрибуте NEXT_HOP, недоступен, то BGP-маршрут считается недействительным.

Атрибут MULTI_EXIT_DISC

Атрибут многоточечного дискриминатора BGP Multiexit Discriminator (MULTI_EXIT_DISC, или MED) представляет собой необязательный нетранзитивный атрибут (код типа 4). Этот атрибут является своего рода подсказкой внешним соседним узлам о предпочтительном маршруте в AS, имеющую несколько точек входа. Атрибут MED известен также как внешняя метрика маршрута. Предпочтение отдается меньшим значениям MED.

В отличие от атрибута LOCAL PREF, обмен атрибутом MED между различными AS допускается, но при этом атрибут MED, однажды принятый AS, уже не должен покидать ее пределов. Когда на AS поступает сообщение об обновлении маршрутов с определенным значением атрибута MED, это значение принимается во внимание в процессе принятия решения внутри AS. Когда передается сообщение об обновлении маршрутов BGP на другую AS, MED сбрасывается в 0 (если явно не задано определенное значение исходящего MED).

Если маршрут формируется в AS, то чаще всего значение MED соответствует метрике IGP-маршрута. Это очень удобно, в особенности, если у клиента несколько соединений с одним провайдером. Метрика IGP в этом случае показывает близость или удаленность сети клиента от точки входа в AS. Сеть, которая находится ближе к точке входа А и дальше от точки входа Б, будет иметь меньшую метрику IGP на граничном маршрутизаторе, подключенном в точку А. При трансляции метрики IGP в MED принимаемый AS трафик должен поступать через соединение, ближайшее к пункту назначения. Подобное поведение является результатом того, что для достижения того же пункта назначения было отдано предпочтение маршруту с меньшим значением MED. Атрибуты MED могут применяться и клиентами и провайдерами для распределения трафика по нескольким каналам между двумя AS.

Если не указано иное, маршрутизатор сравнивает атрибуты MED для различных маршрутов от внешних соседей, которые принадлежат к одной AS. Атрибуты MED от различных AS сравнивать нельзя, так как атрибут MED, связанный с маршрутом, как правило, несет в себе информацию о внутренней топологии AS, о правилах маршрутизации и

о протоколе маршрутизации. Сравнение атрибутов MED от различных AS напоминает сравнение яблок и апельсинов. Однако для администраторов, которые имеют веские причины проводить подобное сравнение, компанией Cisco в маршрутизаторах предусмотрена команда **bgp always-compare-med**, с помощью которой в протоколе BGP можно проводить сравнение MED от различных AS для одного и того же маршрута. Ш См. в главе 11 на с. 308 раздел "Атрибут MULTU_EXIT_DISC".

В примере, приведенном на рис. 6.13, показано, как с помощью атрибута MED одна AS может повлиять на процесс принятия решения в другой AS. На рис. 6.13 сети ANET и YNET пытаются повлиять на выходной трафик сети XNET путем пересылки в эту сеть различных значений атрибута MED.

Сеть XNET принимает маршрут в сеть 121 13.0.0/16 из трех различных источников: SJ (метрика 120), LA (метрика 200) и NY (метрика 50). Маршрутизатор gp будет сравнивать два значения метрик, поступающих из сети ANET, и выберет маршрут на маршрута:затор SJ, так как маршрут к нему был объявлен с меньшей метрикой. Если на маршрутизаторе SF используется команда **bqp always-compale-med,** то он будет сравнивать метрику 120 с метрикой 50, поступившей от маршрутизатора NY, и отдаст предпочтение маршруту на NY в сеть 128.213.0.0/16. На процесс принятия решения на маршрутизаторе SF можно повлиять также с помощью атрибута LOCAL_PREF, задавая его в сети XNET, для того, чтобы "перекрыть" метрики, поступающие от других AS. Тем не менее применение атрибута MED оправдано если в сети XNET принятие решений о маршрутизации по BGP базируется на внешних факторах, что способствует упрощению конфигурации маршрутизатора. Клиенты, у которых имеется несколько соединений с провайдером в различных географических точках, могут обмениваться метриками со своими провайдерами, влияя на потоки трафика друг через друга, что позволяет лучше распределить нагрузку.

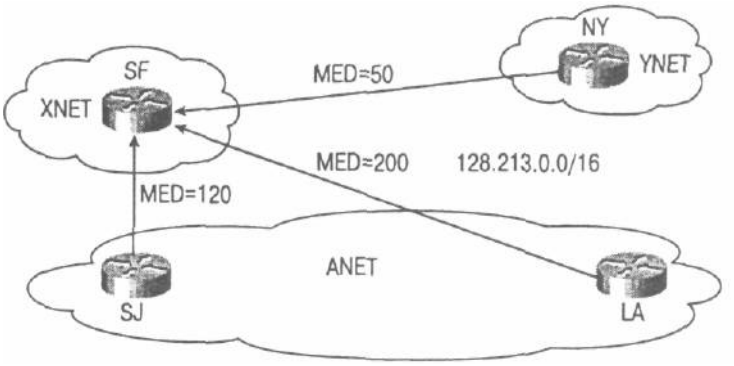


Рис. 6.13. *Действие атрибута MED*

Атрибуты MED являются своего рода препятствием при объединении маршрутов, когда провайдеры анонсируют заданный блок CIDR из нескольких географических точек сети и исключают из блока кратчайшие маршруты. Применение MED в этих случаях может привести к неоптимальной маршрутизации, так как однозначно определенные маршруты из блока CIDR могут рассеиваться между несколькими AS, и атрибуты MED, связанные с более точными маршрутами, будут недействительны.

При использовании атрибутов MED для выполнения *маршрутизации с наилучшим исходом (best-exit routing)* некоторые провайдеры разрешают работать с определенными маршрутами внутри блоков CIDR — благодаря этому появляется возможность выбрать узлы и удалить с их помощью ответвления, возникающие в результате агрегации. Единственная проблема в этом случае — сложность управления анонсированием определенных маршрутов, следствием чего может стать неоптимальная маршрутизация.

При использовании агрегации атрибуты MED не всегда воспринимаются другими узлами не только из-за потенциальной угрозы неоптимальной маршрутизации, но по причине того, что сети получают значение MED из интеллектуальных метрик IGP или не обеспечивается совместимость всех MED с префиксами от различных AS. В любом случае для AS, обладающих соответствующими MED, такой результат лучше, чем слабый процесс принятия решений.

Если провайдеры не используют атрибуты MED или другие приемы для обеспечения

маршрутизации с наилучшим исходом, то такая маршрутизация называется *маршрутизацией по ближайшему выходу (closest exit)* или *маршрутизацией "горячей картошки " (hot potato routing).* Большое количество ответвлений маршрутов является результатом маршрутизации по ближайшему выходу, обычно применяемой при междоменной маршрутизации в точках обмена трафиком.

Атрибут Local_Preference

Атрибут локальных предпочтений Local Preference (сокращенно LOCAL_PREF) является общеизвестным атрибутом, используемым по вашему усмотрению (код типа 5). Атрибут Local Preference показывает степень предпочтения, заданную данному маршруту, по сравнению с другими маршрутами в тот же пункт назначения. Наивысшее значение коэффициента предпочтения указывает на то, что данный маршрут более предпочтителен по сравнению с другими. Коэффициент локального предпочтения, как видно из его названия, является локальным для данной автономной системы, и обмен его значениями разрешен только между узлами IBGP. Если AS имеет соединения с другими AS по протоколу BGP, то она будет получать сведения о маршрутах к одним и тем же пунктам назначения от различных AS. Коэффициент локального предпочтения обычно используется для установки точки выхода из AS для того, чтобы достичь определенного узла получателя. Поскольку этим атрибутом могут обмениваться все BGP-маршрутизаторы внутри AS, то все они будут "знать" об общей точке выхода из AS.

Ш См. в главе 11 на с. 306 раздел "Атрибут LOCAL_PREF"

Давайте рассмотрим вариант инфраструктуры, представленный на рис. 6.14. Допустим, компания ANET, имеющая сеть с одноименным названием, оплатила подключение к Internet через двух сервис-провайдеров — XNET и YNET. Так, ANET подключена к YNET по основному каналу типа T3 и к XNET — по резервному каналу типа T1.

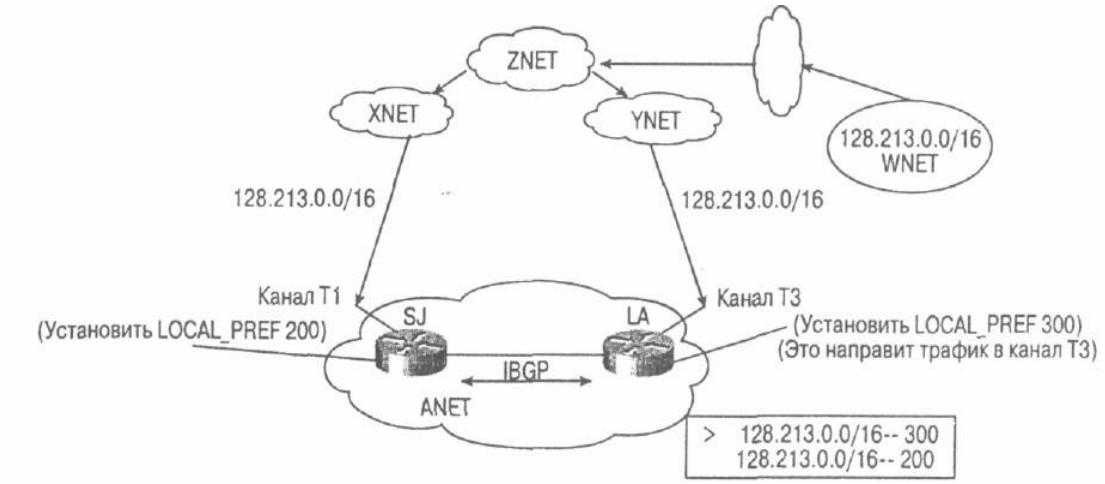


Рис. 6.14. *Пример применения атрибута Local Preference*

Для сети ANET критичным является выбор канала для пересылки выходного трафика. Конечно же, ANET в нормальных условиях предпочтет использовать канал T3 на YNET, так как он является более высокоскоростным.

И здесь вступают в действие коэффициенты предпочтения, задаваемые атрибутом Local Preference: маршрутизатор LA назначает маршрутам, принятым от YNET, коэффициент предпочтения 300. Маршрутизатор SJ присваивает маршрутам в сеть XNET меньший коэффициент предпочтения, например, 200. Так как маршрутизаторы LA и SJ обмениваются между собой маршрутной информацией по IBGP, то они оба соглашаются, что точкой выхода из AS следует считать канал с провайдером YNET, так как он имеет более высокий коэффициент предпочтения. Из рис. 6.14 видно, что сеть ANET получает сведения о маршруте 128.213.0.0/16 через XNET и YNET. Маршрутизаторы LA и SJ согласны использовать канал через YNET как точку выхода для сети 128.213.0.0/16, так как он имеет самый высокий коэффициент предпочтения 300. Поскольку коэффициент локальных предпочтений является локальным только для самой AS (т.е. для сети ANET), все

обсуждаемые нами манипуляции будут влиять только на исходящий трафик, никак не отражаясь на входящем трафике. Входящий трафик по-прежнему может поступать в AS по каналу T1.

Параметр weight (вес), используемый в оборудовании Cisco, имеет сходство с атрибутом Local Preference в том, что он присваивает маршруту с более высоким коэффициентом предпочтения больший вес. Разница между ними состоит в том, что параметр веса является локальным для конкретного маршрутизатора и не передается другим маршрутизаторам, даже если они принадлежат одной AS. Параметр weight влияет на маршруты, принятые от различных провайдеров одним маршрутизатором (например, маршрутизатор с несколькими соединениями от двух и более провайдеров). Параметр **weight** имеет более высокий приоритет, чем любой из атрибутов BGP, и используется в качестве коммутатора при определении предпочтительности того или иного маршрута.

Атрибут ATOMIC_AGGREGATE

Агрегация (объединение) маршрутов приводит к потерям информации, так как объединенные маршруты имеют несколько различных источников с различными атрибутами. Атрибут ATOMIC_AGGREGATE представляет собой общеизвестный атрибут, используемый по собственному усмотрению (код типа 6), который устанавливается в определенное значение при потере информации. По существу, если в системе распространяются объединенные маршруты, приводящие к потере информации, то в такой маршрут необходимо включать атрибут ATOMIC_AGGREGATE.

Атрибут ATOMIC_AGGREGATE не должен устанавливаться, если объединенный маршрут переносит дополнительную информацию, с помощью которой можно установить источник объединения. В качестве примера можно привести объединенный маршрут с параметром AS_SET, как уже обсуждалось ранее. К объединенному маршруту, в котором содержатся данные о сформировавших его AS, атрибут ATOM1Q_AGGREGATE не добавляется.

См. в главе 11 раздел "Только объединенные маршруты, подавление однозначно определенных маршрутов"

Атрибут AGGREGATOR

AGGREGATOR является необязательным транзитивным атрибутом (код типа 7). Он определяет автономную систему и маршрутизатор, сформировавший объединенный маршрут. Спикер BGP, на котором выполняется объединение маршрутов, может добавлять атрибут AGGREGATOR, где содержится номер AS, к которой принадлежит спикер, и его IP-адрес. В оборудовании Cisco IP-адрес -- обычно идентификатор маршрутизатора ROUTER_1D (RID), который представляет собой адрес технологической петли маршрутизатора, если таковой сконфигурирован. Если не задан адрес петли, то самый старший IP-адрес на маршрутизаторе и становится RID. Петельный интерфейс (или, как его еще называют, технологическая петля) — это виртуальный интерфейс, о котором мы говорили ранее в этой главе. На рис. 6.15 представлен пример функционирования атрибута AGGREGATOR.

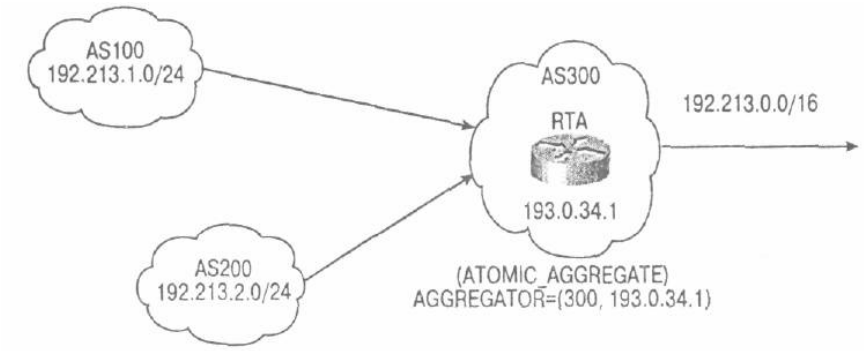


Рис. 6.15. *Пример реализации атрибута AGGREGATOR*

Здесь AS300 получает маршруты 192.213.1.0/24 и 192.213.2.0/24 от AS200.

Генерируя объединенный маршрут 192.213.0.0/16, маршрутизатор RTA может включить атрибут AGGREGATOR, который содержит номер AS 300 и RID маршрутизатора (RTA) 193.0.34.1, сформировавшего объединенный маршрут.

Атрибут COMMUNITY

В контексте протокола BGP *сообществом (community)* называют группу узлов, которые имеют общее свойство. Сообщество не ограничивается рамками одной сети или одной автономной системы, и узлы, входящие в него, необязательно должны быть связаны друг с другом физически. Для примера приведем группу сетей, которые принадлежат образовательному или государственному сообществам. Эти сети могут принадлежать любой автономной системе. Понятие сообщества используется для упрощения правил маршрутизации — маршруты определяются по логическому критерию, а не по префиксу IP или номеру AS. Спикер BGP может использовать атрибут COMMUNITY совместно с другими атрибутами для принятия решения о том, какие маршруты принять, предпочесть и передать другим BGP-системам.

См. в главе 11 раздел "Атрибут COMMUNITY"

COMMUNITY (код типа 8) представляет собой необязательный транзитивный атрибут. Этот атрибут имеет переменную длину и состоит из набора четырехбайтовых значений. Диапазоны значений сообществ для атрибута COMMUNITY от 0x00000000 до 0x0000FFFF и от 0xFFFF0000 до 0xFFFFFFFF зарезервированы. Эти сообщества являются общеизвестными, т.е. они имеют глобальное значение. Ниже приведены примеры общеизвестных сообществ.

- NO_EXPORT (0xFFFFFF01) -- маршрут с этим значением сообщества не должен быть объявлен за пределами AS.
- NO_ADVERTISE (0xFFFFFF) — маршрут с этим значением сообщества не дол жен объявляться другим BGP-узлам.

Кроме атрибутов общеизвестных сообществ, можно определять атрибуты частных (закрытых) сообществ для специального применения. Эти атрибуты описаны в RFC 1998', где определен механизм влияния сообществ на выбор маршрута в BGP в сетях сервис-провайдеров.

Общепринятой практикой считается использование первых двух байт атрибута COMMUNITY для номера AS, а последних двух байт — для определения значения, соотносящегося с этой AS. Например, провайдер (AS256), который желает определить закрытое сообщество с именем my-peer-routers, может воспользоваться атрибутом COMMUNITY со значением 256:1 в десятичной записи. Число 256 указывает на то, что именно этот провайдер задал новое сообщество. А цифра один во второй части записи имеет специальное значение для провайдера. В этом случае она обозначает сообщество с именем my-peer-routers.

Один маршрут может обладать несколькими атрибутами COMMUNITY. Если спикер BGP обнаруживает несколько атрибутов COMMUNITY для одного маршрута, то он может действовать с учетом одного из них, группы или всех этих атрибутов. Маршрутизатор имеет возможность добавлять или модифицировать атрибуты COMMUNITY перед передачей маршрутов другим внешним и внутренним узлам.

На рис. 6.16 представлено применение атрибута COMMUNITY. Узел в сети XNET посылает в сеть YNET сведения о маршрутах X и Y с атрибутом сообщества NO_EXPORT, а маршрут Z оставляет без изменений. Маршрутизатор BGP передает в сеть YNET только маршрут Z в направлении ZNET. Сведения о маршрутах X и Y распространяться не будут, так как им был задан атрибут сообщества NO_EXPORT.

Как вы увидите в последующих главах, благодаря сообществам определение правил маршрутизации становится более гибким.

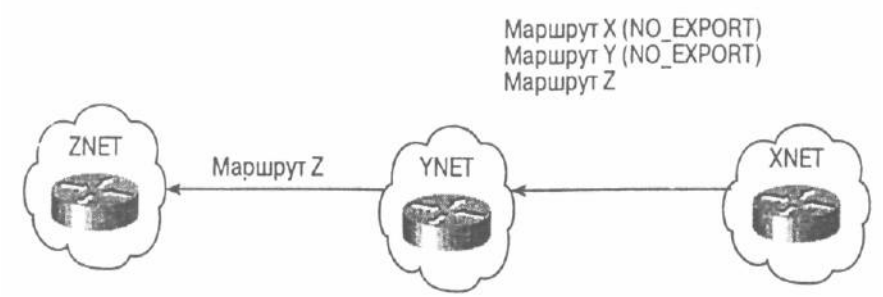


Рис. 6.16. *Применение атрибута COMMUNITY*

Другие атрибуты

Примечание
Атрибуты ORIGINATORJD, CLUSTERJD и CLUSTERJJST обсуждаются в главе 9 "Управление крупномасштабными автономными системами".

Соседние узлы в среде с множественным доступом

Среда, в которую подключены маршрутизаторы, реализующие при обмене данными схему отношений "многие ко многим", называется средой с множественным доступом (МД). Маршрутизаторы в среде МД работают в одной IP-подсети и могут физически получить доступ к другим маршрутизаторам в этой среде через один промежуточный узел (подключенный непосредственно). Сети на базе Ethernet, FDDI, Token Ring, Frame Relay и ATM являются примерами среды с множественным доступом.

Основное правило работы по протоколу IP в среде с МД заключается в том, что маршрутизатор должен всегда объявлять узел, сгенерировавший маршрут, если источник маршрута находится в той же среде с МД, откуда он получил этот маршрут. Другими словами, если маршрутизатор RTA объявляет маршрут, сведения о котором он получил от маршрутизатора RTB, и при этом RTA и RTB находятся в одной среде с МД, то он должен в качестве источника маршрута указать маршрутизатор RTB, а не себя. В противном случае другие маршрутизаторы в той же среде с МД будут при обмене данными использовать лишний узел RTA, чтобы достичь маршрутизатора RTB, который находится в том же сегменте.

На рис. 6.17 маршрутизаторы RTA, RTB и RTC подключены к одной среде с множественным доступом. На маршрутизаторах RTA и RTB поддерживается протокол EBGP, а на маршрутизаторах RTB и RTC - - протокол OSPF. Маршрутизатор RTC получает сведения о сети 11.11.11.0/24 от маршрутизатора RTB по протоколу OSPF. Далее маршрутизатор RTC передает сведения об этом префиксе посредством EBGP на маршрутизатор RTA. Поскольку на маршрутизаторах RTA и RTB поддерживаются различные протоколы маршрутизации, можно решить, что для того, чтобы попасть в сеть 11.11.11.0/24, следующим ближайшим узлом для маршрутизатора RTA должен быть маршрутизатор RTC (10.10.10.2), но на самом деле это не так. В действительности марш руги затор RTA предполагает следующим ближайшим узлом маршрутизатор RTB (10.10.10.3), так как он подключен к той же среде передачи, что и RTC.

В тех случаях, когда среда передачи является широковещательной, такой как Ethernet и FDDI, и между узлами существует физическая взаимосвязь взаимодействие между соседними узлами не вызывает проблем. В ситуациях, когда среда передачи не является широковещательной, например Frame Relay или ATM, следует принять специальные меры, речь о которых пойдет в следующем разделе.

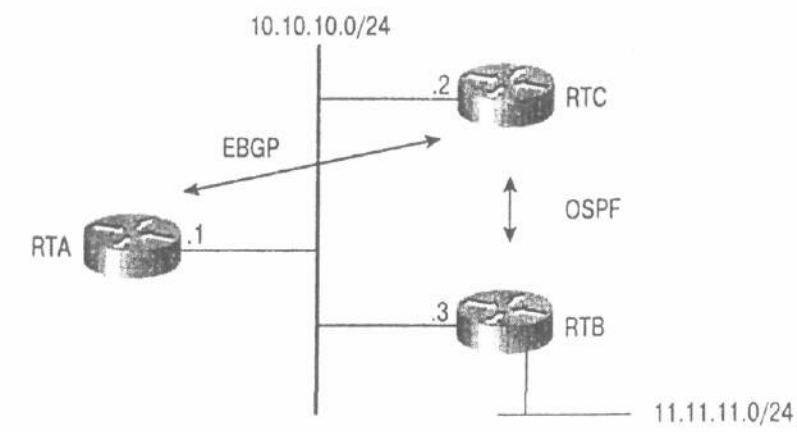


Рис. 6.17. *Пример работы узлов в среде с множественным доступом*

Соседние узлы в нешироковещательной среде с множественным доступом

Такие среды передачи, как Frame Relay и ATM, являются нешироковещательными. При взаимосвязи маршрутизаторов "многие со многими" стабильность обмена данными между ними не гарантируется, если не сконфигурированы виртуальные каналы между маршрутизаторами. Это называется *топологией с полным объединением (full-mesh topology)* и не всегда реализуется на практике по нескольким причинам. Обычно виртуальные каналы Frame Relay и ATM используют несущую за определенную плату, и дополнительные каналы требуют дополнительных финансовых затрат. Кроме высоких финансовых издержек, большинство организаций использует подход центральный сервер —- вспомогательный сервер, когда несколько удаленных узлов объединяют свои виртуальные каналы в одном или нескольких концентрирующих маршрутизаторах на центральном узле, где имеется информация обо всех подобных узлах. На рис. 6.18 показан вариант работы соседних узлов в нешироковещательнон среде передачи.

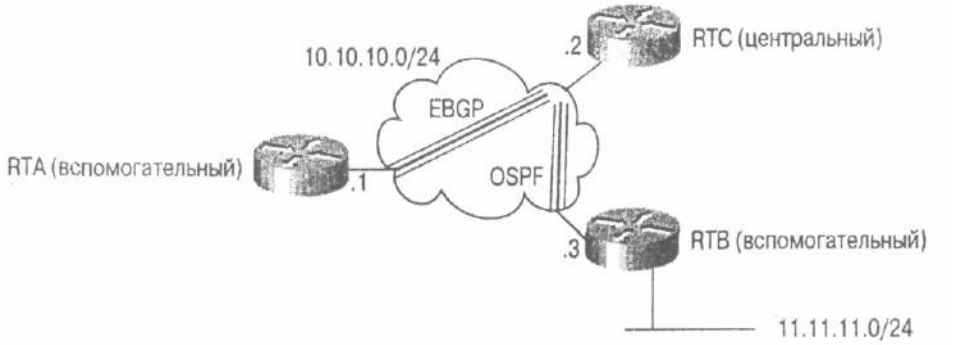


Рис. 6.18. *Вариант функционирования соседних узлов в нешироковещательной среде передачи*

Единственное различие между вариантами функционирования соседних узлов, представленными на рис. 6.17 и 6.18, заключается в том, что на последнем рисунке представлена нешироковещательная среда передачи Frame Relay. Маршрутизатор RTC в этом случае является основным (центральным маршрутизатором), а маршрутизаторы RTA и RTB — вспомогательными. Обратите внимание на виртуальные каналы, образованные между маршрутизаторами RTA и RTC. Как видите, они существуют только между вспомогательными и центральным маршрутизатором, но не между RTA и RTB. Такая топология называется *топологией с частичным объединением (partial-mesh topology).*

Маршрутизатор RTA получает сведения о BGP-маршруте в сеть 11.11.11.0/24 от маршрутизатора RTC, который, в свою очередь, "узнает" об этом маршруте от сгенерировавшего его маршрутизатора RTB. В качестве следующего узла маршрутизатор RTA попытается использовать маршрутизатор RTB (10.10.10.3), т.е. ведет себя так же, как и

в обычной среде с множественным доступом. Однако в этом случае пересылка пакетов состояться не сможет, так как между маршрутизаторами RTA и RTB не существует виртуального канала.

Программное обеспечение Cisco IOS поддерживает специальный параметр, с помощью которого разрешаются подобные конфликтные ситуации. Параметр next-hop-self (задаваемый как часть команды BGP neighbor) заставляет маршрутизатор (в нашем случае RTC) объявлять маршрут в сеть 11.11.11.0/24 со своим адресом (10.10.10.2) в качестве следующего ближайшего узла. Затем маршрутизатор RTA, чтобы достичь сети П. 11.11.0/24, направляет свой трафик на RTC.

Применение команды next-hop-self и объявление зоны демилитаризации

Зоной демилитаризации (demilitarized zone - - DMZ) называют сеть с совместным доступом, лежащую между несколькими AS. IP-сеть, используемая в DMZ, может принадлежать или не принадлежать одной из AS, входящих в сеть DMZ. Как мы уже выяснили ранее, адрес ближайшего следующего узла, получаемый от узла EBGP, сохраняется внутри IBGP. Таким образом, для любого протокола IGP исключительно важно, чтобы был доступен IP-адрес, указанный в атрибуте NEXT_HOP сообщения UPDATE. Одна из возможностей обеспечить это условие — сделать подсеть DMZ частью IGP и объявлять эту подсеть в AS, как и любые другие. Другой способ заключается в "перекрывании" адреса следующего узла, т.е. принудительном задании IP-адреса соседнего граничного IBGP-маршрутизатора в качестве следующего ближайшего узла.

На рис. 6.19 маршрутизатор SJ получает обновление маршрутов с информацией о сети 128.213.1.0/24, где в качестве следующего ближайшего узла указан узел с адресом 1.1.1.1 (часть DMZ). Для того чтобы маршрутизатор SJ мог связаться с этим узлом, сеть 1.1.1.0/24 должна быть объявлена внутри AS граничным маршрутизатором SF,

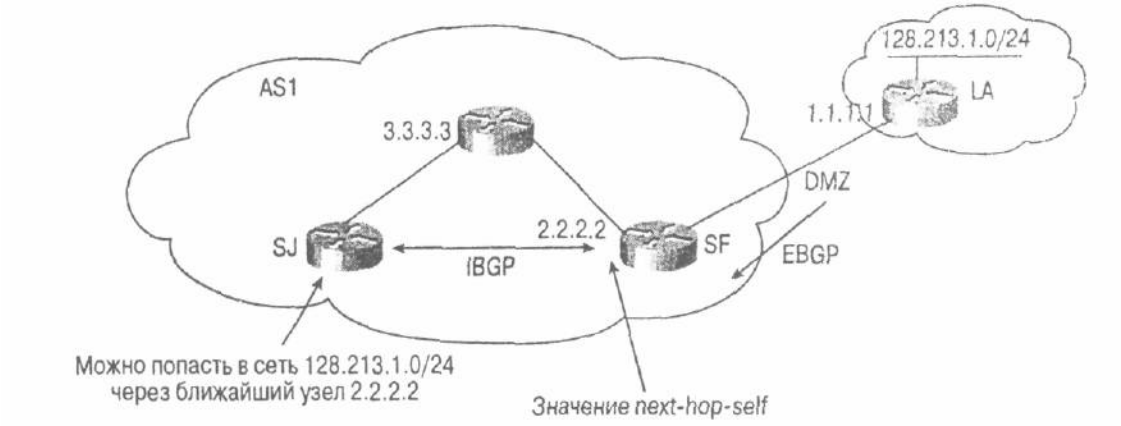


Рис. 6.19. *Применение параметра next-hop-self*

Еше одно условие заключается в том, чтобы на маршрутизаторе SF был установлен параметр next-hop-self для обслуживания соединения IBGP с соседним маршрутизатором SJ. Таким образом, для всех EBGP-маршрутов следующим ближайшим узлом будет узел с адресом 2.2.2.2, который уже входит в IGP. Тогда маршрутизатор SJ без проблем может попасть в свой следующий ближайший узел.

Выбор того или иного метода зависит от необходимости попасть в сеть DMZ. В качестве примера приведем команду ping, задаваемую оператором на интерфейс маршрутизатора внутри AS, который входит в DMZ. Для того чтобы команда ping была выполнена успешно, необходимо включить DMZ в IGP. В других случаях DMZ может быть доступна через какой-либо неоптимальный маршрут вне данной AS. Вместо того чтобы попасть в DMZ изнутри AS, маршрутизатор может попытаться воспользоваться другим EBGP-соединением. В этом случае применение команды next-hop-self обеспечивает возможность достижения ближайшего следующего узла изнутри AS. Во всех других случаях

оба метода обеспечивают нормальную работу по BGP.

Кроме того, важно отметить, что участники точек обмена трафиком в сети Internet часто требуют указания в качестве атрибута NEXT_HOP в сообщении UPDATE IP-адреса взаимодействующего узла и возводят это условие в ранг правила маршрутизации.

Использование частных AS

Для экономного расходования номеров AS InterNIC не распространяет действительные номера AS клиентам, чьи правила маршрутизации являются лишь дополнениями к правилам маршрутизации, установленным их провайдерами. Таким образом, если определенная организация имеет одно или несколько соединений с одним провайдером, то обычно требуется, чтобы этой организацией использовались номера AS из диапазона общедоступных для частного использования (номера с 64512 до 65535). Тогда все сообщения о BGP-маршрутах, получаемые провайдером от узла клиента, буду содержать частные номера AS.

См. в главе 11 на раздел "Использование частных номеров AS"

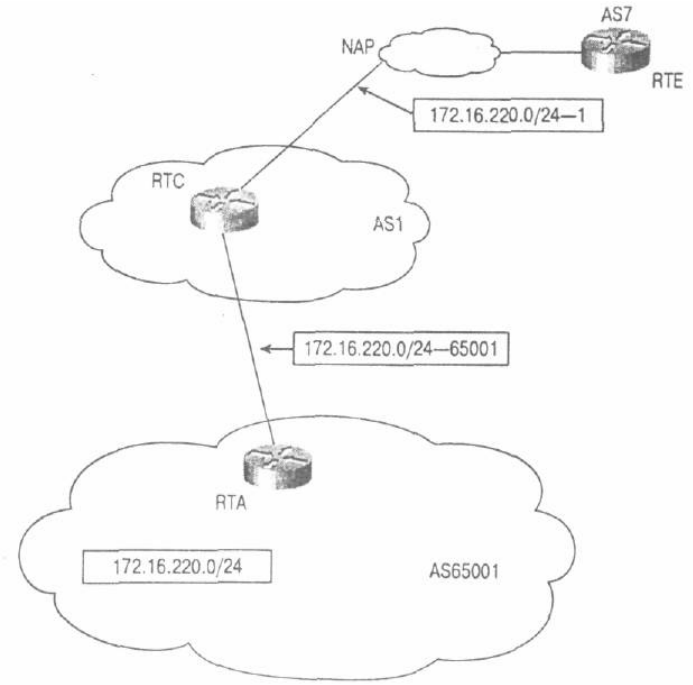


Рис. 6.20. *Изъятие номеров частных AS*

Номера частных AS не могут свободно распространяться по сети Internet, так как они не являются уникальными. По этой причине в своем оборудовании компания Cisco реализовала функцию изъятия номеров частных AS из списка AS_PATH перед его передачей в Internet. Таким образом, происходит своего рода передача полномочий, так называемое проксирование. На рис. 6.20 показана процедура изъятия номеров частных AS из списка AS_PATH.

Примечание

В главе 1, "Эволюция сети Internet", мы говорили о точках доступа к сети (Network Access Point — NAP) и их роли в объединении различных провайдеров. Иногда соединения BGP в NAP организуются с помощью сервера маршрутов, где несколько узлов из различных AS объединяются по протоколу EBGP в единую систему. Сервер маршрутов также имеет собственный номер AS. На рис. 6.20 NAP представлена маршрутом к серверу маршрутов RTE, который имеет номер AS7. Концепция применения сервера маршрутов обычно используется, когда необходимо, чтобы несколько AS для обмена маршрутами по EBGP привязывались к одной точке.

На рис. 6.20 AS1 обеспечивает подключение к Internet сети клиента (AS65001). Так как клиент имеет соединение только с этим провайдером и у него нет в ближайшем будущем планов по подключению через еще один канал к другому провайдеру, то ему выделяется номер частной AS. Если же клиенту потребуется организовать подключение к Internet через еще одного провайдера, то соответствующий его региону реестр сети Internet обязан будет присвоить его AS уникальный номер.

Все префиксы, генерируемые в AS65001, будут иметь атрибут AS_PATH 65001. Рассмотрим, каким образом распространяются сведения о префиксе 172.16.220.0/24 из AS6500I (рис. 6.20). Для того чтобы AS1 передала сведения об этом маршруте в сеть Internet, необходимо избавиться от номера частной AS, Когда префикс попадает в сеть Internet, он будет уже иметь в качестве источника номер AS провайдера. Обратите внимание, что префикс 172.16.220.0/24 приходит в NAP с AS_PATH 1.

Протокол BGP изымает номера частных AS, только если требуется передать информацию об их маршрутах каким-либо внешним узлам. Это означает, что изъятие номеров AS должно быть задано на маршрутизаторе RTC как элемент конфигурации соединения с соседним узлом RTE.

Еще раз обращаем ваше внимание на то, что частные AS должны быть подключены только к одному провайдеру. Если в списке атрибута AS_PATH встречаются номера действительных и частных AS, то протокол BGP расценивает его как неправильно сформированный и не изымает номера частных AS. В этом случае обновление маршрутов будет обрабатываться в обычном порядке. Номера частных AS изымаются только из списков AS_PATH, содержащих хотя бы один номер частной AS в диапазоне от 64512 до 65535.

Дополнительные сведения об архитектуре сетей с использованием частных AS вы можете получить в RFC 2270[2].

Атрибут AS_PATH и объединение маршрутов

Объединение маршрутов предполагает суммирование диапазона префиксов в один или несколько блоков CIDR с целью минимизации количества маршрутов в таблицах маршрутов. Однако информация AS_PATH, содержащаяся в маршрутной информации нескольких маршрутов, будет теряться при одновременном суммировании этих маршрутов. Это может привести к образованию петель маршрутизации, так как маршрут, проходящий через ту же самую AS, может быть воспринят ею как новый маршрут.

Для того чтобы избежать нежелательных последствий объединения маршрутов, в BGP имеется возможность задания типа объекта в AS_PATH параметром AS_SET, где все AS вносятся в список неупорядоченными. В этот набор входят номера AS, через которые проходит маршрут. Объединенные маршруты, в которых переносится информация AS_SET, имеют совокупный набор атрибутов, сформированный индивидуальными маршрутами.

На рис. 6.21 система AS1 объявляет маршрут в сеть 192.213.1.0/24, а AS2 объявляет маршрут в сеть 192.213.2.0/24. Система AS3 объединяет оба маршрута в 192.213.0.0/16. Автономная система, которая объявляет объединенный маршрут, считает саму себя генератором маршрута, независимо от того, откуда поступили составляющие его маршруты. Когда AS3 объявляет объединенный маршрут 192.213.0.0/16, значение в AS_PATH будет 3. Это приведет к потерям маршрутной информации, так как генераторы маршрутов ASI и AS2 уже не будут присутствовать в списке номеров AS__PATH. Если другие AS каким-то образом объявят объединенный маршрут системам AS1 и AS2, они примут этот маршрут, что приведет к возникновению петли маршрутизации. "Зная" значение AS_SET, AS3 может объявлять объединенный маршрут 192.213.0.0/16, сохраняя при этом информацию о составляющих его компонентах. Значение параметра AS_SET {1 2} указывает на то, что объединенный маршрут сформирован из отдельных маршрутов, поступивших от соответствующих AS, без определенного порядка следования. Информация, заключенная в атрибуте AS_PATH, с учетом параметра AS_SET будет выглядеть как 3 {1 2}.

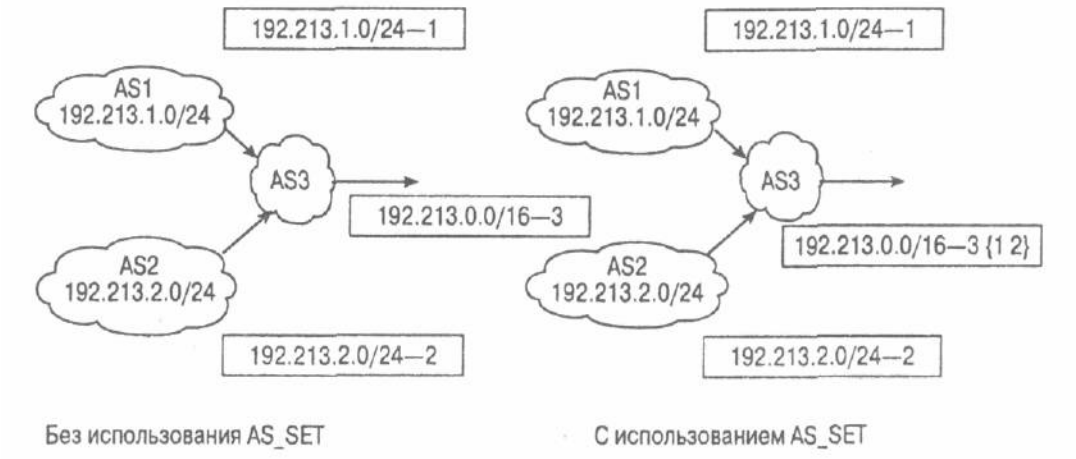


Рис. 6.21. *Влияние параметра AS_SET*

Указание параметра AS_SET перед командой aggregate-address вызовет автоматическую генерацию набора AS, связанного с данным объединенным маршрутом.

Управление атрибутом AS_PATH

Довольно часто целесообразно модифицировать информацию, содержащуюся в атрибуте AS_PATH, для управления междоменной маршрутизацией. Ввиду того что в BGP отдается предпочтение более короткому маршруту, заданному в AS_PATH, операторы сети не могут отказаться от возможности вносить изменения в маршрутную информацию путем включения фиктивных записей в AS_PATH, которые увеличивают длину маршрута и влияют на траекторию движения трафика. В оборудовании Cisco имеется возможность задания номеров AS в начале списка атрибута AS_PATH пользователем, что позволяет удлинить маршрут. Ниже приведен пример, с помощью которого мы поясним, как использовать эту возможность.

См. в главе 11 раздел "Управление атрибутом AS_PATH"

На рис. 6.22 AS50 подключена к двум провайдерам — AS200 и AS100. При этом AS100 имеет непосредственное соединение с NAP, в то время как AS200 для того, чтобы достичь NAP, должна пройти через промежуточный узел AS300.

На рис. 6.22 показан частный случай передачи информации о префиксе 192.213.1.0/24 через различные AS в NAP. Когда префикс 192.213.1.0/24 поступает в NAP через AS300, она имеет AS_PATH 300 200 50. Если этот же префикс поступает в NAP через AS100, то значение атрибута AS_PATH будет 100 50, т.е. короче, чем маршрут через AS300. Так, исходящий из NAP поток "предпочтет" более короткий маршрут согласно AS_PATH, и трафик в сеть 192.213.1.0/24 будет направлен через AS100.

К сожалению, AS50 ведет себя совершенно противоположным образом и старается передать весь входящий трафик по более высокоскоростному каналу Т3 на AS200. Для этого на AS50 в атрибут AS_PATH вставляются фиктивные записи о несуществующих номерах AS на маршруте, и затем эти сведения отправляются с сообщениями об обновлениях маршрутов на AS 100. На практике для формирования фиктивных записей используется повтор собственного номера AS, т.е. AS50 повторяет свой номер столько раз, сколько требуется для того, чтобы изменить баланс в пользу AS200 и представить маршрут через нее менее длинным, чем через AS 100.

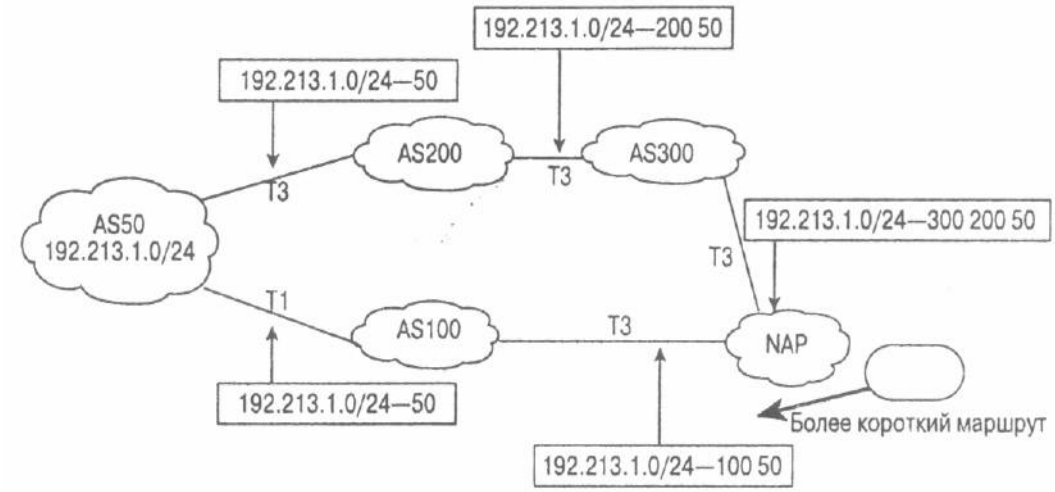


Рис. 6.22. *Пример маршрутизации перед вводом фиктивных записей в AS_PATH*

На рис. 6.23 показано, как AS50 вставляет два номера AS 50 50 в начало атрибута AS_PATH для префикса 192.213.1.0./24, который объявляется AS100. Когда префикс 192.213.1.0.0/24 поступает в NAP через AS100, значение AS_PATH будет 100 50 50 50, что длиннее, чем AS_PATH 300 200 50 через AS300. Таким образом, исходящий из NAP поток отдаст предпочтение более короткому маршруту, и трафик в сеть 192.213.1.0/24 будет проходить через AS300.

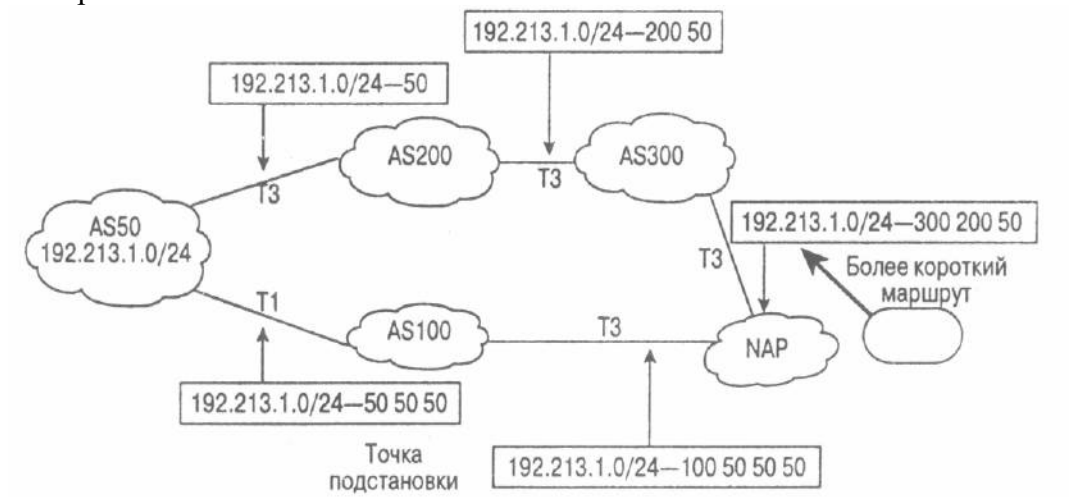


Рис. 6.23. *Пример маршрутизации после вставки фиктивных записей в AS_PATH*

На практике фиктивные номера всегда должны дублировать номер AS, анонсирующей маршрут, или маршрут, полученный от соседнего узла (в том случае, когда AS увеличивает длину маршрута для входящих обновлений). Добавление какого-либо другого номера AS вводит систему в заблуждение и может привести к образованию петель маршрутизации или "черных дыр". Обратите внимание на точку в рис. 6.23, где происходит вставка фиктивных номеров.

Фильтрация маршрутов и управление атрибутами

Концепция фильтрации маршрутов довольно проста. Спикер BGP может выбрать маршруты, которые следует посылать и принимать от других BGP-узлов. Фильтрация маршрутов широко используется при описании правил маршрутизации. Автономная система может идентифицировать входящий трафик, который она согласна принимать от соседних

узлов, путем задания списка маршрутов, объявляемых своим соседним узлам. И наоборот, AS может управлять исходящим трафиком путем указания списка маршрутов, поступающих от соседних узлов, которые она согласна обслуживать.

См. в главе 11 раздел "Фильтрация маршрутов и управление атрибутами"

Фильтрация используется также на уровне протокола для того, чтобы ограничить "перетекание" обновлений маршрутов из одного протокола в другой. Ранее в этой главе мы обсуждали как возможность вложения BGP-маршрутов в IGP, так и IGP-маршрутов, а также статических маршрутов в BGP. Согласно терминологии Cisco, это называется *преобразованием (redistributing)* маршрутной информации между протоколами. В этой главе мы также обсудим опасные последствия взаимного преобразования маршрутов между протоколами. Фильтрация необходима для точного определения информации, поступающей от BGP в IGP, и наоборот.

Маршруты, которым разрешено проходить через фильтр, могут обладать видоизмененными атрибутами. Модифицирование атрибутов влияет на процесс принятия решения в протоколе BGP при определении наилучшего маршрута к заданному пункту назначения.

Входная и выходная фильтрация

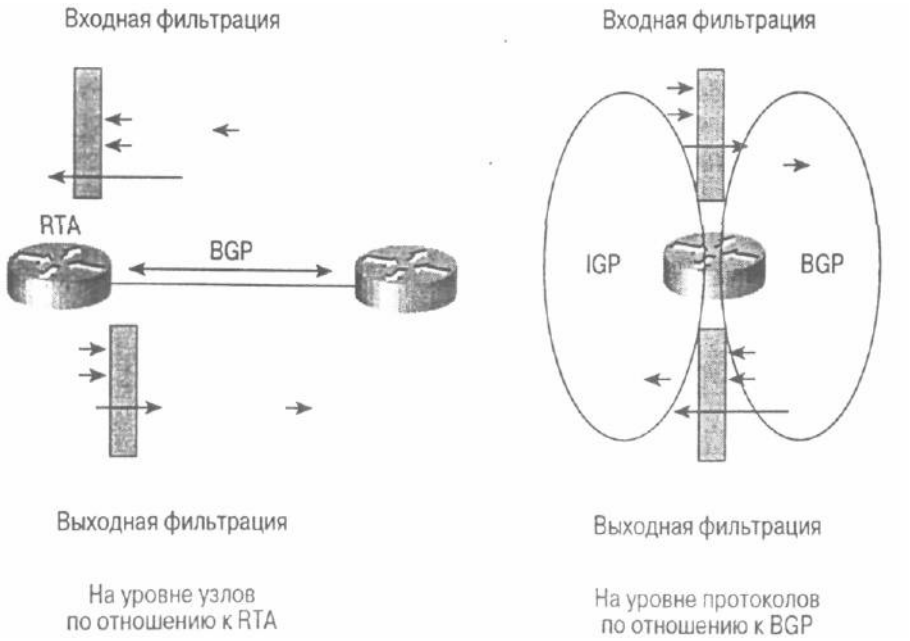


Рис. 6.24. *Пример входной и выходной фильтрации*

Концептуально входная и выходная фильтрация может выполняться и на уровне отдельного узла, и на уровне протоколов. На рис. 6.24 приведен пример входной и выходной фильтрации.

С точки зрения обмена маршрутами между BGP-узлами, входная фильтрация указывает на то, что спикер BGP отсеивает маршрутную информацию, поступающую от других узлов, а при выходной фильтрации ограничения налагаются на маршрутную информацию, распространяемую самим спикером BGP другим узлам. Фильтрация выполняется одинаково, независимо от того, идет ли речь о внешних (EBGP) или о внутренних(1BGP) BGP-узлах.

На уровне протоколов входная фильтрация ограничивает поступление сообщений об обновлениях маршрутов, которые были преобразованы из одного протокола в другой. Выходная фильтрация делает обратное — она ограничивает обратное преобразование маршрутов в исходный протокол.

Фильтрация и управление маршрутами

При фильтрации и управлении маршрутом иди набором маршрутов выполняется три основных действия.

1. Идентифицирование маршрутов.
2. Разрешение или запрещение работы с маршрутами.
3. Внесение изменений в атрибуты.

В оборудовании Cisco применяются списки разрешения доступа (access lists), списки префиксов (prefix lists) или списки разрешения доступа по заданному маршруту (as-path access lists) только для фильтрации. Также в оборудовании Cisco используется концепция карт маршрутов для фильтрации и для выполнения манипуляций с атрибутами. Карты маршрутов подробно обсуждаются в главе 11.

Идентификация маршрутов

Идентификация маршрутов — это установление критерия, по которому можно отличить один маршрут от другого. Подобный критерий может быть определен на основе префикса IP маршрута; автономной системы, сгенерировавшей маршрут; списка AS, через которые пролегает маршрут; значения атрибута маршрута и т.д. Список критериев включается в правила фильтрации, и маршрут сравнивается с первым критерием в списке. Если маршрут не соответствует первому критерию в списке, то он помечается как протиBGPечащий первому критерию в списке. Если маршрут соответствует критерию, он считается идентифицированным и не проверяется на соответствие другим критериям.

Если же маршрут при сравнении не соответствует ни одному из критериев, заданных в списке, то он отвергается.

В общем случае идентификация маршрутов проводится на основе информации сетевого уровня о доступности сети (Network Layer Reachability Information — NLRI), атрибута AS_PATH или на основе обоих этих критериев. Каждый из этих методов идентификация более подробно описан в последующих разделах.

Идентификация маршрутов на основе NLRI

Маршрут BGP можно идентифицировать с помощью блока NLRI, который представляет собой префикс и маску (см. главу 4). Для выполнения фильтрации можно задавать отдельный префикс или диапазон префиксов. Если маршрут входит в диапазон разрешенных префиксов, то он будет идентифицирован.

См. в главе 11 раздел "Идентифицирование и фильтрация маршрутов на основе NLRI"

На рис. 6.25 представлен критерий фильтрации 10.1.0.0 0.0.255.255, который действителен для диапазона маршрутов, идентифицированных префиксом 10.1.0.0 и инверсной маской 0.0.255.255. Нули в маске указывают на соответствие маршрута, в то время как единицы обозначают бит, который не анализируется при фильтрации. Таким образом, диапазон 10.1.0.0 0.0.255.255 будет идентифицировать все маршруты от **10.1.X.X.** Вместе с префиксами, приведенными на рис. 6.25, фильтр будет идентифицировать маршруты 10.1.1.0/24, 10.1.2.0/24 и 10.1.2.2/30 и отвергать маршруты 11.2.0.0/16 и 12.1.1.0/24. Фильтрация, основанная на префиксах, более детально будет рассмотрена в последующих главах.

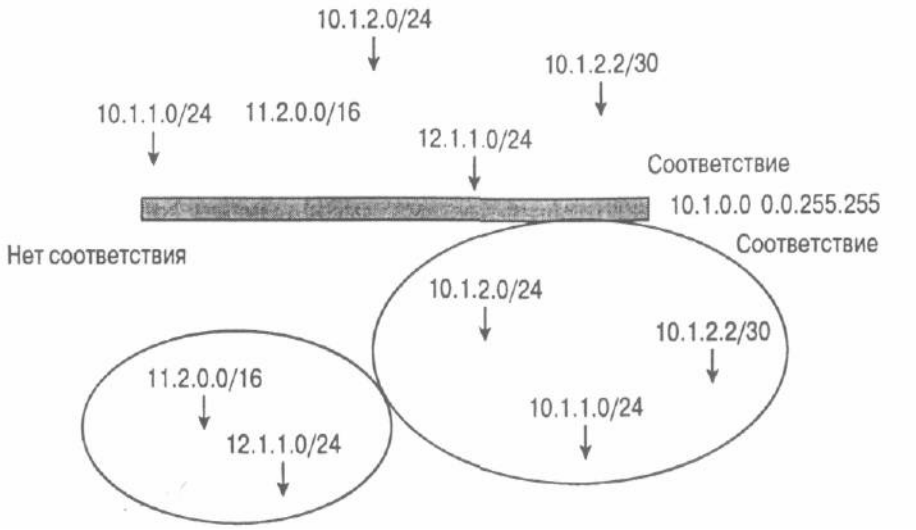


Рис. 6.25. *Пример критерия фильтрации на основе NLRI*

В настоящее время для фильтрации маршрутной информации наиболее широко используются списки префиксов. Их поддержка была включена в IOS совсем недавно, но они уже успели доказать свое превосходство над традиционными списками разрешения доступа.

Списки префиксов обеспечивают более эффективную и менее ресурсоемкую платформу для анализа, где требуется значительно меньше ресурсов системы для обработки маршрутной информации. Сегодня списки префиксов могут использоваться только для фильтрации маршрутной информации.

Одно из ключевых преимуществ списков префиксов — возможность их оперативной модификации. Другими словами, вы можете добавлять, удалять или изменять записи в списке без его полной реконструкции. Кроме того, синтаксис команд для настройки списков префиксов более удобен и интуитивно понятен администраторам, чем используемый для обычных списков разрешения доступа. В главе 11 подробно описаны основные принципы конфигурирования списков префиксов.

Идентификация маршрутов на основе атрибута AS_PATH

Идентификация маршрутов на основе информации, заключенной в атрибуте AS_PATH, немного сложнее, чем идентификация по NLRI. Как вам уже известно, атрибут AS_PATH представляет собой список номеров AS, через которые проходит маршрут на пути к заданной BGP-системе. Сам по себе список — это строка символов из следующего набора:

- Цифры от 0 до 9
- Пробелы
- Левая фигурная скобка {
- Правая фигурная скобка }
- Левая скобка (
- Правая скобка)
- Начало ввода строки
- Конец ввода строки
- Точка .

Например, список атрибута AS_PATH 10 2 в действительности состоит из символа начала ввода строки, за которым следуют символы 1 и 0, пробел и символ 2. И завершает строку символ конца ввода строки.

Ш См. в главе 11 на с. 288 раздел "Идентифицирование и фильтрация маршрутов на основе атрибута AS_PATH"

Идентификация маршрутов по списку AS_PATH заключается в его сравнении с *нормальным выражением (regular expression).* Нормальное выражение представляет собой комбинации символов, представленных формулой вида ^{Л}200 100$. Подобное нормальное

выражение представляет собой список, который начинается с 200, далее следует пробел и список заканчивается значением 100. Символы ^Л и $ обозначают, соответственно, начало и конец ввода строки.

Примечание

Нормальное выражение может быть сформировано как с использованием одиночных символов, так и с использованием групп символов.

Разрешение и запрещение маршрутов

После идентификации маршрута над ним можно провести определенные действия. Маршрут, в зависимости от заданных для соединения правил фильтрации, может быть либо разрешен, либо отвергнут. Критерии для разрешения или запрещения передачи трафика по определенному маршруту зависят от правил маршрутизации, принятых для конкретной AS. Если маршрут разрешен для передачи трафика, то он либо принимается "как есть", либо его атрибуты подвергаются модификации. Не все атрибуты маршрута могут быть модифицированы. Это зависит от атрибутов. Если маршрут оказывается запрещенным, то он просто отвергается без какой-либо последующей обработки.

Управление атрибутами

Если маршрут разрешен, то его атрибуты можно изменять в порядке влияния на процесс принятия решения маршрутизатором. В предыдущих разделах вы уже видели, каким образом такие атрибуты, как Local Preference и MED, могут быть дополнены или изменять свою длину для предпочтения одного маршрута другому. Как вы увидите позже, управление атрибутами является мощным рычагом, с помощью которого можно формировать правила маршрутизации, распределять нагрузку и влиять на симметрию маршрутов.

На рис. 6.26 представлена схема применения различных критериев к набору маршрутов с целью их фильтрации и модификации их атрибутов.

Заметим, что в каждом случае могут применяться один или несколько критериев. Маршрут может проверяться по префиксу и по атрибуту AS_PATH. Например, для того, чтобы была разрешена работа по нему, он должен соответствовать всем заданным критериям.

Отметим также, что после совпадений всех критериев для маршрута сравнение прекращается. Следовательно, имеет значение порядок проверки записей в правилах фильтрации. Например, если маршрут находится в начале списка, то запись, разрешающая все маршруты, будет перекрывать действие всех остальных записей.

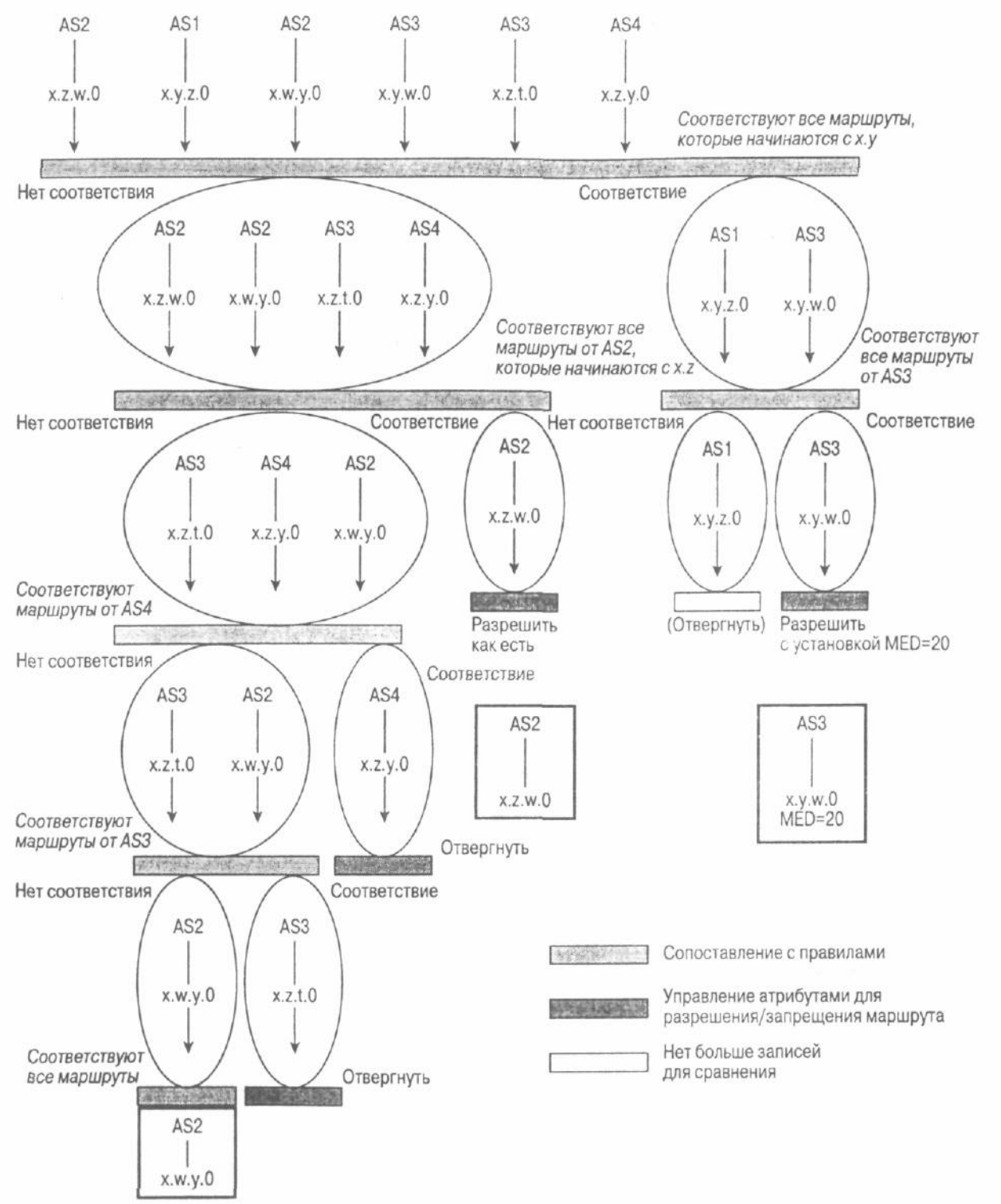


Рис. 6.26. *Пример фильтрации маршрутов и управления их атрибутами*

Односимвольные выражения

В односимвольных выражениях маршрутная информация сравнивается с одним символом. Нормальное односимвольное выражение вида 3 осуществляет сравнение с символом 3 в строке ввода. Имеется возможность задания диапазона одиночных символов, с которыми должно проводиться сравнение. Диапазоны таких символов заключаются в квадратные скобки ([]). Порядок расположения символов внутри диапазона не имеет значения. Например, нормальное выражение, состоящее из диапазона символов [efghEFGH], будет проверяться на совпадение с любым символом из заданного диапазона. Так, при задании строк, содержащих "hello" и "there", выясняется, что обе они соответствуют нормальному выражению, так как содержат символ e.

Диапазоны символов ограничиваются конечной точкой диапазона. Например, можно

задавать диапазоны [a-zl и [0-9], которые охватывают все маленькие буквы от a до z и все цифры от 0 до 9, соответственно.

Задавая знак вставки (Л) в начале диапазона, можно отменять или отрицать соответствие той или иной записи правил фильтрации. Так, например, диапазон [^a-dA-D] будет проверяться на соответствие любым символам, кроме a, Ь, c, d, A, B, C и D. Некоторые символы, такие как знак доллара $ и знак подчеркивания _, имеют специальное назначение (табл. 6.4).

Чтобы специальные символы воспринимались как часть вводимой строки, ставьте перед ними обратный слеш (\). Например, диапазон [abc\$] будет означать, что анализу подвергаются строки, содержащие символы a, Ь, c и $. В табл. 6.4 представлены специальные символы, применяемые в нормальных выражениях.

Таблица 6.4. *Специальные символы для нормальных выражений*

Название символа	Символ	Специальное значение
Точка	.	Соответствует всем символам, включая пробел
Звездочка	*	Соответствует 0 и более выражений
Знак плюс	+	Соответствует 1 или более выражений
Знак вопроса	?	Соответствует 0 или более встречающихся выражений
Знак вставки	^	Соответствует началу строки. Также используется для инвертирования выражения, если используется внутри диапазона символов — например [^диапазон]
Знак доллара	$	Соответствует концу строки
Подчеркивание	_	Соответствует точке (.), правой ({) и левой (}) фигурным скобкам, а также правой и левой обычным скобкам, символам начала и конца строки или пробелу
Квадратные скобки	[диапазон]	Определяет диапазон символов в выражении
Дефис	-	Разделяет конечные точки диапазона

Многосимвольные выражения

Многосимвольные нормальные выражения представляют собой упорядоченную последовательность односимвольных выражений. Выражение может быть комбинацией букв, цифр, других символов и символов специального назначения. Ниже приведен пример нормального многосимвольного выражения: 100 1[0-9]. Это выражение анализируется на соответствие строке, в которой имеется последовательность- 100, затем пробел, затем 1 и любая цифра от 0 до 9. Таким образом, нормальному выражению будут соответствовать любые из строк: 123 100 10 11, **100 19** или 19 **100 11** 200 и т. д.

Создание комплексных нормальных выражений

Для создания комплексных нормальных выражений широко используются специальные символы, представленные в табл. 6.4. Знаки вставки (^) и доллара ($) используются для обозначения начала и конца строки. Другие символы, такие как звездочка (*), знак плюс (+) и знак вопроса (?), позволяют повторять определенные символы внутри нормального выражения.

В приведенном примере отражается соответствие любых сочетаний буквы a, включая ее отсутствие:

a* эквивалентно любому из выражений вида: (пусто), a, aa. aaa, aaaa и т.д.

В следующем примере показано, что в строке должна присутствовать хотя бы одна буква a:

a+ эквивалентно выражениям a, aa, aaa, aaaa и т.д.

В этом примере представлен список, который может содержать, а может и не содержать букву a:

ba?b эквивалентно выражениям bb или bab.

Для повторения записи в многосимвольном выражении оно заключается в круглые скобки. Например, выражение (ab)+ эквивалентно ab или abab.

Символ подчеркивания (_) соответствует символам начала (^) и конца ($) строки, символам скобок, пробелу, квадратным скобкам, точке или самому себе. Символ точки (.) соответствует одиночному символу, включая простой пробел. На рис. 6.27, в табл. 6.5 и 6.6 показано, как символы могут сопрягаться друг с другом при создании нормального выражения.

Давайте рассмотрим топологию сети, представленную на рис. 6.27. Автономные системы AS400, AS300, AS200, AS100 и AS50 генерируют маршруты в сети A, B, C, D и E, соответственно. Маршрутизатор RTA в AS50 получает сведения обо всех маршрутах в эти сети от соседних систем — AS 100 и AS300. После инициализации процесса принятия решения в BGP маршрутизатор RTA подбирает наилучший маршрут к этим сетям согласно табл. 6.5.

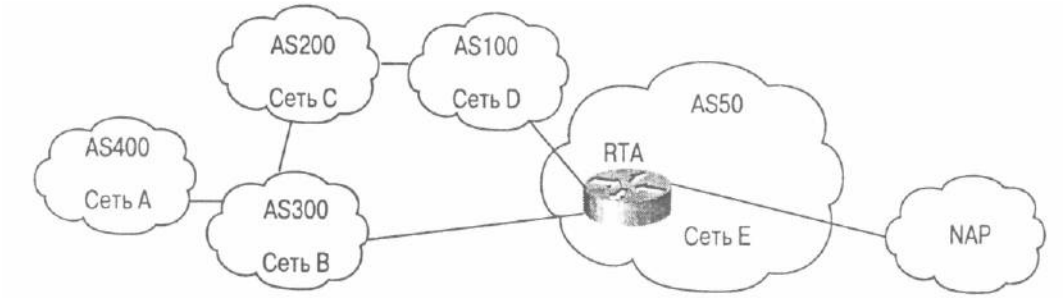


Рис. 6.27. *Пример топологии сети для создания комплексного нормального выражения*

Таблица 6.5*. Выбор наилучшего BGP-маршрута для маршрутизатора RTA*

Сеть	**Значение AS_PATH**
Сеть A	300, 400
Сеть B	300
Сеть C	100, 200
Сеть D	100
Сеть E	нет

В табл. 6.6 представлены нормальные выражения, которые можно было бы использовать при фильтрации маршрутов, объявляемых маршрутизатором RTA для NAP.

Таблица 6.6. Выражения и результаты их воздействия

Маршруты, объявляемые в NAP маршрутизатором RTA	**Выражение**	**Маршрутная информация**	**Результат воздействия**
Только локальные маршруты	^$	Нет	Сеть E
Все маршруты	.*	Все маршруты	Сеть A, Сеть B, Сеть C, Сеть D, Сеть E
Маршруты, сгенерированные непосредственно клиентами	^300^{100}	300 100	Сеть B, Сеть D
Маршруты из сетей клиентов и их клиентов	^300_^100_	300 400 300 100 200 100	Сеть A, Сеть B, Сеть C Сеть D
Маршруты, сгенерированные AS200	_200$	100 200	Сеть C
Маршруты, прошедшие через AS100	_100_	100 200 100	Сеть C, Сеть D

Выражение типа ^$ указывает на пустой список маршрутов, в котором находятся только локальные маршруты. {Вспомните о том, что мы говорили о номере локальной AS при объявлении маршрутов по EBGP. Он не включается в список, т.е. равен нулю). Символы ^ и $ определяют лишь границы пустой строки. Знак подчеркивания в выражении _200$ ограничивает номер AS, т.е. номер AS должен быть точно 200, а не 1200 или 2200.

Фильтрация на основе сведений, заключенных в атрибуте AS_PATH, является исключительно эффективной, так как она одновременно выполняется над всеми маршрутами в AS__PATH. Если бы не было подобного типа фильтрации, то пришлось бы отдельно задавать тысячи маршрутов или становиться членом уже идентифицированного сообщества BGP.

Группы узлов

Группа узлов BGP (peer group) — это труппа соседних BGP-узлов, в которых используется один набор правил маршрутизации. Вместо того чтобы определять правила для каждого отдельного узла, вы можете задать групповое имя и назначить правила группе узлов. Например, администратор, устанавливая правила маршрутизации для взаимодействующих узлов, к большинству из них может применить один и тот же набор правил, следовательно, он может задать эти правила группе узлов, что избавит его от необходимости задания одних и тех же правил каждому узлу в отдельности.

Однако группы узлов организуются не только с целью облегчения труда оператора по настройке отдельных BGP-узлов, они также предохраняют BGP-узел от постоянного уточнения правил работы с соседями. С помощью групп узлов маршрутизатор формирует обновление маршрутной информации всего один раз на основе принятых для группы правил маршрутизации и затем рассылает одинаковые сообщения об обновлении всем взаимодействующим в группе узлам.

На рис. 6.28 показан маршрутизатор RTA с тремя узлами, которые имеют те же внутренние правила маршрутизации, что и RTA. Существует еще три внешних узла, которые также имеют одинаковые правила маршрутизации. Таким образом, при конфигурации маршрутизатора RTA задается две группы узлов — одна внутри AS и одна вне AS. Каждая группа узлов содержит набор правил, который действителен в отношении RTA и определяет порядок его взаимодействия с другими узлами. Эти правила могут представлять собой фильтры на основе набора префиксов IP или фильтры на основе атрибута AS_PATH. Возможны и другие манипуляции с атрибутами. После определения групп узлов все эти правила применяются к узлам, составляющим группы.

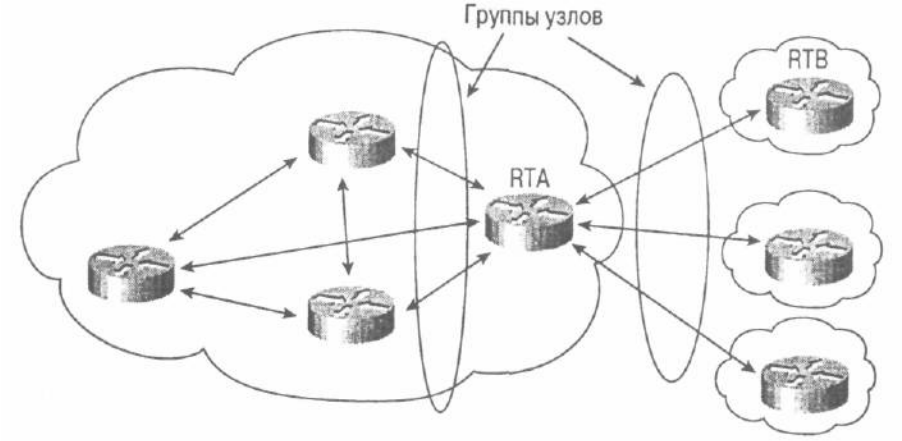


***Рис. 6.28**. Организация групп узлов*

Одно время в Cisco IOS имелись ограничения по работе соседних EBGP-узлов в составе группы узлов. Однако в последующих модификациях эти ограничения были устранены. Так что теперь мы не будем останавливаться на них. Но следует помнить о том, что, если вы пользуетесь старыми версиями IOS, то можете встретиться с подобными ограничениями. Обратитесь к документации на имеющуюся у вас версию Cisco IOS, чтобы прояснить эти вопросы.

Исключение из группы узлов

Исключение узла или узлов из группы происходит, когда один или несколько узлов используют правила маршрутизации, отличающиеся от правил, заданных для группы. В дополнение к набору правил, действующих внутри группы, можно добавлять другие правила для отдельных узлов. Рассмотрим случай, представленный на рис. 6.28, когда маршрутизатору RTA требуется задать набор правил для взаимодействия с маршрутизатором RTB. На маршрутизаторе RTA можно использовать дополнительные фильтры в направлении RTB, сохраняя при этом маршрутизатор RTB в группе внешних узлов.

См. в главе 11 раздел Труппы взаимодействующих узлов"

Агрегация в BGP-4

Одним из основных улучшений, внесенных в протокол BGP-4, является возможность работы бесклассовой междоменной маршрутизации (Classless Interdomain Routing — CIDR) и поддержка суперсетей. Принципы работы CIDR и суперсетей обсуждались нами в главе 3, "IP-адресация и методы распределения адресов".

Процесс агрегации проводится над маршрутами, которые имеются в маршрутной таблице BGP. Это протиBGPечит действию команды network, применяемой к маршрутам в таблице маршрутов IP. Агрегация может выполняться, если в маршрутной BGP-таблице имеется хотя бы один однозначно определенный маршрут.

Компанией Cisco Systems предлагается несколько различных вариантов манипулирования объединенными маршрутами для удовлетBGPения нужд сети Internet. В этом разделе мы рассмотрим несколько простых приемов объединения маршрутов и затем перейдем к более сложным, но и более интересным вариантам объединения маршрутов.

Простое объединение маршрутов с подавлением однозначно определенных маршрутов

При таком способе объединения маршрутов объявляется лишь сам объединенный маршрут, а все составляющие его определенные маршруты подавляются. Обычно это выполняется лишь в случаях, когда однозначно определенные маршруты не несут полезной информации для принятия решения о более эффективной пересылке трафика.

См. в главе 11 раздел "Только объединенные маршруты, подавление однозначно определенных маршрутов"

На рис. 6.29 показана ситуация, когда все обновления маршрутов объединены в один простой маршрут. Предположим, что в AS 100 имеется набор сетей в диапазоне от 172.16.0.0/24 до 172.16.15.0/24. Сюда входят сети 172.16.0.X, 172.16.1.X и т.д. Список определенных префиксов можно объединить в диапазон 172.16.0.0/20. Таким образом, другие узлы оповещаются об объединенном маршруте 172.16.0.0/20, а сведения об остальных определенных префиксах игнорируются.

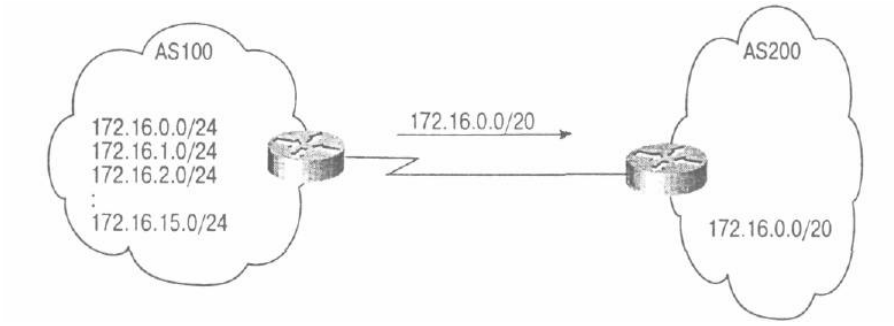


Рис. 6.29. *Пример объединения маршрутов в BGP-4 с подавлением определенных маршрутов*

Объединение маршрутов с однозначно определенными маршрутами

Существует множество ситуаций, в которых AS рассылает информацию об объединенном маршруте и однозначно определенные маршруты. Обычно это происходит в тех случаях, когда клиент подключается к одному провайдеру по нескольким каналам. Тогда провайдер может использовать информацию об однозначно определенных маршрутах для принятия оптимального решения о пересылке трафика в сеть клиента. В то же время провайдер может распространять сведения об объединенном маршруте только в направлении NAP, чтобы минимизировать общее число маршрутов в сети Internet в целом. Подобная ситуация проиллюстрирована на рис. 6.30.

Здесь AS100 подключена через каналы SF и NY к провайдерской системе AS200. Система AS100 может послать на AS200 только объединенный маршрут 172.16.0.0/20 или послать вместе с ним однозначно определенные маршруты. Если объединенный маршрут посылается только по каналам SF и NY, то и весь трафик от AS 100 в AS200 всегда будет направляться по одному из этих каналов. Таким образом, создается несбалансированная нагрузка. (Вопросы распределения нагрузки будут обсуждаться далее в главе 7, "Избыточность, симметрия и распределение нагрузки"). Для распределения нагрузки AS100 посылает сведения не только об объединенном маршруте, но и о других однозначно определенных маршрутах. По каждому каналу для различных маршрутов могут пересылаться различные метрики. На основе номера сети в AS200 принимается решение об использовании канала SF или NY для связи с AS100.

См. в главе 11 раздел "Объединенные и однозначно определенные маршруты"

Чтобы избежать усложнения маршрутных таблиц вне сети провайдера, все одно значно определенные маршруты от клиентских сетей обычно замыкаются на уровн! сети провайдера. Так, AS200 в направлении NAP будет лишь распространять сведена об объединенном маршруте 172.16.0.0/20, подавляя при этом рассылку однозначж определенных маршрутов.

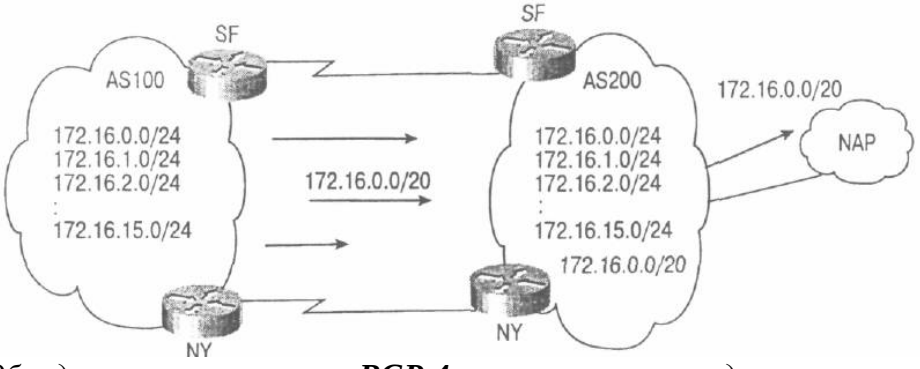


Рис. 6.30. *Объединение маршрутов в* ***BGP-4*** *с использованием однозначно определенных маршрутов*

Как правило, провайдеры стремятся минимизировать мероприятия по настройке и администрированию. В этом случае для того, чтобы остановить распространение однозначно определенных маршрутов в направлении NAP, можно использовать так называемый динамический подход. Он заключается в том, что все однозначно определенные маршруты при выделении из объединенного маршрута помечаются в AS 100 атрибутом сообщества **NO_EXPORT.** Эта процедура представлена на рис. 6.31.

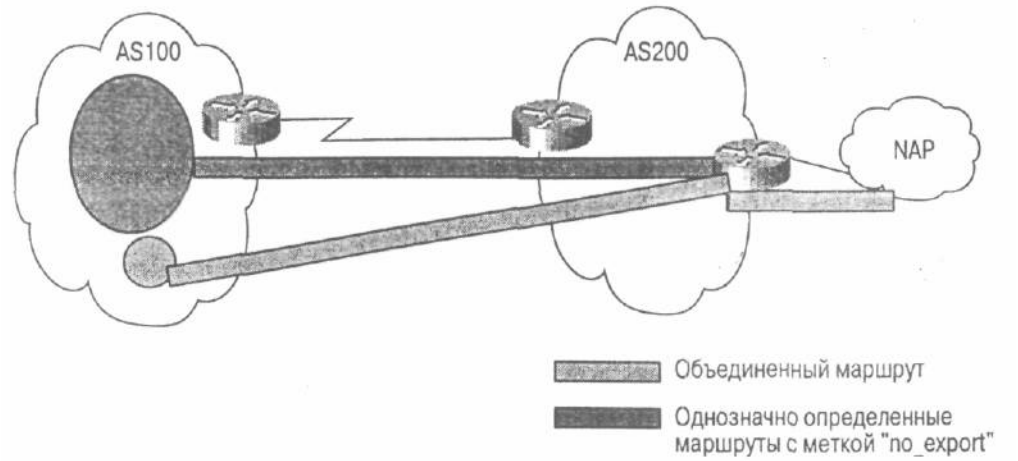


Рис. 6.31. *Пример объединения маршрутов в сообщество с пометкой* ***NO_EXPORT***

Итак, когда AS200 получает обновление маршрутной информации от AS 100, она обнаруживает, что однозначно определенные маршруты последней объединены в сообщество с требованием не пересылать сведения о них внешним узлам. Тогда как обычно в NAP и другим внешним узлам будет рассылаться информация только об объединенном маршруте.

Объединение маршрутов с использованием набора однозначно определенных маршрутов

В некоторых случаях, кроме объединенного маршрута, необходимо объявлять и однозначно определенные маршруты. На рис. 6.32 показана ситуация, когда это необходимо.

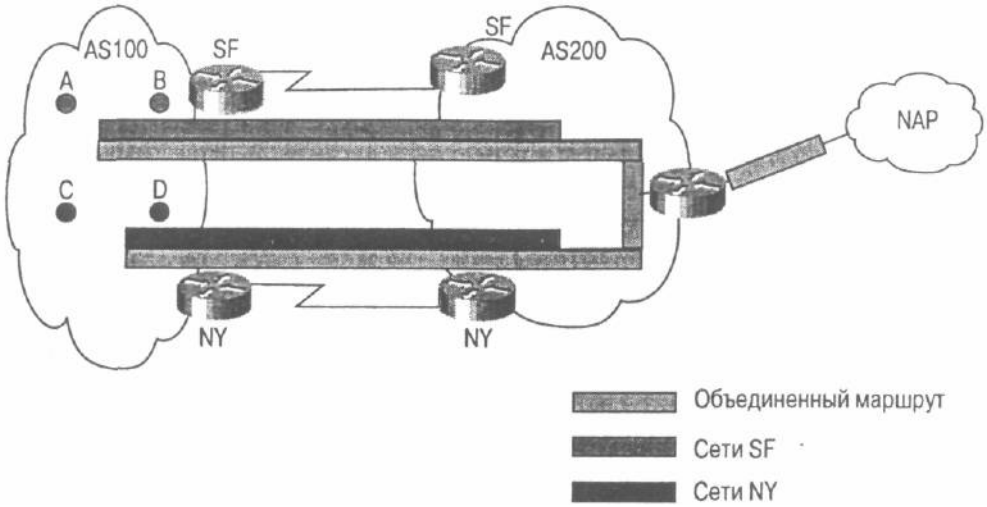


Рис. 6.32. *Пример объединения маршрутов с включением набора однозначно определенных маршрутов*

На рис. 6.32 показана AS100, которая подключена к AS200 по двум каналам. Желательно, чтобы сети, входящие в AS100, которые расположены ближе к каналу SF, вели обмен данными именно через него, а сети, которые территориально ближе к каналу NY, работали бы с каналом NY. Этого можно достичь следующим образом.

- Объявить по каналу SF только объединенный маршрут и маршруты в сети, расположенные вблизи этого канала.
- Объявить по каналу NY только объединенный маршрут и маршруты в сети, расположенные вблизи канала NY.

Таким образом, AS200 будет направлять трафик в сети, расположенные вблизи SF, по каналу SF, а трафик для сетей вблизи NY --по каналу NY. Сети, территориально расположенные в других регионах, могут вести обмен трафиком по обоим каналам либо по одному из них. При выходе из строя одного из каналов связи доступ к сетям можно обеспечить по второму каналу посредством объединенного маршрута, который объявляется по обоим каналам. Для объявления только объединенного маршрута в сторону NAP можно использовать метод пометки маршрутов с помощью NO_EXPORT, который мы обсуждали в предыдущем разделе.

См. в главе 11 на раздел "Объединение маршрутов с использованием набора однозначно определенных маршрутов"

Потери информации внутри объединенного маршрута

Объединение маршрутов приводит к потерям информации, так как атрибуты отдельных маршрутов при объединении теряются. Как мы уже отмечали, протокол BGP определяет набор AS_SET, который представляет собой математический ряд, состоящий из элементов, где содержатся сведения обо всех маршрутах, участвующих в объединении. Примерами таких элементов могут служить атрибуты AS_PATH и COMMUNITY. Применение AS_SET совместно с объединенным маршрутом делают его нестабильным, поскольку изменение атрибутов отдельных маршрутов, входящих в объединенный маршрут, приводит и к изменению свойств самого объединенного маршрута, что может послужить причиной постоянного удаления этого маршрута из таблиц (как нестабильного) или к частым его обновлениям.

См. в главе 11 раздел "Потери информации в объединенном маршруте"

Изменение атрибутов объединенного маршрута

В определенных ситуациях требуется изменить атрибуты объединенного маршрута. Один из таких случаев — наличие в объединенном маршруте нежелательных атрибутов, которые были унаследованы от отдельных маршрутов в процессе их объединения (в случае применения AS_SET). Примером может служить атрибут сообщества NO_EXPORT, унаследованный объединенным маршрутом от одного из однозначно определенных маршрутов. Он может стать причиной того, что объединенный маршрут не будет экспортирован другим AS. Еще одна ситуация, требующая изменения атрибутов объединенного маршрута, -- отражение уровня предпочтения этого объединенного маршрута. Например, необходимость в этом может возникнуть, когда клиент объявляет объединенный маршрут провайдеру по нескольким каналам. Тогда клиент может установить для объединенного маршрута различные значения атрибута MED для различных каналов, чтобы повлиять на выбор точки входа в AS. Компания Cisco разработала методику, позволяющую пользователю модифицировать соответствующим образом атрибуты объединенного маршрута.

Формирование объединенного маршрута на основе набора однозначно определенных маршрутов

Итак, вы уже увидели, что объединенный маршрут с выражением AS_SET содержит набор всех атрибутов (включая и номера AS), которые заданы для отдельных маршрутов, участвующих в объединении. Если в объединении принимают участие маршруты из различных AS, то необходимо указывать, на основе каких именно маршрутов сформирован тот или иной объединенный маршрут. Эта информация используется центральными и вспомогательными AS, когда в каждое ответвление AS ведут отдельные объединенные маршруты от центральной AS. При формировании объединенного маршрута центральная AS будет исключать однозначно определенные маршруты, ведущие в ответвленные AS, и передавать в эти AS лишь объединенный маршрут. Объединенный маршрут, полученный ответвленной AS, уже не содержит номера самой ответвленной AS и поэтому не будет отвергаться (как угрожающий образовать петлю). На рис. 6.33 приведен пример формирования объединенного маршрута подобным образом.

В роли центральной AS выступает AS3, которая принимает информацию о маршрутах 192.68.11.0/24 и 192.68.10.0/24 от вспомогательных (или ответвленных) ASI и AS2. В префиксе 192.68.11.0/24 AS_PATH равен 1, а в 192.68.10.0/24 AS_PATH имеет значение 2. Когда на основе всех однозначно определенных маршрутов в AS3 формируется AS_SET, то информация в AS_PATH будет {1 2}. Однако сам по себе объединенный маршрут при пересылке на AS1 или AS2 будет отвергаться во избежание формирования петли. В AS1 при анализе атрибута AS_PATH будет обнаружен ее собственный номер, и обновление маршрута будет игнорировано. То же самое происходит и в AS2. Если вы можете указать, какие именно маршруты формируют объединенный маршрут, то можно, например, определить, что объединенный маршрут формируется только на основе 192.68.11.0/24. Таким образом, в AS_PATH будет только информация о AS1 и ни слова об AS2. Вследствие такой манипуляции объединенный маршрут может теперь безболезненно передаваться обратно на AS2. В AS2 он может использоваться для передачи трафика на все узлы AS1.

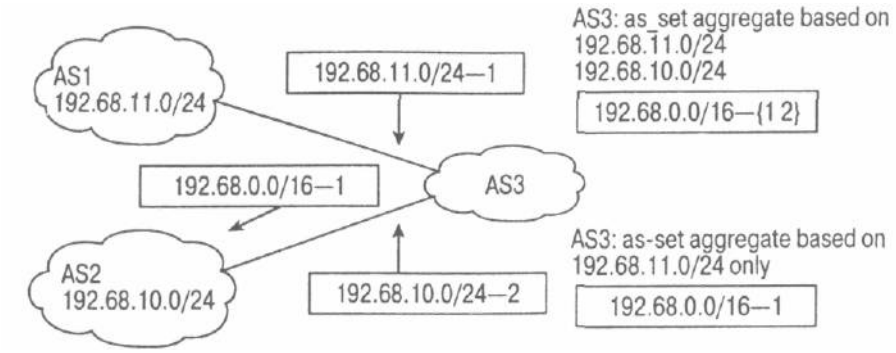


Рис. 6.33. *Формирование объединенного маршрута на основе набора однозначно определенных маршрутов*

См. в главе 11 раздел "Формирование объединенного маршрута на основе наборов однозначно определенных маршрутов"

Забегая вперед

Теперь, обучившись азам работы с протоколами маршрутизации и зная о свойствах атрибутов BGP, вы можете приступить к применению этих инструментов на практике в сетях

с различной топологией. Таким образом, вы убиваете нескольких зайцев — вы сразу можете перейти к изучению резервирования, обеспечения симметрии и распределения нагрузки, значение которых в зависимости от требований к сети может варьироваться. Иногда даже эти требования могут конфликтовать друг с другом. Значение этих задач при разработке схемы сети более детально будет обсуждаться в следующей главе. Атрибуты, описанные нами в этой главе, активно используются в следующей главе для построения системы маршрутизации.

Настройка параметров протокола BGP для удовлетBGPения требований конкретной сети включает как настройку параметров для работы внутри AS, так и вне AS. Другими словами, правила работы устанавливаются еще и для сетей, с которыми взаимодействует ваша сеть. Хотя это и не входит в сферу вашего влияния, но они напрямую связаны с тем, как настроен BGP с вашей стороны. Все эти вопросы мы постараемся осветить в следующей главе, где рассматривается несколько вариантов резервирования, обеспечения симметрии и распределения нагрузки в сетевых архитектурах, которые на сегодняшний день наиболее широко применяются в сети Internet.

Часто задаваемые вопросы

В — *Если взаимодействующие узлы в моей сети, по протоколу IBGP не соединены друг с другом непосредственно, должен ли я использовать мультиузловой EBGP?*

О — Нет, Не существует каких-либо ограничений по связности iBGP-узлов. Мультиузловой EBGP используется только для протокола EBGP.

В — *Требуется ли для обеспечения синхронизации вкладывать BGP- маршруты вIGP?*

О — Нет. Вложение BGP-маршрутов в IGP не рекомендуется. Желательно вообще не выполнять синхронизацию с BGP. Убедитесь, что подобные настройки не приведут к проблемам доступности сетей внутри AS.

В— Обеспечит ли большую стабильность BGP-маршрутам вывод их через IGP-маршруты с помощью команды network, по сравнению с преобразованием IGP-маршрутов в BGP?

О — Нет, В обоих вышеупомянутых методах флуктуации IGP-маршрутов будут перенесены на BGP-маршруты. Команда **network** просто дает лучшие возможности по управлению маршрутами и позволяет не волноваться о том, что IGP-маршруты могут быть преобразованы в BCP, ёсли вы используете преобразование. Комбинируйте ее со статическими маршрутами (с дистанцией, например, 254) в Null0,-если нужно предотвратить флуктуацию маршрутов даже при нестабильной маршрутизации по IGP.

В — *Должен ли я составлять список всех подключенных интерфейсов и объявлять их с помощью команды* **network**?

О — Если нужно обеспечить доступ по BGP в непосредственно подключённые к вашему узлу подсети, то вы можете их объявить с помощью этой команды. В противном случае это не обязательно.

В - *У меня есть два граничных маршрутизатора, которые связаны с провайдером по EBGP, а между собой по lBGP. Если я буду указывать IGP-маршруты с помощью команды* **network** на *обоих маршрутизаторах, приведет ли это к образованию петли?*

О— Нет, таким образом вы не создадите петлю. В принципе; реализуя подобную схему, вы обеспечиваете резервирование в своей сети. Если один из граничных маршрутизаторов выйдет из строя, то второй будет анонсировать те же самые сети.

В — *Мне нужно получать всего несколько маршрутов от соседнего узла. Могу ли я фильтровать канал со своей стороны?*

О — Да, это можно сделать. Однако сначала следует попросить администратора соседнего узла посылать лишь нужные вам маршруты для того, чтобы минимизировать нежелательное использование полосы пропускания и избежать флуктуации маршрутов. В то же время фильтрация только с вашей стороны поможет избежать случаев, когда соседи

посылают вам больше маршрутов, чем вы ожидали.;

В — *Мой провайдер требует, чтобы я установил различные локальные предпочтения для разных соединений. Возможно ли это?*

О — Нет. Локальные предпочтения определяются только внутри AS, и информация о них не передается во время сеансов EBGP. Однако более оптимальное решение в этом случае — попросить провайдера сконфигурировать карту маршрутов на его узле, в которой проводилось бы сопоставление строк для сообщества BGP (подобно описанному в RFC 1998) в течение сеанса с вашим узлом. При этом в его AS устанавливались бы локальные предпочтения для ваших маршрутов. Вы должны будете померить используемые вами маршруты соответствующей строкой, обозначающей их принадлежность к определенному сообществу, чтобы правильно устанавливалось значение локального предпочтения в течение сеанса работы с вашим узлом. Для .этой цели вы можете также использовать атрибут MED.

В — *Получаемый от провайдера атрибут MED конфликтует с моим IGP и влияет на обмен трафиком. Что мне делать в этом случае?*

О — Если получение MED от провайдера создает проблемы, то попросите его не пересылать вам MED. Вы также можете самостоятельно установить MED в 0 со своей стороны.

В — *Я подключен к нескольким провайдерам. Иногда в мою AS поступает огромное количество трафика, не принадлежащего моей.AS. В чем может быть причина?* ;

О— Возможно, вы объявляете маршруты, которое получаете от одного провайдера другим провайдерам. В этом случае другие AS могут использоватъ вашу AS как транзитную. Убедитесь в том, что вы объявляете провайдерам только внутренние маршруты.

В — *У меня имеется* несколько *соединений с одним провайдером. Следует ли мне беспокоиться объявлении маршрутов, полученных по одному каналу, во второй канал?*

О — Предположительно правила работы по BGF на стороне вашего провайдера позволяют обнаруживать маршруты, которые уже были получены от вас его AS, и игнорировать их. Однако это не очень хорошая практика. Поступая таким образом, вы дополнительно нагружаете процессор и перегружаете канал бесполезной информацией. Если это возможно, убедитесь в том, что вы пересылаете сведения только о собственных маршрутах.

В — *Я как провайдер выдал одному из своих клиентов частный номер AS. Теперь клиент хочет организовать еще одно соединение, но уже с другим провайдером. Что произойдет если рн по-прежнему будет использовать выданный мною частный номер AS?*

О – Хотя сегодня в сети internet подобные случаи не редкость, это считается очень нежелательной конфгурацией. После того, как вы объявите сеть клиента в Internet, *вы* будете вырезать номер частной AS и анонсировать эти маршруты так, как будто они были сгенерированы вашей AS. Если второй провайдер выполняет те же самые действия, то сети вашего клиента будут объявляться двумя AS с уникальными.номерами, что может быть причиной образования петель маршрутизации. Кроме того, если перёд анонсированием адреса клиента другим сетям выполняется объединение маршрутов, то это может привести к тому, что однозначно определенные маршруты будут доступны через одного из двух провайдеров, т.е. становится невозможным равномерное распределение нагрузки между двумя каналами. Если клиент по каким-либо причинам не может получить глобальный номер AS от своего регионального реестра сети Internet, то он должен убедиться в том, что оба его провайдера сконфигурировали соединение с ним правильно и обеспечили совместное распределение нагрузки, а также организацию безотказной работы при данной конфигурации.

В — *Я подключен к одному провайдеру в Сан-Франциско и объявляю свои маршруты по протоколу BGP. Кроме того, я подключен к другому провайдеру* в *Лос-Анджелесе. Следует ли мне получить разные номера для этих AS?*

О – Если сети в Сан-Франциско и в Лос-Анджелесе подчинены одной администрации и работают с другими AS по одним и тем же правилам, то они должны быть объединены в одну AS. Помните, что разделение сетей в BGP означает разделение административной ответственности и правил. Определяющим фактором для этого должна быть топология сети.

Ссылки

[1] RFC 1998, "Ал AppUcation of the BGP Community Attribute in Multi-home Routing," www.isi.edu/in-notes/rfcl998.txt

[2] RFC 2270, "Using a Dedicated AS for Sites Homed to a single Provider," www.isi.edu/in-notes/rfc2270.txt

Ключевые темы этой главы:

- **Избыточность.** Обеспечение стабильности сети засчет построения альтенативных маршрутов (маршрутов по умолчанию), по которым в случае отказа основного маршрута будет передаваться весь трафик — одна из важнейших задач при создании сети.
- **Определение маршрутов по умолчанию.** Конфигурирование маршрутов по умолчанию является фундаментом для создания надёжных сетевых соединений с резервированием. Однако когда имеется несколько маршрут по умолчанию, возникает необходимость определения их приоритетности.
- **Симметрия.** Настройка маршрутов таким образом, чтобы определенный трафик поступал в AS и покидал ее в одной и той же точкё, — одна из основных задач при создании системы маршрутизации.
- **Распределение нагрузки.** Распределение трафика между несколькдаи каналами' для оптимизации нагрузки производительности сети.
- **Примеры.** Рассмотрено несколько примеров сетей с использованием избыточности, симметрии и распределения нагрузки. Предложены также различные варианты настройки атрибутов для выполнения вышеуказанных требований кструктуре сети.

Глава 7. Избыточность, симметрия и распределение нагрузки

Избыточность

Несмотря на то что провайдеры и их клиенты предпочитают работать по своим каналам непрерывно, время от времени по различным причинам происходят отказы. Однако проблема обеспечения непрерывности связи не может быть решена в одностороннем порядке. Она требует совместных усилий от всех звеньев цепи, обеспечивающей подключение организации к сети. Так, соединение маршрутизатора с сетькК Internet организуется с привлечением устройств сопряжения с каналом связи (DSU/CSU), системы электропитания, соединительных кабелей, физической линии, по которой осуществляется привязка к провайдеру, и огромного количества администраторов, каждый из которых отвечает за различные участки соединения. В любой момент человеческая, программная или физическая ошибка или другие непредвиденные обстоятельства (такие как стихийные бедствия или аварии линий электропередачи) могут повлиять на соединение с провайдером.

Принимая во внимание все эти причины, обеспечение избыточности в сети более чем желательно. Однако критическим в этой ситуации является нахождение равновесия между избыточностью и симметрией. Задачи обеспечения избыточности и симметрии могут довольно часто вступать в конфликт. Чем большей избыточностью характеризуется сеть, тем более непредсказуемыми являются точки входа и выхода трафика из сети. Если у клиента имеется несколько соединений с провайдером — одно в точке присутствия (Point Of Presence -- POP) в Сан-Франциско, и другое в POP в Нью-Йорке, то трафик, вышедший из Сан-Франциско может вернуться в Нью-Йорк. Добавление еше одного соединения с POP в Далласе повышает общую надежность подключения компании к сети Internet, но в то же время усложняет задачу обеспечения симметрии. Сетевые администраторы должны учитывать эти нюансы при разработке правил маршрутизации для своих сетей.

Давление географических ограничений

При создании сетей и реализации избыточности, кроме надежности, должен учитываться еще и географический фактор. Большинство компаний является одновременно и национальными, и международными. С их точки зрения, автономная система представляет собой совокупность объединенных физически географических точек. Корпорация со своей AS, охватывающей несколько географических точек, может воспользоваться услугами как одного, так и нескольких провайдеров в разных регионах. На рис. 7.1 представлен офис в Сан-Франциско, обслуживаемый ASK который подключен к POP провайдера JSP1 в Сан-

Франциско, и офис в Нью-Йорке, подключенный к POP провайдера ISP2 в Нью-Йорке. При таком построении сети трафик может проходить по более короткому маршруту путем прохождения через ближайшую прилегающую POP.

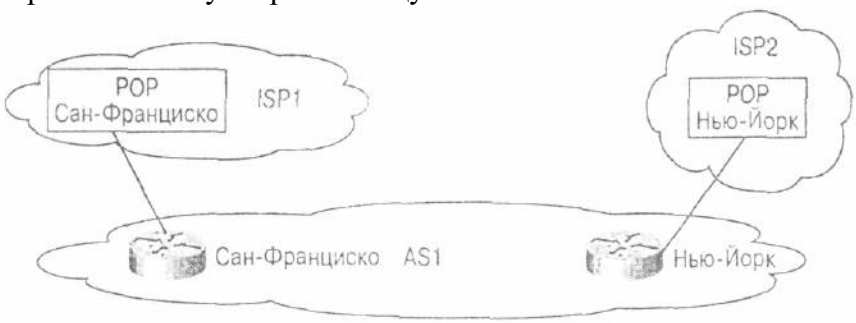


Рис. 7.1. *Географически разнесенное подключение организации к Internet по нескольким каналам*

Так как избыточность зиждется на существовании альтернативных маршрутов в сеть и из нее, то необходимо обеспечить хранение дополнительной маршрутной информации в таблицах маршрутов. Чтобы избежать ненужной нагрузки на систему маршрутизации, практичнее воспользоваться механизмами маршрутизации по умолчанию. Маршрутизация по умолчанию позволяет организовать резервные маршруты, которыми можно воспользоваться в случае выхода из строя основных маршрутов. В следующем разделе делается попытка осветить различные аспекты применения маршрутизации по умолчанию для простых систем маршрутизации.

Определение маршрутов по умолчанию

Организация маршрутов по умолчанию — мощнейший инструмент для минимизации количества маршрутов, которые должен изучить маршрутизатор, и для обеспечения избыточности в сетях на случай аварий и отказов каналов связи. В терминологии компании Cisco, маршрут по умолчанию называется *шлюзом по умолчанию* или *резервным шлюзом (gateway of last resort)*. Очень важно, чтобы вы уяснили, как осуществляется маршрутизация по умолчанию, и помнили, что она облегчает жизнь администратора только в том случае, если сконфигурирована правильно.

По определению маршрут по умолчанию — это маршрут в таблице пересылки IP, который используется в случае когда нет сведений о маршруте в заданный пункт назначения. Другими словами, маршрут по умолчанию представляет собой крайний случай, к которому обращаются, если нет никакой информации о маршруте в заданный пункт назначения.

Маршруты по умолчанию, распространяемые динамически

Универсальный общеизвестный маршрут по умолчанию обычно представляется комбинацией сети и маски подсети вида 0.0.0.0/0.0.0.0 (или сокращенно 0/0). Этим маршрутом маршрутизаторы могут обмениваться во время динамических обновлений маршрутной информации. Система, объявляющая этот маршрут, подразумевает, что сама она является по умолчанию для других систем. На рис. 7.7 показан пример объявления

такого маршрута.

Рис. 7.2. *Динамическое объявление маршрута по умолчанию*

Динамические маршруты по умолчанию (0/0) могут распространяться с помощью протокола BGP или IGP, в зависимости от того, какой из протоколов используется двумя доменами маршрутизации. С целью обеспечения избыточности и исключения возможности отказа следует получать маршруты по умолчанию от разных источников.

В контексте протокола BGP для маршрута по умолчанию можно установить локальное предпочтение, что позволяет устанавливать приоритет для маршрутов по умолчанию и интерпретировать их как основной и резервный. В таком случае при выходе из строя одного из маршрутов вместо него используется другой.

В левой части рис. 7.2 с помощью одного маршрутизатора по двум каналам организуется соединение AS1 и AS2. Если в AS1 принимается несколько возможных маршрутов от AS2, то в ней можно принимать только один маршрут по умолчанию 0/0. В приведенном примере AS1 получает сведения о маршруте 0/0 по двум каналам и присваивает каждому маршруту степень предпочтения, задавая для основного канала значение локального предпочтения 100 и 50 (или другое значение меньше 100) для резервного. При нормальной работе этой схемы для AS1 будет установлен шлюз по умолчанию 1.1.1.1.

Эти же процедуры можно проделать и при работе по нескольким маршрутизаторам (см. правую часть рис. 7.2), если в AS поддерживается протокол IBGP. Основной и резервный каналы определяются в этом случае с помощью величин локальных предпочтений, которыми обмениваются маршрутизаторы по IBGP.

См. в главе **12** раздел "Динамические маршруты по умолчанию"

Статические маршруты по умолчанию

Многие операторы прибегают к фильтрации получаемых динамически маршрутов по умолчанию, чтобы избежать ситуации, когда трафик приходит в ту точку сети, куда он не должен был попасть. Существует определенный механизм задания собственных маршрутов по умолчанию в отдельной AS статически путем установления собственного маршрута 0/0. Статически задаваемые маршруты по умолчанию дают администратору возможность полного контроля над поведением маршрутов, так как в его власти определить шлюз по умолчанию, а не получать эти сведения от другой стороны.

См. в главе 12 раздел "Статические маршруты по умолчанию"

Оператор может установить статический маршрут по умолчанию 0/0 в следующих случаях.

- Для указания IP-адреса следующего ближайшего шлюза.
- Для указания интерфейса маршрутизатора.

- Для указания сетевого адреса.

На рис. 7.3 представлены первые два пункта. Маршрутизатор слева указывает статически собственный маршрут по умолчанию 0/0 в направлении IP-адреса 1.1.1.1. Тот же маршрутизатор в правой части рисунка указывает маршрут по умолчанию 0/0 в направлении своего интерфейса Ethernet. В последнем из приведенных случаев дальнейшая обработка нужна для того, чтобы выяснить, кому следует посылать заданный сегмент трафика. Подобная процедура обычно выполняется при отправке пакетов протокола разрешения адреса (Address Resolution Protocol — ARP)[1], с помощью которых проводится идентификация физического адреса следующего ближайшего маршрутизатора.

В системе можно также устанавливать маршрут по умолчанию на основе сетевого адреса, полученного от другой системы. На рис. 7.4 AS1 получает от AS2 динамически маршрут по умолчанию 192.213.0.0/16. Если AS1 укажет для своей сети маршрут по умолчанию 192.213.0.0/16, то эта сеть автоматически становится резервным шлюзом. При таком подходе используется рекурсивный просмотр возможных маршрутов с целью определения IP-адреса следующего ближайшего шлюза. В этом примере в ходе рекурсивного просмотра маршрутов определяется, что маршрут 192.213.0.0/16 был получен от ближайшего соседнего узла 1.1.1.1 и весь трафик будет направлен соответствующим образом.

Для маршрутов по умолчанию очень важно динамически самоустраняться, если объекты в сети, на которые они указывают, становятся недостижимыми. В оборудовании Cisco предусмотрена поддержка маршрута по умолчанию с учетом существования точки сети, на которую он указывает. Например, если маршрут по умолчанию указывает на сетевой адрес и эта сеть уже недостижима (она отсутствует в таблице IP-маршрутов), то маршрут по умолчанию также будет исключен из таблицы IP-маршрутов. Подобные действия необходимы в тех случаях, когда имеется несколько маршрутов по умолчанию. Один из них может использоваться как основной, а другие как резервные (работа по ним может проводиться, когда основной маршрут становится недоступным).

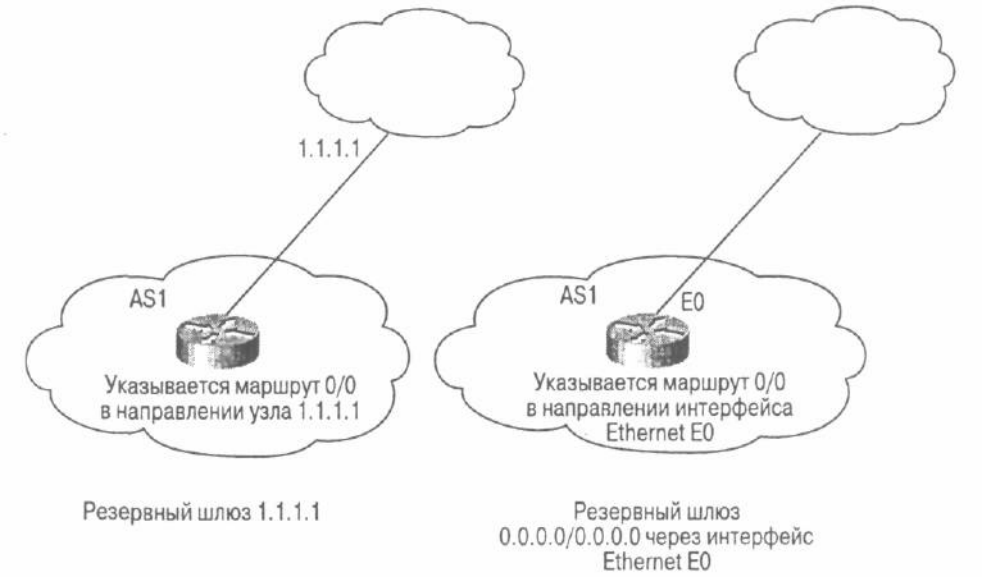


Рис. 7.3. *Статические маршруты по умолчанию*

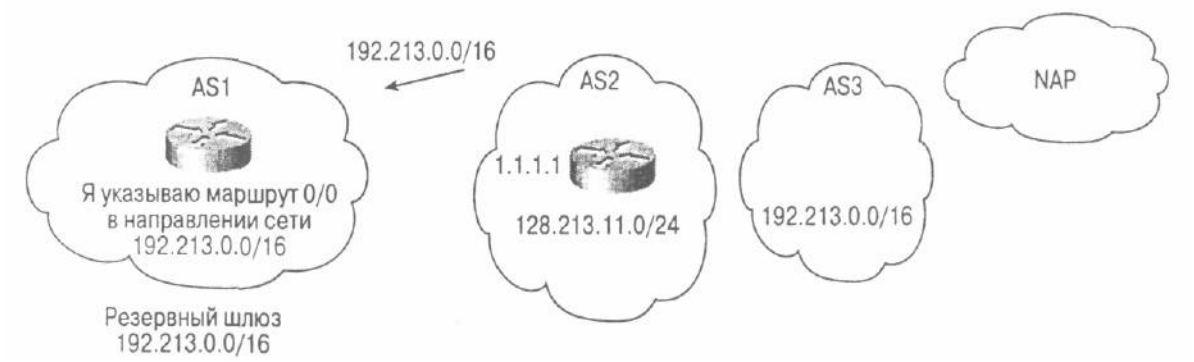


Рис. 7.4. *Указание в качестве маршрута по умолчанию сетевого адреса*

Сети по умолчанию следует выбирать исходя из их близости к сети Internet. Тогда повышается возможность того, что они имеют непосредственное соединение с NAP или с сетями других сервис-провайдеров. Очень важно, когда AS, к которой подключается ваша сеть, имеет непосредственное соединение с NAP. На рис. 7.4 показана ситуация, когда в ASI устанавливается маршрут по умолчанию в сеть провайдера (AS2) путем указания префикса 128.213.11.0/24 или суперсети 192.213.0.0/16. Указание в качестве маршрута по умолчанию маршрута в сеть 128.213.11.0/24 устанавливает зависимость от стабильности лишь части связи (от AS1 до AS2), а не всей связи (от AS1 до AS3) в направлении NAP. Если соединение между AS2 и AS3 по какой-либо причине пропадает, то AS1 по прежнему будет пересылать трафик в направлении AS2, а не направлять его по другим маршрутам по умолчанию (предполагая, что AS1 подключена еще к каким-либо провайдерам). Наилучшим выбором в качестве маршрута по умолчанию является маршрут в суперсеть 192.213.0.0/24, так как его существование более типично для всего соединения п направлении NAP и не зависит от отдельных промежуточных соединений.

Сети, выбираемые в качестве сетей по умолчанию, не должны оыть подсетями какой-либо сети. Подсеть, которая то появляется, то исчезает, вызовет флуктуацию маршрута по умолчанию -- он будет периодически возникать и пропадать. Наилучшим решением является выбор в качестве маршрута по умолчанию основного объединенного маршрута или суперсети, которая отражает стабильность работы всей сети провайдера, а не отдельного соединения.

Одновременно можно использовать несколько статических маршрутов по умолчанию. Один из методов установки нескольких маршрутов по умолчанию заключается в указании нескольких сетей (если возможно, с использованием объединенных маршрутов для обеспечения более высокой стабильности) и установлении степеней предпочтения для них с использованием атрибута локального предпочтения из протокола BGP. Все это действительно как для одного маршрутизатора, подключенного к провайдеру по нескольким каналам, так и для нескольких маршрутизаторов, взаимодействующих внутри AS no 1BGP. Оба эти варианта представлены на рис. 7.5. Единственное их отличие от того, что представлено на рис. 7.4, состоит в том, что в последнем случае клиент сам устанавливает маршрут по умолчанию, а не ожидает получения от провайдера маршрута по умолчанию 0/0. В этом примере клиент выбирает маршрут в сеть 128.213.0.0/16 с локальным предпочтением 100 в качестве основного канала для передачи трафика. Второй канал с локальным предпочтением 50 является резервным маршрутом по умолчанию, и работа по нему начинается в случае выхода из строя основного канала.

Рис. 7.5. *Статическое задание нескольких сетей по умолчанию*

Еще один путь статического задания маршрутов по умолчанию основан на использовании параметра дистанции (он описан в табл. 6.1 в главе 6, "Настройка параметров BGP"). чтобы установить степень предпочтения. Поскольку обмен этим параметром между маршрутизаторами не проводится, такой вариант может работать только в случае одного маршрутизатора, подключенного по нескольким каналам к провайдеру.

Если два статических маршрута заданы с различными значениями дистанции, то побеждает тот из них, у кого меньшее значение дистанции. Если маршрут с лучшей дистанцией становится недоступным, то маршрутизатор автоматически переходит к использованию следующего маршрута по умолчанию. Если маршруты по умолчанию имеют одинаковую дистанцию, то трафик будет распределен между двумя этими маршрутами с использованием механизмов преимущественной коммутации.

На рис. 7.6 приведен пример использования параметра дистанции для установки нескольких маршрутов по умолчанию. Система AS1 подключена к AS2 по двум каналам и устанавливает собственные маршруты по умолчанию в AS2. Один из каналов в AS1 используется как основной благодаря присвоению ему значения дистанции 50, которое меньше значения дистанции для резервного канала — 60. В случае отказа основного канала весь трафик пойдет по резервному каналу.

Рис. 7.6. *Статические маршруты по умолчанию, указывающие на несколько соединений*

Запомните следующее: если маршрут привязан к определенному интерфейсу, то интерфейс должен стать недоступным до того, как маршрут станет недействительным. Например, по умолчанию в оборудовании Cisco по HDLC-соединению ведется обмен специальными сообщениями для того, чтобы удостовериться в нормальной работе соединения. Если в течение заданного времени эти сообщения не поступают на одну из сторон, то интерфейсное соединение разрывается. В результате этих действий маршрут удаляется из таблицы маршрутов как недействительный. С другой стороны, в виртуальных каналах, образуемых с помощью технологии Frame Relay и ATM, обмен подобными сообщениями между двумя маршрутизаторами не проводится. Это означает, что, если

виртуальный канал по какой-либо причине выходит из строя, интерфейсы будут находиться в активном состоянии и связанный с ними маршрут будет считаться действительным.

Симметрия

Симметрия означает, что трафик, покидающий AS из заданной точки выхода, будет возвращаться в ту же самую точку. Этого легко достичь, если в AS существует всего одна точка входа и точка выхода. Однако, согласно требованию избыточности, необходимо наличие нескольких соединений, вследствие чего трафик имеет тенденцию к асимметричности. Когда трафик является асимметричным, все — и клиенты, и провайдеры - отмечают недостаточный уровень контроля за входом и выходом трафика из их AS. Так, трафик из восточного побережья США может попасть по "декоративному маршруту" на западное побережье и пройти через несколько промежуточных узлов внутри одной AS, прежде чем попасть на узел, сгенерировавший его. Как отмечалось в главе 6, подобные ситуации, как правило, являются результатом маршрутизации по ближайшему выходу.

На самом деле не все так плохо, как кажется. В некоторых ситуациях вполне приемлемо работать с асимметричным трафиком, в зависимости от используемых в сети приложений и общей физической топологии (скорости каналов и количества промежуточных узлов между заданными точками). Вообще клиенты и провайдеры предпочитают, чтобы их трафик возвращался примерно или точно в ту же точку, из которой он покидал AS, чтобы минимизировать потенциальные задержки, обязательно проявляющиеся в противном случае. Кроме того, клиенты, возможно, захотят переносить свой трафик как можно дальше по своей сети, чтобы избежать появления задержек или перегрузки во взаимодействующей сети.

Для реализации симметрии вам потребуется обозначить основной канал и приложить все усилия к тому, чтобы весь трафик проходил именно через этот канал. Несмотря на то что мы будем рассматривать несколько методов достижения симметрии через задание определенных правил маршрутизации, важно понять, что на практике очень редко удается достичь полной симметрии, однако это не должно вызывать особого беспокойства.

Распределение нагрузки

Распределение нагрузки заключается в оптимальном разделении трафика данных между несколькими соединениями. Часто считают, что распределение нагрузки должно быть равномерным. Равномерное распределение трафика довольно иллюзорная вещь даже в сетях, которые находятся под управлением одной администрации. В большинстве случаев добиться равномерного распределения трафика очень тяжело, ввиду многочисленности игроков, от которых зависит решение этой проблемы. Распределение нагрузки проводится для того, чтобы наилучшим образом задействовать все существующие избыточные соединения. При этом необходимо четкое понимание того, какой трафик вы собираетесь распределять — входящий или исходящий.

Трафик следует воспринимать как два связанных субъекта — входящий и исходящий. С точки зрения автономной системы, трафик, принимаемый от других AS является входящим, а трафик, передаваемый другим AS, — исходящим.

Предположим, ваша сеть подключена к двум ISP, и канал в направлении провайдера ISP1 испытывает перегрузки. Прежде всего спросите себя: "Какой трафик создает проблемы, входящий или исходящий? Получаю ли я весь трафик от ISP1 и посылаю ли я весь свой трафик ISP1?"

Структура входящего и исходящего трафика напрямую связана с тем, каким образом вы объявляете свои маршруты и с тем, как вы получаете сведения о маршрутах от других AS. На входящий трафик влияет то, как AS объявляет маршруты к своим сетям внешнему миру,

в то время как исходящий трафик зависит от обновлений маршрутов, поступающих от других AS. Разберитесь в структуре вашего трафика, так как эта информация будет ложится в основу всех последующих решений. С этой минуты, когда мы говорим о влиянии на входящий трафик, мы подразумеваем применение атрибутов к исходящим объявлениям маршрутов, потому что, то, как сведения о наших маршрутах будут получаться другими, воздействует на маршрутизацию входящего трафика. Точно так же, говоря о способах воздействия на исходящий трафик, мы будем иметь в виду применение атрибутов ко входящим объявлениям маршрутов, так как способ получения маршрутной информации нашей сетью влияет на маршрутизацию исходящего трафика. На рис. 7.7 представлены модели поведения входящего и исходящего трафика. Как видите путь, по которому исходящий трафик может попасть в сеть A, зависит от того, где получены сведения о сети A. Поскольку сведения о маршруте в сеть A были получены и от узла SF, и от NY, то ваш исходящий трафик в направлении сети A может быть направлен либо через узел SF, либо через NY.

С другой стороны, маршрут для входящего трафика, предназначенного сетям B и C, зависит от того, каким образом вы объявите маршруты к этим сетям. Если вы объявите маршрут в сеть C только через узел NY, то входящий трафик для сети C будет передаваться только по каналу с узлом NY. Точно так же, если вы объявите маршрут в сеть B только через узел SF, то входящий трафик для сети B будет передаваться только через узел SF. Хотя подобная схема работы обеспечивает, казалось бы, оптимальное распределение трафика, входящего в AS, однако она не обеспечивает избыточности для объявляемых сетей.

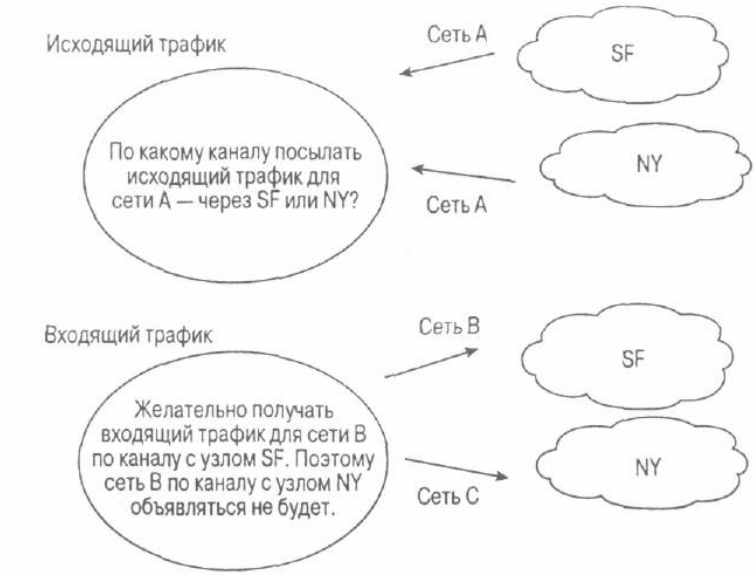


Рис. 7.7. *Принятие решений о направлении входящего и исходящего трафика*

Примеры реализации избыточности, симметрии и распределения нагрузки в сетях

Итак, к этому моменту вы уже в общих чертах знакомы с основными подходами построения сетей — избыточностью, симметрией и распределением нагрузки, которые потенциально конфликтуют друг с другом. Каким же образом сбалансировать трафик между несколькими каналами и при этом обеспечить одну точку входа и выхода, как этого требует симметрия? Эта задача еще более усложняется, когда несколько каналов разбросаны между различными маршрутизаторами в автономной системе. Атрибуты маршрутов, описанные в

главе 6, предстаапяют собой инструменты для реализации требований избыточности, симметрии и распределения нагрузки. Для достижения нужного результата оператору необходимо выбрать правильные атрибуты и схемы фильтрации.

В этом разделе представлены примеры конфигурации сетей с оптимизацией требований избыточности, симметрии и распределения нагрузки. Эти примеры, естественно, не могли охватить все возможные конфигурации сетей, и технологические решения, приведенные здесь, также не являются единственными в своем роде. Однако уроки, которые вы можете извлечь из этих примеров, помогут вам в реализации более эффективных схем.

Первый пример представляет собой простейший случай, в последующих — степень сложности возрастает. Обратите внимание, что между клиентом и провайдером в большинстве случаев прослеживается четкая граница, так как провайдер может, в свою очередь, являться клиентом другого провайдера. Принципиальное отличие заключается в следующем: клиенты получают доступ в Internet путем подключения к провайдерам и не предлагают подключение другим клиентам. Провайдеры же, предлагая доступ в Internet клиентам, сами могут выступать в роли клиентов для других провайдеров.

В последующих разделах рассматриваются примеры построения сетей, в которых клиенты принимают от своих провайдеров минимальное количество информации о маршрутах или вообще не получают сведения о маршрутах, получают частичные сведения о маршрутах, полные сведения маршрутах, или все это получается в комбинированном виде. Когда клиент принимает от провайдера минимальное количество маршрутов или вообще ни одного маршрута (только *с использованием маршрута по умолчанию),* вы можете предположить, что клиент может получать сведения о маршруте 0/0 или о парс объединенных маршрутов, позволяющих ему статически задать маршруты по умолчанию. *Частичная маршрутизация* обычно включает в себя локальные маршруты провайдера и маршруты от провайдера в сети других клиентов. *Полная маршрутизация* представляет собой совокупность всех существующих в сети Internet маршрутов — это около 75000 маршрутов на начало 2000 года. В случае комбинации этих вариантов клиент может получать маршрут по умолчанию и частичные маршруты от одного и того же провайдера иди частичные маршруты от одного провайдера it полные маршруты от другого и т.д.

Вариант 1: одноканальное соединение

Клиенты с одноканальным подключением к Internet имеют всего одно соединение с провайдером. На рис. 7.8 показана подобная схема подключения.

Подобные клиенты обычно получают необходимое обслуживание путем указания маршрутов по умолчанию в сеть провайдера. Провайдер, со своей стороны, также может установить статические маршруты в сеть своего клиента. Этот метод является самым дешевым и наиболее эффективным. Технически ни провайдер, ни клиент не нуждаются в работе по протоколу BGP. Таким образом, маршрутизатору на стороне клиента нет необходимости получать сведения обо всех маршрутах в сети Internet- Это существенно снижает использование памяти и нагрузку на процессор маршрутизатора. В этом случае нет необходимости заботиться о симметрии, так как трафик имеет лишь одну точку входа и точку выхода.

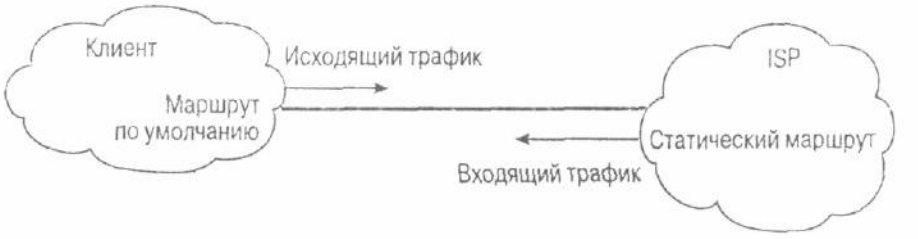


Рис. 7.8. *Пример однокаиалыюго соединения*

Узлы с одноканальным подключением обычно имеют один канал с Internet. В этом

случае о резервировании можно забыть. Если соединение по какой-либо причине пропадает, то клиент вынужден ожидать, пока оно будет восстановлено. Очевидно, что подобная схема не может удовлетBGPять требованиям непрерывности передачи данных. Узды с одноканальным подключением без резервных каналов не могут использоваться в приложениях с высокими требованиями к надежности соединений. Следует также отметить, что одним из преимуществ этой схемы является стандартная маршрутизация по умолчанию, которая значительно упрощает систему.

Вариант 2: многоканальное соединение с одним провайдером

Клиент, у которою имеется несколько соединений с Internet через одного провайдера, называется клиентом с многоканальным соединением через одного провайдера. Для обеспечения многоканального соединения с одним провайдером предположим, что в качестве протокола маршрутизации используется протокол BGP. Хотя и большинстве случаев это не требуется, он обеспечит большую гибкость, чем какой-либо другой протокол маршрутизации. Более того, в этом случае (многоканальное соединение с одним провайдером) можно воспользоваться частными номерами AS, так что нет необходимости запрашивать уникальный номер AS. Подобная схема работы предложена в главе 6. Дополнительную информацию по этим вопросам вы сможете найти в RFC 2270. В последующих разделах мы обсудим следующие вопросы.

- Маршрутизация только по умолчанию; один канал основной и один резервный.
- Маршрутизация по умолчанию: основной и резервный каналы, плюс частичная маршрутизация.
- Маршрутизация по умолчанию: основной и резервный каналы, плюс полная и частичная маршрутизация.
- Автоматическое распределение нагрузки.
- Распределение нагрузки между двумя маршрутизаторами, которые совместно используют несколько каналов.

Маршрутизация только по умолчанию: один канал основной и один резервный

При такой схеме подключения клиент конфигурирует маршрутизацию по умолчанию в сторону провайдера и не принимает никакой информации о частичных или полных маршрутах. Клиент может по умолчанию работать с двумя каналами сразу. На рис. 7.9 клиент может использовать один канал как основной для всего трафика, а другой — иметь в качестве резервного и работать по нему при отказе основного канала. Если имеется более двух соединений с провайдером, клиент может устанавливать несколько маршрутов по умолчанию, задавая им различные уровни предпочтения.

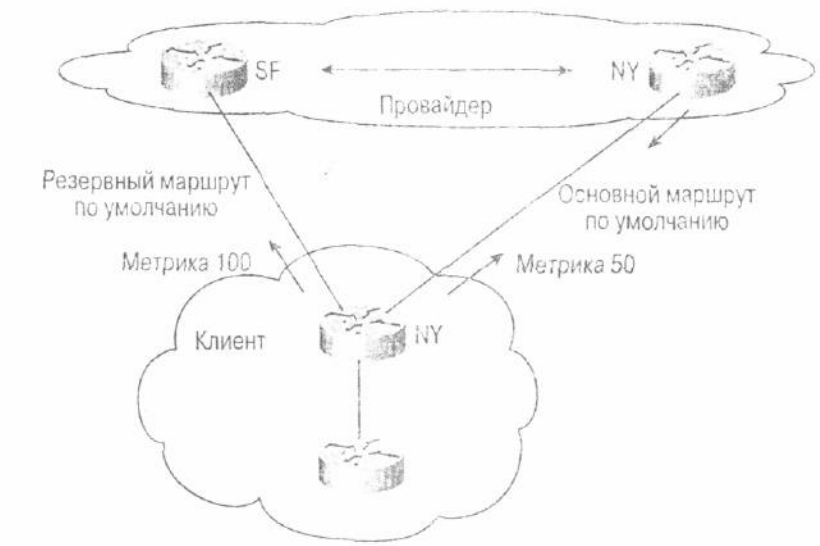


Рис. 7.9. *Вариант многоканального соединения с одним провайдером*

См. в главе 12 раздел "Маршрутизация только по умолчанию: один основной и один резервный канал". В последующих разделах мы обсудим управление входящим и исходящим трафиком.

Исходящий трафик клиента

В варианте сети, представленном на рис. 7.9, когда при подключении к провайдеру в нескольких географических точках используется один маршрутизатор, имеется возможность использовать несколько статических маршрутов по умолчанию с различными значениями дистанции. При этом маршрут по умолчанию с меньшей дистанцией будет основным. Кроме того, сведения о маршруте по умолчанию 0/0 или нескольких объединенных маршрутах также могут быть получены от провайдера динамически, что позволяет затем клиенту установить маршрут по умолчанию. Для выбора приоритетности одного маршрута по умолчанию можно использовать локальные предпочтения.

Предположим, что маршрут по умолчанию (рис. 7.9) через узел NY более предпочтителен, чем маршрут через узел SF. При нормальной работе клиент будет использовать канал с узлом NY в качестве основного, а канал с узлом SF как резервный.

Для исходящего трафика распределение нагрузки некритично, так как весь трафик пересылается по основному каналу, а второй канал выступает в качестве резервного.

Отсутствие распределения нагрузки компенсируется тем, что маршрутизатору клиента требуется меньше памяти и ресурсов процессора.

Входящий трафик клиента

Клиент может объявлять свои сети провайдеру посредством протокола BGP. Провайдер имеет два соединения с клиентом. Выбор провайдером соединения для связи с клиентом будет определяющим для входящего трафика клиента. Обычно провайдер старается направить весь трафик (предполагая, что все атрибуты BGP одинаковы) через ближайшую к себе точку выхода из клиентской AS. Если трафик в направлении клиента окажется вблизи канала с узлом NY, то он будет поступать в AS клиента именно через этот узел. Если трафик окажется поблизости узла SF, то он будет направлен клиенту через узел SF.

Все эти факторы неподвластны клиенту. Клиенты, которые хотят преодолеть эти ограничения и самостоятельно контролировать свой входящий трафик на одном из каналов, должны объявлять свои маршруты с различными метриками. Провайдер в этом случае будет направлять трафик в AS клиента, руководствуясь значениями метрик. На рис. 7.9 показана ситуация, когда клиент объявляет свои маршруты с метрикой 50 в направлении узла NY и с метрикой 100 в направлении узла SF. Таким образом, трафик в сеть клиента будет направляться по маршруту через узел NY.

Маршрутизация по умолчанию, основной и резервный каналы, плюс частичная маршрутизация

Этот вариант подключения представляет собой тот же самое, что и предыдущий, за исключением того, что клиент способен принимать частичную маршрутизацию от провайдера. Такая схема подключения показана на рис. 7.10.

В случае использования подхода, представленного на рис. 7.10, клиент имеет дополнительную гибкость в выборе точки выхода, так как в этом случае ему предоставляется больше маршрутной информации. Далее мы обсудим работу входящего и исходящего трафика в подобной схеме.

См. в главе 12 раздел "Маршрутизация по умолчанию: основной и резервный каналы плюс частичная маршрутизация"

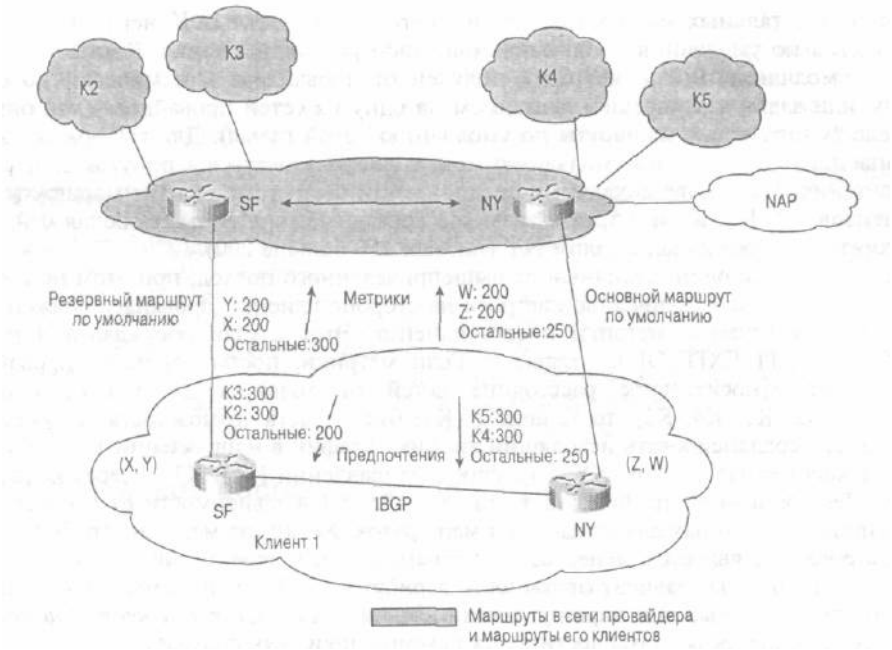


Рис. 7.10. *Вариант многоканального соединения с одним провайдером с использованием частичной маршрутизации*

Исходящий трафик клиента

Рассмотрим ситуацию, когда Клиент 1 подключается к провайдеру через два отдельных маршрутизатора. У клиента в этом случае есть выбор, какой из маршрутов использовать для каждого из частичных маршрутов, полученных от провайдера. Эта задача решается установкой различных локальных предпочтений для маршрутов, поступающих в AS клиента. Локальные предпочтения могут быть установлены на основе атрибута AS_PATH, префикса или того и другого. Если набор предпочтений установлен на основе AS_PATH, то локальные предпочтения будут применяться ко всем префиксам, содержащимся в AS. Если решение о маршрутизации должно приниматься на основе префиксов, локальные предпочтения могут быть установлены для каждого префикса. На рис. 7.10, показано, что в зависимости от физического местоположения AS или префиксов с учетом AS провайдера, клиент может пересылать трафик на Клиента 2 и Клиента 3 (К2 и К3) по каналу с узлом SF, и трафик Клиенту 4 и Клиенту 5 (К4 и К5) — по каналу с узлом NY. Клиент может реализовать подобную схему следующим образом.

- Маршрутам, получаемым по каналу от узла NY, назначить локальное предпочтение 300 в сторону клиентов К4 и К5. Задать всем другим маршрутам значение 250. (Таким образом, будут описаны маршруты к К2 и К3).
- Маршрутам, сведения о которых были получены по каналу от узла SF, назначить локальное предпочтение 300 для маршрутов к клиентам К2 и К3. Задать всем

остальным маршрутам локальное предпочтение 200. (Таким образом, будут описаны маршруты к К4 и К5).

Когда к одному узлу имеется несколько маршрутов (через внешний и внутренний BGP), клиент, скорее всего, будет использовать маршруты к К4 и К5 по каналу с узлом NY (так как 300 больше 200). Подобным образом клиент воспользуется маршрутами с К2 и КЗ по каналу с узлом SF (так как 300 больше 250). Для всех других клиентов, кроме К2, КЗ, К4 и К5, будет использоваться канал через узел NY (так как 250 больше 200).

Для всех остальных маршрутов в сети Internet, неизвестных Клиенту I, будет использоваться по умолчанию либо основной, либо резервный канал. Кроме того, маршрут по умолчанию 0/0 может быть получен от провайдера динамически по обоим каналам или задан статически с указанием на одну из сетей провайдера (как описано в разделе "Статические маршруты по умолчанию" этой главы). Для выбора основного или запасного маршрута по умолчанию можно воспользоваться атрибутом локального предпочтения. На основе локальных предпочтений были установлены маршруты в сети клиентов К2, КЗ, К4 и К5, при этом все остальные маршруты, включая 0/0, будут обслуживаться через канал с узлом NY (так как 250 больше 200).

Существует еще один, отличный от вышеприведенного подход, при этом не требуется проводить большое количество настроек на стороне клиента, провайдер должен лишь посылать используемые метрики в сеть клиента. Этот случай обсуждался в разделе "Атрибут MULTI_EXIT_DISC" главы 6. Если метрики, поступающие от провайдера, представляют относительное расстояние сетей от точек входа в сети клиентов (например, К2, КЗ, К4, К5), то Клиент 1 (К1) будет иметь возможность соответствующим образом сбалансировать исходящий трафик. Трафик в направлении К4 и К5 будет выходить через канал с узлом NY, а трафик в направлении К2 и КЗ — через канал с узлом SF. Весь остальной трафик будет покидать AS К1 в зависимости от соответствующих метрик, полученных для каждого из маршрутов. Хотя этот метод не требует серьезных настроек, он является менее однозначным, так как траектория трафика от К1 в этом случае полностью зависит от точности атрибутов MED, получаемых от провайдера. Помните, что подобный тип маршрутизации называется *маршрутизацией по ближайшему выходу.* Наилучшие результаты достигается комбинацией двух подходов.

Входящий трафик клиента

Клиент может повлиять на входящий трафик путем объявления для соединений различных метрик. Некоторые провайдеры даже поощряют своих клиентов к отправке сведений о маршрутах с внутренними метриками 1GP (эти вопросы также рассматривались в главе 6). Таким образом, провайдер будет доставлять трафик клиенту через соединение, которое находится наиболее близко к пункту назначения. На рис. 7.10 клиент вручную устанавливает все метрики, чтобы получить такой результат.

- Маршруты, сведения о которых пересылаются по каналу через узел NY, несут информацию о префиксах Z и W с MED, равным 200. Всем остальным префиксам присваивается метрика 250. (Сюда входят префиксы X и Y).
- Маршруты, сведения о которых пересылаются по каналу через узел SF, несут информацию о префиксах X и Y со значением 200. Остальным префиксам задается MED 300. (К ним относятся Z и W).

Когда к одному и тому же пункту назначения существует несколько маршрутов, провайдер обеспечивает доступ к префиксам Z и W через соединение с узлом NY (так как 200 меньше 300). Точно так же провайдер будет получать доступ к префиксам X и Y по каналу от узла SF (так как 200 меньше, чем 250). Для всех других префиксов, отличных от X, Y, W и Z, провайдер будет выбирать канал с узлом NY (так как 250 меньше 300).

В RFC I998[2] описан еще один метод влияния на входящий трафик. Хотя мы не будем обсуждать этот вариант, советуем вам ознакомиться и с таким подходом к проблеме управления входящим трафиком.

Маршрутизация по умолчанию, основной и резе^ный каналы, плюс полная и частичная маршрутизация

Для клиентов, которые имеют несколько соединений с одним провайдером, имеется возможность получать сведения обо всех маршрут к провайдеру через один канал, а по другим каналам либо вообще не получать маршрутной информации, либо получать частичные сведения о маршрутах. Подобный подход описан и в предыдущих разделах. Для управления исходящим трафиком клиента используется совокупность локальных предпочтений, а для упраачения входящим трафиком — набор метрик (или процедуры, описанные в RFC 1998). Кроме того, если имеется возможность обмена внутренними метриками между клиентом и провайдером, то вы можете обеспечить определенный уровень распределения нагрузки.

Внимание!

При обработке исходящего трафика изменение точек выхода для заданных маршрутов очень опасно. Это может привести к образованию петель маршрутизации в исходящем трафике, который будет по умолчанию следовать по маршруту IGP в направлении клиентского маршрутизатора BGP, а тот, в свою очередь, будет направлять трафик по умолчанию на другой маршрутизатор BGP. Другими словами, при работе с маршрутами по умолчанию все маршруты в домене маршрутизации должны вести себя одинаково, особенно при пропадании одного из маршрутов по умолчанию. Возможно, сейчас это может показаться не-! много запутанным, но мы более подробно рассмотрим эти вопросы в следующей главе.

Автоматическое распределение нагрузки

Как вы уже, очевидно, поняли, распределение нагрузки — не такая простая задача и требует тщательного планирования. В программном обеспечении Cisco IOS поддерживается динамическое распределение нагрузки на отдельном маршрутизаторе для узлов, сведения о маршрутах с которыми получены по EBGP и находящимися в той же автономной системе. Таким образом, от администратора требуется гораздо меньше усилий для настройки.

Примечание

Отметим, что реальный эффект от распределения нагрузки зависит от режимов пакетной коммутации, разрешенных для протоколов маршрутизации. Хотя обсуждение принципов коммутации не входит в круг рассматриваемых нами вопросов, понимание основ коммутации в сетях просто необходимо.

На рис. 7.11 приведен пример, когда маршрутизатор на узле NY подключается к провайдеру по двум каналам и получает по ним идентичные обновления маршрутов.

Маршрутизатор Cisco будет (локально в своей таблице IP-маршрутов) содержать шесть идентичных BGP-маршрутов к одному пункту назначения. Однако при передаче EBGP-обновлений другим IBGP-узлам он будет анонсировать только один (наилучший) маршрут в заданный пункт назначения. Адрес маршрута к следующему ближайшему узлу будет автоматически изменяться, чтобы отражать собственный адрес маршрутизатора NY, а не транслировать адрес ближайшего следующего узла из EBGP в IBGP. Обратите внимание, что эта операция выполняется автоматически только в том случае, когда задано динамическое распределение нагрузки.

По умолчанию маршрутизатор Cisco будет распределять нагрузку на основе информации о пунктах назначения (т.е. хостах, на которые интенсивней всего направляется трафик). Распределение нагрузки, таким образом, выполняется циклически. Один хост будет закреплен за одни;*-! маршрутом (интерфейсом), следующий хост — за другим маршрутом (интерфейсом) и т.д.

В схеме, приведенной на рис. 7.11, предполагается, что клиент получает два идентичных маршрута в сеть 192.213.10.0/24. Без автоматического распределения нагрузки в процессе принятия решения в BGP выбирается только один из существующих маршрутов. Для распределения трафика между маршрутами сетевой администратор должен соответствующим образом модифицировать атрибуты BGP.

См. в главе 12 раздел "Распределение нагрузки в BGP при работе по нескольким каналам

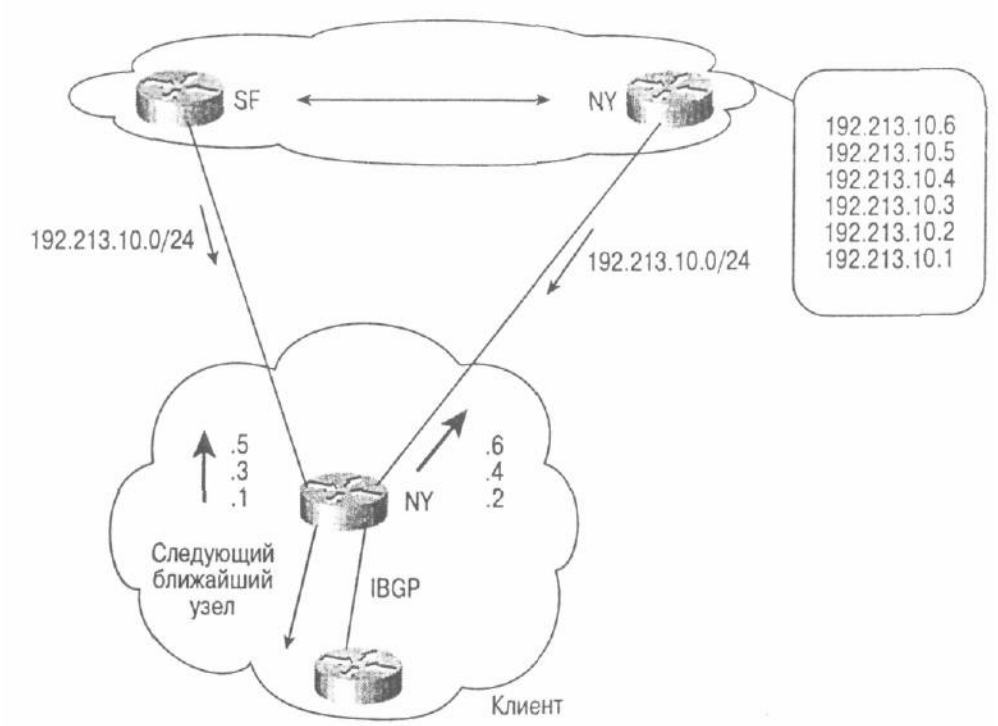


Рис. 7.11. *Маршрутизатор получает обновления маршрутов из двух источников*

При автоматическом распределении нагрузки для префикса 192.213.10.0/24 в протоколе BGP хранится две записи — одна для канала с узлом SF, а вторая для канала с узлом NY. Исходящий из сети клиента трафик распределяется между двумя каналами на циклической основе, если клиенту требуется передать трафик в пункт назначения 192.213.10.1 через 192.213.10.6. При этом по каналу с узлом SF можно попасть на узел 10.1, на 10.2 — по каналу с узлом NY, на 10.3 — по каналу с узлом SF и т.д.

Примечание

Как уже отмечалось, в протоколе BGP по умолчанию устанавливается только один, наилучший, маршрут к каждому пункту назначения, присутствующему в таблице маршрутов. Однако для организации нескольких маршрутов в таблице IP-маршрутов, если они объявляются через одни и те же соседние AS, можно использовать многонаправленный BGP (BGP Multipath). Для установления до шести различных маршрутов к одной сети на маршрутизаторе может быть использована команда maximum-paths. В главе 12 вы найдете дополнительные сведения о настройке BGP Multipath.

Примечание

Распределение нагрузки вышеописанным образом функционирует только в тех случаях, когда идентичные обновления маршрутов поступают на один и тот же маршрутизатор от одного провайдера. Эта схема не будет работать, если клиентский узел подключен к нескольким провайдерам.

Распределение нагрузки между двумя маршрутизаторами, которые совместно используют несколько каналов

В некоторых случаях два маршрутизатора сЪвместно используют несколько физических каналов в качестве запасных или для обеспечения высокопроизводительных

соединений, как это показано на рис. 7.12. Хотя автоматическое распределение нагрузки наиболее применимо к исходящему трафику, если вы хотите повлиять на входящий трафик. В этом случае вам следует прибегнуть к манипулированию метриками маршрутов.

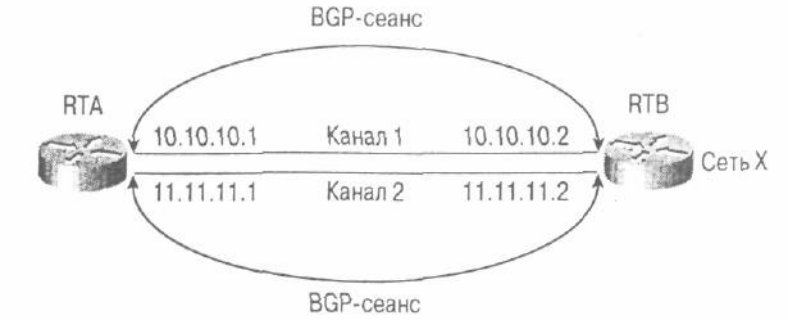


Рис. 7.12. *Распределение нагрузки между двумя маршрутизаторами, которые совместно используют несколько каналов*

Одним из условий распределения трафика в примере, приведенном на рис. 7.12, является обеспечение этого динамическим путем. На самом деле это всего лишь частный случай автоматического распределения нагрузки, о котором мы говорили ранее. Однако, с другой стороны, динамическое распределение нагрузки может привести к созданию дополнительной нагрузки на маршрутизатор. В этом случае каждый маршрутизатор будет принимать дублированные обновления маршрутов от другого маршрутизатора. При реализации полной маршрутизации вы можете оказаться в ситуации, когда по каждому из каналов будет передаваться около 70000 маршрутов. Чтобы избежать подобного, желательно (и даже необходимо) реализовать распределение нагрузки для схемы, приведенной на рис. 7.12, на основе статической маршрутизации.

При нормальной работе в протоколе BGP на каждый префикс хранится информация о наиболее подходящем ближайшем узле. Как показано в табл. 7.1, маршрутизатор RTA получает два идентичных BGP-маршрута для сети X.

Таблица 7.1. *Таблица BGP-маршрутов маршрутизатора RTA — сеть X доступна через узел 10.10.10.2*

Пункт назначения	Следующий ближайший узел
СетьX	10.10.10.2 (наилучший)
СетьX	11.11.11.2

Протокол BGP выбирает наилучший маршрут и вносит его в таблицу IP-маршрутов. В нашем случае выбирается маршрут через узел 10.10.10.2. В табл. 7.2 представлена таблица IP-маршрутов маршрутизатора RTA, где доступ к следующему ближайшему узлу 10.10.10.2 осуществляется по каналу 1. Следовательно, в этом случае не обеспечивается распределение нагрузки.

Таблица 7.2. *Таблица IP-маршрутов маршрутизатора RTA — сеть X доступна по каналу 1*

Пункт назначения	Следующий ближайший узел
СетьX	10.10.10.2
10.10.10.0/24	Канал 1

См. в главе 12 раздел "Распределение нагрузки между двумя маршрутизаторами, которые совместно используют несколько каналов"

С целью обеспечения более интеллектуального распределения нагрузки имеется возможность "обмануть" BGP путем установки в качестве следующего узла виртуаль-нпро интерфейса, а не физического соединения. Тогда таблица IP-маршрутов будет использоваться для выполнения реального распределения нагрузки через IP-адрес виртуального интерфейса по нескольким интерфейсным IP-адресам, которые подключены

непосредственно. На рис. 7.13 маршрутизатору RTB назначается петельный интерфейс (виртуальный интерфейс), который маршрутизатор RTA может использовать при соединении по протоколу BGP с соседним узлом. Таким образом, в качестве следующего узла будет использоваться IP-адрес петельного интерфейса, а не физического соединения. Для распределения нагрузки в каналах без участия протокола BGP может использоваться как динамическая, так и статическая маршрутизация по IGP.

Как показано в табл. 7.3, маршрутизатор RTA будет получать сведения о BGP-маршрутах от своего соседа 12.12.12.12 и сможет попасть в сеть X именно через узел 12.12.12.12.

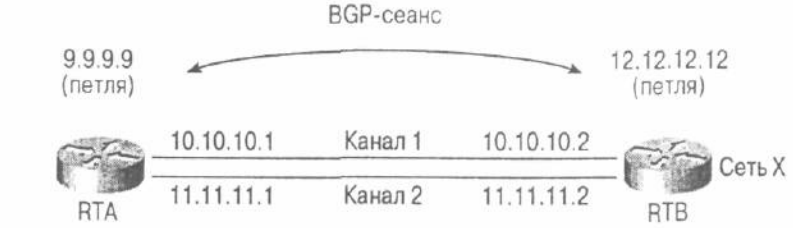


Рис. 7.13. *Простой сеанс BGP по нескольким физическим каналам*

Таблица 7.3. *Таблица BGP маршрутизатора RTA — сеть X доступна через узел 12.12.12.12*

Пункт назначения	Следующий ближайший узел
СетьX	12.12.12.12

В табл. 7.4 описана таблица IP-маршрутов маршрутизатора RTA. Ближайший узел с сетевым адресом 12.12.12.12 может быть достигнут и по каналу 1, и по каналу 2. Попасть в сеть 12.12.12.0/24 можно либо с использованием протокола IGP, либо путем указания нескольких статических маршрутов через каналы 1 и 2. Теперь маршрутизатор сможет распределить трафик между этими каналами. Вследствие рекурсивного опроса маршрутов в подобной схеме распределение нагрузки выполняется между сетями, а не между отдельными узлами. Так, трафик в сети от маршрутизатора RTB также может распределиться по нескольким каналам.

Таблица 7.4. *Таблица маршрутов маршрутизатора RTA — сеть X доступна по каналам 1 и 2*

Пункт назначения	Следующий ближайший узел
СетьX	12.12.12.12
12.12.12.0/24	Канал 1
12.12.12.0/24	Канал 2

Вариант 3: многоканальное соединение с различными провайдерами

Клиент, подключенный к нескольким провайдерам, называется клиентом с многоканальным соединением с различными провайдерами. Подобные схемы подключения имеют право на существование при необходимости обеспечить избыточность и работу корпоративной сети в различных географических областях. Поведение исходящего трафика зависит от каждого конкретного случая. Поведение входящего трафика ко всех случаях одинаково, мы рассмотрим его в конце раздела.

В последующих разделах мы затронем следующие вопросы, касающиеся организации многоканальных подключений к различным провайдерам.

- Маршрутизация только по умолчанию; один основной и один резервный каналы.
- Маршрутизация по умолчанию: один основной и один резервный каналы, плюс

частичная маршрутизация.

- Маршрутизация по умолчанию: один основной и один резервный каналы, плюс частичная и полная маршрутизация.
- Входящий трафик клиента (управление атрибутом AS_PATH).

Маршрутизация только по умолчанию: один основной и один резервный каналы

При такой структуре подключения клиент может использовать маршруты по умолчанию, установленные с узлом провайдера. Один канал используется как основной, а второй — как резервный. На рис. 7.14 представлена такая схема подключения.

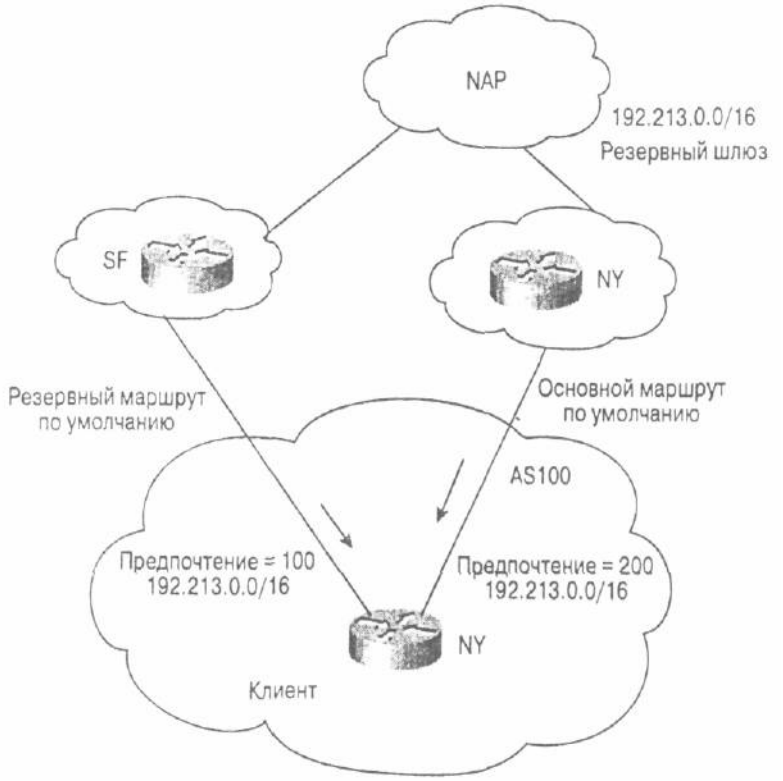


Рис. 7.14. *Многоканальное подключение к двум провайдерам*

Клиент может самостоятельно устанавливать или получать сведения о маршрутах по умолчанию от двух провайдеров. Причем в последнем случае маршруты по умолчанию могут быть получены как статически, так и динамически. Клиент может выбирать один из маршрутов по умолчанию в качестве основного путем задания административной дистанции или локального предпочтения. Хорошо зарекомендовало себя указание маршрутов по умолчанию к обоим провайдерам, позволяющее принимать сведения об одной сети от двух провайдеров. На основании маршрута в эту сеть клиент будет залапать свой мпршрут по умолчанию 0/0и для выбора того или иного канала манипулировать значениями локальных предпочтений. Если один из маршрутов вследствие отказа канала одного из провайдеров исключается из таблицы маршрутов, то его место занимает другой маршрут по умолчанию. Клиент может договориться с провайдерами о пересылке только одной записи о сети по умолчанию или со своей стороны фильтровать маршруты, исключая все, кроме одного.

На рис. 7.14 клиент указывает маршрут по умолчанию в направлении префикса 192.213.0.0/16, который он получает от обоих провайдеров. Канал с узлом NY будет основным, а канал с узлом SF -- резервным. Так, клиент устанавливает наибольшее

локальное предпочтение для канала с NY (200).

Маршрутизация по умолчанию: один основной и один резервный каналы, плюс частичная маршрутизация

В рассмотренной выше схеме имеется возможность организовать частичную маршрутизацию, что повлияет на поведение исходящего трафика. На рис. 7.15 показана такая схема подключения. Клиент принимает частичные маршруты от одного или обоих провайдеров. В то же время клиент должен определить или сконфигурировать маршруты по умолчанию в направлении обоих провайдерских узлов, причем выбрать из них основной и резервный.

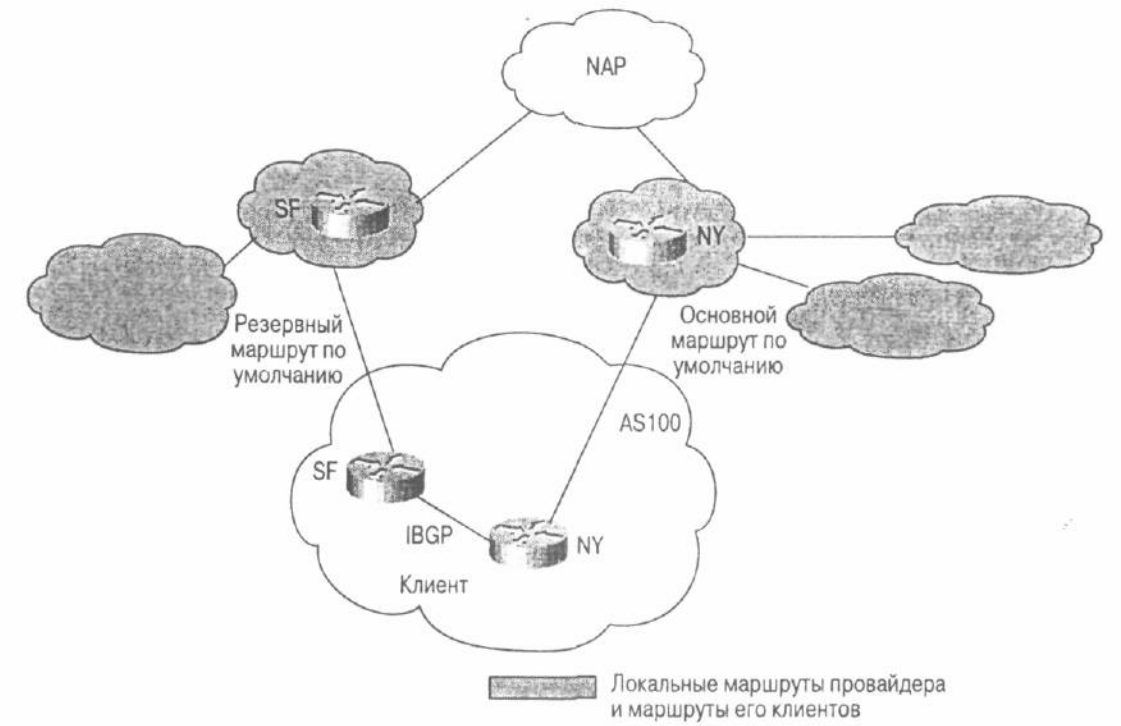


Рис. 7.15. *Многоканальное подключение к двум провайдерам с применением час-тинной маршрутизации*

Принимая частичные маршруты от провайдеров, клиенту не требуется получать сведения обо всех маршрутах в сети Internet, и он может самостоятельно принимать решения о наилучших маршрутах при взаимодействии с провайдерами. (Для некоторых крупных провайдеров частичные маршруты составляют существенную долю всех маршрутов, проходящих через их узлы). В ситуации, представленной на рис. 7.15, в протоколе BGP делается правильный выбор маршрута, и узел клиента выбирает тот канал с провайдером, который находится ближе всего к сети назначения (т.е. выбирается кратчайший маршрут через AS). Для других маршрутов в сети Internet будет применяться принцип основного и резервного маршрутов. Клиент может указать определенную сеть в качестве сети по умолчанию и принимать сведения о маршрутах к ней от обоих провайдеров с использованием механизма локальных предпочтений для выбора того или иного маршрута.

Маршрутизация по умолчанию: один основной и один резервный каналы, плюс частичная и полная маршрутизация

При подключении одного узла к различным провайдерам нет необходимости в получении полных маршрутов от одного или обеих провайдеров, если только клиент не планирует сам стать провайдером и передавать через свой узел все маршруты своим клиентам (т.е. чтобы его AS работала в транзитном режиме). На рис. 7.16 представлена схема подобного подключения.

Клиент может принимать полные сведения о маршрутах от одного или от обоих провайдеров, в зависимости от требований, предъявляемых к распределению нагрузки. Если организуется полная маршрутизация от двух (или более) провайдеров, клиент может

воспользоваться локальными предпочтениями, чтобы принимать решение об организации доступа в ту или иную сеть через определенного провайдера. Это решение может приниматься на основании номера AS, префикса или информации о сообществах. В некоторых случаях клиенту может понадобиться принимать полные маршруты от одного провайдера, а с другим провайдером реализовать взаимодействие через маршрутизацию по умолчанию и частичную маршрутизацию. При такой схеме работы клиент может получить наибольшие преимущества от обоих провайдеров без необходимости управления всеми маршрутами на различных каналах. Как вы увидите позже, нестабильность работы Internet может быть следствием того, что у какого-либо из провайдеров слишком высокая нагрузка на процессор маршрутизатора.

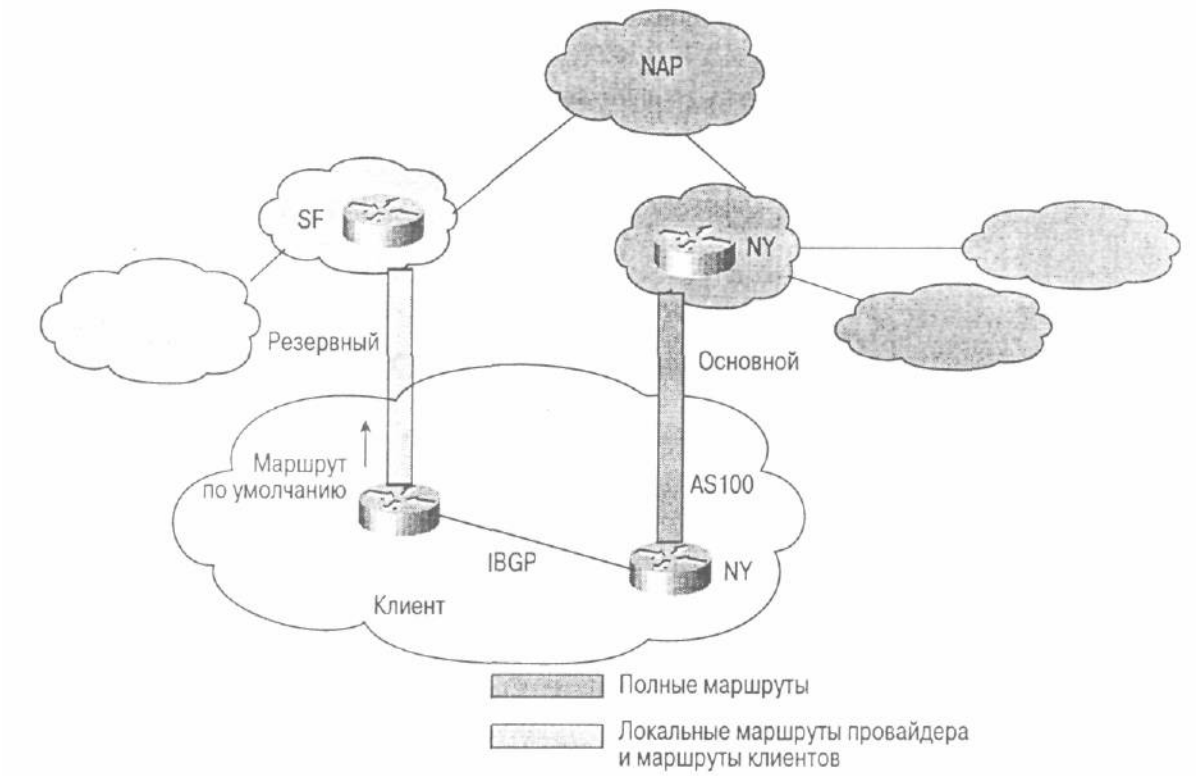


Рис. 7.16. *Многоканальное подключение к двум провайдерам с использованием полной и частичной маршрутизации*

На рис. 7.16 показано, что клиент получает полные сведения о маршрутах от провайдера на узле NY и частичные — от провайдера на узле SF. На клиентском узле также указывается маршрут по умолчанию в направлении провайдера SF. Канал с узлом SF будет использоваться для маршрутов клиента и локальных маршрутов провайдера SF вследствие более короткого маршрута AS. Для всех других маршрутов будет использоваться канал с узлом NY, так как канат с узлом SF обеспечивает только частичную маршрутизацию. При выходе из строя канала с узлом SF все сети могут быть доступны по каналу с узлом NY. Есди же пропадает канал с узлом NY, то вес маршруты, используемые в сети Internet могут быть получены по умолчанию по каналу с узлом SF.

См. в главе 12 раздел "Подключение к различным провайдерам по нескольким каналам"

Входящий трафик клиента

Входящий трафик клиента зависит от того, каким образом его узел объявляет сведения о своих сетях провайдерам. Отметим, что в случае подключения к нескольким провайдерам рассылка различных метрик' от клиентского узла не принесет никаких результатов. Причиной этого является нетранзитивность значений MED. Другими словами, обмен значениями MED может проводиться только между клиентом и провайдером, а провайдеры между собой уже не могут обмениваться этой информацией.

Чтобы динамически влиять на маршрутизацию со стороны провайдера, клиент

может использовать атрибут AS_PATH. Вставляя фиктивные записи в этот атрибут, клиент может изменять длину AS_PATH. Так, провайдеры, получая информацию об одних и тех же префиксах, но с различной длиной маршрутов, будут выбирать маршрут с наименьшей длиной (если все остальные более приоритетные атрибуты одинаковы). Обратите внимание, что при подключении к нескольким провайдерам недостаточно напрямую воздействовать на провайдера только из-за отсутствия гарантий того, что соседний провайдер будет самостоятельно получать трафик от других провайдеров для заданных сетей клиента. Управление маршрутами должно распространяться на провайдеров вплоть до точки обмена трафиком, так как именно в ней баланс (равно как и длина маршрута) будут меняться в ту или иную сторону.

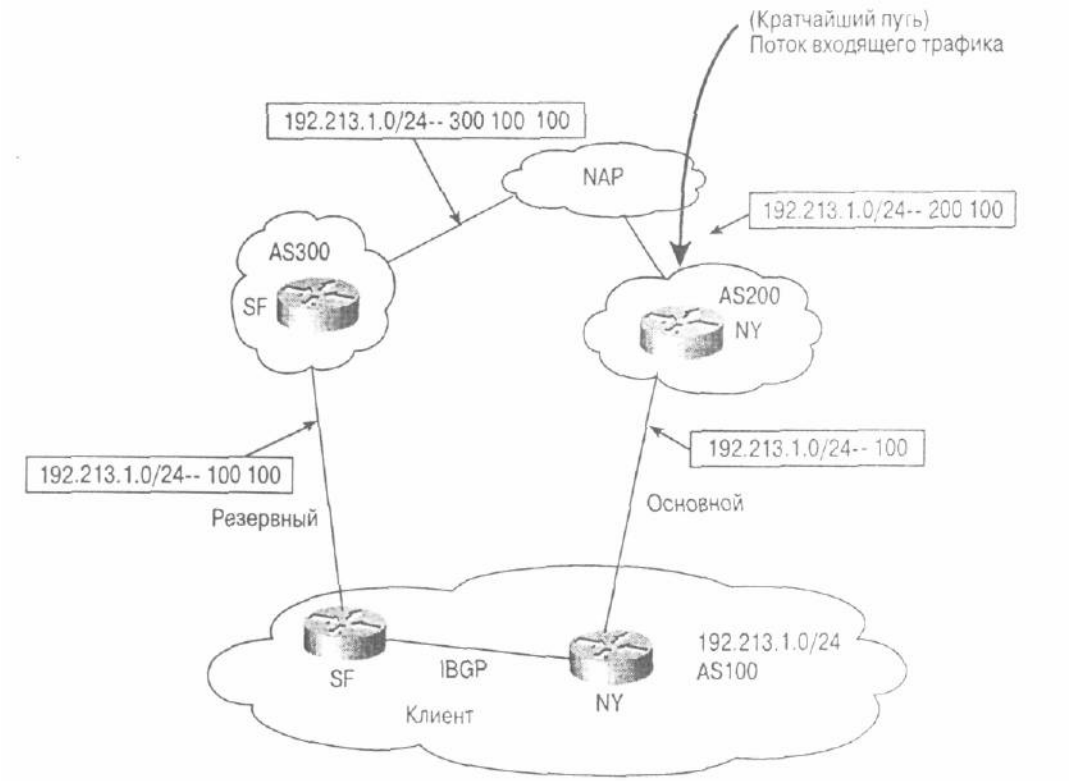


Рис. 7.17. *Использование фиктивных записей в AS_PATH с целью управления маршрутизацией*

На рис. 7.17 представлен пример воздействия фиктивных записей в AS_PATH на систему маршрутизации. Клиент (AS 100)/добавляет фиктивную запись (100) в свой атрибут AS PATH, который затем передается на AS300. Провайдеры в NAP получат сведения об одних и тех же префиксах, но с различной длиной маршрута (300 100 100 и 200 100) и выберут более короткий из них через AS200 (предполагается, что все остальные более приоритетные атрибуты одинаковы). В качестве фиктивной записи следует задавать номер той AS, которая ее генерирует (в нашем случае 100).

Вариант 4: подключение клиентов к одному провайдеру с резервным каналом между ними

В определенных ситуациях клиенты, объединенные общими интересами, могут по договоренности организовывать соединение между собой и резервное соединение через сеть Internet. То есть клиенты могут быть подключены к одному провайдеру и в то же время обладать альтернативным соединением между собой. В этом случае обычно имеются два варианта.

- Частное соединение между двумя клиентами может использоваться как вторич ное (резервное) в случае выхода из строя одного из каналов с Internet.
- Частное соединение может использоваться как основное для внутреннего трафика

между двумя компаниями и как резервное в случае выхода из строя одного из каналов с Internet. Если требуется реализовать резервирование, то клиенты должны объявлять маршруты в свои сети провайдеру. При этом AS одного из клиентов должна быть готова к работе в транзитном режиме, если выйдет из строя канал с Internet другого клиента.

Использование частного канала только в качестве резервного

На рис. 7.18 приведена схема подключения систем AS2 и AS3 к одному провайдеру — AS1. Клиентские системы AS2 и AS3 имеют также канал между собой, который используется только как резервный. Отметим, что в AS2 и AS3 введены одинаковые правила маршрутизации.

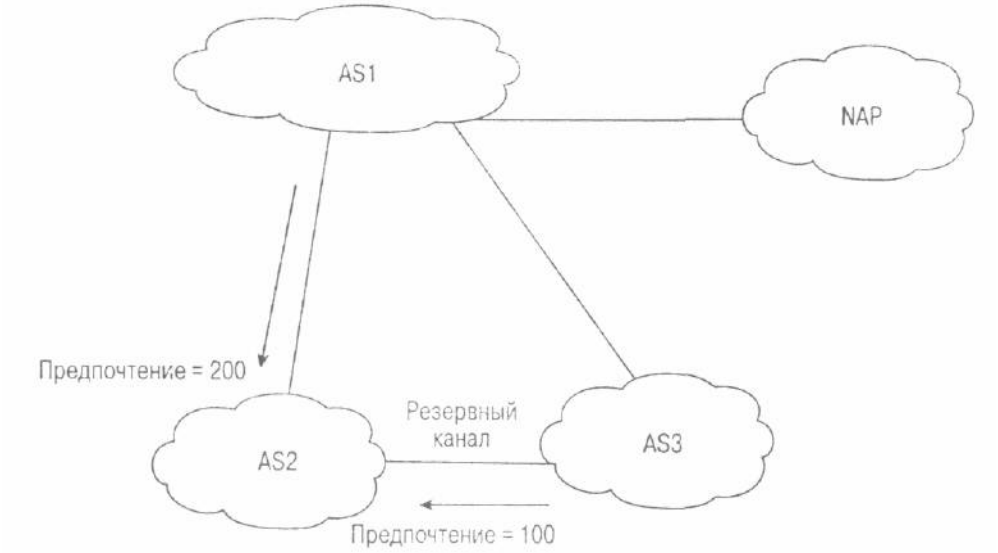


Рис. 7.18. *Использование частного капала в качестве резервного*

В схеме, представленной на рис. 7.18, особый интерес вызывает исходящий из AS2 трафик. Во всех случаях: когда AS2 получает полные сведения о маршрутах или работает с комбинацией маршрутов по умолчанию, полной и частичной маршрутизации от провайдера и AS3, -- в ней необходимо устанавливать локальные предпочтения для поступающих от AS1 маршрутов (200). выше, чем те, что принимаются от AS3 (100). В результате наивысший приоритет всегда будет у соединения с провайдером (AS1), а частный канал между двумя клиентами будет действовать только как резервный. Если AS2 позволяет работать с частичной маршрутизацией, то в ней можно установить маршруты по умолчанию к узлу провайдера (AS1) и к AS3. Установив соответствующие локальные предпочтения, вы обеспечите направление всего трафика провайдеру. Если же AS2 работает по схеме полной маршрутизации, получая все маршруты от провайдера, и частичной маршрутизации с AS3, то в ней может оставаться маршрут по умолчанию в AS3, на случай, если канал с провайдером выйдет из строя. В AS2 маршруту по умолчанию, полученному от AS3, не следует задавать более низкое локальное предпочтение, чем полным маршрутам, полученным от провайдера.

Частный канал используется как основной для обмена трафиком между AS2 и AS3

На рис. 7.19 представлена ситуация, когда канал между AS2 и AS3 используется как основной для обмена трафиком между локальными сетями и клиентами AS2 и AS3. Для остального трафика будет использоваться канал с провайдером AS1. В этой схеме оба канала (канал с провайдером и частный канал) являются резервными по отношению друг к другу.

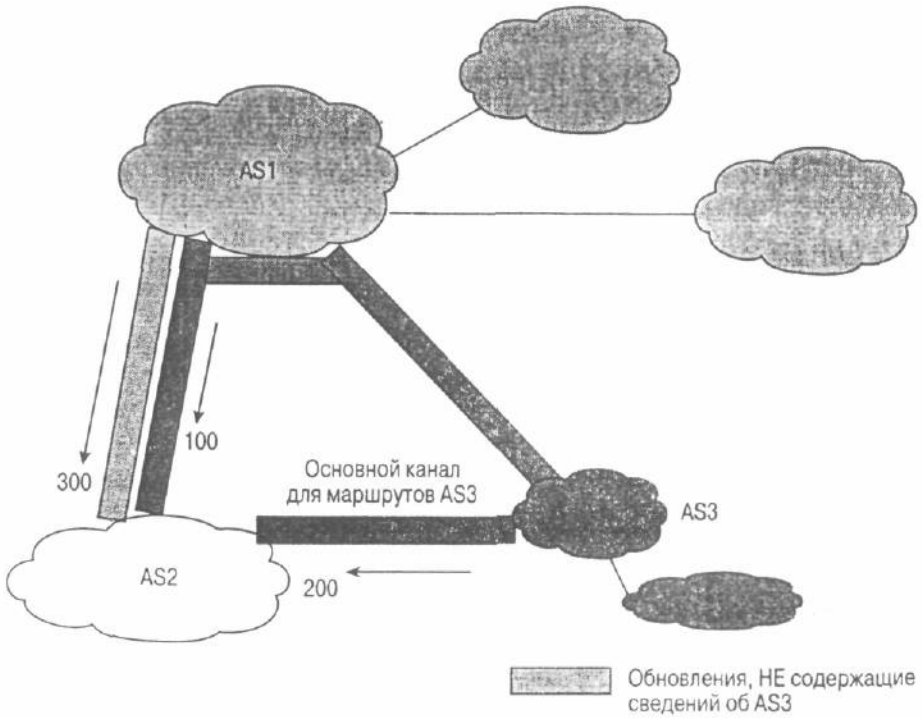


Рис. 7.19. *Частный канал используется как основной для обмена трафиком между двумя клиентскими AS*

Предположим, что в рассматриваемой ситуации (см. рис. 7.19) имеется возможность использовать как маршрутизацию по умолчанию, так и полную и частичную маршрутизацию. Рассмотрим все варианты маршрутизации применительно ко входящему и исходящему трафику для AS2. Обычно в подобных случаях системы ведут себя согласно схемам, заданным по умолчанию в протоколе BGP. Так как всегда отдается предпочтение кратчайшему маршруту (если все более приоритетные атрибуты одинаковы), то AS2 и AS3 всегда будут использовать канал между собой, чтобы попасть в сети друг друга. Ради эксперимента с правилами маршрутизации BGP мы попытаемся с помощью атрибута локальных предпочтений повлиять на исходящий трафик.

См. в главе 12 раздел "Клиенты одного провайдера с резервным каналом между ними"

Исходящий трафик AS2

Всем обновлениям маршрутов, где не упоминается AS3, следует установить значение локального предпочтения выше, чем для других обновлений. Это совершенно не зависит от того, используется ли в AS2 частичная, полная или маршрутизация по умолчанию с провайдером и AS3. В случае, показанном на рис. 7.19, всем маршрутам, которые сгенерированы не в AS3, будет присвоено значение предпочтения 300. А обновлениям, посылаемым через частный канал с AS3 (включая маршруты, которые сгенерированы самой AS3), будет задано значение предпочтения 200. Поступающие от провайдера обновления маршрутов, в которых содержатся маршруты, сгенерированные AS3, будут пересылаться на AS2, где им по умолчанию будет присвоено значение предпочтения 100. Таким образом, будет выбираться частный канал между AS2 и AS3 (так как 200 больше 100).

Весь остальной трафик AS2 будет направляться по маршрутам по умолчанию либо на узел провайдера, либо на AS3, причем предпочтение будет отдаваться маршруту на узел провайдера.

Можно также добиться, чтобы AS2 принимала только маршруты, сгенерированные AS3, и не принимала никаких маршрутов от провайдерского узла. В таком случае AS2 будет устанавливать маршруты по умолчанию и на AS3, и на узел провайдера, указывая маршрут к провайдеру как наиболее предпочтительный. Тогда трафик, направленный в AS3, будет проходить по частному каналу, а остальной — направляться в канал с провайдером (как

наилучший по умолчанию). В случае выхода из строя канала с провайдером, по умолчанию, работа будет осуществляться по частному каналу.

Входящий трафик AS2

Во всех рассмотренных случаях состояние входящего трафика практически не применялось. И вариант 4 не исключение. Вследствие меньшей длины маршрута поступающий из сети Internet трафик всегда будет проходить по каналу между провайдером и AS2. Остальной трафик, генерируемый в AS3 или ее клиентами, будет проходить по частному каналу ввиду меньшей длины маршрута. А это как раз то, чего мы и добивались.

Вариант 5: подключение клиентов к различным провайдерам с резервным каналом между ними

Иногда возникает необходимость в соединении AS через Internet посредством различных провайдеров. Однако поддержка такой схемы является очень сложной задачей. Мы пойдем даже несколько дальше и рассмотрим, как реализовать эту схему с точки зрения провайдера.

На рис. 7.20 представлены AS1, являющая клиентом ISP1, и AS2, которая является клиентом провайдера ISP2. Согласно двустороннему соглашению, AS1 и AS2 организовали частный канал между своими узлами и решили использовать его в качестве резервного на случай отказа основного канала с Internet.

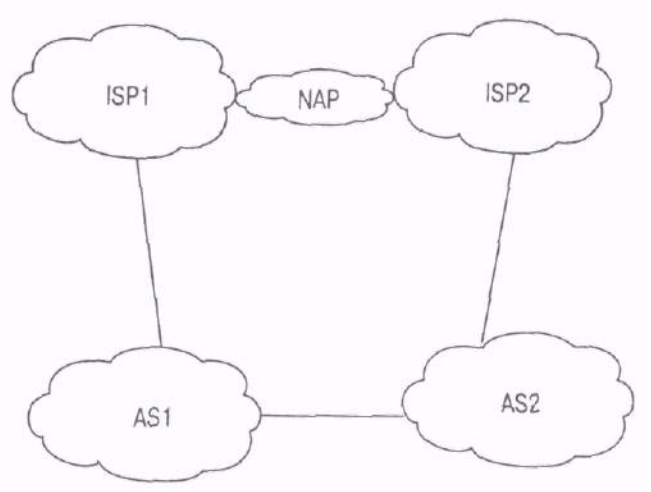


***Рис. 7.20**. Подключение клиентов к нескольким провайдерам с использованием резерв ного канала*

Как правило, отдельные AS нежелательно использовать в качестве транзитных для других AS. В рассматриваемой ситуации (см. рис. 7.20) в ASI требуется настроить маршрутизацию от провайдера ISP1 таким образом, чтобы он мог попасть в AS2 через ISP2. Точно так же в AS2 следует настроить маршрутизацию от провайдера ISP2 таким образом, чтобы он попадал в AS1 через ISP1. В этом случае для обеспечения работы резервного канала AS1 объявляет маршруты в сети AS2 провайдеру ISP1, а AS2 объявляет маршруты в сети ASI провайдеру ISP2.

См. в главе 12 раздел "Клиенты различных провайдеров с резервным каналом между собой"

Распределение ролей между каналами проводится здесь так же, как и в предыдущем случае, описанном в разделе "Вариант 4: подключение клиентов к одному провайдеру с резервным каналом между ними". Частный канал может выступать только как резервный либо использоваться для обмена трафиком между двумя клиентскими узлами.

Условие, чтобы провайдером не использовался узел одного клиента для доставки трафика другому клиенту, выполнить довольно сложно. Провайдеру JSPI придется

установить более высокие локальные предпочтения для маршрутов в AS2, поступающих от провайдера ISP2, чем для маршрутов в AS2, поступающих из AS1. Таким образом, при нормальных условиях провайдер ISP1, чтобы достичь AS2, будет использовать соединение с провайдером ISP2. Те же условия действительны и для ISP2 в отношении AS1.

Провайдеры склонны минимизировать усилия по управлению своей сетью насколько это возможно. Когда провайдер имеет дело с сотнями, а иногда и тысячами клиентов, установка локальных предпочтений вручную весьма обременительное занятие. Кроме того, провайдеры часто устанавливают правила маршрутизации на основе номеров AS, а не на базе отдельных сетей.

Для внедрения нужных правил маршрутизации вы можете использовать два подхода. Первый — заключается в строгой координации между провайдерами и их клиентами. Он основан на работе с сообществами BGP. Второй подход — управление атрибутом AS_PATH. Его намного легче реализовать, но не все производители оборудования и программного обеспечения поддерживают работу с ним. Управление атрибутом AS_PATH также может потребовать координации между провайдером и клиентами, если на узле провайдера используются фильтры атрибута AS__PATH.

Подход с использованием атрибута COMMUNITY

Подход с использованием атрибута COMMUNITY доказал свою высокую эффективность. Провайдеры требуют преобразования значений атрибута COMMUNITY в соответствующие значения локальных предпочтений (как оговорено в RFC 1998). Провайдер автоматически будет назначать обновлениям маршрутов, поступающим от клиентов, соответствующие значения локальных предпочтений.

Чтобы обеспечить в этой схеме определенный уровень управления, нужно принять только маршрутизацию и правила от провайдера ISP1. То же самое будет справедливо и в отношении ISP2. Поток трафика в случае, приведенном на рис. 7.21, можно разбить как минимум на три части.

В зависимости количества соединений между клиентом и провайдером, можно разбивать поток трафика и на большее количество частей, но приведенный пример иллюстрирует основные три формы следования трафика.

Образцы потоков, с точки зрения ISP1, выглядят следующим образом.

- **Образец 1** - маршруты, сгенерированные клиентской AS1, или локальные маршруты клиента.
- **Образец 2** - транзитные маршруты через AS1. Эти маршруты поступают от AS2 и включают в себя все остальные маршруты, которые AS2 получает от 1SP2. Эту информацию использует ISP1 для того, чтобы попасть в AS2 через AS1 в случае отказа канала между провайдеро*Г15Р2 и AS2. Этот образец трафика относится к так называемым транзитным клиентским маршрутам.
- **Образец 3** -- все остальные маршруты, поступающие от ISP2, или маршруты 1SP. К ним относятся маршруты, полученные от AS2.

Разделив все маршруты по категориям, провайдер ISP1 будет назначать значения атрибутов COMMUNITY для каждого образца трафика и динамически транслировать их в локальные предпочтения. Подобное преобразование приведено в табл. 7.5.

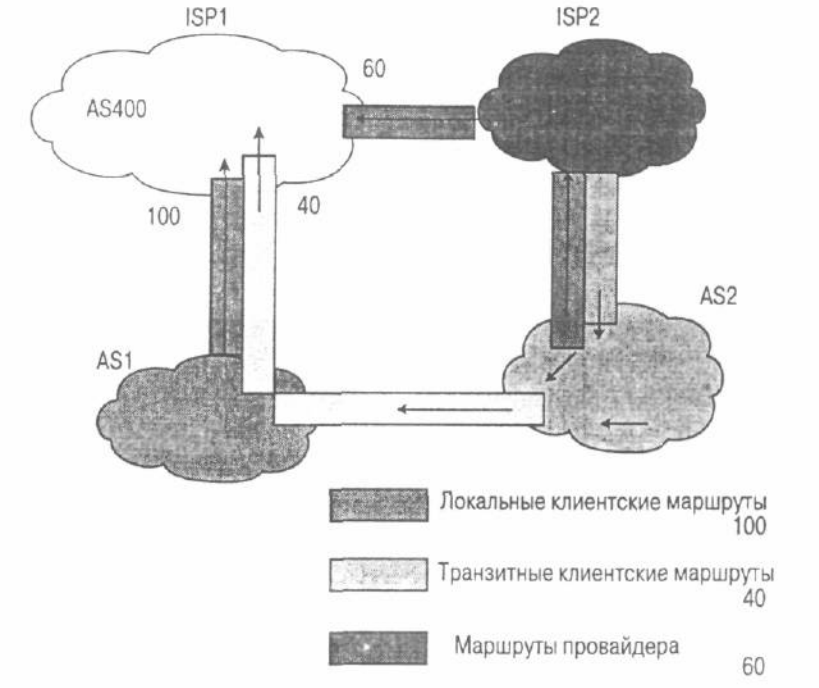


Рис. 7.21. *Использование атрибута COMMUNITY*

Таблица 7.5. *Динамическое преобразование в локальные предпочтения*

Образец	**Атрибут COMMUNITY**	**Локальное предпочтение**
Локальные клиентские маршруты	Нет	100
Транзитные клиентские маршруты	400:40	40
Маршруты ISP	400:60	60

Провайдер ISP1 будет информировать своих клиентов и взаимодействующих с ним провайдеров о том, что его локальные предпочтения динамически устанавливаются согласно табл. 7.5. Клиенты, в свою очередь, могут динамически влиять на принятие решений путем пересылки соответствующих значений атрибута COMMUNITY. Как видно из рис. 7.21, AS1 будет посылать свои локальные маршруты без атрибута COMMUNITY и транзитные маршруты со значением COMMUNITY 400:40. Провайдер ISP2 в этом случае будет свои маршруты посылать со значением COMMUNITY 400:60.

В соответствии с предпочтениями, полученными в табл. 7.5, провайдер ISP1 предпочитает работать с локальными маршрутами ASI через прямой канал (так как он имеет наивысшее значение предпочтения 100). Все остальные маршруты, включая маршруты от AS2, будут обслуживаться через ISP2 (предпочтение 60 больше 40).

См. в главе 12 раздел "Управление маршрутами с помощью атрибута COMMUNITY"

Подход с использованием атрибута AS_PATH

Этот подход в точности совпадает с тем, который уже обсуждался для случая подключения к различным провайдерам по нескольким каналам в разделе "Входящий трафик клиента (Управление атрибутом AS_PATH)". Используя этот подход, можно эффективно воздействовать на принятие провайдерами решений об использовании того или иного маршрута. На рис. 7.22 представлен пример сети, в которой при маршрутизации используется управление атрибутом AS_PATH.

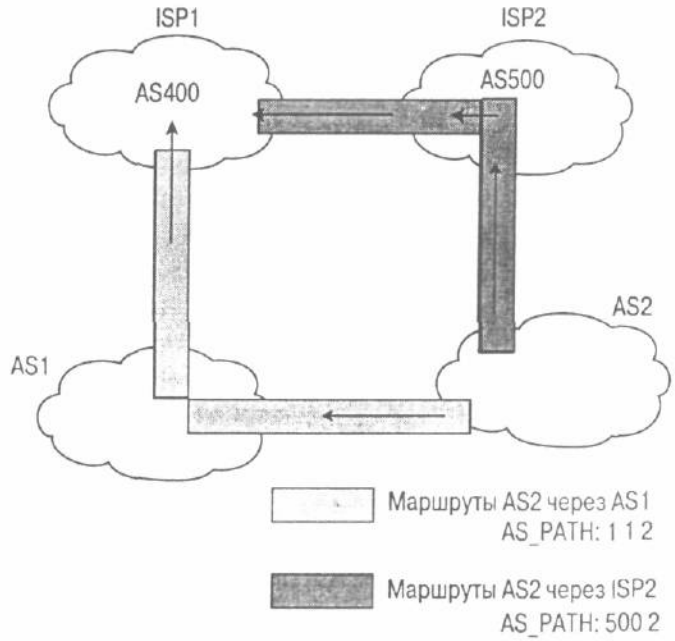


Рис. 7.22. *Пример управления атрибутом AS_PATH*

Предположим, что все атрибуты локальных предпочтений имеют значение по умолчанию, т.е. одинаковы и не "перекрывают" действие атрибута AS_PATH. В этом случае атрибут AS_PATH будет изменяться таким образом, чтобы ISP1 использовал прямое соединение с AS1 для трафика, адресованного в AS1, и прямое соединение с ISP2 для трафика, адресованного ISP2. Причем эти решения будут приняты на основании кратчайшего маршрута (наименьшего AS_PATH).

Для трафика, предназначенного AS2, у ISP1 есть два равнозначных маршрута: через ISP2 и AS1. Значение AS_PATH у провайдера JSP1 через AS1 для AS2 равно 1 2, а для маршрута через ISP2 — 500 2, т.е. они имеют одинаковую длину.

Чтобы повлиять на принятие провайдером ISP1 решения, в AS1 необходимо увеличить длину AS_PATH при объявлении маршрутов ISP1 путем добавления дополнительного номера AS в список AS_PATH. Как правило, AS1 просто повторно указывает собственный номер AS. Новый атрибут AS_PATH для провайдера ISP1 от AS2 будет равен 1 1 2, т.е. длиннее маршрута через ISP2 500 2. В результате, чтобы попасть в AS2, провайдер ISP1 будет использовать ISP2.

См. в главе 12 на раздел "Управление маршрутами с помощью атрибута AS_PATH"

Забегая вперед

Овладение маршрутизацией на границах домена позволяет получить полный контроль над трафиком, поступающим в автономную систему и покидающим ее. Еше один элемент головоломки-- течение трафика внутри AS. Не все маршрутизаторы внутри одной AS поддерживают работу по протоколу BGP. IGP-маршрутизаторы обычно не могут переносить информацию обо всех маршрутах в сети Internet из-за недостаточного объема памяти. Для обеспечения соединений между внутренними маршрутизаторами и пунктами назначения вне AS обычно используют маршрутизацию по умолчанию. Однако с использованием маршрутов по умолчанию появляется угроза образования петель маршрутизации, если имеются конфликты между правилами работы по BGP и IGP. В следующей главе мы рассмотрим сопряжение правил маршрутизации BGP с использованием IGP-маршрутов по умолчанию. В ней также будет обсуждаться использование правил маршрутизации для полного контроля над системой маршрутизации на основе IP-адресов источников вместо традиционной модели с использованием адресов пунктов назначения.

Часто задаваемые вопросы

В — Я *задал статически маршрут по умолчанию в направлении моего провайдера, указав сеть, сведения о которой я получил по BGP. Что может произойти, если эта сеть то появляется, то исчезает?*

О — В этом случае и ваш маршрут по умолчанию будет появляться и исчезать. Вот почему **не** следует указывать в качестве маршрута по умолчанию маршрут в определенную подсеть. Всегда старайтесь указывать объединенный маршрут или маршрут в суперсеть, так как маловероятно, что они будут часто выключаться.

В — У *меня есть выбор: указать маршрут по умолчанию 0/0 по BGP или задать статический маршрут по умолчанию. Что предпочесть?*

О — Для граничного маршрутизатора оба метода одинаковы, если только объединенный маршрут, на который вы указываете, является стабильным. С другой стороны, после того, как вы приняли маршрут 0/0 по BGP, он сразу же распространяется между всеми вашими IBGP-маршрутизаторами, и существует вероятность того, что вы перешлете его другим EBGP-узлам. Задав статически маршрут по умолчанию, вы получаете более полный контроль над ситуацией.

В — *Моя AS подключена к двум провайдерам* — в *Сан-Франциско и в Нью-Йорке. Мне необходимо, чтобы весь трафик в направлении узла в Сан-Хосе и от него передавался через соединение с провайдером в Сан-Франциско. Весь остальной трафск я хочу направить через канал с провайдером в Нью-Йорке. Как обеспечить работу подобной схемы?*

О — Так как вы собираетесь работать с двумя разными провайдерами, то придется использовать MED. При этом следует прибегнуть к манипулированию атрибутом AS_PATH (или воспользоваться методами, предложенными в RFC 1998) для влияния на входящий трафик и изменением локальных предпочтений для исходящего трафика. Для входящего трафика в направлении Сан-Хосе вы можете изменить длину AS_PATH, сделав ее больше для всех маршрутов, объявляемых по каналу с Нью-Йорком. Основная проблема — ваш исходящий трафик. Если вам доподлинно известно, с какими сетями чаще всего работают пользователи в Сан-Хосе, то вы можете задать для них на выходе из узла в Сан-Франциско наиболее подходящие значения локальных предпочтений. Если пользователям с узла в Сан-Хосе нужно будет попасть в какую-либо сеть, то установка наиболее подходящего локального предпочтения для соединения с провайдером в Сан-Франциско приведет к тому, что весь исходящий трафик будет проходить именно через это соединение. Однако это не соответствует вашему желанию, чтобы весь остальной трафик направлялся через соединение с провайдером в Нью-Йорке.

Еще один путь реализации вашего сценария— использование специфических правил маршрутизации, согласно которым маршрутизатор может отслеживать адреса источников и соответствующим образом направлять трафик по тому или иному соединению. Этот процесс подробно описан в главе 8, "Управление маршрутизацией в автономной системе".

В — Я *добавляю номера AS к своим маршрутам, чтобы распределить свой трафик. Но я не вижу никакого эффекта от этого. Почему?*

О — Помните о том, что обновления ваших маршрутов распространяются между несколькими провайдерами. Любой провайдер на пути к пункту назначения, может использовать локальные предпочтения, корректируя таким образом исходную длину вашего маршрута. Проконсультируйтесь на этот счет со своим провайдером.

В — *Должен ли я устанавливать правила работы по BGP? Почему я не могу предоставить протоколу BGP самостоятельно определять правильный маршрут?*

О — На самом деле нет необходимости устанавливать правила маршрутизации. Хотя следует помнить, что в протоколе BGP не принимается во внимание скорость каналов и ваши требования к траектории трафика. Если вас удовлетBGPяет траектория вашего трафика, то нет необходимости вносить изменения в атрибуты.

Ссылки

1 RFC 826, "Address Resolution Protocol (ARP)," www.isi.edu/in-notes/rfc826.txt

2 RFC 1998, "An Application of the BGP Community Attribute in Multi-home Routing," www.isi.edu/in-notes/rfcl998.txt

Ключевые темы этой главы:

- **Взаимодействие маршрутизаторов, не поддерживающих BGP, с маршрутизаторами под управлением BGP.** Дается краткий обзор методов, с помощью которых маршрутизаторы в AS без поддержки BGP могут общаться с внешним миром.
- **Маршруты по умолчанию внутри AS: правила маршрутизации по основному и запасному маршруту.** Рассматриваются различные методы**,** с помощью которых можно избежать образования петель маршрутизации, если таблица маршрутов внутри AS конфликтует с требованиями по обеспечению основного и запасного маршрутов на внешнюю AS. :
- **Маршруты по умолчанию внутри AS: другие правила маршрутизации для работы по BGP.** Представлены правила маршрутизации, отличные от правил маршрутизации по основному и запасному маршруту, которые также могут стать причиной образования петель маршрутизации внутри AS.
- **Маршрутизация по правилам.** Дается определение и приводятся примеры управления маршрутами на основе IP-адресов источников или на основе и адресов источников и адресов пунктов назначения.

Глава 8. Управление маршрутизацией в автономной системе

В предыдущей главе мы подробно рассматривали вопросы взаимодействия между несколькими AS и способы "обеспечения избыточности, симметрии и распределения нагрузки с помощью атрибутов BGP. Мы коснулись поведения граничных BGP-маршрутизаторов, через которые организовывалось соединение между различными AS. Маршрутизаторы у провайдеров Internet обычно работают по протоколу BGP с использованием отдельных узлов, где поддерживаются только протоколы внутреннего шлюза (Interior Gateway Protocols — IGP). У клиентов в большинстве случаев имеется всего один-два маршрутизатора с BGP, а основную массу составляют внутренние маршрутизаторы, работающие с протоколами IGP, которые обеспечивают маршрутизацию по умолчанию в направлении BGP-маршрутизаторов. Таким образом, очень важно, чтобы правила маршрутизации для BGP не противоречили правилам маршрутизации внутри AS. Конфликт правил маршрутизации может привести к образованию петель маршрутизации. В этой главе мы как раз и будем обсуждать взаимодействие IGP-маршрутов внутри одной AS. Также вашему вниманию будут предложены различные параметры для управления маршрутами посредством правил маршрутизации.

Взаимодействие маршрутизаторов, не поддерживающих BGP, с маршрутизаторами под управлением BGP

,

Маршрутизаторы внутри AS без поддержки протокола BGP могут общаться с внешним миром следующими способами:

- преобразованием BGP в IGP;
- использованием маршрутов по умолчанию внутри AS.

Преобразование из BGP в IGP

Полное преобразование всех BGP-маршрутов в IGP не рекомендуется. Проделав подобную операцию, вы создадите избыточную нагрузку на систему маршрутизации, которая перегрузит любой протокол IGP. Протоколы внутренней маршрутизации не предназначены для обслуживания большего количества сетей, чем может теоретически поддерживаться в пределах одной AS. Разве что в качестве дополнения они способны обслуживать ограниченное число внешних маршрутов, информация о которых поступает от других IGP.

Это вовсе не означает, что BGP-маршруты вообще нельзя преобразовывать в IGP. В

зависимости от количества BGP-маршрутов и того, насколько критично преобразование их в IGP, трансляция частичных BGP-маршрутов может быть вполне оправданной. Если вы будете следовать этим курсом, то вам следует внимательно выполнять операции по преобразованию маршрутов в IGP. Описание подробностей механизма преобразования BGP в IGP не входит в перечень вопросов, рассматриваемых в этой книге. Однако есть несколько моментов, которые необходимо принимать во внимание, — размер доступной памяти, свободные ресурсы процессора для вычисления маршрутов и обработки обновлений маршрутов, нагрузку на канал вследствие управления маршрутизацией, влияние на конвегренцию, ограничения отдельных протоколов IGP и топологию конкретной сети.

Преобразование частичных BGP-маршрутов в IGP в заданных точках AS может помочь направить соответствующий исходящий трафик в определенные точки выхода. Исходящий трафик по другим маршрутам по-прежнему будет направляться по умолчанию на BGP-маршрутизаторы. Хотя преобразование BGP-маршрутов в IGP в некоторых случаях кажется довольно оптимальным решением, оно обладает рядом серьезных недостатков. Так, если протоколы IGP по сути являются классовыми (такие как RIP-1 или протокол внутреннего шлюза (Interior Gateway Protocol — IGRP), то вся информация о блоках бесклассовой междоменной маршрутизации (CIDR) будет теряться. Еше одной проблемой является потенциальная нестабильность преобразуемых BGP-маршрутов, которая передается при преобразовании IGP-маршрутам. В основном отказы сетей происходят именно по причине отказов IGP-маршрутов в результате флуктуации, вызванных большим количеством внешних маршрутов.

Использование в AS маршрутов по умолчанию

Более практичным решением для взаимодействия с внешним миром маршрутизаторов, не поддерживающих BGP, является использование существующих внутри AS маршрутов по умолчанию к маршрутизаторам, выполняющим функции ближайшего внешнего шлюза, через который вы можете выйти за пределы локальной AS. Маршрут по умолчанию может быть получен AS от каждого из граничных маршрутизаторов автономной системы. Каждый IGP-маршрутизатор может получать маршрут по умолчанию от одного или нескольких маршрутизаторов. Каждый IGP-маршрутизатор выбирает наилучший маршрут в пункт назначения, который находится за пределами AS, на основе внутреннего весового коэффициента или метрики маршрута по умолчанию. После того как трафик доходит до BGP-маршрутизаторов, он распространяется согласно наилучшим маршрутам, выделенным в BGP. На рис. 8.1 представлена схема взаимодействия маршрутизаторов без BGP внутри одной AS. Как видите, они используют маршруты по умолчанию, чтобы достичь ближайшего BGP-маршрутизатора.

Здесь маршрутизаторы RTC и RTD являются граничными BGP-маршрутизаторами, которые посылают в AS1 маршрут по умолчанию вида 0/0. Маршрутизатор RTB представляет собой внутренний транзитный маршрутизатор, который полностью поддерживает IBGP и взаимодействует с маршрутизаторами RTC и RTD. Внутренние маршрутизаторы без BGP, такие как RTA, могут получать сведения о маршруте по умолчанию от различных источников с помощью протоколов IGP, при этом они будут использовать маршрут по умолчанию с наименьшей метрикой IGP. На рис. 8.1 маршрутизатор RTA получает маршрут 0/0 от RTB с метрикой 10, от RTE - с метрикой 20 (RTA-RTB: 10 + RTB-RTE: 10) и or RTF - с метрикой 30 (RTA-RTF: 10 + RTF-RTG: 10 + RTG-RTB или RTC: 10). В этой ситуации маршрутизатор RTA воспользуется каналом с маршрутизатором RTB, так как последний имеет наименьшую внутреннюю метрику (10). После того как трафик поступает на маршрутизатор RTB, вступает в действие таблица BGP-маршрутов, по которой определяются окончательные маршруты в пункт назначения, которые находятся в других AS.

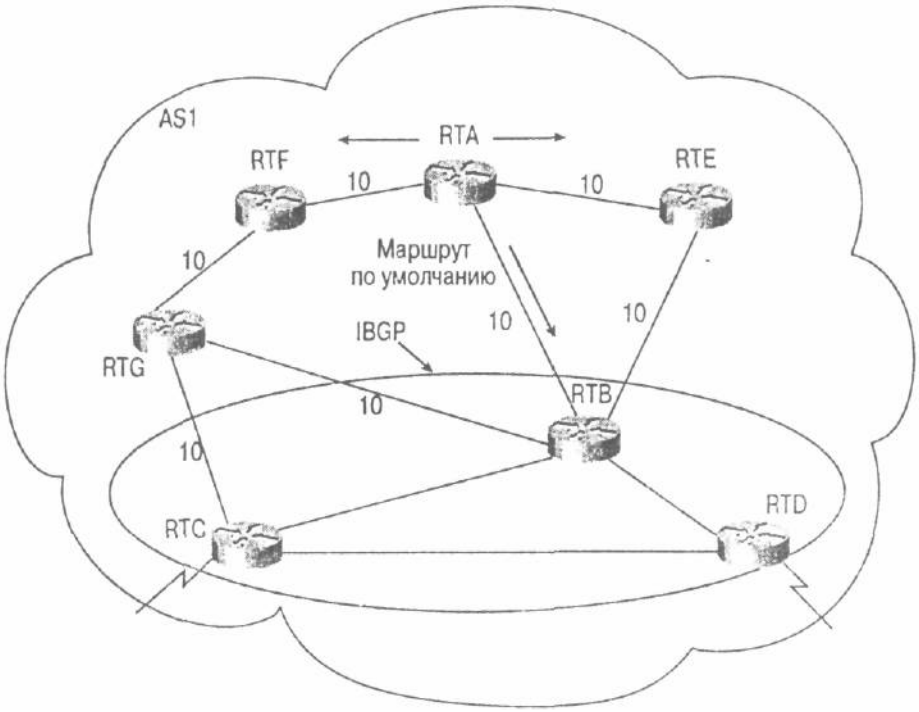


Рис. 8.1. *Пример использования маршрутов по умолчанию*

См. в главе 12 раздел "Установка маршрутов по умолчанию"

Работа по протоколу IBGP внутри AS позволяет направлять трафик в точки выхода из AS и обрабатывать транзитный трафик в случаях, рассмотренных нами ранее (таких как поддержка резервного канала для другой AS при выходе из строя ее основного канала). Кроме того, большинство методов обеспечения симметрии, о которых мы говорили в предыдущей главе, не могут быть реализованы, если BGP-маршрутизаторами не поддерживается IBGP.

Конфликт правил маршрутизации BGP с внутренними маршрутами по умолчанию

В зависимости от физической топологии AS и установленных правил маршрутизации BGP могут возникать определенные конфликтные ситуации. Так. трафик внутри AS. следующий согласно маршрутам по умолчанию протокола IGP на граничный BGP-маршрутизатор, может образовать петлю, если граничные BGP-маршрутизаторы имеют правила маршрутизации, противBGPечащие правилам для протокола IGP, что в результате приведет к возвращению трафика в исходную точку.

См. в главе 12 раздел "Маршрутизация по правилам"

В этом разделе мы рассмотрим ситуации, приводящие к образованию петель в маршрутизации, а также возможные пути решения этих проблем. Мы рассмотрим следующие два случая.

- Маршруты по умолчанию внутри AS и их взаимодействие с правилами маршру

тизации BGP с использованием основного и резервного канала.

- Маршруты по умолчанию внутри AS и их взаимодействие с другими правилами маршрутизации BGP.

Маршруты по умолчанию внутри AS: правила маршрутизации BGP по основному и запасному маршруту

Рассмотрим вариант организации маршрутизации, представленный на рис. 8.2. Здесь AS1 подключается к сети Internet по двум каналам. На маршрутизаторе RTC, который находится в Сан-Франциско, поддерживается протокол внешнего граничного шлюза (Exterior Border Gateway Protocol — EBGP) с одним провайдером, в то же время в Нью-Йорке через маршрутизатор RTD организовано еще одно соединение также с использованием протокола EBGP. Внутри AS маршрутизаторы RTC и RTD общаются по протоколу IBGP. Однако, как видите, они не имеют между собой прямого физического соединения, т.е. весь трафик между RTC и RTD будет передаваться через маршрутизаторы RTA и RTB.

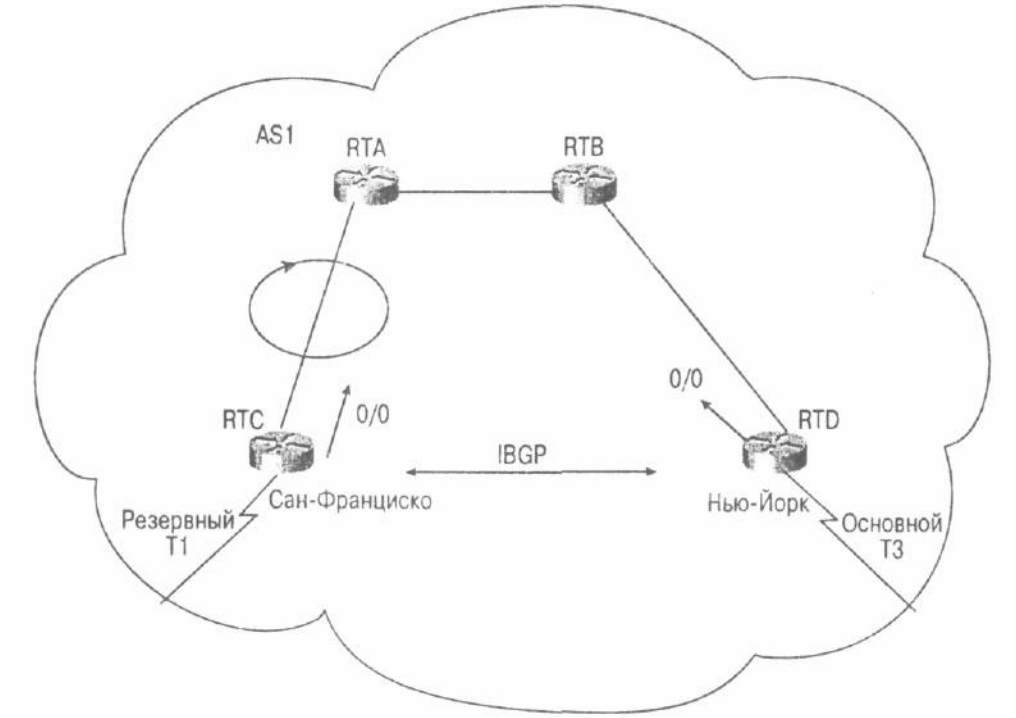


Рис. 8.2. *Образование петли маршрутизации при использовании маршрутов по умолчанию*

Предположим, что маршрутизаторы RTC и RTD принимают полные маршруты от провайдеров. Кроме того, они посылают в AS1 маршрут по умолчанию 0/0, который будет использоваться протоколом IGP. Также представим, что в AS1 требуется обеспечить схему работы с использованием основного и запасного канала, что позволяет выбрать в качестве основного канал T3 с провайдером в Нью-Йорке. Таким образом, для маршрутов, которые используют канал в Нью-Йорке, в ASI будут устанавливаться наивысшие локальные предпочтения, что и позволит использовать его как основной. Канал T1 с провайдером в Сан-Франциско будет использоваться как резервный для всего исходящего трафика, который приходит на маршрутизатор RTC. Тогда он будет направлен обратно на маршрутизатор RTD.

Маршрутизаторы RTA и RTB являются внутренними, и они не поддерживают работу по протоколу BGP. Эти маршрутизаторы обмениваются маршрутами между собой и другими маршрутизаторами внутри AS посредством одного из протоколов IGP. Они не могут работать с внешними маршрутами. Все, что они могут делать, — посылать трафик, сведения о пункте назначения которого неизвестны, согласно маршрутам по умолчанию в направлении маршрутизаторов RTC или RTD, в зависимости от того, какой из этих

маршрутов будет иметь меньшую метрику IGP. Так, трафик из внешних сетей, попав на маршрутизатор RTA, будет направлен по умолчанию на RTC, а трафик, пришедший на маршрутизатор RTB, будет возвращен на RTD.

Когда маршрутизатор RTC принимает какой-либо трафик, он изменяет его траекторию в направлении маршрутизатора RTD, так как согласно оговоренным нами правилам работы по BGP для данной AS, канал с провайдером в Нью-Йорке является основным. Ввиду того что маршрутизатор RTC не имеет прямого соединения с RTD он посылает весь трафик на маршрутизатор RTA. Маршрутизатор RTA принимает трафик и ... направляет его обратно на RTC! Как видите, мы получили петлю маршрутизации между RTA и RTC.

Для того чтобы избежать возникновения подобной ситуации при использовании маршрутов по умолчанию внутри AS с использованием схемы работы по основному и резервному каналу, могут быть использованы следующие варианты решений.

Вариант 1: управление метрикой IGP

В этом варианте мы попытаемся предотвратить образование петли, направляя весь трафик для внешних узлов через маршрут по умолчанию в направлении маршрутизатора RTD. Это можно осуществить, заставив маршрутизатор RTC распространять маршрут по умолчанию 0/0 внутри AS с очень высокой метрикой, что сделает маршрут 0/0 от маршрутизатора RTD более коротким, а следовательно, и более привлекательным для любого внутреннего маршрутизатора. Трафик в этом случае ни при каких обстоятельствах не должен поступать на маршрутизатор RTC в поисках маршрута к неизвестному пункту назначения (на внешний узел), если только не выйдет из строя канал с провайдером в Нью-Йорке.

Вариант 2: IBGP-маршрут короче IGP-маршрута

Существование более короткого маршрута между IBGP-маршрутизаторами дает уверенность в том, что трафик не будет возвращен обратно на IGP-маршрутизаторы, чтобы попасть в пункт назначения. Это необходимо только в тех случаях, если правила маршрутизации по BGP требуют перенаправлять трафик с одного BGP-маршрутизатора на другой. Подобные ситуации имеют место когда IBGP-маршрутизатор не имеет внешнего соединения, по которому он мог бы направить трафик. Если же у него есть подобное внешнее соединение, то оно не используется в качестве наилучшего маршрута (см. случай с маршрутизатором RTC на рис. 8.2).

На схеме, представленной на рис. 8.2, избежать образования петли можно, если граничные маршрутизаторы RTC и RTD, которые поддерживают работу по протоколу IBGP, также совместно используют один и тот же физический сегмент, такой как канал связи. В таком случае трафик, поступающий от маршрутизатора RTA на RTC, будет направлен по физическому каналу (он не показан на рис. 8.2), который обеспечивает в данном случае более короткий прямой маршрут между маршрутизаторами RTC и RTD, а также вносит элемент избыточности, так как по нему будет проходить весь трафик в случае выхода из строя всех остальных каналов.

Вариант 3: поддержка BGP в транзитных маршрутизаторах

Поддержка протокола BGP во всех транзитных маршрутизаторах позволяет направлять трафик за пределы AS, как только он поступит на эти маршрутизаторы. Как показано на рис. 8.2, если бы маршрутизаторами RTA и RTB совместно с RTC и RTD поддерживался протокол IBGP, то для всего трафика, поступающего на эти маршрутизаторы, можно было бы элегантно определить подходящий маршрут за пределы сети. Хотя легче всего просто нагрузить маршрутизаторы RTA и RTB полными BGP-маршрутами, вы можете также реализовать обработку на этих маршрутизаторах частичных маршрутов и/или BGP-маршрутов по умолчанию, что потребует дополнительной конфигурации. Отметим, что, хотя AS1 может и не являться транзитной AS, маршрутизаторы RTA и RTB *могут* использоваться для обработки трафика между граничными маршрутизаторами. Внутренние IGP-маршруты могут использовать облако IBGP, чтобы выйти во внешний мир, как показано на рис. 8.2.

Вариант 4: кто и как генерирует маршрут по умолчанию?

В этом варианте избежать образования петли можно, если основной маршрутизатор генерирует IGP-маршрут по умолчанию, а второй маршрутизатор не делает этого. В этом случае маршрутизатор RTD будет генерировать IGP-маршрут по умолчанию 0/0, а RTC не будет. Тогда весь трафик будет следовать согласно маршруту по умолчанию на маршрутизатор RTD.

Однако такая схема работает только в нормальных условиях, а что делать, если основной канал или маршрутизатор выйдут из строя[9] Если канал с провайдером в Нью-Йорке выйдет из строя, IGP-маршрутизаторы потеряют единственный маршрут по умолчанию 0/0. А так как маршрутизатор RTC не генерирует маршрута по умолчанию, то трафик, адресованный узлам за пределами AS, будет потерян, так как маршрутизаторы не будут "знать" что с ним делать.

Идеальным решением в этом случае является анонсирование маршрута по умолчанию от RTC только тогда, когда канал в Нью-Йорке выходит из строя. Если канал в Нью-Йорке выходит из строя, то маршрутизатор RTD должен прекратить объявление IGP-маршрута по умолчанию, а RTC, наоборот, должен начать эту процедуру. Для реализации такого механизма маршрутизаторы должны принять на себя следующие обязательства.

- **BGP-маршрутизатор должен прекратить объявление IGP-маршрута по умолчанию, если его внешний канал выходит из строя.** Это требование довольно легко реализовать, если в IGP поддерживается преобразование внешнего маршрута по умолчанию 0/0. Когда маршрут 0/0 прекращает свое существование, то IGP-маршрут по умолчанию также исчезает вместе с ним. Возможность преобразования и поведение маршрута по умолчанию зависит от протокола IGP, который вы используете, и от производителя конкретного сетевого оборудования. Так, например, способ преобразования в оборудовании компании Cisco может отличаться от способов, реализованных в устройствах других производителей.
- **BGP-маршрутизатор должен объявлять IGP-маршруг по умолчанию только в том случае, если маршрут по умолчанию будет использовать локальный внешний канал. Таким** образом, это правило обязывает любой маршрутизатор прекратить генерирование маршрута по умолчанию, если маршрут по умолчанию, который он предпочитает использовать, приходит от другого маршрутизатора внутри AS, а не из другой AS. То есть, если второй маршрутизатор предпочитает использовать маршрут по умолчанию, поступивший от другого маршрутизатора внутри AS, то это означает, что основной канат исправен и для работы следует использовать именно его. Однако когда основной канал выходит из строя, второй маршрутизатор воспользуется IGP-маршрутом по умолчанию из другой AS и самостоятельно объявит его. Эту ситуацию дня лучшего понимания лете рассмотреть на примере.

См. в главе 12 раздел "Установка маршрутов по умолчанию"

Следующие два примера подчеркивают отличия между маршрутами по умолчанию, сгенерированными протоколами RIP и OSPF на базе оборудования компании Cisco.

Маршрут по умолчанию, сгенерированный протоколом RIP

На рис. 8.3 маршрутизаторы RTC и RTD могут получить сведения о маршруте по умолчанию 0/0 или статически сконфигурировать его в направлении соответствующих узлов провайдеров. При нормальных условиях работы маршрутизатор RTD автоматически (или через управляемое преобразование) вставляет маршрут 0/0 в протокол RIP. Маршрутизатор RTC обнаруживает присутствие маршрута по умолчанию, поступающего от RTD, и прекращает генерировать свой собственный маршрут по умолчанию. Весь трафик направляется на маршрутизатор RTD.

В случае выхода из строя канала с провайдером в Нью-Йорке, маршрутизатор RTD прекращает генерировать RIP-маршрут по умолчанию. Маршрутизатор RTC обнаруживает

отсутствие RIP-маршрута 0/0 и подставляет собственный маршрут по умолчанию. Обратите внимание, что RTC получает маршрут по умолчанию 0/0 посредством EBGP (от внешней сети, подключенной к этому маршрутизатору), RIP и, возможно, IBGP, если в течение IBGP-сеанса маршрутизатор RTD передает сведения о маршруте 0/0. Вследствие более высокого значения локального предпочтения маршрута через RTD, маршрутизатор RTC выбирает IBGP-маршрут 0/0. Так как административная дистанция на RTC (200) больше, чем административная дистанция RIP, которая равна 120 (см. табл. 6.1), то будет выбираться RIP-маршрут по умолчанию 0/0 с меньшей административной дистанцией.

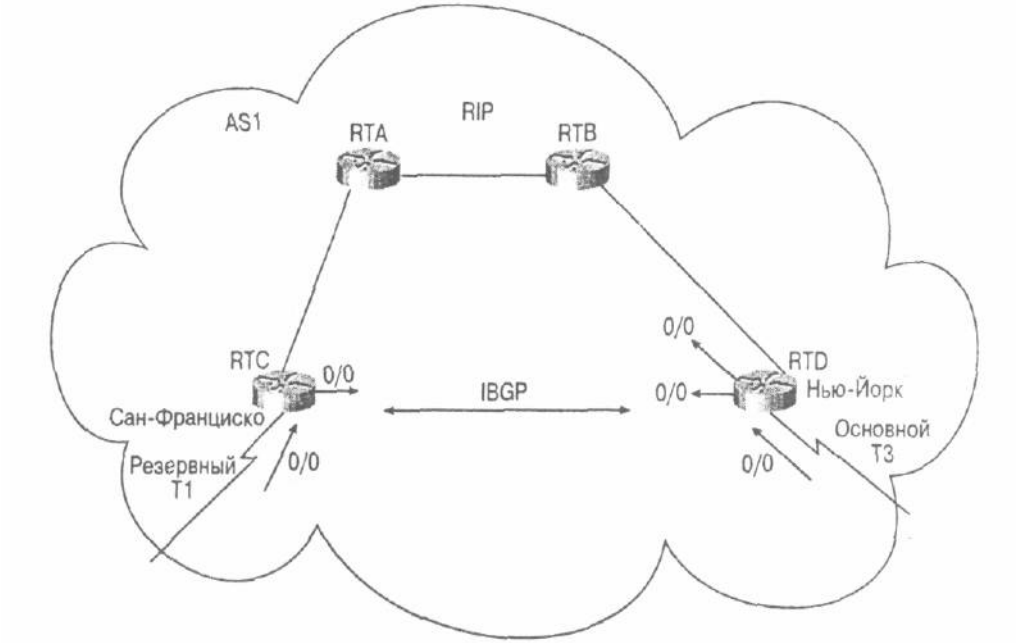


Рис. 8.3. *Преобразование маршрута по умолчанию 0/0 в протоколе RIP*

Маршрут по умолчанию, сгенерированный протоколом OSPF

Протокол OSPF ведет себя совершенно по-другому. BGP-маршрут 0/0 не может быть напрямую преобразован в OSPF. В протоколе OSPF предусмотрены различные приемы, позволяющие в любое время сгенерировать маршрут 0/0 в OSPF, поэтому даже лучше, если обнаружено его присутствие в таблице IP-маршрутов. Теперь рассмотрим это все на примере, представленном на рис. 8.4.

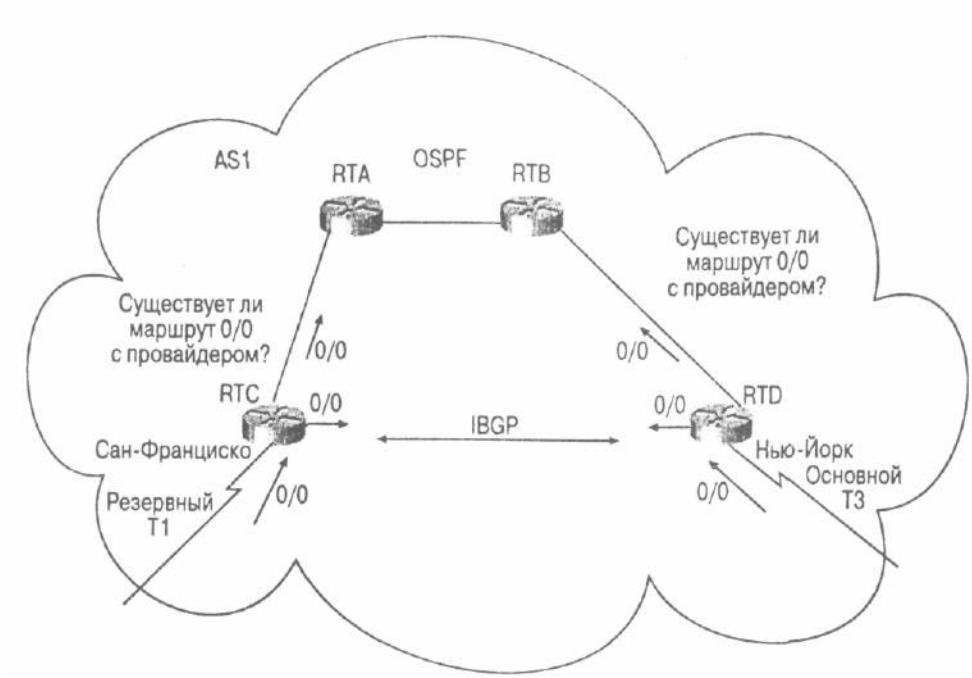


Рис. 8.4. *Преобразование маршрута по умолчанию 0/0 в протоколе OSPF*

Маршрутизаторы RTD и RTC получают EBGP-маршрут 0/0 или самостоятельно статически указывают маршрут по умолчанию на узел провайдера. Если RTD и RTC сконфигурированы так, что маршрут по умолчанию вида 0/0 преобразуется для работы по OSPF при появлении в их маршрутных IP-таблицах, модель с использованием основного и резервного каналов больше не будет работать. Это происходит потому, что и RTD и RTC получают IBGP-маршрут 0/0 друг от друга. То есть маршрутизатор RTC всегда будет вставлять префикс 0/0 в OSPF, независимо от состояния канала с провайдером в Нью-Йорке. К тому же, в отличие от схемы работы по протоколу RIP, маршрутизатор RTC будет игнорировать OSPF-маршрут по умолчанию, поступающий от RTD, так как он сконфигурирован самостоятельно генерировать собственный маршрут по умолчанию.

Чтобы исправить эту ситуацию необходимо прибегнуть к дальнейшему конфигурированию маршрутизаторов RTD и RTC. Их нужно настроить таким образом, чтобы они генерировали OSPF-маршрут 0/0, только если их собственный маршрут по умолчанию указывает на узел провайдера.

См. в главе 12 раздел "Применение OSPF в качестве протокола IGP"

По существу, если маршрутизатор RTD среди других выбирает маршрут по умолчанию, который указывает на узел провайдера, то он преобразует его в OSPF-маршрут. Точно так же маршрутизатор RTC предпочитает маршрут по умолчанию, который указывает на узел его провайдера и преобразует маршрут 0/0 в OSPF-маршрут.

В подобной модели происходят следующие процессы.

- Нормальная работа обеспечивается, когда канал с провайдером в Нью-Йорке исправен.
- Маршрутизатор RTD предпочитает внешний маршрут по умолчанию всем другим и преобразует маршрут 0/0 в OSPF-маршрут.
- Маршрутизатор RTC, работающий по EBGP (от своего провайдера), IBGP и OSPF получает маршрут по умолчанию 0/0. Он игнорирует маршрут по умолчанию, полученный по OSPF, как мы уже отмечали ранее.
- Маршрутизатор RTC использует маршрут 0/0, полученный от RTD no IBGP, так как последний имеет более высокое локальное предпочтение.
- Поскольку маршрут 0/0 получен маршрутизатором RTC не от провайдера, маршрутизатор не анонсирует никакого OSPF-маршрута по умолчанию.
- Если канал в Нью-Йорке выходит из строя, то маршрутизатор RTD прекращает получать маршрут 0/0 от своего провайдера, но продолжает принимать маршрут 0/0 по IBGP и генерировать его для OSPF, так как он не получен от провайдера.
- Маршрутизатор RTC прекращает получать маршрут 0/0 по IBGP и использует маршрут по умолчанию через своего провайдера. Затем, он начинает анонсировать маршрут по умолчанию 0/0 в протокол OSPF.

Маршруты по умолчанию внутри AS: другие правила маршрутизации в BGP

Итак, как вы уже убедились, образование петель может произойти в любой момент, если IGP-маршруты по умолчанию вступают в конфликт с правилами маршрутизации, принятыми в BGP. В схемах с использованием основного и резервного каналов вы можете выбирать какой из граничных маршрутизаторов должен генерировать маршрут по умолчанию, так как вы приняли решение, какой из них будет основным маршрутизатором для всего внешнего по отношению к AS трафика. В некоторых случаях на правила маршрутизации для вашей AS могут налагаться определенные ограничения, обусловленные внешними факторами. В других случаях нор-мальная маршрутизация по IBGP/EBGP может

сделать точку выхода из AS неопределенной, что приведет к конфликту между маршрутами по умолчанию.

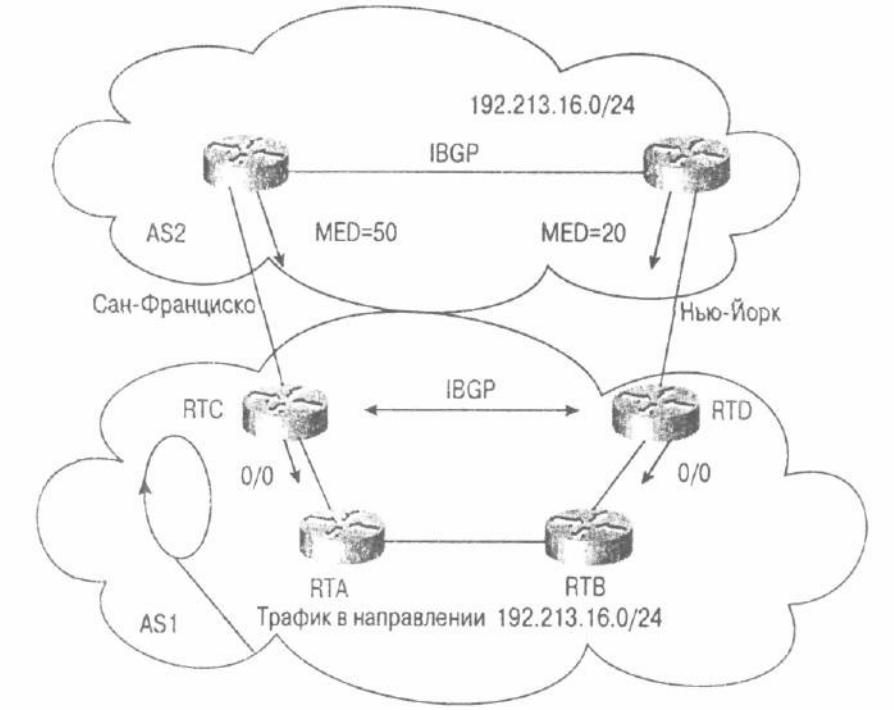


Рис. 8.5. *Правила маршрутизации, подверженные влиянию внешних факторов*

Рассмотрим рис. 8.5. Здесь AS1 подключена к провайдеру AS2 и получает от него полные или частичные маршруты в двух географических точках — Нью-Йорке и Сан-Франциско. При этом в AS1 маршруты по умолчанию объявляются и маршрутизатором RTC в Сан-Франциско и маршрутизатором RTD в Нью-Йорке таким образом, что внутренние узлы будут направлять исходящий трафик в ближайшую точку выхода из AS.

Предположим также, что в AS1 очень тонко настроен механизм подстановки маршрутов по умолчанию. Маршрутизатор в Сан-Франциско никогда не будет подставлять свой маршрут по умолчанию, если канал с провайдером вышел из строя, та же картина наблюдается и с маршрутизатором в Нью-Йорке. Все прекрасно и работает нормально до тех пор пока провайдер AS2 не начнет объявлять в направлении AS1 свои метрики (MED).

Представим, что AS2 посылает свои обновления маршрутов на ASI с определенными внутренними IGP-метриками, такими как MED. Клиентская система AS1 получает сведения об одних и тех же сетях по каналам в Сан-Франциско и в Нью-Йорке, но с различными значениями MED. Для каждой сети в протоколе BGP выбирается маршрут с наименьшей метрикой. Если, например, маршрутизатор RTC получает сведения о маршруте в сеть 192.2И.16.0/24 с MED равным 50 по каналу в Сан-Франциско и с MED 20 по каналу в Нью-Йорке, то он выберет канал в Нью-Йорке. Зто означает, что маршрутизатор RTA может следовать внутреннему маршруту по умолчанию в направлении RTC и затем будет инструктирован о необходимости переслать трафик на маршрутизатор RTD, чтобы попасть в сеть 192.213.16.0/24. Точно так же маршрутизатор RTB может использовать маршрут по умолчанию на RTD и затем перенаправить трафик на RTC. В обоих случаях это приведет к образованию петли маршрутизации.

Как видите, точка выхода для всех сетей в данном случае не может быть определена однозначно, как в случае с использованием основного и резервного каналов. Для решения этой проблемы необходимо выполнить следующие действия.

- Игнорировать MED и использовать при маршрутизации модель с основным и резервным каналами.
- Организовать более короткий маршрут непосредственно между RTC и RTD, что позволит направлять трафик между точками выхода из AS по кратчайшему пути между IBGP-маршрутизаторами.
- Включить поддержку IBGP между маршрутизаторами RTA, RTB, RTC и RTD.

Но и в других нормальных случаях могут появляться петли маршрутизации. Всегда

есть опасность появления петель, когда вы работаете с несколькими каналами и одновременно используете несколько маршрутов по умолчанию внутри AS. Если ваша сеть подключена к двум провайдерам, то. возможно, вы пожелаете использовать для одних узлов сети подключение к одному провайдеру, а для узлов, ближайших ко второму провайдеру, подключение через другого провайдера. Если у вас имеются 1GP-маршруты по умолчанию, то вы можете оказаться в ситуации, когда маршрут заканчивается в точке выхода, а обратный маршрут отсутствует.

Итак, решение проблемы образования петель маршрутизации сводится либо к более четкому определению точек выхода из AS с помощью протоколов BGP и IGP, либо к предотвращению возврата трафика между IBGP-маршругизаторами через IGP-маршрутизаторы. Чем больше мер предосторожности вы примете для обеспечения нормальной циркуляции трафика, тем меньше вероятность появления петель маршрутизации.

Маршрутизация по правилам

Маршрутизация по правилам (policy routing) является средством управления маршрутами на основе адреса отправителя (источника) или адресов источника и пункта назначения, или лишь на основе адреса пункта назначения. Маршрутизация по правилам может применяться для управления трафиком как внутри одной AS, так и между несколькими AS. Маршрутизация по правилам уходит корнями в статическую маршрутизацию. Она в основном используется, чтобы изменить нормальное поведение, диктуемое динамическими протоколами маршрутизации.

Статическая маршрутизация позволяет направлять трафик определенным узлам на основе пунктов назначения, заданных для него. Так, трафик в пункт назначения 1 может быть послан через точку A, а трафик в пункт назначения 2 — через точку B,

Маршрутизация по правилам также позволяет направлять трафик на основе адреса источника или комбинации адресов источника и пункта назначения (или даже на основе других атрибутов). Трафик, сгенерированный в сети 1, может быть передан через точку A; трафик, сгенерированный в сети 1 и адресованный сети 2, может быть передан через точку B.

См. в главе 12 раздел "Маршрутизация по правилам"

В последующих разделах мы рассмотрим маршрутизацию по правилам на основе адреса источника и на основе адресов источника и пункта назначения, а также различные способы применения этого вида маршрутизации.

Маршрутизация по правилам на основе адреса источника

Рассмотрим пример, представленный на рис. 8.6. Предположим, AS1 получила сетевые адреса от обоих провайдеров. Диапазон 10.10.10.0/24 был получен от AS3, а 11.11.11.0/24-- от AS4. В AS1 требуется обеспечить передачу всего трафика, сгенерированного сетями диапазона 10.10.10.0/24, в направлении AS3, а трафика от 11.11.11.0/24 -- на AS4, причем независимо от конечного пункта назначения. Чтобы обеспечить выполнение этих требований, в AS1 можно использовать маршрутизацию по правилам и заставить весь трафик с IP-адресами отправителя, принадлежащими к диапазону 10.10.10.0/24. направлять на узел 1.1.1.1, а трафик с IP-адресами из диапазона 11.11.11.0/24 -- на узел 2.2.2.2.

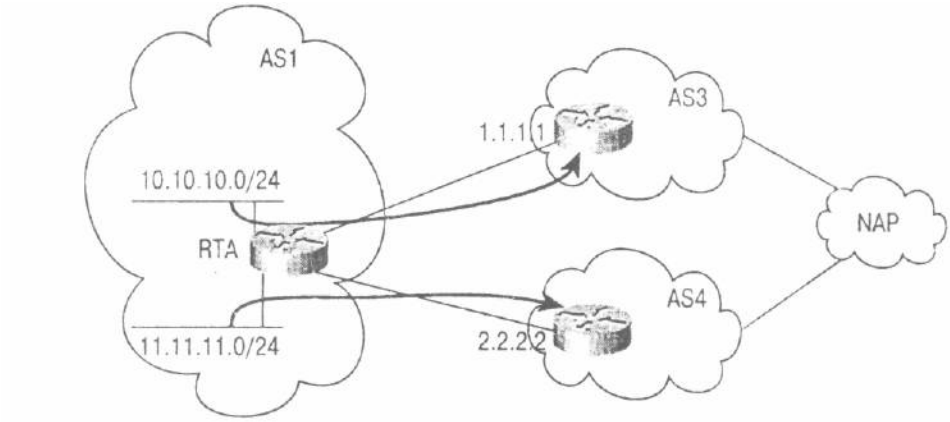


Рис. 8.6. *Вариант маршрутизации по правилам на основе адреса отправителя*

Маршрутизация по правилам на основе адресов источника и пункта назначения

Маршрутизация по правилам может также выполняться на основе комбинации адресов отправителя и получателя, как это показано на рис. 8.7.

Предположим, что маршрутизатор RTA будет передавать любой график, поступающий из сети 10.10.10.0/24 в сеть 12.12.12.12/24 в Нью-Йорке, по каналу с Сан-Франциско. Для всего трафика, поступающего из сети 10.10.10.0/24 в сеть 13.13.13.13/24 в Нью-Йорке, маршрутизатор RTA будет использовать канал с Сан-Хосе. Для различных комбинаций адресов отправителя и получателя можно сформировать правила маршрутизации. При этом для комбинации (адрес источника=10.10.10.10/24, адрес получателя 12.12.12.12/24) следующим ближайшим узлом будет задан маршрутизатор с адресом 1.1.1.1. Трафик с комбинацией адресов (адрес источника=10.10.10.10/24, адрес получателя^! 3.13.13.13/24) будет направляться на узел 2.2.2.2.

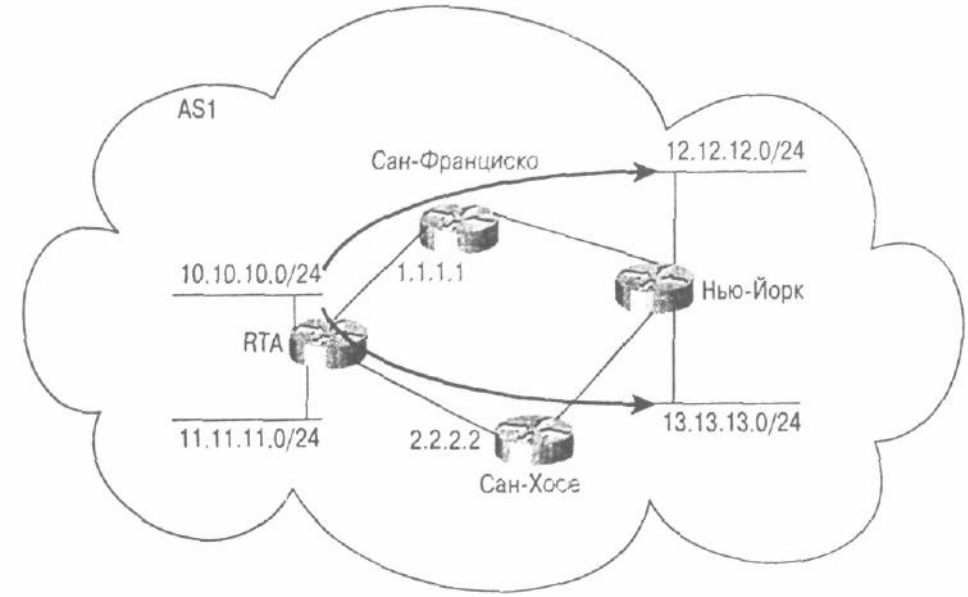


Рис. 8.7. *Вариант маршрутизации по правшам на основе адресов отправителя и получателя*

Маршрутизация по правилам с использованием маршрута по умолчанию и динамическая маршрутизация

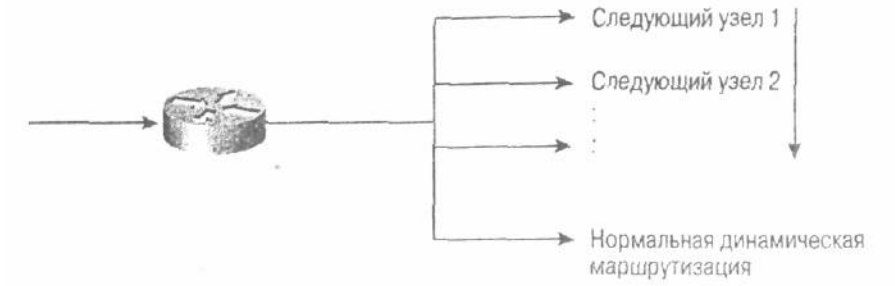


***Рис. 8.8**. Маршрутизация по правилам с использованием маршрута по умолчанию и динамическая маршрутизация*

При принудительной орынизации модели статического поведения на передний план выходит проблема обеспечения резервного канала. При выходе из строя одного узла очень важно обеспечить доставку трафика по альтернативному маршруту на другой узел. Компания Cisco предлагает тBGPчески подойти к маршрутизации по правилам путем задания нескольких ближайших узлов для передачи трафика. Если первый из узлов в списке отключен от сети или временно недоступен, то делается попытка отправить данные на второй ближайший узел и так далее. Если ни один из статически заданных ближайших узлов не оказался доступен, маршрутизатор может быть сконфигурирован для отправки трафика согласно нормальной динамической маршрутизации (т.е. на основе адреса получателя), как это показано на рис. 8.8.

Другие применения маршрутизации по правилам

Одно из практических применений маршрутизации по правилам — построение на ее основе брандмауэров. *Брандмауэрами (firewalls)* называют устройства, с помощью которых обеспечивается безопасность трафика. Реализации брандмауэров включают в себя фильтрацию пакетов, аутентификацию и шифрование. В зависимости от конфигурации сети администраторы могут направлять определенную часть или весь входящий (или исходящий) трафик на брандмауэр, как это показано на рис. 8.9.

В рассмотренной ситуации мы можем контролировать весь трафик, поступающий в сеть организации через внешние соединения. Возможно, организации требуется, чтобы пользователи с удаленных узлов подключались к Internet через брандмауэр. Если брандмауэр находится на траектории движения трафика — это не проблема. Любой входящий и исходящий трафик будет проходить через брандмауэр, чтобы попасть в пункт назначения. Однако в некоторых случаях (таких как представленный на рис. 8.9) трафик обходит брандмауэр. Маршрутизация по правилам может быть задана для маршрутизатора, граничащего с внешними сетями, чтобы входящий трафик направлялся на брандмауэр. После того как брандмауэр выполнит необходимые процедуры или шифрование, трафик посылается в конечный пункт назначения.

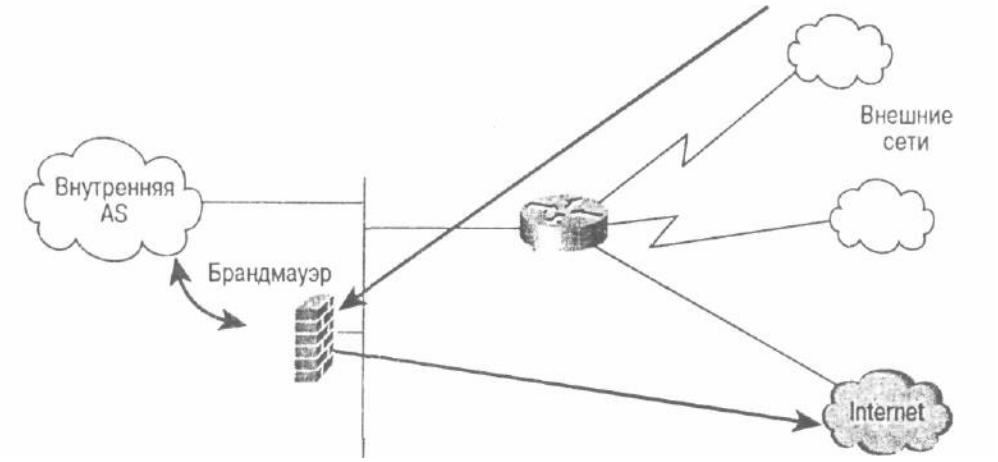


Рис. 8.9. *Входящий и исходящий трафик может проходить через брандмауэр*

Примечание

При маршрутизации по правилам адреса получателя не изменяются. Маршрутизация по правилам влияет лишь на направление трафика тому или иному ближайшему узлу, с которого и будет проводиться пересылка в конечный пункт назначения.

Маршрутизация по правилам может также применяться при работе по коммутируемым каналам для более оптимального распределения нагрузки, как показано на рис. 8.10.

Пользователи с доступом по коммутируемым линиям в определенную точку присутствия направляются в сеть провайдера, соответственно их IP-адресам. Как показано на рис. 8.10 пользователи с доступом по коммутируемым линиям в регионе 1 направляются к провайдеру 1, а пользователи в регионе 2 направляются к провайдеру 2.

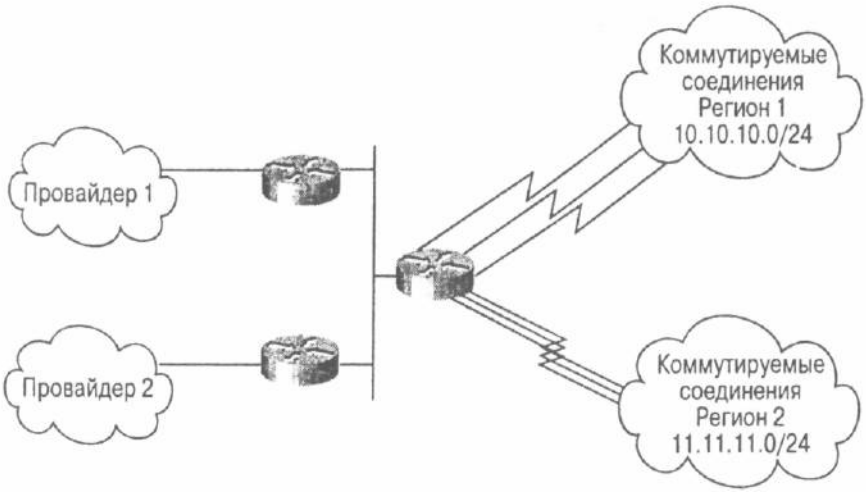


Рис. 8.10. *Распределение нагрузки на коммутируемых линиях на основе адреса отправителя*

Маршрутизация по правилам не должна использоваться вместо обычной динамической маршрутизации, а лишь дополнять ее. Маршрутизация по правилам имеет и ряд недостатков.

- Для идентификации адресов отправителя или комбинации адресов отправителя и получателя требуется задание дополнительной статической конфигурации. Следует очень осторожно проводить подобные настройки, чтобы не сорвать обработку другого трафика и одновременно указать альтернативные пути обработки заданного трафика при выходе из строя основного канала.
- Маршрутизация по правилам создает дополнительную нагрузку на процессор, так как она основана на вычислении IP-адресов отправителей, в отличие от обычной динамической и статической маршрутизации, которые основаны на работе с IP-

адресами получателей. При обработке трафика на пути к пункту назначения реализованы сложные методы кэширования и коммутации. Большинство этих методов маршрутизации и кэширования IP-пакетов еще не оптимизированы. Кроме того, при маршрутизации по правилам требуются дополнительные такты процессора для определения адресов отправителя. Подобных издержек можно избежать, поняв структуру потоков IP-трафика, что позволит нормально отслеживать информацию об источниках и пунктах назначения пакетов. Применение этих новых методов кэширования ослабит нагрузку на процессор маршрутизатора, избавив его от постоянного вычисления соответствующих правилам адресов отправителей в потоке IP-трафика, и сделает маршрутизацию по правилам более эффективной.

Забегая вперед

Автономные системы имеют свойство разрастаться до гигантских размеров вопреки усилиям администраторов. Провайдеры Internet могут в один прекрасный день обнаружить, что их сети представляют собой огромные BGP-конклавы, громоздкие и неэффективные в управлении. С другой стороны, сети предприятий могут расти в направлении, которое предполагает жесткую конкуренцию протоколов внутреннего шлюза в борьбе с нестабильностью. Управление крупномасштабными автономными системами состоит в искусстве разделения этих огромных доменов на более мелкие части, которые уже поддаются управлению. В следующей главе мы предложим вашему вниманию концепции и методы, которые помогут провайдерам и клиентам реализовать архитектуру сети с нормальной структурированной маршрутизацией внутри доменов.

Часто задаваемые вопросы

В — *Между своими граничными маршрутизаторами* я не *поддерживаю протокол IBGP. Имеется пи возможность образования петель маршрутизации?*

О — При взаимодействии IGP и BGP петли не могут появиться. Если ваши внутренние маршрутизаторы работают по умолчанию с граничными BGP-маршрутизаторами, то после того как трафик попадет на граничный маршрутизатор, у него есть единственный путь наружу — сеанс EBGP.

В — *У меня есть два граничных маршрутизатора, которые взаимодействуют по IBGP и соединены между собой физическим каналом. Я использую локальные предпочтения для управления,точками выхода. Что случится, если этот канал выйдет, из строя?*

О — Если вы устанавливаете правила работы по BGP, согласно которым трафик будет направляться между двумя граничными BGP-маршрутизаторами, то это равнозначно отсутствию связи между ними. При выходе из строя прямого канала между ними ваш трафик может образовать петлю внутри AS.

В — *Если я для перенаправления трафика от одного маршрутизатора на другой использую прямой канал между двумя граничными маршрутизаторами, которые взаимодействуют по IBGP, нужно ли обеспечить скорость работы по этому каналу такую же, как и скорость работы по каналам с провайдерами?*

О— Ваш канал будет обслуживать только тот трафик, которым обмениваются граничные маршрутизаторы, и частично входящий трафик. Попробуйте выяснить, какую .часть в процентном соотношении от общего трафика составляет именно эта нагрузка ; и исходя из этого заказывайте полосу пропускания.

В — *Мне нужно направить трафик в пункт назначения X по прямому каналу между ,маршрутизаторами и в пункт назначения Y*— через *Ethernet*. *Могу ли я обеспечить эти требования с помощью маршрутизации по правилам?*

О — Это можно выполнить с помощью статической маршрутизации, которая работает на основе именно адреса пункта назначения. В вашем случае нет необходимости прибегать к маршрутизации по правилам, которая использует адрес источника или адреса источника и пункта назначения.

.В – *Могу ли я применить маршрутизацию по правилам на входящем и исходящем интерфейсах моего маршрутизатора?*

О — Маршрутизация по правилам основана на проверке адреса отправителя, поступающего на интерфейс. Настройте соответствующим образом входящий интерфейс.

Ключевые темы этой главы:

- **Отражатели маршрутов.** Представлен метод управления крупномасштабными автономными системами (autonomous system — AS) с помощью выбранных маршрутов, которые выступают в качестве узловых точек для внутренних BGP-сеансов.
- **Конфедерации.** Рассматривается метод управления крупными AS с помощью разделения их на более мелкие AS.
- **Управление ростом инфраструктуры IGP.** Приведены различные методы управления сетями, в которых рост инфраструктуры характеризуется использованием нескольких протоколов IGP.
- **Виртуальные частные сети с отражателями маршрутов.** Обсуждается метод ограничения доступа к сети внутри AS с использованием отражателей маршрутов.

Глава 9. Управление крупномасштабными автономными системами

Управление маршрутизацией автономных систем, состоящих из сотен узлов, может представлять для сетевых администраторов серьезную проблему. При обслуживании крупномасштабных сетей провайдеры и их клиенты сталкиваются с различными проблемами. Со стороны сервис-провайдера большинство маршрутизаторов работает под управлением протокола граничного шлюза (Border Gateway Protocol - BGP). По-' скольку протоколом BGP предусмотрено правило, согласно которому один спикер протокола внутреннего граничного шлюза (Interior Border Gateway Protocol -- IBGP) не может объявлять маршрут, полученный от другого спикера IBGP, инфраструктура IBGP может быстро и бесконтрольно расти. Однако у клиентов, как правило, большинство маршрутизаторов работают под управлением протоколов внутреннего шлюза (Interior Gateway Protocols — IGP), инфраструктура которых также может бурно разрастаться без контроля со стороны администратора.

В этой главе обсуждаются методы и приемы для улучшения эффективности управления протоколами BGP и IGP внутри крупномасштабных автономных систем. Вы должны сами решить, использовать ли методы, описываемые нами в этой главе, и какой из них вам наиболее подходит. Помните, что внедрение любого метода, любой новой технологии всегда сопряжено с определенными трудностями. Навязывание сложных приемов и методов в случаях, где это реально не требуется, может принести больше вреда, чем пользы. Заблаговременное планирование может значительно уменьшить количество проблем по мере развития инфраструктуры сети.

Отражатели маршрутов

В сетях некоторых сервис-провайдеров Internet (Internet Service Providers -- ISP) структура внутреннего BGP может приобрести огромные размеры (более 100 внутренних BGP-сеансов, на маршрутизатор). В этой ситуации настоятельно рекомендуется реализовать один из механизмов координации взаимодействующих узлов. Концепция *отражателя маршрута (route reflector)*[1] основана на идее выделения отдельного мар-;шрутизатора, который называют *маршрутизатором сосредоточения (concentration router),* в качестве узловой точки при проведении внутренних BGP-сеансов. Несколько BGP-маршрутизаторов (клиентов) могут взаимодействовать с центральным сервером (отражателем маршрутов), а затем отражатели маршрутов уже будут взаимодействовать друг с другом. Хотя согласно протоколу BGP нельзя объявлять маршруты, полученные от одного IBGP-спикера другому, процедура отражения маршрутов позволяет серверам отражателей маршрутов "отражать" маршруты, как будет описано позже, делая своего рода исключение из ограничений протокола IBGP.

Отражатели маршрутов рекомендуется применять только в AS, где имеется крупная внутренняя инфраструктура протокола BGP. Работа в режиме отражателя маршрутов

производит дополнительную нагрузку на ресурсы сервера отражателя маршрутов, что при неправильной настройке может привести к образованию петель маршрутизации и обшей нестабильности системы маршрутизации. Принимая во внимание эти факты, не рекомендуется использование отражателей маршрутов в сетях любой топологии.

Как видите, отражение маршрутов обеспечивает некоторые преимущества и для серверов отражателей маршрутов, и для их клиентов. Например, с помощью сервера отражателя маршрутов передача сообщений UPDATE может осуществляться нескольким узлам одновременно, т.е. нет необходимости генерировать уникальные сообщения для каждого узла отдельно. Кроме того, клиенты обычно взаимодействуют только с локальным сервером отражателя маршрутов, что значительно сокращает количество сеансов, которые они будут обслуживать.

См. в главе 12 раздел "Отражатели маршрутов"

Внутренние узлы без отражателей маршрутов

Без отражателей маршрутов BGP-спикеры в AS будут иметь логическую структуру. Ранее мы уже обсуждали их работу в этой книге; приведенный ниже рисунок лишь напоминание для вас. На рис. 9.1 маршрутизаторы RTA, RTB и RTC формируют внутреннюю логическую BGP-сеть. Каждый маршрутизатор работает с двумя другими как обычная полноценная BGP-система. Маршрутизаторы RTA и RTB, а также RTB и RTC имеют между собой физическое соединение. А маршрутизаторы RTA и RTC не соединены между собой физически.

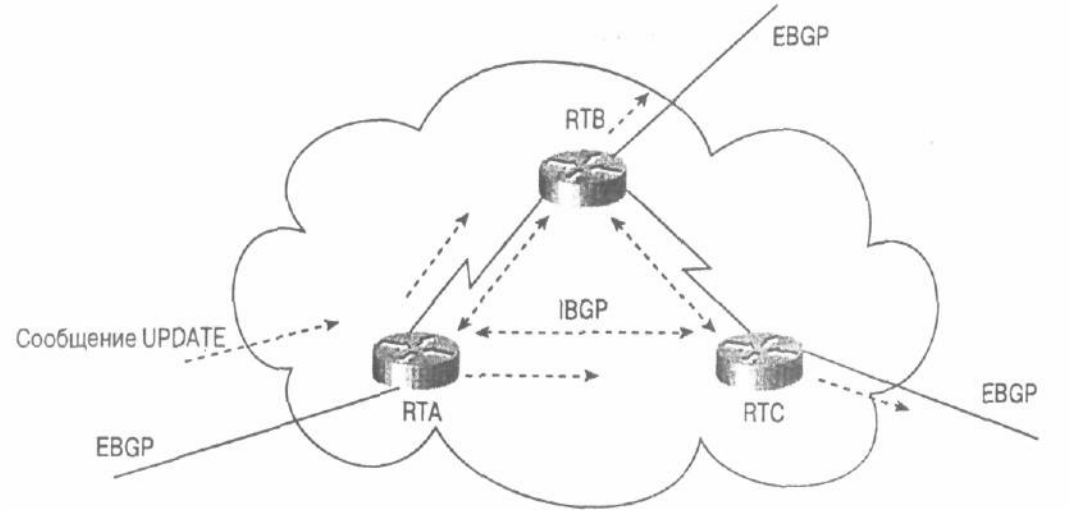


Рис. 9.1. *Внутренние узлы в нормальной полносвязной среде*

Когда маршрутизатор RTA получает сообщение об обновлении маршрутов от внешнего узла, он пересылает его своим взаимодействующим внутренним узлам RTB и RTC. Обратите внимание, что между RTA и RTC физическое соединение отсутствует. Маршрутизатор RTA передает данные об обновлении маршрутов на RTC во время внутреннего BGP-сеанса. В свою очередь, маршрутизаторы RTB и RTC передают обновление своим внешним узлам.

Сообщение UPDATE, которое RTB получает от RTA, не объявляется повторно для RTC, так как этот узел является внутренним и сообщение UPDATE маршрутизатором RTB было получено от внутреннего узла (RTA). Без организации внутреннего сеанса по протоколу BGP между маршрутизаторами RTA и RTC последний никогда не получил бы обновления маршрута, следовательно требуется полносвязная сеть 1BGP.

Внутренние узлы с отражателями маршрутов

Отражатель маршрутов для других маршрутизаторов, которые называют *клиентами,* действует как точка сосредоточения. Клиенты взаимодействуют с отражателем маршрутов и ведут с ним обмен маршрутной информацией. В свою очередь, отражатель маршрутов передает (или, как говорят, отражает) информацию между клиентами и другими IBGP- и EBGP-узлами.

На рис. 9.2 маршрутизатор RTB настроен для работы в качестве отражателя маршрутов между двумя клиентами -- маршрутизаторами RTA и RTB. Маршрутизатор RTA получает сообщение об обновлении от внешнего узла и передает его RTB. Тот отражает обновление от клиента RTA клиенту RTC. В таком случае нет необходимости в организации отдельного сеанса между RTA и RTC, так как отражатель маршрутов распространяет BGP-информацию от RTA к RTC.

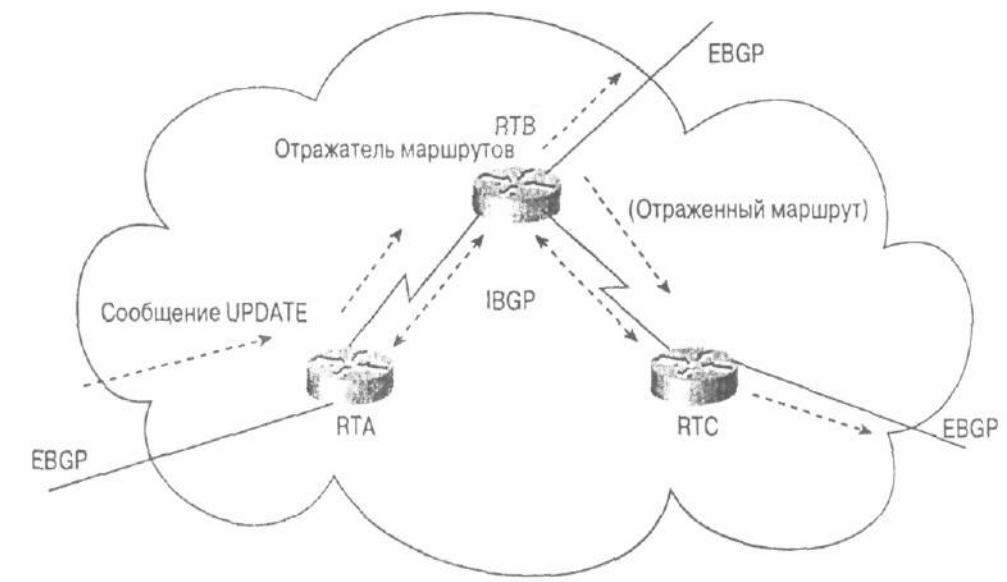


Рис. 9.2. *Внутренние узлы с использованием отражателя маршрутов*

В AS, когда администратору необходимо организовывать значительное число BGP-сеансов между маршрутизаторами, концепция отражателя маршрутов обеспечивает очень полезное и масштабируемое решение проблемы.

Соглашения об именах и правила работы

По отношению к отражателю IBGP-узлы делятся на две категории — *клиенты (clients)* и *неклиенты (noncliems).* Отражатель маршрутов вместе со своими клиентами формирует *кластер (cluster).* Все узлы, не вошедшие в кластер, квалифицируются отражателем маршрутов как неклиенты. На рис. 9.3 отображены все описанные нами компоненты.

Неклиенты (стандартные IBGP-спикеры) должны соединяться друг с другом и с отражателем маршрутов, так как они работают в соответствии с нормальными правилами объявления маршрутов по 1BGP, хотя им уже и не требуется взаимодействовать с узлами-клиентами отражателя маршрутов. Клиенты не должны взаимодействовать со спикерами вне кластера, к которому они принадлежат. Все эти условия для клиентов и неклиентов отображены на рис. 9.3.

Функция отражения маршрутов реализована только в самом отражателе маршрутов; все остальные клиенты и неклиенты представляют собой обычные BGP-узлы, в которых отсутствуют какие-либо настройки отражателя маршрутов. Клиенты отражателя маршрутов

являются таковыми лишь потому, что сам отражатель воспринимает их как клиентов (т.е. указывает их в своем списке клиентов).

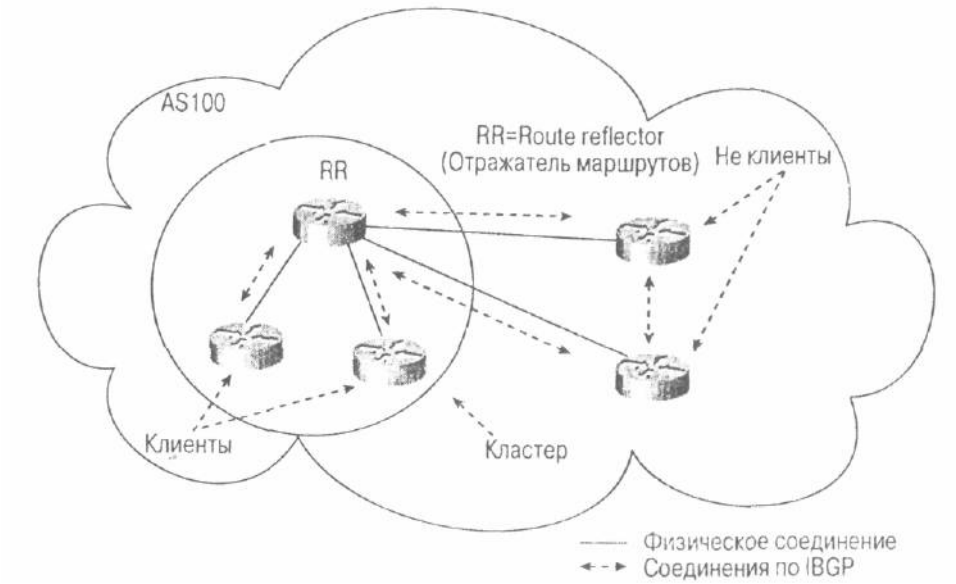


Рис. 9.3. *Компоненты, участвующие в процессе отражения маршрутов*

Отражатель маршрутов, который получает несколько маршрутов для одного и того же пункта назначения, использует обычный в BGP процесс принятия решения для выбора наилучшего маршрута. Далее наилучший маршрут будет распространяться внутри AS на основе следующих правил.

- Если маршрут получен не от клиента, то отражать его только клиентам.
- Если маршрут получен от клиента, то отражать его всем узлам; и клиентам, и неклиентам.
- Если маршрут получен от внешнего EBGP-узла, также отражать его всем узлам: и клиентам, и неклиентам.

Ввиду того что концепция отражения маршрутов применяется только внутри AS, внешние по отношению к AS маршрутизаторы, которые получают сообщение UPDATE по протоколу EBGP, считаются неклиентами и при приеме и передаче обновлений маршрутов ведут себя, как обычные неклиенты.

Вопросы обеспечения избыточности при работе с несколькими отражателями маршрутов в AS

В неполносвязной **BGP-сети** внутри AS избыточность и надежность становятся важными параметрами. Если выходит из строя отражатель маршрутов, то клиенты оказываются изолированными. Требования избыточности обязывают нас организовать несколько отражателей маршрутов в кластере, где клиенты могут одновременно взаимодействовать с несколькими маршрутизаторами. Если один из отражателей выходит из строя, то другой (или другие) смогут выполнять его функции.

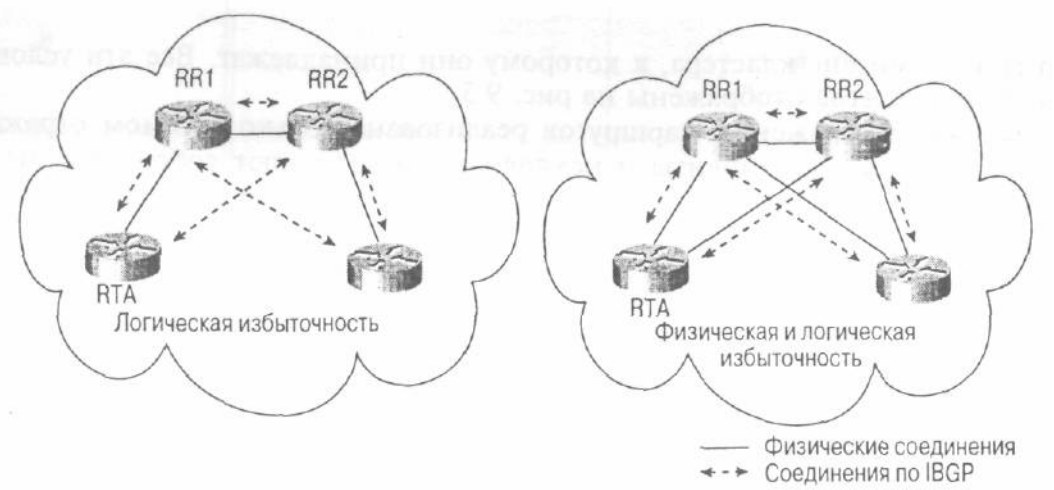


Рис. 9.4. *Сравнительная характеристика обеспечения логической и физической избыточности*

Нельзя преувеличивать важность дополнения логической избыточности по сравнению с физической избыточностью. Нет смысла обеспечивать избыточность отражателей маршрутов, если отсутствует физическая избыточность, В левой части рис. 9.4 показана схема реализации логической избыточности, где маршрутизатор RTA является клиентом отражателей маршрутов RR1 и RR2, (Здесь и далее отражатели маршрутов будут обозначаться как RR от англ, — route reflector). Маршрутизатор RTA взаимодействует с обоими отражателями маршрутов для создания избыточного соединения. К сожалению, если соединение с отражателем маршрутов RR1 по каким-либо причинам будет разорвано или выйдет из строя отражатель RR1, то маршрутизатор RTA окажется изолированным. Таким образом, в данной схеме логическое соединение RTA и отражателя маршрутов RR2 не имеет практической пользы, оно лишь потребляет больше памяти и создает дополнительную нагрузку на процессор.

В правой части рис, 9.4 представлена схема реализации физической избыточности, в которой можно обеспечить также и физическую избыточность. В случае выхода из строя соединения с отражателем маршрутов RR1 маршрутизатор RTA может работать с RR2.

Модели топологии с элементами отражения маршрутов

Национальные телекоммуникационные сети, как правило, планируются в соответствии с географическими регионами и имеют точки сосредоточения в городах. Провайдеры имеют собственные точки присутствия (Points of Presence — POP), которые также иногда называют *хабами (hubs),* в различных регионах США с высокоскоростными каналами, соединяющими различные узлы смешанной топологии. Концепция отражателя маршрутов может использоваться, чтобы логически объединять маршрутизаторы под управлением протокола BGP в структуры, соответствующие физической топологии сети. На рис. 9.5 показана комплексная инфраструктура с использованием отражателей маршрутов.

Исключая тот факт, что отражатель маршрутов должен обрабатывать больше IBGP-сеансов, чем маршрутизаторы-клиенты, любой маршрутизатор может быть сконфигурирован как отражатель маршрутов. Физическая топология вашей сети определяет, какой маршрутизатор будет наилучшим отражателем маршрутов. Поскольку клиентам не требуется взаимодействовать с большим количеством других IBGP-маршрутизаторов, для обработки соединений по EBGP у него имеется больше свободных ресурсов.

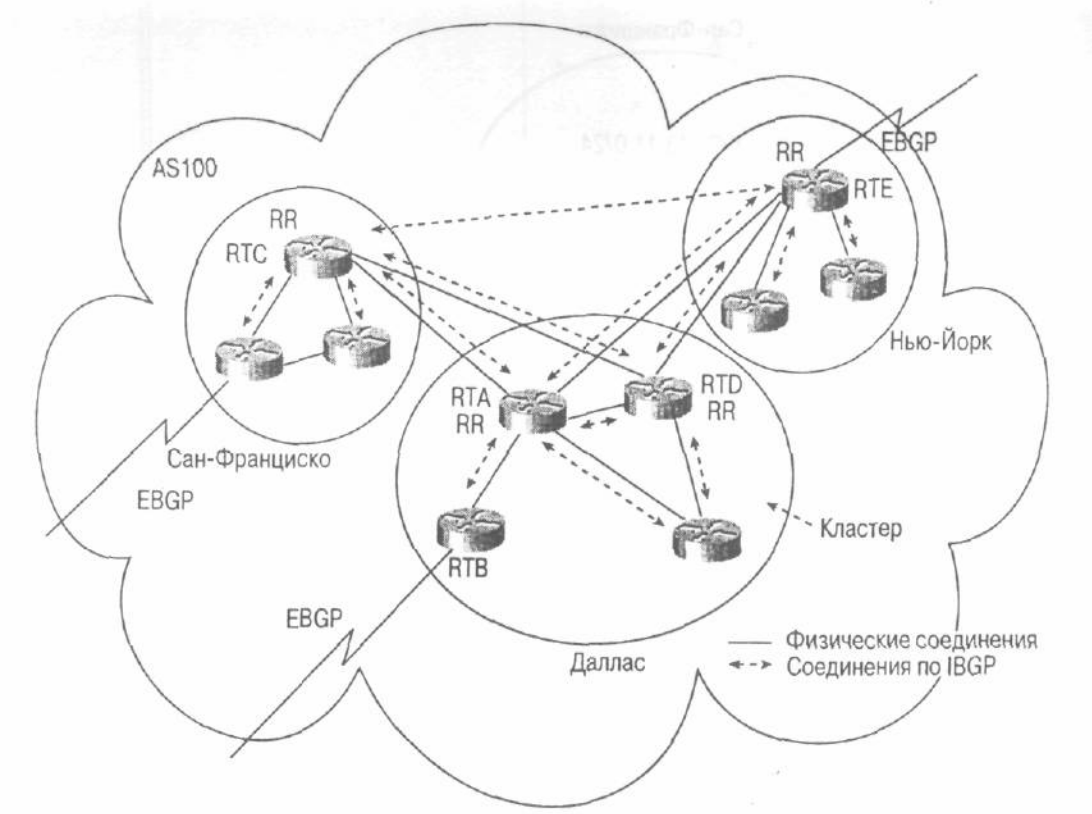


Рис. 9.5. *Инфраструктура сети с несколькими отражателями маршрутов*

На рис. 9.5 AS100 разделена на три кластера: Сан-Франциско, Даллас и Нью-Йорк. Кластер Далласа в целях обеспечения избыточности имеет несколько отражателей маршрутов. Маршрутизаторы RTA и RTD физически соединяют Сан-Франциско и Нью-Йорк. При выборе отражателей маршрутов имеет смысл просто следовать физическому потоку трафика, так что маршрутизаторы RTA и RTD являются наилучшими кандидатами на эту роль в кластере Далласа.

В Сан-Франциско маршрутизатор RTC обеспечивает физическое соединение с Далласом, так что именно он и будет наилучшим кандидатом для отражателя маршрутов. То же самое справедливо и для кластера Нью-Йорка: маршрутизатор RTE физически соединяет Нью-Йорк с Далласом и, следовательно, является наилучшим кандидатом для работы в качестве отражателя маршрутов.

Отражатель маршрутов и IBGP-атрибуты

Концепция использования отражателя маршрутов не влияет на работу по протоколу IBGP, т.е. отражатель маршрутов не может изменять атрибуты отражаемых IBGP-маршрутов. Например, атрибут NEXT_HOP остается неизменным, когда два отражателя маршрутов обмениваются IBGP-маршрутом. Это необходимо, чтобы избежать образования петель внутри AS.

На рис. 9.6 показано, почему отражатель маршрутов не должен модифицировать атрибуты. В качестве примера приводится атрибут NEXT_HOP. Здесь представлен элемент сети, объединяющий Даллас и Сан-Франциско.

Хотя правила отражения маршрутов гласят о том, что отражатель маршрутов не должен модифицировать любой из атрибутов, большинство реализаций позволяет выполнять своего рода фильтрацию. Будьте предельно внимательны при выполнении подобных настроек.

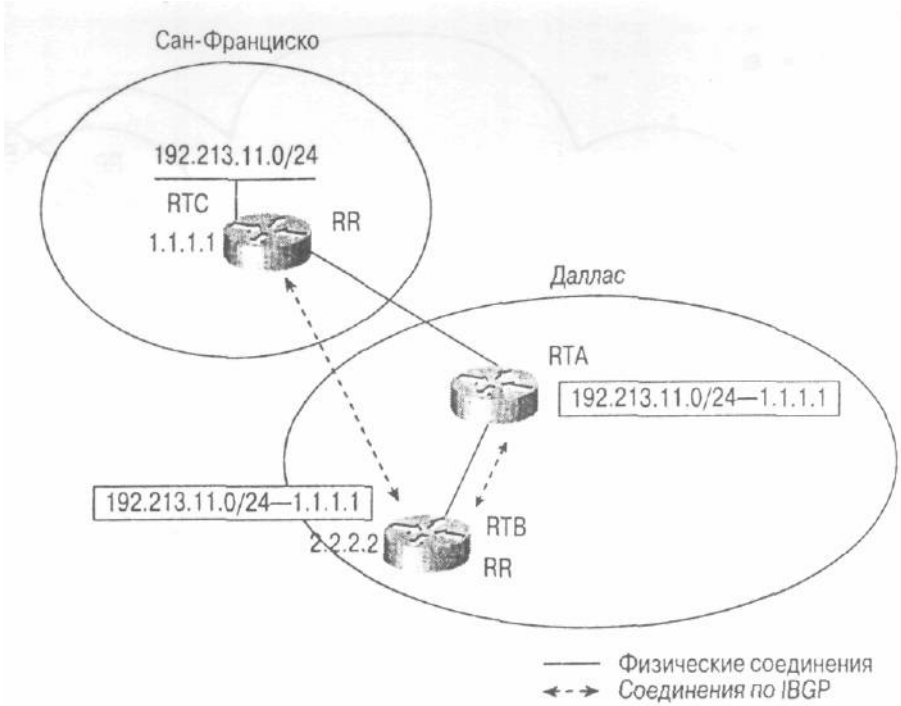


Рис. 9.6. *Отражатель маршрутов сохраняет информацию об IBGP-атрибутах*

Предположим, что маршрутизатор RTB указан в качестве отражателя маршрутов раньше, чем маршрутизатор RTA, и между маршрутизаторами RTB (2.2.2.2) и RTC (1.1.1.1) установлен lBGP-сеанс. Это выглядит не совсем правильно, так как трафик физически проходит через маршрутизатор RTA, в то время как RTB лишь отражает BGP-маршруты между RTA и RTC. Маршрутизатор RTB будет получать префикс 192.213.11.0/24 от своего соседа по IBGP-маршрутизатора RTC с адресом RTB следующего узда (1.1.1.1). Маршрутизатор RTB будет отражать маршрут своему клиенту-маршрутмзатору RTA также с адресом следующего узла (1.1.1.1). Это стандартное и наиболее желательное отражение маршрутов.

С другой стороны, если маршрутизатор RTB изменит IP-адрес следующего узла на свой IP-адрес (2.2.2.2), то RTA попытается использовать RTB, чтобы попасть в сеть 192.213.11.0/24. В результате между маршрутизаторами RTA и RTB возникнет петля, так как RTA будет использовать в качестве следующего узла RTB. Вследствие этого маршрутизатор будет пересылать трафик обратно RTA, чтобы доставить его в конечный пункт назначения. Эта гипотетическая ситуация служит примером того, почему отражателю маршрутов нельзя изменять IBGP-атрибуты.

Стоит упомянуть еще раз о том, что отражатели маршрутов могут распространять только наилучшие маршруты к заданному пункту назначения. Другими словами, если отражатель маршрутов получает сведения об одном и том же префиксе от различных клиентов, то только один маршрут будет передаваться на другие узлы. Следовательно, при использовании отражателей маршрутов число возможных маршрутов к заданному пункту назначения, возможно, будет меньше, чем для конфигурации полносвязной сети. Вследствие такого поведения еще раз рекомендуем вам, чтобы топологии с использованием отражателей маршрутов соответствовали физической топологии сети, иначе имеется опасность неоптимальной маршрутизации.

Как избежать образования петель маршрутизации

Для обнаружения петель протокол BGP использует информацию, заключенную в атрибуте AS_PATH. BGP-обновление, пытающееся повторно поступить в AS, которая его сгенерировала, будет игнорироваться граничным маршрутизатором исходной AS.

С введением в вашу сеть отражателей маршрутов опасность образования петель маршрутизации внутри AS возрастает. Сообщение об обноаании маршрутов, которое покинуло кластер, может повторно в него поступить. Петли внутри AS не могут быть обнаружены традиционным методом с использованием AS_PATH, так как обновления

маршрутов не имеют подписи AS, сгенерировавшей атрибут. Таким образом, при использовании отражателей маршрутов в BGP можно использовать два дополнительных атрибута, с помощью которых можно избежать образования петель внутри AS, — ORIGINATOR_ID и CLUSTER_LIST.

Применение атрибута ORIGINATORJD

Атрибут *ORIGINATOR_ID* представляет собой четырехбайтовый необязательный нетранзитивный атрибут BGP (код типа 9). В этом атрибуте переносится информация об идентификаторе маршрутизатора ROUTER_ID, который сгенерировал маршрут в локальной AS, и добавляется в сообщение UPDATE отражателем маршрутов. Если в результате неправильной конфигурации обновление поступает обратно на сгенерировавший его узел, то оно должно быть отвергнуто эти узлом.

Атрибут CLUSTER_LIST

Атрибут *CLUSTER_LIST* является необязательным нетранзитивным атрибутом BGP (код типа 10). Каждому кластеру назначается идентификатор CLUSTER_ID.

Атрибут CLUSTER LIST представляет собой последовательность идентификаторов CLUSTER_ID, в которых содержится маршрутная информация о списке кластеров, через которые прошло сообщение UPDATE. Когда отражатель маршрутов посылает маршрут от своих клиентов неклиентам вне кластера, он добавляет свой локальный CLUSTER_ID в CLUSTER_LIST или создает новый список CLUSTER_LIST, если таковой отсутствует. Если отражатель маршрутов принимает сообщение UPDATE, в CLUSTER_LIST которого уже имеется CLUSTER_ID, то он игнорирует это сообщение. Таким образом, в CLUSTER_LIST реализован механизм предотвращения образования петель маршрутизации внутри AS, в то время как список AS_PATH, который мы обсуждали ранее, предотвращает образование петель для сообщений UPDATE, прошедших через несколько внешних AS.

Более подробно об этом читайте в разделе "Отражатели маршрутов" в главе 12.

Отражатели маршрутов и группы взаимодействующих узлов

Вернемся к главе 6, "Настройка параметров BGP", где в качестве группы взаимодействующих узлов выступает группа соседних BGP-маршрутизаторов, которым заданы одни и те же правила маршрутизации. Ранее отражатели маршрутов могли использоваться только в группах взаимодействия, когда все клиенты внутри кластера соединены по полносвязной схеме. Причины этого лучше всего объяснить на примере. В обычной ситуации с использованием отражателя маршрутов маршрутизатор А получает префикс от маршрутизатора В. Затем маршрутизатор А посылает сообщение UPDATE, содержащее удаленные маршруты (в поле WITHDRAWN ROUTES), обратно на маршрутизатор В для того, чтобы нейтрализовать этот маршрут. Другими словами, маршрутизатор А информирует маршрутизатор В о том, что заданный префикс недоступен через маршрутизатор А. Это предотвращает образование петли маршрутизации, когда маршрутизатор А требует, чтобы префикс был доступен через маршрутизатор В, а В указывает, что префикс доступен через А.

В группе взаимодействующих узлов одно и то же сообщение UPDATE (с той же самой информацией WITHDRAWN ROUTES) будет разослано всем членам группы. При наличии в группе взаимодействующих узлов отражателя маршрутов последний, получив префикс от одного из клиентов и попытавшись нейтрализовать этот маршрут, будет удалять такой префикс, если он поступит от других клиентов. Так как клиенты не общаются друг с другом по BGP, этот префикс теряется. Следовательно, существует необходимость в обеспечении полносвязной схемы для работы по протоколу IBGP между клиентами отражателя маршрутов, чтобы и другие клиенты могли получать сведения о префиксе прямо от узла, который его генерирует. Однако даже при такой схеме сетевые администраторы избегают построения полносвязной IBGP-сети между всеми IBGP-маршрутизаторами в AS.

Они лишь выполняют соединения по полносвязной схеме между отражателями маршрутов и клиентами (в отличие от организации соединений между клиентами в кластере).

К счастью, в операционной системе IOS отсутствуют требования к организации полносвязных соединений между клиентами отражателя маршрутов. В настоящее время клиенты отражателя маршрутов, объединенные в группу взаимодействующих узлов, не нуждаются в организации соединений по схеме "каждый с каждым".

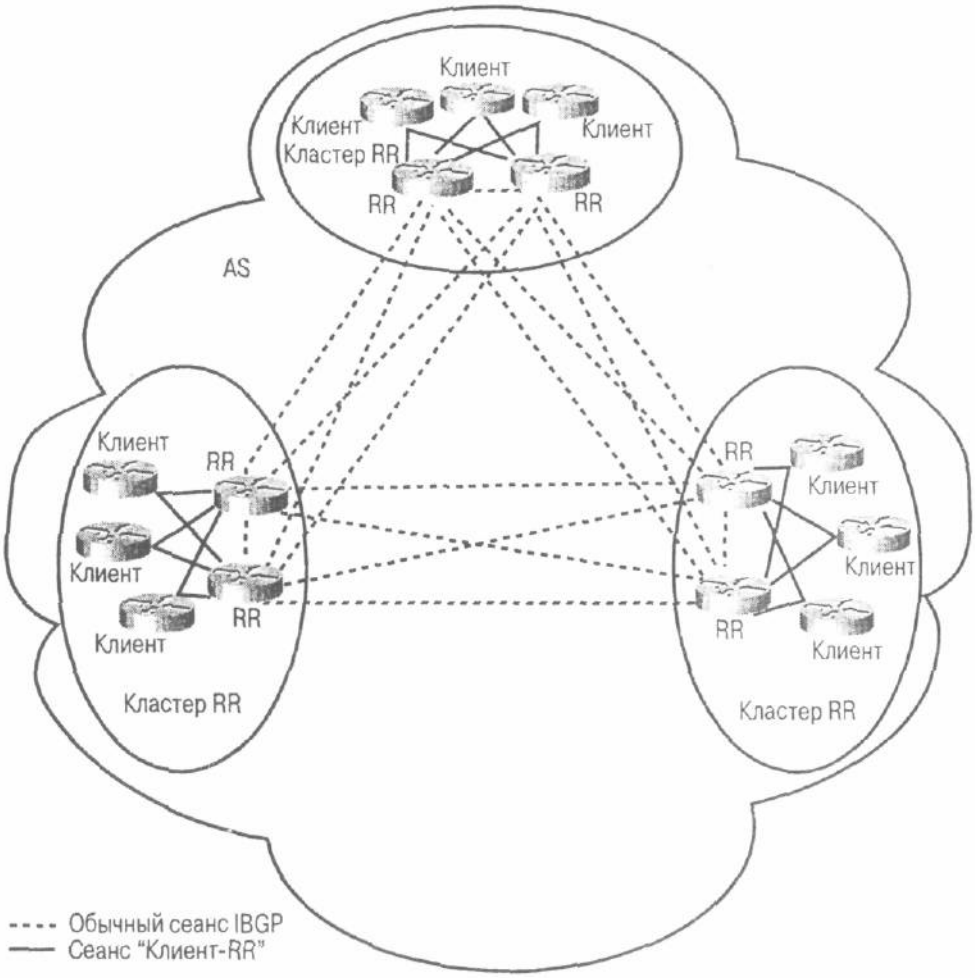


Рис. 9.7. *Типовая топология сети с применением отражения BGP-маршрутов*

С использованием групп взаимодействующих узлов структура AS будет представлять собой систему колец, объединенных по полносвязной схеме. Отражатели маршрутов при этом соединены каждый с каждым, а клиенты должны лишь подключаться к отражателям маршрутов. На рис. 9.7 представлена структура такой сети: каждая выделенная группа узлов представляет собой определенную группу взаимодействующих узлов и кластер отражателя маршрутов. В отличие от этой схемы, на рис. 9.8 представлена структура, которую необходимо реализовать, если отражатели маршрутов не используются.

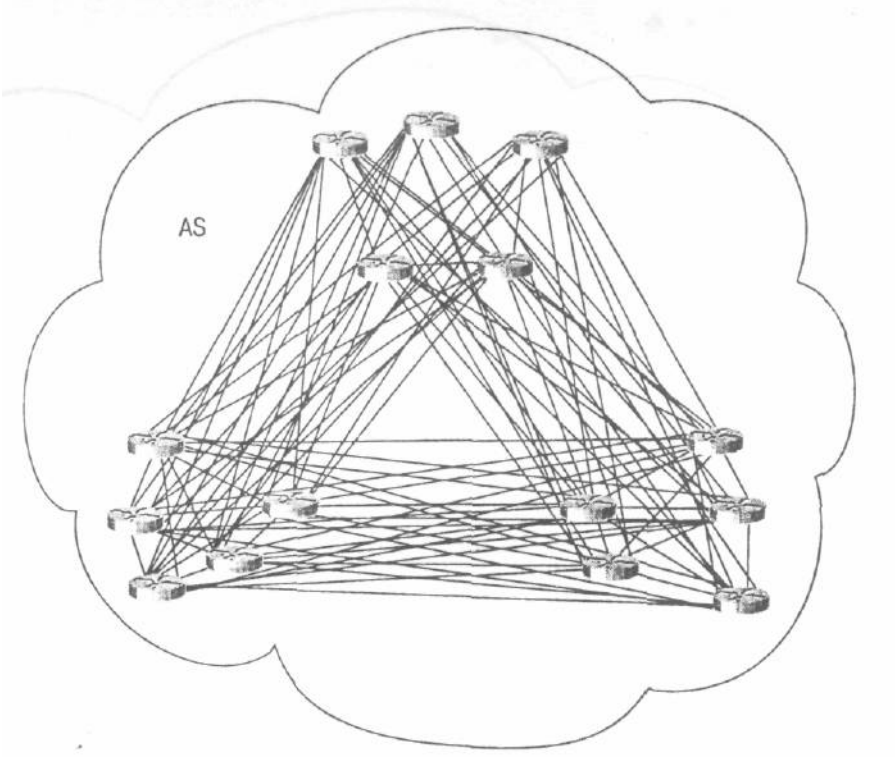


Рис. 9.8. *Полносвязная топология сети*

В заключение отметим, что концепция отражателя маршрутов приобретает вес большую популярность в крупных сетях вследствие своей простоты и масштабного подхода, который не требует существенных затрат. Более того, переход к структуре сети с применением отражателя маршрутов происходит очень легко. Изменения придется внести лишь в те маршрутизаторы, которые будут выполнять роль отражателей маршрутов, а все остальные — будут функционировать, как обычно. К тому же маршрутизаторы, в которых не заложены функции отражения маршрутов (клиенты или неклиенты), по-прежнему могут быть частью AS без какого-либо ущерба для маршрутной информации, передаваемой с использованием BGP.

Конфедерации

Конфедерация (confederation)[2] -- еще один метод борьбы с полносвязностью IBGP внутри AS. Организовывать конфедерации рекомендуется только в тех случаях, когда в работу по протоколу IBGP вовлечено большое количество узлов, что вызывает лавинообразное нарастание числа JBGP-сеансов на отдельном маршрутизаторе.

См. в главе 12 раздел "Конфедерации"

BGP-конфедерации основаны на концепции, согласно которой AS может разбиваться на несколько более мелких AS. Внутри каждой такой подсистемы AS действительны все правила маршрутизации по IBGP. Например, все BGP-маршрутизаторы внутри подсистемы AS должны быть соединены друг с другом по полносвязной схеме. Так как каждая подсистема AS имеет собственный номер, то они будут взаимодействовать по внешнему протоколу BGP. Однако несмотря на то, что между подсистемами AS используется протокол EBGP, маршрутизация внутри конфедерации напоминает маршрутизацию по IBGP внутри отдельной AS. Другими словами, при пересечении подсистемы AS информация о следующем ближайшем узле, локальные предпочтения и MED сохраняются. Для внешней глобальной сети конфедерация выглядит, как одна отдельная AS.

На рис. 9.9 приведен пример конфедерации.

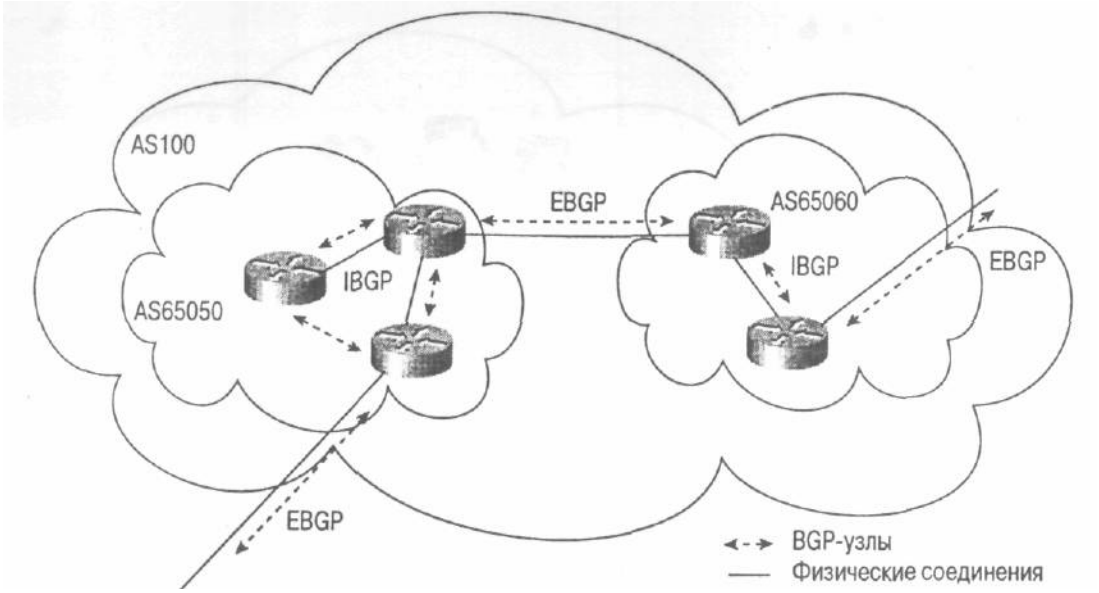


Рис. 9.9. *Пример структуры конфедерации*

Здесь AS 100 разделена на две подсистемы AS — AS65050 и AS65060. Теперь вся AS тредставляет собой одну большую конфедерацию, которая идентифицируется номе-эом 100. Все подсистемы AS "невидимы" для внешней сети, и им в принципе можно задавать любые номера. Эти номера вы можете назначать из диапазона частных номе-ов AS (от 64512 до 65534, как это оговорено в RFC I930[3]) для того, чтобы не использовать уникальные номера AS.

Как уже отмечалось, внутри подсистем AS используется полносвязная схема соединений для работы по IBGP. Протокол EBGP используется для организации работы чежду несколькими подсистемами AS, а также между самой конфедерацией и внешними AS. В конфедерациях относительно легко выявить петли маршрутизации внутри \S, благодаря протоколу EBGP, который поддерживается для обеспечения взаимодей-:твия между подсистемами AS. Механизм применения списка AS в AS_PATH, пре-ютвращаюший образование петель маршрутизации, используется для обнаружения обновлений маршрутов, покидающих одну подсистему AS и пытающихся повторно топасть в эту же подсистему AS. Сообщение об обновлении маршрута, которое по-jTopHO пытается попасть в подсистему AS, где оно сгенерировано, немедленно обна-зуживается, так как подсистема AS "увидит" в AS_PATH собственный номер AS.

Недостатки конфедераций

Основной недостаток при работе с конфедерациями заключается в том, что миграция к ггруктуре конфедерации требует существенной перенастройки маршрутизаторов и внесе-•шя изменений в логическую топологию. Кроме того, маршрутизация через конфедерацию ложет быть неоптимальной, если не установить вручную некоторые правила маршрутизация по BGP. На рис. 9.10 представлена схема маршрутизации в конфедерации.

Итак, конфедерация 100 состоит из трех подсистем AS: 65010, 65020 и 65030. Маршрут AS_PATH внутри конфедерации 100 представляет собой последовательность юмеров AS и подсистем AS, через которые пролегает маршрут. В стандартном BGP кратчайший маршрут через AS определяет наилучший путь для передачи трафика. Однако в конфедерации подсистема AS «е влияет на общую длину маршрута через VS. Например, одному префиксу соответствуют два равных по длине маршрута VS__PATH, в каждом из которых, в свою очередь, имеются AS_PATH различной дли-ш от разных подсистем AS, что в результате может привести к неоптимальной мар-ирутизации внутри AS, так как неясно, какой из маршрутов наилучший. Со стороны подсистемы AS 65030 маршрут с AS_PATH

(65010) имеет ту же длину, что и AS_PATH (65020 65010). При этом трафик внутри конфедерации может передаваться по любому из этих маршрутов. Чтобы повлиять на работу системы маршрутизации, можно дополнительно установить соответствующие правила маршрутизации. Например, для того, чтобы маршруту с AS_PATH (65020 65010) предпочесть маршрут с AS_PATH (65010), можно сконфигурировать для них локальные предпочтения.

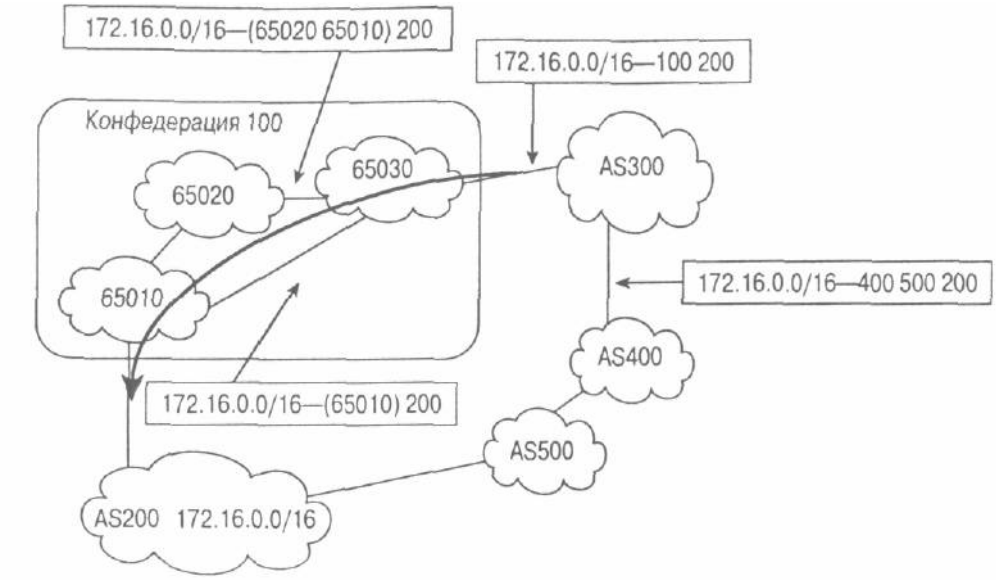


Рис. 9.10. *Внутренняя и внешняя маршрутизация в конфедерации*

Так как конфедерация — это по сути отдельная AS, маршрут, избранный внешней AS для прохождения через конфедерацию, неизвестен. Это обстоятельство вводит в заблуждение те AS, в которых правила маршрутизации основаны на длине атрибута AS_PATH. Чтобы попасть в AS200, AS300, скорее всего, выберет маршрут через конфедерацию 100, так как этот маршрут выглядит короче, чем маршрут через AS400 и AS500. В действительности, конечно же, маршрут через конфедерацию 100 не самый короткий, так как на самом деле придется пересечь три подсистемы AS (65030 65020 65010), в то время как альтернативный маршрут предлагает пересечь только две (AS400 AS500). Однако AS300 никогда не узнает об этом "подвохе", если не будет раскрыта структура конфедерации AS 100.

Обмен маршрутами и принятие решений BGP в конфедерациях

Хотя между подсистемами AS обмен маршрутами проводится посредством протокола EBGP, к ним применимы все правила работы по IBGP в порядке обеспечения действия всей AS как единого домена маршрутизации. При этом внутри AS, наряду со значениями атрибутов MED и LOCAL_PREF, по EBGP передается и атрибут NEXT_HOP.

Согласно алгоритму принятия решения в протоколе BGP, только изменения BGP-маршрутов к внешним конфедерациям можно сравнивать с маршрутами внутри самой конфедерации. Если нет конфедераций, то EBGP-маршруты перекрывают IBGP-маршруты. При использовании конфедераций появляется новая разновидность EBGP-маршрутов -- маршруты между подсистемами AS, которые называют внешними маршрутами конфедерации. В протоколе BGP предпочтение тем или иным маршрутам отдается в следующем порядке:

EBGP-маршруты от конфедерации к другим AS > внешние маршруты конфедерации > IBGP-маршруты.

Следовательно, если имеется два маршрута к одному и тому же пункту назначения:

один-- ведущий за пределы конфедерации, и один- внутри конфедерации, протокол BGP выберет внешний маршрут. Далее, если в BGP имеется возможность выбора между двумя маршрутами к одному и тому же пункту назначения: один -внутри подсистемы AS и второй -- из подсистемы AS, то BGP выберет опять же внешний маршрут, который ведет за пределы подсистемы AS. Все это верно для ситуации, в которой все остальные атрибуты одинаковы.

Рекомендуемая структура конфедерации

Выделение подсистем AS и организация соединений между ними случайным образом внутри конфедераций может привести к серьезным осложнениям. Может проявиться нежелательная нагрузка на вычислительные ресурсы маршрутизаторов в результате обработки одной и той же информации, получаемой от различных подсистем AS по одному маршруту. Кроме того, вследствие одинаковой длины маршрутов внутри AS появляются неоптимальные маршруты, на что мы уже обращали ваше внимание.

Опыт показывает, что централизованная структура конфедерации позволяет получить наиболее оптимальные схемы маршрутизации. Централизованная структура подразумевает, что все подсистемы AS обмениваются информацией друг с другом через центральный магистральный канал подсистемы AS.

На рис. 9.11 показано, что каждая подсистема AS взаимодействует только с одной подсистемой AS. Результат может быть более унифицированным с учетом длины атрибута AS_PATH и схемы обмена маршрутами внутри конфедерации.

Рис. 9.11. *Централизация в конфедерации*

Конфедерации против отражателей маршрутов

Выбор между использованием конфедераций или отражателей маршрутов — нелегкая задача. Хотя различные организации эмпирическим путем получили определенные результаты от использования этих схем, компания Cisco рекомендует использовать отражатели маршрутов для обеспечения полносвязности по IBGP в крупных сетях. Реальные схемы доказали, что отражатели маршрутов более гибкие в настройке и обслуживании. С другой стороны, конфедерации могут применяться для обеспечения работы протоколов IGP в отдельно взятой подсистеме AS, независимо от IGP в других подсистемах AS, что помогает избежать нестабильности в крупных AS.

В некоторых случаях отражатели маршрутов могут применяться совместно с конфедерациями. В частности, AS можно разделить на подсистемы, в каждой из которых один отражатель маршрутов будет обслуживать только подсистему AS.

Какой бы из методов вы не использовали, всегда следует помнить о его ограничениях и влиянии на систему маршрутизации. С учетом этого вам следует планировать дальнейшее развитие сети.

Ограничение роста IGP-инфраструктуры

Необходимость ограничения роста сетей объясняется тем, что в результате бесконтрольного разрастания сети стало трудно управлять протоколами IGP. Причем не имеет значения, имеете ли вы дело с устаревшим RIP версии I или с новейшими протоколами кратчайшего открытого маршрута Open Shortest Path First (OSPF) и взаимодействия промежуточных систем Intermediate System-to-Intermediate System (IS-IS), но рано или поздно вы столкнетесь с проблемой масштабируемости вашей сети. До сих пор в этой главе мы обсуждали отражатели маршрутов и конфедерации с точки зрения решения проблемы управления ростом IBGP. Один из способов управления ростом IGP-инфраструктуры заключается в сегментировании AS между несколькими регионами, в каждом из которых поддерживается работа по определенному протоколу IGP. В свою очередь, отдельные регионы должны объединяться по BGP. При такой схеме построения сети стабильность работы сети в одном регионе не будет влиять на стабильность работы другого.

Какими же критериями при выборе точек сегментирования должны руководствоваться разработчики сетей? Очевидно одно: Internet представляет собой одну огромную сеть, которая не может обслуживаться ни одним из протоколов IGP, поэтому она сегментируется с помощью протокола BGP.

Итак, чем же определяется размер сети? Количеством маршрутизаторов или маршрутов? И какое это количество должно быть? На эти вопросы вы услышите различные ответы, которые в основном зависят от устойчивости к ошибкам протокола IGP, средств, используемых для управления ростом инфраструктуры сети и контроля стабильности, а также того, предоставляет ли BGP-сегментирование больше преимуществ при меньших затратах (в долларах и усилиях), чем средства IGP.

Такие протоколы, как OSPF и IS-IS, предоставляют определенные методы контроля стабильности маршрутов и средства для суммирования маршрутов. Но даже несмотря на все эти возможности, lGP-инфраструктура имеет тенденцию к разрастанию. На сегодняшний день можно ориентироваться на предельный объем таблиц IP-маршрутов порядка 2000-3000 внутренних IGP-маршрутов. При достижении этой границы нужно внимательно проанализировать, не продолжает ли их количество расти. Приведенное нами количество маршрутов не может создавать серьезные проблемы, так как на сегодняшний день транзитные BGP-маршрутизаторы в сети Internet практически без проблем обрабатывают до 75000 маршрутов. Проблемные ситуации возникают вследствие отказов оборудования или нестабильности линий доступа, когда информация о маршрутах не достигает пунктов назначения и они периодически стремятся сойтись в одну точку, что вызывает так называемое "зависание" сети.

Означает ли это, что сети с 3000 IGP-маршрутов нуждаются в сегментировании с помощью BGP? Необязательно. В большинстве случаев реструктуризация самого IGP с использованием технологий сегментирования и суммирования маршрутов может способствовать повышению степени управляемости сети.

Чтобы понять, почему необходима такая осторожность при сегментировании с помощью BGP, следует представлять, какому риску подвергается AS при сегментировании. Основная мощь протоколов IGP, особенно тех, что основаны на анализе состояния канала, заключалась в конвергенции — их способности быстро адаптироваться к изменениям сети. Еще одна сильная сторона IGP заключается в их способности поддерживать определенный уровень избыточности и распределять нагрузку в сети.

С другой стороны, протокол BGP был специально разработан для поддержки правил маршрутизации между различными AS без акцента на конвергенцию. При сегментировании доменов маршрутизации с помощью BGP конвергенция может быть реализована внутри полученных мелких сегментов, но она снижается при пересечении локально администрируемых AS вследствие зависимости BGP от TCP-сеансов, с помощью которых проводится передача обновлений маршрутов.

Еще один недостаток: для задания и соблюдения правил маршрутизации BGP внутри сегментов требуется дополнительное вмешательство со стороны пользователя. Как уже говорилось, единственным средством влияния на работу системы маршрутизации BGP является манипулирование атрибутами. Очевидно, что соблюдать правила маршрутизации в нескольких сегментированных подсистемах AS намного труднее, чем в отдельном протоколе IGP. Понимание всех этих аспектов необходимо разработчикам при создании сетей. Более подробное освещение этой проблемы не входит в круг вопросов, рассматриваемых в этой книге. Для получения полной информации по этой теме обратитесь к книге *"Решения для крупномасштабных IP-сетей " ("Large-Scale IP Network Solutions*'/•

В этом разделе мы рассмотрим два метода сегментирования AS:

- разделение нескольких регионов с помощью IBGP;
- разделение нескольких регионов с помощью EBGP.

Сегментирование AS с несколькими регионами, разделяемыми по IBGP

Автономная система может быть разделена на регионы, в каждом из которых используются различные протоколы IGP. Регионы логически взаимосвязаны по полносвязной схеме посредством IBGP. Для обеспечения избыточности регионы могут быть объединены физически с использованием схемы подключения "каждый с каждым", как это показано на рис. 9.12.

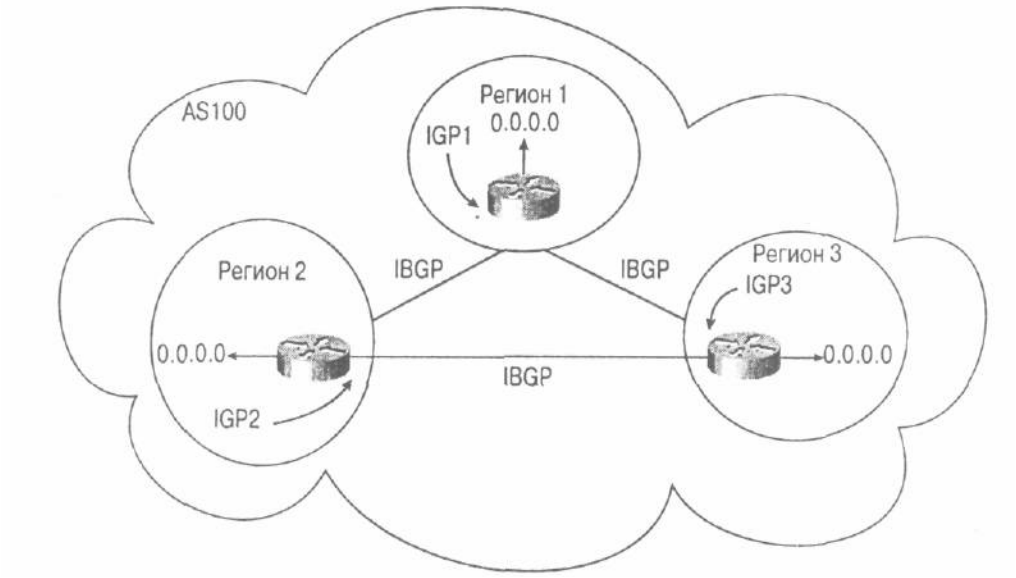


Рис. 9.12. *Объединение регионов с помощью IBGP*

Каждый регион будет преобразовывать свои IGP-маршругы в IBGP-маршругы и анонсировать внутри региона маршрут по умолчанию. В результате в каждом регионе будет сформирован маршрут по умолчанию, указывающий на граничный BGP-маршрутизатор, для всех пунктов назначения, не принадлежащих к данному региону. Региональные граничные маршрутизаторы будут передавать все маршруты от всех регионов (по IBGP) и соответствующим образом направлять трафик. Каждый регион надежно защищен от нестабильностей, порождаемых другими регионами, так как внутренние маршрутизаторы без поддержки BGP открыты только для маршрутов, которые принадлежат соответствующему региону. При необходимости организовать доступ с помощью динамического протокола маршрутизации между региональными граничными маршрутизаторами, участвующими в полносвязной схеме IBGP, вы можете использовать IGP отдельно.

Однако при таком варианте построения сети отсутствует одна важная деталь — соединение с Internet. Подключение подобной схемы к Internet требует дополнительного планирования. Как показано на рис. 9.12, каждый регион уже имеет маршрут по умолчанию

в другие регионы внутри AS. Проблема возникает, если BGP-маршрутизатор (в одном из регионов) не поддерживает синхронизацию с реальными маршрутами, полученными через точку соединения с Internet. В этой ситуации внутренние маршрутизаторы без поддержки BGP должны выбрать между маршрутом по умолчанию в Internet и маршрутом по умолчанию в другие регионы, как это показано на рис. 9.13.

Чтобы избежать этого, всем регионам необходимо всегда по умолчанию указывать на региональный граничный BGP-маршрутизатор или пытаться достичь заданных пунктов назначения в Internet или других регионах AS. Согласно этому требованию, соединения с Internet должны быть частью полносвязной IBGP-сети, как это показано на рис. 9.14.

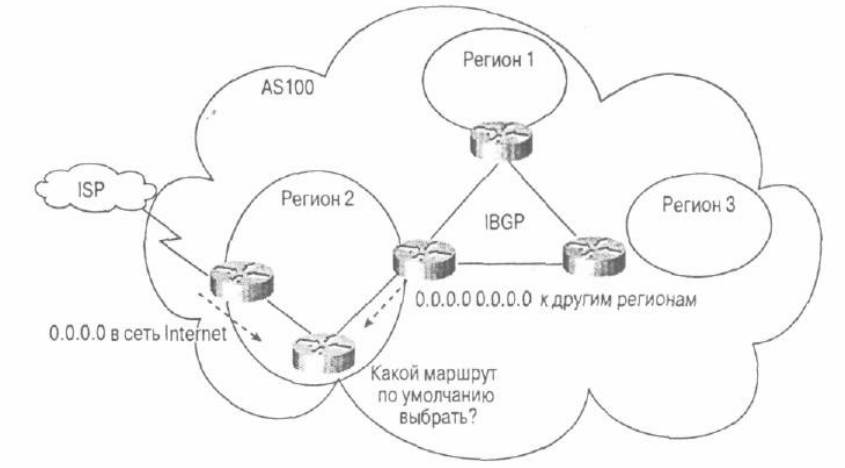


Рис. 9.13. *Конфликты маршрутов по умолчанию*

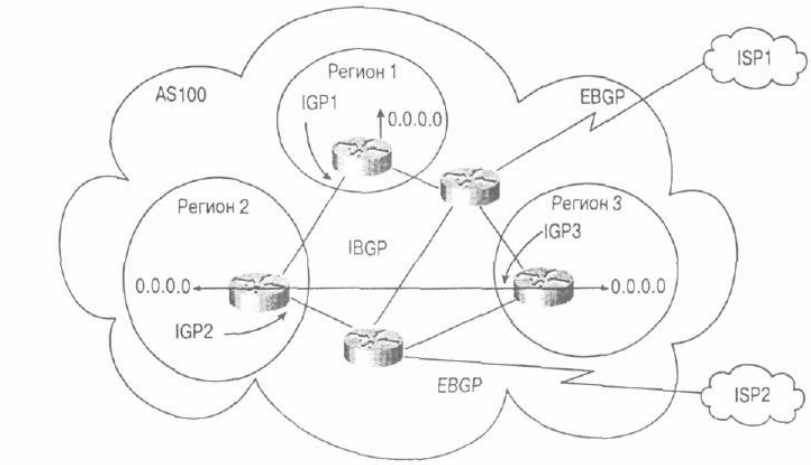


Рис. 9.14. *Несколько регионов с подключением к сети Internet*

Регионы 1, 2 и 3 объединены по полносвязной схеме с помощью протокола IBGP, который обслуживает и соединение с Internet. Внутренние маршрутизаторы без поддержки BGP в каждом регионе направляют трафик по умолчанию на граничный BGP-маршрутизатор, в котором содержатся все маршруты. Если пункт назначения принадлежит к другому региону, то трафик просто направляется в этот регион. В противном случае трафик будет пересылаться в Internet согласно правилам маршрутизации BGP.

Однако при использовании этого метода отсутствует гибкость в определении правил маршрутизации для каждого из регионов. Так как все регионы находятся в одной автономной системе, префиксы, принадлежащие различным регионам, отличить с помощью BGP не так-то легко. В более сложных структурах сети для того, чтобы отличить префиксы различных регионов, можно воспользоваться атрибутом сообщества COMMUNITY. Эта схема может применяться вместе с иерархической конфигурацией отражения маршрутов, что позволяет создавать крупные виртуальные частные сети. Об этом более подробно читайте в конце главы.

Сегментирование AS с несколькими регионами, разделяемыми по EBGP

Если для обработки трафика между регионами требуется определить гибкие правила маршрутизации, то каждый регион, например, может быть представлен как отдельная автономная система. Как известно для обеспечения взаимодействия между AS требуется протокол EBGP, а внутри AS — протокол IBGP. В каждой AS вы можете использовать определенный протокол IGP, если есть необходимость в дополнительном динамическом протоколе маршрутизации, который бы обеспечивал взаимодействие между различными EBGP-узлами. Этот метод сегментирования AS представлен на рис. 9.15.

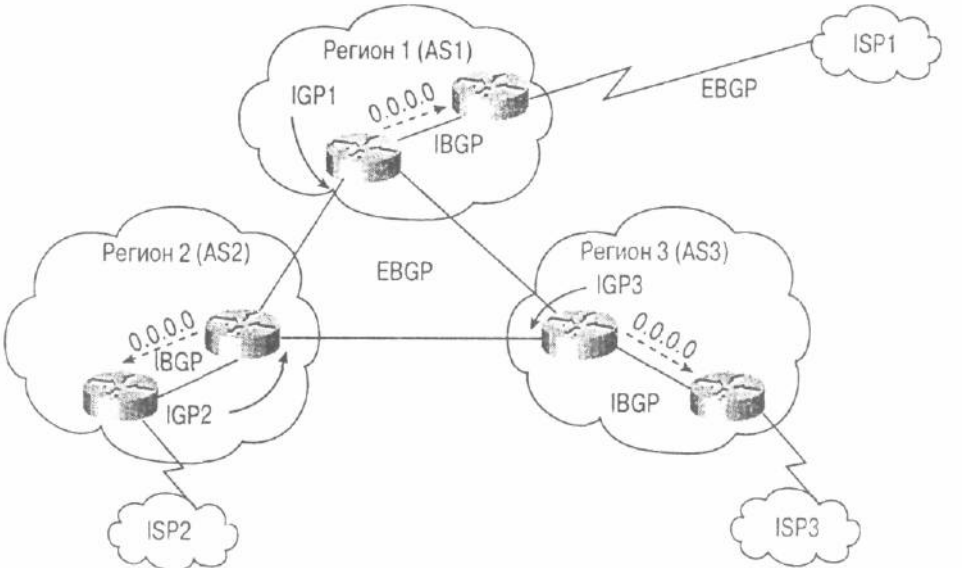


Рис. 9.15. *Несколько регионов, разделенных с помощью EBGP*

На первый взгляд кажется, что это наиболее оптимальное решение для управления ростом IGP-инфраструктуры. Теперь AS разделена на несколько более мелких AS, где каждая AS представляет собой отдельный регион со своим номером AS и имеет выделенное соединение с Internet. Каждая AS "вкладывает" в протокол BGP собственный набор IGP-маршрутов, которые передаются всем остальным регионам и в сеть Internet. Внутренние маршрутизаторы в каждом регионе по умолчанию направляют трафик на граничные BGP-маршрутизаторы, где содержатся все маршруты. Правилами маршрутизации BGP могут быть установлены локальные предпочтения для префиксов, поэтому региональные граничные маршрутизаторы для связи между регионами будут предпочитать соединениям с Internet использование внутренних маршрутов. Для подключения к сети Internet в регионах отдается предпочтение собственному провайдеру. Однако если канал с провайдером выходит из строя, то регион может воспользоваться доступом в Internet через провайдеров других регионов.

Подобная схема выглядит прекрасно на бумаге, но все изменится, когда вы попытаетесь воплотить ее в жизнь. Очень трудно обосновать в Американском реестре адресов Internet (American Registry for Internet Numbers — AR1N) или других региональных реестрах сети Internet (Regional Internet Registries — RIR) необходимость получения номеров AS для развертывания подобной сети. Так как количество номеров AS ограничено, они быстро расходуются. Скорее всего потребуется реальное обоснование, прежде чем в реестрах ARIN и RIR сетевому администратору выделят несколько номеров AS. К сожалению, аргумента о необходимости использования нескольких AS для улучшения стабильности работы IGP обычно недостаточно для такого обоснования. Вам следует проконсультироваться с сотрудниками локального регионального реестра сети Internet относительно того, что будет считаться "достаточным основанием".

Другие альтернативы включают использование частных номеров AS или BGP-конфедераций для управления ростом IGP-инфраструктуры. Мы обсудим эти способы в

следующих разделах.

Использование частных номеров AS

Использование частных номеров AS представляет собой ешс один способ деления крупных AS с несколькими регионами, связанных по EBGP. В регионах поддерживается протокол IBGP, а для связи между ними используется протокол EBGP. К тому же каждый региональный домен будет преобразовывать свои префиксы к виду префиксов BGP для облегчения межрегионального взаимодействия. Внутренние маршрута заторы в каждом регионе без поддержки BGP будут по умолчанию отправлять весь трафик на региональные граничные маршрутизаторы, где содержатся все маршруты. И наконец, вы можете организовать работу по одному из протоколов IGP, если есть необходимость в динамическом протоколе, который обеспечивал бы взаимодействие по EBGP между регионами.

Этот сценарий хорошо работает без подключения к сети Internet. Однако когда появляется соединение с Internet все частные номера AS должны быть скрыты от внешнего мира. Для сокрытия частных номеров AS требуется усложнение сетевой архитектуры. На рис. 9.16 показан один из вариантов решения этой проблемы, который может использоваться при подключении AS к сети Internet и использовании частных номеров AS.

На рис. 9.16 представлено, каким образом AS раздроблена на несколько частных подсистем AS. Теперь регион 1 ~ это AS65001, регион 2 — AS65002, а регион 3 — AS65003. В каждой частной AS могут поддерживаться взаимоисключающие протоколы IGP.

Для обеспечения соединений с Internet в качестве точки объединения всех частных AS используется AS 100. Очень важно отметить, что это — центральная магистральная AS с легальным номером AS, уникальным для сети Internet. Все частные AS будут посредством EBGP взаимодействовать с магистральной AS100 для обеспечения межрегиональных и Internet-соединений.

С целью недопущения передачи информации о частных номерах AS внешним Internet-провайдерам сетевой администратор может воспользоваться процедурой очистки AS_PATH. о которой мы говорили в главе 6. В AS 100 изымаются частные номера AS, поступающие от каждого региона, перед передачей BGP-обновлений внешним провайдерам Internet.

На рис. 9.16 AS65001 генерирует префикс 192.213.16.0/24. Для всех остальных частных AS этот префикс будет поступать с AS_PATH 100 65001. При передаче его внешним AS200 и AS300 частный номер AS изымается и AS_PATH будет иметь вид 100. Таким образом, все ваши сети будут в Internet объявлены как сгенерированные в AS100.

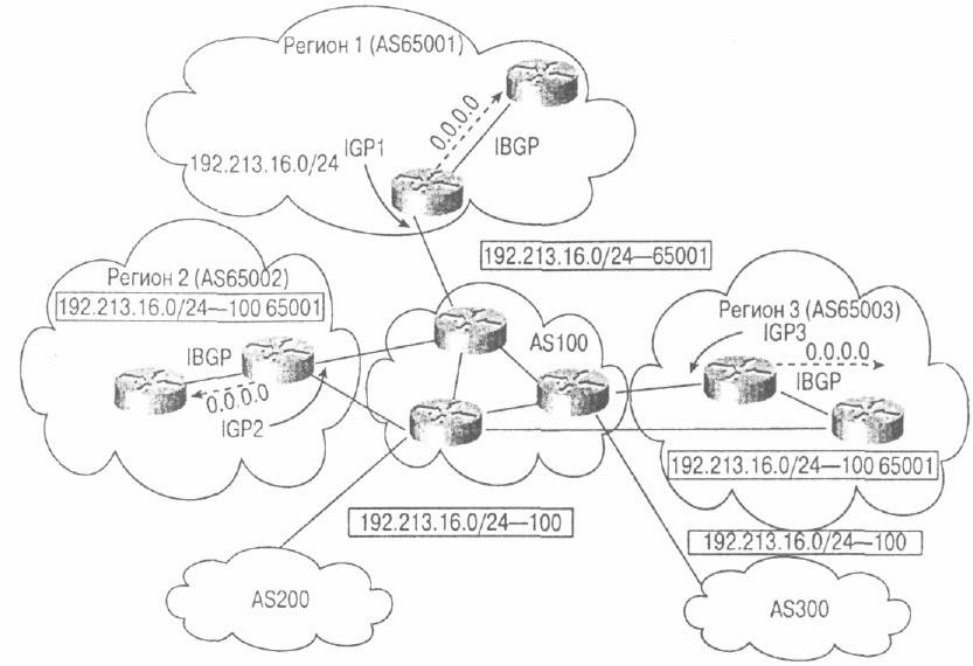


Рис. 9.16. *Частные AS при подключении к различным провайдерам*

Использование конфедераций для управления ростом IGP-инфраструктуры

Для обуздания роста **IGP-инфраструктуры** можно использовать конфедерации. Вы уже видели, каким образом конфедерации могут разделять AS на более мелкие подсистемы. Если в каждой подсистеме AS обеспечить поддержку различных протоколов IGP, то централизованная схема, представленная на рис. 9.16, приобретет определенную жизнеспособность. Итак, теперь 1GP используются независимо друг от друга, однако из внешнего мира AS воспринимается как единое целое. Для обеспечения межрегионального взаимодействия каждый IGP может подставляться в BGP. По-прежнему внутренние маршрутизаторы в каждом регионе, не поддерживающие BGP, будут по умолчанию ссылаться на региональный граничный BGP-маршрутизатор, где имеются все маршруты. Д1Я обеспечения доступа в Internet всем регионам по умолчанию будет использоваться центральная AS. Таким образом, вырисовывается схема подобная, приведенной на рис. 9.16.

Однако имеется и обратная сторона медали: конфедерации требуют "тонкой" настройки и не обеспечивают одинаковых возможностей по установке правил между подсистемами AS, так как вся AS рассматривается как единое целое. К тому же любая схема с использованием нецентрализованных конфедераций может привести к таким осложнениям, как неоптимальная маршрутизация, и, следовательно, — к затруднениям при выборе маршрута для передачи трафика внутри конфедерации.

Забегая вперед

До сих пор мы рассматривали протокол BGP, который в умелых руках может быть мощнейшим инструментом, делающим маршрутизацию более структурированным процессом. Теперь вы знаете, как управлять трафиком и как сегментировать AS на более мелкие управляемые элементы. Однако остается еще один аспект, который требует тщательного изучения, — это нестабильность маршрутов в сети Internet. Большое количество факторов обуславливает появление флуктуации (колебаний) маршрутов, которые, в свою очередь, порождают флуктуации трафика. Некоторые причины вы можете устранить самостоятельно, другие - нет. В настоящее время невозможно представить себе мир без Internet, и в ваших интересах уважать и защищать целостность этой сети. В следующей главе мы обсудим некоторые причины, порождающие нестабильность маршрутов и меры, которые могут быть приняты, чтобы устранить или по крайней мере снизить эффект от их воздействия.

Часто задаваемые вопросы

В — *У меня есть два узла: один — в Сан-Франциско, другой* — в Сан-Хосе. *Не лучше ли мне разделить их на две AS и организовать между ними взаимодействие по BGP вместо IGP, по которому они работают сейчас?*

О— Ваша схема не нуждается в сегментировании по BGP. Помните, что, хотя сегментирование и дает возможность гибкого управления, и повышает структурирование, оно требует соблюдения правил BGP. В небольших сетях, каковой является ваша сеть, вы можете добиться тех же результатов путем выполнения иерархической маршрутизации с помощью IGP.

В — *В моей сети недостаточно BGP-узлов для организации отражателя маршрутов. Что будет, если я буду использовать хотя бы один из них в качестве отражателя?*

О — Вы получите нормальную маршрутизацию с своей сети. Вам лишь нужно

понять, что в подобной модели необходимо полагаться на централизацию маршрутизаторов, которые будут пересылать BGP-маршруты в другую часть вашей сети. Однако отражатель маршрутов не только создает дополнительную нагрузку на процессор маршрутизатора, но является и точкой возможного отказа. Следовательно, необходимо предусмотреть дополнительные меры предосторожности, установив для обеспечения избыточности хотя бы два отражателя маршрутов. Кроме того, вам придется столкнуться с такими вещами, как группы взаимодействующих узлов и изменение атрибутов, о чем мы говорили в этой главе. Если в вашем случае нет повышенных требований к нагрузке на процессор маршрутизатора, то смело можете конфигурировать отражатель маршрутов.

В — *При работе с конфедерациями внешний EBGP-маршрут более предпочтителен, чем внешний маршрут между конфедерациями. Означает ли это, что я никогда не смогу использовать другую подсистему AS в качестве точки выхода?*

О — Нет. Вы всегда можете использовать атрибуты, как, например, атрибут локального предпочтения, для указания желаемой точки выхода.

В — *Так как локальные предпочтения не передаются между AS, они не будут передаваться и между подсистемами AS, не так ли?*

О — Это не совсем так. При внесении определенных изменений подсистемы распознают, что они взаимодействуют с внешними узлами внутри конфедерации, и обрабатывают все атриб^ы, которые обычно обрабатываются только с помощью протокола IBGP.

В — *Мне нужно сконфигурировать отражатели маршрутов, но программное обеспечение используемых маршрутизаторов не поддерживает этих функций. Нужно ли мне модернизировать все маршрутизаторы?*

О — Нет. Вам потребуется модернизировать лишь маршрутизаторы, которые планируется использовать в качестве отражателей маршрутов. Остальные маршрутизаторы будут выполнять роль обычных IBGP-спикеров. Это поможет вам постепенно реорганизовать всю вашу сеть.

Ссылки

[1] RFC 1966, "BGP Route Reflection: An Alternative to Full Mesh IBGP," www.isi.edu/in-notes/rfc 1966.txt

[2] RFC 1965, "Autonomous System Confederations for BGP,"www.isi.edu/in-notes/rfcl965.txt

- RFC 1930, "Guidelines for Creation, Selection, and Registration of an Autonomous System (AS)," www.isi.edu/in-notes/rfcl930.txt

[4] Raza, Khalid and Mark Turner. *Large-Scale IP Network Solutions* (Indianapolis, Ind.: Cisco Press, 2000)

Ключевые темы этой главы:

- **Источники нестабильности маршрутов в Internet.** Дается крпткий обзор причин, вызывающих нестабильность маршрутов в сети Internet.
- **Управление маршрутами и аннулирование содержимого кэша.** Компанией Cisco предлагается программное решение, позволяющее администраторам переконфигурировать атрибуты "на ходу", с минимальными последствиями для маршрутов и содержимого кэша.
- **Обновление BGP-маршрутов.** Обновление маршрутов протокола граничного шлюза обеспечивает механизм применения новых правил маршрутизации путем запроса обновленного набора маршрутов от другого узла, что не требует **сброса** соединения по BGP или прекращения анонсирования маршрутов.
- **Разгрузка маршрутов.** Процедура разгрузки маршрутов предполагает выявление и блокирование нестабильных маршрутов с целью обеспечения более высокой надежности как внутри, так и вне автономной системы.

Глава 10. Проектирование стабильных сетей на базе TCP/IP

Организация и поддержка стабильности маршрутов внутри и между сетями TCP/IP является решающим фактором для создания надежной связности в сети Internet и в других сетях на базе TCP/IP. Многочисленные ошибки при организации сетей и другие проблемы могут способствовать дестабилизации работы в Internet. В этой главе мы рассмотрим основные источники нестабильности маршрутов в сетях на базе TCP/IP и методы ее снижения.

Источники нестабильности маршрутов в Internet

Основным симптомом нестабильности маршрутов является исчезновение маршрута из маршрутной таблицы. Такой маршрут может периодически появляться и исчезать; это состояние называют *колебанием* или *мерцанием (flapping)* маршрута. На уровне протокола маршрутизации в этом случае происходит следующее: в BGP посылается сообщение об обновлении маршрута, а затем этот маршрут сразу же удаляется. Маршрутизатор, который постоянно получает сообщения UPDATE и WITHDRAWN, должен пересылать эти сообщения своим соседям. Эти сообщения приходят во все сети, поддерживающие в глобальной сети Internet протокол BGP. Очевидно, что постоянное поступление этих сообщений в сеть значительно снижает ее производительность.

Ниже приведены факторы, которые вызывают нестабильность маршрутов в сети Internet.

- Нестабильность протокола внутреннего шлюза (Interior Gateway Protocol — IGP).
- Дефектны оборудования.
- Ошибки в программном обеспечении.
- Недостаточная мощность процессора.
- Недостаточный объем памяти.
- Модернизация и техническое обслуживание сети.
- Человеческие ошибки.
- Перегруженность соединений.

Нестабильность IGP

Динамическое преобразование IGP-маршрутов в BGP-маршруты может привести к нежелательным колебаниям маршрутов. В этом случае проблемы внутри домена маршрутизации могут распространиться и за его пределы. Как уже отмечалось в главе 6,

"Настройка параметров BGP", статическое преобразование маршрутов в BGP может облегчить решение этой проблемы.

Объединение маршрутов на граничных или центральных маршрутизаторах также может снизить потенциальную опасность, связанную с преобразованием маршрутов из IGP в BGP. При объединении маршруты переходят в BGP как один объединенный маршрут. Нестабильность какого-либо из маршрутов, входящих в состав объединенного маршрута, не влияет на стабильность всего объединенного маршрута.

Однако и по сей день некоторые разработчики сетей вынуждены прибегать к динамической маршрутизации по ряду причин.

- В реализациях протокола BGP разрешается статически объявлять только ограниченное число сетей. Лимит статически объявляемых маршрутов зависит от производителя оборудования и может колебаться. Однако независимо от размеров лимита, если требуется дополнительно задать несколько маршрутов к сетям сверх этого лимита, администраторам придется обратиться к преобразованию маршрутов из IGP в BGP.
- Некоторые администраторы побаиваются, что сети, маршруты к которым они статически объявляют, могут стать недоступными для маршрутизатора. Эти опасения вполне понятны, особенно для случаев, когда маршруты объявляются из различных точек AS. Объявление недоступного маршрута может привести к образованию "черной дыры".

Дефекты оборудования

Дефектные интерфейсы, бракованные узлы или некачественные линии связи также могут влиять на стабильность маршрутов. Интерфейс, который постоянно то отключается, то включается в работу может привести к изменению маршрутной информации. Некоторые дефекты оборудования можно исправить, но в основном они не могут быть устранены силами обслуживающего персонала. Избыточность систем и соединений -- одно из важнейших средств обеспечения нормальной работы сети при отказах оборудования или при выходе из строя канала связи, но при физической неисправности узла прерывается процесс маршрутизации, а это влечет за собой лавинообразные отказы в обслуживании на всей протяженности маршрута.

Ошибки в программном обеспечении

Ошибки в программном обеспечении (на жаргоне специалистов "баги" - от англ. bugs) могут привести к отказам различных систем и повлиять на стабильность работы всей сети. Разработчики протоколов маршрутизации постарались выявить все дефекты в программном обеспечении перед тем как их продукты поступили на рынок. Однако практически невозможно предусмотреть все ситуации, которые могут возникать в реальных сетях. Администраторам следует всесторонне протестировать новое программное обеспечение или его новые функции в тестовых лабораториях или в наименее критичных к простоям сегментах сети для того, чтобы составить определенное представление о программном обеспечении, прежде чем внедрять его во всей сети.

Недостаточная мощность процессора

Чем больше обновлений маршрутов и сеансов связи обслуживает маршрутизатор, тем большая мощность процессора требуется для выполнения этих операций. Представьте себе маршрутизатор грузовиком с колесной схемой 4x4, а процесс маршрутизации и перегрузку трафика — перевозимым грузом. Удивит ли вас, если такой грузовик испытывает

определенные трудности при движении с грузом в 20 тонн? Правильно подобрать соответствующую мощность процессора очень важно для нормальной маршрутизации в сети.

Недостаточный объем памяти

Кроме объема памяти, который требуется маршрутизатору для нормальной работы операционной системы, значительный объем памяти необходим для хранения таблиц маршрутов, их кэширования, ведения баз данных и других задач. Маршрутизатор, который исчерпал свою память, может прекратить работу, что приведет к потере информации обо всех маршрутах, которая была получена от других узлов или посылалась на другие узлы.

В терминах протокола BGP, запись с информацией о маршруте состоит из записи в таблице пересылки IP-маршрутов и из соответствующей записи в таблице BGP-маршрутов. Сегодня в таблицах маршрутов сети Internet указывается более 75000 маршрутов, и это число ежемесячно увеличивается. Системы, получающие все маршруты из Internet через одного или нескольких провайдеров, едва справляются с таким объемом информации (если вообще справляются) при наличии 32 Мбайт памяти (для хранения BGP-информации и другой маршрутной информации). Большинство провайдеров старается нарастить объем памяти в своих системах до 96, 128 и даже до 256 Мбайт только для обслуживания маршрутных таблиц. Недостаточный объем памяти очень часто может служить причиной нестабильности, так как маршрутизатор, работавший с недостаточным количеством памяти, в дальнейшем может собирать разрозненную информацию из фрагментированной памяти и стать постоянным (до перезагрузки) источником колебаний маршрутов.

Модернизация и техническое обслуживание сетей

Сети, по сути, — динамические образования. И улучшение производительности, и укрупнение телекоммуникационных узлов, и расширение сети — все это требует внесения изменений в структуру сети. Изменения могут быть также связаны с переходом на новые версии программного обеспечения или сетевого оборудования, добавлением новых соединений, с расширением полосы пропускания или переделкой топологии сети (переезд организации в другие помещения и т.п.).

По очевидным причинам администраторы предпочитают останавливать работу сети в то время, когда она наименее интенсивно используется. В некоторых сетях время простоя не может превышать одного часа даже в ночное время суток, вследствие различий временных зон. Несмотря на эти трудности период модернизации — не то время, когда пользователи испытывают наибольшие проблемы при работе в сети, так как администраторы обычно разрабатывают запасной план и всегда могут вернуться к старой схеме, если новая схема не заработала, как ожидалось. Этот период может длиться и в течении следующего дня, если ошибки в конфигурации или проблемы с программным и аппаратным обеспечением возникнут у пользователя, который только на следующий день включился в работу. Возвращение к старой схеме на этом этапе не является решением проблемы. К сожалению, чтобы исправить подобную ситуацию администраторы иногда начинают добавлять или изменять параметры настройки как говорится "на ходу", что может не только не улучшить, а даже усугубить ситуацию.

Чтобы снизить вероятность нанесения ущерба сети, следует вначале смоделировать изменения, которые вы собираетесь вносить, на нерабочем участке сети, если это возможно. Кроме того, нельзя одновременно вносить сразу несколько изменений, действуйте постепенно шаг за шагом. Например, для провайдера неблагоразумно модернизировать программное обеспечение на основном маршрутизаторе, параллельно заменяя кабели и другое сетевое оборудование. Тщательное планирование и моделирование сети — ключ к проведению успешной модернизации.

Человеческие ошибки

Большинство проблем, связанных с нестабильностью сети, обычно вызваны человеческим фактором, так как администраторы прибегают к различным уловкам при выполнении правил администрирования или перенастраивают систему, не зная возможных последствий этого. Работая с сетью сложной структуры, очень легко допустить ошибку. Фильтрация одного маршрута или сервиса может привести к тому, что вся AS будет изолирована от глобальной сети. Администраторы должны предвидеть проблемы еше до того, как они появятся.

Вот лишь один пример ошибки такого рода. Маршрутизатор может посылать своим соседям BGP-маршрут по умолчанию вида 0.0.0.0. Если вы недостаточно осторожны, трафик может пойти по нежелательному маршруту. Точно так же, как в чьи-то обязанности входит рассылка правильных маршрутов по умолчанию, ваша обязанность -обезопасить себя, т.е. убедиться в том, что вы 'фильтруете любые нежелательные маршруты по умолчанию, или в том, что процесс протекает именно так, как вы предполагали. Список возможных человеческих ошибок достаточно длинный: кто-нибудь может непреднамеренно объявлять чужие сети, провайдер может неожиданно прекратить объявление маршрута к вашей сети или кто-то объединить сети, не подлежащие объединению. Основное правило в этом случае — не думать, что все будут играть по вашим правилам. Другие администраторы {обычно по невнимательности) руководствуются правилами, которые напрямую конфликтуют с вашими правилами, что может привести к серьезным нарушениям в работе сети и к падению производительности.

Перегруженность соединений

В некоторых случаях выход из строя какого-либо соединения может повлечь за собой перегрузку трафиком другого соединения. Это происходит вследствие того, что через соединение пропускается весь дополнительный трафик и оно работает на пределе своей мощности. Даже если такое соединение имеет достаточную пропускную способность, маршрутизатор может не справляться с дополнительной нагрузкой, если для этого у него нет вычислительных ресурсов.

В процессе борьбы с нестабильностью в реализациях протокола BGP появилось несколько полезных функций. Хотя-эти функции и не решают всех проблем, они значительно снижают нестабильность маршрутов.

Функции по обеспечению стабильности в BGP

Конечно разработка эффективных правил маршрутизации и правильная их настройка на маршрутизаторах являются основой построения стабильной сети. Кроме атрибутов BGP, имеется несколько функций, которые помогут в борьбе с нестабильностью.

- Управление маршрутами и аннулирование (очистка) содержимого кэша.
- Обновление BGP-маршрутов.
- Разгрузка BGP-маршрутов.

Управление маршрутами и аннулирование содержимого кэша

Б основе сеанса, проводимого по протоколу BGP, лежит определенное соединение, организованное средствами транспортного протокола между двумя соседними узлами. Само по себе такое соединение создается с помощью передачи сообщения OPEN, которое содержит такие параметры, как номер версии BGP. При обмене сообщениями об обновлениях маршрутов передается информация о различных их атрибутах - - метрике, сообществах и списке номеров AS, через которые проходит маршрут. Если администратор вносит изменения в атрибуты или правила, традиционные реализации протокола BGP требуют, чтобы TCP-сеанс с соседним узлом был сброшен (т.е. разорван и снова восстановлен) — тогда внесенные изменения вступят в силу.

К сожалению, каждый раз при сбросе TCP-сеанса прерывается и процесс маршрутизации. Когда сеанс сброшен, аннулируется все содержимое кэша, т.е. все маршруты исчезают, что порождает в Internet своего рода цепную реакцию. К моменту восстановления сеанса все маршруты и содержимое кэша должно быть восстановлено. Как видите, налицо реальный ущерб от подобных действий.

Компания Cisco Systems предлагает механизм, который называется *"мягкой перенастройкой" (soft reconfiguration),* позволяющий администраторам изменять значения атрибутов "на ходу", без разрыва TCP-сеанса, или вручную имитировать колебания маршрута. В этом случае кэш маршрутов не подвергается очистке, и влияние на всю систему маршрутизации минимально.

См. в главе 12 раздел "Управление маршрутами и аннулирование содержимого кэша"

Однако этот метод имеет недостаток, который заключается в установлении немодифицированных маршрутов {по отношению к базе Adj-RIB-In, которая должна быть сбалансирована с Adj-RIB-Out) от заданных узлов и хранении их в локальной памяти маршрутизатора. При использовании мягкой перенастройки на крупных узлах может значительно возрастать потребление ресурсов памяти. Согласно выведенному правилу для хранения каждого маршрута, поступающего от взаимодействующего узла, требуется 250 байт памяти.

Обновление маршрутов в BGP

Недавно было найдено еще одно решение, устраняющее недостаток связанный с повышением потребления памяти при мягкой перенастройке. Этот альтернативный метод называется *способностью к обновлению маршрута (mute refresh capability)* и основан на правилах ведения переговоров в сеансе BGP-4 (эти вопросы освещены в главе 5, "Протокол граничного шлюза Border Gateway Protocol версии 4") с целью облегчения запроса о повторном объяаае-ниц всех префиксов, полученных от взаимодействующего узла (т.е. его Adj-RIB-Out).

Разгрузка маршрутов

Еще один механизм, способствующий снижению нестабильности сети, -- *разгрузка маршрутов (route dampening).* Постоянно появляющийся и исчезающий маршрут побуждает BGP-узел генерировать сообщения UPDATE и WITHDRAWN, которые периодически поступают в сеть Internet. Огромное количество генерируемого трафика маршрутизации может загрузить всю наличную полосу пропускания и процессоры маршрутизаторов.

При разгрузке маршруты подразделяются на категории и могут быть *нормальными (behaved)* либо *аномальными (ill behaved).* Нормальные маршруты ведут себя с высокой степенью стабильности в течение заданного периода времени. Аномальные маршруты крайне нестабильны даже в течение короткого периода времени и на них должны быть наложены "штрафы" пропорционально ожидаемой нестабильности. Нестабильный маршрут должен подавляться (точнее, просто не объявляться) до тех пор, пока он не стабилизируется.

Для оценки ожидаемой стабильности маршрута используется его так называемая "история", т.е. статистика его поведения в недавнем прошлом. С целью сбора "истории" маршрута отслеживается количество колебаний маршрута за определенный период времени. С целью разгрузки маршрута при каждом колебании на него налагается штраф. Когда штраф достигает установленной заранее границы, маршрут просто подавляется. Однако даже после подавления маршрут может накапливать штрафы. Чем чаще колебания маршрута, тем быстрее он будет подавлен.

Похожие критерии приняты и для прекращения подавления и повторного объявления маршрута. Алгоритм для снижения штрафов является экспоненциальным выражением. Этот алгоритм основан на наборе параметров, которые пользователь может самостоятельно изменять. В оборудовании Cisco вы можете изменять значения следующих терминов и параметров.

- **Штраф (Penalty)** - увеличивающееся числовое значение, которое назначается маршруту каждый раз при его колебании.
- **Половина "штрафного" времени (Half-life)** -- настраиваемое числовое значение, определяющее период времени, в течение которого штраф должен уменьшиться наполовину.
- **Граница подавления (Suppress limit)** — числовое значение, которое сравнивается с величиной штрафа. Если штраф больше, чем граница подавления, то маршрут подавляется.
- **Граница повторного использования (Reuse limit)** -- настраиваемое числовое зна чение, которое сравнивается с величиной штрафа. Если штраф меньше, чем граница повторного использования, то подавление маршрут прекращается.
- **Подавляемый маршрут (Suppressed route)** -- маршрут, который не объявляется, даже если он доступен. Маршрут становится подавляемым, если значение штрафа превышает границу подавления.
- **"Историческая" запись (History entry)** — запись, которая используется **для** хранения информации о колебаниях маршрута. С целью мониторинга и вычисления уровня колебаний маршрута очень важно хранить в маршрутизаторе подобную информацию. Когда маршрут стабилизируется, "историческая" запись становится бесполезной и должна удаляться с маршрутизатора.

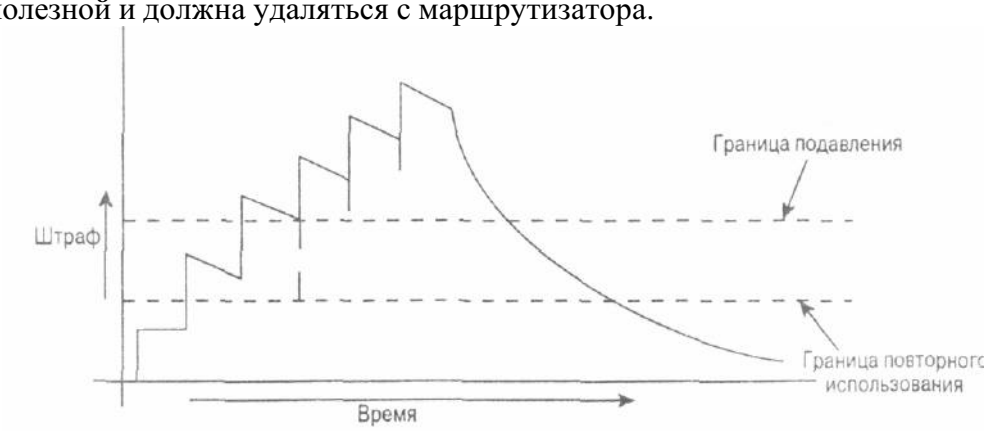


Рис. 10.1. *Определение размеров штрафа при разгрузке маршрута*

На рис. 10.1 представлен процесс получения штрафов при каждом колебании маршрута. Как видите, величина штрафов экспоненциально нарастает согласно параметру половины "штрафного" времени. С целью отражения сведений о колебаниях маршрута параметр половины "штрафного" времени может изменяться администратором. Большие величины

половины "штрафного" времени могут вызвать более медленное повышение штрафов, что, в свою очередь, повлияет на время, в течение которого маршрут будет подавляться.

Обеспечение стабильности внутри AS

Преимущества от разгрузки маршрутов заметны как внутри, так и за пределами автономной системы. Когда BGP-маршруты преобразуются в IGP-маршруты, очень важно, чтобы нестабильность отдельных маршрутов не влияла на внутреннюю маршрутизацию и не приводила к "зависанию" внутри AS. Именно здесь приходит на помощь механизм разгрузки маршрутов. Все маршруты, подверженные колебаниям, будут подавлены и не поступят в AS до тех пор, пока не достигнут определенной степени стабильности. На рис. 10.2 сравниваются сети, в которых наблюдается колебание EBGP-маршрутов, с использованием и без использования разгрузки маршрутов.

На рис. 10.2 маршруты Rl, R2 и R3 транслируются из протокола BGP в IGP и поступают в AS. Стрелки возле маршрута R2, указывающие вверх и вниз, обозначают его колебания. Маршруты передаются по IBGP и/или IGP, в зависимости от того, каким образом администратор выполняет преобразование маршрутов в AS. В любом случае мерцание маршрута R2 создает большую нагрузку на граничный маршрутизатор и на внутренние маршрутизаторы. Система, работающая под управлением протокола IGP, будет наводняться сообщениями об удалении и восстановлении нестабильного маршрута. В случае применения разгрузки маршрутов, аномальные маршруты будут подавляться (по достижении границы подавления) и сведения о них не будут поступать в AS.

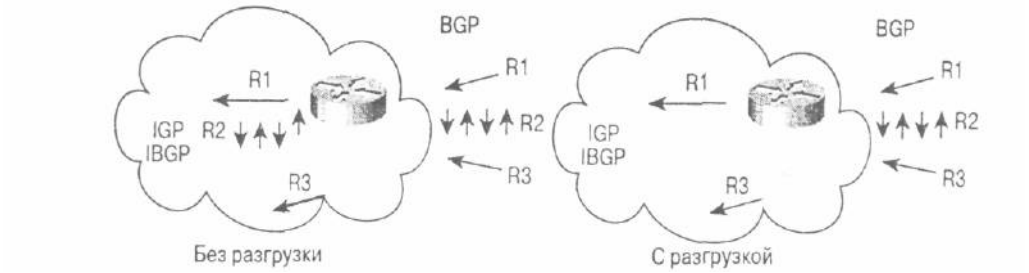


Рис. 10.2. *Влияние колебаний EBGP-маршрутов на IGP-маршруты*

Факторы нестабильности за пределами AS

Благодаря разгрузке маршрутов вы можете не допустить передачу нестабильных EBGP-маршрутов на другие узлы. Эти мероприятия позволяют сократить использование полосы пропускания канала и вычислительных мощностей граничных маршрутизаторов. Если вы являетесь провайдером для нескольких клиентов, то очень важно не нагружать собственную сеть (да и другие сети) нестабильными маршрутами, которые затем будут поступать в сети ваших клиентов. Случай, когда провайдер объявляет сети клиента как часть объединенного маршрута, является исключением. Объединенный маршрут всегда будет стабильным, даже если большинство его элементов нестабильны. Тем не менее вызывает тревогу нестабильность клиентских маршрутов внутри AS провайдера. Когда сеть клиента по каким-либо причинам нельзя объединить (вследствие подключения к нескольким провайдерам или использования адресного пространства, которое не входит в диапазон адресов провайдера), то нестабильные маршруты могут попадать за пределы сети.

При использовании процедуры разгрузки граничный маршрутизатор подавляет колеблющиеся клиентские маршруты. Подавление маршрутов проводится согласно правилам разгрузки и параметрам, которые мы обсуждали в этом разделе. На рис. 10.3 представлен пример разгрузки маршрутов у провайдера.

Одно из возможных последствий применения разгрузки маршрутов заключается в том, что клиент, даже после стабилизации маршруга, будет наблюдать несколько коротких отказов в обслуживании. На рис. 10.3 показаны колебания клиентского маршрута R2. Если провайдер на своем узле организовал разгрузку маршрутов, то маршрут R2 будет оштрафован и подавлен на определенный промежуток времени, величина которого зависит от интенсивности колебаний маршрута. Маршрут R2 может подвергаться разгрузке в течение нескольких минут. Но и после прекращения колебаний маршрута R2 величина полученных им штрафов может превышать границу повторного использования, поэтому он

не может повторно использоваться, пока не погасит все штрафы в пределах границы повторного использования. Иногда администраторы клиентских сетей тщетно пытаются выяснить причину, по которой их подсети недоступны для внешнего мира. Если администраторы не подозревают о том, что их маршруты подвергаются разгрузке, они могут попытаться исправить ситуацию другими средствами, вызывая тем самым еше более интенсивное колебание маршрутов и подвергая их еще большим штрафам. Наилучшее решение -- выяснить у провайдера, получает ли он информацию о ваших маршрутах, и если — да, то проверить, почему они не объявляются. Обычно провайдеры руководствуются строгими правилами и по запросу клиента могут и не отменить разгрузку маршрутов. Единственное, что в такой ситуации может предпринять провайдер, -- это сбросить информацию об "истории" маршрутов, которые подвергались разгрузке. Конечно, эта операция выполняется при условии, что клиент испытывает определенные проблемы в работе системы маршрутизации, вызывающие флуктуацию маршрутов.

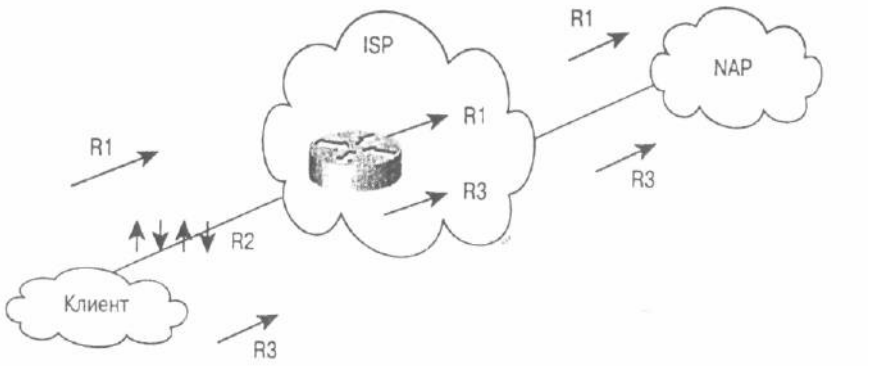


Рис. 10.3. *Разгрузка маршрутов на узле провайдера*

С другой стороны, провайдеры сами могут быть источниками нестабильности, и их влияние на маршруты клиента может быть намного шире. Если при передаче всех маршрутов между клиентом и провайдером или между двумя провайдерами происходят колебания маршрутов, то граничные маршрутизаторы примут на себя первый удар.

Предположим, что вы получаете полные маршруты из сети Internet (в настоящее время около 75000) от нескольких провайдеров. Теперь представьте, что 5 % этих маршрутов (около 3750) мерцают каждые 2 минуты. Ваш граничный маршрутизатор попросту не справится с такой нагрузкой.

Без разгрузки маршрутов в вышеописанном случае очень трудно определить, что происходит на самом деле. Все, что вам известно, — нагрузка на ваш граничный маршрутизатор лавинообразно возрастает. При использовании разгрузки маршрутов все нестабильные маршруты генерируют "исторические" записи, в которых фиксируется уровень стабильности того или иного маршрута. После выявления всех нестабильных маршрутов, легко можно определить, откуда они поступают (по адресу следующего узла). Хотя разгрузка маршрутов и не решает проблемы, она позволяет выявить ее источник. Когда вы обнаружите виновного, можно временно прекратить BGP-сеанс с провайдером и. позвонив в службу технической поддержки провайдера, начать жаловаться.

В заключение хотелось бы отметить, что нестабильность маршрутов в Internet будет в той или иной степени влиять на работу сети. Обязанность каждого администратора минимизировать колебания своих маршрутов. Это достигается четким представлением о том, что они делают и для чего. Провайдеры также становятся более суровыми к источникам колебаний маршрутов. Некоторые провайдеры, например, применяют жесткие штрафы к маршрутам с длинными масками. Это, возможно, выглядит чрезмерным, однако с каждым месяцем все тяжелее контролировать все, что происходит в Internet. Возникает необходимость в создании своего рода "патруля маршрутизации", который бы следил за нарушителями правил маршрутизации и выписывал им штрафы.

Забегая вперед

Мы уже достаточно много говорили о структуре сетей и работе системы маршрутизации. Если вы хотите для себя составить перспективу развития этой области путем изучения современных образцов архитектуры систем маршрутизации, то лучшее еше впереди. В последующих главах мы коснемся рассмотренных нами ранее схем построения сетей и организации маршрутизации в них на примерах реальных конфигураций с использованием программного кода операционной системы для маршрутизаторов Cisco 1OS.

Примеры конфигураций и листинги приводятся с полными пояснениями тех или иных действий и результатов их выполнения. Настоящие примеры листингов взяты с маршрутизаторов Cisco, и с их помощью мы рассмотрим функционирование различных атрибутов BGP, а также влияние различных конфигураций маршрутизатора на таблицы маршрутов. Мы надеемся, что примеры конфигураций, приведенные в двух следующих главах, помогут вам достичь более высокого профессионального уровня в интеграции ваших сетей в глобальную сеть Internet.

Часто задаваемые вопросы

В — *Если я указываю максимальное количество принимаемых от взаимодействую щего узла префиксов, например 100, и разрешаю мягкую перенастройку узла, отразится ли этот лимит на количестве маршрутов, которые могут храниться в памяти (т.е. не больше 100)?*

О — Нет. Все маршруты, сведения о которых получены от другого узла, (т.е. в базе Adj-RIB-in еще до применения входных правил маршрутизации) могут храниться в памяти.

В — *У меня проблемы с граничным маршрутизатором, так как канал с провайдером постоянно пропадает и появляется, что вызывает колебания моих маршрутов. Поможет ли настройка процедуры разгрузки на моем маршрутизаторе стабилизировать его работу?*

О — В определенной степени — да. В этом случае разгрузка маршрутов может стаби-I лизировать вашу AS, если при этом проводится преобразование BGP-маршрутов в IGP-маршруты. Однако ваш граничный маршрутизатор по-прежнему будет обнаруживать нестабильность маршрутов, поступающих от вашего провайдера. Лучше всего в : этом случае связаться с провайдером и выяснить у него причины нестабильности.

В — *Я объявляю IGP-маршруты с помощью BGP. Однако мои IGP-маршруты очень | нестабильны, что влияет и на BGP-маршруты, заставляя их колебаться. Поможет ли в такой ситуации разгрузка маршрутов?*

О — Нет. Если вы настроите разгрузку маршрутов на своем граничном маршрутизаторе, то ваши маршруты вообще не будут объявляться. Вам нужно найти источник нестабильности IGP-маршрутов.

В — *Некоторые из моих внутренних маршрутов мерцают, что побуждает моего провайдера проводить их разгрузку. Я не могу выяснить, в чем проблема. Что мне делать в этом случае?*

О — Вы всегда можете задать маршруты статически и транслировать их в протокол BGP. Таким образом, они всегда будут объявляться, независимо от того, доступны они или нет. Лучше (если есть такая возможность) задать ваши маршруты как часть объединенного маршрута. Объединенные маршруты не исчезают из-за колебаний в их элементах. Они могут исчезнуть, если все составляющие их маршруты окажутся недоступными. I

В — *Являясь провайдером, я не хочу "штрафовать" все префиксы одинаково. Существует ли механизм назначения различным префиксам различных параметров разгрузки?*

О— Да. Компания Cisco предлагает вам достаточную гибкость при выборе параметров разгрузки по таким критериям, как префикс IP, AS_PATH или сообщество.

Часть IV. Настройка маршрутизаторов для работы в сетях TCP/IP

В этой части...

В предыдущих частях мы рассмотрели концепции и схемы маршрутизации, не останавливаясь на настройке оборудования. В главах 11 и 12 вы найдете примеры конфигураций для большинства рассмотренных нами ранее в частях II и III концепций и процедур. В главе 11 основное внимание уделяется настройке BGP-атрибутов. А в главе 12 мы рассмотрим примеры конфигурации для более сложных схем сети, с которыми наиболее часто придется иметь дело администраторам при определении правил маршрутизации. Однако помните, что вы не сможете просто переложить приведенные нами примеры на ваши собственные правила маршрутизации. Они скорее являются моделями определенных решений для различных схем маршрутизации, с которыми вы можете столкнуться на практике при разработке правил маршрутизации по мере расширения сети. Вам необходимо экстраполироваться и привести эти модели в соответствие с вашей ситуацией.

Ключевые темы этой главы:

- **Сеанс связи между взаимодействующими маршрутизаторами.** Приведены примеры конфигураций для первичной настройки маршрутизации. В этом разделе рассмотрен также синтаксис основных команд конфигурации маршрутизаторов.
- **Фильтрация маршрутов и управление атрибутами.** Обсуждается создание карт маршрутов в BOP, фильтрация на основе NLRI и на основе атрибута AS_PATH.
- **Группы взаимодействующих узлов.** Приводятся примеры настройки и использования групп взаимодействующих узлов.
- **Источники обновления маршрутов.** Описаны настройки динамического и статического вложения информации в BOP на примерах.
- **Наложение протоколов ("черные ходы").** Приводятся примеры настроек для изменения параметра дистанции с целью задания порядка приоритетности различным маршрутам.
- **Атрибуты протокола** BGP. Рассматриваются примеры настройки атрибутов NEXT_HOP, AS_PATH, LOCAL PREFERENCE, MED и COMMUNITY.
- **Агрегация в BGP-4.** Приводятся примеры нескольких вариантов агрегации маршрутов.

Глава 11. Настройка основных функций и атрибутов BGP

Эта глава первая из двух, где мы рассматриваем примеры конфигурации маршрутизаторов. Если вы уже знакомы с используемыми концепциями, то можете перейти к анализу примеров настройки и научиться задавать команды для настройки основных функций и атрибутов BGP. В этой главе мы сосредоточимся на настройке маршрутизаторов для работы в составе простейших структур, а в следующей — перейдем к более сложным схемам сети.

Даже если вы уже воспользовались ссылками из предыдущих глав на соответствующие примеры конфигурации, настоятельно рекомендуем еще раз проанализировать их, чтобы лучше закрепить теоретический материал из предыдущих глав. В дополнение к кодам настройки маршрутизаторов мы включили в эту главу множество таблиц маршрутов, которые, надеемся, помогут глубже понять смысл ваших действий и их возможные последствия.

Главы 11 и 12 не призваны заменить руководство пользователя по работе с маршрутизаторами Cisco, и в них вы не найдете все команды и варианты настройки маршрутизаторов. В этих главах представлены примеры конфигурации маршрутизаторов для общих ситуаций, которые чаще всего возникают при подключении сетей к Internet. В вашей сети возможно потребуется скомбинировать несколько вариантов или способов подключения с целью выработки наиболее эффективного набора правил маршрутизации.

Далее мы увидим, что AS может играть роль клиента, провайдера или даже обе эти роли сразу. Постарайтесь не запутаться в этих ролях AS или в IP-адресации, которая может показаться вам несколько необычной. Все приведенные примеры — всего лишь упражнения, которые помогут вам понять, как функционирует протокол BGP, и оптимально реализовать его возможности.

Сеанс связи между взаимодействующими маршрутизаторами

В этом примере демонстрируются различные типы BGP-сеансов между взаимодействующими узлами, с которыми вы можете столкнуться на практике. Рассмотрим рис. 11.1.

Сеанс взаимодействия по 1BGP происходит внутри AS3 между физическим адресом маршрутизатора RTF и петельным адресом маршрутизатора RTA. Кроме того, между AS3 и AS1 сформирован сеанс по EBGP с использованием двух IP-адресов из одного сегмента на маршрутизаторах RTA и RTC. Еще один EBGP-сеанс сформирован между маршрутизаторами RTF в AS3 и RTD в AS2 с использованием IP-адресов, которые находятся в разных сегментах (мультиузловое соединение).

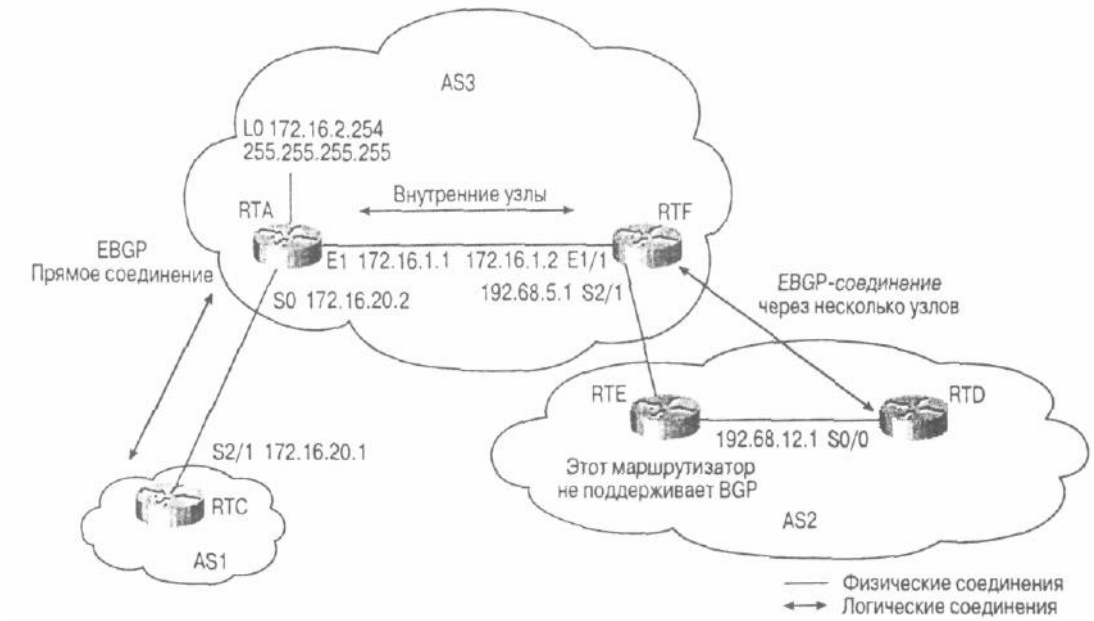


Рис. 11.1. *Сеанс между несколькими узлами*

Очень важно помнить, что TCP-соединение не будет установлено в BGP до тех пор, пока между двумя узлами не будет установлена связь по протоколу ЮР или если они соединены напрямую. В качестве протокола ЮР мы будем использовать OSPF, с помошью которого и будет обеспечиваться требуемая внутренняя связность. В листинге 11.1 показана конфигурация маршрутизатора RTA.

Листинг 11.1. *Конфигурация маршрутизатора RTA*

```
ip subnet-zero

interface Loopback0
 ip address 172/16/2/254 255/255/255.255

interface Ethernetl
 ip address 172.16.1.1 255.255.255.0

interface Serial0
 ip address 172.16.20.2 255.255.255.0

router ospf 10
 network 172.16.0.0 0.0.255.255 area 0

router bgp 3
 no syncronization
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.20.1 remote-as 1
 no auto-summary

ip classless
```

В примере конфигурации маршрутизатора RTA в листинге 11.1 вам, возможно, незнаком синтаксис команд. Синтаксис всех этих команд приведен в табл. 11.1. В ней этим командам даются пояснения применительно к схеме маршрутизации, представленной на рис. 11.1. В последующих примерах, которые вы найдете в этой главе, все команды маршрутизатора будут касаться в основном настройки работы по BGP, IGP или статической маршрутизации. Команды, с помощью которых проводится назначение интерфейсам IP-адресов, не будут рассматриваться, поскольку объем книги ограничен.

Таблица 11.1 *Команды конфигурации из листинга 11.1*

Команда	Пояснение
ip subnet-zero	Это команда глобальной настройки, если требуется настроить интерфейсы, которые подпадают под нулевые подсети (какой является 192.168.1.0/30). С переходом к бесклассовой междоменной маршрутизации применение нулевых подсетей стало довольно распространенным явлением и рекомендуется в качестве конфигурации по умолчанию
interface *type slot/port*	С помощью этой команды задается тип интерфейса и его номер на маршрутизаторе. Настройки, следующие за ней, будут относиться именно к этому интерфейсу. (Синтаксис *slot/port* может немного отличаться в зависимости от производителя оборудования). Обратите внимание, что на RTA задаются три команды interface — для каждого из трех соединений. Интерфейс *loopback* представляет собой программный интерфейс, эмулирующий интерфейс, который всегда находится в активном состоянии
ip address /p-*address mask* [secondary]	С помощью этой команды интерфейсу задается пара IP-адреса и маски. Например, IP-адрес порта Ethernet на RTA задается так: ip address 172.16.1.1 255.255.255.0
router process [process-id]	Эта глобальная команда определяет процесс (режим работы) маршрутизатора, такой как OSPF, RIP или BGP, и присваивает ему идентификатор процесса. Некоторым процессам, например RIP, нет необходимости присваивать идентификатор процесса. Например, при настройке маршрутизатора RTA команда router ospf 10 определяет, что процесс OSPF имеет идентификатор 10, а команда router bgp 3 указывает на BGP-процесс в автономной системе 3
network	Эта команда определяет сеть или (в случае работы по OSPF) интерфейсы, принимающие участие в маршрутизации
inverse mask	Обратите внимание на команду network, заданную на маршрутизаторе RTA. Как видите, там имеется запись 0.0.255.255 — т.е. за нулями следуют единицы. Это и есть инверсная маска, где нули точно определяют размер сети, а единицы представляют незначащие биты. Например, запись 172.16.0.0 0.0. 255. 255 указывает на любую сеть или IP-адрес вида 176.16.X.X. Инверсные маски можно использовать при организации списков разрешения доступа, а также совместно с командой network. В табл. 1 1 .2 приведена схема преобразования нормальных масок в инверсные в десятичной записи
area *area-number*	Эта команда используется для организации области с заданным номером в протоколе OSPF
neighbor	Этой командой задаются параметры соединения с соседними BGP-узлами и правила маршрутизации, которые будут использоваться между маршрутизатором и другими узлами. При настройке маршрутизатора RTA команда neighbor 172.16.1.2 remote-as 3 указывает на то, что BGP-сеанс будет устанавливаться между RTA и взаимодействующим узлом l 72. 1 6. 1 .2 в AS3
no syncronization	Эта команда отключает синхронизацию между BGP и IGP. Более подробно она описана в главе 6, "Настройка параметров BGP"
no auto-summary	Эта команда выключает в BGP режим автосуммирования по классам на границе основных сетей. Без этой команды BGP не будет рассылать маршруты в подсети основной сети, которые были преобразованы из IGP. Другими словами, сообщения об обновлениях маршрутов 172.16.1.0/24, 172.16-2.0/24 и т.д. будут рассылаться в виде одного сообщения о сети класса B 172.16.0,0/16. Суммирование на границе основных сетей необходимо выполнять только в том случае, если в AS объединены сети, принадлежащие одной основной сети (т.е. сети одного класса). Если нет явной необходимости в суммировании сетей, по умолчанию желательно запрещать эту процедуру
ip classless	Эта команда позволяет маршрутизатору пересылать пакеты, которые адресованы в неизвестные сети, соседним сетям, подключенным непосредственно к маршрутизатору. По умолчанию, когда маршрутизатор получает пакеты для подсети, которая правильно задана сточки зрения IP-адресации, но отсутствует в его таблицах маршрутов и для этой сети нет маршрута по умолчанию, он отвергает пакеты. Задав команду ip classless, вы разрешаете маршрутизатору пересылать эти пакеты по наилучшему маршруту в суперсеть

update-source *interface*	Эта команда, заданная совместно с выражением neighbor, определяет номер интерфейса, который будет использоваться в качестве IP-адреса источника во время BGP-сеанса с соседним узлом. Так, при настройке маршрутизатора RTA вторая строка с выражением neighbor указывает на то, что в качестве IP-адреса источника будет использоваться интерфейс LoopbackO
remote -as	Эта команда, заданная совместно с выражением neighbor, определяет номер AS удаленного взаимодействующего BGP-узла. При настройке RTA первое выражение neighbor указывает на то, что внутренний соседний BGP-узел 1 72. 1 6. 1 .2 принадлежит к локальной AS3. Третье выражение выражением neighbor указывает на то, что BGP-узел 1 72. 1 6.20. 1 принадлежит к AS 1

Таблица 11.2 *Преобразование префиксов C1DR в десятичные маски подсети*

Префикс CIDR	Десятичная запись	Инверсная запись
/1	128.0.0.0	127.255.255.255
/2	192.0.0.0	63.255.255.255
/3	224.0.0.0	31.255.255.255
/4	240.0.0.0	15255.255,255
/5	248.0.0.0	7.255.255.255
/6	252.0.0.0	3.255.255.255
/7	254.0.0.0	1.255.255.255
/8	255.0.0.0	0.255.255.255
/9	255.128.0.0	0.127.255.255
/10	255.192.0.0	0.63.255.255
/11	255.224.0.0	0.31.255.255
/12	255.240.0.0	0.15.255.255
/13	255.248.0.0	0.7.255.255
/14	255.252.0.0	0.3.255.255
/15	255.254.0.0	0.1.255.255
/16	255.255.0.0	0.0.255.255
/17	255.255.128.0	0.0.127.255
/18	255.255.192.0	0.0.63.255
/19	255.255.224.0	0.0.31.255
/20	255.255.240.0	0.0.15.255
/21	255.255.248.0	0.0.7.255
/22	255.255.252.0	0.0.3.255
/23	255.255.254.0	0.0.1.255
/24	255.255.255.0	0.0.0.255
/25	255.255.255.128	0.0.0.127
/26	255.255.255.192	0.0.0.63
/27	255.255.255.224	0.0.0.31
/28	255.255.255.240	0.0.0.15
/29	255.255.255.248	0.0.0.7
/30	255.255.255.252	0.0.0.3
/31	255.255.255.254	0.0.0.1
/32	255.255.255.255	0.0.0.0

Рассмотрим теперь конфигурацию маршрутизатора RTF (листинг 11.2).

Листинг 11.2. *Конфигурация маршрутизатора RTF*

```
ip subnet-zero

interface Ethernetl
 ip address 172.16.1.2 255.255.255.0

interface Serial2/l
 ip address 192.68.5.1 255.255.255.0

router ospf 10
 network 172.16.0.0 0.0.255.255 area 0
 network 192.68.0.0 0.0.255.255 area 0

router bgp 3
no syncronization
 neighbor 172.16.2.254 remote-as 3
 neighbor 192.68.12.1 remote-as 2
 neighbor 192.68.12.1 ebgp-multihop 2
 no аисо-summary

ip classless
```

Вы, наверное, обратили внимание на команду ebgp-multihop 2, которая используется как часть команды neighbor при настройке маршрутизатора RTF. Она показывает, что внешний BGP-узел не имеет непосредственного соединения с маршрутизатором RTF и доступен не более чем через два промежуточных узла. Помните, что команда применяется только в EBGP и не будет работать с IBGP. Величина, задаваемая в конце команды ebgp-multihop (в нашем примере 2), определяет значение времени жизни (Time To Live -- TTL), которое задается в заголовке IP-пакета. В листингах 11.3 и 11.4 представлены примеры конфигурации маршрутизаторов RTC и RTD.

Листинг 11.3. *Конфигурация маршрутизатора RTC*

```
ip subnet-zero

interface Serial2/l
 ip address 172.16.20.1 255.255.255.0

router bgp 1
no synchronization
 neighbor 172.16.20.2 remote-as 3
 no auto-summary

ip classless
```

Листинг 11.4. *Конфигурация маршрутизатора RID*

```
ip subnet-zero

interface SerialO/0
 ip address 192.68.12.1 255.255.255.0

router ospf 10
 network 192.68.0.0 0.0.255.255 area 0

router bgp 2
 neighbor 192.68.5.1 remote-as 3
 neighbor 192.68.5.1 ebgp-multihop 2
 no auto-summary

ip classless
```

В листинге 11.5 представлен сеанс между взаимодействующими узлами после того, как они установили соединение.

Листинг 11.5. *Соединение между взаимодействующими узлами (маршрутизатор RTF)*

```
RTF#show ip bgp neighbor
 BGP neighbor is 172.16.2.254, remote AS 3, internal link
 BGP version 4, remote router ID 172.16.2.254
 BGP state = Established, table version = 2, up for 22:36:09
 Last read 00:00:10, hold time is 180, keepalive interval is 60 seconds
 Minimum time between advertisement runs is 5 seconds
 Received 1362 messages, 0 notifications, 0 in queue
 Sent 1362 messages, 0 notifications, 0 in queue
 Connections established 2; dropped 1
Connection state is ESTAB, I/O status: 1, uread input bytes: 0
Local host: 172.16.1.2, Local port: 11008
Foreign host: 172.16.2.254, Foreign port: 179

BGP neighbor is 192.68.12.1, remote AS 2, external link
 BGP version 4, remote router ID 192.68.5.2
 BGP state = Established, table version = 2, up for 22:13:01
 Last read 00:00:00, hold time is 180, keepalive interval is 60 seconds
 Minimum time between advertisement runs is 30 seconds
 Received 1336 messages, 0 notifications, 0 in queue
 Sent 1336 messages, 0 notifications, 0 in queue
 Connections established 1; dropped 0
 External BGP neighbor may be up to 2 hops away.
Connection state is ESTAB, I/O status: 1, uread input bytes: 0
Local host: 192.68.5.1, Local port: 11016
Foreign host: 192.68.12.1, Foreign port: 179
```

Для маршрутизатора RTF соседний узел 172.16.2.254 является внутренним соседним узлом, который принадлежит AS3. Соединение между этими узлами организовано по BGP-4 с использованием версии таблицы 2. Версия таблицы изменяется каждый раз, когда происходит ее обновление.

Другим соседним маршрутизатором по отношению к RTF является узел 192.68.12.1, с которым также установлено соединение. Этот маршрутизатор является внешним соседним узлом, принадлежащим к AS2. Обратите внимание; в листинге указано, что этот узел доступен через два промежуточных узла (согласно команде ebgp-multihop).

Фильтрация маршрутов и управление атрибутами

Фильтрация маршрутов и управление атрибутами являются основой для реализации правил маршрутизации в BGP. В этом разделе мы рассмотрим следующее:

- карты BGP-маршрутов;
- списки префиксов;
- идентифицирование и фильтрация маршрутов на основе NLRI;
- идентифицирование и фильтрация маршрутов на основе анализа AS_PATH.

Для введения в действие новых правил маршрутизации в BGP, таких как фильтрация маршрутов или управление атрибутами, для входных правил маршрутизации (Input Policy Engine) нужно представить набор соответствующих маршрутов. Это можно сделать несколькими способами. Самый грубый из них ~ сбросить BGP-сеанс с помощью команды:

```
clear ip bgp [ * / address / peer-group] [soft [in / out]]
```

Однако с помощью мягкой перенастройки или механизма обновления BGP-маршрутов возможно более элегантное решение этой проблемы. Более детально об этих методах читайте в главе 12, "Настройка эффективных правил маршрутизации в сети Internet".

Карты BGP-маршрутов

Карты маршрутов используются в протоколе BOP для управления и внесения изменений в маршрутную информацию, а также для определения условий, согласно которым осуществляется распределение маршрутов между маршрутизаторами и процессами маршрутизации.

Формат команды для задания карты маршрутов следующий:

route-map map-tag [**permit | deny**] [*sequence-number]*

Здесь *map-tag* (метка карты) - имя, идентифицирующее карту маршрутов, а *sequence-number* (номер последовательности) указывает место, начиная с которого запись в карте маршрутов будет взаимосвязана с другими записями в этой же карте маршрутов. Записи в картах делаются последовательно.

Рассмотрим на примере задание карты маршрутов с именем MYMAP (листинг 11.6).

Листинг 11.6. *Карта маршрутов MYMAP*

```
route-map MYMAP permit 10
! First set of conditions goes here.
route-map MYMAP permit 20
! Second set of conditions goes here.
```

Когда в BGP к обновлениям маршрутов применяется карта MYMAP, вначале применяется наименьшая запись (в нашем случае запись 10). Если первый набор условий не выполнен, то применяется вторая запись и т.д. до тех пор, пока один из наборов условий не будет выполнен или пока не закончатся возможные наборы условий.

Часть карты маршрутов, где задаются условия, описывается командами match и set. Команда match задает критерий, по которому проводится сопоставление, а команда set определяет действие, которое должно быть выполнено, если обновление маршрута встречает условие, заданное командой match.

В листинге 11.7 показан пример конфигурации простой карты маршрутов.

Листинг 11.7. *Конфигурация карты маршрутов*

```
router bgp 3
neighbor 1.1.1.1 route-map MYMAP in
!
route-map MYMAP permit 10
match ip address 1
set metric 5
i
route-map MYMAP permit  20
set  local-preference  200

ip  access-list  1  permit  1.1.1.0   0.0.0.255
```

Из листинга 11.7 видно, что в карту маршрутов MYMAP включены маршруты, полученные от соседнего узла 1.1.1.1. Если маршрут в сеть 1.1.1.0/24 объявлен соседним узлом 1.1.1.1, то его метрике будет присвоено значение 5 (как указано в последовательности с номером 10) и больше никаких действий предприниматься не будет. Все остальные префиксы, полученные от соседнего узла, также будут представлены в карте маршрутов MYMAP в последовательности с номером 10. Однако если они не совпадают с условиями, заданными в этой последовательности, то будут сопоставляться последовательности условий с номером 20. Если же ни одно из условий не будет выполнено, то атрибут LOCAL_PREF для всех префиксов будет равен 200.

Существует два типа списков разрешения доступа — стандартные и расширенные. Основное их различие заключается в том, что стандартные списки разрешения доступа

применяются к IP-адресам источника, а расширенные списки — к адресам источника и пункта назначения или к адресам источника и сетевой маске. Для объявления стандартного списка разрешения доступа применяется глобальная команда access-list, а объявление расширенного списка мы рассмотрим далее в этой главе.

```
access-list access-list-number {deny / permit} source [source-wildcard]
```

Стандартный список разрешения доступа используется для сопоставления определенного IP-адреса хоста или сети с заданным в списке с целью разрешения или запрещения определенного маршрута. Рассмотрим пример стандартного списка разрешения доступа, представленного в листинге 11.8.

Листинг 11.8. *Стандартный список разрешения доступа*

```
router bgp  3
neighbor   1.1.1.1   route-map  MYMAP   in
!
route-map MYMAP  permit   10
match   ip   address   1
set metric   5
!
ip  access-list  1  permit  1.1.1.0   0.0.0.255
```

В листинге 11.8 список разрешения доступа с номером 1 идентифицирует все маршруты, поступающие от 1.1.1.X. (Обратите внимание на инверсную запись маски 0.0.0.255 вместо явного указания префикса 1.1.1.0/24, как это было сделано в предыдущем листинге, — таким образом разрешается префикс именно этой длины). Маршрут в форме 1.1.1.X будет совпадать с указанным в списке разрешения доступа и распространяться (так как задано ключевое слово permit, т.е. разрешить) с метрикой 5. Согласно логике, эти маршруты будут распространяться благодаря тому, что выполнены условия, заданные в карте маршрутов. Списки префиксов являются новым, более эффективным и интуитивно понятным способом фильтрации маршрутов. Более детально мы обсудим его позже.

Если сообщение об обновлении маршрутов не соответствует критерию, заданному в записи карты маршрутов, то в протоколе BGP последовательно применяются все записи, пока не будут предприняты какие либо действия или не исчерпаются записи в карте маршрутов. Если обновление маршрутов не соответствует ни одному из критериев, то оно не будет распространяться или модифицироваться каким-либо образом. Это обновление маршрутов будет просто отвергнуто.

Карты маршрутов могут использоваться как со входящими (in), так и с исходящими (out) обновлениями PGP-маршрутов. В листинге 11.9 представлена карта маршрутов MYMAP, которая используется применительно к исходящим обновлениям BGP-маршрутов, поступающих от узла с адресом 172.16.20.2.

Листинг 11.9. *Карта маршрутов MYMAP и исходящие обновления маршрутов*

```
router bgp 1
neighbor 172.16.20.2 remote-as 3
neighbor 172.16.20.2 route-map MYMAP out
```

Списки префиксов

Как уже отмечалось в главе 6, списки префиксов являются новейшим, более эффективным и интуитивно понятным способом идентификации маршрутов для их последующего сравнения с критериями или для фильтрации в протоколах маршрутизации. Давайте рассмотрим основные рекомендации по настройке списков префиксов.

В отличие от списков разрешения доступа, списки префиксов могут обновляться не

полностью, а частично. Это означает, что каждая запись в списке префиксов идентифицируется номером последовательности. Если необходимо добавить, удалить или изменить заданную в списке префиксов последовательность, то нужно лишь указать ее номер. Основное преимущество такой операции заключается в том, что, в отличие от списков разрешения доступа, не нужно удачять список и затем снова настраивать его при внесении каких-либо изменений.

В списках префиксов также имеет значение порядок следования записей, т.е. первая запись, которая соответствует условиям, и будет приниматься как руководство к действию. Как и в списках разрешения доступа, вы можете разрешить (permit) или запретить (deny) распространение сведений о том или ином префиксе. Списки префиксов также позволяют организовать фильтрацию не только отдельных префиксов, но и диапазонов префиксов.

Настройку списков префиксов мы обсудим позже. Отметим, что наши рекомендации ни в коем случае не должны использоваться вместо тех, что имеются в документации на вашу версию 1OS.

Списки префиксов могут именоваться любой строкой, содержащей буквы и цифры, а также могут снабжаться специальным описанием. Описание списка префиксов выглядит примерно так:

```
ip prefix-list list-name description text
```

Для того чтобы добавить или удалить записи из списка префиксов, понадобится следующий синтаксис:

```
ip prefix-list list-name seq seq-value deny/permit network/len [ge ge-value] [le le-value]
```

Здесь *list-name* — строка буквенно-цифровых символов. Для обозначения списков префиксов вы можете использовать стандартную схему с последовательной нумерацией или задать более понятное вам название:

```
ip prefix-list list-name <до 80 символов, пробелы разрешены>
```

В табл. 11.3 представлены примеры фильтрации префиксов.

Таблица 11.3. *Фильтрация заданных префиксов*

Критерий фильтрации	Выполняемая команда
Разрешить все кроме префикса 192.68.0.0/16	ip prefix-list sample permit 192.68.0.0/16
Запретить маршрут по умолчанию	ip prefix-list sample deny 0.0.0.0/0
Разрешить все	ip prefix-list sample permit 0.0.0.0/0 le 32
Запретить все	ip prefix-list sample deny 0.0,0.0/0 le 32
Запретить /25+ во всем адресном пространстве	ip prefix-list sample deny 0.0.0.0/0 ge 25
В префиксе 192.68.0.0/24 запретить /25+	ip prefix-list sample deny 192.68.0.0/24 ge 25
Разрешить все адреса с префиксами от /8 до /24	ip prefix-list sample permit 0.0.0.0/0 ge 8 le 24

Обратите внимание, что в табл. 11.3 не включены номера последовательностей. Если вы не указываете номер последовательности, то начальное значение последовательности по умолчанию — 5 и будет увеличивается на 5 в каждой последующей записи, если явно не указан другой номер последовательности.

Для пошаговой настройки списка префиксов вам просто нужно указать номер последовательности для записи, которую вы собираетесь добавить, удалить или изменить в списке. В качестве примера рассмотрим список префиксов в листинге 11.10.

Листинг 11.10. *Пример списка префиксов*

```
ip prefix-list   sample  seq 5   permit   1.1.1.0/24
ip prefix-list   sample  seq 10  permit   2.2.2.0/24
ip prefix-list   sample  seq 15  permit   3.3.3.0/24
ip prefix-list   sample seq   20  deny     0.0.0.0/0   le   32
```

Пример списка префиксов, приведенный в листинге 11.10, разрешает префиксы 1.1,1.0/24, 2.2.2.0/24 и 3.3.3.0/24 и запрещает все остальные маршруты. Предположим, что необходимо разрешить также маршрут 4.4.4.0/24. Соответствующая настройка записи будет выполняться таким образом:

```
ip prefix-list sample seq 18 permit 4.4.4.0/24
```

При внесении этой записи в список префиксов последний приобретет вид, представленный в листинге 11.11.

Листинг 11.11. *Список префиксов после добавления записи, разрешающей работу с маршрутом 4.4.4.0/24*

```
ip prefix-list sample seq 5  permit 1.1.1.0/24
ip prefix-list sample seq 10 permit 2.2.2.0/24
ip prefix-list sample seq 15 permit 3.3.3.0/24
ip prefix-list sample seq 18 permit 4.4.4.0/24
ip prefix-list sample seq 20 deny   0.0.0.0/0  le 32
```

Теперь допустим, что необходимо запретить маршрут 4.4.4.0/24. Вам потребуется внести следующую запись в список префиксов:

```
no ip prefix-list sample seq 18
```

Теперь наш список префиксов будет выглядеть так, как показано в листинге 11.12.

Листинг 11.12. *Список префиксов после добавления записи, запрещающей работу с маршрутом 4.4.4.0/24*

```
ip prefix-list sample seq 5  permit 1.1.1.0/24
ip prefix-list sample seq 10 permit 2.2.2.0/24
ip prefix-list sample seq 15 permit 3.3.3.0/24
ip prefix-list sample seq 20 deny   0.0.0.0/0  le 32
```

Как видите, пошаговое внесение изменений в списки префиксов предоставляет администратору огромные преимущества по сравнению с традиционными списками разрешения доступа. В последующих листингах главы 11 мы будем в основном применять списки разрешения доступа. Примеры использования списков префиксов вы найдете в следующей главе.

Идентифицирование и фильтрация маршрутов на основе NLRI

Чтобы ограничить маршрутную информацию, получаемую или распространяемую вашим маршрутизатором, вы можете фильтровать ее на основе сообщений об обновлении маршрутов, отправляемых на соседний взаимодействующий узел или поступающих от него. Фильтр представляет собой совокупность списков префиксов (или списков разрешения доступа), которые применяются к сообщениям об обновлении маршрутов, поступающим или

отправляемым взаимодействующему узлу. На рис. 11.2, маршрутизатор RTD в AS2 объявляет маршрут в сеть 192.68.10.0/24 и посылает его маршрутизатору RTF. Затем маршрутизатор RTF передает сведения о маршруте маршрутизатору RTA по IBGP, который, в свою очередь, будет распространять сведения о нем в AS1. Таким образом, объявляя сеть 192.68.10.0/24, AS3 может стать транзитной.

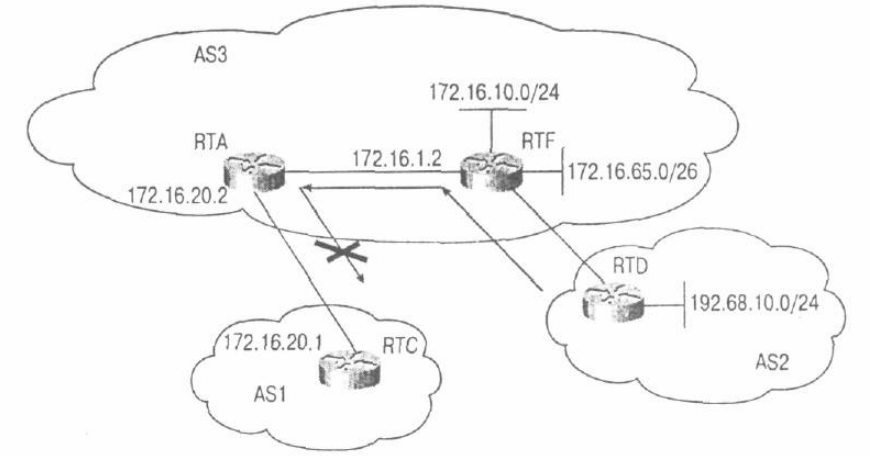


Рис. 11.2. *Идентифицирование и фильтрация префиксов*

Чтобы избежать подобной ситуации, на маршрутизаторе RTA можно настроить фильтр, который бы не допускал распространения префикса 192.68.10.0/24 в AS1. В листинге 11.13 показан пример организации такого фильтра на маршрутизаторе RTA.

Листинг 11.13. *Фильтр префикса на маршрутизаторе RTA*

```
router bgp 3
 no synchronization
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 prefix-list 1 out
 no auto-summary
!
ip prefix-list 1 seq 5 deny 192.68.10.0/24
ip prefix-list 1 seq 10 permit 0.0.0.0/0 le 32
```

В листинге 11.13 комбинация команд настройки маршрутизатора neighbor prefix-list и списка префиксов 1 не дает возможности маршрутизатору RTA распространять сведения о префиксе 192.68.10.0/24 в AS1. В части настроек маршрутизатора, где описывается список префиксов, когда последний используется совместно с BGP-командой neighbor, задается фильтрация исходящих обновлений маршрутов в направлении указанного соседнего узла (обратите внимание на ключевое слово out). Заметьте, что список префиксов 1 заканчивается логическим выражением, которое разрешает распространение всех обновлений маршрутов **(permit 0.0.0.0/0 1e** 32). Когда с целью фильтрации используются списки префиксов или списки разрешения доступа, если в конце не определено какое-либо действие, то по умолчанию будет применяться логическое выражение "запретить все остальное". Это означает, что маршруты, не соответствующие ранее заданным условиям, будут отвергаться. Вот почему очень важно определить действия по умолчанию. В нашем примере маршрут 192.68.10.0/24 будет запрещен, а все остальные маршруты разрешены к распространению. Для этой цели задание в явном виде выражения "запретить все остальное" (если вам действительно это необходимо) поможет избавиться от головной боли.

Хотя из примера этого и не видно, наиболее распространенная на сегодня практика заключается в подготовке списка маршрутов, которые будут разрешены к распространению, и последующей настройке в явном виде запрещенных маршрутов. Этим гарантируется, что маршрутная информация от взаимодействующего узла приниматься не будет, если она не прошла через фильтр. Конечно, это может оказаться не совсем удобным, если от взаимодействующего узла вы получаете большое количество маршрутной информации.

Карты маршрутов (которые ссылаются на списки разрешения доступа, списки

префиксов или на другие правила маршрутизации) могут применяться для фильтрации обновлений маршрутов в предыдущем примере. Для того чтобы познакомить вас с различными параметрами, применяемыми при фильтрации, был выбран метод с использованием прямого списка префиксов.

Применение списков разрешения доступа для фильтрации маршрутов к суперсетям или диапазонов маршрутов — довольно сложное дело, поэтому обычно прибегают к различным хитростям. Предположим, что, например, к марш руги затору RTF на рис. 11.2 подключены различные подсети из диапазона 172.16.X.X, а вам необходимо объявить объединенный маршрут только в форме 172.16.0.0/16. Тогда стандартный список разрешения доступа вида: access-list I permit 172.16.0.0 0.0.255.255

работать не будет, так как он разрешает больше, чем вам необходимо. В стандартном списке разрешения доступа анализируется только IP-адрес источника и не проверяется длина сетевой маски. Таким образом, приведенный список разрешения доступа допускает объявление маршрутов 172.16.0.0/16, 172.16.0.0/17, 172.16.0.0/18, 172.16.0.0/24 и т.д.

Чтобы ограничить распространение маршрутов только одним 172.16.0.0/16, необходимо воспользоваться расширенным списком разрешения доступа: access-list *access-list-number* {deny | permit} protocol source source-wildcard destination destination-wildcard I mask mas.fc-ivild.card

Этим выражением вы определяете расширенный список разрешения доступа, в котором сопоставление проводится по паре адресов источника и пункта назначения или **по** адресу источника и маске. На основе этих данных принимается решение о разрешении или запрещении распространения определенного обновления маршрутов. Номер списка разрешения доступа назначается из диапазона между 100 и 199. Если проверка осуществляется по протоколу IP и сопоставление проводится по паре адрес источника — маска, то строка конфигурации может быть задана в виде:

```
access-list access-list-number permit ip network-number network-do-not-care-bits mask mask-do-not-care-bits
```

Например:

```
access-list 101 permit ip 172.16.0.0 0.0.255.255 255.255.0.0 0.0.0.0
```

Здесь 0 — это точно совпавший бит, а 1 — незначащий бит.

В приведенном выше расширенном списке разрешения доступа показано, что объединенный маршрут 172.16.0.0/16 будет распространяться лишь потому, что его маска в точности совпадает с 255.255.0.0. А обновление маршрутов, содержащее 172.16.0.0/17. фильтр уже не пропустит.

Вы можете достичь этого и другими путями. Например:

```
access-list 101 permit ip host 172.16.0.0 host 255.255.0.0
```

Или использовать список префиксов в качестве фильтра:

```
ip prefix-list 1 seq 5 permit 172.16.0.0/16
```

Помните: по умолчанию предполагается, что все, не описанное в правилах, запрещено, если вы не задали обратное.

Идентифицирование и фильтрация маршрутов на основе атрибута AS_PATH

Фильтрация маршрутов на основе информации, которая содержится в атрибуте AS_PATH, довольно удобна, когда требуется фильтровать все маршруты в одной или между

несколькими AS. Этот метод является достаточно эффективной альтернативой созданию списков из сотен маршрутов, например, в случае фильтрации на основе списков префиксов. Вы также можете указать список префиксов для входящих и исходящих обновлений маршрутов на основе значений атрибута AS_PATH.

Давайте снова вернемся к рис. 11.2. Если требуется не допустить чтобы AS3 стала транзитной, то нужно сконфигурировать ее граничные маршрут заторы RTA и RTF так, чтобы они объявляли только локальные сети. Имеются в виду локальные сети, сгенерированные в самой AS. В листинге 11.14 приведена конфигурация маршрутизатора RTA, где выполняются эти действия (подобным образом нужно будет настроить и маршрутизатор RTF).

Листинг 11.14. *Как не допустить, чтобы AS3 стала транзитной (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no synchronization neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 1 out
 no auto-summary
ip as-path access-list 1 permit ^$
```

В листинге 11.14 список разрешения доступа 1, работающий на основе анализа AS_PATH, идентифицирует только те обновления, которые были сгенерированы в AS3. Субкоманда filter-list при фильтрации обновлений маршрутов работает совместно со списком разрешения доступа на основе анализа AS_PATH. В данном примере эта субкоманда применяется к исходящим обновлениям маршрутов (обратите внимание на ключевое слово out). Нормальное выражение вида ^$ говорит о том, что атрибут AS_PATH пустой, т.е. не содержит каких-либо данных. Знак вставки (^) указывает на начало атрибута AS_PATH, а символ $ обозначает конец атрибута AS_PATH. Так как все сети, сгенерированные в AS3, будут иметь пустые атрибуты AS_PATH (вспомните, что локальный номер AS не включается в атрибут, пока маршрут не передан другому EBGP-узлу), то они будут объявляться дальше. Все остальные префиксы будут отвергаться.

Если вы хотите убедиться в корректной работе вашего нормального выражения, воспользуйтесь для этого командой EXEC:

show ip bgp regexp *regular-expression*

По этой команде маршрутизатор выдает все маршруты, которые соответствуют заданному нормальному выражению. Значения символов в нормальных выражениях описаны в табл. 6.4 в главе 6.

Примечание

Карты маршрутов (в частности те, что ссылаются на списки разрешения доступа на основе атрибута AS_PATH) могли также применяться и для фильтрации обновлений маршрутов в предыдущем примере. Однако для того, чтобы показать различные способы организации фильтров, было предложено использовать субкоманду filter-list.

Группы взаимодействующих узлов

В протоколе BGP *группой взаимодействующих узлов (peer group)* называют группу соседних BGP-узлов, которые совместно используют один набор обновлений правил маршрутизации. Обновление правил маршрутизации обычно проводится на основе карт маршрутов, списков распределения маршрутов, списков префиксов и списков фильтров. Так, вместо того, чтобы задавать одни и те же правила для каждого узла, вы задаете имя группе узлов и указываете для нее правила маршрутизации.

Применение групп взаимодействующих узлов показано на рис. 11.3.

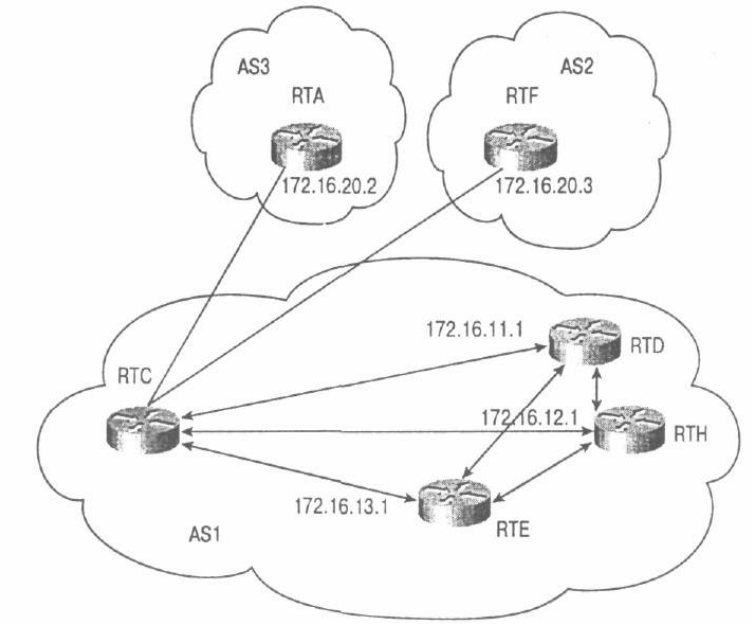


Рис. 11.3. *Группы взаимодействующих узлов в BGP*

Здесь маршрутизатор RTC формирует одинаковые сеансы с маршрутизаторами RTD, RTE и RTH. Вместо того чтобы формулировать и задавать подобные правила отдельно каждому узлу, на маршрутизаторе RTC создается группа взаимодействующих узлов, в которой действуют единые правила маршрутизации для всех узлов, являющихся членами этой группы. В листинге 11.15 приведена конфигурация маршрутизатора для работы в группе.

Листинг 11.15. *Настройка группы взаимодействующих BGP-узлов*

```
router bgp 1
 neighbor INTERNALMAP peer-group
 neighbor INTERNALMAP remote-as 1
 neighbor INTERNALMAP route-map INTERNAL out
 neighbor INTERNALMAP filter-list 1 out
 neighbor INTERNALMAP filter-list 2 in
 neighbor  172.16.11.1 peer-group  INTERNALMAP
 neighbor  172.16.13.1 peer-group  INTERNALMAP
 neighbor  172.16.12.1 peer-group  INTERNALMAP
neighbor  172.16.12.1 filter-list  3  in
```

Конфигурация, приведенная в листинге 11,15, описывает группу узлов с именем INTERNALMAP, в которой содержатся следующие правила маршрутизации. (Отметим, что имя INTERNALMAP мы выбрали произвольно, так что здесь вы совершенно вольны в выборе. Как правило, более практично давать группам имена описательного характера.)

- Используется карта маршрутов с именем INTERNALMAP.
- Задан список фильтров для исходящих обновлений маршрутов (filter list 1).
- Задан список фильтров для входящих обновлений маршрутов (filter list 2).

Эта конфигурация объединяет в группу все соседние узлы внутри AS — маршрутизаторы RTD, RTE и RTH.

Члены группы наследуют все параметры, заданные группе узлов. Члены группы могут быть сконфигурированы таким образом, что правила группы будут "перекрываться", если вносимые изменения не повлияют на исходящие обновления маршрутов. Другими словами, члены группы взаимодействующих узлов могут быть сконфигурированы так, что они будут действовать вопреки заданным в группе правилам, а это может повлиять на входные правила маршрутизации. Так, например, при настройке маршрутизатора RTC задается также список фильтров с номером 3 для обновлений, поступающих от соседнего узла с IP-адресом 172.16.12.1 (RTH). Список фильтров 3 будет "перекрывать" действие входных правил маршрутизации, которые установлены для группы INTERNALMAP в

отношении маршрутизатора RTH.

В листинге 11.16 показано, как настроить группу взаимодействующих BGP-узлов с именем EXTERNALMAP на маршрутизаторе RTC с участием внешних взаимодействующих узлов из AS3 и AS2.

Листинг 11.16. *Группа взаимодействующих BGP-узлов*

```
router bgp 1
neighbor EXTERNALMAP peer-group
neighbor EXTERNALMAP route-map  SETMED  out
neighbor EXTERNALMAP filter-list  1  out
neighbor EXTERNALMAP filter-list 2   in
neighbor 172.16.20.2 remote-as   3
neighbor 172.16.20.2 peer-group EXTERNALMAP
neighbor 172.16.20.3 remote-as  2
neighbor 172.16.20.3 peer-group EXTERNALMAP
neighbor 172.16.20.3 filter-list  3  in
ip as-path access-list 1 permit  ^$3
```

В конфигурации, представленной в листинге 11.16, команды **neighbor remote-as** задаются после команд конфигурации маршрутизатора **neighbor peer-group,** так как требуется описать несколько различных внешних AS. Обратите внимание, что в этой конфигурации настраивается список фильтров 3, который можно использовать для подавления на узле с IP-адресом 172.16.20.3 (RTF) параметров конфигурации, заданных для входящих обновлений маршрутов.

Два последующих абзаца относятся только к старым версиям IOS. В новых версиях эти ограничения отсутствуют.

Обратите внимание, что внешние BGP-узлы — маршрутизаторы RTA и RTF, которые также входят в группу узлов EXTERNALMAP, принадлежат к одной подсети 172.16.20.0. Это ограничение вводится, чтобы не допустить потерь информации. Поместив внешние взаимодействующие узлы в различные подсети, вы можете спровоцировать маршрутизатор RTC посылать обновления маршрутов на взаимодействующие узлы (RTA и RTF) по неопределенному IP-адресу. При нормальном функционировании EBGP, маршруты при отсутствии непосредственного соединения со следующим ближайшим узлом будут игнорироваться (вспомните, чтобы не допустить этого мы использовали команду **ebgp-multihop).** Таким образом, будут потеряны и обновления маршрутов.

Еще одно ограничение заключается в том, что группы взаимодействующих узлов не следует составлять из EBGP-узлов, если маршрутизатор между этими узлами действует как транзитный. Если маршрутизатор (RTC) пересылает обновления маршрутов с одного внешнего узла на другой, то помещение внешних узлов в группу может привести к тому, что маршруты будут ошибочно удаляться. Отметим, что список фильтров 1 был задан для того, чтобы разрешить распространение локальных маршрутов AS1 только в направлении соседних узлов RTA и RTF. Таким образом, маршрутизатор RTC не сможет работать как чисто транзитный между RTA и RTF.

Заранее заданные группы взаимодействующих узлов могут значительно облегчить конфигурирование новых узлов (пока имеет место снижение вероятности ошибок). Кроме того, они также могут значительно сократить время настройки протокола BOP на отдельных маршрутизаторах. Сообщения UPDATE для членов группы BGP-узлов генерируются только один раз и затем просто копируются на все узлы группы, что позволяет значительно сократить потребление ресурсов процессора, которые были бы задействованы более интенсивно, если бы каждое сообщение UPDATE генерировалось отдельно.

Источники обновления маршрутов

Как вы уже знаете, маршруты в протокол BGP могут поступать динамически или

статически. Выбор способа поступления маршрутов в BGP зависит от количества и стабильности маршрутов.

Динамическое вложение информации в BGP

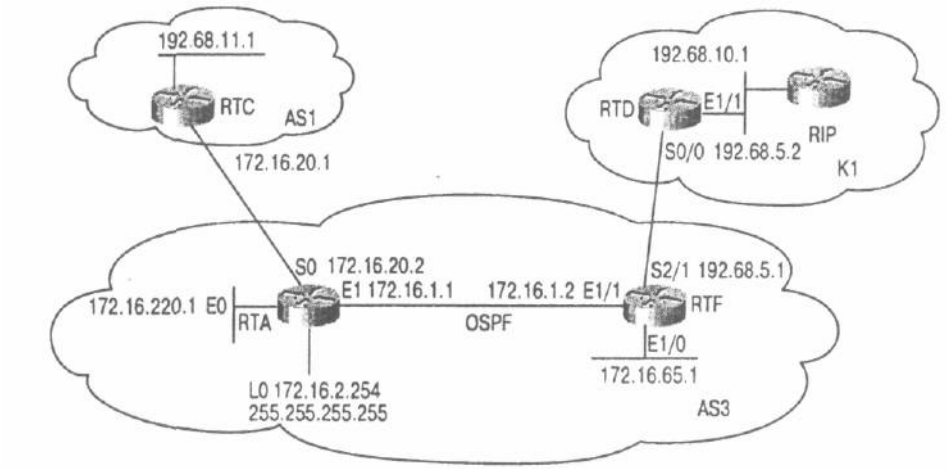


Рис. 11.4. *Вложение маршрутов в протокол BGP*

В приведенном ниже примере демонстрируется динамическое вложение маршрутной информации в протокол BGP. Рассмотрим рис. 11.4

Предположим, что AS3 подключена к сети Internet через AS1. Внутри AS3 в качестве протокола IGP используется OSPF, а с ASI она взаимодействует по EBGP. С другой стороны, у AS3 есть клиент К1, со следующими характеристиками:

- К1 указывает маршрут по умолчанию в направлении AS3;
- К1 объявляет все свои маршруты в AS3 с использованием протокола RIP.

На маршрутизаторе RTF организовано два процесса маршрутизации —OSPF и RIP. Кроме того, маршрутизатор RTF будет только ожидать появления RIP-маршрутов через соединение с К1 и преобразовывать полученные маршруты в OSPF-маршруты. С другой стороны, на маршрутизаторе RTA также запущено два процесса — OSPF и BOP. Маршрутизатор RTA будет динамически преобразовывать все свои локальные маршруты и маршруты, полученные от своего клиента К1, в BGP-маршруты. Все это видно из листинга 11.17.

Листинг 11.17. *Преобразование маршрутов в маршрутизаторе RTF*

```
interface Ethernetl/0
 ip address 172.16.65.1 255.255.255.192

interface Etnernetl/1
 ip address 172.16.1.2 255.255.255.0

interface Serial2/l
 ip address 192.68.5.1 255.255.255.0

router ospf 10
 redistribute rip subnets
 network 172.16.0.0 0.0.255.255 area 0

router  rip
 passive-interface   Serial2/l
 network   192.68.5.0
```

В конфигурации маршрутизатора RTF вы встретите две новых команды.

- **passive-interface type number** — запрещает пересылку обновлений маршрутов на указанный интерфейс. В приведенном примере при использовании совместно с RIP эта команда не допускает пересылку обновлений RIP-маршрутов на

интерфейс S2/1. Это необходимо в том случае, когда несколько клиентов, подключенных к маршрутизатору RTF, не должны "видеть" сети друг друга.
При работе по протоколу OSPF эта команда запрещает передачу пакетов hello на определенный интерфейс, что, в свою очередь, не допускает обмена информацией о состоянии соединения через этот интерфейс.

- **redistribute protocol** ***[process-id]*** осуществляет преобразование (подстановку) маршрутов из одного процесса маршрутизации в другой. В рассматриваемом случае в маршрутизаторе RTF RIP-маршруты преобразуются в OSPF-маршруты (процесс 10). Для команды redistribute существует огромное количество различных расширений (например, subnets). Все эти расширения будут рассмотрены позднее.

С помощью ключевого слова subnets можно удостовериться, вся ли информация о подсетях была в точности передана процессу OSPF. Это требуется только при вложении маршрутов в протокол OSPF. Подобный случай приведен в листинге 11.18.

Листинг 11.18. *Статическое преобразование маршрутов в RIP на маршрутизаторе RTD*

```
interface Ethernetl/1
 ip address 192.68.10.1 255.255.255.0

interface Serial0/0
 ip address 192.68.5.2 255.255.255.0

router rip
 redistribute static
 network 192.68.5.0
 network 192.68.10.0
 default-metric 1

ip route 0.0.0.0 0.0.0.0 192.68.5.1
```

Обратите внимание, что маршрутизатор RTD сконфигурирован со статическим маршрутом по умолчанию в направлении маршрутизатора RTF. Для всех пунктов назначения вне сети клиента К1 маршрутизатор RTD будет направлять трафик на RTF. Кроме того, RTD будет также ретранслировать информацию о статическом маршруте по умолчанию во внутренний домен RIP, так что все остальные маршруты могут по умолчанию указывать на AS3. Команда маршрутизатора default-metric задает метрику преобразуемых маршрутов. В этом случае метрика по умолчанию устанавливается равной *]*, так как имеется всего один промежуточный узел для Маршрута 0/0 и этот маршрут преобразуется в RIP-маршрут. Этот случай описан в листинге 11.19.

Листинг 11.19. *Преобразование OSPF-маршрутов на маршрутизаторе RTA*

```
interface Ethernet0
 ip address 172.16.220.1 255.255.255.0

interface Ethernetl
 ip address 172.16.1.1 255.255.255.0

interface Serial0
 ip address 172.16.20.2 255.255.255.0

router ospf 10
 passive-interface Serial 0
 network 172.16.0.0 0.0.255.255 area 0

router  bgp  3
 redistribute ospf 10 match external 1 external 2
 neighbor 172.16.20.1 remote-as 1
 no auto-summary
```

На маршрутизаторе RTA имеется комбинация OSPF-маршрутов, которые

принадлежат к AS3 и к другим внешним маршрутам, поступающим из RIP-домена К1. При применении команды маршрутизатора redistribute маршрутизатор RTA будет динамически подставлять все эти маршруты в свой процесс BGP. Отметим, что в маршрутизаторе RTA с этой командой используются также ключевые слова match external I external 2, поскольку протоколом OSPF не предусмотрено вложение внешних маршрутов в BGP, если явно не указано, что он должен это делать. Это было сделано, чтобы избежать образования петель маршрутизации в тех случаях, когда информация о внешних OSPF-маршрутах может быть получена из BGP-маршрутов.

В листинге 11,20 показан фрагмент таблицы IP-маршрутов на маршрутизаторе RTA после выполнения вышеуказанных действий.

Листинг 11.20. *Таблица маршрутов маршрутизатора RTA*

```
RTA#show ip route
Codes:C -- connected, S -- static, I - IGRP, R - RIP
      M -- mobile, B - BGP, D - EIGRP, EX - EIGRP
      external, O -- OSPF, IA - OSPF inter area
      N1 •- OSPF NSSA external type 1, N2 •- OSPF NSSA
      external type 2, El - OSPF external type 1
      E2 -- OSPF external type 2, E - EGP, i - IS-IS
      LI - IS-IS level-1, L2 - IS-IS level-2,
            * - candidate default U - per-user static
            route, o – ODR

Gateway of last resort is not set
O E2 192.68.5.0/24 [110/20] via 172.16.1.2, 2dl3h, Ethernetl
O E2 192.68.10.0/24 [110/20] via 172.16.1.2, 2dl3h, Ethernetl
B    192.68.11,0/24 [20/0] via 172.16.20.1, 2dl3h
     172.16.0.0/16 is variably subnetted, 5 subnets, 3 masks
C      172.16.2.254/32 is directly connected, LoopbackO
C      172.16.220.0/24 is directly connected, EthernetO
C      172.16.20.0/24 is directly connected, SerialO
C      172.16.1.0/24 is directly connected, Ethernetl
0      172.16.65.0/26 [110/20] via 172.16.1.2, 2dl3h, Ethernetl
```

Обратите внимание, что в таблице IP-маршрутов маршруты в сети 192.68.10.0/24 и 192.68.5.0/24 описаны как внешние OSPF-маршруты (O E2). Динамическое преобразование приведет к тому, что маршруты в эти сети будут вложены в BGP. В листинге 11.21 представлен внешний вид таблицы BGP-маршрутов на маршрутизаторе RTC после преобразования.

Листинг 11.21. *Таблица BGP-маршрутов на маршрутизаторе RTC*

```
RTCttshow  ip bgp
BGP table version is 20, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
1 - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network                Next Hop      Metric      LocPrf      Weight   Path
 *>   172.16.1.0/24     172.16.20.2   0           0           3        ?
 *>   172.16.2.254/32   172.16.20.2   0           0           3        ?
 *>   172.16.20.0/24    172.16.20.2   0           0           3        ?
 *>   172.16.65.0/26    172.16.20.2   20          0           3        ?
 *>   172.16.220.0/24   172.16.20.2   0           0           3        ?
 *>   192.68.5.0        172.16.20.2   20          0           3        ?
 *>   192.68.10.0       172.16.20.2   20          0           3        ?
 *>   192.68.11.0       0.0.0.0       0                       32768    i
```

Обратите внимание, что все маршруты к сетям в AS3, где поддерживается OSPF, становятся BGP-маршругами в AS1. В BGP нужно передавать сведения не обо всех сетях, принадлежащих к вашей AS. Дело в том, что довольно часто внутри AS имеются сети с общедоступными IP-адресами или используются IP-адреса, незарегистрированные официально. Нельзя допустить, чтобы эти данные попадали во внешние сети. Посмотрите, каким образом проводится преобразование в BGP адреса петли 172.16.2,254/32. Провайдеры

обычно не разрешают объявление столь длинных префиксов (например, /32) и настаивают, чтобы вы их фильтровали или сами фильтруют их на своих узлах. Это ограничение вводится с целью сдерживания роста глобальных таблиц IP-маршрутов. Нет необходимости и в преобразовании в BOP сведений о сети 172.16.20.0/24, которая является зоной демилитаризации (demilitarized zone — DMZ). Вот почему совместно с преобразованием маршрутов следует использовать фильтрацию. Благодаря этому вы сможете точно определить маршруты, которые следует объявлять.

Маршрутизатор **RTA,** сконфигурированный, как показано в листинге 11.22, позволяет выполнить такую фильтрацию.

Примечание

Начиная с этого момента и далее в целях экономии места мы будем рассматривать в листингах только те команды, которые непосредственно связаны с обсуждаемой темой.

Листинг 11.22. *Фильтрация преобразуемых маршрутов*

```
router ospf 10
 passive-interface Serial0
 network 172.16.0.0 0.0.255.255 area 0

router bgp  3
 redistribute  ospf  10 match external   1  external  2
 neighbor  172.16.20.1  remote-as   1
 neighbor  172.16.20.1  route-map BLOCKROUTES  out
 no  auto-summary

access-list  1  permit  172.16.2.254   0.0.0.0
access-list  1  permit  172.16.20.0   0.0.0.255

route-map BLOCKROUTES  deny  10
 match  ip address  1

route-map BLOCKROUTES permit 20
```

Фильтрация в листинге 11.22 выполняется с помощью карты маршрутов, которая определяет набор действий, предпринимаемых в случае необходимости анализа ситуации на основе определенного критерия. Критериями в нашем случае является выявление соответствия маршрута на хост 172.16.2.254/32 и к сети 172.16.20.0/24, чтобы не допустить их вложения в BGP. В списке разрешения доступа **access-list 1** вы можете найти соответствия этим маршрутам, а команда **route-map BLOCKROUTES** определяет дальнейшие действия (в нашем случае маршруты должны быть отвергнуты). Вторая запись в карте (20) разрешает преобразование всех остальных маршрутов в BGP. (Если нужно более детальное описание этого процесса, обратитесь к главе 6, где обсуждалась фильтрация).

В листинге 11.23 представлено, как будет выглядеть таблица BGP-маршрутов на маршрутизаторе RTC после фильтрации. Как видите, маршруты на хост 172.16.2.254/32 и в сеть 172.16.20.0/24 отсутствуют.

Листинг 11.23. *Таблица BGP-маршрутов на маршрутизаторе RTC после фильтрации*

```
RTC#show ip bgp
BGP table version is 34, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
      i -• internal Origin codes: i - IGP, e - EGP, ? - incomplete
 Network                Next Hop      Metric      LocPrf      Weight   Path
 *>   172.16.1.0/24     172.16.20.2   0           0           3        ?
 *>   172.16.65.0/26    172.16.20.2   20          0           3        ?
 *>   172.16.220.0/24   172.16.20.2   0           0           3        ?
 *>   192.68.5.0        172.16.20.2   20          0           3        ?
 *>   192.68.10.0       172.16.20.2   20          0           3        ?
 *>   192.68.11.0       0.0.0.0       0                       32768    i
```

Чтобы лучше контролировать, что преобразуется из ЮР в BOP вы можете

использовать команду **network.** Команда **network** — еще один способ отдельно задавать префиксы, которые нужно пересылать по BOP. С помощью этой команды указывается префикс, который будет пересылаться дальше (адрес сети и ее маска). Например, выражение network **172.16.1.0 mask** 255.255.255.0 говорит о том, что должен пересылаться префикс 172.16.1.0/24. С сетями, которые приходятся на опорные связки (типа 255.0.0.0, 255.255.0.0 или 255.255.255.0), указывать маску не обязательно. Так, например, чтобы послать сведения о префиксе 172.16.0.0/16 достаточно выражения **network 172.16.0.0.** Такого рода сети выводятся в таблице BGP-маршрутов без префикса /x. Например, сеть класса C с адресом 192.68.11.0 может быть записана как 192.68.11.0/24.

Согласно схеме, представленной на рис. 11.4, в конфигурации маршрутизатора RTA будут указаны адреса сетей, преобразуемые в BOP (листинг 11.24).

Листинг 11.24. *Конфигурация маршрутизатора RTA, где задаются адреса сети, требующие преобразования в BGP*

```
router ospf 10
passive-interface Serial0
 network 172.16.0.0 0.0.255.255 area 0

router bgp  3
 network  172.16.1.0 mask  255.255.255.0
 network   172.16.65.0  mask  255.255.255.192
 network   172.16.220.0  mask  255.255.255.0
 network  192.68.5.0
 network  192.68.10.0
 neighbor  172.16.20.1  reniote-as   1
 no  auto-summary
```

В листинге 11.25 показан внешний вид таблицы BGP-маршрутов на маршрутизаторе RTC после выполнения описанных действий.

Листинг 11.25. *Таблица BGP-маршрутов на маршрутизаторе RTC*

```
RTC#show  ip bgp
BGP  table  version   is   34,   local   routerID   is   192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric      LocPrf      Weight  Path
 *>   172.16.1.0/24     172.16.20.2   0           0           3       i
 *>   172.16.65.0/26    172.16.20.2   20          0           3       i
 *>   172.16.220.0/24   172.16.20.2   0           0           3       i
 *>   192.68.5.0        172.16.20.2   20          0           3       i
 *>   192.68.10.0       172.16.20.2   20          0           3       i
 *>   192.68.11.0       0.0.0.0       0                       32768   i
```

Итак, в BGP преобразуются все маршруты, кроме 172.16.2.254/32 и 172.16.20.0/24. Обратите внимание, что таблица BGP-маршрутов очень похожа на ту, которая была сформирована после преобразования OSPF-маршрутов в BGP-маршруты и применения к ним фильтров. Единственная заметная разница заключается в коде источника i, который ставится в конце записи с информацией о маршруте. Код источника i означает, что источник сведений об этих сетях — внутренний по отношению к AS (в данном случае это протокол IGP). Если вы просмотрите еще раз предыдущий снимок таблицы BGP-маршрутов, сделанный с BGP-маршрутизатора (листинг 11.23), то увидите, что в нем указывается код источника ?, т.е. незавершенный. Это говорит о том, что сведения об этих сетях были получены какими-то другими средствами. Маршруты, поступающие в BGP после преобразования, будут иметь незавершенный код источника, если явно не указано другое.

Команда network действует только в том случае, если заданные в ней префиксы известны маршрутизатору. Другими словами, в BGP невозможно объявлять префиксы вслепую, только на основании того, что они указаны в команде. Прежде чем объявить тот или иной префикс, маршрутизатор сопоставит его с содержимым таблицы IP-маршрутов. Так, в листинге 11.25, если вы задаете команду network 172.16.192.0 mask 255.255.255.0,

маршрут в эту сеть не будет сгенерирован, так как она не известна маршрутизатору.

Статическое вложение информации в BGP

Указание префиксов в команде network имеет те же недостатки, что и динамическое преобразование маршрутов. Если маршрут, заданный командой network становится недоступным, то протокол BGP уведомит об этом остальные узлы, послав сообщение об обновлении маршрута; когда маршрут восстановится, BGP снова формирует и посылает обновление маршрута. Если подобная ситуация многократно повторяется, то нестабильность IGP-маршрута перейдет в нестабильность BGP-маршрута. Единственный способ избежать этого — использовать комбинацию статических префиксов совместно с командой network. Таким образом достигается гарантированное присутствие префиксов в маршрутных IP-таблицах, и, следовательно, они всегда будут объявляться.

В предыдущем примере при необходимости проверить, не приводят ли колебания маршрута 192.68.10.0/24 к колебаниям BGP-маршрутов, вы могли бы задать на маршрутизаторе RTA статический маршрут в виде:

```
ip route 192.68.10.0  255.255.255.0  Ethernetl
```

При использовании статического вложения маршрутов запись с нужным префиксом всегда будет присутствовать в таблице IP-маршрутов и будет объявляться, пока интерфейс Ethernetl находится в активном состоянии. Недостаток этого метода состоит в том, что, даже если маршрут станет недоступным, он все равно будет объявляться по BGP. Полученная таким образом стабильность сети, по сравнению с ущербом, который может нанести один или несколько аномальных маршрутов, является для сетевых администраторов весомым аргументом в пользу этого метода вложения маршрутов. В листинге 11.26 представлена конфигурация маршрутизатора RTA, которая гарантирует, что маршрут в сеть 192.68.10.0/24 всегда будет посылаться другим узлам.

Листинг 11.26. *Конфигурация с постоянно объявляемым маршрутом*

```
router bgp 3
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 network 192.68.5.0
 network 192.68.10.0
 neighbor 172.16.20.1 remote-as 1
 no auto-summary

ip route 192.68.10.0 mask 255.255.255.0     Ethernetl
```

Обратите внимание, что маршрутизатор RTA самостоятельно генерирует префикс 192.68.10.0/24, а не получает его от RTF. Если через статический маршрут объявляется объединенный маршрут, то с целью предотвращения образования петли маршрутизации следует перенаправлять статический маршрут в "битовую корзину" **null 0.**

Наложение протоколов ("черные ходы")

В примере, приведенном ниже, показан вариант использования команды backdoor с целью изменения дистанции EBGP. Это делается для того, чтобы назначить более высокий приоритет протоколам ЮР по сравнению с протоколами EBGP для определенных сетевых адресов. На рис. И.5 приведена схема топологии сети, иллюстрирующая эту ситуацию.

Как видно из рис. 11.5, между AS2 и AS 1 на частном канале поддерживается протокол IGP (OSPF), а между AS2 и AS3 — протокол EBGP. Маршрутизатор в AS1 будет получать объявления о маршруте 192.68.10.0/24 от AS3 по EBGP с дистанцией 20 и от AS2 по OSPF с дистанцией НО. Так как, согласно правилу, предпочтение отдается маршруту с меньшей дистанцией, маршрутизатор RTC, чтобы достичь сети 192.68.10.0/24, воспользуется соединением с AS3 по BGP. Обратите внимание, что EBGP-маршрут имеет дистанцию 20.

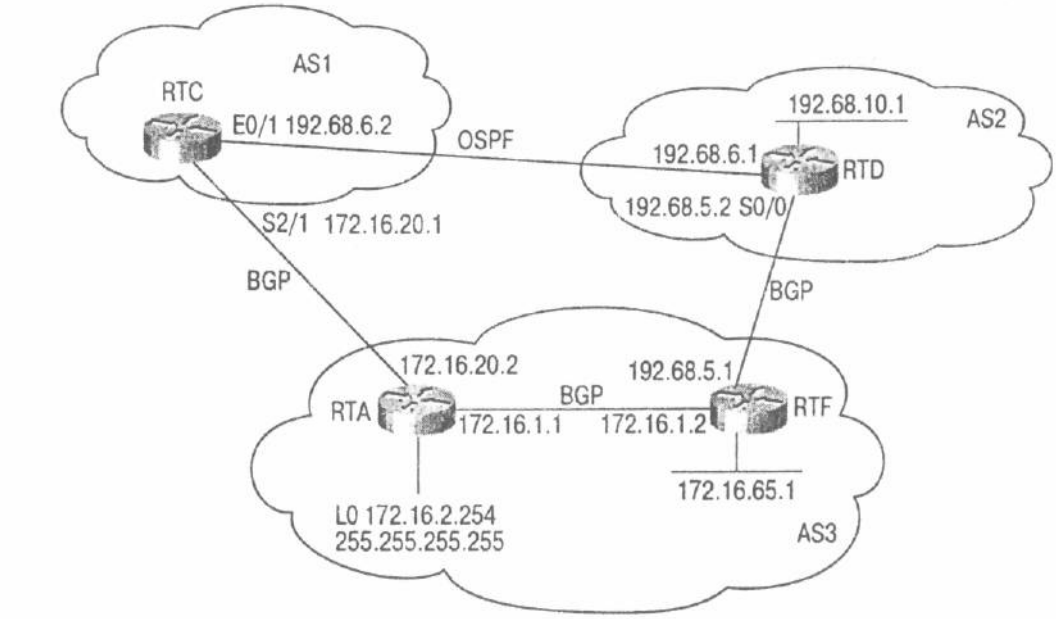


Рис. 11.5. *Наложение маршрутов в BGP*

Листинг 11.27. *Таблица IP-маршрутов на маршрутизаторе RTC*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
       * - candidate default U - per-user static route
Gateway of last resort is not set
C   192.68.6.8/24 is directly connected, EthernetO/1
B    192.68.10.0/24 [20/0] via 172.16.20.2, 00:21:36
     172.16.8.0/16 is variably subnetted, 3 subnets, 2 masks
C    172.16.20.0/24 is directly connected, Serial2/l
B    172.16.1.0/24 [20/0] via 172.16.20.2, 00:21:37
B    172.16.65.0/26 [20/20] via 172.16.20.2, 00:21:37
```

Если вы хотите, чтобы RTC предпочел OSPF-маршрут, то нужно настроить маршрутизатор так, как это показано в листинге 11.28.

Листинг 11.28. *Конфигурация маршрутизатора RTC для предпочтения OSPF-маршрута*

```
router bgp  I
 neighbor 172.16.20.2 remote-as 3
 network 192.68.10.0 backdoor
 no auto-summary
```

В листинге 11.28 с помощью команды **network 192.68.10.0** backdoor вы изменяете дистанцию BGP-маршрута 192.68.10.0/24 с 20 на 200, что делает OSPF-маршрут более предпочтительным, так как он имеет дистанцию 110. Заметьте, команда network **192.68.10.0 backdoor** не дает BGP сгенерировать объявление этой сети.

В листинге 11.29 приведены результаты выполнения описанных настроек в таблице RTC-маршрутов. Как видите, сведения о маршруте 192.68.10.0/24 теперь получены по протоколу OSPF с дистанцией **[110],** т.е. в этом случае будет задействован частный канал между AS1 и AS2.

Листинг 11.29. *Новая таблица маршрутов маршрутизатора RTC*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
       * - candidate default U - per-user static route
Gateway of last resort is not set
C    192.68.6.0/24 is directly connected, EthernetO/1
O IA 192.68.10.0/24 [110/20] via 192.68.6.1, 00:00:21, EthernetO/1
     172.16.0.0/16 is variably subnetted, 3 subnets, 2 masks
C   172.16.20.0/24 is directly connected, Serial2/l
B       172.16.1.0/24 [20/0] via 172.16.20.2, 00:29:07
B       172.16.65.0/26 [20/20] via 172.16.20.2, 00:29:07
```

Атрибуты BGP

В этом разделе мы будем работать с топологией сети, представленной на рис. 11.6, которая поможет нам выяснить, как работают различные атрибуты протокола BGP.

В листингах 11.30—11.33 представлены базовые конфигурации маршрутизаторов RTA, RTF, RTC и RTD. Дополнительные настройки мы будем вносить по мере обсуждения отдельных атрибутов.

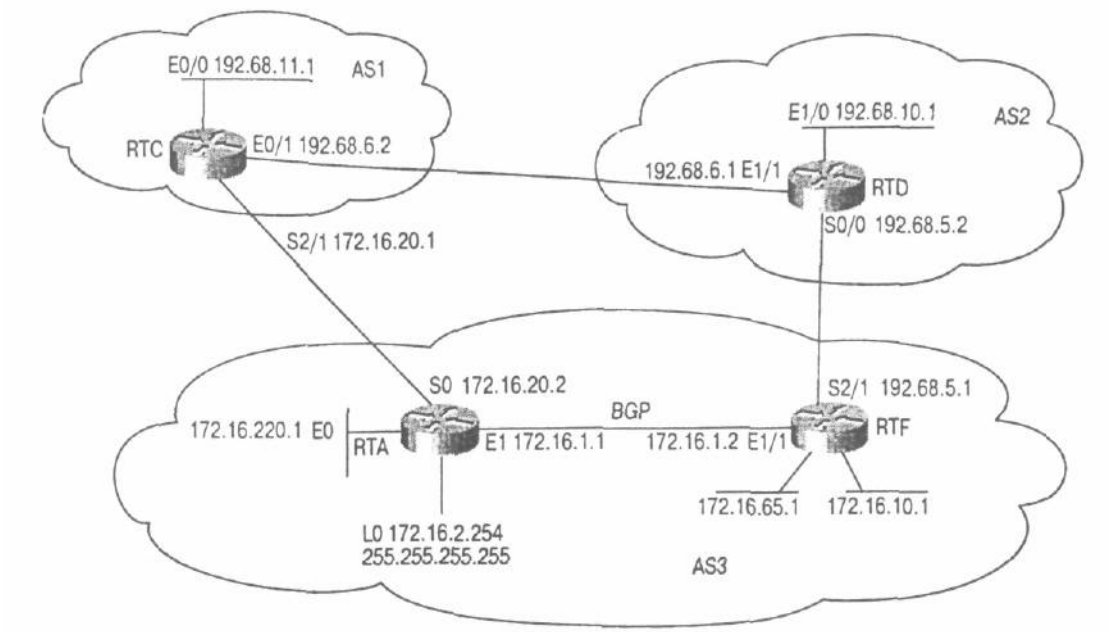


Рис.11. 6. *Применение атрибутов BGP*

Листинг 11.30. *Базовая конфигурация маршрутизатора RTA (рис. 11.6)*

```
ip subnet-zero

interface LoopbackO
 ip address 172.16.2.254 255.255.255.255

interface EthernetO
 ip address 172.16.220.1 255.255.255.0

interface Ethernetl
 ip address 172.16.1.1 255.255.255.0

interface SerialO
```

```
 ip address 172.16.20.2 255.255.255.0

router ospf 10
 passive-interface SerialO
 network 172.16.0.0 0.0.255.255 area 0

router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16,220.0 mask 255.255.255.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source LoopbackO
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out no auto-summary

ip classless
ip as-path access-list 10 permit ^$
```

Листинг 11.31. *Базовая конфигурация маршрутизатора RTF (рис. 11.6)*

```
ip subnet-zero
interface EthernetO/0
 ip address 172.16.10.1 255.255.255.0

interface Ethernet 1/0
 ip address 172.16.65.1 255.255.255.192

interface Ethernetl/1
 ip address 172.16.1.2 255.255.255.0

interface Serial2/l
 ip address 192.68.5.1 255.255.255.0

router ospf 10
 network 172.16.0.0 0.0.255.255 area 0

router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.2.254 remote-as 3
 neighbor 192.68.5.2 remote-as 2
 neighbor 192.68.5.2 filter-list 10 out
 no auto-summary

ip classless
ip as-path access-list 10 permit ^$
```

Листинг 11.32. *Базовая конфигурация маршрутизатора RTC (рис. 11.6)*

```
ip subnet-zero
interface EthernetO/0
 ip address 192.68.11.1 255.255.255.0

interface EthernetO/1
 ip address 192.68.6.2 255.255.255.0

interface Serial2/l
 ip address 172.16.20.1 255.255.255.0

router bgp 1
 network 192.68.11.0
 neighbor 172.16.20.2 remote-as 3
 neighbor 192.68.6.1 remote-as 2
```

```
 no auto-summary

ip classless
```

Листинг 11.33. *Базовая конфигурация маршрутизатора RTD (рис. 11.6)*

```
ip subnet-zero
interface Ethernetl/0
 ip address 192.68.10.1 255.255.255.0

interface Ethernetl/1
 ip address 192.68.6.1 255.255.255.0

interface SerialO/0
 ip address 192.68.5.2 255.255.255.0

router bgp 2
 network 192.68.10.0
 neighbor 192.68.5.1 remote-as 3
 neighbor 192.68.6.2 remote-as 1
 no auto-suranary

ip classless
```

Предположим также, что AS3 не является транзитной. Именно поэтому задается команда filter-list 10, которая заставляет генерировать AS3 только локальные маршруты. Маршруты, полученные от AS1 и AS2, не будут распространяться вне AS3. Отметим также, что некоторые сети, такие как 172.16.10.0/24, объявляются с помощью команды **network** как на маршрутизаторе RTA, так и на маршрутизаторе RTF. Так можно проверить, будут ли объявляться эти сети при отсутствии связи между AS3 и AS1 или AS3 и AS2.

Атрибут NEXT_HOP

Используя последующие примеры, мы обсудим функционирование атрибута BGP NEXT_HOP и покажем методы управления его значениями.

В листинге 11.34 показана BGP-таблица маршрутизатора RTF.

Листинг 11.34. *BGP-таблица маршрутизатора RTF*

```
RTC#show  ip bgp
BGP  table  version   is   34,   local   routerID   is   192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric      LocPrf     Weight  Path
 *>   i172.16.1.0/24    172.16.2.254  0           100        0       i
 *>                     0.0.0.0       0                      32768   i
 *>   i172.16.10.0/24   172.16.2.254  20          100        0       i
 *>                     0.0.0.0       0                      32768   i
 *>   i172.16.65.0/24   172.16.2.254  20          100        0       i
 *>                     0.0.0.0       0                      32768   i
 *>   i172.16.220.0/24  172.16.2.254  20          100        0       i
 *>                     172.16.1.1    20                     32768   i
 *>   192.68.10.0       192.68.5.2    0                      0       2  i
 *>   192.68.11.0       192.68.5.2                           0       2 1 i
 *>  i                  172.16.20.1   0                      0       1 i
```

Сведения о сети 192.68.11.0/24 получены по IBGP (обратите внимание на символы i справа) от узла NEXT_HOP 172.16.20.1, а это IP-адрес внешнего соседнего узла по отношению к RTA. Как правило. IP-адрес EBGP-узла хранится в домене маршрутизации, вот почему очень важно иметь внутренний маршрут к соседнему ближайшему узлу (NEXT_HOP). В другом случае BGP-маршрут окажется просто бесполезным. Существует несколько способов удостовериться в том, что вы нормально получите доступ по EBGP к

ближайшему соседнему узлу. Первый из них заключается в том, что ближайший следующий узел поддерживает работу по 1GP. Это видно из таблицы маршрутов на маршрутизаторе RTA, куда включен интерфейс SerialO, поддерживающий OSPF; так, маршрутизатор RTF сможет получить сведения о 172.16.20.1. Даже если OSPF поддерживается на RTA через интерфейс SerialO, маршрутизатору нет необходимости постоянно обмениваться пакетами hello через этот интерфейс, так как задана команда passive-interface.

Второй способ заключается в применении команды next-hop-self neighbor (см. листинг 11.31), которая заставит маршрутизатор объявлять маршрут к самому себе в качестве ближайшего соседнего узла. В конфигурации маршрутизатора RTF в листинге 11.31 обратите внимание, как в конец выражения neighbor в направлении RTA добавляется команда next-hop-self. Таким образом, когда маршрутизатор RTF объявляет внешние сети, такие как 192.68.10.0/24 в направлении RTA, он в качестве следующего ближайшего узла будет подставлять себя. Посмотрите на таблицу BGP-маршрутов маршрутизатора RTA, представленную в листинге 11.35. Из нее видно, что префикс 192.68.10.0/24 получен от следующего ближайшего узла 172.16.1.2, который вместе с маршрутизатором RTF является внутренним узлом. Так как узел 172.16.1,2 уже является частью маршрута OSPF, то нет никаких проблем чтобы попасть на него.

Листинг 11.35. *Таблица BGP-маршрутов маршрутизатора RTA*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop       Metric     LocPrf     Weight  Path
 *>   i172.16.1.0/24  172.16.1.20    0          100        0       i
 *>                   0.0.0.0        0                     32768   i
 *>   i172.16.10.0/24 172.16.1.20    0          100        0       i
 *>                   172.16.1.20    20                    32768   i
 *>   i172.16.65.0/24 172.16.1.20    0          100        0       i
 *>                   172.16.1.20    20                    32768   i
 *>   i172.16.220.0/24 172.16.1.20   20         100        0       i
 *>                   0.0.0.0        0                     32768   i
 *>   i192.68.10.0    172.16.1.2     0          100        0       2  i
 *>                   172.16.20.1                          0       1 2 i
 *> 192.68.11.0       172.16.20.100  0                     0       1 i
```

Внимательно просмотрите листинг 11.35 и вы увидите, что на самом деле сведения о маршруте 192.68.10.0/24 поступают двумя различными путями, в то время как маршрут в сеть 192.68.11.0/24 распространяется одним путем. Это может показаться немного запутанным, но в действительности маршрутизация происходит именно так, как необходимо. В этой ситуации маршрутизатор RTF выбирает наилучший маршрут в сеть 192.68.11.0/24 через маршрутизатор RTA (см. листинг 11.34). Именно поэтому маршрутизатор RTF не объявляет сеть 192.68.11.0/24 маршрутизатору RTA, а на RTA хранится отдельная запись о 192.68.11.0/24.

Атрибут AS_PATH

Рассматривая таблицу BGP-маршрутов на маршрутизаторе RTF, представленную в листинге 11.36, обратите внимание на информацию о маршруте через AS (AS_PATH) в конце каждой строки. Так, сведения о сети 192.68.П.0/24 получены с использованием IBGP с AS_PATH I, а по EBGP — с AS_PATH 2 1. Это означает, что при необходимости отправить данные посредством IBGP в сеть 192.68.11.0/24 маршрутизатор RTF может сделать это через AS1, а при применении EBGP — через AS2 и AS1. Однако, как мы знаем, в протоколе BGP предпочтение отдается кратчайшему маршруту, поэтому будет использоваться IBGP-маршрут с AS_PATH 1. Символ ">" в левой части строки означает, что из двух возможных маршрутов в сеть 192.68.11.0/24 протокол BGP в качестве "наилучшего" выбирает второй

маршрут.

Листинг 11.36. *Таблица BGP-маршрутов на маршрутизаторе RTI*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *>   i172.16.1.0/24  172.16.2.254  0           100        0       i
 *>                   0.0.0.0       0                      32768   i
 *>   i172.16.10.0/24 172.16.2.254  20          100        0       i
 *>                   0.0.0.0 0                            32768   i
 *>   i172.16.65.0/24 172.16.2.254  20          100        0       i
 *>                   0.0.0.0 0                            32768   i
 *>   i172.16.220.0/24 172.16.2.254 0           100        0       i
 *>                   172.16.1.1    20                     32768   i
 *>   192.68.10.0     192.68.5.2    0                      0       2  i
 *  192.68.11.0       192.68.5.2                           0       2 1 i
 *>i                  172.16.20.1   0           100        0       1 i
```

Управление атрибутом AS_PATH

Давайте рассмотрим подробно BGP-таблицу на маршрутизаторе RTF, приведенную в листинге 11.36. Как видите, маршрутизатор RTF выбирает кратчайший маршрут к сети 192.68.11.0/24 через AS1. В листинге 11.37 показан набор команд, с помощью которых вы можете управлять содержимым атрибута AS_PATH. Длину маршрута можно увеличить, добавив дополнительные номера AS в маршрут AS_PATH. Вернемся к сети, показанной на рис. 11.6. С помощью операций, указанных в листинге 11.37 вы добавляете два дополнительных номера AS в атрибут AS_PATH, который посылается от маршрутизатора RTC на RTA, и таким образом изменяете решение об использовании маршрута в сеть 192.68.11.0/24, принимаемое на маршрутизаторе RTF.

Листинг 11.37. *Управление атрибутом AS_PATH путем добавления дополнительных номеров AS*

```
router bgp 1
 network 192.68.11.0
 neighbor 172.16.20.2 remote-as 3
 neighbor 172.16.20.2 route-map AddASnumbers out
 neighbor 192.68.6.1 remote-as 2 no auto-summary

route-map AddASnumbers permit 10
 set as-path prepend 1 1
```

В примере конфигурации маршрутизатора, представленном в листинге 11.37, показано, как проводится вставка двух дополнительных номеров AS 1 и 1 в атрибут AS_PATH, посылаемый от маршрутизатора RTC на RTA. Теперь давайте посмотрим, как это отразилось на BGP-таблице маршругизатора RTF. Из листинга 11.38 видно, что маршрутизатор RTF теперь может обмениваться данными с сетью 192.68.11.0/24 через ближайший узел 192.68.5.2, т.е. по маршруту 2 1. Маршрутизатор RTF будет использовать этот маршрут, так как он короче, чем прямой маршрут через AS1, который теперь имеет длину (1 1 1).

Листинг 11.38. *BGP-таблица на маршрутизаторе RTF после внесения изменений в атрибут AS_PATH*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *>   i172.16.1.0/24  172.16.2.254  0           100        0       i
```

```
*>                    0.0.0.0          0                  32768   i
*>   i172.16.10.0/24  172.16.2.254     20        100      0       i
*>                    0.0.0.0 0        0                  32768   i
*>   i172.16.65.0/24  172.16.2.254     20        100      0       i
*>                    0.0.0.0 0        0                  32768   i
*>   i172.16.220.0/24 172.16.2.254     0         100      0       i
*>                    172.16.1.1       20                 32768   i
*>   192.68.10.0      192.68.5.2       0                  0       2  i
*>   192.68.11.0      192.68.5.2                          0       2 1 i
*>i                   172.16.20.1      0         100      0       1 1 1 i
```

Использование частных номеров AS

Приведенный пример иллюстрирует, как сконфигурировать BGP на маршрутизаторе, чтобы не допустить попадания информации о частных номерах AS в сеть Internet. На рис. 11.7 представлена топология сети, на которую мы будем опираться в этом разделе.

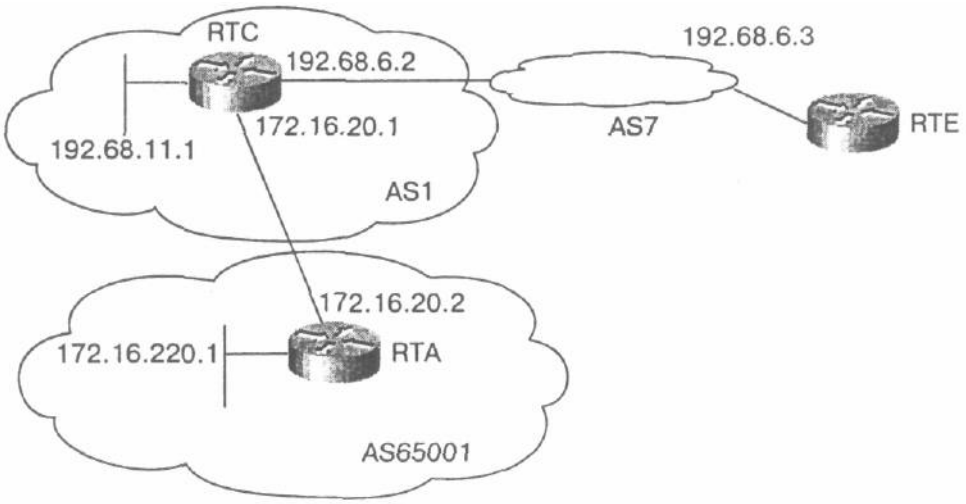


Рис. 11.8. *Подавление частных номеров AS*

Настроив маршрутизаторы RTA и RTC так, как указано в листингах 11.39 и 11.40, вы не допустите распространения частного номера AS 65001 в сеть Internet через AS1 при распространении BGP-маршрутов.

Листинг 11.39. *Конфигурация маршрутизатора RTA, не допускающая распространения частных номеров AS*

```
router bgp 65001
 network 172.16.220.0 mask 255.255.255,0
 neighbor 172.16.20.1 remote-as 1
 no auto-summary
```

Листинг 11.40. *Конфигурация маршрутизатора RTC, не допускающая распространения частных номеров AS 1*

```
router bgp 1
network 192.68.11.0 mask 255.255.255.0
 neighbor 172.16.20.2 remote-as 65001
 neighbor 192.68.6.3 remote-as 7
 neighbor  192.68.6.3  remove-private-AS
 no auto-summary
```

Обратите внимание, как в листинге 11.40 используется ключевое слово **remove-**private-AS при настройке соединения с соседним узлом в AS7. В листинге 11.41 представлены таблицы BGP на маршрутизаторах RTC и RTE.

Листинг 11.41. *BGP-таблицы на маршрутизаторах RTC и RTE*

```
RTC# show ip bgp
```

```
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *>   172.16.220.0/24   172.18.20.2   0                      0       65001 i
 *>   192.68.11.0       0.0.0.0       0                      32768   i

RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *>   172.16.220.0/24   192.68.6.2                           0       0 1 i
 *>   192.68.11.0       192.68.6.2    0                      0       0 1 i
```

Как видите, префикс 172.16.220.0/24 в BGP-таблице на маршрутизаторе RTC имеет AS_PATH 65001, а в BGP-таблице на маршрутизатое RTE AS_PATH —равный 1. Таким образом, на маршрутизаторе RTC при передаче маршрутной информации в AS7 извлекаются частные номера AS из атрибута AS_PATH. Обратите внимание, что команда **remove-private-AS** применяется к исходящим маршрутам, т.е. она используется только в точках выхода из сети.

Атрибут LOCAL_PREF

Установка локальных предпочтений (с помощью атрибута LOCAL_PREF) также влияет на процесс принятия решения об использовании того или иного маршрута в протоколе BGP. Если к одному и тому же префиксу существует насколько маршрутов, то предпочтение отдается тому из них, который имеет наибольшее значение локального предпочтения. Атрибут LOCAL_PREF действителен для всех узлов внутри одной AS и относится к высшему уровню процесса принятия решения в BGP (он следует сразу за параметром weight, который используется только в оборудовании компании Cisco и является локальным для маршрутизатора). Этот атрибут рассматривается перед атрибутом AS_PATH. Так, длинному маршруту с большим значением AS_PATH и LOCAL_PREF будет отдаваться предпочтение перед коротким маршрутом, но имеющим меньшее значение LOCAL_PREF. В листинге 11.42 (помните, что мы при рассмотрении конфигурации опираемся на рис. 11.7) представлена конфигурация маршрутизатора RTF, где устанавливаются более высокие локальные предпочтения для всех обновлений BGP-маршрутов, поступающих от маршрутизатора RTD.

Листинг 11.42. *Конфигурация маршрутизатора RTF, устанавливающая более высокие локальные предпочтения для обновлений BGP-маршрутов, поступающих от RTD*

```
router bgp 3
no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.2.254 remote-as 3
 neighbor 172.16.2.254 next-hop-self
 neighbor 192.68.5.2 remote-as 2
 neighbor 192.68.5.2 filter-list 10 out
 neighbor 192.68.5.2 route-map SET/LOCAL in
 no auto-summary

ip as-path access-list 10 permit ^$

route-map SETLOCAL permit 10
set local-preference 300
```

Командой route-map SETLOCAL всем маршрутам, поступающим от маршрутизатора RTD (обратите внимание на ключевое слово in), присваивается локальное предпочтение 300 (локальное предпочтение по умолчанию — 100). Посмотрите, как в листинге 11.43 протоколом BGP выбираются маршруты к префиксам 192.68.10.0/24 и 192.68.11.0/24 через ближайший узел 192.68.5.2 с локальным предпочтением 300.

Листинг 11.43. *BGP-таблица на маршрутизаторе RTF*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric     LocPrf     Weight  Path
 *>   172.16.1.0/24     0.0.0.0       0                     32768   i
 *> i                   172.16.2.254  0          100        0       i
 *>   172.16.10.0/24    0.0.0.0       0                     32768   i
 *> i                   172.16.2.254  20         100        0       i
 *>   172.16.65.0/24    0.0.0.0       0                     32768   i
 *> i                   172.16.2.254  20         100        0       i
 *>   i172.16.220.0/24  172.16.1.1    20                    32768   i
 *> i                   172.16.2.254  0          100        0       i
 *>   192.68.10.0       192.68.5.2    0          300        0       2  i
 *>   192.68.11.0       192.68.5.2                300        0       2 1 i
```

Так как атрибут LOCAL_PREF передается внутри AS, маршрутизатор RTF передает локальные предпочтения маршрутизатору RTA, как показано в листинге 11.44 (BGP-таблица маршрутизатора RTA).

Листинг 11.44. *BGP-таблица на маршрутизаторе RTA с указанием локальных предпочтений*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric     LocPrf     Weight  Path
 *>   i172.16.1.0/24    172.16.1.2    0          100        0       i
 *>                     0.0.0.0       0                     32768   i
 *>   i172.16.10.0/24   172.16.1.2    0          100        0       i
 *>                     172.16.1.2    20                    32768   i
 *>   i172.16.65.0/24   172.16.1.2    0          100        0       i
 *>                     172.16.1.2    20                    32768   i
 *>   i172.16.220.0/24  172.16.1.2    20         100        0       i
 *>                     0.0.0.0       0                     32768   i
 *>   i192.68.10.0      172.16.1.2    0          300        0       2 i
 *>                     172.16.20.1                         0       1 2 i
 *>   i192.68.11.0      192.68.5.2                300        0       2 1 I
 *>                     172.16.20.1                         0       1 i
```

Как видите, даже несмотря на то, что длина AS_PATH при работе по EBGP меньше, при работе по IBGP будет выбираться префикс 192.68.11.0/24 с локальным предпочтением 300. Другие префиксы, получаемые по протоколу IBGP, такие как 172.16.10.0/24, имеют локальное предпочтение по умолчанию равное 100.

Атрибут MULTU_EXIT_DISC

В этом разделе мы рассмотрим, как использовать метрики одной AS, чтобы повлиять на процесс принятия решения о выборе маршрута в другой AS. Это осуществляется с помощью атрибута MULTI_EXIT_DISC (сокращенно MED). На рис. 11.8 AS3 является клиентом провайдера AS1. Для того чтобы влиять на входящий трафик, нужно в AS3

генерировать набор метрик и отсылать его в AS1, Если все BGP-атрибуты одинаковы, в протоколе BGP будет отдаваться предпочтение тем маршрутам, которые имеют меньшую метрику по сравнению с другими.

На маршрутизаторах RTA и RTF поддерживается для внутреннего обмена протокол IBGP, а для работы с провайдером AS1 — протокол EBGP. Маршрутизатор RTG является внутренним и поддерживает только протокол OSPF. Предположим, что маршрутизаторы RTA и RTF должны в направлении AS1 посылать MED для достижения следующих целей.

- Пересылать входящий трафик для сети 172.16.1.0/24 по каналу SF.
- Входящий трафик для остальных сетей пересылать с применением граничного маршрутизатора, который обладает информацией о маршрутах в эти сети с меньшими внутренними метриками. Например, входящий трафик в направлении сети 172.16.112.0/24 должен пересылаться по каналу SF, если у RTA есть маршрут с меньшей метрикой к этой сети, чем у RTF.

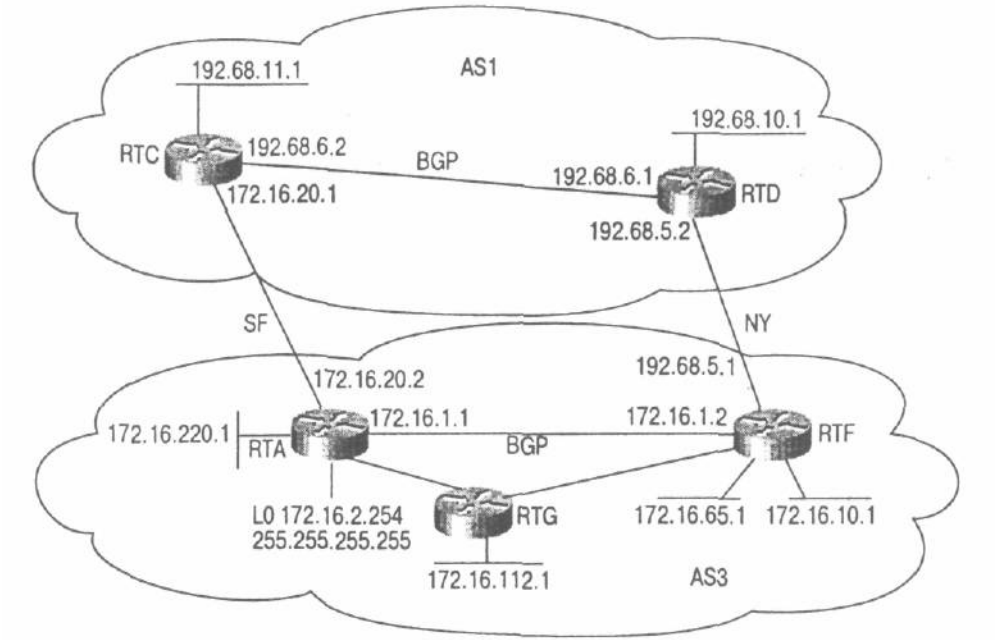


Рис. 11.8. *Установка атрибута MED*

В листингах 11.45 и 11.46 показаны конфигурации маршрутизаторов RTA и RTF, удовлетворяющие указанным критериям.

Листинг 11.45. *Установка атрибута MED - конфигурация маршрутизатора RTA*

```
router ospf 10
 passive-interface Serial0
 network 172.16.0.0 0.0.255.255 area 0

router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 network 172.16.112.0 mask 255.255.255.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 no auto-summary

ip as-path access-list 10 permit ^$
```

Листинг 11.46. *Установка атрибута MED — конфигурация маршрутизатора RTF*

```
router ospf 10
 network 172.16.0.0 0.0.255.255 area 0

router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 network 172.16.112.0 mask 255.255.255.0
 neighbor 172.16.2.254 remote-as 3
 neighbor 172.16.2.254 next-hop-self
 neighbor 192.68.5.2 remote-as 1
 neighbor 192.68.5.2 route-map SETMETRIC out
 neighbor 192.68.5.2 filter-list 10 out
 no auto-summary

ip as-path access-list 10 permit ^$
access-list 1 permit 172.16.1.0 0.0.0.255

route-map SETMETRIC
 permit 10 match ip address 1
 set metric 50

 route-map SETMETRIC permit 20
```

Маршрутизатор RTF, сконфигурированный, как показано в листингах 11.45 и 11.46, будет генерировать префикс 172.16.1.0/24 с MED равным 50. Получив сведения об этом префиксе, ASI сравнивает метрику 50 с метрикой 0, поступающей от маршрутизатора RTA, и, естественно, будет задействовать канал SF. Все остальные сети будут объявляться в BGP с внутренними метриками, и AS1 будет выбирать ту точку входа, которая обладает наименьшей метрикой. В листинге 11.47 представлена BGP-таблица на маршрутизаторе RTD после внесения описанных изменений.

Листинг 11.47. *BGP-таблица на маршрутизаторе RTD*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric     LocPrf     Weight  Path
 *>   172.16.1.0/24   192.68.5.1    50                    0       3 i
 *> i                 192.68.6.2    0          100        0       3 i
 *>   172.16.10.0/24  192.68.5.1    0                     0       3 i
 *>   172.16.65.0/26  192.68.5.1    0                     0       3 i
 *   172.16.112.0/24  192.68.5.1    84                    0       3 i
 *> i                 192.68.6.2    74         100        0       3 i
 *>   172.16.220.0/24 192.68.5.1    20                    0       3 i
 *>                   192.68.6.2    0          100        0       3 i
 *>   192.68.10.0     0.0.0.0       0                     32768   i
 *> i192.68.11.0      192.68.6.2    0          100        0       i
```

Посмотрите как в таблице BGP-маршрутов, приведенной в листинге 11.47, выбирается маршрут в сеть 172.16.1.0/24 через узел 192.68.6.2, т.е. через маршрутизатор RTC (он использует самого себя как следующий узел с помощью команды next-hop-self). Этот маршрут выбирается благодаря тому, что имеет меньшую метрику (0 меньше 50). Для всех остальных сетей маршрутизатор RTD также будет использовать маршруты с наименьшими метриками. Обратите внимание, что сведения о сети 172.16.112.0/24 получены от маршрутизатора RTA по маршруту с метрикой 74, а от маршрутизатора RTF — с метрикой 84. Таким образом, чтобы доставить данные в сеть 172.16.112.0/24, маршрутизатор RTD будет передавать их по каналу SF.

Для маршрутов, сведения о которых были получены посредством BGP, имеется возможность пересылки в другую AS с внутренними метриками IGP, перенесенными при преобразовании маршрутов в BGP-маршруты. Это выполняется с помощью команды set metric-type internal, которая описывает часть карты маршрутов в направлении соседнего узла. Таким образом, BGP-маршруты будут переносить в атрибуте MED информацию о внутренних метриках протокола IGP.

Атрибут COMMUNITY

Этот раздел посвящен атрибуту COMMUNITY, с помощью которого можно динамически влиять на процесс принятия решения о выборе маршрута в другой AS. Для обеспечения работы сети, представленной на рис. 11.9, в листинге 11.48 показан пример конфигурации маршрутизатора RTA, в котором AS3 объявляет маршрут в сеть 172.16.65.0/26 автономной системе AS1 и динамически инструктирует ее о запрете на объявление этого маршрута другим AS. Для этого при объявлении маршрута 172.16.65.0/26 в AS1 в системе AS3 ему присваивается атрибут COMMUNITY со значением **no-export.**

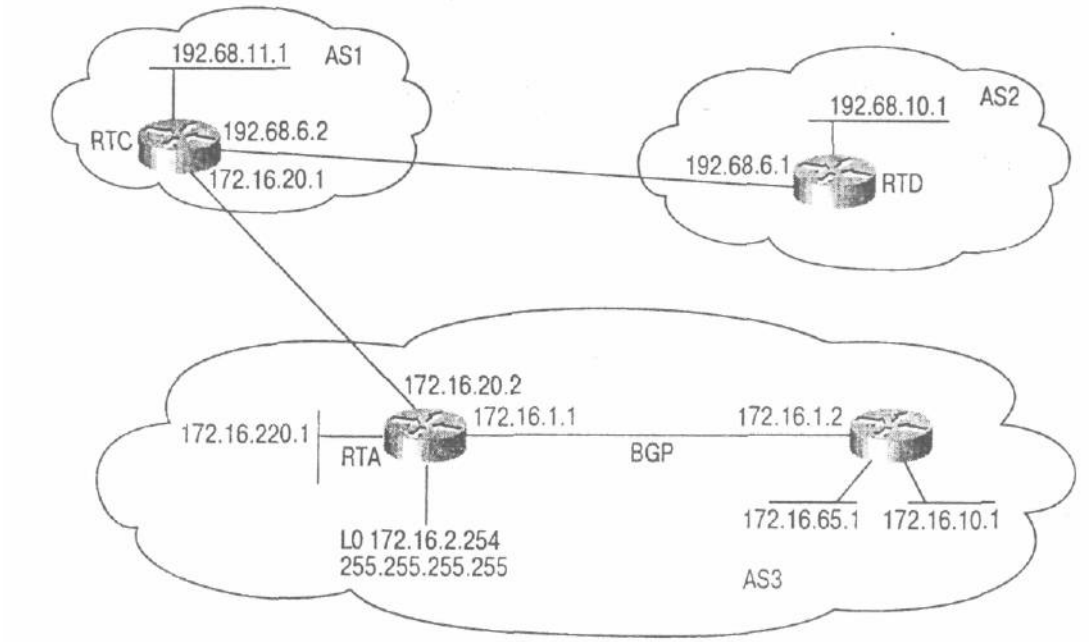


Рис. 11.9. *Установка атрибута COMMUNITY*

Листинг 11.48. *Конфигурация маршрутизатора RTA с использованием атрибута COMMUNITY*

```
router bgp 3
no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 send-community
 neighbor 172.16.20.1 route-map SETCOMMUNITY out
 no auto-summary

access-list  1 permit 172.16.65.0   0.0.0.63

route-map  SETCOMMUNITY permit 10
match  ip  address   1
set community no-export
route-map  SETCOMMUNITY permit  20
```

Так как по умолчанию сведения о сообществах не распространяются между внутренними и внешними соседними BGP-системами, чтобы принудительно послать этот атрибут на другой узел, необходимо в команде neighbor задать параметр send-community.

В примере конфигурации, представленном в листинге 11.48, описана карта маршрутов SETCOMMUNITY в направлении соседнего узла 172.16.20.1 (RTC). Запись в карте маршрутов с номером 10 будет обрабатывать префикс 172.16.65.0/26 и устанавливать атрибут COMMUNITY в no-export. Ключевое слово send-community необходимо задавать при сеансе связи с соседним узлом, что позволяет пересылать указанному узлу атрибут COMMUNITY. В записи карты маршрутов под номером 20 разрешается передавать сведения обо всех остальных сетях без изменений.

В листинге 11.49 показана BGP-запись на маршрутизаторе RTC для сети 172.16.65.0/26.

Листинг 11.49. *Вывод сведений о BGP-маршруте в сеть 172.16.65.0/26*

```
RTC#show ip bgp 172.16.65.0 255.255.255.192
BGP routing table entry for 172.16.65.0/26, version 3
Paths: (1 available, best #1, not advertised to EBGP peer)
 3
   172.16.20.2 from 172.16.20.2 (172.16.2.254)
   Origin IGP, metric 20, valid, external, best
   Community: no-export
```

Как видите, в атрибуте COMMUNITY было задано значение no-export, что не позволяет передавать данное обновление маршрутов на другие EBGP-узлы. Таким образом, маршрутизатор RTC не будет объявлять этот маршрут своему внешнему соседу — маршрутизатору RTC. Следует отметить, что в таблице BGP-маршрутов на маршрутизаторе RTD, представленной в листинге 11.50, нет сведений о маршруте в сеть 172.16.65.0/26.

Листинг 11.50. *Таблица BGP-маршрутов на маршрутизаторе RTD*

```
RTC# show ip bgp
BGP table version is 34, local routerID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network          Next Hop      Metric      LocPrf     Weight  Path
 *>   172.16.1.0/24     192.68.6.2                       0       1 3 i
 *>   172.16.10.0/24    192.68.6.2                       0       1 3 i
 *>   172.16.220.0/24   192.68.6.2                       0       1 3 i
 *>   192.68.10.0       0.0.0.0       0                  32768   i
 *>   192.68.11.0       192.68.6.2    0           100    0       1 i
```

Агрегация в BGP-4

В приведенных ниже примерах описываются различные методы агрегации (объединения) маршрутов, которые вы можете встретить в сети Internet. Способ формирования и объявления объединенных маршрутов, а также то, распространяются ли они самостоятельно или совместно с однозначно определенными маршрутами, будет влиять на потоки трафика в сети и на размер маршрутных BGP-таблиц. Помните, что объединяются только те маршруты, сведения о которых есть в маршрутной BGP-таблице. Объединенный маршрут может посылаться другим узлам, если в таблице BGP-маршрутов присутствует по крайней мере один однозначно определенный маршрут.

Только объединенные маршруты, подавление однозначно определенных маршрутов

В этом разделе мы покажем, как сгенерировать объединенный маршрут без распространения однозначно определенных маршрутов, которые входят в его состав. В сети, представленной на рис. 11.10, маршрутизатор RTA посылает информацию о префиксах 172.16.220.0/24, 172.16.1.0/24 и 172.16.65.0/26 на маршрутизатор RTC.

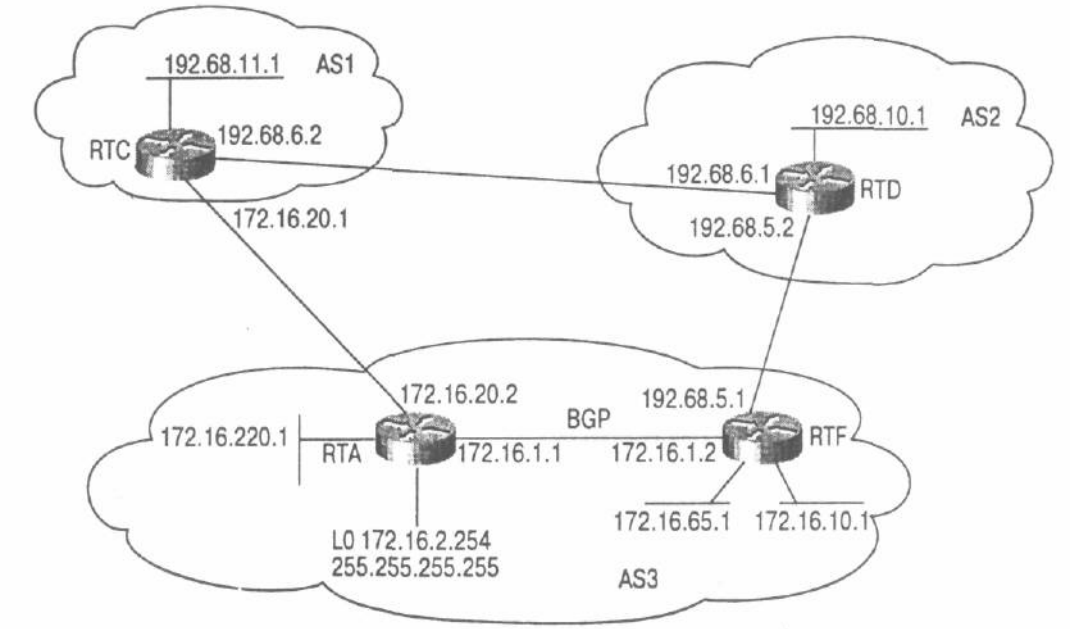


Рис. 11.10. *Объединение маршрутов в BGP (с подавлением однозначно определенных маршрутов).*

В примере конфигурации, приведенном в листинге 11.51, показано каким образ< на маршрутизаторе RTA проводится объединение всех этих маршрутов в один отде^ ный префикс 172.16.0.0/16, который и посылается на маршрутизатор RTC. Таким с разом, мы полагаем, что AS3 является единоличным владельцем сети класса 172.16.0.0/16. На маршрутизаторе RTF выполняется такое же объединение, и его ко фигурация подобна конфигурации маршрутизатора RTA.

Листинг 11.51. *Конфигурация маршрутизатора RTA для объединения маршрутов в оди! .префикс с целью их пересылки на маршрутизатор RTC*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 aggregate-address 172.16.0.0 255.255.0.0 summary-only
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1   filter-list  10  out
 no  auto-summary

ip  as-path access-list  10  permit  A$
```

При настройке маршрутизатора RTA для объединения однозначно определенных маршрутов в один 172.16.0.0/16 используется команда aggregate-address.

В таблице BGP-маршрутов на маршрутизаторе RTA (листинг 11.52) показано, как генерируется новый объединенный маршрут 172.16.0.0/16. Как видите, этот маршрут сгенерирован на маршрутизаторе RTA (NEXT_HOP 0.0.0.0), а все остальные однозначно

определенные префиксы подавляются (обратите внимание на символ s (т.е. suppressed — подавляемые) в начале строки). В этом случае такой же результат может быть достигнут при выполнении автосуммирования.

Листинг 11.52. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 14, local routerID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network          Next Hop      Metric      LocPrf     Weight  Path
 *>   172.16.0.6  0.0.0.0                               32768   i
 * i              172.16.1.2                 100        0       i
 s>   172.16.1.0/24   0.0.0.0       0                   32768   i
 s>   172.16.10.0/24  172.16.1.2    20                  32768   i
 s>   172.16.65.0/26  172.16.1.2    20                  32768   i
 s>   172.16.220.0/24 0.0.0.0       0                   32768   i
 *   192.68.10.0     172.16.20.1                        0       1 2 i
 *> i                172.16.1.2    0           100      0       2 i
 *>   192.68.11.0     172.16.20.1   0                    0       1 i
```

Из таблицы BGP-маршрутов на маршрутизаторе RTC, представленной в листинге 11.53, видно, что от маршрутизатора RTA получены сведения только о префиксе 172.16.0.0/16. Однако вследствие того, что маршрутизатор RTF выполняет то же объединение маршрутов, RTC получит этот же префикс и от него (через маршрутизатор RTD в AS2).

Листинг 11.53. *Таблица BGP-маршрутов на маршрутизаторе RTC*

```
RTC# show ip bgp
BGP table version is 22, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network          Next Hop      Metric      LocPrf     Weight  Path
 *>   172.16.0.0  172.16.20.2                          0       0 3 i
 *                192.68.6.1                           0       0 2 3 i
 *>   192.68.10.0 192.68.6.1    0                      0       0 2 i
 *>   192.68.11.0 0.0.0.0       0                      32768   i
```

Рассмотрим более подробно запись об объединенном маршруте 172.16.0.0/16 (листинг 11.54).

Листинг 11.54. *Вывод сведений о BGP-маршруте 172.16.0.0 (маршрутизатор RTC)*

```
RTC#show ip bgp 172.16.0.0
BGP routing table entry for 172.16.0.0/16, version 22
Paths: (2 available, best #1, advertised over EBGP)
 3, (aggregated by 3 172.16.2.254)
   172.16.20.2 from 172.16.20.2 (172.16.2.254)
     Origin IGP, valid, external, atomic-aggregate, best
 2 3, (aggregated by 3 192.68.5.1)
from 192.68.6.1 (192.68.10.1)
  Origin IGP, valid, external, atomic-aggregate
```

Обратите внимание на присутствие в записи атрибута ATOMIC_AGGREGATE, который указывает на то, что префикс 172.16.0.0/16 является объединенным. Отметим также выражения aggregated by 172.16.2.254 и aggregated by 192.68.5.1, которые описывают атрибут AGGREGATOR. Атрибут AGGREGATOR (мы его обсуждали в главе 6) несет информацию о номере AS и идентификаторе маршрутизатора ROUTER_ID, который сгенерировал маршрут — в нашем случае это AS3 и ROUTER_ID — RTA и RTF.

Объединенные маршруты также могут быть сгенерированы статически, как можно видеть по конфигурации маршрутизатора RTA в листинге 11.55 и маршрутизатора RTF в листинге 11.56.

Листинг 11.55. *Генерирование объединенных маршрутов с использованием статических маршрутов (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no synchronization
 network 172.16.0.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.15.1.2 update-source Loopback0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 no auto-summary

ip route 172.16.0.0 255.255.0.0 null0
 ip as-path access-list 10 permit ^$
```

Листинг 11.56. *Генерирование объединенных маршрутов с использованием статических маршрутов (конфигурация маршрутизатора RTF)*

```
router bgp  3
 no  synchronization
 network   172.16.0.0
 neighbor  172.16.2.254   remote-as  3
 neighbor  172.16.2.254  next-hop-self
 neighbor   192.68.5.2   remote-as  2
 neighbor  192.68.5.2   filter-list  10  out
 no auto-summary

ip  route  172.16.0.0  255.255.0.0  null0
ip  as-path  access-list  10 permit  ^$
```

В листингах 11.55 и 11.56 в таблицу маршрутов вводится статическая запись о статическом маршруте 172.16.0.0/16. Заметьте, что статический маршрут указывает на нулевое устройство null0 (битовую корзину). Если у маршрутизаторов RTA или RTF нет сведений об однозначно определенных маршрутах, которые составляют объединенный маршрут 172.16.0.0, то весь трафик будет игнорироваться. Так предотвращается образование петель маршрутизации, и трафик в RTA и RTF направляется по умолчанию их провайдерам (см. раздел "Бесклассовая междоменная маршрутизация" в главе 3, "IP-адресация и методы распределения адресов").

Объединенные и однозначно определенные маршруты

В некоторых случаях в дополнение к объединенному маршруту желательно передавать в соседние AS однозначно определенные маршруты. Как правило, это делается в AS, подключенных по нескольким каналам к одному провайдеру. Тогда AS провайдера, получая сведения об однозначно определенных маршрутах, сможет принимать более обоснованное решение о выборе маршрута. (Вы уже видели, как AS, принимающая различные метрики маршрутов, может соответствующим образом направлять трафик). На рис. 11.11 представлена топология сети, где реализуется этот вариант объединения маршрутов.

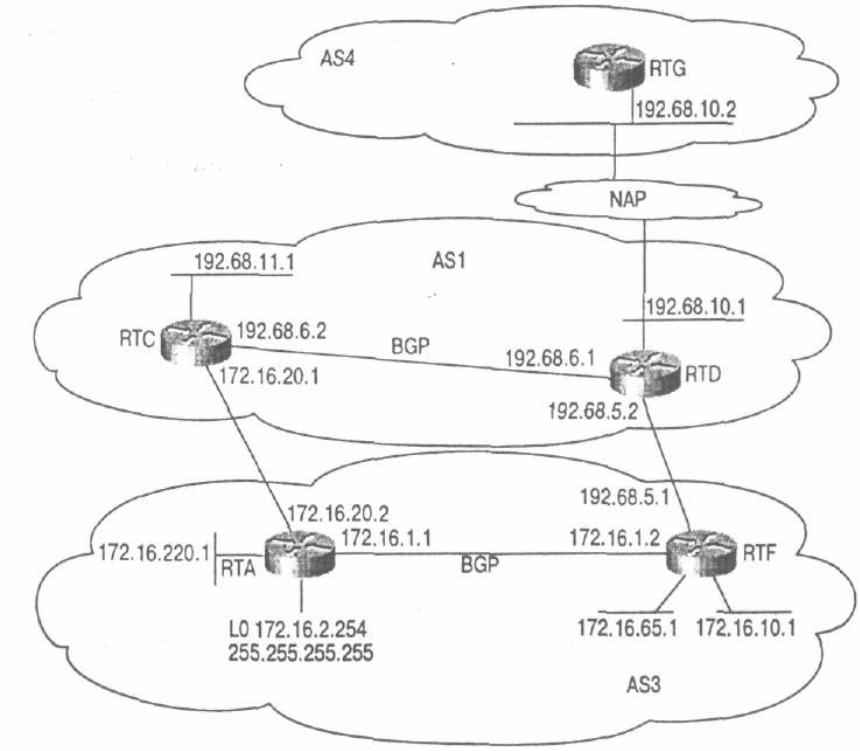


Рис. 11.11. *Еще один вариант объединения в BOP*

На рис. 11.11 показана AS3, которая подключена к одному провайдеру (AS1) по нескольким каналам. Через маршрутизаторы RTA и RTF, принадлежащие AS3, в AS1 посылается объединенный маршрут 172.16.0.0/16 и однозначно определенные маршруты (см. конфигурацию маршрутизаторов в листингах П.57 и 11.58).

Листинг 11.57. *Подключение к одному провайдеру по нескольким каналам (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no synchronization
network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.22B.O mask 255.255.255.0
 aggregate-address 172.16.0.0 255.255.0.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source LoopbackO
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 no auto-summary

ip as-path access-list 10 permit ^$
```

Листинг 11.58. *Подключение к одному провайдеру по нескольким каналам (конфигурация маршрутизатора RTF)*

```
router bgp  3
 no  synchronization
 network  172.16.1.0 mask  255.255.255.0
 network  172.16.10.0  mask  255.255.255.0
 network  172.16.65.0 mask  255.255.255.192
 network  172.16.220.0 mask 255.255.255.0
 aggregate-address   172.16.0.0  255.255.0.0
 neighbor  172.16.2.254  remote-as  3
 neighbor  172.16.2.254  next-hop-self
 neighbor  192.68.5.2   remote-as   1
 neighbor  192.68.5.2  filter-list  10  out
 no  auto-summary

ip as-path access-list 10 permit  ^$
```

Обратите внимание на команду **aggregate-address** в конфигурации маршрутизаторов RTA и RTF. Эта команда задается без параметра **summary-only,** так что и объединенный и однозначно определенные маршруты будут объявляться совместно.

В BGP-таблице на маршрутизаторе RTC (листинг 11.59) видно, что RTC получил сведения не только об объединенном маршруте 172.16.0.0/16, но и однозначно определенные маршруты. Маршрутизатор RTD получает ту же самую информацию.

Листинг 11.59. *Подключение к одному провайдеру по нескольким каналам*

```
RTC# show ip bgp
BGP table version is 28, local router ID is 192.68,11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric      LocPrf     Weight  Path
 * i172.16.0.0          192.8.6.1                 100        0       3 i
 *>                     172.16.20.2                          0       3 i
 * i172.16.1.0/24       192.8.6.1     0           100        0       3 i
 *>                     172.16.20.2   0                      0       3 i
 * i172.16.10.0/24      192.8.6.1     0           100        0       3 i
 *>                     172.16.20.2   20                     0       3 i
 * i172.16.65.0/26      192.8.6.1     0           100        0       3 i
 *>                     172.16.20.2   20                     0       3 i
 * i172.16.220.0/24     192.8.6.1     20          100        0       3 i
 *>                     172.16.20.2   0                      0       3 i
 *>i192.68.10.0         192.8.6.1     0           100                0 i
 *> 192.68.11.0         0.0.0.0       0                      32768   i
```

Используя атрибут COMMUNITY со значением no-export, маршрутизаторы RTA и RTF могут инструктировать RTC и RTD не распространять сведения об однозначно определенных маршрутах и посылать на AS4 только объединенный маршрут 172.16.0.0/16. Эта процедура очень эффективно сдерживает рост таблиц маршрутов, так как AS4 получает только сведения об объединенном маршруте. В листинге 11.60 представлена конфигурация маршрутизатора RTF, где выполняется эта процедура. Маршрутизатор RTA будет сконфигурирован аналогично.

Листинг 11.60. *Подавление экспорта однозначно определенных маршрутов (конфигурация маршрутизатора RTA)*

```
router bgp  3
 no  synchronization
 network  172.16.1.0 mask  255.255.255.0
 network  172.16.10.0  mask  255.255.255.0
 network   172.16.65.0  mask  255.255.255.192
 network   172.16.220.0  mask  255.255.255.0
 aggregate-address   172.16.0.0  255.255.0.0
 neighbor  172.16.2.254   remote-as  3
 neighbor  172.16.2.254  next-hop-self
 neighbor   192.68.5.2   remote-as   1
 neighbor   192.68.5.2   send-community
 neighbor  192.68.5.2  route-map  SETCOMMUNITY out
 neighbor  192.68.5.2   filter-list  10  out
 no  auto-summary

ip  as-path  access-list  10  permit  ^$
access-list  101  permit  ip  172.16.0.0   0.0.255.255  host  255.255.0.0
route-map   SETCOMMUNITY permit   10
 match  ip  address   101
route-map   SETCOMMUNITY permit  20
set  community no-export
```

Из листинга 11.60 видно, что для настройки маршрутизатора RTF используется несколько записей **route-map SETCOMMUNITY,** с помощью которых описываются однозначно определенные маршруты 172.16.1.0/24, 172.16.220.0/24, 172.16.10.0/24 и

172.16.65.0/26 с параметром **community no-export.** Эти записи запрещают маршрутизатору RTD посылать сведения об указанных маршрутах внешним AS (в данном случае AS4). С другой стороны, объединенный маршрут 172.16.0.0/16 передается без изменений и без атрибута COMMUNITY, что позволяет распространять сведения о нем и внешней AS4.

В карте маршрутов запись с номером 10 указывает на использование списка разрешения доступа 101, который будет пропускать только объединенный маршрут 172.16.0.0/16. Посмотрите, как в части списка разрешения доступа с записью о хосте описывается маска подсети 255.255.0.0. Она задана так, чтобы отделять однозначно определенные маршруты, начинающиеся с 172.16, от нужного объединенного маршрута. В записи с номером 10 не устанавливаются значения атрибута COMMUNITY, следовательно, объединенный маршрут будет передаваться именно в том виде, как он был сгенерирован.

Запись 20 запрещает распространение всех однозначно определенных маршрутов, включая параметр **no-export** в атрибуте COMMUNITY.

В листингах 11.61 и 11.62 представлены необходимые конфигурации маршрутизаторов RTC и RTD.

Листинг 11.61. *Конфигурация маршрутизатора RTC*

```
router bgp 1
 no synchronization
 network 192.68.11.0
 neighbor 172.16.20.2 remote-as 3
 neighbor 192.68.6.1 remote-as 1
 neighbor 192.68.6.1 next-hop-self
 neighbor 192.68.6.1 send-community
 no auto-summary
```

Листинг 11.62. *Конфигурация маршрутизатора RTD*

```
router bgp  1
 no  synchronization
 network   192.68.10.0
 neighbor  192.68.5.1   reroote-as  3
 neighbor  192.68.6.2  remote-as  1
 neighbor  192.68.6.2 next-hop-self
 neighbor  192.68.10.2  remote-as  4
 no  auto-summary
```

Обратите внимание на параметр send-community, заданный в команде neighbor, для конфигурирования маршрутизатора RTC в листинге 11.61. Так как маршрутизатор RTA также выполняет такое же объединение маршрутов, то при IBGP-сеансе с RTC маршрутизатор RTD будет получать сведения об однозначно определенных маршрутах. Если маршрутизатор RTC передает их без параметра no-export, то маршрутизатор RTD будет объявлять эти маршруты внешним узлам.

В листинге 11.63 выделены интересующие нас записи в таблице BGP-маршрутов на маршрутизаторе RTD. Первая запись указывает, что префикс 172.16.220.0/24 не будет объявляться соседним EBGP-узлам. Это происходит вследствие того, что маршрутизаторы RTA и RTF помечают этот префикс (и все остальные "однозначные" маршруты) атрибутом COMMUNITY с параметром no-export. Вторая запись показывает, что объединенный маршрут был сгенерирован маршрутизаторами RTA и RTF без изменений и может быть послан в AS4.

Листинг 11.63. *Вывод сведений о BGP-маршрутах на маршрутизаторе RTD*

```
RTD#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 5
Paths; (2 available, best #2, not advertised to EBGP peer)
 3
  192.68.5.1      from 192.68.5.1
    Origin IGP, metric 20, valid, external
    Community: no-export
 3
```

```
  192.68.6.2      from 192.68.6.2 (192.68.11.1)
    Origin IGP, metric 0, localpref 100, valid, internal, best
    Community: no-export

 RTD#show ip bgp 172.16.0.0
BGP routing table entry for 172.16.8.0/16, version 8
Paths: (2 available, best #1, advertised over IBGP, EBGP)
 3, (aggregated by 3 192.68.5.1)
  192.68.5.1      from 192.68.5.1
    Origin IGP, valid, external, atomic-aggregate, best
 3, (aggregated by 3 172.16.2.254)
192.68.6.2  from 192.68.6.2 (192.68.11.1)
Origin  IGP,   localpref  100,   valid,   internal,   atomic-aggregate
```

Советуем вам подробно рассмотреть BGP-таблицу на маршрутизаторе RTG, представленную в листинге 11.64. В ней вы увидите, что от AS3 на AS4 посылаются сведения только об объединенном маршруте 172.16.0.0/16. Все однозначно определенные маршруты в ней отсутствуют.

Листинг 11.64. *Таблица BGP-маршрутов на маршрутизаторе RTG*

```
RTC# show ip bgp
BGP table version is 14, local router ID is 192.68.10.2
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *> 173.16.0.0        192.68.10.1                          0       1 3 i
 *> 192.68.10.0       192.68.10.1   0                      0       1 i
 *> 192.68.11.0       192.68.10.1                          0       1 i
```

Объединение маршрутов с использованием набора однозначно определенных маршрутов

На рис. 11.12 показано, как AS3 может использовать комбинацию из объединенного и однозначно определенных маршрутов для оказания влияния на поведение AS1 при выборе канала связи для доступа к сетям в AS3.

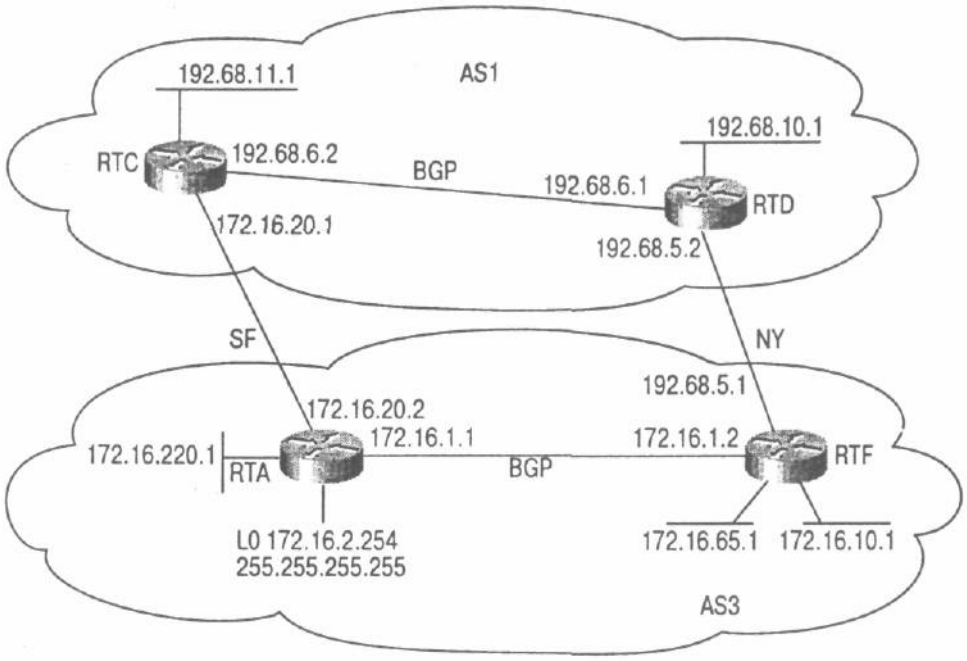


Рис. 11.12. *Объединение маршрутов в BGP с набором однозначно определенных маршрутов*

Маршрутизатор RTA через прямое соединение с AS1 будет посылать объединенный маршрут 172.16.0.0/16 и однозначно определенные маршруты 172.16.1.0/24, 172.16.10.0/24 и

172.16.65.0/26. Маршрутизатор RTF через прямое соединение с AS3 будет пересылать объединенный маршрут 172.16.0.0/16 и только один однозначно определенный маршрут 172.16.220.0/24. В результате AS1 будет вынуждена направлять часть трафика в сеть 172.16.220.0/24 через маршрутизатор RTF, а трафик по другим маршрутам — через маршрутизатор RTA.

Карта подавления маршрутов (suppress map) является частным случаем обычной *карты маршрутов (route map)* и позволяет указывать, какие однозначно определенные маршруты могут быть подавлены, а какие разрешены к распространению. Когда в карте подавления маршрутов указывается определенный маршрут, это означает, что он должен быть подавлен. Если маршрут не включен в карту подавления маршрутов (отклонен), то он не подавляется, т.е., другими словами, он разрешен к распространению. Обратите внимание: если вы запрещаете маршрут в карте подавления маршрутов, то это не означает, что сведения о нем не будут распространяться, а наоборот, этот маршрут запрещается подавлять. В листинге 11.65 приведен пример настройки карты подавления маршрутов.

Листинг 11.65. *Применение харты подавления маршрутов в BGP*

```
router bgp 3
 no  synchronization
 network  172.16.1.0 mask  255.255.255.0
 network  172.16.10.0 mask  255.255.255.0
 network  172.16.65.0 mask  255.255.255.192
 network  172.16.220.0 mask 255.255.255.0
 aggregate-address   172.16.0.0   255.255.0.0   suppress-map  SUPPRESS
 neighbor  172.16.1.2  remote-as  3
 neighbor  172.16.1.2  update-source Loopback0
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1   filter-list  10  out
 no  auto-summary

ip as-path access-list  10 permit   ^$
access-list 1 permit  172.16.220.0  0.0.0.255
access-list 1 deny any

route-map  SUPPRESS permit  10
 match  ip address   1
```

Карта подавления маршрутов используется на маршрутизаторе RTA с именем SUPPRESS. С ее помощью запрещается объявление маршрута 172.16.220.0/24 и разрешается объявление всех остальных маршрутов. В результате маршрутизатор RTA анонсирует объединенный маршрут 172.16.0.0/16 и однозначно определенные маршруты 172.16.1.0/24, 172Л6.10.0/24 и 172.16.65.0/26. В листинге 11.66 приведена таблица BGP-маршрутов на маршрутизаторе RTA. Снова обращаем ваше внимание на символ s в начале некоторых строк.

Листинг 11.66. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 17, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 * 1172.16.0.0        172.16.1.2                100        0       i
 *>                   0.0.0.0                              32768   i
 *> 172.16.1.0/24     0.0.0.0       0                      32768   i
 *> 172.16.10.0/24    172.16.1.2    20                     32768   i
 *> 172.16.65.0/26    172.16.1.2    20                     32768   i
 s> 172.16.220.0/24   0.0.0.0       0                      32768   i
 * i192.68.10.0       172.16.1.2    0           100        0       1 i
 *>                   172.16.20.1                          0       1 i
 * i192.68.11.0       172.16.1.2    0           100        0       1 i
 *>                   172.16.20.1   0                      0       1 i
```

Маршрутизатор RTF, руководствуясь похожей логикой, тоже будет объявлять объединенный и однозначно определенный маршрут 172.16.220.0/24. В листинге 11.67 показана конфигурация маршрутизатора RTF.

Листинг 11.67. *Конфигураций маршрутизатора RTF*

```
router bgp  3
 no  synchronization
 network  172.16.1.0 mask  255.255.255.0
 network  172.16.10.0 mask  255.255.255.0
 network  172.16.65.0 mask 255.255.255.192
 network  172.16.220.0 mask  255.255.255.0
 aggregate-address   172.16,0.0  255.255.0.0   suppress-map ALLOW
 neighbor  172.16.2.254  remote-as  3
 neighbor  172.16.2.254  next-hop-self
 neighbor  192.68.5.2   remote-as   1
 neighbor  192.68.5.2  filter-list 10 out
 no  auto-summary

ip as-path access-list 10 permit  ^$
access-list  1  deny  172.16.220.0   0.0.0.255
access-list  1  permit any

route-map ALLOW permit   10
 match  ip address   1
```

В конфигурацию маршрутизатора RTF, представленную в листинге 11.67, включена карта подавления маршрутов с именем ALLOW, с помощью которой разрешается распространение префикса 172.16.220.0/16 и подавляются все остальные маршруты. В результате AS1 для получения доступа в сеть 172.16.220.0/24 будет вынуждена использовать маршрутизатор RTF. Названия карт подавления маршрутов — SUPPRESS и ALLOW - отражают основную функцию карты маршрутов. Так, на маршрутизаторе RTA более удобно подавить один однозначно определенный маршрут и разрешить все остальные, так как количество *разрешенных к распространению* маршрутов достаточно велико. При настройке маршрутизатора RTF более целесообразно разрешить один маршрут и подавить все остальные, так как число *подавляемых* маршрутов достаточно велико.

Конфигурация маршрутизатора RTF, приведенная в листинге 11.67, позволяет объявлять объединенный маршрут 172.16.0.0/16 и однозначно определенный маршрут 172.16.220.0/24, а все остальные маршруты подавлять. В листинге 11.68 показана таблица BGP-маршрутов на маршрутизаторе RTF.

Листинг 11.68. *Таблица BGP-маршрутов на маршрутизаторе RTF*

```
RTC# show ip bgp
BGP table version is 17, local router ID is 192.68.5.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *> 172.16.0.0        0.0.0.0                               32768   i
 * i                  172.16.2.254                100        0       i
 s> 172.16.1.0/24     0.0.0.0       0                       32768   i
 s> i                 172.16.2.254  0             100       0       i
 s> 172.16.10.0/24    0.0.0.0       0                       32768   i
 s> i                 172.16.2.254  20            100       0       i
 s> 172.16.65.0/26    0.0.0.0       0                       32768   i
 s> i                 172.16.2.254  20            100       0       i
 *> 172.16.220.0/24   172.16.1.1    20                      32768   i
 *> i192.68.10.0      192.68.5.2    0                       0       1 i
 * i                  172.16.20.1                 100       0       1 i
 * 192.68.11.0        192.68.5.2                            0       1 i
 * i                  172.16.20.1                 100       0       1 i
```

Учитывая представленные в листингах 11.65 и 11.67 конфигурации маршрутизаторов RTA и RTF, AS1 сможет попасть в сеть 172.16.220.0/24 только по каналу

RTF-RTD, а в сети 172.16.1.0/24, 172.16.65.0/26 и 172.16.10.0/24- только по каналу RTC-RTA. Это видно из BGP-таблицы на маршрутизаторе RTD, представленной в листинге 11.69.

Листинг 11.69. *Таблица BGR-маршрутов на маршрутизаторе RTD*

```
RTC# show ip bgp
BGP table version is 19, local router ID is 192.68.10.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network                Next Hop      Metric      LocPrf     Weight  Path
 * i172.16.0.0          192.68.6.2                100        0       3 i
 *>                     192.68.5.1                           0       3 i
 *> i172.16.1.0/24      192.68.6.2    0           100        0       3 i
 *> i172.16.10.0/24     192.68.6.2    20          100        0       3 i
 *> i172.16.65.0/26     192.68.6.2    20          100        0       3 i
 *>  172.16.220.0/24    192.68.5.1    20                     0       3 i
 *>  192.68.10.0        0.0.0.0       0                      32768   i
 *> i192.68.11.0        192.68.6.2    0           100        0       i
```

Для маршрутизатора RTD имеется только один путь достичь сети 172.16.220.0/24 - через канал между RTD и RTF. В случае пропадания канала объединенный маршрут по-прежнему будет объявляться по обоим каналам, так что трафик будет следовать по объединенному маршруту.

В отдельных случаях администраторы требуют, чтобы несколько соседних узлов получали информацию о некоторых уже подавленных однозначно определенных маршрутах. Подавление могло выполняться с помощью команды neighbor с параметром summary-only. В этом случае компания Cisco предоставляет еще одну форму управления маршрутами, которая называется *карта неподавленных маршрутов (unsuppress map).* Эта карта применяется только между соседними узлами. Карта неподавленных маршрутов позволяет объявлять ранее подавленные маршруты. Например, если вам нужно, чтобы маршрут от RTA 172.16.220.0/24 в направлении маршрутизатора RTF (I72.16.1.2) не подавлялся, то используйте конфигурацию RTA, предлагаемую в листинге 11.70.

Листинг 11.70. *Конфигурация маршрутизатора RTA*

```
neighbor  172.16.1.2  unsuppress-map AllowSpecifics

route-map AllowSpecifics  permit   10
 match  ip  address   1

access-list  1  permit   172.16.220.0   0.0.0.255
```

Конфигурация маршрутизатора, приведенная в листинге 11.70, позволяет объявлять префикс 172.16.220.0/24 в направлении маршрутизатора RTF.

Потери информации в объединенном маршруте

При объединении маршрутов наблюдаются потери маршрутной информации. Подробные сведения о маршруте, которые переносятся в однозначно определенных префиксах, будут теряться при суммировании в объединенный маршрут. Параметр AS_SET используется для сохранения атрибутов отдельных однозначно определенных маршрутов в виде математического набора, который позволял бы получить больше информации об элементах объединенного маршрута. На рис. 11.13 представлена топология сети, на которую мы будем опираться в этом разделе.

На рис. 11.13 маршрутизатор RTA объединяет префиксы 192.68.10.0/24 и I92.68.1I.O/24, поступающие, соответственно, от AS2 и AS1. Без использования параметра AS_SET объединенный маршрут 192.68.0.0/16 рассматривается как сгенерированный в AS3, и вся информация об атрибутах отдельных префиксов 192.68.10.0/24 и 192.68.11.0/24 теряется рассмотрим два варианта конфигурации маршрутизатора RTA - без использования

AS_SET (листинг 11.71) и с применением AS_SET (листинг 11.73). Вы увидите, как будет вести себя объединенный маршрут 192.68.0.0/16 в этих случаях.

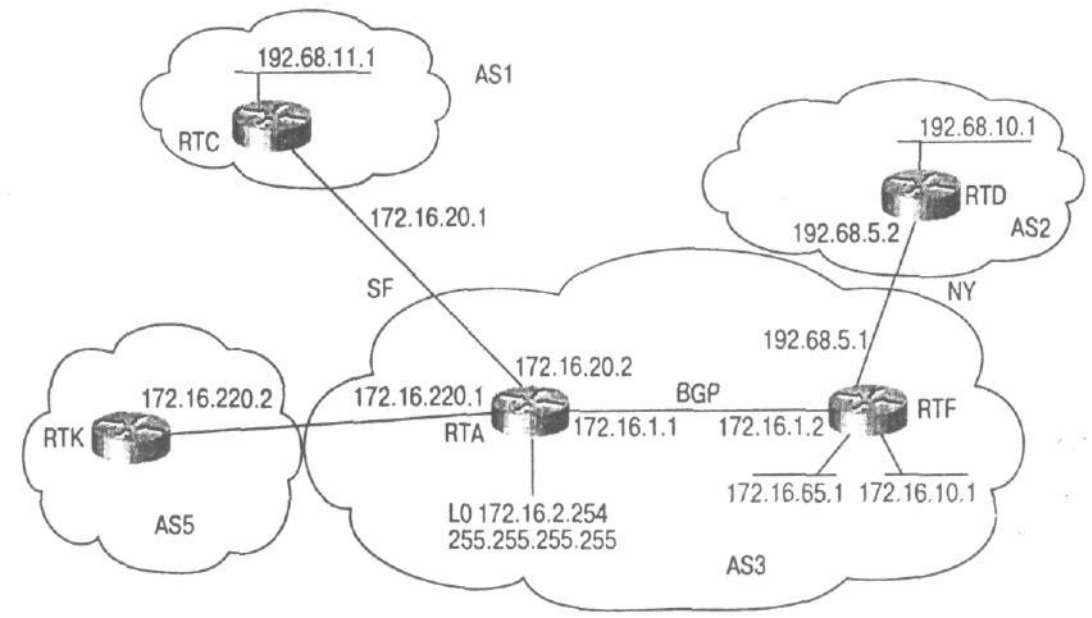


***Рис. 11.13**. Борьба с потерями маршрутной информации в объединенных маршрутах*

***Листинг 11.71.** Конфигурация маршрутизатора ЙТАбез применения AS_SET*

```
router bgp  3
 no  synchronization
 network  172.16.1.0  mask  255.255.255.0
 network  172.16.10.0  mask  255.255.255.0
 network  172.16.65.0 mask  255.255.255.192
 network  172.16.220.0  mask  255.255.255.0
 aggregate-address   192.68.0.0  255.255.0.0
 neighbor  172.16.1.2   remote-as   3
 neighbor  172.16.1.2  update-source  Loopback0
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1   filter-list  10  out
 neighbor  172.16.220.2  remote-as  5
 no  auto-summary

ip  as-path  access-list  10  permit   ^$
```

Из BGP-таблицы на маршрутизаторе RTK (листинг 11.72) видно, как будет выглядеть объединенный маршрут 192.68.0.0/16. Обратите внимание, что в объединенном маршруте отсутствуют сведения об индивидуальных маршрутах, так как в AS_PATH представлен только один номер AS — 3.

***Листинг 11.72.** ТаблицаBGP-маршрутов на маршрутазаторе RTK*

```
RTC# show ip bgp
BGP table version is 8, local router ID is 172.16.220.2
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf     Weight  Path
 *> 172.16.1.0/24     172.16.220.1  0                      0       3 i
 *> 172.16.10.0/24    172.16.220.1  20                     0       3 i
 *> 172.16.65.0/26    172.16.220.1  20                     0       3 i
 *> 172.16.220.0/24   172.16.220.1  0                      0       3 i
 *> 192.68.0.0/16     172.16.220.1                         0       3 i
 *> 192.68.10.0       172.16.220.1                         0       3 2 i
 *> 192.68.11.0       172.16.220.1                         0       3 1 i
```

Применяя параметр AS_SET, как это показано в листинге 11.73, вы сможете посылать объединенный маршрут от RTA с набором определенной маршрутной информации о его элементах.

Листинг 11.73. *Конфигурация маршрутизатора RTA с использованием AS_SET*

```
router bgp 3
 no synchronization
 network  172.16.1.0 mask 255.255.255.0
 network  172.16.10.0 mask 255.255.255.0
 network  172.16.65.0 mask 255.255.255.192
 network  172.16.220.0 mask  255.255.255.0
 aggregate-address  192.68.0.0  255.255.0.0  as-set
 neighbor  172.16.1.2  remote-as  3
 neighbor  172.16.1.2  update-source  Loopback0
 neighbor  172.16.20.1  remote-as  1
 neighbor  172.16.20.1  filter-list  10  out
 neighbor 172.16.220.2 remote-as 5
 no auto-summary

ip as-path access-list  10 permit  л$
```

Обратите внимание, как в таблице BGP-маршрутов, приведенной в листинге 11.74, изменяется объединенный маршрут 192.68.0.0/16 при введении в маршрутную информацию параметра SET {2,1}- Таким образом указывается, что в объединенный маршрут включены маршруты, которые прошли через AS1 или AS2. Информация из AS_SET крайне необходима, чтобы избежать появления петель маршрутизации, так как дает возможность узнать, через какие узлы прошел маршрут.

Листинг 11.74. *Таблица BGP-маршрутов на маршрутизаторе RTK*

```
RTC# show ip bgp
BGP table version is 12, local router ID is 172.16.220.2
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric     LocPrf   Weight  Path
 *> 172.16.1.0/24     172.16.220.1  0                   0       3 i
 *> 172.16.10.0/24    172.16.220.1  20                  0       3 i
 *> 172.16.65.0/26    172.16.220.1  20                  0       3 i
 *> 172.16.220.0/24   172.16.220.1  0                   0       3 i
 *> 192.68.0.0/16     172.16.220.1                      0       3 {2,1}
                                                                i
 *> 192.68.10.0       172.16.220.1                      0       3 2 i
 *> 192.68.11.0       172.16.220.1                      0       3 1 i
```

Если объединенный маршрут попадает на AS1 или AS2, то для обнаружения петель маршрутизации в протоколе BGP он будет подвергнут анализу, в результате которого будет выявлен набор маршрутной информации AS_SET, и объединенный маршрут будет отвергнут.

Предполагая, что объединенный маршрут с параметром AS_SET содержит сведения о каждом входящем в его состав маршруте, необходимо представлять, что внесение изменений в маршрутную информацию отдельного маршрута приведет к обновлению маршрутной информации объединенного маршрута в целом. Например, если маршрут 192.68.11.0/24 становится недоступным, маршрутная информация объединенного маршрута будет изменяться с 3 {2,1} на 3 2, т.е. объединенный маршрут будет обновляться. Если объединенный маршрут сформирован из десятков или даже сотен маршрутов, он будет испытывать постоянные колебания, если какие-либо из составляющих его маршрутов окажутся нестабильными.

Изменение атрибутов объединенного маршрута

В некоторых случаях требуется внести изменения в атрибуты объединенного маршрута, В этом разделе мы приведем один пример, когда такие действия могут оказаться полезными и даже необходимыми.

Как вы уже убедились, объединенный маршрут может переносить информацию об отдельных входящих в него элементах, если ему задан параметр AS_SET. Если один или несколько маршрутов, формирующих в объединенном маршруте AS_SET, сконфигурированы с атрибутом COMMUNITY **no-export,** то объединенный маршрут будет иметь такой же атрибут. Но в этом случае его нельзя передавать в другие сети. Чтобы исправить ситуацию, имеется возможность изменить атрибут COMMUNITY объединенного маршрута с помощью так называемой **карты атрибутов (atribute map),** которая является по сути еше одной формой карты маршрутов, но применяемой только к атрибутам.

В сети, показанной на рис. 11.13, маршрутизатор RTC помечает маршрут 192.68.11.0/24 атрибутом COMMUNITY со значением **no-export.** Если на маршрутизаторе RTA проводится объединение маршрутов 192.68.11.0/24 в 192.68.0.0/16 с использованием AS_SET, то в объединенный маршрут также перейдет атрибут COMMUNITY со значением **no-export.** В листингах 11.75 и 11.76 представлены конфигурации маршрутизаторов RTC и RTA, которые реализуют именно такую схему.

Листинг 11.75. *Конфигурация маршрутизатора RTC*

```
router bgp 1
 network 192.68.11.0
 neighbor 172.16.20.2 remote-as 3
 neighbor 172.16.20.2 send-community
 neighbor 172.16.20.2 route-map SET-COMMUNITY out
 no auto-summary

access-list 1 permit 192.68.11.0 0.0.0.255

route-map SETCOMMUNITY permit 10
 match ip address 1
 set community no-export

route-map SETCOMMUNITY permit 20
```

Листинг 11.76. *Конфигурация маршрутизатора RTA*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
network  172.16.10.0 mask  255.255.255.0
 network  172.16.65.0 mask  255.255.255.192
 network  172.16.220.0 mask  255.255.255.0
 aggregate-address  192.68.0.0  255.255.0.0  as-set
 neighbor 172.16.1.2   remote-as  3
 neighbor  172.16.1.2  update-source  Loopback0
 neighbor 172.16.20.1  remote-as  1
 neighbor  172.16.20.1  filter-list  10  out
 neighbor  172.16.220.2  remote-as   5
 no  auto-summary

ip as-path access-list 10 permit  ^$
```

Поскольку маршрутизатор RTA объединяет маршруты с использованием AS_SET, то объединенный маршрут будет содержать все элементы, присущие индивидуальным маршрутам, в частности атрибут COMMUNITY со значением **no-export,** поступающий вместе с префиксом 192.68.11.0/24 (сгенерированным маршрутизатором RTC). Обратите внимание как в таблице BGP-маршрутов на маршрутизаторе RTA, представленной в листинге 11.77, показан объединенный маршрут 192.68.0.0/16. Как видите, его запрещено

объявлять внешним EBGP-узлам.

Листинг 11.77. *Информация о BGP-маршруте на маршрутизаторе RTA*

```
RTA#show ip bgp 192.68.0.0
BGP routing table entry for 192.68.0.0 255.255.0.0, version 22
Paths: (2 available, best #2, not advertised to EBGP peer, advertised
over IBGP
    Local (aggregated by 3 192.68.5.1)
from 172.16.1.2 (192.68.5.1)
        Origin IGP, localpref 100, valid, internal, atomic-
aggregate
    {2,1} (aggregated by 3 172.16.2.254)
      0.0.0.0
        Origin IGP, localpref 100, weight 32768, valid,
aggregated,
 local, best
        Community: no-export
```

Используя карты атрибутов, вы можете манипулировать атрибутами объединенного маршрута. В примере, приведенном ниже, показано, как "очистить" атрибут COMMUNITY и разрешить объявление маршрутов EBGP-узлам. В конфигурации маршрутизатора RTA, приведенной в листинге **11.78,** встречается карта атрибутов с именем SET_ATTRIBUTE, с помощью которой объединенному маршруту устанавливается атрибут COMMUNITY без параметров.

Листинг 11.78. *Конфигурация маршрутизатора RTA*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 aggregate-address 192.68.0.0 255.255.0.0 as-set attribute-map
SET_ATTRIBUT?
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.220.2 remote-as 5
 no auto-summary

ip as-path access-list 10 permit ^$

route-map SET_ATTRIBUTE permit 10
 set community none
```

Посмотрите, как теперь в таблице BGP-маршрутов (листинг 11.79) представлен объединенный маршрут 192.68.0.0/16. Здесь уже есть возможность объявлять его другим EBGP-узлам.

Листинг 11.79. *Вывод сведений о BGP-маршруте на маршрутизаторе RTA*

```
RTA#show ip bgp 192.68.0.0
BGP routing table entry for 192.68.0.0 255.255.0.0, version 10
Paths: (2 available, best #2, advertised over IBGP, EBGP)
  Local (aggregated by 3 192.68.5.1)
    172.16.1.2 from 172.16.1.2 (192.68.5.1)
      Origin IGP, localpref 100, valid, internal, atomic-aggregate
 {2,1} (aggregated by 3 172.16.2.254)
      0.0.0.0
      Origin IGP, localpref 180, weight 32768, valid, aggregated,
 local, best
```

Формирование объединенного маршрута на основе наборов однозначно определенных маршрутов

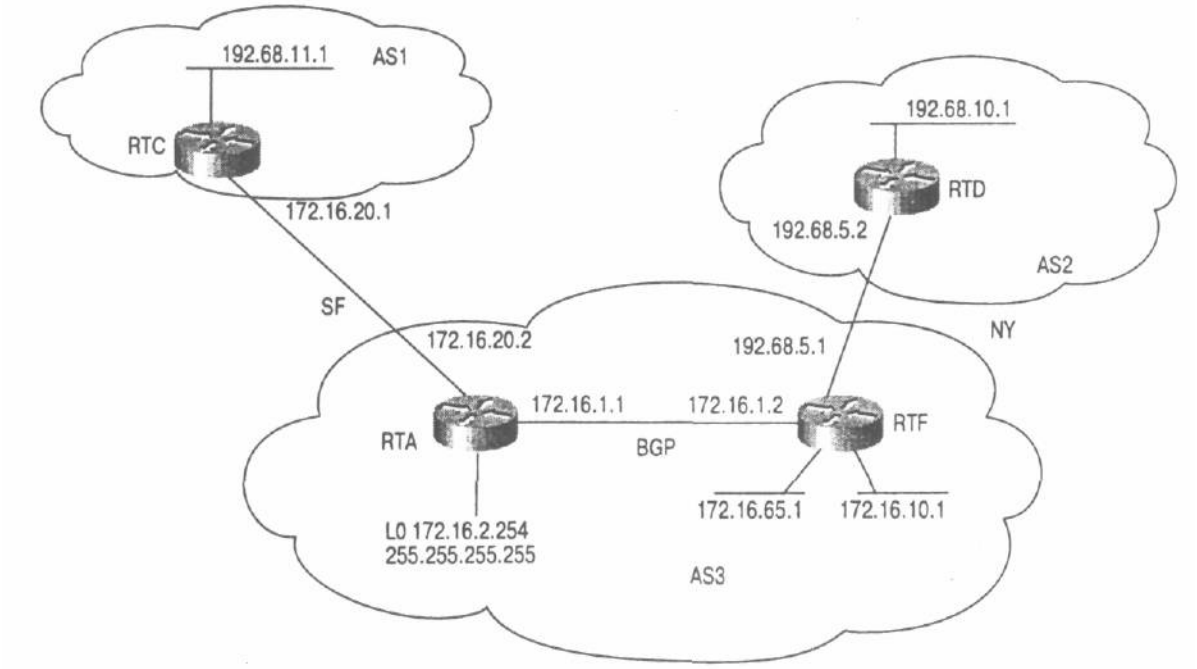


Рис. 11.14. *Объединение маршрутов с помощью карт объявления маршрутов*

Наличие информации об отдельных префиксах, формирующих объединенный маршрут, позволит определить, какие атрибуты будет переносить объединенный атрибут. Если бы была возможность в примере предыдущего раздела исключить префикс 192.68.11.0/24 из набора префиксов, формирующих объединенный маршрут, то в объединенный маршрут не попал бы атрибут COMMUNITY со значением no-export.

Существует еще одна форма карты маршрутов - - *карта объявления маршрутов (advertising map), с* помощью которой можно сформировать объединенный маршрут на основе ограниченного набора однозначно определенных маршрутов. На рис. 11.14 показана топология сети, которую мы рассмотрим в качестве примера.

Как видно из рис. 11.14, маршрутизаторы RTA и RTF получают сведения о маршрутах 192.68.11.0/24 и 192.68.10.0/24, соответственно, от AS1 и AS2. Если на маршрутизаторах RTA и RTF проводится объединение этих маршрутов в 192.68.0.0/16 с параметром as-set, то объединенный маршрут не может попасть обрагно в AS1 или AS2, так как в AS_PATH {1 2} уже есть сведения о маршрутах в эти автономные системы. Это происходит благодаря механизму обнаружения петель маршрутизации в протоколе BGP.

Допустим, необходимо посылать сведения о маршруте 192.68.0.0/16 не в AS2, а обратно в AS1. Решение этой проблемы — не включать AS1 в атрибут AS_PATH, тогда AS1 не будет отвергать объединенный маршрут. Этого можно добиться, сформировав с помощью параметра advertise-map объединенный маршрут на маршрутизаторах RTA и RTF только на основе префикса 192.68.10.0/24.

Чтобы достичь нужных результатов, маршрутизатор RTA следует сконфигурировать так, как показано в листинге 11.80. На маршрутизаторе RTF конфигурация будет лишь незначительно отличаться.

Листинг 11.80. *Конфигурация маршрутизатора RTA*

```
router bgp  3
 no  synchronization
 network   172.16.1.0  mask  255.255.255.0
 network  172.16.10.0 mask 255.255.255.0
 network  172.16.65.0 mask 255.255.255.192
 aggregate-address   192.68.0.0  255.255.0.0  as-set advertise-map
```

```
    SE LECT_MORE_S PE CI F_ROUTES
 neighbor  172.16.1.2   remote-as   3
 neighbor  172.16.1.2  update-source  LoopbackO
 neighbor   172.16.20.1   remote-as   1
 neighbor   172.16.20.1   filter-list  10   out
 no  auto-summary

ip as-path  access-list  10 permit  A$

access-list  1  permit  192.68.10.0   0.0.0.255

route-map  SELECT_MORE_SPECIF_ROUTES  permit   10
 match  ip  address   1
```

Разрешая распространение префикса 192.68.10.0/24, карта объявления маршрутов вынуждает маршрутизатор RTA формировать объединенный маршрут только на базе этого маршрута. Таким образом, префикс 192.68.11.0/24 в формировании объединенного маршрута не участвует.

Воспользовавшись командой show ip bgp, можно получить сведения о заданном BGP-маршруте. Как видите, AS_PATH объединенного маршрута равен 2 , а не {1 2} (листинг 11.81). Это означает, что объединенный маршрут может теперь объявляться в AS1, так как в AS_PATH отсутствуют сведения о его прохождении через эту автономную систему. Однако AS2 по-прежнему не сможет принимать объединенный маршрут.

Листинг 11.81. *Сведения о BGP-маршруте на маршрутизаторе RTA*

```
RTA#show ip bgp 192.68.0.0
BGP routing table entry for 192.68.0.0 255.255.0.0, version 31
Paths: (2 available, best #2, advertised over IBGP)
 2 (aggregated by 3 192.68.5.1)
   172.16.1.2 from 172.16.1.2 (192.68.5.1)
     Origin IGP, localpref 100, valid, internal, atomic-aggregate
 2 (aggregated by 3 172.16.2.254)
   0.0.0.0
     Origin IGP, localpref 100, weight 32768, valid, aggregated,
 local, atomic-aggregate,best
```

Забегая вперед

Атрибуты BGP являются основными элементами при организации взаимодейств* различных сетей. Комбинирование и манипулирование различными атрибутами п< может сформировать уникальные для вашей автономной системы правила маршрут! зации. В следующей главе все изученное будет рассмотрено на примерах реальнь схем. Будут затронуты проблемы, с которыми постоянно сталкиваются сетевые адм! нистраторы при создании сетей. В следующей главе показаны примеры управлен* стабильностью в сети Internet с использованием технологий подавления колебани маршрутов, мягкой перенастройки и обновления BGP-маршрутов. Мы также обсуди возможность организации так называемого фильтра исходящих маршрутов (Outbour Route Filter — ORF) и работы с мультипротокольным расширением протокола BG (Multiprotocol BGP -- MBGP). И закончим мы книгу рассмотрением практически подходов к формированию правил маршрутизации.

Ключевые темы этой главы:

- **Избыточность, симметрия и распределение нагрузки.** Рассмотрены примеры конфигураций для динамических и статических маршрутов по умолчанию, для организации подключения по нескольким каналам к одному и нескольким провайдерам, для распределения нагрузки и создания схем, где клиенты совместно используют резервный канал.
- **Установка маршрутов по умолчанию.** Представлены примеры конфигурирования маршрутизаторов для определения-маршрутов по умолчанию 5 различных сетевых архитектурах. В частности подробно анализируется проблема выбора однго маршрута из множества IGP-маршрутов.
- **Маршрутизация по правилам.** Приводятся примеры Застройки маршрутизации по адресу источника.
- **Отражатели маршрутов.** Приводятся действующие схемы применения отражателей маршрутов в группах взаимодействующих узлов.
- **Конфедерации.** Приводятся примеры применения конфедераций.
- **Управление маршрутами и аннулированием содержимого кэша.** Рассматриваются синтаксис и примеры настройки протокола BGP с использованием мягкой перенастройки и возможности обновления BGP-маршрутов (BGP Route Refresh capability).
- **Работа маршрутизатора в режиме фильтра исходящих BGP-маршрутов (Outbound Route Filter — ORF).** Фильтр исходящих BGP-маршрутов является одной из дополнительных возможностей протокола BGP и применяется с целью "проталкивания" префиксов от соседнего BGP-узла через его собственный фильтр на основе списков префиксов. Мы обсудим преимущества использования BGP ORF и представим вашему вниманию несколько вариантов его настройки.
- **Разгрузка маршрутов.** Приводится синтаксис и примеры настройки функции разгрузки маршрутов BGP.

Глава 12. Настройка эффективных правил маршрутизации в сети Internet

В главе 11, "Настройка основных функций и атрибутов BGP", мы рассматривали примеры конфигурации основных функций и атрибутов протокола BGP. В этой главе мы рассмотрим примеры конфигурации, которые отражают наиболее сложные и потенциально конфликтные схемы маршрутизации. Кроме того, здесь вы найдете примеры конфигурации сети, которые помогут вам контролировать рост сложных сетей.

Возможно, одной из наиболее сложных задач является разработка правил маршрутизации ля используемой сетевой архитектуры. Конечно, это первое, что вы должны сделать перед тем, как начнете процесс настройки этих правил. И для этого существует метод жесткого определения правил маршрутизации, где четко описан порядок операций. Тщательный анализ функционирования сети и потенциальных возможностей ее роста поможет выявить имеющиеся проблемы и оптимальные пути их решения.

Избыточность, симметрия и распределение нагрузки

Примеры конфигурации, приведенные в этом разделе, описывают способы реализации избыточности маршрутов, симметрии и различные варианты распределения нагрузки в сети. Пожалуйста, помните, что все, предложенные нами варианты не являются истиной в последней инстанции. Но мы надеемся, что большинство наших рекомендаций может быть вами использовано. Примеры, которые мы приводим здесь, помогут вам уяснить порядок и способы установки правил маршрутизации. Мы начнем с нескольких примеров реализации маршрутизации по умолчанию.

Динамические маршруты по умолчанию

В протокол BGP очень важно контролировать маршруты по умолчанию, так как, будучи сгенерированы случайным образом, они могут привести к появлению серьезных проблем, во всей системе маршрутизации. Подобные случаи могут быть причиной того, что BGP-спикер, который намеревался сгенерировать маршрут по умолчанию для определенного узла, переполняется информацией о маршрутах по умолчанию, поступающей от соседних узлов. В оборудовании Cisco предусмотрена возможность точного указания маршрута по умолчанию в направлении определенного узла.

На рис. 12.1 маршрутизатор RTA генерирует маршрут по умолчанию 0.0.0.0/0 только в направлении маршрутизатора RTC. Соседние IBGP-узлы, такие как RTF, не получают сведений об этом маршруте по умолчанию.

В листинге 12.1 представлена конфигурация маршрутизатора RTA.

Листинг 12.1. *Конфигурация маршрутизатора PTAдля работы с динамическими маршрутами по умолчанию*

```
router bgp  3
 no  synchronization
 network  172.16.1.0  mask  255.255.255.0
 neighbor   172.16.20.1   remote-as   1
 neighbor  172.16.20.1  default-originate
 no  auto-summary
```

Параметр **default-originate,** который используется в команде маршрутизатора **neighbor,** осуществляет пересылку маршрута по умолчанию 0/0 на маршрутизатор RTC. Это видно из таблиц IP- и BGP-маршрутов, приведенных в листинге 12.2.

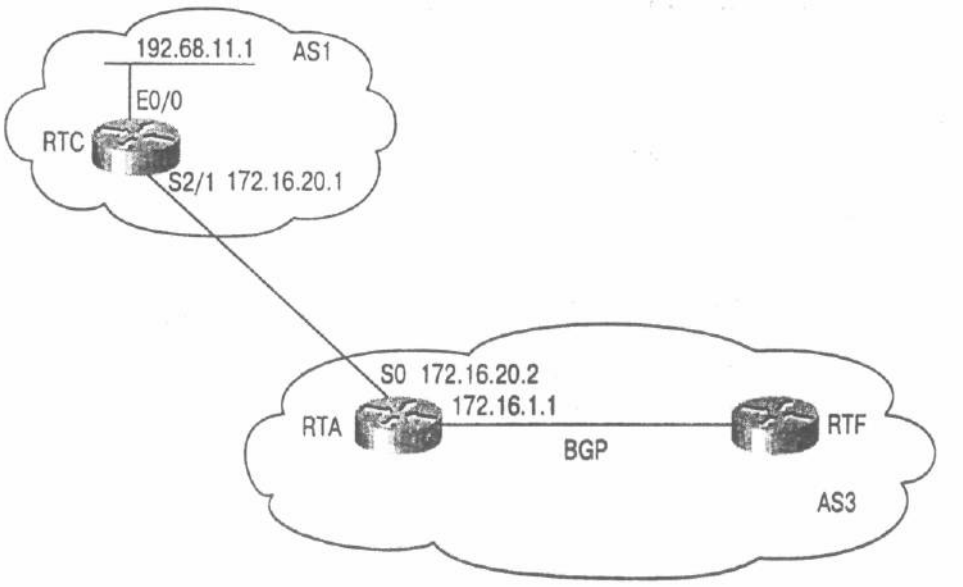


Рис. 12.1. *Динамические маршруты по умолчанию*

Листинг 12.2. *Таблицы IP- и BGP-маршрутов н маршрутизаторе RTC*

```
RTC# show ip bgp
BGP table version is 14, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, i - incomplete
 Network              Next Hop      Metric      LocPrf    Weight   Path
 *> 0.0.0.0           172.16.20.2   0                     0        3 i
 *> 172.16.1.0/24     172.16.20.1   0                     0        3 i
 *> 192.68.11.0       0.0.0.0       0                     32768    i

RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       El•- OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route

Gateway of  last  resort  is  172.16.20,2  to network  0.0.0.0
C    192.68.11.0/24   is  directly connected,   EthernetO/0
C    172.16.20.0/24  is directly connected,   Serial2/1
B*   0.0.0.0/0   [20/0]   via  172.16.20.2,   00:04:40
```

Как видно из таблицы маршрутов на маршрутизаторе RTC, представленной в листинге 12.2, маршрут по умолчанию 0/0 получен динамически от маршрутизатора RTA и установлен резервный шлюз — 172.16.20.2 (т.е. маршрутизатор RTA).

Маршруты по умолчанию могут также генерироваться для всех BGP-узлов с использованием команды маршрутизатора network 0.0.0.0 до тех пор, пока маршрутизатор будет объявлять этот маршрут по умолчанию как свой собственный. Предполагая, что у RTA

есть маршрут по умолчанию (он может быть установлен и статически), вы можете воспользоваться конфигурацией, приведенной в листинге 12.3.

Листинг 12.3. *Генерирование маршрутов по умолчанию для всех BGP-узлов (конфигурация маршрутизатора RTA).*

```
router bgp 3
 no synchronization
 network 0,0.0.0
 network 172.16.1.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 no auto-summary
```

Статические маршруты по умолчанию

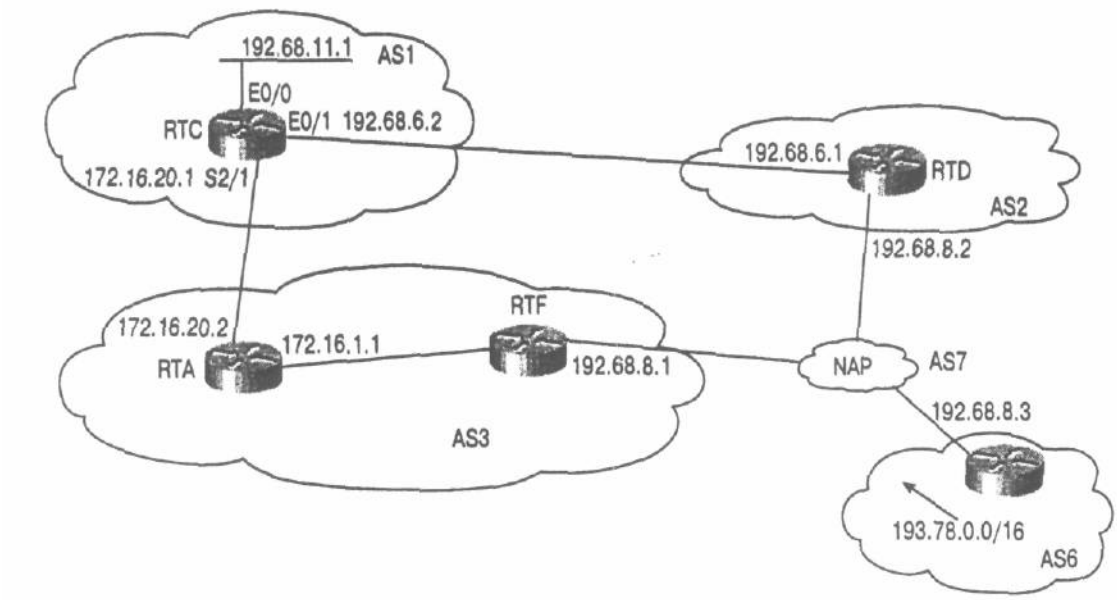


Рис. 12.2. *Обработка маршрута по умолчанию 0/0*

Вместо того чтобы получать сведения о маршруте по умолчанию 0/0 динамически, маршрутизатор может установить собственный статический маршрут по умолчанию. На рис. 12.2 показано, как это делается.

Вы можете на маршрутизаторе RTC использовать следующую команду:

ip route prefix mask {address \ interface} [distance]

Статический маршрут по умолчанию 0/0 при этом может указывать на сетевой адрес, адрес шлюза или физический интерфейс. Здесь параметр *distance* является средством, с помощью которого вы можете задать статическому маршруту определенный уровень предпочтения, если существует несколько точек входа в сеть. Маршруты с меньшей дистанцией (значением *distance)* предпочитаются маршрутам с более длинной дистанцией.

В листинге 12.4 показано, как на маршрутизаторе RTC установить маршрут по умолчанию в направлении сети 193.78.0.0/16.

Листинг 12.4. *Конфигурирование маршрутизатора RTC для установки маршрута по умолчанию в сеть 193.78.0.0/16*

```
router bgp  1
 network 192.68.11.0
 neighbor 172.16.20.2  remote-as  3
 neighbor 192.68.6.1  remote-as  2
no auto-summary ip  route  0.0.0.0   0.0.0.0   193.78.0.0
```

Из таблицы BGP-маршрутов на маршрутизаторе RTC, представленной в листинге

12.5, видно, что маршрут в сеть 193.78.0.0/16 был получен двумя путями — через AS3 и через AS2. В протоколе BGP, как мы знаем, предпочтение отдается наилучшему маршруту. (Чтобы повлиять на выбор маршрута, можно использовать атрибуты BGP).

Листинг 12.5. *Таблица BGP-маршрутов на маршрутизаторе RTC*

```
RTC# show ip bgp
BGP table version is 8, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric     LocPrf   Weight   Path
 *> 192.68.11.0       0.0.0.0       0                   32768    i
 *> 193.78.0.0/16     172.16.20.2                       0        3 7 6 i
 *                    192.68.6.1                        0        2 7 6 i
```

Из таблицы IP-маршрутов на маршрутизаторе RTC, приведенной в листинге 12.6, видно, каким образом установлен шлюз по умолчанию, чтобы достичь сети 193.78.0.0./16. При рекурсивном просмотре таблицы IP-маршрутов видно, что в сеть 193.78.0.0/16 можно попасть через узел с адресом 172.16.20.2 (т.е. RTA).

Листинг 12.6. *Таблица IP-маршрутов на маршрутизаторе RTC*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default u - per-user static route

Gateway of last resort is 193.78.0.0 to network 0.0.0.0
C  192.68.6.0/24 is directly connected, Ethernet0/1
C  192.68.11.0/24 is directly connected, Ethernet0/0
B  193.78.0.0/16 [20/0] via 172.16.20.2, 00:32:32
C  172.16.20.0/24 is directly connected, Serial2/l
S* 0.0.0.0/0 [1/0] via 193.78.0.0
```

Если вы не хотите, чтобы трафик следовал по одному маршруту, то можно воспользоваться командой **ip route 0.0.0.0 0.0.0.0,** чтобы направить его в несколько сетей сразу или по нескольким IP-адресам одновременно. Ключевое слово **distance** дает вам в этом случае возможность устанавливать для маршрутов уровни предпочтения, как это показано в листинге 12.7.

Листинг 12.7. *Применение distance для предпочтения одного маршрута по умолчанию другому (конфигурация маршрутизатора RTC)*

```
router bgp 1
 network 192.68.11.0
 neighbor  172.16.20.2   remote-as   3
 neighbor  192.68.6.1  remote-as  2
 no  auto-summary

ip route  0.0.0.0   0.0.0.0  172.16.20.2   40
ip route  0.0.0.0   0.0.0.0  192.68.6.1    50
```

Обратите внимание, что на маршрутизаторе RTC указываются два различных IP-адреса. Это могут быть также два адреса сетей, которые присутствуют в маршрутной IP-таблице. Дистанция первого статического маршрута 40 говорит о том, что до тех пор, пока будет доступен узел 172.16.20.2, этот маршрут будет предпочитаться второму. Если же узел 172.16.20.2 выходит из строя, то запись о статическом маршруте к нему изымается и активным становится второй маршрут (см. таблицу маршрутов на маршрутизаторе RTC в листинге 12.8).

Листинг 12.8. *Таблица маршрутов RTC*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R -- RIP, M -- mobile, B - BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E – EGP
       i – IS-IS, L1 – IS-IS level-1, L2 – IS-IS level-2,
         *- candidate default U - per-user static route
Gateway of last resort is 172.16.20.2 to network 0.0.0.0
C     192.68.6.0/24 is directly connected, EthernetO/1
C     192.68.11.0/24 is directly connected, EthernetO/0
B     193.78,0.0/16 [20/0] via 172.16.20.2, 00:45:08
C     172.16.20.0/24 is directly connected, Serial2/1
S*    0.0.0.0/0 [40/0] via 172.16.20.2
```

В листинге 12.9 показаны результаты выполнения той же команды, если пропадает канал между RTA и RTC.

Листинг 12.9. *Таблица маршрутов на маршрутизаторе RTC после пропадания канала RTC.RTA*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R -- RIP, M -- mobile, B - BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E – EGP
       i – IS-IS, L1 – IS-IS level-1, L2 – IS-IS level-2,
         *- candidate default U - per-user static route
Gateway of last resort is 192.68.6.1 to network 0.0.0.0
C  192.68.6.0/24 is directly connected, EthernetO/1
C  192.68.11.0/24 is directly connected, EthernetO/0
B  193.78.0.0/16 [20/0] via 192.68.6.1, 00:01:14
S* 0.0.0.0/0 {60/0} via 192.68.6.1
```

Подключение к одному провайдеру по нескольким каналам

Для случаев подключения клиента к одному провайдеру мы рассмотрим следующие примеры конфигурации.

- Маршрутизация только по умолчанию: один основной и один резервный канал.
- Маршрутизация по умолчанию: один основной и один резервный канал плюс частичная маршрутизация.
- Автоматическое распределение нагрузки.

Маршрутизация только по умолчанию: один основной и один резервный канал

На рис. 12,3 представлена AS3, которая подключена по двум каналам к автономной системе AS1. Система AS3 не получает сведений о BGP-маршрутах от AS1, но сама посылает сведения о собственных BGP-маршрутах. На маршрутизаторе RTA будут описаны маршруты по умолчанию в направлении AS1, при этом канал с узлом AS1 в Нью-Йорке (NY) будет основным, а канал с другим узлом в AS1 в Сан-Франциско (SF) — резервным.

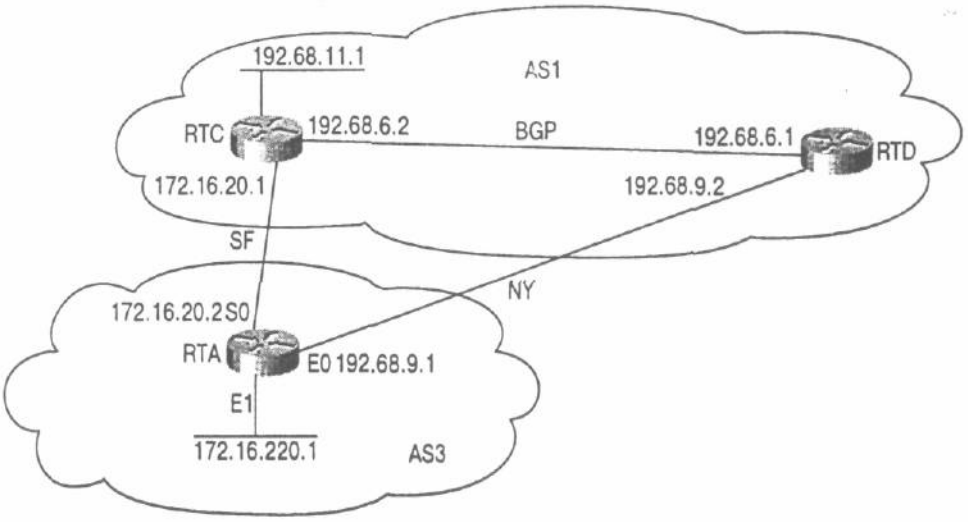


Рис. 12.3. *Подключение к одному провайдеру по нескольким каналам (маршрутизация только по умолчанию: один основной и один резервный канал}*

При такой организации сети в ней требуется установить следующие правила маршрутизации.

- Исходящий трафик AS3 всегда должен передаваться по каналу NY, если он исправен. В противном случае происходит перекоммугация трафика на резервный канал SF.

  Это осуществляется путем настройки на маршрутизаторе RTA двух статических маршрутов, которые по умолчанию направляют трафик в соответствующие каналы. При этом нужно установить маршрут по умолчанию через канал NY с меньшей дистанцией, тогда при выборе маршрута ему будет отдаваться предпочтение.

- Входящий трафик AS3 всегда должен поступать по каналу NY, если он исправен. В противном случае происходит перекоммутация трафика на резервный канал SF.

  Для этого на маршрутизаторе RTA при пересылке маршрутов в AS1 используются различные метрики. Меньшая метрика присваивается маршруту, который использует для передачи трафика канал NY. Таким образом, входящий трафик AS3, поступающий от AS1, всегда будет передаваться по каналу NY. В этих целях вы можете также использовать и другие атрибуты (такие как BGP-сообщества и с их помощью влиять на правила маршрутизации на удаленных узлах).

- Запретить поступление в AS3 сообщений об обновлениях BGP-маршрутов.

  Для того чтобы выполнить это условие, нужно настроить в AS3 карту маршрутов или список префиксов, которые бы блокировали поступающие сообщения об обновлении BGP-маршрутов. Провайдер (в нашем случае это AS1) по вашему требованию не будет посылать вам сообщений об обновлениях маршрутов. В любом случае вам следует позаботиться о защите собственной AS от разного рода неожиданностей. По ошибке провайдер может послать на ваш маршрутизатор все свои маршруты и тогда сложно даже предсказать, как поведет себя ваша AS.

В листинге 12.10 показана конфигурация маршрутизатора RTA с использованием маршрутизации только по умолчанию, с одним основным и одним резервным каналом.

Листинг 12.10. *Маршрутизация только по умолчанию: один основной и один резервный канал (конфигурация маршрутизатора RTA*

```
router bgp  3
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 route-map BLOCK in
 neighbor 172.16.20.1 route-map SETMETRIC1 out
 neighbor 192.68.9.2 remote-as 1
```

```
 neighbor 192.68.9.2 route-map BLOCK in
 neighbor 192.68.9.2 route-map SETMETRIC2 out
 no auto-summary

ip  route  0.0.0.0   0.0.0.0   172.16.20.1   50
ip  route  0.0.0.0   0.0.0.0   192.68.9.2   40
route-map  SETMETRIC1  permit  10
 set metric 100

route-map SETMETRIC2 permit 10
 set metric  50

route-map BLOCK deny 10
```

Как видите, здесь в AS3 используются статические маршруты, которые являются маршрутами по умолчанию в AS1. Маршруту 0/0 в направлении маршрутизатора RTD • задана дистанция 40, а маршруту к RTC — 50. Таким образом, канал NY будет работать как основной. Можно также разрешить получение AS3 одного маршрута по умолчанию от AS1 ив дальнейшем использовать его.

Карты маршрутов SETMETRIC1 и SETMETRIC2 используются здесь для установки метрик маршрутов: 50 — для маршрута на RTD и 100 — для маршрута на RTC, вынуждая пересылать трафик по каналу NY.

Карта маршрутов BLOCK применяется для блокирования всех обновлений маршрутов BGP, поступающих от AS1.

Таблица IP-маршрутов на маршрутизаторе RTA (листинг 12.11) показывает, как установлен маршрут по умолчанию. Обратите внимание, что маршруту 0/0 с дистанцией 40 отдается предпочтение перед маршрутом с дистанцией 50, а в качестве шлюза по умолчанию используется ближайший узел с адресом 192.68.9.2.

Листинг 12.11. *Таблица IP-маршрутов на маршрутизаторе RTA*

```
RTC#show ip route
Codes: C - connected, S - static, I - IGRP, R -- RIP, M -- mobile, B - BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       El - OSPF external type 1, E2 - OSPF external type 2, E – EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2,
         *- candidate  default  U   -  per-user   static   route
Gateway, of last resort is 192.68.9.2 to network 0.0.0.0
C 192.68.9.0 is directly connected, EthernetO
255.255.255.0 is subnetted, 2 subnets
C   172.16.220.0 is directly connected, Ethernetl
C   172.16.20.0 is directly connected, SerialO
S*  0.0.0.0 0.0.0.0 [40/0] via 192.68.9.2
```

В листинге 12.12 представлена таблица BGP-маршрутов на маршрутизаторе RTC, которая отражает тот факт, что AS3 всегда работает по каналу между RTA и RTD, так как этот маршрут обладает меньшей метрикой (50). Префикс 172.16.220.0/24 доступен как по IBGP, так и по EBGP. Однако в качестве наилучшего маршрута избирается IBGP-маршрут. Посмотрите: в этой таблице, чтобы попасть в сеть 172.16.220.0/24, используется следующий ближайший узел 192.68.6.1, поскольку соединение с соседним узлом RTC на маршрутизаторе RTD сконфигурировано с применением команды next-hop-self.

Листинг 12.12. *Таблица BGP-маршрутов на маршрутизаторе RTC*

```
RTC# show ip bgp
BGP table version is 11, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric      LocPrf   Weight   Path
 *> i171.16.220.0/24  192.68.6.1    50          100      0        3 i
 *                    172.16.20.2   100                  0        3 i
 *> 192.68.11.0       0.0.0.0       0                    32768    i
```

Маршрутизация по умолчанию: основной и резервный каналы плюс частичная маршрутизация

В этом разделе мы рассмотрим вопросы управления течением трафика в случае, если в вашей AS, наравне с маршрутизацией по умолчанию, допускается частичная маршрутизация, т.е. получение части маршрутов провайдера и их использование внутри AS. Частичные маршруты (partial routes) — это обычно локальные маршруты провайдера и его клиентов. На рис. 12.4 показана AS3, внутри которой взаимодействие между узлами сети организовано по протоколу IBGP, а также имеется два соединения с провайдером (AS1) в различных местах по протоколу EBGP.

В сети, представленной на рис. 12.4, следует установить следующие правила маршрутизации.

- Система AS3 может принимать только локальные маршруты от ASI и ее клиен тов, т.е. AS6. Также AS3 может получать один маршрут непосредственно из сети Internet для организации маршрута по умолчанию в сторону провайдера (AS1).
- Весь исходящий трафик в направлении AS1 и AS6 (т.е. частичные маршруты) должен передаваться по каналу SF, а в случае выхода его из строя — по резервному каналу.
- Для остального исходящего трафика, адресованного в сеть Internet, в AS3 будет по умолчанию использоваться канал NY. В случае выхода его из строя будет использоваться резервный канал.
- Для входящего трафика AS3 укажет AS1 использовать для сети 172.16.220.0/24 канал SF.
- Остальной входящий трафик будет поступать через основной канал NY.
-

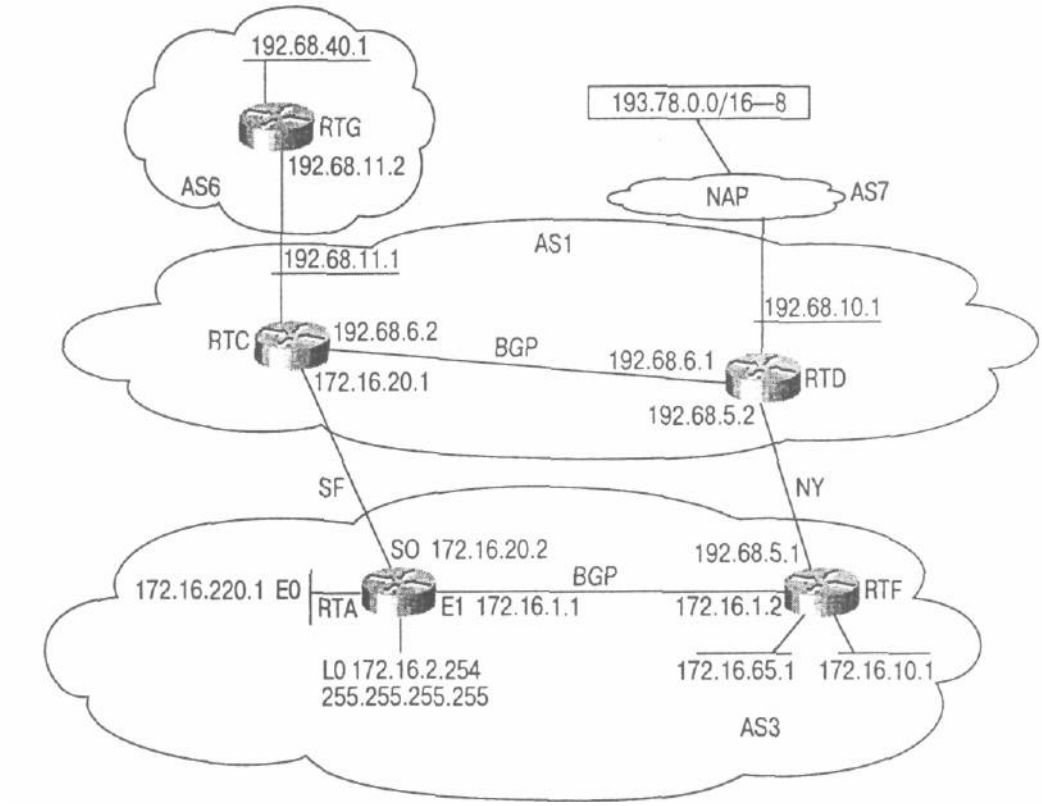


Рис. 12.4. *Маршрутизация по умолчанию с одним основным и одним резервным каналом плюс частичная маршрутизация*

В листингах 12.13 и 12.14 показаны варианты настройки маршрутизаторов RTA и RTF для работы с частичными маршрутами.

Листинг 12.13. *Маршрутизация по умолчанию: один основной и один резервный канал плюс частичная маршрутизация (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.1.2 next-hop-self
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 route-map SET_OUTBOUND_TRAFFIC in
 neighbor 172.16.20.1 route-map SET_INBOUND_TRAFFIC out
 neighbor 172.16.20.1 filter-list 10 out
 no auto-summary

ip route 0.0.0.0 0.0.0.0 193.78.0.0
ip as-path access-list 10 permit ^$
ip as-path access-list 4 permit ^1 6$
ip as-path access-list 4 permit ^1$

access-list 2 permit ip 172.16.220.0 0.0.0.255
access-list 101 permit ip 193.78.0.0 0.0.255.255 255.255.0.0 0.0.0

route-map SET_OUTBOUND__TRAFFIC permit 10
 match ip address 101
 set local-preference 200
route-map SET_OUTBOUND_TRAFFIC permit 20
 match as-path 4
 set local-preference 300

route-map SET_INBOUND_TRAFFIC permit 10
 match ip address 2
 set metric 200

route-map SET_INBOUND_TRAFFIC permit 20
 set metric 300
```

Листинг 12.14. *Маршрутизация по умолчанию: один основной и один резервный канал плюс частичная маршрутизация (конфигурация маршрутизатора RTF)*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.2.254 remote-as 3
 neighbor 172.16.2.254 next-hop-self
 neighbor 192.68.5.2 remote-as 1
neighbor 192.68.5.2 route-map SET_OUTBOUND_TRAFFIC in
 neighbor 192.68.5.2 route-map SET_INBOUND_TRAFFIC out
 neighbor 192.68.5.2 filter-list 10 out
 no auto-summary

ip route 0.0.0.0 0.0.0.0 193.78.0.0
 ip as-path access-list 10 permit ^$
 ip as-path access-list 4 permit ^1 6$
 ip as-path access-list 4 permit ^1$
access-list 101 permit ip 193.78.0.0  0.0.255.255  255.255.0.0  0.0.0.0
route-map SET_OUTBOUND TRAFFIC permit 10
 match ip address 101
 set local-preference 250

route-map SET_OUTBOUND_TRAFFIC permit 20
```

```
 match as-path 4
 set local-preference 250

route-map SET_INBOUND_TRAFFIC permit 10
 set metric 250
```

Итак, изучив пример конфигурации маршрутизатора RTA в листинге 12.13, можно сделать следующие выводы:

- Карта маршрутов SET_OUTBOUND_TRAFFIC используется в EBGP-сеансе между маршрутизатором RTA и AS1. Эта карта маршрутов помогает указать, какой тип исходящего трафика по какому каналу следует передавать. Первая запись в ней (с номером 10) разрешает передавать трафик из Internet только в сеть с адресом 193.78.0.0/16. Маршрут в эту сеть и будет использоваться в качестве маршрута по умолчанию. Ему будет присвоено значение локального предпочтения 200, которое меньше локального предпочтения 250, назначенного трафику от маршрутизатора RTF. Таким образом, весь трафик в направлении сети Internet будет по умолчанию пересылаться по каналу NY.

  Вторая запись (20) устанавливает всем префиксам, поступающим от AS1 и AS6 локальное предпочтение равное 300, т.е. большее 250, которое имеет трафик от маршрутизатора RTF. Таким образом, канал SF становится основным для трафика, предназначенного AS1 и AS6. Обратите внимание, что эта карта маршрутов разрешает поступление в AS3 только частичных маршрутов от AS1 и AS6 путем фильтрации по атрибуту AS_PATH либо (^{Л}1$) для AS 1, либо ("1 6$) для AS6.

  Вместо того чтобы последовательно указывать всех клиентов AS1, как это делается в списке разрешения доступа на основе анализа атрибута AS_PATH с номером 4, можно просто задать нормальное выражение вида ^{Л}1?[0—9]*$, с помощью которого охватываются все атрибуты AS_PATH, начинающиеся с 1 и имеющие длину 2 (т.е. в AS_PATH указывается два узла) — а это AS1 и все ее возможные клиенты. Тогда список разрешения доступа с учетом вышесказанного приобретет вид:

  ip as-path access-list 4 permit ^{A}l ?[0—9]*$ (чтобы ввести символ ?, нажмите сначала клавиши Ctrl-V).

  Будьте предельно внимательны. Если AS1, в свою очередь, соединяется с каким-либо более крупным провайдером по прямому каналу (т.е. не через NAP), то вышеприведенное нормальное выражение "захватит" и локальные маршруты этого провайдера.
- Карта маршрутов SET_INBOUND_TRAFFIC также применяется во время EBGP-сеанса с AS1. Первая запись в ней (10) вынуждает посылать префикс 172.16.220.0/24 с метрикой 200, которая меньше метрики 250, заданной такому же маршруту, но посылаемому на маршрутизатор RTF. Таким образом, трафик от AS1 в заданный пункт назначения будет передаваться по каналу SF. Все остальные маршруты будут посылаться с метрикой 300, которая, в свою очередь, больше метрики 250, посылаемой на маршрутизатор **RTF.** Тогда остальной входящий трафик будет передаваться по каналу NY.
- Список фильтрации с номером 10 не допускает, чтобы AS3 стала транзитной.
- Выражение **ip route 0.0.0.0 0.0.0.0** устанавливает маршрут по умолчанию в сеть 193.78.0.0/16.

В таблице BGP-маршрутов на маршрутизаторе RTA будут присутствовать следующие записи (см. листинг 12.15).

Листинг 12.15. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 19, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? - incomplete
 Network          Next Hop      Metric     LocPrf    Weight   Path
 *> i171.16.220.0/24  192.68.6.1    50         100       0        3 i
 *                    172.16.20.2   100                  0        3 i
 *> 192.68.11.0       0.0.0.0       0                    32768    i
```

Нетрудно заметить, что маршрутизатор RTA "видит" только сети, принадлежащие AS1 и ее клиенту — AS6 (исключая маршрут по умолчанию). Для сети 193.78.0.0./16, маршрут в которую является маршрутом по умолчанию, трафик будет пересылаться по каналу NY, так как этот маршрут имеет локальное предпочтение 250, Трафик в направлении AS1 и AS6 маршрутизатор RTA будет пересылать по каналу RTA-RTC (локальное предпочтение 300). Из таблицы IP-маршрутов на маршрутизаторе RTA, приведенной в листинге 12.16, видно, что на RTA устанавливается собственный маршрут по умолчанию в сеть 193.78.0.0/16, который пролегает через узел 172.16.1.2.

Листинг 12.16. *Таблица IP-маршрутов на маршрутизаторе RTA*

```
RTA#show ip route
Codes: C -- connected, S •- static, I •• IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPF external type 1, E2 - OSPF external type 2,
       E - EGP i - IS-IS, LI •- IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U -• per-user static route, o - ODR
Gateway of last resort is 193.78.0.0 to network 0.0.0.0
B  192.68.10.0/24 [20/0] via 172.16.20.1, 00:07:34
B  192.68.11.0/24 [20/0] via 172.16.20.1, 00:07:34
B  192.68.40.0/24 [20/0] via 172.16.20.1, 00:07:34
   172.16.0.0/16 is variably subnetted, 6 subnets, 3 masks
C    172.16.2.254/32 is directly connected, LoopbackO
C    172.16.220.0/24 is directly connected, EthernetO
C    172.16.20.0/24 is directly connected, SerialO
0    172.16.10.0/24 [110/20] via 172.16.1.2, 01:39:52, Ethernetl
C    172.16.1.0/24 is directly connected, Ethernetl
O    172.16.65.0/26 [110/20] via 172.16.1.2, 01:39:52, Ethernetl
S*  0.0.0.0/0 [1/0] via 193.78.0.0
B  193.78.0.0/16 [2QQ/0] via 172.16.1.2, 00:03:07
```

В листинге 12.17 приведена таблица BGP-маршрутов на маршрутизаторе RTD.

Листинг 12.17. *Таблица BGP-маршрутов на маршрутизаторе RTD*

```
RTC# show ip bgp
BGP table version is 14, local router ID is 192.68.10.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? - incomplete
 Network          Next Hop      Metric     LocPrf    Weight   Path
 *> 172.16.1.0/24     192.68.5.1    250                  0        3 i
 *> 172.16.10.0/24    192.68.5.1    250                  0        3 i
 *> 172.16.65.0/26    192.68.5.1    250                  0        3 i
 *>i172.16.220.0/24   192.68.6.2    200        100       0        3 i
 *                    192.68.5.1    250                  0        3 i
 *> 192.68.10.0       0.0.0.0       0                    32768    i
 *>i192.68.11.0       192.68.6.2    0          100       0        i
 *>i192.68.40.0       192.68.6.2    0          100       0        6 i
 *> 193.78.0.0/16     192.68.10.2                        0        7 8 i
```

Как видите, маршрутизатор RTD может получить доступ ко всем сетям в AS3 через прямой канал RTD-RTF, кроме сети, описываемой префиксом 172.16.220/24, доступ к которой можно получить по каналу RTC-RTA, так как ее метрика меньше (200).

Распределение нагрузки в BGP при работе по нескольким каналам

Прежде чем перейти к обсуждению вопросов распределения нагрузки, хотелось бы подчеркнуть, что в действительности все функции по коммутации пакетов, выполняемые маршрутизатором, полностью зависят от режима коммутации, который вы задаете в вашем оборудовании Cisco (CEF по пакетам, CEP по пункту назначения, быстрая коммутация, ступенчатая коммутация и т.д.). Детальное обсуждение режимов коммутации не входит в круг вопросов, затрагиваемых в этой книге, однако о них следует помнить, приступая к распределению нагрузки. Дополнительную информацию о режимах коммутации в оборудовании компании Cisco вы можете найти в документации или в книге *"Внутренняя архитектура программного обеспечения Cisco IOS" ("Inside Cisco IOS Software Architecture ")*.

Итак, давайте рассмотрим проблему распределения нагрузки. В обычных условиях, когда BGP-спикер принимает сведения об идентичных маршрутах от прилежащей AS, то выбирается и помещается в таблицу маршрутов только один лучший из них (обычно выбирается маршрут с наименьшим значением ROUTER_ID). Если в протоколе BGP разрешена работа по нескольким каналам, то в таблицу IP-маршрутов может включаться несколько маршрутов к одному пункту назначения (до шести маршрутов).

На рис. 12.5 показана реализация режима динамического распределения нагрузки для идентичных маршрутов, предлагаемая компанией Cisco.

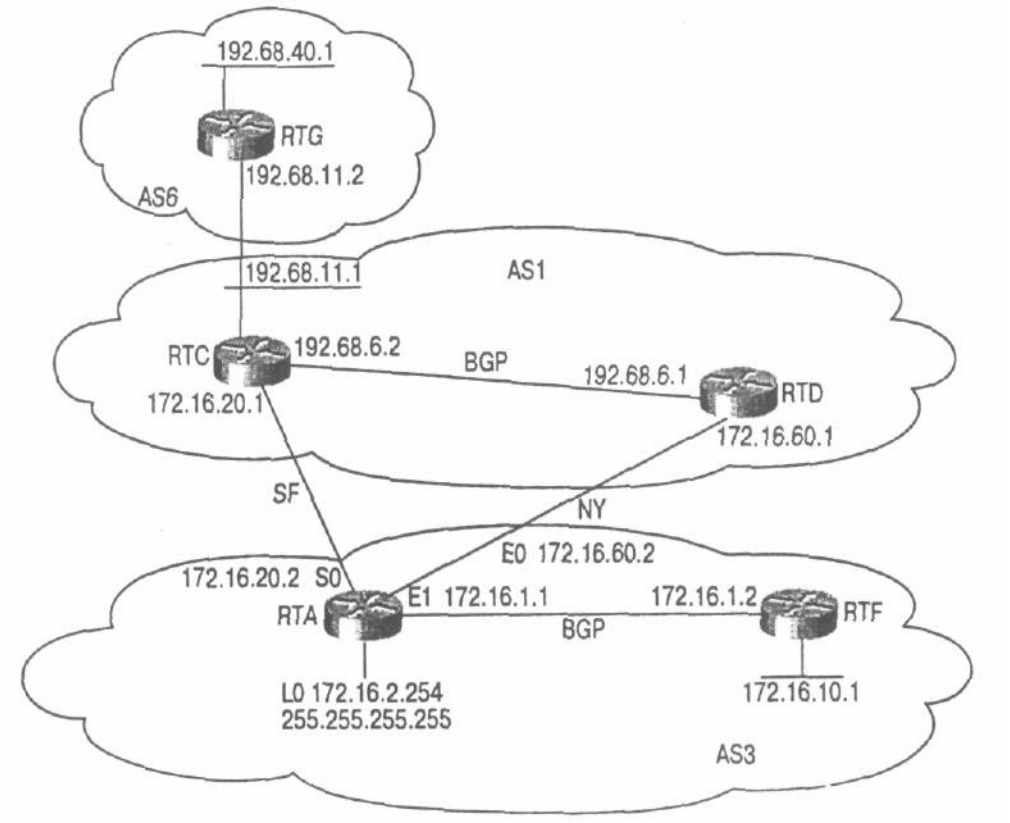


Рис. 12.5. *Автоматическое распределение нагрузки*

Маршрутизатор RTA взаимодействует по протоколу EBGP в AS1 с маршрутизаторами RTC и RTD. Он получает идентичные сообщения об обновлении маршрутов в сети 192.68.11.0/24 и 192.68.40.0/24 по двум каналам. Вы можете настроить маршрутизатор RTA с помощью команды **maximum-paths** таким образом, чтобы он выполнял автоматическое распределение нагрузки между шестью каналами. Как видно из листинга 12.18, в нашем случае с помощью команды **maximum-paths** количество каналов установлено равным 2.

Листинг 12.18. *Распределение нагрузки с помощью протокола IP (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no  synchronization
 neighbor  172.16.1.2   remote-as   3
 neighbor  172.16.1.2 update-source  LoopbackO
 neighbor  172.16.20.1  remote-as   1
 neighbor  172.16.20.1  filter-list 10  out
 neighbor  172.16.60.1  remote-as  1
 neighbor  172.16.60.1  filter-list  10  out
 maximum-paths  2
 no  auto-summary

ip as-path access-list 10 permit  A$
```

Посмотрим теперь на таблицу BGP-маршрутов маршрутизатора RTA, которая представлена в листинге 12.19. В ней вы увидите, что теперь у RTA есть идентичная информация о маршрутах в 192.68.11.0/24 и 192.68.40.0/24. Как правило, в качестве наилучшего в протоколе BGP избирается один из маршрутов. Затем этот маршрут передается в таблицу IP-маршрутов.

Листинг 12.19. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 8, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric      LocPrf    Weight   Path
 *>i172.16.10.0/24    172.16.1.2    0           100       0        i
 *>i192.68.11.0       172.16.20.1   0                     0        1 i
 *                    172.16.60.1                         0        1 i
 *> 192.68.40.0       172.16.20.1                         0        1 6 i
 *                    172.16.60.1                         0        1 6 i
```

С помощью команды **maximum-paths** BGP уведомляется о том, что для передачи трафика будут использоваться все идентичные маршруты из таблицы IP-маршрутов (до шести, в зависимости от заданной величины). Обратите внимание, что основным условием для подобных маршрутов является их принадлежность к одной AS.

В листинге 12.20 представлена таблица IP-маршрутов, в которой вы видите несколько маршрутов к одному пункту назначения. Посмотрите, как получены сведения о префиксах 192.68.11.0/24 и 192.68.40.0/24.

Листинг 12.20. *Таблица IP-маршрутов на маршрутизаторе RTA*

```
RTA#show ip route
Codes: C -- connected, S •- static, I •• IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPF external type 1, E2 - OSPF external type 2,
       E - EGP i - IS-IS, LI •- IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U -• per-user static route, o - ODR
Gateway of last resort is 193.78.0.0 to network 0.0.0.0
B  192.68.11.0/24 [20/0] via 172.16.60.1, 00:03:20
                  [20/0] via 172.16.20.1, 00:03:18
B  192.68.40.0/24 [20/0] via 172.16.60.1, 00:03:20
                  [20/0] via 172.16.20.1, 00:03:18
   172.16.0.0/16 is variably subnetted, 5 subnets, 2 masks
C    172.16.2.254/32 is directly connected, LoopbackO
C    172.16.16.0/24 is directly connected, EthernetO
C    172.16.20.0/24 is directly connected, SerialO
0    172.16.10.0/24 [110/20] via 172.16.1.2, 00:20:23, Ethernetl
C    172.16.1.0/24 is directly connected, Ethernetl
```

Если рассматривать узлы, взаимодействующие по протокоу IBGP, то маршрутизатор

RTA из множества идентичных маршрутов будет объявлять только один BGP-маршрут, а именно тот, которому задан параметр **next-hop-self**. Так как маршрутизатор RTA взаимодействует по IBGP с RTF, то он объявлять ему только один маршрут в сеть 192.68.11.0/24 и один маршрут в сеть 192.68.40.0/24 со значением NEXT_HOP – 172.16.2.254, а не внешним узлом. Это видно из таблицы BGP-маршрутов, приведенной в листинге 12.21. Внешним узлам, как обычно будет посылаться наилучший маршрут.

Листинг 12.21. *Таблица BGP-маршрутов на маршрутизаторе RTF*

```
RTC# show ip bgp
BGP table version is 56, local router ID is 172.16.10.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric     LocPrf    Weight   Path
 *> 172.16.10.0/24    0.0.0.0       0                    32768    i
 *>i192.68.11.0       172.16.2.254  0          100       0        1 i
 *>i192.68.40.0       172.16.2.254               100       0        1 6 i
```

Распределение нагрузки между двумя маршрутизаторами, **которые совместно используют несколько каналов**

Этот раздел посвящен вопросам организации распределения нагрузки между двумя маршрутизаторами, которые взаимодействуют по нескольким каналам без дублирования идентичной маршрутной информации. Мы рассмотрим ситуацию с двумя каналами.

В схеме сети, предложенной на рис. 12.6, необходимо настроить на маршрутизаторах RTA и RTC петельные интерфейсы (листинги 12.22 и 12.23) и организовать между этими маршрутизаторами сеанс связи. С помощью статических маршрутов вы сможете явно указать маршруты через физические интерфейсы к обоим петельным интерфейсам. Тогда таблица IP-маршрутов будет содержать два маршрута к ближайшему соседнему узлу, и появится возможность распределения нагрузки.

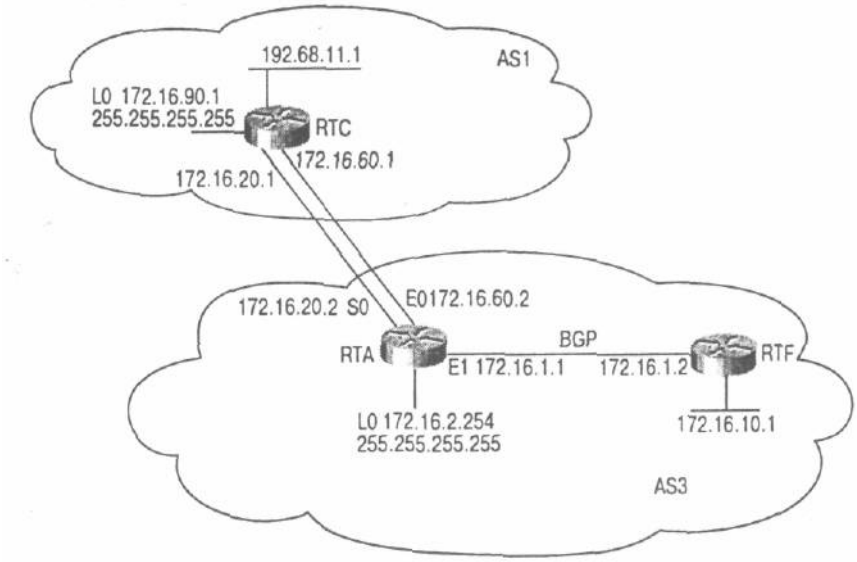


Рис. 12.6. *Распределение нагрузки между двумя маршрутизаторами, которые совместно используют два канала связи*

Листинг 12.22. *Распределение нагрузки между двумя маршрутизаторами, которые совместно используют два канала связи (конфигурация маршрутизатора RTA)*

```
interface Loopback0
 ip address 172.16.2.254 255.255.255.255
router bgp 3
 no synchronization
 neighbor 172.16.1.2 next-hop-self
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 update-source Loopback0
 neighbor 172.16.90.1 remote-as 1
```

```
 neighbor 172.16.90.1 ebgp-multlhop 2
 neighbor 172.16.90.1 update-source LoopbackO
 no auto-summary

ip route 172.16.90.1 255.255.255.255 172.16.20.1
ip route 172.16.90.1 255.255.255.255 172.16.60.1
```

Листинг 12.23. *Распредепение нагрузки между двумя маршрутизаторами, которые совместно используют два канала связи (конфигурация маршрутшатора RTC)*

```
interface LoopbackO
 ip address 172.16.90.1 255.255.255.255
router bgp 1 .
 network 192.68.11.0
 neighbor  172.16.2.254   remote-as   3
 neighbor  172.16.2.254  ebgp-multihop  2
 neighbor  172.16.2.254  update-source  LoopbackO
 no  auto-summary

ip route  172.16.2.254   255.255.255.255   172.16.20.2
ip  route  172.16.2.254   255.255.255.255   172.16.60.2
```

В листинге 12.24 представлено, как теперь маршрутизатор RTA получает от RTC сведения о BGP-маршрутах. Это происходит через узел с адресом 172.16.90.1 (адрес петельного интерфейса). Обратите внимание, что адрес источника update-source сконфигурирован как параметр команды neighbor. Так устанавливается IP-адрес источника в TCP-соединении с заданным интерфейсом. Если он не определен, то для достижения взаимодействующего узла будет использоваться выходной интерфейс и узел разорвет соединение.

Листинг 12.24. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 4, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? - incomplete
 Network              Next Hop      Metric      LocPrf    Weight   Path
 *>i172.16.10.0/24    172.16.1.2    0           100       0        i
 *>i192.68.11.0       172.16.90.1   0                     0        1 i
```

В таблице маршрутов на маршрутизаторе RTA два статических маршрута позволяют вести работу по двум каналам для обеспечения доступа на узел 172.16.90.1 (листинг 12.25). Так что маршрутизатор будет распределять нагрузку между двумя каналами.

Листинг 12.25. *Таблица IP-маршрутов на маршрутизаторе RTA*

```
RTAftshow ip route
Codes: C - connected, S - static, I - IGRP, R -- RIP, M -- mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPF external type 1, E2 - OSPF external type 2,
       E - EGP i •- IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR
Gateway of last resort is not set
B  192.68.11.0/24 [20/0] via 172.16.90.1, 00:00:41
   172.16.0.0/16 is variably subnetted, 6 subnets, 2 masks
C     172.16.2.254/32 is directly connected, LoopbackO
C     172.16.60.0/24 is directly connected, EthernetO
C     172.16.20.0/24 is directly connected, SerialO
0     172.16.10.0/24 [110/20] via 172.16.1.2, 02:17:34, Ethernetl
C     172.16.1.0/24 is directly connected, Ethernetl
S     172.16.90.1/32   [1/0] via  172.16.20.1
                       [1/0] via  172.16.60.1
```

Отметим, что все рассмотренные варианты распределения нагрузки действительны только для исходящего трафика. Подобным образом нужно настроить маршрутизатор и с другой стороны. Мы обсудим эти вопросы в последующих разделах.

Подключение к различным провайдерам по нескольким каналам

Для случая, когда один клиент подключается к нескольким провайдерам, мы рассмотрим вариант маршрутизации, где сведения об обновлении маршрутов распространяются как с использованием маршрутизации по умолчанию, так и с использованием частичной и полной маршрутизации.

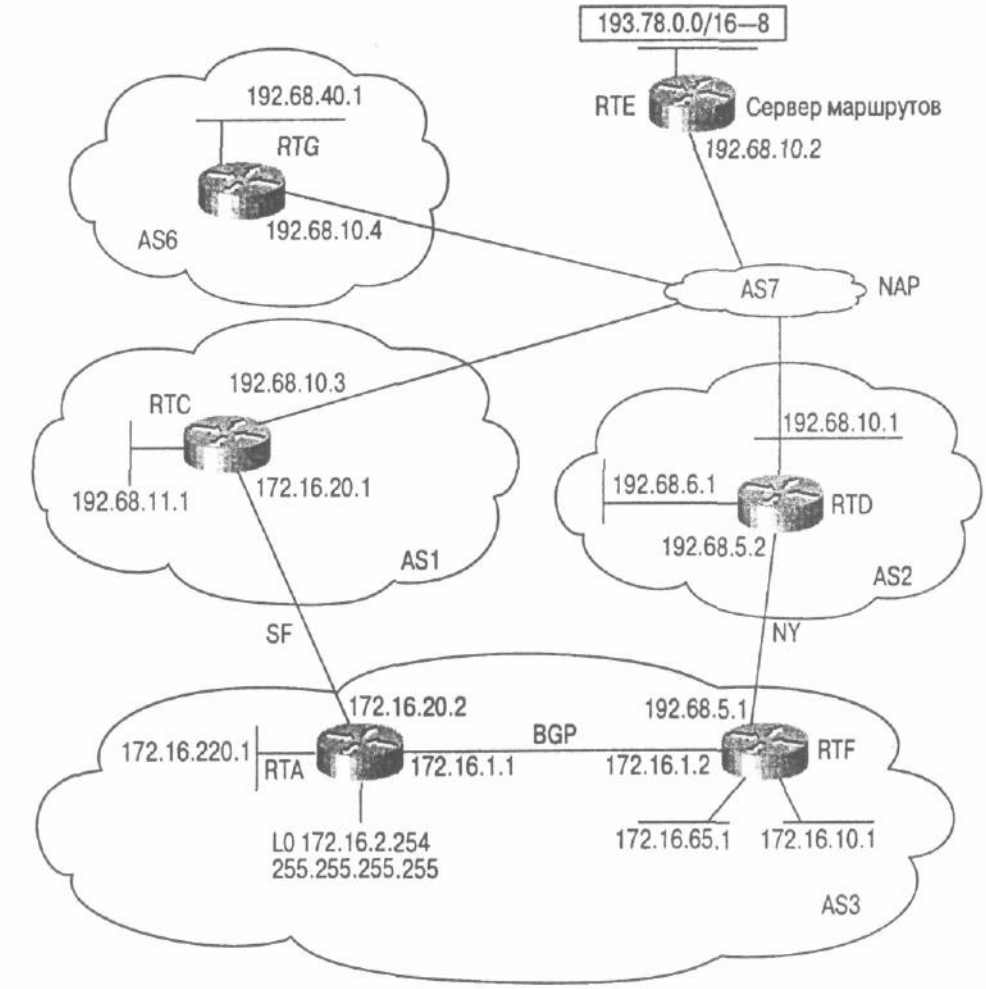


Рис. 12.7. *Подключение к нескольким провайдерам (маршрутизация по умолчанию, основной и резервный каналы, полная и частичная маршрут изация)*

На рис. 12.7 показана AS3, подключенная к двум провайдерам — AS1 и AS2. Они, в свою очередь, обмениваются маршрутной информацией и передают трафик в AS6, а также друг другу через точку обмена трафиком. Автономные системы AS6, AS2 и AS1 взаимодействуют с маршрутизатором RTE, который действует в этой схеме как сервер маршрутов, пропускающий только маршрутную информацию между этими тремя AS. В этом случае нужно установить следующие правила маршрутизации.

- Вся информация о локальных маршрутах от AS1 должна поступать в AS3 только через канал SF. Вся остальная маршрутная информация из сети Internet должна приниматься через канал NY, который является основным.
- В AS3 информация о маршруте по умолчанию от AS1 будет приниматься толь ко в случае выхода из строя канала NY.

В AS3 сеть 172.16.220.0/24 из внешнего мира может быть доступна по каналу SF, а сети 172.16.10.0/24 и 172.16.65.0/26 - по каналу NY.

- Система AS3 не может быть транзитной для AS1 и AS2, т.е. ни при каких об стоятельствах AS1 не должна использовать AS3 для того, чтобы попасть в AS2.

В листингах 12.26—12.29 показано, как реализовать эту схему маршрутизации. В листинге 12.26 представлена конфигурация маршрутизатора RTA.

Листинг 12.26. *Подключение к нескольким провайдерам (конфигурация маршрутизатора RTA): маршрутизация по умолчанию, основной и резервный каналы, полная и частичная маршрутизация*

```
router bgp  3
 no  synchronization
 network  172.16.1.0 mask  255.255.255.0
 network   172.16.10.0  mask  255.255.255.0
 network   172.16.65.0 mask  255.255.255.192
 network  172.16.220.0  mask  255.255.255.0
 neighbor  172.16.1.2   remote-as   3
 neighbor  172.16.1.2  update-source  Loopback0
 neighbor  172.16.1.2  next-hop-self
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1  route-map ACCEPT_LOCAL  in
 neighbor  172.16.20.1   route-map   PREPEND_PATH  out
 no  auto-summary

ip  as-path access-list  1  permit   ^1 ?[0-9]*$
ip  as-path access-list  2  permit  ^$

access-list  1  permit  172.16.65.0   0.0.0.63
access-list  1  permit  172.16.10.0   0.0.0.255

route-map  PREPEND_PATH permit   10
 match ip address  1
 set  as-path prepend  3

route-map  PREPEND_PATH permit 20
 match  as-path  2

route-map ACCEPT_LOCAL permit  10
 match as-path  1
```

На маршрутизаторе RTA используется карта маршрутов ACCEPT_LOCAL, которая позволяет принимать информацию о частичных маршрутах от AS1. Как видите, в карте маршрутов делается попытка выделить все маршруты, которые соответствуют выражению *^1 ?[0-9]*$,* что, как уже отмечалось, позволяет принимать информацию о локальных маршрутах от AS1 и ее клиентов,

Для взаимодействия с маршрутизатором RTC на RTA используется карта маршрутов PREPEND_PATH, с помощью которой добавляются номера AS во все префиксы, передаваемые по каналу NY (такие как 172.16.10.0/24 и 172.16.65.0/26). Таким образом, эти префиксы будут иметь меньшую длину AS_PATH при передаче по каналу NY. Однако добавление значений в атрибут AS_PATH необходимо скоординировать с провайдером. Дело в том, что провайдер может устанавливать правила, согласно которым ваш префикс будет связан с маршрутной информацией. Например, в AS1 может быть установлено правило объявлять сведения о маршрутах в AS3 в NAP с помощью выражения $^{A}3$, которое представляет AS_PATH, начинающийся с номера 3 и оканчивающийся номером 3. Так, если AS3 начнет посылать свой AS_PATH в виде 3 3 3 3, то провайдер будет отвергать маршруты с таким AS_PATH как не соответствующие правилам маршрутизации.

Обратите внимание, что в записи 20 в карте маршрутов PREPEND_PATH разрешается объявлять только локальные маршруты в AS3. Это достигается путем сопоставления локальных префиксов с пустым AS-PATH, который описывается нормальным

выражением ^$.

Точно так же в листинге 12.27 настраивается маршрутизатор RTF. Он должен анонсировать префиксы, поступающие через канал SF, по каналу NY с внесением дополнений в атрибут AS_PATH. Таким образом, трафик, входящий в эти сети, будет поступать по каналу SF (так как длина AS_PATH в этом случае будет меньшей).

Листинг 12.27. *Подключение к нескольким провайдерам (конфигурация маршрутизатора RTF): маршрутизация по умолчанию, основной и резервный каналы, полная и частичная маршрутизация*

```
router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.2.254 remote-as 3
 neighbor 172.16.2.254 next-hop-self
 neighbor 192.68.5.2 remote-as 2
 neighbor 192.68.5.2 route-map PREPEND_PATH out
 no auto-summary

ip as-path access-list 2 permit ^$
access-list 1 permit 172.16.220.0 0.0.0.255

route-map PREPEND_PATH permit 10
 match ip address 1
 set as-path prepend 3

route-map PREPEND_PATH permit 20
 match as-path 2
```

Как видите, на маршрутизаторе RTF от AS2 принимаются все маршруты, а объявляются только локальные маршруты (^S) с дополнительным номером AS в AS_PATH для маршрута 172.16.220.0/24.

В листинге 12.28 приведена таблица BGP-маршрутов на маршрутизаторе RTA.

Листинг 12.28. *Таблица BGP-маршрутов на маршрутизаторе RTA*

```
RTC# show ip bgp
BGP table version is 13, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? - incomplete
 Network          Next Hop      Metric     LocPrf   Weight  Path
 *> 0.0.0.0       172.16.20.1                       0       1 i
 *> 172.16.1.0/24 0.0.0.0       0                  32768   i
 *i               172.16.1.2    0          100     0       i
 *> 172.16.10.0/24 172.16.1.2   20                 32768   i
 *i               172.16.1.2    0          100     0       i
 *> 172.16.65.0/26 172.16.1.2   20                 32768   i
 *i               172.16.1.2    0          100     0       i
 *>i172.16.220.0/24 0.0.0.0     0                  32768   i
 *i               172.16.1.2    20         100     0       i
 *>i192.68.6.10   172.16.1.2    0          100     0       2 i
 *> 192.68.11.0   172.16.20.1   0                  0       1 i
 *>i193.78.0.0/16 172.16.1.2                100     0       2 7 8 i
```

Итак, маршрутизатор RTA получает маршрут по умолчанию (0.0.0.0) от RTC. Кроме этого он получает также сведения о локальных маршрутах AS1 (таких как 192.68.11.0/24) и может посылать трафик в эти сети непосредственно по каналу SF. Для всех остальных маршрутов маршрутизатор RTA будет использовать канал NY.

С другой стороны, входящий трафик должен следовать по кратчайшему пути. Из таблицы BGP-маршрутов на маршрутизаторе RTG, приведенной в листинге 12.29, показана

внешняя AS, которая работает через NAP (такая как AS6) и должна передать трафик в сети AS3.

Листинг 12.29. *Таблица BGP-маршрутов на маршрутизаторе RTG*

```
RTC# show ip bgp
BGP table version is 9, local router ID is 192.68.40.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric      LocPrf    Weight   Path
 *> 172.16.1.0/24     192.68.10.1                         0        7 2 3 i
 *> 172.16.10.0/24    192.68.10.1                         0        7 2 3 i
 *> 172.16.65.0/26    192.68.10.1                         0        7 2 3 i
 *> 172.16.220.0/24   192.68.10.3                         0        7 1 3 i
 *> 192.68.6.0        192.68.10.1                         0        7 2 i
 *? 192.68.11.0       192.68.10.3                         0        7 1 i
 *> 192.68.40.0       0.0.0.0       0                     32768    i
 *> 193.78.0.0/16     192.68.10.2                         0        7 8 i
```

Обратите внимание, что префиксы 172.16.10.0/24 и 172.16.65.0/26 могут быть доступны по каналу NY (маршрут 7 2 3). Префикс 172.16.220.0/24 доступен по каналу SF (маршрут 7 1 3).

Клиенты одного провайдера с резервным каналом между ними

По обоюдной договоренности клиенты одного провайдера могут организовать между собой соединение в частном порядке. Такой частный канал будет использоваться в качестве резервного в том случае, если соединение с провайдером у одного из клиентов выйдет из строя. В этом разделе мы рассматриваем случай, когда частный канал используется как основной для обмена трафиком между AS двух клиентов и как резервный для работы в сети Internet в случае выхода из строя канала с провайдером у одного из клиентов.

В этом примере мы немного изменим роли. На рис. 12.8 показана AS3, которая является провайдером и предоставляет доступ в Internet двум клиентам — AS1 и AS2. Эти клиенты по взаимной договоренности решили использовать каналы друг друга для доступа в сеть Internet в том случае, если один из каналов выходит из строя. В нормальных условиях, когда все каналы исправны, частный канал используется только для обмена трафиком между AS1 и AS2, весь остальной трафик в и из Internet передается по каналам с провайдером AS3.

Предположим также, что и AS1 и AS2 получают полные сведения о маршрутах из сети Internet, т.е. выполняется маршрутизация по полной схеме. Тогда AS1 и AS2 должны объявлять свои маршруты в AS3, так как для обеспечения работы по резервной схеме в AS3 должны быть сведения о том, как попасть в сети AS1 через AS2, и наоборот. Обычно это происходит автоматически, согласно схеме функционирования протокола BOP. Согласно правилу кратчайшего маршрута, AS1 и AS2 всегда смогут получить доступ к сетям друг друга через частный канал между ними. Чтобы попрактиковаться в установке правил маршрутизации, попытаемся решить эту проблему с использованием атрибута LOCAL_PREF. В листинге 12.30 представлена конфигурация маршрутизатора RTC. Подобным образом следует настроить и маршрутизатор RTD.

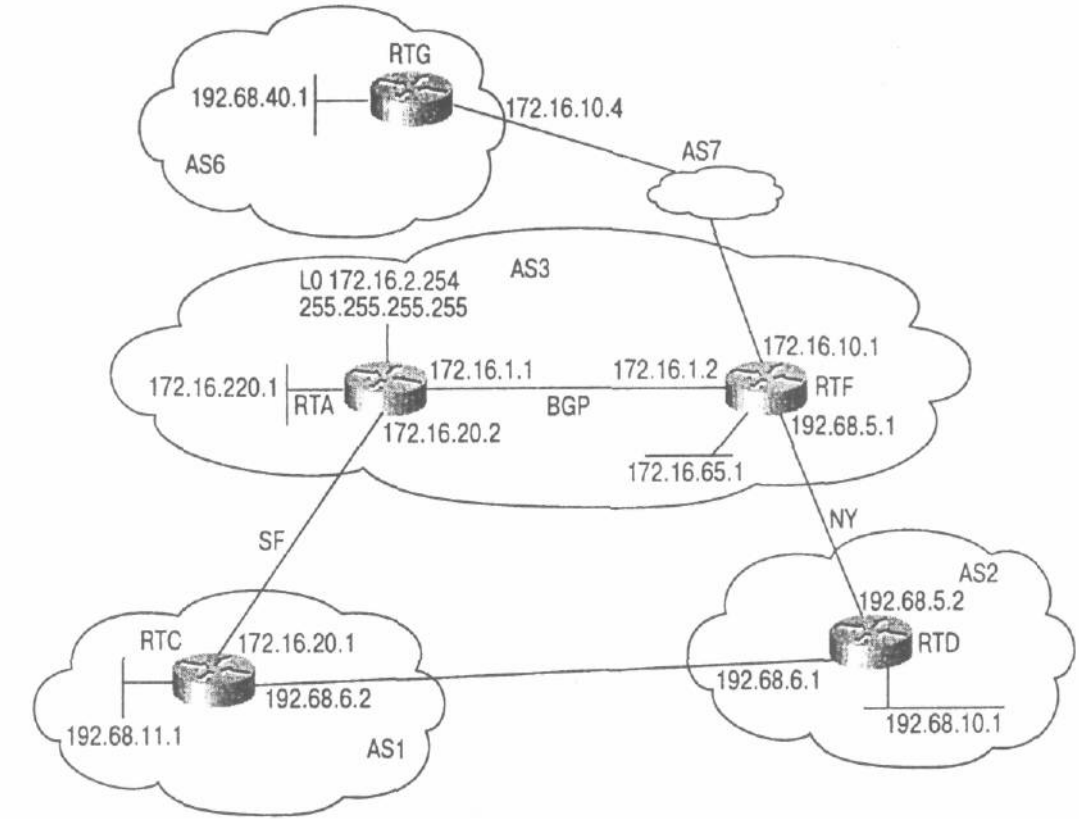


Рис. 12.8. *Использование частного канала между двумя клиентами в качестве основного*

Листинг 12.30. *Использование частного канала в качестве основного (конфигурация маршрутизатора RTC)*

```
router bgp 1
 network 192.68.11.0 neighbor 172.16.20.2 remote-as 3
 neighbor 172.16.20.2 route-map PREF_FROM_AS3 in
 neighbor 192.68.6.1 remote-as 2
 neighbor 192.68.6.1 route-map PREF_FROM_AS2 in
 no  auto-summary

ip as-path  access-list  1  permit _2_

route-map  PREF_FROM_AS3 permit  10
 match  as-path   1
 set  local-preference  100

route-map  PREF__FROM_AS3  permit 20
 set  local-preference  300

route-map  PREF__FROM_AS2  permit  10
 set  local-preference  200
```

В примере конфигурации, представленном в листинге 12.30, описывается карта маршрутов PREF_FROM_AS2, с помощью которой все маршрутам, поступающим от AS2 присваивается локальное предпочтение 200. Другая карта маршрутов — PREF_FROM_AS3 присваивает всем маршрутам от AS3 локальное предпочтение 300. Вы уже, вероятно, обратили внимание на нормальное выражение вида _2_, которое указывает на маршруты, прошедшие через AS2. Согласно этой конфигурации, сведения обо всех сетях, поступающие от AS2 или ее клиентов, будут передаваться по частному каналу. Все остальные сведения о маршрутах будут передаваться через провайдера AS3. В листинге 12.31 представлена таблица BGP-маршрутов на маршрутизаторе RTC.

Листинг 12.31. *Использование частного канала в качестве основного (таблица BGP-маршрутов на маршрутизаторе RTC)*

```
RTC# show ip bgp
BGP table version is 11, local router ID is 192.68.11.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric     LocPrf   Weight   Path
 *> 172.16.1.0/24     172.16.20.2   0          300      0        3 i
 *                    192.68.6.1                200      0        2 3 i
 *> 172.16.10.0/24    172.16.20.2   20         300      0        3 i
 *                    192.68.6.1                200      0        2 3 i
 *> 172.16.65.0/26    172.16.20.2   20         300      0        3 i
 *                    192.68.6.1                200      0        2 3 i
 *>i172.16.220.0/24   172.16.20.2   0          300      0        3 i
 *                    192.68.6.1                200      0        2 3 i
 * 192.68.10.0        172.16.20.2               100      0        3 2 i
 *>                   192.68.6.1    0           200      0        2 i
 *> 192.68.11.0       0.0.0.0       0                    32768    i
 *> 192.68.40.0       172.16.20.2               300      0        3 6 i
 *                    192.68.6.1                200      0        2 3 6 i
```

Как видите, префиксу 192.68.10.0/24, поступающему от AS3, установлено локальное предпочтение 100, так как его атрибут AS_PATH 3 2 содержит номер автономной системы 2. Всем остальным маршрутам, сведения о которых поступают от AS3, устанавливается локальное предпочтение равное 300.

Клиенты различных провайдеров с резервным каналом между собой

При добавлении или удалении клиентов провайдеры стараются изменять как можно меньше параметров, так как множественные изменения конфигурации могут негативно сказаться на масштабируемости их сети. Каждый раз, когда добавляется или удаляется клиент, провайдеру необходимо соответствующим образом подстраивать свои правила маршрутизации согласно требованиям клиента. В последующих примерах вы увидите, как в AS можно использовать атрибут COMMUNITY или другие методы управления маршрутами для того, чтобы новый клиент мог динамически согласовывать свои правила маршрутизации с правилами, установленными провайдером.

Управление маршрутами с помощью атрибута COMMUNITY

На рис. 12.9 клиент AS1 подключается к провайдеру AS4. А клиент AS3 получает доступ в Internet через провайдера AS3. При этом между клиентами AS1 и AS2 имеется частный канал, который используется для обмена трафиком между этими AS. Весь остальной трафик должен передаваться по каналам с соответствующими провайдерами: AS1 -- через AS4, а AS2 — через AS3. В том случае, если частный канал выходит из строя, клиенты должны иметь возможность обмениваться между собой трафиком через сеть Internet. Если у одного из клиентов выходит из строя канал связи с провайдером, то он может через частный канал воспользоваться соединением с Internet другого клиента.

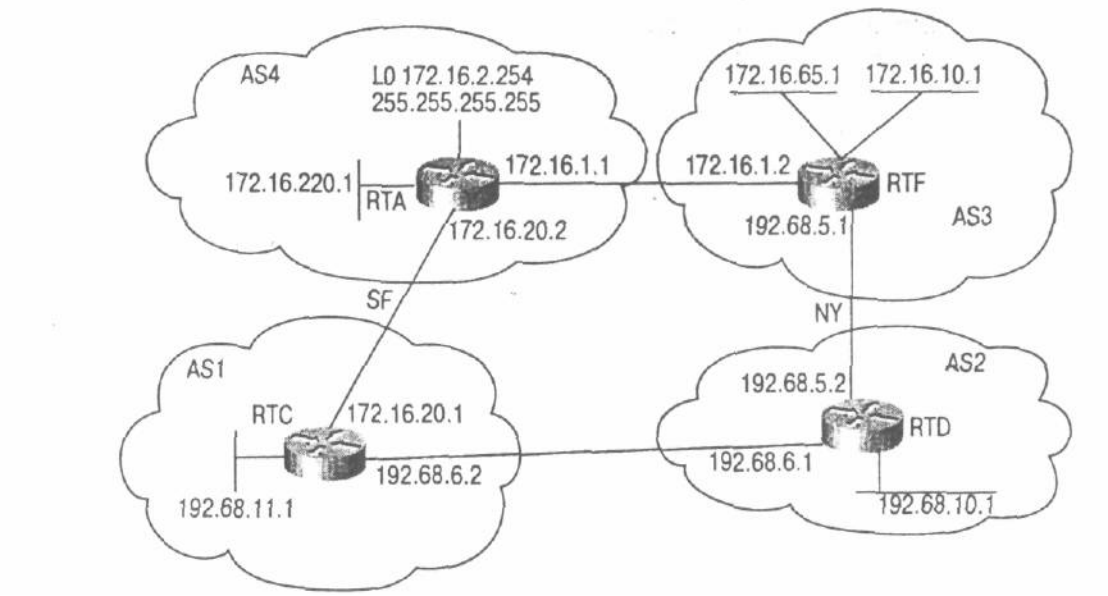


Рис. 12.9. *Несколько AS, подключенных к различным провайдерам*

В листингах с 12.32 по 12.34 показаны соответствующие конфигурации маршрутизаторов RTA, RTC и RTF. Маршрутизатор RTD должен быть зеркальным отражением RTC.

Листинг 12.32. *Использование частного канала между двумя AS, подключенных к различным провайдерам, с применением атрибута COMMUNITY (конфигурация маршрутизатора RTA)*

```
router bgp 4
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.1.2 remote-as 3
 neighbor 172.16.1.2 route-map CHECK_COMMOKITY in
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 route-map CHECK_COMMUNITY in
 no auto-summary

ip community-list 2 permit 4:40
ip community-list 3 permit 4:60

route-map CHECK_COMMUNITY permit 10
 match  community 2
 set  local-preference  40

route-map  CKECK__COMMUNITY permit  20
 match  community  3
 set  local-preference  60

route-map  CHECK__COMMUNITY permit  30
 set  local-preference  100
```

Итак, как видите, на маршрутизаторе RTA создана карта маршрутов CHECK_ COMMUNITY. Согласно этой карте маршрутов, атрибут COMMINUTY проверяется на соответствие списку (команда **match community)** значений, представленных с помощью команды ip **community-list.** В карте маршрутов описывается следующее.

- Запись 10 — для всех маршрутов со значением COMMUNITY 4:40 установить локальное предпочтение 40.
- Запись 20 — для всех маршрутов со значением COMMUNITY 4:60 установить локальное предпочтение 60.
- Запись 30 — для всех остальных маршрутов установить по умолчанию локаль ное предпочтение 100.

Листинг 12.33. *Использование частного канала между двумя AS, подключенных к различным провайдерам, с применением атрибута COMMUNITY (конфигурация маршрутизатора RTC)*

```
router bgp 1
 network  192.68.11.0
 neighbor  172.16.20.2   remote-as   4
 neighbor  172.16.20.2   send-community
 neighbor  172.16.20.2   route-map  set community out
 neighbor  172.16.20.2   filter-list  10  out
 neighbor  192.68.6.1  remote-as  2
 no  auto-summary

ip  as-path access-list  2  permit _2_

ip as-path access-list 10 permit  ^$
ip as-path access-list 10 permit  ^$

route-map  setcommunity permit  10
 match  as-path  2
set  community  4:40

route-map  setcommunity permit 20
```

Давайте рассмотрим карту маршрутов setcommunity, которая создана для обслуживания маршрутов, поступающих от AS4. В этой карте маршрутов устанавливается следующее.

- Запись 10 — для всех маршрутов, прошедших через AS2 (_2_) установить атрибут COMMUNITY 4:40.
- Запись 20 — все остальные маршруты должны проходить без изменений и не содержать атрибута COMMUNITY.

На маршрутизаторе RTC используется также список фильтров filter-list out, благодаря которому AS4 не получает сведений об AS3 через AS1. Этот фильтр разрешает только обмен маршрутами между AS1 и AS2. Если же канал между AS4 и AS3 выходит из строя, то AS4 не сможет использовать AS1, чтобы отправить информацию о маршрутах в AS3.

Листинг 12.34. *Использование частного канала между двумя AS, подключенных к различным провайдерам, с применением атрибута COMMUNITY (конфигурация маршрутизатора RTF)*

```
router bgp  3
 network   172.16.10.0 mask  255.255.255.0
 network  172.16.65.0  mask  255.255.255.192
 neighbor  172.16.1.1   remote-as   4
 neighbor  172.16.1.1   send-community
 neighbor  172.16.1.1  route-map  setcommunity out
 neighbor  192.68.5.2   remote-as  2
 no auto-summary

route-map  setcommunity permit  10
 set  community  4:60
```

На маршрутизаторе RTF карта маршрутов **setcommunnity,** которая используется совместно с BGP-командой **neighbor,** устанавливает всем маршрутам в направлении AS4 атрибут COMMUNITY 4:60.

Посмотрим, чего же мы добились. Рассмотрим таблицу BGP-маршрутов на маршрутизаторе RTA, приведенную в листинге 12.35.

Листинг 12.35. *Использование частного канала между двумя AS, подключенных*

к различным провайдерам, с применением атрибута COMMUNITY (таблица BGP-маршрутов на маршрутизаторе RTA)

```
RTC# show ip bgp
BGP table version is 7, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network                Next Hop      Metric      LocPrf    Weight   Path
 *> 172.16.10.0/24      172.16.1.2    0           60        0        3 i
 *> 172.16.65.0/26      172.16.1.2    0           60        0        3 i
 *> 172.16.220.0/24     0.0.0.0       0                     32768    i
 *> 192.68.10.0         172.16.1.2                60        0        3 2 i
 *                      172.16.20.1               40        0        1 2 i
 *> 192.68.11.0         172.16.20.1   0           100       0        1 i
```

Как видите, маршрутизатор RTA динамически установил локальное предпочтение равное 60 для всех маршрутов провайдера AS3. Все маршруты, поступающие от AS2 через AS1, имеют локальное предпочтение 40, а все локальные маршруты AS1 -- 100.

Все маршруты, сгенерированные AS1 (т.е. локальные маршруты клиента), такие как **192.68.11.0/24,** AS4 будет направлять прямо в AS1. Маршруты, принадлежащие AS2 (транзитные клиентские маршруты), такие как 192.68.10.0/24, будут передаваться AS4 другому провайдеру, т.е. на AS3. Все остальные маршруты, объявляемые провайдером (маршруты провайдера) в AS4, также будут направляться провайдеру AS3.

Управление маршрутами с помощью атрибута AS_PATH

В качестве альтернативы управлению маршрутами с помощью атрибута COMMUNITY вы можете использовать для этой же цели атрибут AS_PATH (см. рис. 12.9). Поскольку LOCAL_PREF является нетранзитивным атрибутом, управляя маршрутами с помощью AS_PATH, можно повлиять на процесс принятия решений о выборе маршрута не только в двух взаимодействующих AS, но и в других AS.

В листинге 12.36 показано, как маршрутизатор RTC будет добавлять во все маршруты, полученные от AS2, дополнительную запись об AS в обновления маршрутов, посылаемые на AS4, Система AS4, анализируя атрибут AS_PATH, обнаружит обновления маршрутов с более длинным AS_PATH через AS1 и будет пересылать их далее на AS3.

Листинг 12.36. *Управление маршрутами с помощью атрибута AS_PATH (конфигурация маршрутизатора RTC)*

```
router bgp  1
 network  192.68.11.0
 neighbor  172.16.20.2  remote-as   4
 neighbor  172.16.20.2  route-map  setpath  out
 neighbor  172.16.20.2  filter-list  10  out
 neighbor  192.68.6,1  remote-as  2
 no  auto-summary

ip  as-path  access-list 2  permit _2_

ip as-path access-list  10 permit  ^$
ip  as-path access-list  10  permit  ^2$

route-map setpath permit 10
 match  as-path  2
 set  as-path prepend  1

route-map  setpath permit 20
```

Как видите, маршрутизатор RTC добавил еще один номер AS 1 в обновления маршрутов, посылаемые на RTA. В листинге 12.37 показано, как будет выглядеть после этих манипуляций таблица BGP-маршрутов на маршрутизаторе RTA.

Листинг 12.37. *Управление маршрутами с помощью атрибута AS.PATH : (таблица BGP-маршрутов на маршрутизаторе RTA)*

```
RTC# show ip bgp
BGP table version is 9, local router ID is 172.16.2.254
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network                Next Hop      Metric      LocPrf    Weight   Path
 *> 172.16.10.0/24      172.16.1.2    0                     0        3 i
 *> 172.16.65.0/26      172.16.1.2    0                     0        3 i
 *> 172.16.220.0/24     0.0.0.0       0                     32768    i
 *> 192.68.10.0         172.16.1.2                          0        3 2 i
 *                      172.16.20.1                 100     0        1 1 2 i
 *> 192.68.11.0         172.16.20.1   0             100     0        1 i
```

Обратите внимание: теперь, чтобы попасть в сеть 192.68.10.0/24, на маршрутизаторе RTA будет выбираться маршрут через AS3. Вам следует позаботиться о том, чтобы . провайдер в AS1 не использовал списков разрешения доступа на основе анализа атрибута AS_PATH и принимал сведения о маршрутах только от вашей AS в форме ^{Л}1$ или ^{Л}1 2$. В противном случае добавление номеров AS в AS_PATH может привести к тому, что провайдер отфильтрует ваши маршруты. Еще раз подчеркиваем: всегда согласовывайте подобные действия со своим провайдером.

Установка маршрутов по умолчанию

В последующих примерах мы покажем, как граничные маршрутизаторы преобразуют маршруты по умолчанию внутри вашей AS. Эти маршруты затем используются при организации работы по протоколу IGP. На рис. 12.10 показана следующая схема: AS3 подключена по нескольким каналам к двум провайдерам — AS1 и AS2. Между маршрутизаторами RTA и RTC, а также между RTF и RTD поддерживается протокол EBGP. Внутри AS3 маршрутизаторы RTA и RTF взаимодействуют по протоколу IBGP. Мы рассмотрим два варианта. Первый: RTA и RTF имеют между собой непосредственное физическое соединение. Второй: они работают через посредника. Последний вариант приведен, чтобы показать последствия, к которым может привести неправильная установка маршрута по умолчанию или конфликт маршрута по умолчанию с установленными правилами маршрутизации. И наконец, RTG — внутренний маршрутизатор, который поддерживает только работу по протоколу IGP. Этот маршрутизатор, чтобы получить доступ к сетям вне AS2, должен использовать маршрут по умолчанию 0/0.

Вариант с непосредственным соединением между RTA и RTF довольно прост. Здесь очень сложно сделать что-нибудь неправильно. Пока оба маршрутизатора выдают сведения об IGP-маршрутах по умолчанию, трафик, попавший на любой из этих BGP-узлов, найдет свой маршрут из AS во внешний мир. Важно отметить, что исходящий трафик, поступивший на BGP-маршрутизатор, во избежание образования петли не должен возвращаться на маршрутизаторы, которые не поддерживают BGP.

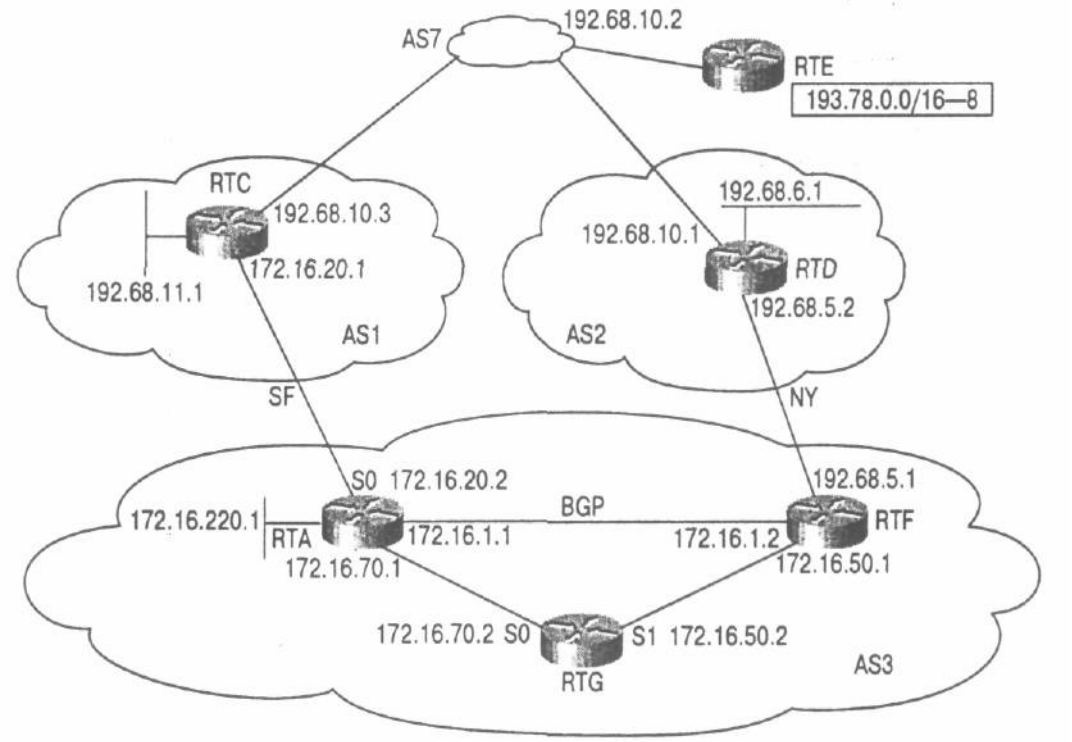


Рис. 12.10. *Установка маршрутов по умолчанию внутри AS (граничные маршрутизаторы имеют непосредственное соединение между собой)*

Если граничные маршрутизаторы работают не с полными маршрутами, они могут принимать маршрут по умолчанию от одного из провайдеров. Оба канала с провайдерами могут использоваться одновременно, или один канал используется как основной, а второй — в качестве резервного. Несмотря на используемые правила маршрутизации трафик все равно найдет путь наружу.

В примерах конфигурации, представленных в листингах 12.38 и 12.39, маршрутизаторы RTA и RTF получают полные сведения о маршрутах от своих провайдеров. Затем они посылают в свою AS сведения о маршрутах по умолчанию (сами они не получают маршрутов по умолчанию, так как имеют полные сведения о маршрутах в сети Internet). В качестве протокола IGP здесь используется протокол OSPF (варианты применения других протоколов 1GP мы рассмотрим позднее). Обратите внимание на присутствие в обоих конфигурациях субкоманды OSPF default-information originate.

Листинг 12.38. *Установка маршрутов по умолчанию внутри AS с помощью граничных маршрутизаторов, соединенных между собой физически (конфигурация маршрутизатора RTA)*

```
router ospf 10
passive-interface Serial0
 network 172.16.8.9 8.0.255.255 area 0
default-information originate always

router bgp 3
 no synchronization
 network 172.16.1.0 mask 255.255.255.0
 network 172.16.70.0 mask 255.255.255.0
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.1.2 remote-as 3
 no auto-summary

ip as-path access-list 10 permit ^$
```

Листинг 12.39. *Установка маршрутов по умолчанию внутри AS с помощью граничных маршрутизаторов, соединенных между собой физически (конфигурация маршрутизатора RTF)*

```
router ospf 10
 network 172.16.0.0 0.0.255.255 area 0
default-information originate always

router bgp  3
 no   synchronization
 network   172.16.1.0  mask  255.255.255.0
 network  172.16.50.0 mask  255.255.255.0
 neighbor  172.16.1.1  remote-as   3
 neighbor  172.16.1.1  next-hop-self
 neighbor   192.68.5.2   remote-as  2
 neighbor   192.68.5.2   filter-list  10  out
 no   auto-summary

ip as-path access-list  10 permit  ^$
```

В листинге 12.40 приведена конфигурация маршрутизатора RTG.

Листинг 12.40. *Установка маршрутов по умолчанию внутри AS с помощью граничных*

```
router  ospf  10
 network  172.16.0.0   0.0.255.255  area  0
```

Обратите внимание, что при конфигурации маршрутизаторов RTA и RTF вместе с командой router ospf в субкоманде default-information originate используется ключевое слово always. Так вы принуждаете протокол OSPF в OSPF-домене всегда рассылать маршрут по умолчанию 0/0. При этом внутренний маршрутизатор RTG, который поддерживает только OSPF, получая сведения о маршруте по умолчанию из нескольких источников, будет руководствоваться при выборе маршрута наименьшей метрикой. Маршрутизаторы, которые ближе (согласно метрикам) к RTA будут по умолчанию использовать его в качестве шлюза в Internet, а маршрутизаторы, которые расположены ближе к RTF, будут использовать маршрутизатор RTF.

В таблице IP-маршрутов на маршрутизаторе RTG, представленной в листинге 12.41, показано, как выбирается шлюз по умолчанию. Вы видите, что им оказался маршрутизатор RTA (172.16.70.1), который имеет меньшую метрику, чем RTF.

Листинг 12.41. *Установка маршрутов по умолчанию внутри AS с помощью граничных маршрутизаторов, соединенных между собой физически (таблица IP-маршрутов на маршрутизаторе RTG)*

```
RTG#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX -• EIGRP external, 0 - OSPF, IA -• OSPF inter area
       N1 - OSPF NSSA external type 1, N2 •- OSPF NSSA external type 2
       El - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, LI -- IS-IS level-1, L.2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR
Gateway of last resort is 172.16.70.1 to network 0.0.0.0
   172.16.0.0/16 is subnetted, 5 subnets
O    172.16.220.0/24 [110/74] via 172.16.70.1, 00:03:27, SerialO
C    172.16.50.0/24 is directly connected, Seriall
O    172.16.20.0/24 [110/74] via 172.16.70.1, 00:03:27, SerialO
O    172.16.1.0/24 [110/74] via 172.16.70.1, 00:03:27, SerialO
C    172.16.70.0/24 is directly connected, SerialO
0*E2 0.0.0.0/0 [110/1] via 172.16.70.1, 00:03:27, SerialO
```

Конфликт между правилами маршрутизации BGP и внутренним маршрутом по умолчанию

Всегда при работе с маршрутами по умолчанию, которые используются для доставки трафика в пункты назначения с неизвестными для данной AS маршрутами, следует соблюдать предельную осторожность, так как имеется потенциальная возможность образования петель маршрутизации. Петля маршрутизации возникает, когда маршругизатор X по умолчанию посылает данные на маршрутизатор Y, который, в свою очередь, использует маршрутизатор X, чтобы доставить их в пункт назначения. Тогда трафик будет передаваться по замкнутому кругу (петле) от X к Y и обратно.

Маршрут по умолчанию 0/0 по-разному преобразуется из BGP в IGP, в зависимости от используемого протокола ЮР. Мы рассмотрим различные варианты такого преобразования для протоколов ЮР — OSPF, RIP, EIRGP, а также IS-IS.

Ниже представлен случай, когда маршрутизаторы RTA и RTF не имеют непосредственного соединения между собой. Как вы убедитесь позднее, это намного усложняет **их** конфигурацию, и повышает возможность образования петли маршрутизации. Такая конфигурация может использоваться только в отдельных случаях.

Предположим, что в AS3 (рис. 12.11) установлены правила маршрутизации для работы по основному и запасному каналам, где канал NY — основной, а SF — резервный. Тогда маршрутизатор RTA получает сведения о IBGP-маршрутах с более высокими локальными предпочтениями, чем у EBGP-маршрутов, и будет пересыпать трафик взаимодействующему с ним по IBGP узлу RTF. Если же маршрутизатор RTG получает сведения о маршруте по умолчанию 0/0 и от RTA, и от RTF, то он будет выбирать маршрут по умолчанию через маршрутизатор RTF (как основной), в противном случае может возникнуть петля маршрутизации. Приведенная ниже последовательность событий объясняет, почему.

Шаг 1. Маршрутизатор RTG пытается переслать трафик в пункт назначения за пределами AS3.
Шаг 2. Маршрутизатор RTG выбирает маршрут по умолчанию в направлении RTA.
Шаг 3. На маршрутизаторе RTA правилами установлено, что в качестве точки выхода из AS нужно использовать маршрутизатор RTF.
Шаг 4. Чтобы переслать трафик на маршрутизатор RTF в качестве следующего промежуточного узла на RTA используется RTG (они ведь не имеют непосредственного соединения).
Шаг 5. Маршрутизатор RTG принимает трафик, адресованный за пределы AS, и посылает его обратно на RTA, т.е. образуется петля.

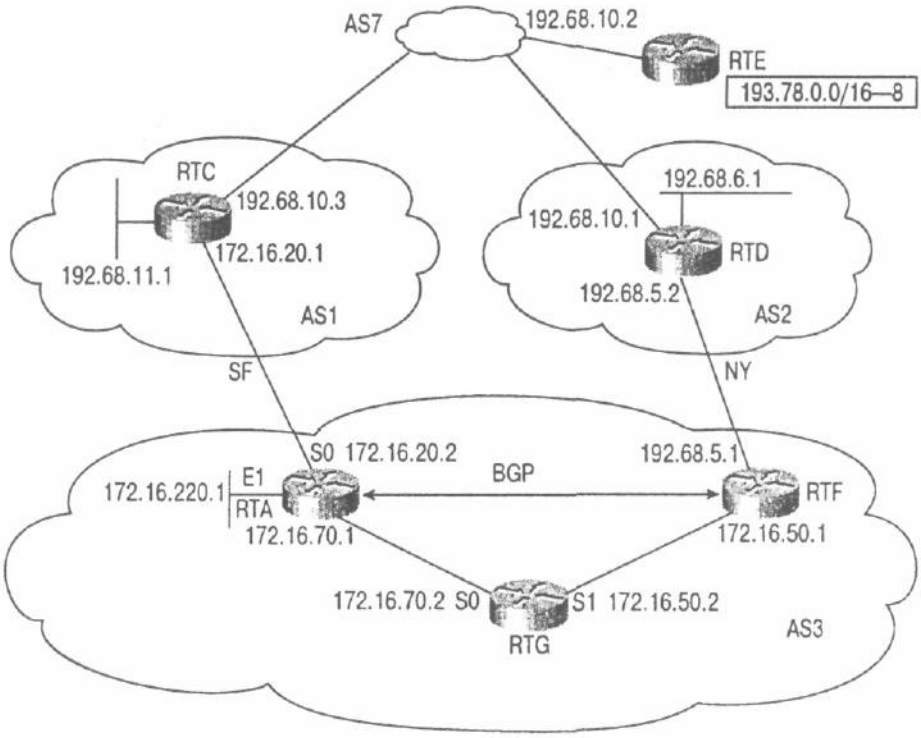


Рис. 12.11 *. Установка маршрутов по умолчанию внутри AS без непосредственного соединения между граничными маршрутизаторами*

Чтобы не допустить развития событий по такому сценарию, вы можете воспользоваться любым из таких методов.

1. Убедитесь, что маршрутизатор RTA не посылает сведения о маршруте по умолчанию 0/0 в протокол IGP до тех пор, пока не выйдет из строя основной канал. Когда все соединения работают исправно, весь трафик пересылается по умолчанию на маршрутизатор RTF и покидает пределы AS. Если же выходит из строя основной канал NY, то маршрутизатор RTA должен рассылать сведения о маршрутах по умолчанию в IGP.

   Этот метод лучше всего работает именно в среде с основным и резервным каналами. В случаях, когда точка выхода из AS не определена, очень сложно определить, какой маршрутизатор должен посылать сведения о маршруте по умолчанию. В таких ситуациях граничный маршрутизатор должен обладать возможностью посылать трафик по внешнему каналу.
2. Убедитесь в том, что граничный маршрутизатор (RTA) не посылает трафик обратно на внутренний маршрутизатор (RTG), который уже использует RTA по умолчанию. Это можно проконтролировать путем указания кратчайшего пути (на основе метрик) через BGP-маршрутизаторы. Например, установив непосредственное соединение между RTA и RTF. Если RTG использует RTA по умолчанию, то последний через физическое соединение будет пересылать весь трафик на маршрутизатор RTF.
3. Организовать обмен маршрутной информацией по протоколу IBGP с использованием полносвязной схемы между RTA, RTG и RTF. Маршрутизатор RTG будет получать сведения обо всех маршрутах по протоколу BGP.
4. Установить метрики таким образом, чтобы внутренний маршрутизатор (RTG) всегда имел маршрут через основной канал с самой низкой метрикой.

В представленном примере мы воспользовались вторым методом. Третий метод несложен, поэтому не нуждается в подробном описании. В каждом из вариантов, приведенных ниже, мы рассмотрим различные протоколы IGP и будем использовать либо первый, либо четвертый метод для решения нашей проблемы. Даже если в каждом конкретном случае вы будете использовать только один из рассмотренных выше методов, помните, что методы 1 и 4 можно использовать с любым из протоколов IGP.

Чтобы упростить проблему, предположим, что маршрутизаторы RTA и RTF динамически получают сведения о маршруте по умолчанию 0/0 от своих провайдеров (независимо от того, нужна им эта информация или нет). Итак, в последующих разделах мы увидим, как обрабатываются сведения о маршрутах по умолчанию в оборудовании компании Cisco.

Применение OSPF в качестве протокола IGP

Маршрут по умолчанию вводится в OSPF с помощью следующей команды маршрутизатора:

default-information originate *[always]* ***[metric*** *metric-value]* ***[metric-type*** *type-value]* ***[route-map*** *map-name]*

По команде **default-information originate** (без ключевого слова **always**) маршрут по умолчанию 0/0 преобразуется в OSPF-маршрут, но только в том случае, если маршрутизатор сам имеет маршрут по умолчанию. Здесь нельзя использовать ключевое слово **always**, так как при выходе из строя канала связи граничный маршрутизатор будет продолжать преобразование маршрута по умолчанию в IGP-маршрут, даже если он не сможет больше доставлять трафик по назначению. (Помните, что прямой связи между граничными маршрутизаторами нет!)

Если маршрутизаторы RTA и RTF сконфигурированы с командой **default-information originate**, то произойдет следующее,

Шаг 1. Маршрутизатор RTA получает сведения о маршруте по умолчанию по EBGP и по IBGP.

Шаг 2. Так как на RTA задано, что все должно проводиться через RTF (маршрут с более высоким локальным предпочтением), то он отдает предпочтение маршруту 0/0 по IBGP.

Шаг 3. Так как у RTA есть маршрут по умолчанию (по BGP), то он начинает преобразование его в IGP.

Итак, вы оказались в ситуации, когда оба маршрутизатора генерируют маршруты по умолчанию и вполне вероятно образование петли.

Вы можете сказать, что, поскольку канал NY является основным, то RTA не должен посылать сведения о маршрутах по умолчанию. Но подумайте, что произойдет, если выйдет из строя канал NY. При этом маршрутизатор RTF перестает объявлять маршрут по умолчанию 0/0 в IGP. И маршрутизатор RTA также не посылает никаких сведений о маршрутах по умолчанию, так что трафик не сможет покинуть пределы AS.

Решение этих проблем заключается в том, чтобы RTA и RTF объявляли маршрут по умолчанию, *только* если у них имеется этот маршрут и *только* в том случае, если он получил сведения о нем по EBGP. Когда маршрутизатор RTA обнаруживает, что маршрут по умолчанию 0/0 получен по EBGP, а не по IBGP, он сообщает, что с каналом NY что-то случилось и начинает самостоятельно посылать маршрут по умолчанию. Это можно сделать с помощью карты маршрутов совместно с командой default-information originate, как показано в листинге 12.42.

Листинг 12.42. *Использование маршрута по умблчанию только в определенных l обстоятельствах (конфигурация маршрутизатора RTA)*

```
router  ospf 10
 passive-interface   Serial0
 network  172.16.0.0  0.0.255.255 area 0
 default-information  originate  route-map  SEND_DEFAULT_IF

router bgp  3
 no  synchronization
 network  172.16.220.0  mask  255.255.255.0
 network   172.16.70.0  mask  255.255.255.0
```

```
 neighbor   172.16.20.1   remote-as   1
 neighbor  172.16.20.1   filter-list  10  out
 neighbor  172.16.50.1  remote-as   3
 neighbor  172.16.50.1   route-map  setlocalpref  in
 no  auto-summary

ip as-path  access-list  10  permit  ^$

access-list  1  permit  0.0.0.0
access-list  2  permit  172.16.20.1

route-map  setlocalpref permit  10
 set  local-preference  300

route-map  SEND_DEFAULT_IF permit  10
 match  ip  address   1
 match  ip next-hop  2
```

Обратите внимание, на карту маршрутов SEND_DEFAULT_IF, которая описывается в команде маршрутизатора default-information originate. Эта карта маршрутов выполняет проверку на соответствие условию, когда у маршрута по умолчанию 0/0 (access-list 1) в качестве следующего узла используется хост с адресом 172.16.20.1 (access-list 2). Это соответствует тому, что сведения о маршруте 0/0 получены по EBGP, а не по IBGP. Маршрутизатор RTA определяет, что канал NY вышел из строя и начинает передавать в OSPF собственный маршрут по умолчанию 0/0.

Вторая карта маршрутов setlocalpref присваивает значение локального предпочтения 300 всем IBGP-маршрутам на маршрутизаторе RTA. Таким образом, всем IBGP-маршрутам будет отдаваться предпочтение по сравнению с EBGP-маршрутами.

Как видно из листинга 12.43, на маршрутизаторе RTF также генерируется маршрут по умолчанию и посылается в OSPF только при условии, что сведения о нем получены по внешнему каналу (NEXT_HOP 192.68.5.2). В случае выхода из строя канала NY маршрутизатор RTF прекращает объявление маршрута 0/0, даже если он может получать его по IBGP от RTA.

Листинг 12.43. *Прекращение объявления маршрута по умолчанию при определенных условиях (конфигурация маршрутизатора RTF)*

```
router ospf  10
 network  172.16.0.0   0.0.255.255  area  0
 default-information originate  route-map  SEND_DEFAULT_IF

router bgp  3
 no  synchronization
 network  172.16.50.0 mask  255.255.255.0
 neighbor   172.16.70.1   remote-as   3
 neighbor  172.16.70.1  next-hop-self
 neighbor  192.68,5.2   remote-as  2
 neighbor  192.68.5.2  filter-list 10  out
 no  auto-summary

ip as-path access-list  10 permit  ^$

access-list  1  permit  0.0.0.0
access-list 2  permit  192.68.5.2

route-map  SEND_DEFAULT_IF permit  10
 match  ip  address   1
 match  ip next-hop 2
```

Как видно из листинга 12.44, на маршрутизаторе RTG поддерживается только протокол OSPF и установлен маршрут по умолчанию 0/0 для всех пунктов назначения за пределами AS.

Листинг 12.44. *Установка маршрутов по умолчанию для пунктов назначения за пределами AS (конфигурация маршрутизатора RTG)*

```
router ospf 10
 network 172.16.0.0 0.0.255.255 area 0
```

В листинге 12.45 представлена таблица IP-маршрутов на маршрутизаторе RTA. Как видите, здесь предпочтение отдается маршруту по умолчанию 0/0 через маршрутизатор RTF (взаимодействующий по IBGP узел) с NEXT_HOP 172.16.50.1. Так как NEXT_HOP отличается от адреса внешнего взаимодействующего узла— 172.16.20.1, то RTA не будет передавать сведения о маршруте по умолчанию в OSPF.

Листинг 12.45. *Таблица IP-маршрутов на маршрутизаторе RTA*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is 172.16.58.1 to network 0.0.0.0

B  192.68.6.0/24 [200/0] via 172.16.50.1, 00:03:06
B  192.68.11.0/24 [200/0] via 172.16.50.1, 00:03:06
B  193.78.0.0/16   [200/0]   via   172.16.50.1,   00:03:06
   172.16.0.0/16  subnetted,   4   subnets
C    172.16.20.0/24   is  directly connected,   SerialO
C    172.16.220.0/24   is  directly  connected,   Ethernetl
O    172.16.50.0/24    [110/164]   via   172.16.70.2,   02:17:37,   Seriall
C    172.16.70.0/24   is  directly connected,   Seriall
B* 0.0.0.0/0   [200/0]   via   172.16.50.1,   00:03:07
```

В листинге 12.46 представлена таблица IP-маршрутов на маршрутизаторе RTG. Посмотрите внимательно, как на RTG устанавливается маршрут по умолчанию в направлении RTF. Теперь все правила BGP-маршрутизации и маршруты по умолчанию в IGP синхронизированы (т.е. не конфликтуют).

Листинг 12.46. *Таблица IP-маршрутов на маршрутизаторе RTG*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway  of  last  resort   is   172.16.50.1  to  network  0.0.0.0

   172.16.0.0/16  is  subnetted,   4   subnets
O    172.16.20.0/24   [110/128]   via   172.16.70.1,   02:21:04,   SerialO
O    172.16.220.0/24   [110/74]   Via   172.16,70.1,    02:21:04,   SerialO
C    172.16.50.0/24 is directly connected, Serial1
C    172.16.70.0/24 is directly connected, SerialO
O*E2 0.0.0.0/0 [110/13 via 172.16.50.1, 00:41:26, Seriall
```

В случае выхода из строя канала NY маршрутизатор RTA будет получать сведения о маршруте 0/0 через внешний канал с промежуточным узлом 172.16.20.1 и посылать их дальше в OSPF.

Преобразование BGP-маршрута 0/0 в OSPF-маршрут с помощью команды redistribute в этом случае делать не следует.

Применение RIP в качестве протокола IGP

Реализация протокола RIP в оборудовании Cisco существенно отличается от OSPF, особенно при обработке маршрута по умолчанию вида 0/0. Маршрут по умолчанию 0/0, сведения о котором получены по BGP, автоматически переносится в RIP. Процесс RIP должен лишь назначить с помощью команды default-metric соответствующую метрику (число промежуточных узлов) маршруту по умолчанию. Предположим, что в нашем примере (см. рис. 12.11) на маршрутизаторах RTA, RTF и RTG поддерживается протокол RIP. Необходимо на маршрутизаторе RTA установить такую метрику для маршрута 0/0, который передается в RIP, чтобы внутренний маршрутизатор (RTG) всегда отдавал предпочтение маршрутизатору RTF.

С помощью конфигурации RTA, приведенной в листинге 12.47, маршруту по умолчанию 0/0 устанавливается метрика 5. Как видите, здесь не требуется преобразования маршрутов, чтобы передавать данные из BGP в RIP.

Листинг 12.47. *Применение RIP в качестве протокола IGP (конфигурация маршрутизатора RTA)*

```
router  rip
 passive-interface  Serial0
 network  172.16.0.0
 default-metric  5

router bgp  3
 no  synchronization
 network  172.16.220.0  mask  255.255.255.0
 network  172.16.70.0  mask  255.255.255.0
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1  filter-list 10  out
 neighbor  172.16.50.1   remote-as   3
 neighbor 172.16.50.1   route-map  setlocalpref  in
 no auto-summary

ip as-path  access-list  10  permit   ^$

route-map  setlocalpref permit  10
 set  local-preference   300
```

С помощью конфигурации маршрутизатора RTF, представленной в листинге 12.48, объявляется RIP-маршрут 0/0 со счетчиком промежуточных узлов установленным в 1.

Листинг 12.48. *Применение RIP в качестве протокола IGP (конфигурация маршрутизатора RTF)*

```
router  rip
 network  172.16.0.8  default-metric  1

router  bgp   3
 no  synchronization
 network  172.16.50.0 mask  255.255.255.0
 neighbor  172.16.70.1   remote-as  3
 neighbor  172.16.70.1  next-hop-self
 neighbor  192.68.5.2   remote-as   2
 neighbor  192.68.5.2   filter-list  10  out
 no  auto-summary

ip  as-path access-list  10 permit ^$
```

Конфигурация маршрутизатора RTG, приведенная в листинге 12.49, позволяет работу на нем только по протоколу RIP, а весь трафик за пределы AS3 отправляется согласно маршруту по умолчанию 0/0.

Листинг 12.49. *Применение RIP в качестве протокола IGP (конфигурация маршрутизатора RTG)*

```
router rip
 network 172.16.0.0
```

В листинге 12.50 показано, как выглядит таблица IP-маршрутов на маршрутизаторе RTG. Как видите, на RTG установлен маршрут по умолчанию в направлении RTF, так как этот маршрут обладает более низкой метрикой 1.

Листинг 12.50. *Таблица IP-маршрутов на маршрутизаторе RTG*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is 172.16.50.1 to network 0.0.0.0

   172.16.0.0/16 is subnetted, 4 subnets
R   172.16.220.0/24 [120/1] via 172.16.70.1, 00:00:03, SerialO
C   172.16.50.0/24 is directly connected, Seriall
R   172.16.20.0/24 [120/1] via 172.16.70.1, 00:00:03, SerialO
C   172.16.70.0/24 is directly connected, SerialO
R* 0.0.0.0/0 [120/1] via 172.16.50.1, 00:00:22, Seriall
```

Примечание

Если требуется обеспечить еще несколько условий, согласно которым сведения о маршруте 0/0 будут пересылаться в RIP, в качестве альтернативных средств можно использовать преобразование из BGP в RIP и карты маршрутов.

Применение EIGRP в качестве протокола IGP

Маршруты по умолчанию, сведения о которых распространяются с помощью BGP, поступают в EIGRP посредством преобразования. Метрики маршрута 0/0 при этом необходимо преобразовать в EIGRP-совместимые метрики с помощью субкоманды default-me trie.

Конфигурация маршрутизатора RTA, представленная в листинге 12.51, позволяет проводить подстановку маршрута по умолчанию с более высокой метрикой таким образом, что внутренний маршрутизатор (RTG) всегда будет получать маршрут в направлении RTF с меньшей метрикой.

Листинг 12.51. *Применение EIGRP в качестве протокола IGP (конфигурация маршрутизатора RTA)*

```
router eigrp 1
 redistribute bgp 3 route-map DEFAULT_ONLY
 passive-interface SerialO
 network 172.16.0.0
 default-metric 5 100 250 100 1500

router bgp 3
 no synchronization
 network 172.16.70.0 mask 255.255.255.0
 network 172.16.220.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.50.1 remote-as 3
 neighbor 172.16.50.1 route-map setlocalpref in
 no auto-summary
```

```
ip as-path access-list 10 permit  ^$

access-list 5 permit 0.0.0.0

route-map  setlocalpref permit 10
 set  local-preference  300

route-map DEFAULT_ONLY permit  10
 match  ip address   5
```

На маршрутизаторе RTA используется карта маршрутов DEFAULT_ONLY, с помощью которой и выделяется маршрут по умолчанию 0/0. Все остальные обновления маршрутов не будут преобразованы в EIGRP. Кроме того, на RTA с помощью субкоманды default-metric устанавливается метрика для этого маршрута.

Аналогично выполняется и конфигурация маршрутизатора RTF (листинг 12.52), где в EIGRP преобразуется только маршрут 0/0 с применением карты маршрутов DEFAULT_ONLY.

Листинг 12.52. *Применение EIGRP в качестве протокола IGP (конфигурация маршрутизатора RTF)*

```
router eigrp  1
 redistribute bgp  3   route-map DEFAULT_ONLY
 network   172.16.0.0
 default-metric  1000   100   250   100   1500

router bgp  3
 no  synchronization
 network  172.16.50.0 mask  255.255.255.0
 neighbor  172.16.70.1  remote-as  3
 neighbor  172.16.70.1  next-hop-self
 neighbor  192.68.5.2  remote-as  2
 neighbor  192.68.5.2   filter-list  10  out
 no  auto-summary

ip as-path access-list  10 permit  ^$

access-list  5 permit  0.0.0.0

route-map DEFAULT__ONLY permit  10
 match  ip  address   5
```

На маршрутизаторе RTF используется выражение default-metric 1000 100 250 100 1500, с помощью которого устанавливается метрика маршрута по умолчанию в приемлемом для протокола EIGRP виде. Обратите внимание на часть полосы пропускания (1000), выделяемую в выражении default-metric на маршрутизаторе RTF, которая является намного большей, чем полоса пропускания, выделяемая на RTA (5). Таким образом, метрика, получаемая от RTF, намного меньше, чем получаемая от RTA.

Как показано в листинге 12.53, на маршрутизаторе RTG поддерживается только протокол EIGRP, и доступ ко всем маршрутам за пределами AS3 осуществляется через маршрут по умолчанию.

Листинг 12.53. *Применение EIGRP в качестве протокола IGP (конфигурация маршрутизатора RTG)*

```
router eigrp 1
 network 172.16.0.0
```

В листинге 12.53 приведена таблица IP-маршрутов на маршрутизаторе RTG. Обратите внимание, что на RTG поддерживается маршрут по умолчанию в направлении маршрутизатора RTF.

Листинг 12.54. *Таблица IP-маршрутов на маршрутизаторе RTG*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is 172.16.50.1 to network 0.0.0.0

   172.16.0.0/16 is subnetted, 4 subnets
D    172.16.220.0/24 [90/2195456] via 172.16.70.1, 00:12:17, SerialO
C    172.16.50.0/24 is directly connected, Seriall
D    172.16.20.0/24 [90/2681856] via 172.16.70.1, 00:12:17, SerialO
C    172.16.70.0/24 is directly connected, SerialO
D*EX 0.0.0.0/0 [170/3097600] via 172.16.50.1, 00:07:40, Seriall
```

Применение IGRP в качестве протокола IGP

В протоколе IGRP не поддерживается маршрут по умолчанию 0.0.0.0. Чтобы установить в IGRP маршрут по умолчанию, необходимо на маршрутизаторах RTA и RTF задать глобальную команду ip default-network. Эта сеть по умолчанию должна переводиться под управление IGRP и устанавливать на внутренних маршрутизаторах маршруты по умолчанию. Для успешного преобразования ей нужно задать метрику по умолчанию.

Как видно из конфигурации маршрутизатора RTA, приведенной в листинге 12.55, в нем в качестве сети по умолчанию устанавливается сеть 192.68.6.0/24 (или любая другая классовая сеть, сведения о которой получены по BGP). Далее на RTA эта сеть переводится под управление IGRP.

Листинг 12.55. *Применение IGRP в качестве протокола IGP (конфигурация маршрутизатора RTA)*

```
router igrp 1
 passive-interface SerialO
 redistribute bgp 3 route-map DEFAULT_ONLY
 network 172.16.0.0
 default-metric 5 100 250 100 1500

router bgp 3
 no synchronization
 network 172.16.70.0 mask 255.255.255.0
 network 172.16.220.0 mask 255.255.255.0
 neighbor   172.16.20.1   remote-as   1
 neighbor  172.16.20.1  filter-list  10  out
 neighbor  172.16.50.1  remote-as   3
 neighbor   172.16.50.1  route-map  setlocalpref in
 no  auto-summary

ip  default-network   192.68.6.0
ip as-path access-list  10 permit  ^$

access-list  5 permit  192.68.6.0  0.0.0.255

route-map setlocalpref permit 10
 set  local-preference  300

route-map DEFAULT_ONLY permit 10
 match ip  address   5
```

Как видно из конфигурации маршрутизатора RTF, представленной в листинге 12.56, на нем также устанавливается маршрут по умолчанию в сеть 192.68.6.0/24 и преобразуется с лучшей метрикой в IGRP.

Листинг 12.56. *Применение IGRP в качестве протокола ЮР (конфигурация маршрутизатора RTF)*

```
router  igrp  1
redistribute bgp  3   route-map DEFAULT_ONLY
 network   172.16.0.0
 default-metric  1000   100  250  100   1500

router bgp  3
 no   synchronization
 network  172.16.50.0 mask 255.255.255.0
 neighbor  172.16.70.1   remote-as   3
 neighbor  172.16.70.1  next-hop-self
 neighbor  192.68.5.2   remote-as   2
 neighbor  192.68.5.2  filter-list  10  out
 no  auto-summary

ip  default-network  192.68.6.0
ip as-path access-list 10 permit  ^$

access-list  5  permit 192.68.6.0   0.0.0.255

route-map DEFAULT_ONLY permit  10
 match  ip  address   5
```

Как видите, на маршрутизаторе RTG поддерживается только протокол IGRP и маршрут по умолчанию за пределы AS3 (листинг 12.57).

Листинг 12.57. *Применение IGRP в качестве протокола ЮР (конфигурация маршрутизатора RTG)*

```
router  igrp   1
 network   172.16.0.0
```

В листинге 12.58 представлена таблица IP-маршрутов на маршрутизаторе RTG. Как видите, на RTG установлен маршрут по умолчанию в направлении RTF.

Листинг 12.58. *Таблица IP-маршрутов на маршрутизаторе RTG*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is 172.16.50.1 to network 192.68.6.0

I*   192.68.6.0/24 [100/8576] via 172.16.50.1, 00:00:32, Seriall
   172.16.0.0/16 is subnetted, 4 subnets
I    172.16.220.0/24 [100/8576] via 172.16.70.1, 00:00:32, SerialO
C    172.16.50.0/24 is directly connected, Seriall
I    172.16.20.0/24 [100/10476] via 172.16.70.1, 00:00:32, SerialO
C    172.16.70.0/24 is directly connected, SerialO
```

Применение IS-IS в качестве протокола IGP

Протокол IS-IS настраивается подобно OSPF. Для этого используется команда маршрутизатора default-information originate.

В конфигурации маршрутизатора RTA, представленной в листинге 12.59, указано, что он генерирует маршрут по умолчанию в IS-IS только при условии получения сведений о нем по внешнему каналу.

Листинг 12.59. *Применение IS-IS в качестве протокола IGP (конфигурация маршрутизатора RTA)*

```
router isis 100
 redistribute connected
 default-information originate route-map SEND_DEFAULT_IF
net 49.0001.0000.0cOO.OOOa.OO

router bgp 3
 no synchronization
 network 172.16.220.0 mask 255.255.255.0
 network 172.16.70.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.50.1 remote-as 3
 neighbor 172.16.50.1 route-map setlocalpref in
 no auto-summary

ip as-path access-list 10 permit ^A$

access-list 1 permit 0.0.0.0
access-list 2 permit 172.16.20.1

route-map SEND_DEFAULT_IF permit 10
 match ip address 1
 match ip next-hop 2
```

В конфигурации, представленной в листинге 12.60, маршрутизатор RTF генерирует маршрут по умолчанию в IS-IS при условии, что он получил информацию о нем по внешнему каналу.

Листинг 12.60. *Применение IS-IS в качестве протокола IGP (конфигурация маршрутизатора RTF)*

```
router  isis  100
 default-information  originate  route-map  SEND_DEFAULT_IF
 net  49.0001.0000.OcOO.OOOc.OO

router bgp  3
 no   synchronization
 network   172.16.50.0 mask  255.255.255.0
 neighbor  172.16.70.1  remote-as   3
 neighbor  172.16.70.1  next-hop-self
 neighbor  192.68.5.2  remote-as   2
 neighbor  192.68.5.2   filter-list  10  out
 no  auto-summary

ip  as-path  access-list  10 permit   ^A$

access-list 1  permit  0.0.0.0
access-list 2  permit  192.68.5.2

route-map   SEND_DEFAULT_IF permit   10
 match ip address  1
 match ip next-hop 2
```

Маршрутизатор RTG сконфигурирован таким образом, что на нем поддерживается только IS-IS и выход за пределы AS3 возможен только с помощью маршрута по умолчанию 0/0 (листинг 12.61).

Листинг 12.61. *Применение IS-IS в качестве протокола IGP (конфигурация маршрутизатора RTG)*

```
router  isis   100
 net  49.0001.0000.0c00.000b.00
```

В листинге 12.62 представлена таблица IP-маршрутов на маршрутизаторе RTG. Как видите, на RTG установлен маршрут по умолчанию в направлении к RTF.

Листинг 12.62. *Таблица IP-маршрутов на маршрутизаторе RTG*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is 172.16.50.1 to network 0.0.0.0

    172.16.0.0/16 is subnetted, 4 subnets
i LI     172.16.220.0/24 [115/20] via 172.16.70.1, Serial0
i LI     172.16.20.0/24 [115/20] via 172,16.70.1, Serial0
C        172.16.50.0/24 is directly connected, Seriall
C        172.16.70.8/24 is directly connected, Serial0
i*L2 0.0.0.0/0 [115/10] via 172.16.50.1, Seriall
```

Маршрутизация по правилам

В этом разделе показано, как использовать маршрутизацию по правилам для пересылки трафика на основе IP-адреса источника, а не IP-адреса пункта назначения. На рис. 12.12 показан маршрутизатор RTA, который поддерживает работу по протоколу BOP с провайдерами AS1 и AS2. На внутренних маршрутизаторах, таких как RTG и RTF, поддерживается только протокол IGP (в данном случае OSPF) и установлены маршруты по умолчанию в направлении RTA.

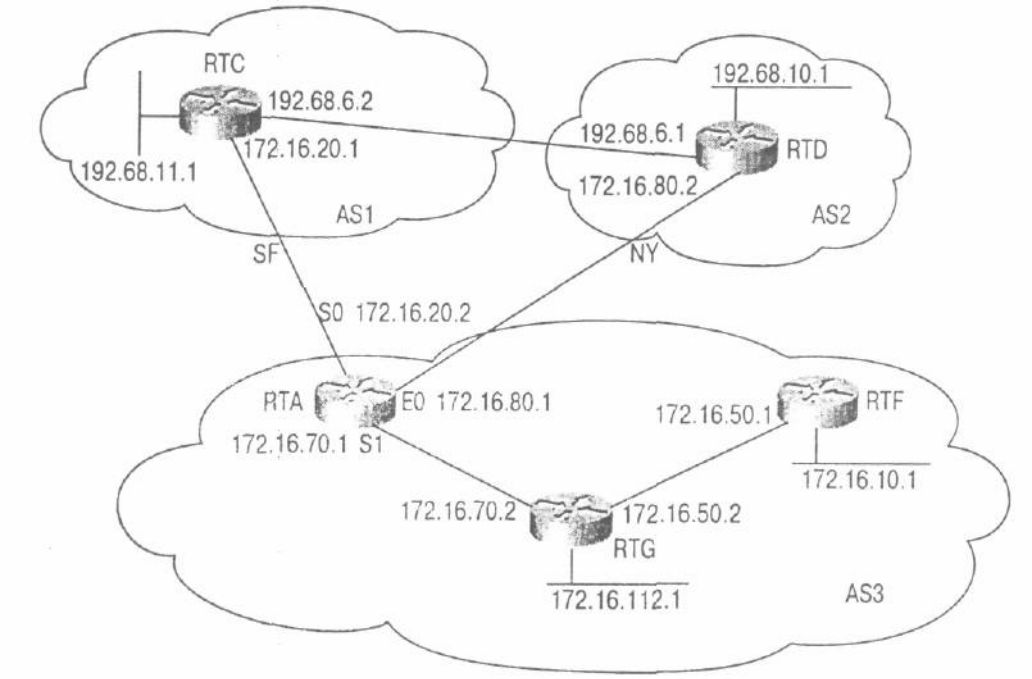


Рис. 12.12. *Вариант организации маршрутизации по правилам*

Итак, на маршрутизаторе RTA нужно установить такие правила маршрутизации, чтобы трафик, поступающий через интерфейс S1 от RTG направлялся бы в AS2, если он поступил из сети 172.16.10.0/24. Трафик, поступающий от RTG из сети 172.16.112.0/24, следует направлять в AS1, а в случае выхода из строя канала с провайдером — в AS2. Все остальные IP-адреса источников обрабатывать без изменений. В листинге 12.63 представлена

конфигурация маршрутизатора RTA, удовлетворяющая критериям вышеприведенных правил маршрутизации.

Листинг 12.63. *Маршрутизация по правилам (конфигурация маршрутизатора RTA)*

```
interface Ethernet0
 ip address 172.16.80.1 255.255.255.0

interface Seriall
 ip address 172.16.70.1 255.255.255.0
 ip policy route-map CHECK_SOURCE

router ospf 10
 passive-interface Serial0
 passive-interface Ethernet0
 network 172.16.0.0 0.0.255.255 area 0
 default-information originate always

router bgp 3
 network 172.16.50.0 mask 255.255.255.0
 network 172.16,70.0 mask 255.255.255.0
 network 172.16.10.0 mask 255.255.255.0
 network 172.16.112.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.80.2 remote-as 2
 neighbor 172.16.80.2 filter-list 10 out
 no auto-summary

ip as-path access-list 10 permit ^$

access-list 1 permit 172.16.10.0 0.0.0.255
access-list 2 permit 172.16.112.0 0.0.0.255

route-map CHECK__SOURCE permit 10
 match ip address 1
 set ip next-hop 172.16.80.2

route-map CHECK__SOURCE permit 20
 match ip address 2
 set ip next-hop 172.16.20.1 172.16.80.2
```

Маршрутизация по правилам используется всегда в отношении входящего интерфейса. Как видите, интерфейс Serial 1 настраивается с помощью команды **ip policy route-map** *map-name.* В нашем случае ко всему входящему на интерфейс Serial 1 трафику применяется карта маршрутов CHECK^SOURCE. Эта карта работает следующим образом.

- Запись 10 - для всех IP-адресов источника, которые принадлежат сети 172.16.10.0/24, назначить следующим узлом 172.16.80.2. Если узел 172.16.80.2 недоступен, пакет отвергается.
- Запись 20 - для всех IP-адресов источника, принадлежащих сети 172.16.112.0/24, установить в качестве следующего узла 172.16.20.1. Если узел 172.16.20.1 недоступен, то попытаться направить трафик на узел 172.16.80.2.
- Все остальные IP-адреса обрабатываются как обычно.

В данном случае маршрутизация по правилам с помощью карт маршрутов позволяет выбирать следующий узел, на который будет пересылаться трафик. Необходимо всегда иметь резервный маршрут. Весь трафик, на который не распространяется действие карт маршрутов, должен обрабатываться без изменений согласно протоколу маршрутизации. Чтобы проиллюстрировать, что мы получили в результате последних манипуляций, на маршрутизаторе RTG мы задали команду traceroute 192.68.10.1 с IP-адресом источника 172.16.112.1. В листинге 12.64 представлена таблица IP-маршрутов на маршрутизаторе RTA.

Листинг 12.64. *Маршрутизация по правилам (таблица маршрутов на маршрутизаторе RTA)*

```
RTA#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B -BGP
       D - EIGRP, EX - EIGRP external, 0 - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       El - OSPP external type 1, E2 -- OSPF external type 2, E - EGP
       i - IS-IS, LI - IS-IS level-1, L2 - IS-IS level-2,
         * - candidate default U - per-user static route, o - ODR

Gateway of last resort is not set

B  192.68.10.0/24 [20/0] via 172.16.80.2, 00:30:09
B  192.68.11.0/24 [20/0] via 172.16.20.1, 00:30:14
   172.16.0.0/16 is subnetted, 5 subnets
O    172.16.50.0/24 [110/69] via 172.16.70.2, 00:27:27, Seriall
C    172.16.20.0/24 is directly connected, SerialO
C    172.16.80.0/24 is directly connected, EthernetO
C    172.16.70.0/24 is directly connected, Seriall
```

В листинге 12.65 показан процесс выполнения команды traceroute на маршрутизаторе RTG с адресом источника 172.16.112.1 и адресом пункта назначения 192.68.10.1.

Листинг 12.65. *Маршрутизация по правилам (выполнение команды traceroute на маршрутизаторе RTG)*

```
RTG#traceroute
Protocol [ip]:
Target IP address: 192.68.10.1
Source address: 172.16.112.1
Numeric display [n]:
Timeout in seconds [3]:
Probe count [3]:
Minimum Time to Live [1]:
Maximum Time to Live [30]:
Port Number [33434]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Type escape sequence to abort.
Tracing the route to gateway.aeg-aas.de (192.68.10.1}
172.16.70.1 4 msec 4 msec 0 msec
172.16.20.1 4 msec 4 msec 4 msec
192.68.6.1 4 msec 4 msec 4 msec
```

Обратите внимание, что маршрутизатор RTA связывается с сетью 192.68.10.0/24 через промежуточный узел 172.16.20.1 (вторая строка в выводе результатов по команде **traceroute).** Хотя даже таблица маршрутов на RTA указывает, что в сеть 192.68.10.0/24 можно попасть через узел 172.16.80.2.

В этой второй попытке показано, что может случиться, если интерфейс Serial 0 вышел из строя, а узел 172.16.20.1 недоступен. Теперь давайте выполним команду **traceroute** на маршрутизаторе RTG с адресом источника 172.16.112.1 и пунктом назначения 192.68.10.1, когда интерфейс Serial 0 неактивен, как это показано в листинге 12.66.

Листинг 12.66. *Маршрутизация по правилам (повторное выполнение команды traceroute на маршрутизаторе RTG)*

```
RTG# traceroute
Protocol [ip]:
Target IP address: 192.68.10.1
Source address: 172.16.112.1
Numeric display [n]:
Timeout in seconds [3]:
Probe count [3] :
```

```
Minimum Time to Live [1]:
Maximum Time to Live [30]:
Port Number [33434]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Type escape sequence to abort.
Tracing the route to gateway.aeg-aas.de (192.68.10.1}
172.16.70.1 0 msec 4 msec 4 msec
172.16.80.2 8 msec 4 msec 4 msec
```

Как видите, вывод результатов выполнения traceroute, представленный в листинге 12.66, демонстрирует, что маршрутизатор RTA выбирает альтернативный следующий узел 172.16.80.2.

Прежде чем приступить к реализации решений на основе маршрутизации по правилам, ознакомьтесь с документацией на вашу версию IOS и уточните, поддерживаются ли те процедуры, которые вы собираетесь использовать, и если — да, то на каком оборудовании. Кроме того, вы можете получить дополнительную информацию по этим вопросам в справочнике по командам Cisco (Cisco Guide Reference), руководстве по конфигурированию (Configuration Guidelines) и в информационных бюллетенях для вашего оборудования и версий программного обеспечения (Release Notes).

Отражатели маршрутов

В этом разделе мы рассмотрим варианты практического применения отражателей мари ipyroB (route reflectors) и групп взаимодействующих узлов (peer groups). На рис. 12.13 представлены маршрутизаторы RTA и RTG, которые формируют кластер отражателя маршрутов, где RTG является отражателем маршрутов. Маршрутизаторы RTF, RTE и RTD формируют еще один кластер, где в качестве отражателя маршрутов используется маршрутизатор RTF. Кроме того, маршрутизаторы RTG и RTF являются частью группы взаимодействующих узлов с названием REFLECTORS. Если существуют другие отражатели маршрутов, то все они должны соединяться по полносвязной схеме с использованием протокола 1BGP. Маршрутизатор RTF помещает всех своих клиентов в группу с именем CLIENTS, где используются общие правила маршрутизации.

Клиенты идентифицируются серверами отражателей маршрутов путем задания !"imic-reflector-client с соответствующим параметром neighbor **x.x.x.x.** Традиционно ложилось определенное ограничение в IOS, согласно которому необходимо, чтобы .ражение маршрутов между клиентами было запрещено, если они являются членами . ;п ппы взаимодействующих узлов. Однако теперь это ограничение снято. Внутри кла-гера RTF-RTD-RTE поддерживается полносвязная работа по протоколу IBGP, и ".•тенты взаимодействуют только с соответствующим отражателем маршрутов. В листинге 12.67 приведена конфигурация маршрутизатора RTF.

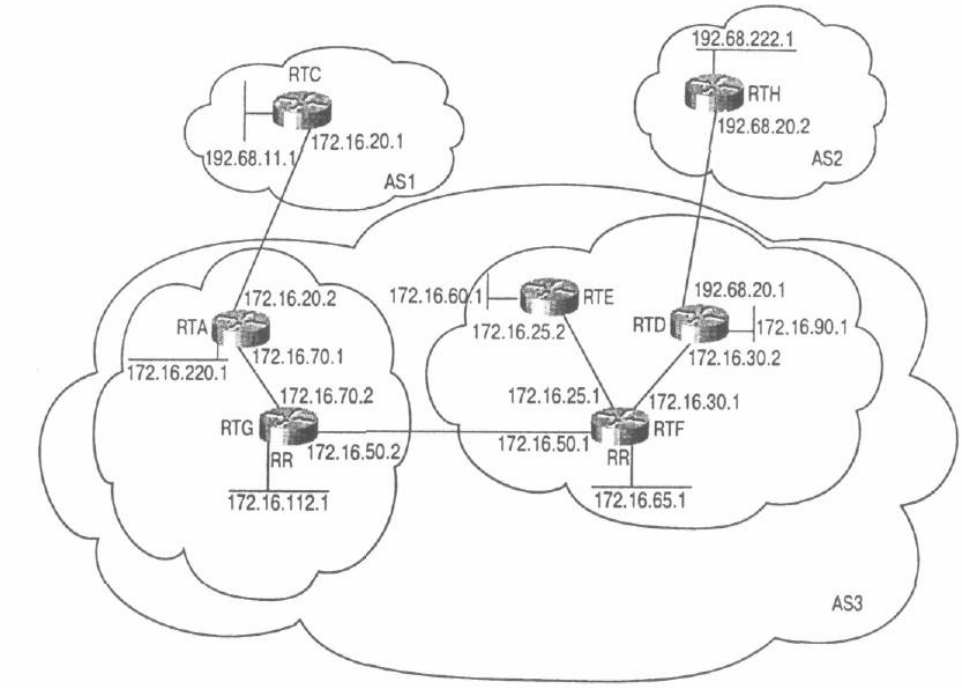


***Рис. 12.13**. Отражатели маршрутов*

***Листинг 12.67.** Отражатели маршрутов (конфигурация маршрутизатора RTF)*

```
router bgp 3
 no  synchronization
 network   172.16.65.0 mask  255.255.255.192
 network  172.16.50.0 mask  255.255.255.0
 network  172.16.25.0  mask  255.255.255.0
 network  172.16.30.0 mask  255.255.255.0
 neighbor REFLECTORS peer-group
 neighbor REFLECTORS  remote-as   3
 neighbor  CLIENTS peer-group
 neighbor  CLIENTS  remote-as   3
 neighbor  CLIENTS   route-reflector-client
 neighbor  172.16.25.2  peer-group  CLIENTS
 neighbor   172.16.30.2  peer-group  CLIENTS
 neighbor  172.16.50.2  peer-group REFLECTORS
 no  auto-summary
```

Так как на RTF имеется единственный возможный маршрут за пределы кластера, маршрутизаторы RTE и RTD сконфигурированы как клиенты отражателя маршрутов. Вместе маршрутизаторы составляют так называемый *кластер (cluster)*. Маршрутизаторы RTE и RTD ведут с RTF обычный IBGP-сеанс. Другими словами, клиент может и не знать, что он является клиентом. (Заметим, что это было одно из требований при организации отражения маршрутов -- клиенты не должны знать, что они клиенты). В листингах 12.68—12.70 представлены конфигурации маршрутизаторов RTD, RTG и RTA.

***Листинг 12.68.** Отражатели маршрутов (конфигурацияі маршрутизатора RTD)*

```
router bgp 3
 no synchronization
 network 172.16.90.0 mask 255.255.255.0
 network 172.16.30.0 mask 255.255.255.0
 neighbor 172.16.30.1 remote-as 3
 neighbor 172.16.30.1 next-hop-self
 neighbor 192.68.20.2 remote-as 2
 neighbor 192.68.20.2 filter-list 10 out
 no auto-summary

ip as-path access-list 10 permit ^$
```

Листинг 12.69. *Отражатели маршрутов (конфигурация маршрутизатора RTG)*

```
router bgp 3
 no synchronization
 network 172.16.112.0 mask 255.255.255.0
 network 172.16.50,0 mask 255.255.255.0
 network 172.16.70.0 mask 255.255.255.0
 neighbor 172.16.50.1 remote-as 3
 neighbor 172.16.70.1 remote-as 3
 neighbor 172.16.70.1 route-reflector-client
 no auto-summary
```

Листинг 12.70. *Отражатели маршрутов (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no  synchronization
 network  172.16.220.0  mask  255.255.255.0
 network  172.16.70.0 mask  255.255.255.0
 neighbor  172.16.20.1  remote-as  1
 neighbor  172.16.20.1  filter-list 10  out
 neighbor  172.16.70.2  remote-as  3
 neighbor  172.16.70.2   next-hop-self
 no  auto-summary

ip as-path access-list 10 permit ^$
```

В таблице BGP-маршругов, приведенной в листинге 12.71, вы можете видеть, как некоторые из маршрутов на маршрутизаторе RTD отражаются в его собственный кластер.

Листинг 12.71. *Отражатели маршрутов (таблица BGP-маршрутов на маршрутизаторе RTD)*

```
RTD#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 52
Paths: (1 available, best #1)
 Local
(metric 192} from 172.16.30.1 (172.16.220.1)
 Origin IGP, metric 0, localpref 100, valid, internal, best
 Originator : 172.16.220.1, Cluster list: 172.16.65.1, 172.16.112.1
```

Обратите внимание, что RTD воспринимает узел, сгенерировавший маршрут 172.16.220.0/24, как 172.16.220.1, т.е. ROUTERJD маршрутизатора RTA. Маршрут также несет в себе информацию о списке кластеров, куда входят ROUTER_1D всех отражателей маршрутов, которые он пересекает.

Когда несколько отражателей маршрутов сконфигурированы внутри кластера, всем им должен быть присвоен одинаковый CLUSTER_ID. Это необходимо, чтобы предотвратить появление петель маршрутизации, которые могут возникать между кластерами. Например, если маршрутизатор RTE был сконфигурирован как отражатель маршрутов, то на маршрутизаторах RTF и RTE следует задать дополнительную команду bgp cluster-id *number*. В листинге 12.72 показана соответствующая конфигурация маршрутизатора RTF.

Листинг 12.72. *Несколько отражателей маршрутов (конфигурация маршрутизатора RTF)*

```
router bgp 3
 no synchronization
 network 172.16.65.0 mask 255.255.255.192
 network 172.16.50.0 mask 255.255.255.0
 network 172.16.25.0 mask 255.255.255.0
 network 172.16.30.0 mask 255.255.255.0
 neighbor REFLECTORS peer-group
 neighbor REFLECTORS remote-as 3
 neighbor CLIENTS peer-group
 neighbor CLIENTS remote-as 3
```

```
neighbor CLIENTS route-reflector-client
neighbor 172.16.25.2 peer-group CLIENTS
neighbor 172.16.30.2 peer-group CLIENTS
neighbor 172.16.50.2 peer-group REFLECTORS
bgp cluster-id 1000
no auto-summary
```

Параметр CLUSTER_ID — это специальный номер, идентифицирующий кластер. Каждому кластеру нужно назначить уникальный номер. Это также необходимо, чтобы предотвратить образование петель маршрутизации в том случае, если и RTF, и RTE настроены как отражатели маршрутов в одном кластере. Отметим, что "хорошим тоном" считается назначение CLUSTERED, даже если в кластере используется только один отражатель маршрутов. После задания CLUSTER_ID на маршрутизаторе нельзя добавлять новых клиентов. Следовательно, если вы упустили при конфигурировании одного из клиентов, вам придется удалить всех клиентов из списков, а затем снова создать их и повторно установить CLUSTER_ID. Так что определение CLUSTER_1D во время начальной загрузки маршрутизатора — очень хорошая идея.

Конфедерации

На схеме, представленной на рис. 12.14, показано, как разделить AS3 на две меньших подсистемы — AS65050 и AS65060. Номера подсистем AS выбраны из диапазона частных номеров (64512—65535). В каждой из подсистем AS в качестве протокола IGP используется протокол OSPF. Причем OSPF внутри AS65050 не зависит от OSPF в AS65060, т.е. номера областей, используемые в AS65050, можно повторно использовать в AS65060. Таким образом, вы получаете определенное преимущество по сравнению даже с BGP, если IGP в одной подсистеме AS не зависит от 1GP в другой подсистеме AS.

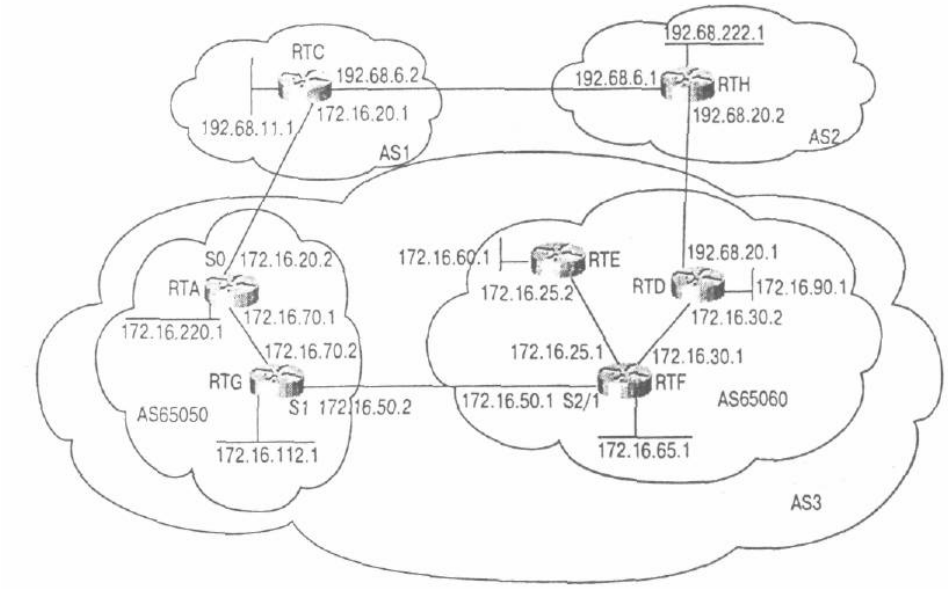


Рис. 12.14. *Конфедерации*

Из конфигурации маршрутизатора RTA, приведенной в листинге 12.73, **видно,** что все его интерфейсы находятся в области OSPF с номером 5. На маршрутизаторе RTA для взаимодействия с RTC в AS1 поддерживается протокол EBGP, а для работы с RTG в AS65050 - протокол IBGP.

Обратите внимание, что на маршрутизаторе RTA субкоманда **bgp configuration identifier** 3 помогает представить его маршрутизатору RTC как члена конфедерации 3.

Листинг 12.73. *Конфедерации (конфигурация маршрутизатора RTA)*

```
router ospf 10
 passive-interface Serial0
 network 172.16.0.0 0.0.255.255 area 5

router bgp 65050
 no synchronization
 bgp confederation identifier 3
 network 172.16.220.0 mask 255.255.255.0
 network 172.16.70.0 mask 255.255.255.0
 neighbor 172.16.20.1 remote-as 1
 neighbor 172.16.20.1 filter-list 10 out
 neighbor 172.16.70.2 remote-as 65050
 no auto-summary

ip as-path access-list 10 permit ^$
```

Как видно из листинга 12.74, на маршрутизаторе **RTC** при взаимодействии с RTA используется обычный протокол EBGP. С точки зрения RTC, маршрутизатор RTA принадлежит AS3. Он также не имеет сведений о подсистемах AS внутри конфедерации 3. И с маршрутизатором RTH в AS2 RTC также взаимодействует по **EBGP.**

Листинг 12.74. *Конфедерации (конфигурация маршрутизатора RTC)*

```
router bgp  1
 network    192.68.11.0
 neighbor   172.16.20.2   remote-as   3
 neighbor   192.68.6.1   remote-as  2
 no  auto-summary
```

Маршрутизатор RTG является граничным для подсистемы AS65050 и взаимодействует с RTF в подсистеме AS65060 по *EBGP для конфедераций (confederation EBGP).* Кроме того, с маршрутизатором RTA он взаимодействует с помощью протокола IBGP. К тому же RTG является граничным маршрутизатором, который имеет общую с RTA область OSPF 5, а остальные его интерфейсы лежат в области 0. Обратите внимание, что в RTG запрещена обработка данных, поступающих по протоколу OSPF через интерфейс Serial I (passive-interface Serial 1), который к тому же является общим с маршрутизатором RTF. В этом соединении поддерживается только протокол EBGP. В листинге 12.75 приведена конфигурация маршрутизатора RTG.

Листинг 12.75. *Конфедерации (конфигурация маршрутизатора RTG)*

```
router  ospf  10
 passive-interface  Seriall
 network  172.16.70.2   0.0.0.0  area  5
 network  172.16.0.0   0.0.255.255   area   0

router bgp   65050
 no  synchronization
 bgp  confederation  identifier  3
 bgp  confederation peers   65060
 network   172.16.112.0  mask  255.255.255.0
 network   172.16.50.0  mask  255.255.255.0
 network  172.16.70.0  mask  255.255.255.0
 neighbor  172.16.50.1   remote-as   65060
 neighbor  172.16.50.1  next-hop-self
 neighbor   172.16.70.1   remote-as   65050
 no  auto-summary
```

Как видите, RTG также идентифицирует себя как часть конфедерации 3 (bgp confederation identifier 3). Чтобы сохранить все атрибуты, такие как LOCAL_PREF и NEXT_HOP, при проведении EBGP-сеанса с AS65060 на RTG используется команда bgp confederation peers 65060. С помощью этой команды EBGP-сеанс для конфедерации с подсистемой AS65060 будет выглядеть как сеанс по протоколу IBGP. С помощью команды

neighbor 172.16.50.1 next-hop-self для маршрутов, поступающих от RTG на RTF, в качестве адреса следующего ближайшего узла будет устанавливаться IP-адрес марш руги затора RTG. Без этой команды адрес следующего ближайшего узла для всех EBGP-маршрутов от AS1 будет посылаться на RTF равным 172.16.20.1, что действительно только при наличии к нему доступа из конфедерации маршрутизаторов в подсистеме AS65060.

Как видно из листинга 12.76, та же конфигурация применяется и на маршрутизаторе RTF, который также является граничным для подсистемы AS65060. Кроме того, RTF является граничным маршрутизатором для областей 0 и 5. Области 0 и 5 в AS65060 абсолютно не зависят от таких же областей в AS65050. Таким образом, оба протокола IGP "защищены" друг от друга посредством BGP. При этом между маршрутизаторами RTE, RTD и RTF поддерживается полносвязная схема взаимодействия с использованием протокола IBGP и группы узлов с именем SUB_AS_65060.

Листинг 12.76. *Конфедерации (конфигурация маршрутизатора RTF)*

```
router ospf  10
 passive-interface  Serial2/l
 network  172.16.25.1   0.0.0.0  area  5
 network  172.16.0.0   0.0.255.255  area  0

router bgp   65060
 no  synchronization
 bgp  confederation  identifier  3
 bgp  confederation peers   65050
 network  172.16.65.0 mask  255.255.255.192
 network  172.16.50.0 mask  255.255.255.0
 network  172.16.25.0  mask  255.255.255.0
 network  172.16.30.0 mask  255.255.255.0
 neighbor  SUB_AS_65060 peer-group
 neighbor  SUB_AS_65060  remote-as   65060
 neighbor  172.16.25.2  peer-group  SUB_AS_65060
 neighbor  172.16.30.2  peer-group  SUB_AS_65060
 neighbor  172.16.50.2   remote-as   65050
 neighbor  172.16.50.2  next-hop-self
```

Маршрутизатор RTD является граничным *для* конфедерации 3. Как видно из листинга 12.77, RTD взаимодействует с RTH в AS2 по протоколу EBGP и по полносвязной схеме — с RTE и RTF в подсистеме AS65060. Все интерфейсы маршрутизатора RTD находятся в области 0. На RTD не поддерживается OSPF во внешнем канале с AS2. Вот почему для внешних обновлений, поступающих на RTD, в качестве следующего ближайшего узла следует задать его же, прежде чем распространять эти сведения в направлении RTF и RTE.

Листинг 12.77. *Конфедерации (конфигурация маршрутизатора RTD)*

```
router  ospf  10
 network  172.16.0.0   0.0.255.255  area  0.0.0.0

router bgp   65060
 no  synchronization
 bgp  confederation  identifier  3
 network  172.16.90.0 mask  255.255.255.0
 network  172.16.30.0 mask  255.255.255.0
 neighbor  172.16.25.2   remote-as   65060
 neighbor   172.16.25.2  next-hop-self
 neighbor   172.16.30.1   remote-as   65060
 neighbor  172.16.30.1   next-hop-self
 neighbor  192.68.20.2   remote-as   2
 neighbor  172.16.20.2   filter-list  10  out
 no  auto-summary

ip  as-path  access-list  10  permit  ^$
```

Как видно из листинга 12.78, все интерфейсы маршрутизатора RTE лежат в области 5 протокола OSPF, и он поддерживает полносвязную схему IBGP с RTF и RTD.

Листинг 12.78. *Конфедерации (конфигурация маршрутизатора RTE)*

```
router  ospf  10
 network  172.16.0.0   0.0.255.255  area  5

router bgp  65060
 no  synchronization
 bgp  confederation  identifier  3
 network  172.16.60.0 mask  255.255.255.0
 network  172.16.25.0 mask  255.255.255.0
 neighbor  172.16.25.1  remote-as   65060
 neighbor  172.16.30.2   remote-as   65060
 no  auto-summary
```

Как показано в листинге 12.79, маршрутизатор RTH является граничным для AS2 и на нем поддерживается протокол EBGP для обеспечения взаимодействия между AS1 и AS3. При этом RTH не имеет сведений о подсистемах AS в конфедерации 3.

Листинг 12.79. *Конфедерации (конфигурация маршрутизатора RTH)*

```
router bgp 2
 network  192.68.222.0
 neighbor  192.68.6.2   remote-as   1
 neighbor   192.68.20.1   remote-as   3
 no  auto-summary
```

Теперь давайте рассмотрим несколько фрагментов таблиц BGP-маршрутов. Посмотрите, как на маршрутизаторе RTH отображаются маршруты (листинг 12.80). Это делается через AS1 и через AS3. Как видите, все подсистемы AS "невидимы" для RTH.

Листинг 12.80. *Конфедерации (таблица BGP-маршрутов на маршрутизаторе RTH)*

```
RTC# show ip bgp
BGP table version is 477, local router ID is 192.68.222.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network                Next Hop      Metric      LocPrf    Weight   Path
 *> 172.16.25.0/24      192.68.20.1                         0        3 i
                        192.68.6.2                          0        1 3 i
 *> 172.16.30.0/24      192.68.20.1   0                     0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.50.0/24      192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.60.0/24      192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.70.0/24      192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.90.0/24      192.68.20.1   0                     0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.65.0/26      192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.112.0/24     192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 172.16.220.0/24     192.68.20.1                         0        3 i
 *                      192.68.6.2                          0        1 3 i
 *> 192.68.11.0         192.68.6.2    0                     0        1 i
 *                      192.68.20.1                         0        3 1 i
 *> 192.68.222.0        0.0.0.0       0                     32768    i
```

Рассмотрим теперь таблицу BGP-маршрутов на маршрутизаторе RTA, которая представлена в листинге 12.81. Из нее видно, что все подсистемы AS заключены в скобки.

Любой маршрут между подсистемами AS имеет длину 0. Посмотрим, как двумя способами получаются сведения о префиксе 192.68.222.0/24: через внутренний маршрут (65060) 2 и через внешний маршрут 1 2. Длина внутреннего маршрута (65060) 2 считается меньшей, так как подсистемы AS при вычислении длины маршрута не учитываются. Именно поэтому был выбран внутренний маршрут.

Листинг 12.81. *Конфедерации (таблица BGP-маршрутов на маршрутизаторе RTA)*

```
RTC# show ip bgp
BGP table version is 13, local router ID is 172.16.220.1
Status codes: s suppressed, d damped, h history, * valid, > best,
    i - internal Origin codes: i - IGP, e - EGP, ? – incomplete
 Network              Next Hop      Metric     LocPrf  Weight  Path
 *>i172.16.25.0/24    172.16.50.1   0          100     0       (65060) i
 *>i172.16.30.0/24    172.16.50.1   0          100     0       (65060) i
 *>i172.16.50.0/24    172.16.70.2   0          100     0       i
 *>i172.16.60.0/24    172.16.50.1   0          100     0       (65060) i
 *> 172.16.70.0/24    0.0.0.0       0                  32768   i
 * i                  172.16.70.2   0          100     0       i
 *>i172.16.90.0/24    172.16.50.1   0          100     0       (65060) i
 *>i172.16.65.0/26    172.16.50.1   0          100     0       (65060) i
 *>i172.16.112.0/24   172.16.70.2   0          100     0       i
 *> 172.16.220.0/24   0.0.0.0       0                  32786   i
 *> 192.68.11.0       172.16.20.1   0                  0       1 i
 * 192.68.222.0       172.16.20.1                      0       1 2 i
 *>i                  172.16.50.1                100     0       (65060) 2
                                                               i
```

В таблице BGP на маршрутизаторе RTF, представленной в листинге 12.82, все маршруты от подсистемы AS65050 считаются внешними маршрутами конфедерации (внешние конфедераты). Внутри конфедерации решение принимается на основе следующих положений: маршруты EBGP более предпочтительны, чем внешние конфедераты, которые, в свою очередь, более предпочтительны, чем внутренние маршруты.

Листинг 12.82. *Конфедерации (BGP-таблица на маршрутизаторе RTA)*

```
RTF#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 22
Paths: (1 available, best #1, advertised over IBGP)
 (65050)
from 172.16.50.2 (172.16.112.1)
Origin IGP, metric 0, localpref 100, valid, confed-external, best
```

Управление маршрутами и аннулирование содержимого кэша

Традиционным требованием в протоколе BGP был сброс TCP-соединения между взаимодействующими узлами для того, чтобы изменения, внесенные в правила маршрутизации возымели действие (**clear ip bgp** [* / *address peer-group]).* При сбрасывании сеансов подобным образом фаза переговоров повторяется с самого начала, при этом аннулируется все содержимое кэша для пересылки IP-пакетов, что существенно влияет на функционирование сети.

Основные причины этого, как уже отмечалось в главе 6, заключаются в том, что маршруты, сведения о которых получены от других узлов, помещаются в базу Adj-RIB-In. Затем эта информация передается входным правилам маршрутизации, соответствующим образом модифицируется и передается процессу принятия решения в BGP. Так как

немодифицированная копия маршрута (которая изначально хранится в Adj-RIB-In) обычно недоступна, но требуется для вступления в действие новых правил маршрутизации, то можно поступить следующим образом.

1. Источник BGP-маршрута можно вручную заставить повторить сведения о маршруте.
2. Может произойти полный сброс сеанса TCP.
3. Чтобы сохранить данные из Adj-RIB-In в памяти, можно воспользоваться *мягкой перенастройкой (soft reconfiguration)* BGP.
4. Чтобы запросить удаленную сторону о повторе ее Adj-RIB-Out, можно использовать обновление BGP-маршрутов (BGP Route Refresh).

Что касается первого исхода, то маловероятно, что в реальной жизни вы будете иметь доступ к маршрутизатору, который сгенерировал исходный маршрут. Второй подход есть проявление грубой силы, что повлечет за собой непредсказуемые последствия, повышающие нестабильность сети. Мягкая перенастройка — очень хороший и элегантный способ решения проблемы, но он требует большого количества памяти. Обновление маршрутов BGP — относительно новое решение, которое со временем будет самым эффективным. Ниже мы обсудим варианты 3 и 4 более подробно.

Мягкая перенастройка BGP

Мягкая перенастройка (soft reconfiguration) BGP — один из методов, позволяющих настройку и введение в действие правил маршрутизации без сброса TCP-сеанса. Это позволяет применять новые правила маршрутизации практически незаметно для сети. Мягкая перенастройка может применяться двумя путями — по входящей и исходящей базам маршрутной информации. Для этого используется следующая команда EXEC:

```
clear ip bgp [* / address / peer-group] [soft [in\out]]
```

Мягкая перенастройка по информационной базе исходящих маршрутов

Всегда при использовании мягкой перенастройки по информационной базе исходящих маршрутов новые правила маршрутизации автоматически вступают в силу и генерируются соответствующие обновления маршрутов (из базы Adj-RIB-Out). Дополнительных ресурсов памяти этот тип мягкой перенастройки не требует. Для осуществления настройки выполняется следующая команда EXEC:

```
clear ip bgp [* / address / peer-group] soft out
```

Мягкая перенастройка по информационной базе входящих маршрутов

Мягкая перенастройка по информационной базе входящих маршрутов проводится немного сложнее. Все входящие обновления маршрутов (из Adj-RIB-In) от заданного узла хранятся в памяти без изменений. Когда вводятся новые правила маршрутизации, то база Adj-RIB-In, которая находится в памяти просто еще раз передается для обработки с помощью входных правил маршрутизации. Для этого используется дополнительная субкоманда:

```
neighbor {address / peer-group} soft-reconfiguration inbound
```

Эта команда необходима для сохранения обновлений маршрутов, поступающих от заданного узла или группы узлов. Для мягкой перенастройки по базе входящих маршрутов нужно использовать команду EXEC:

```
clear ip bgp [* / address / peer-group] soft in
```

Чтобы избежать излишней нагрузки на память, тот же результат может быть

достигнут и путем использования мягкой перенастройки по информационной базе исходящих маршрутов, но на другом конце соединения, что также вызовет повторное объявление базы Adj-RIB-Out.

Если параметр **in/out** не указан (**clear ip bgp** [* | *address* | *peer-group]* **soft**), то выполняется мягкая перенастройка и по входящим, и по исходящим базам маршрутной информации.

Ниже мы рассмотрим пример, где покажем, в чем заключается разница между сбросом BGP-сеанса между двумя маршрутизаторами с использованием мягкой перенастройки BGP и без нее. Пока выполняется сброс сеанса, вы сможете видеть отчет об установлении сеанса и об обмене маршрутной информацией.

Опираясь на схему, приведенную в рис. 12.14, рассмотрим, как сконфигурирован маршрутизатор RTA, посылающий сообщения об обновлении маршрутов на RTC с метрикой 5000 (листинг 12.83).

Листинг 12.83. *Мягкая перенастройка по информационной базе входящих маршрутов (конфигурация маршрутизатора RTA)*

```
router bgp   65050
 no  synchronization
 bgp  confederation  identifier  3
 network  172.16.220.0  mask  255.255.255.0
 network  172.16.70.0 mask  255.255.255.0
 neighbor  172.16.20.1   remote-as   1
 neighbor  172.16.20.1   soft-reconfiguration  inbound
 neighbor  172.16.20.1   filter-list 10  out
 neighbor   172.16.20.1   route-map  setmetric out
 neighbor   172.16.70.2   remote-as   65050
 no auto-summary

ip  as-path  access-list  10  permit   A$

route-map  setmetric permit  10
 set metric   5000
```

Обратите внимание на команду neighbor 172.16.20.1 soft-reconfiguration inbound в листинге 12.83. Она необходима только в том случае, если требуется, чтобы сброс сеанса оказал влияние на информационную базу входящих маршрутов, т.е. если вы не можете повлиять на соседний маршрутизатор, чтобы на нем был сброшен сеанс по информационной базе исходящих маршрутов.

Чтобы новые правила маршрутизации вступили в силу, необходимо сбросить BGP-сеанс между двумя маршрутизаторами, как это показано в листинге 12.84.

Листинг 12.84. *Мягкая перенастройка по информационной базе входящих маршрутов (сброс BGP-сеанса между двумя маршрутизаторами)*

```
RTA#clear ip bgp 172.16.20.1
BGP: 172.16.20.1 reset requested BGP: no valid path for 192.68.11.0/24
BGP: 172.16.20.1 reset by 0x27B740
BGP: 172.16.20.1 went from Established to Idle
BGP: nettable_walker 192.68.11.0/255.255.255.0 no best path selected
BGP: 172.16.20.1 went from Idle to Active
BGP: 172.16.70.2 computing updates, neighbor version 21, table version 23, starting at 0.0.0.0
BGP: 172.16.70.2 send UPDATE 192.68.11.0/24 - unreachable
BGP: 172.16.70.2 1 updates enqueued (average=27, maximum=27)
BGP: 172.16.70.2 update run completed, ran for 0ms, neighbor version 21, start version 23,
throttled to 23, check point net 0.0.0.0
BGP: scanning routing tables BGP  172.16.20.1 went from Active to OpenSent
BGP: 172.16.20.1 went from OpenSent to OpenConfirm
BGP: 172.16.20.1 went from OpenConfirm to Established
BGP: 172.16.20.1 computing updates, neighbor version 0, table version 23, starting at 0.0.0.0
BGP: 172.16.20.1 send UPDATE 172.16.25.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.30.0/24, next 172.16.20.2, path (65060)
BGP: 172.16.20.1 send UPDATE 172.16.50.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.60.0/24, next 172.16.20.2, path (65060)
BGP: 172.16.20.1 send UPDATE 172.16.70.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.90.0/24, next 172.16.20.2, path (65060)
BGP: 172.16.20.1 send UPDATE 172.16.65.0/26, next 172.16.20.2, path (65060)
BGP: 172.16.20.1 send UPDATE 172.16.112.0/24, next 172.16.20.2, path
```

```
BGP: 172.16.20.1 send UPDATE 172.16.220.0/24, next 172.16.20.2, path
BGP: 172.16.20.1 send UPDATE 192.68.222.0/24, next 172.16.20.2, metric 5000, path 3 2
BGP: 172.16.20.1 4 updates enqueued (average=58, maximum=68) BGP: 172.16.20.1 update run
completed, ran for 24ms, neighbor version 0, start version 23, throttled to 23, check point net
0.0.0.0
BGP: 172.16.20.1 rev UPDATE about 192.68.11.0/24, next hop 172.16.20.1, path 1 metric 2000
BGP: 172.16.20.1 rev UPDATE about 192.68.222.0/24, next hop 172.16.20.1, path 1 2 metric 2000
BGP: 172.16.20.1 rev UPDATE about 172.16.25.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.30.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.50.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.60.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.70.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.90.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.65.0/26 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.112.0/24 - denied
BGP: 172.16.20.1 rev UPDATE about 172.16.220.0/24 - denied
BGP: nettable_walker 192.68.11.0/255.255.255.0 calling revise_route
BGP: revise route installing 192.68.11.0/255.255.255.0 -> 172.16.20.1
BGP: 172.16.70.2 computing updates, neighbor version 23, table version 24, starting at 0.0.0.0
BGP: NEXT_HOP part 1 net 192.68.11.0/24, neigh 172.16.70.2, next 172.16.20.1
BGP: 172.16.70.2 send UPDATE 192.68.11,0/24, next 172.16.20.1, metric 2000, path 1
BGP: 172.16.70.2 1 updates enqueued (average=59, maximum=59)
BGP: 172.16.70.2 update run completed, ran for 4ms, neighbor version 23, start version 24,
throttled to 24, check point net 0.0.0.0
BGP: 172.16.20.1 rev UPDATE about 172.16.25.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.30.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.50.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.60.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.70.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.90.0/24 •- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.65.0/26 -- withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.112.0/24 - withdrawn
BGP: 172.16.20.1 rev UPDATE about 172.16.220.0/24 - withdrawn
BGP: 172.16.20.1 computing updates, neighbor version 23, table version 24, starting at 0.0.0.0
BGP: 172.16.20.1 update run completed, ran for Oms, neighbor version 23, start version 24,
throttled to 24, check point net 0.0.0.0
BGP: scanning routing tables
```

Как видно из листинга 12.84, при сбросе TCP-сеанса между двумя маршрутизаторами ведется обмен огромными объемами информации, и сеанс начинает повторно устанавливаться с самого начала.

Примечание

Подобная нагрузка на маршрутизатор также будет заметна, если вы просмотрите нагрузку на процессор или текущую информацию о BGP с помощью команд show process cpu или show ip bgp sum. Вы сможете увидеть реальную нагрузку на процессор маршрутизатора и изменения в BGP-таблице, происходящие с маршрутами. Кроме того, вы сможете наблюдать активный обмен маршрутной информацией со всеми взаимодействующими узлами (это происходит обновление маршрутных таблиц).

Далее в отчете о работе системы вы увидите, что BGP-сеанс сброшен, затем выбор взаимодействующего узла переходит из состояния ожидания в фазу "Соединение установлено", после чего происходит обмен информацией об обновлении маршрутов.

В листинге 12.85 показан тот же самый сброс сеанса, но уже с использованием мягкой перенастройки. Обратите внимание, что метрика 5000 посылается без сброса сеанса BGP, и общая нагрузка на маршрутизатор намного ниже.

Листинг 12.85. *Мягкая перенастройка по информационной базе входящих маршрутов (сброс BGP-сеанса между двумя маршрутизаторами с применением мягкой перенастройки)*

```
RTA#clear ip bgp 172.16.29.1 soft out
BGP: start outbound soft reconfiguration for 172.16.20.1
BGP: 172.16.20.1 computing updates, neighbor version 0, table version 24, starting at 0.0.0.0
BGP: 172.16.20.1 send UPDATE 172.16.25.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.30.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.50.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.60.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.70.0/24, next 172.16,20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.90.0/24, next 172.16.20.2, netric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.65.0/26, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.112.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 172.16.220.0/24, next 172.16.20.2, metric 5000, path 3
BGP: 172.16.20.1 send UPDATE 192.68.11.0/24 — unreachable
```

```
BGP: 172.16.20.1 send UPDATE 192.68.222.0/24, next 172.16.20.2, metric 5000, path 3 2
BGP: 172.16.20.1 5 updates enqueued (average=52, maximum=68)
BGP: 172.16.20.1 update run completed, ran for 24ras, neighbor version 0, start version 24,
throttled to 24, check point net 0,-0.0.0
BGP: scanning routing tables
```

Обновление BGP-маршрутов

Функция обновления BGP-маршрутов (BGP Route Refresh) входит в состав возможностей протокола BGP (которые мы обсуждали в главе 5). При установлении соединения по протоколу BGP взаимодействующие узлы объявляют о поддержке функции обновления маршрутов. Если они поддерживают эту функцию, то могут динамически запрашивать удаленный узел о повторном объявлении его базы Adj-RIB-Out (что обычно происходит и при выполнении мягкой перенастройки по информационной базе исходящих маршрутов). Так как при этом для хранения информации дополнительной памяти не требуется, этот способ более эффективен, чем мягкая перенастройка, и не вызывает дополнительных колебаний маршрутов, как при сбросе BGP-сеанса. Когда запрашивающая сторона получает базу Adj-RIB-Out, то передает ее входным правилам маршрутизации (и новым правилам маршрутизации).

Из листинга 12.86 видно, что поддержка функции обновления маршрутов определяется путем проверки выражения Neighbor capabilities в результате выполнения команды show ip bgp neighbor x.x.x.x.

Листинг 12.86. *Подтверждение о выполнении обновления BGP-маршрутов*

```
R1#show ip bgp n 1.1.2.2
BGP neighbor is 1.1.2.2,  remote AS 2, external link
 BGP version 4, remote router ID 3.3.3.1
 BGP state = Established, up for 2wOd
 Last read 00:00:15, hold time is 180, keepalive interval is 60 seconds
 Neighbor capabilities:
  Route refresh: advertised and received
  Address family IPv4 Unicast: advertised and received
 Received 20674 messages, 0 notifications, 0 in queue
 Sent 20675 messages, 0 notifications, 0 in queue
 Route refresh request: received 1, sent 2
 Minimum time between advertisement runs is 30 seconds

For address family: IPv4 Unicast
 BGP table version 6, neighbor version 6
 Index 1, Offset 0, Mask 0x2
 NEXT_HOP is always this router
 Community attribute sent со this neighbor
 1 accepted prefixes consume 36 bytes
 Prefix advertised 4, suppressed 0, withdrawn 0

 Connections established 1; dropped 0
 Last reset never
Connection state is ESTAB, I/O status: 1, unread input bytes: 0
Local host: 1.1.2.1, Local port: 179
Foreign host: 1.1.2.2, Foreign port: 11000
Enqueued packets for retransmit: 0, input: 0 mis-ordered: 0 (0 bytes)

Event Timers (current time is 0x49ED7420):
Timer Starts      Wakeups     Next
Retrans     20675 0     0x0
TimeWait    0     0     0x0
AckHold     20674 19530 0X0
SendWnd     0     0     0x0
KeepAlive   0     0     0x0
GiveUp      0     0     0x0
PmtuAger    0     0     0x0
DeadWait    0     0     0x0
```

```
iss: 1081723559 snduna: 1082116474 sndnxt: 1082116474    sndwnd: 15567
irs: 1087514066 rcvnxt: 1087906928 rcvwnd:      15605  delrcvwnd:    779

SRTT: 301 ms, RTTO: 621 ms, RTV: 9 ms, KRTT: 0 ms
minRTT: 4 ms, maxRTT: 600 ms, ACK hold: 200 ms
Flags: passive open, nagle, gen tcbs

Datagrams (max data segment is 1460 bytes):
Rcvd: 39791 (out of order: 0), with data: 20674, total data bytes: 392861
Sent: 40473 (retransmit: 0), with data: 20674, total data bytes: 392914
```

Как видите, обновление маршрутов поддерживается только одной стороной. Это видно из записи Route Refresh, где указано количество объявленных **(advertised)** и/или принятых маршрутов **(received).**

В листинге 12.87 показано, как запросить удаленный узел о повторном объявлении его базы Adj-RIB-Out.

Листинг 12.87. *Принуждение удаленного узла к повторному объявлению базы Adj-RIB-Out*

```
rl# clear ip bgp 1.1.2.2 soft in
rl#
2w0d: BGP: 1.1.2.2 sending REFRESH^REQ for afi/safi: 1/1
2w0d: BGP: 1.1.2.2 send message type 128, length (incl. header) 23
2w0d: BGP: 1.1.2.2 send message type 4, length {incl. header) 19
2w0d: BGP: 1.1.2.2 rev message type 4, length (excl. header) 0
```

Таким образом вы посылаете взаимодействующему узлу запрос на обновление маршрутов и на повторное объявление базы Adj-RIB-Out.

Из листинга 12.86, где представлены результаты выполнения команды **show ip bgp neighbor,** видно, что существует также несколько счетчиков, которые фиксируют количество запросов на обновление маршрутов, поступивших и переданных узлом.

Организация работы маршрутизатора в режиме фильтра исходящих BGP-маршрутов

Функция фильтра исходящих BGP-маршрутов (Outbound Route Filter — ORF) также предусмотрена возможностями BGP (которые мы обсуждали в главе 5). Эта функция способствует сбережению ресурсов системы во время обработки обновлений BGP-маршрутов. Если функция ORF объявляется соседним узлом в течение фазы установки сеанса, то это означает, что локальный BGP-спикер разрешит соседнему узлу передавать сведения через его фильтр исходящих префиксов. После того как эти данные получены, локальный BGP-спикер устанавливает фильтр дополнительно к исходящим фильтрам, установленным ранее для соседнего узла. Использование ORF BOP имеет неоспоримые преимущества.

- Локальный BGP-спикер не тратит ресурсы на генерацию обновлений маршрутов, которые фильтруются соседним узлом.
- Обновления маршрутов не занимают полезную полосу пропускания.
- Соседнему маршрутизатору также не нужно затрачивать ресурсы на обработку обновлений маршрутов, которые просто отвергаются при анализе фильтром.

Примечание

Очень важно понимать, что применение фильтра BGP ORF может оказать очень серьезное влияние на работу сети. Например, если BGP-узел поспал глобальную таблицу маршрутов сети Internet (а это более 75000 маршрутов) на низкопроизводительный маршрутизатор, у которого недостаточно памяти, то это может привести к "зависанию" последнего вследствие недостатка памяти даже при установленных входных фильтрах. Кроме того, такая операция может вызвать перегрузку полосы пропускания на низкоскоросгных соединениях.

Функция фильтра исходящих запросов идеально подходит для провайдеров Internet, у которых развиты инструменты автоматизированной генерации фильтров маршрутов и маршрутизаторы связаны с огромным количеством BGP-узлов. Если фильтр сгенерирован и развернут на границе сети провайдера, его действие распространяется и на маршрутизатор клиента, что позволяет сберегать ресурсы и провайдера, и его клиента.

По умолчанию функция ORF не объявляется взаимодействующим узлам. Ее также нельзя реализовать на узле, который уже является членом группы взаимодействующих узлов.

В результате того, что синтаксис для задания BGP ORF в различных версиях IOS очень сильно изменялся, мы решили включить в книгу Приложение В, "Фильтр исходящих маршрутов BGP", где содержатся все информационные бюллетени касающиеся поддержки BGP ORF в оборудовании Cisco. Дополнительную информацию о возможности использования этой функции вы можете найти в документации на вашу версию Cisco IOS.

Разгрузка маршрутов

Разгрузка маршрутов (route dampening) — это механизм, который используется в целях минимизации нестабильностей, вызванных колебаниями маршрутов в сети. Для управления процессом разгрузки маршрутов используется следующая команда:

```
bgp dampening [[route-map map-name] [half-life-time reuse-value
sup-press-value maximum-suppress-time] ]
```

Ниже приведены диапазоны изменения различных параметров.

- *half-life-time* — интервал времени от 1 до 45 минут. Значение по умолчанию — 15 минут.
- *reuse-value* — от 1 до 20000. По умолчанию — 750.
- *suppress-value* — от 1 до 20000. По умолчанию — 2000.
- *maximum-suppress-time* — максимальное время, в течение которого может подавляться маршрут. Оно изменяется в диапазоне от 1 до 255. По умолчанию оно равно 4х *half-life-time.*

Карта маршрутов может быть связана с разгрузкой маршрутов для избирательной применения параметров разгрузки по заданному критерию. Примером такого критери; может служить определенный IP-маршрут, атрибуты AS_PATH или COMMUNITY.

На рис. 12.15 показаны две AS, AS1 и AS3. Маршрутизатор RTA в AS3 поддержи вает протокол 1BGP при взаимодействии с маршрутизатором RTG и протоко; EBGP — для работы с RTC в AS1. Информация, получаемая от AS3 по EBGP, переда ется протоколу OSPF в AS1.

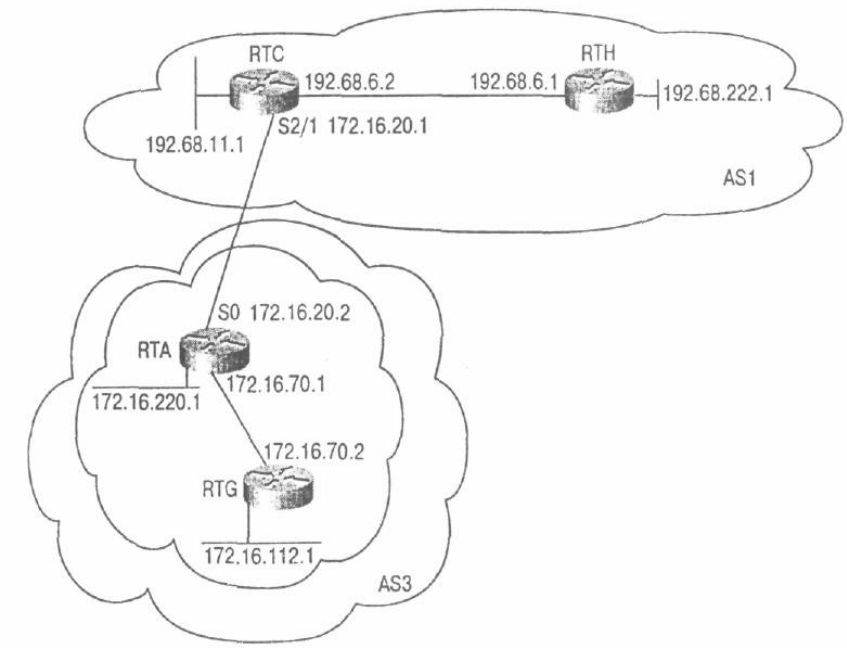


Рис. 12.15. *Разгрузка маршрутов*

Допустим, что маршрутизатор RTC отмечает большое число колебании в сеп 172.16.220.0/24, сведения о которой поступают от AS3, что вызывает колебания в BGP н; самом RTC и, следовательно, в OSPF. При этом маршрут в сеть 172.16.220.0/24 постоянн< появляется и исчезает из маршрутной таблицы маршрутизатора RTH. Чтобы исправит положение, на RTC применяется разгрузка BGP-маршрутов с использованием картъ маршрутов, благодаря чему разгружается только маршрут 172.16.220.0/24. В листингах 12.8; и 12.89 приведены конфигурации маршрутизаторов RTG и RTAjuw этого случая.

Листинг 12.88. *Разгрузка маршрутов (конфигурация маршрутизатора RTG)*

```
router bgp 3
 no synchronization
 network 172.16.112.0 mask 255.255.255.0
 neighbor 172.16.70.1 remote-as 3
no auto-summary
```

Листинг 12.89. *Разгрузка маршрутов (конфигурация маршрутизатора RTA)*

```
router bgp 3
 no synchronization
 network 172.16.220.0 mask 255.255.255.0
 network 172.16.70.0 mask 255.255.255.0
 neighbor   172.16.20.1 remote-as   1
 neighbor  172.16.70.2  remote-as  3
 neighbor   172.16.70.2 next-hop-self
no  auto-summary
```

Итак, маршрутизатор RTC взаимодействует по протоколу EBGP с RTA и по IBGP — с RTH. Все принимаемые маршруты RTC далее передает с помощью протокола OSPF по всей AS1. На этом маршрутизаторе используется карта маршрутов SELECTIVE_DAMPENING, с помощью которой параметры разгрузки применяются только к маршруту в сеть 172.16.220.0/24. Все остальные маршруты, такие как 172.16.112.0/24, разгружаться не будут.

В конфигурации маршрутизатора RTC, приведенной в листинге 12.90, задаются следующие параметры разгрузки.

- Интервал времени *half-life-time* — 20 минут.
- Повторное использование маршрута ограничено 950.
- Маршруты будут подавляться, если сумма штрафов превысит 2500.
- Маршрут может подавляться не более 80 минут.

Листинг 12.90. *Разгрузка маршрутов (конфигурация маршрутизатора RTC)*

```
router ospf 10
 redistribute bgp 1 subnets
 network 192.68.0.0 0.0.255.255 area 0

router bgp 1
 bgp dampening route-map SELECTIVE_DAMPENING
 network 192.68.11.0
 neighbor 172.16.20.2 remote-as 3
 neighbor 192.68.6.1 remote-as 1
 no auto-summary

access-list 1 permit 172.16.220.0 0.0.0.255

route-map SELECTIVEJDAMPENING permit 10
 match ip address 1
 set dampening 20 950 2500 80

route-map SELECTIVE_DAMPENING permit 20
```

Как видно из листинга 12.90, маршрутизатор RTC обрабатывает только колебания маршрута 172.16.220.0/24. Колебанием маршрута считается любое изменение информации о нем. В листинге 12.91 представлена BGP-таблица до колебания маршрута.

Листинг 12.91. *Разгрузка маршрутов (BGP-таблица на маршрутизаторе RTC перед колебанием маршрута)*

```
RTC#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 326
Paths: (1 available, best #1, advertised over IBGP)
 3
172.16.20.2 from 172.16.20.2 (172.16.220.1)
 Origin IGP, metric 0, valid, external, best
```

В листинге 12.92 показано состояние маршрута после колебания. Маршрут не используется и переведен в состояние **history.** По умолчанию маршруту задан штраф 1000, который уже "возмещен"[1] в размере 997.

Листинг 12.92. *Разгрузка маршрутов (BGP-таблица на маршрутизаторе RTC после первого колебания маршрута)*

```
RTC#show ip bgp 172.16.220.8
BGP routing table entry for 172.16.220.0/24, version 327
Paths: (1 available, no best path, advertised over IBGP)
 3 (history entry)
 172.16.20.2 from 172.16.20.2 (172.16.220.1)
 Origin IGP, metric 0, external
Dampinfo: penalty 997, flapped 1 times in 00:00:06
```

В листинге 12.93 показано, как выглядит информация о маршруте после второго колебания (он снова возвращается в рабочее состояние). Здесь снова добавляется штраф 1000, и после его частичного погашения суммарная величина штрафа составляет 1454.

Листинг 12.93. *Разгрузка маршрутов (BGP-таблица на маршрутизаторе RTC после второго колебания маршрута)*

```
RTC#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 328
Paths: (1 available, best #1, advertised over IBGP)
 3
 172.16.20.2 from 172.16.20.2 (172.16.220.1)
 Origin IGP, metric 0, valid, external, best
Dampinfo: penalty 1454, flapped 2 times in 00:01:20
```

В листинге 12.94 показано состояние маршрута после четырех колебаний. Теперь величина штрафа составляет 2851, что превышает установленный лимит 2500. Теперь маршрут подавляется (разгружается) и не передается маршрутизатору RTH. Маршрут будет снова доступен для использования через 31 минуту и 40 секунд. В то же время штраф будет снижать величину интервала повторного использования до 950.

Листинг 12.94. *Разгрузка маршрутов (BGP-таблица на маршрутизаторе ЯТС после четырех колебаний маршрута)*

```
RTC#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 329
Paths: (1 available, no best path, advertised over IBGP}
 3, (suppressed due to dampening)
  172.16.20.2 from 172.16.20.2 (172.16.220.1)
  Origin IGP, metric 0, valid, external
Dampinfo: penalty 2851, flapped 4 times in 00:03:05, reuse in 00:31:40
```

Листинг 12.95 отображает состояние маршрута после шести колебаний. Различие состоит в том, что время *half-life-time* здесь составляет уже 5 минут, а не 20 минут, А время *maximum-suppress-time* составляет 20, а не 80 минут. За более короткое время *half-fife-time* штраф будет возмещен намного быстрее, и маршрут может использоваться гораздо раньше. Обратите внимание, что интервал времени повторного использования теперь составляет всего 8 минут 10 секунд.

Листинг 12.95. *Разгрузка маршрутов (BGP-таблица на маршрутизаторе RTC после шести колебаний маршрута)*

```
RTC#show ip bgp 172.16.220.0
BGP routing table entry for 172.16.220.0/24, version 336
Paths: (1 available, no best path, advertised over IBGP)
 3, (suppressed due to dampening)
  172.16.20.2 from 172.16.20.2 (172.16.220.1)
  Origin IGP, metric 0, valid, external
Dampinfo: penalty 2939, flapped 6 times in 00:08:21, reuse in 00:08:10
```

Изменение настроек таймеров, управляющих разгрузкой, становится крайне необходимым, если администраторы не могут позволить долго отсутствовать тому или иному маршруту. Разгрузка маршрутов в BGP с помощью карт маршрутов представляет собой мощнейшее средство для избирательного подавления аномальных маршрутов, которое позволяет управление и настройку со стороны пользователя.

Забегая вперед

Сеть Internet прошла длинный путь от магистральных каналов NSFNET к информационным трансконтинентальным магистралям 21 века, причем нет никаких признаков снижения темпов ее роста. Да и откуда они могли бы появиться, если каждый день в режиме он-лайн находятся тысячи пользователей, привлеченные приложениями, создаваемыми с помощью информационных технологий?

Протоколы маршрутизации, начиная от раннего EGP до последних версий BGP, подвергаются серьезным испытаниям, поскольку запросы постоянно растут. Протокол BGP также задумывался как простой протокол для внешней маршрутизации, но со временем превратился в стандарт де-факто, который по сути "склеивает" Internet в одну сеть. На самом деле все приемы и уловки, которые предлагаются в BGP, уже давно были использованы, но каждый день мы сталкиваемся с необходимостью реализации новых возможностей на его базе. В результате рождаются новые протоколы и новые технические приемы. Делают ли они

маршрутизацию более легкой и универсальной, пока невозможно сказать с полной уверенностью. Единственное можно заявить совершенно точно: пока в основе ваших разработок лежат эти протоколы, вы всегда будете хозяином в своем домене.

Так как мы, к сожалению, и не смогли бы рассмотреть здесь все потенциальные преимущества и недостатки каждой команды IOS для BGP, рекомендуем вам подробно изучить соответствующую документацию на оборудование компании Cisco или обратиться в ее службу технической поддержки, если возникшая проблема не решается. В Приложение А, "Справочник по командам BGP", мы включили перечень параметров IOS BGP, которые доступны в настоящее время. В Приложении Б, "Ссылки для дальнейшего изучения", вы найдете ссылки на интересные ресурсы и книги. Приложение В, " Фильтр исходящих BGP-маршрутов (Outbound Route Filter — ORF)", содержит информацию об организации фильтра исходящих BGP-маршрутов. И, наконец, в Приложении Г, " Мультипротокольные расширения BGP (Multiprotocol BGP — MBGP)", приводятся данные по мультипротокольным расширениям BGP (Multiprotocol BGP — MBGP).

Часть V. Приложения

В этой части...

Приложение А. Справочник по командам BGP
Приложение Б. Ссылки для дальнейшего изучения
Приложение В. Фильтр исходящих BGP-маршрутов (Outbound Route Filter - ORF)
Приложение Г. Мультипротокольные расширения BGP (Multiprotocol BGP — MBGP)

В приложениях мы приводим дополнительные материалы для использования в качестве справочного пособия. Здесь вы найдете справочник по BGP-командам для Cisco IOS™ и сведения о различных модификациях этой операционной системы.

Приложение А. Справочник по командам BGP

Все команды, приведенные в табл. А-1, представляют собой команды для настройки и управления маршрутизацией по протоколу граничного шлюза (Border Gateway Protocol — **BGP).** Полное описание этих команд, в том числе их полный синтаксис (аргументы, параметры и т.п.), значения по умолчанию, режимы использования, основы и примеры применения, а также связанные с ними команды вы сможете найти в Internet на сервере компании Cisco по адресу:

www.cisco.com/uniyercd/cc/td/doc/product/software/iosl21/121cgcr/ip_r/ iprpt2/lrdbgp.htm#xtocidl41910.

Таблица А-1. *BGP-команды*

Команда	Описание
aggregate-address	Эта команда применяется для создания записи об объединенном маршруте в таблице BGP-маршрутов. Чтобы прекратить действие этой команды, добавьте к ней ключевое слово **no**
auto-summary (BGP)	Эта команда служит для восстановления автоматического суммирования маршрутов подсетей в сетевые маршруты. Чтобы предотвратить автосуммирование и разрешить пересылку маршрутной информации о подсетях через классовые сети добавьте ключевое слово **no**
bgp always-compare-med	Чтобы разрешить сравнение атрибутов Multi Exit Discriminator (MED) для маршрутов, поступающих от соседних узлов, которые принадлежат различным автономным системам, используется команда **bgp always-compare-med**. Запретить сравнение этих атрибутов можно, добавив к команде слово **no**
bgp bestpath as-path ignore	Команда **bgp bestpath as-path ignore** позволяет не допустить, чтобы маршрутизатор опирался при выборе маршрута на значение **as-path**. Отменить ее действие вы можете, указав ключевое слово **no**
bgp bestpath med-confed	Эта команда разрешает проводить сравнение MED маршрутов, полученных от различных узлов, входящих в конфедерацию. Чтобы запретить анализ атрибута MED при сравнении маршрутов добавьте ключевое слово **no**
bgp bestpath missing-as -worst	Чтобы заставить Cisco IOS обрабатывать отсутствие атрибута MED в маршруте как значение равное бесконечности, что позволяет рассматривать маршрут как наименее желательный, используется команда конфигурации **bgp bestpath missing-as-worst**. Чтобы вернуть настройки маршрутизатора в исходное состояние, задайте эту команду с ключевым словом **no**
bgp client-to-client reflection	Для восстановления отражения маршрутов от отражателя BGP-маршрутов к клиентам следует воспользоваться командой **bgp client-to-client reflection**. Чтобы запретить ее действие, задайте эту команду с ключевым словом **no**
bgp cluster-id	Для конфигурирования идентификатора кластера, если в кластере BGP имеется более одного отражателя маршрутов, используйте команду **bgp cluster-id**. Чтобы прекратить ее действие, задайте эту же команду с ключевым словом **no**
bgp confederation identifier	С целью указания идентификатора конфедерации используется команда **bgp confederation identifier**. Чтобы удалить идентификатор конфедерации, задайте эту команду с ключевым словом **no**

bgp confederation peers	Для настройки автономных систем, принадлежащих к одной конфедерации, используйте команду **bgp confederation peers**, Чтобы удалить из конфедерации автономную систему, задайте эту команду с ключевым словом **no**
bgp dampening	Чтобы разрешить разгрузку маршрутов или изменить параметры разгрузки, можно воспользоваться глобальной командой конфигурации **bgp dampening**. Если нужно запретить разгрузку или восстановить параметры по умолчанию, задайте эту команду с ключевым словом **no**
bgp default local-preference	Чтобы изменить значение локального предпочтения, заданное по умолчанию, используйте команду **bgp default local-preference**. Для возвращения значения по умолчанию задайте эту команду с ключевым словом **no**
bgp deterministic med	Для того чтобы Cisco IOS при выборе маршрута из набора маршрутов, объявляемых подсистемой автономной системы внутри конфедерации, могла сравнивать атрибуты MED, используйте команду **bgp deterministic med**. Запретить сравнение атрибутов вы можете, задав эту команду с ключевым словом **no**
bgp fast-external-failover	Чтобы немедленно сбросить BGP-сеансы с соседними узлами в случае выхода из строя канала связи с ними, воспользуйтесь командой **bgp fast-external-failover**
bgp log-nieghbor-changes	Разрешить ведение отчетов о сбросах сеансов с соседними BGP-узлами можно с помощью команды **bgp log-nieghbor-changes.** Запретить ведение отчетов о работе с соседними узлами можно, воспользовавшись этой же командой с ключевым словом **no**
clear ip bgp	Сброс BGP-соединения с мягкой перенастройкой выполняется с помощью выполняемой команды **EXEC clear ip bgp**
Команда clear ip bgp dampening	Сброс информации о разгрузке маршрутов и прекращение подавления маршрутов выполняется с помощью команды EXEC **clear ip bgp dampening**
clear ip bgp flap-statistics	Сброс имеющейся статистики о колебаниях маршрутов выполняется с помощью команды **EXEC clear ip bgp flap-statistics**
clear ip bgp peer-group	Удаление всех членов из группы взаимодействующих BGP-узлов выполняется с помощью команды **EXEC clear ip bgp peer-group**
clear Ip prefix-list	Сброс счетчика записей списка префиксов выполняется с помощью команды **EXEC clear ip prefix-list**
default-Information originate (BGP)	Чтобы в сети 0.0.0.0 применялся протокол BGP, воспользуйтесь командой конфигурации маршрутизатора **default-information originate**. Запретить выполнение этой функции можно с помощью этой же команды, заданной с ключевым словом **no**
default-metric (BGP)	Чтобы установить значения метрик по умолчанию для протокола BGP, используйте команду конфигурации **default-metric**. Вернуть исходные значения можно, задав эту команду с ключевым словом **no**
distance bgp	Для разрешения использования внешних, внутренних и локальных административных дистанций, которые могут обеспечить выбор лучшего маршрута к заданному узлу, используйте команду **distance bgp**. Ее действие прекращается при задании ключевого слова **no**
distribute-list in	С целью фильтрации сведений о сетях используйте команду distribute-list in. Запретить эту функцию можно с помощью этой же команды, заданной с ключевым словом **no**

dIstribute-list out	Чтобы не допустить объявления определенных сетей в исходящих обновлениях маршрутов, воспользуйтесь командой **distribute-list out**. Прекратить действие этой функции можно, задав эту команду с ключевым словом **no**
Ip as-path access-list	Для задания списка разрешения доступа по BGP используйте глобальную команду конфигурации маршрутизатора **ip as-path access-list**. Чтобы запретить действие списка разрешения доступа, задайте эту команду с ключевым словом **no**
ip bgp-community new-format	Чтобы отображать BGP-сообщества в формате AA:NN (номер сообщества автономных систем/2-байтовый номер), используйте команду **ip bgp-community new-format**. Чтобы вернуться к предыдущему формату отображения номеров сообществ, задайте эту команду с ключевым словом **no**
ip community-list	Для создания списка сообществ в BGP и управления доступом к ним воспользуйтесь командой **ip community-list** Чтобы удалить список сообществ, задайте эту команду с ключевым словом **no**
ip prefix-list	Для создания записи в списке префиксов используйте команду **ip prefix-list**. Удалить запись вы можете с помощью этой же команды, заданной с ключевым словом **no**
ip prefix-list description	Чтобы добавить текстовое описание списка префиксов, используйте команду конфигурации **ip prefix-list description**. Чтобы удалить текстовое описание списка, задайте эту же команду с ключевым словом **no**
ip prefix-list sequence-number	Чтобы разрешить генерирование последовательности номеров для записей в списке префиксов, используйте глобальную команду конфигурации **ip prefix-list sequence-number**
match as-path	Для сопоставления маршрута автономной BGP-системы со списком разрешения доступа воспользуйтесь командой конфигурации карты маршрутов **match as-path**. Чтобы удалить запись о маршруте из списка, задайте эту команду с ключевым словом **no**
match community-list	Чтобы провести сопоставление BGP-сообщества, используйте команду **match community-list**. Чтобы удалить запись из списка сообществ, задайте эту команду с ключевым словом **no**
neighbor advertisement-interval	Установить минимальный интервал между посылкой обновлений маршрутов в BOP можно с помощью команды n**eighbor advertisement-interval.** Чтобы удалить запись, задайте эту команду с ключевым словом **no**
neighbor default-originate	Разрешить BGP-спикеру (локальному маршрутизатору) посылать маршрут по умолчанию 0.0.0.0 соседним узлам можно с помощью команды **neighbor default-originate**. Запретить посылать маршрут по умолчанию можно, задав затем эту команду с ключевым словом **no**
neighbor description	Чтобы задать описание соседнего узла, используйте команду **neighbor description**. Удалить описание узла можно с помощью этой же команды, заданной с ключевым словом **no**
neighbor distribute-list	Чтобы распределить информацию, как указано в списке разрешения доступа, используйте команду **neighbor distribute-list**. Чтобы удалить запись из списка, задайте эту команду с ключевым словом **no**
neighbor ebgp-multihop	Чтобы разрешить организацию BGP-соединений с внешними узлами, которые находятся в сетях, не имеющих непосредственного соединения, воспользуйтесь командой **neighbor ebgp-multihop**. Вернуть настройки по умолчанию можно с помощью этой же команды, заданной с ключевым словом **no**

neighbor filter-list	Установить BGP-фильтр можно с помощью команды **neighbor filter-list**. Запретить его использование можно этой командой с ключевым словом **no**
neighbor maximum-prefix	Для ограничения количества префиксов, принимаемых от соседнего узла, воспользуйтесь командой конфигурации neighbor maximum-prefix. Запретить действие этой функции можно, задав эту команду с ключевым словом **no**
neighbor next-hop-self	Чтобы запретить обработку обновлений маршрутов на основе адреса следующего ближайшего узла, используйте команду конфигурации **neighbor next-hop-self**. Прекратить действие этой команды можно, задав ее с ключевым словом **no**
neighbor password	Разрешить аутентификацию TCP-соединения между двумя BGP-узлами с использованием алгоритма Message Digest 5 (MD5) можно с помощью команды конфигурации маршрутизатора **neighbor password**. Запретить эту функцию можно, задав эту команду с ключевым словом **no**
neighbor peer-group (назначение членов группы)	Чтобы сконфигурировать соседний BGP-узел в качестве члена группы взаимодействующих узлов, воспользуйтесь командой **neighbor peer-group**. Удалить узел из группы можно с помощью этой команды, заданной с ключевым словом **no**
neighbor peer-group (создание)	Для создания группы BGP-узлов используйте команду **neighbor peer-group**. Чтобы удалить группу узлов и всех ее членов, задайте эту команду с ключевым словом **no**
neighbor prefix-list	Для распределения информации, как указано в списке префиксов, воспользуйтесь командой **neighbor prefix-list.** Чтобы удалить запись из списка, задайте эту команду с ключевым словом **no**
neighbor remote-as	Чтобы добавить запись в таблицу соседних BGP-узлов, воспользуйтесь командой **neighbor remote-as.** Удалить запись из таблицы можно, воспользовавшись этой же командой с ключевым словом **no**
neighbor route-map	Чтобы ко входящим или исходящим маршрутам применить карту маршрутов, задайте команду **neighbor route-map**. Удалить карту маршрутов можно этой же командой с ключевым словом **no**
neighbor route-reflector-client	Чтобы настроить маршрутизатор для работы в качестве отражателя маршрутов и сконфигурировать заданный соседний узел как его клиента, используйте команду **neighbor route-reflector-client**. Чтобы отметить, что соседний узел не является клиентом, задайте эту команду с ключевым словом **no**. Когда все возможные клиенты будут запрещены, локальный маршрутизатор перестанет выполнять функцию отражателя маршрутов
neighbor send-community	Для задания пересылки атрибута COMMUNITY на соседний узел используйте команду **neighbor send-community**. Чтобы удалить запись, задайте эту же команду с ключевым словом **no**
neighbor shutdown	Чтобы запретить работу с соседним узлом или с группой взаимодействующих узлов, воспользуйтесь командой **neighbor shutdown**. Разрешить работу с соседним узлом или с группой можно с помощью этой же команды, заданной с ключевым словом **no**
neighbor soft-reconfiguration	Настроить Cisco IOS для сохранения информации о принимаемых обновлениях маршрутов можно с помощью команды **neighbor soft-reconfiguration**. Прекратить сохранение принимаемых обновлений маршрутов можно, задав эту команду с ключевым словом **no**

neighbor timers	Установить таймеры для определенных BGP-узлов или группы узлов можно с помощью команды **neighbor timers**. Сбросить все таймеры для заданных узлов или группы узлов можно, задав эту команду с ключевым словом **no**
neighbor update-source	Чтобы разрешить Cisco IOS использовать для внутренних BGP-сеансов любой рабочий интерфейс для организации TCP-соединений, воспользуйтесь командой **neighbor update-source**. Чтобы восстановить первоначальные настройки ближайшего интерфейса, который также называется *наилучшим локальным адресом (best local address),* задайте эту команду с ключевым словом **no**
neighbor version	Чтобы Cisco 1OS работала только с определенной версией протокола BGP, задайте команду конфигурации **neighbor version**. Использовать версию протокола, установленную по умолчанию, можно, задав эту команду с ключевым словом **no**
neighbor weight	Для назначения веса соединению с соседним узлом воспользуйтесь командой **neighbor weight**. Чтобы отменить назначение веса, задайте эту команду с ключевым словом **no**
network (BGP)	Чтобы задать список сетей, которые будут участвовать в процессе маршрутизации по протоколу BGP, воспользуйтесь командой **network**. Чтобы удалить адрес сети из списка, задайте эту команду с ключевым словом **no**
network backdoor	Для задания обходного маршрута к граничному BGP-маршрутизатору, который обеспечил бы более полную информацию о сети, воспользуйтесь командой **network backdoor**. Удалить адрес сети из списка можно с помощью этой команды, заданной с ключевым словом **no**
network weight	Чтобы назначить BGP-сети абсолютный вес, используйте команду **network weight**. Удалить запись можно с помощью этой же команды, заданной с ключевым словом **no**
router bgp	Для настройки маршрутизации по BGP воспользуйтесь глобальной командой конфигурации **router bgp**. Чтобы запретить работу по BGP, задайте эту команду с ключевым словом **no**
set as-path	Чтобы модифицировать атрибут AS_PATH для BGP-маршрута, воспользуйтесь командой **set as-path**. Отменить модификацию AS_PATH можно с помощью этой же команды, заданной с ключевым словом **no**
set comm-list delete	Чтобы удалить сообщества, заданные в атрибуте COMMUNITY, из входящих или исходящих обновлений маршрутов, используйте **команду set comm-list delete.** Аннулировать действие этой команды можно, задав ее с ключевым словом **no**
set community	Для установки атрибута COMMUNITY необходимо использовать команду конфигурирования карт маршрутов **set community**. Чтобы удалить запись об этом атрибуте, следует задать команду с ключевым словом **no**
set dampening	Чтобы задать факторы, при которых следует начинать разгрузку маршрутов в BGP, воспользуйтесь командой **set dampening**. Запретить эту функцию можно, задав эту команду с ключевым словом **no**
set ip next-hop (BGP)	Для указания того, куда отправлять исходящие пакеты, которые соответствуют правилам маршрутизации, описанным в карте маршрутов, задайте команду **set ip next-hop**. Удалить запись можно, задав эту команду с ключевым словом **no**

set metric-type internal	Чтобы установить значение MED для префиксов, объявленных по протоколу внешнего граничного шлюза (Exterior Gateway Protocol — EBGP), совпадающее с метрикой следующего узла по протоколу внутреннего шлюза (Interior Gateway Protocol — IGP), воспользуйтесь командой **set metric-type internal.** Чтобы вернуть значения по умолчанию, задайте эту команду с ключевым словом **no**
set origin (BGP)	Установить код источника в BGP можно с помощью команды конфигурации карты маршрутов **set origin**. Удалить запись из карты маршрутов можно, воспользовавшись этой же командой с ключевым словом **no**
set weight	Указать вес для таблицы BGP-маршрутов можно с помощью команды **set weight**. Чтобы удалить запись о весе таблицы, задайте эту команду с ключевым словом **no**
show ip bgp	Чтобы отобразить все записи из таблицы BGP-маршрутов, задайте команду **EXEC show ip bgp**
show ip bgp cidr-only	Чтобы отобразить маршруты с ненатуральными масками (т.е. участвующие в бесклассовой междоменной маршрутизации -CIDR), воспользуйтесь привилегированной командой **EXEC show ip bgp cidr-only**
show ip bgp community	Отображение маршрутов, принадлежащих к каким-либо BGP-сообществам, осуществляется командой **EXEC show ip bgp community**
show ip bgp community-list	Отображение маршрутов, разрешенных в списке BGP-сообществ, осуществляется командой **EXEC show ip bgp community-list**
show ip bgp dampened-paths	Для отображения разгружаемых BGP-маршрутов задайте команду **EXEC show ip bgp dampened-paths**
show ip bgp filter-list	Для отображения маршрутов, которые соответствуют какому-либо списку фильтров, используйте привилегированную команду **EXEC show ip bgp filter-list**
show ip bgp flap-statistics	Для отображения статистики колебаний BGP-маршрутов используйте команду **EXEC show ip bgp flap-statistics**
show ip bgp inconsistent-as	Для отображения маршрутов с несогласованными источниками используйте команду **EXEC show ip bgp inconsistent-as**
show ip bgp neighbors	Для отображения информации о TCP- и BGP-соединениях с соседними узлами используйте команду **EXEC show ip bgp neighbors**
showip bgp paths	Для отображения всех BGP-маршрутов из базы данных используйте команду **EXEC show ip bgp paths**
show ip bgp peer-group	Для отображения информации о группах взаимодействующих BGP-узлов используйте команду **EXEC show ip bgp peer-group**
Show ip bgp summary	Для отображения маршрутов, соответствующих нормальному выражению, используйте привилегированную команду **EXEC show ip bgp summary**
Show ip prefix-list	Для отображения информации о списке префиксов или записей из списка префиксов используйте команду **EXEC show ip prefix-list**
Synchronization	Для синхронизации между протоколами BGP и IGP используйте команду **synchronization**. Чтобы разрешить в Cisco IOS объявление маршрутов к сетям без необходимости ожидания их формирования в IGP, можно задать эту команду с ключевым словом **no**

Table-map	Для модификации значений метки и метрики при обновлении таблиц IP-маршрутов при получении сведений о BGP-маршрутах используйте команду **table-map**. Запретить действие этой функции можно, задав эту команду с ключевым словом **no**
Timers bgp	Для подстройки таймеров в BGP можно использовать команду конфигурации маршрутизатора **timers bgp**. Сбросить таймеры в значения по умолчанию можно с помощью этой же команды с ключевым словом **no**

Приложение Б. Ссылки для дальнейшего изучения

Организации, регламентирующие работу в Internet

Таблица Б-1.

Русское название	Английское название	Адрес в Internet
Группа инженеров по развитию Г сети Internet	Internet Engineering Task Force (IETF)	www.ietf.org
Североамериканская группа по обслуживанию сети	North American Network Operations Group (NANOG)	www.nanog.org
Американский реестр адресов , Internet	American Registry for Internet Numbers (ARIN)	www.arin.net
Азиатско-тихоокеанский сетевой информационный центр	Asian;Pasific Network Information Center(APNIC)	www.apnic.net
Европейский сетевой координационный центр RIPE	RIPE Network Coordination Centre (NCC)	www.ripe.net
Корпорация по назначению имен и адресов в сети Internet	The Internet Corporation for Assigned Names and Numbers (ICANN)	www.icann.org
Организация по распределению адресов в сети Internet	Internet Assigned Numbers Authority (IANA)	www.iana.org
Объединенная ассоциация по анализу данных в сети Internet	Cooperative Association for Internet Data Analysis (CAIDA)	www.caida.org

Исследовательские и образовательные учреждения

Таблица Б-2. Исследовательские и образовательные учреждения

Русское название	Английское название	Адрес в Internet
Инициатива "Internet следующего поколения"	Next Generation Internet (NGI) Initiative	www.ngi.gov
Internet2	Internet2	www.internet2.edu
Проект Abilene	Abilene	www.ucaid.edu/abilene
Высокоскоростная магистральной сетевая служба	Very High Speed Backbone Network Service (vBNS)	www.vbns.net
Национальный научный фонд США	National Science Foundation	www.nsf.gov

Другие организации и документы

Таблица Б-3.

Русское название	Английское название	Адрес в Internet
Точка обмена информацией	Exchange Point information	www.ep.net
Отчет CIDR	The CIDR Report	www.employees.org/~tbates/cidr-report.html
Отчет о маршрутизации в Азиатско-Тихоокеанском регионе	Asian Pasific Routing Report	www.apnic.net/stats/bgp
Онлайновая служба поддержки соединений компании Cisco	Cisco Connection Online (CCO)	www.cisco.com

Книги

Книги по протоколу TCP/IP

Douglas E. Comer, *Internetworking with TCP/IP, Volume 1, Fourth Edition,* April 2000, Prentice Hall; ISBN 0130183806.
Jeff Doyle, *Routing TCP/IP, Volume I,* 1998, Cisco Press; ISBN 1578700418
W. Richard Stevens, *TCP/IP Illustrated, Volume 1,* January 1994, Addison-Wesley Publishing Company; ISBN 0201633469.
X. Richard Stevens and Gary R. Wright, *TCP/IP Illustrated, Volume* //, January 1995, Addison-Wesley Publishing Company; ISBN: 020163354X.
W. Richard Stevens, *TCP/IP Illustrated, Volume III,* January 1996, Addison-Wesley Publishing Company; ISBN: 0201634953.

Книги по организации маршрутизации

C. Huitema, *Routing in me Internet, Second Edition,* 1999, Prentice Hall; ISBN 0130226475,
OMoy, *OSPF, Anatomy of an Internet Routing Protocol,* 1998, Addison- Wesley Pub Co;

ISBN 0201634724.
R. Perlman, *Interconnections, Second Edition*, 1999, Addison-Wesley Pub Co; ISBN 0201634481.
Alvaro Retana, Don Slice, and Russ White, *Advanced IP Network Design*, 1999, Cisco Press; ISBN 1578700973.
J. Stewart, *BGP4: Inter-Domain Routing in the Internet*, 1998, Addison-Wesley Pub Co; ISBN 0201379511.

Документы, регламентирующие работу в сети Internet (Request for Comments — RFC)

Все документы RFC можно получить на Web-сервере организации инженеров сети Internet (IETF) по адресу www.ietf.org. Вся информация в этих документах представлена в двух видах — информационном (для сведения) и в виде стандартов. В настоящее время работа протокола BOP регламентируется следующим списком RFC:

K. Lougheed and Y. Rekhter, "A Border Gateway Protocol 3 (BGP-3)," RFC 1267, October 1991.
S. Willis and J. Burruss. "Definitions of Managed Objects for the Border Gateway Protocol (Version 3)," RFC 1269, October 1991.
Y. Rekhter, "Experience with the BGP Protocol," RFC 1266, October 1991.
Y. Rekhter, "BGP Protocol Analysis," RFC 1265, October 1991.
D. Haskin, "Default Route Advertisement in BGP2 And BGP3 Versions of the Border Gateway Protocol," RFC 1397, January 1993.
K. Varadhan, "BGP OSPF Interaction," RFC 1403, January 1993.
S. Willis, J. Burruss and J. Chu, "Definitions of Managed Objects for the Fourth Version of the Border Gateway Protocol (BGP-4) using SMIv2," RFC 1657, July 1994.
K. Varadhan, S. Hares and Y. Rekhter, "BGP4/IDRP for IP OSPF Interaction," RFC 1745, December 1994.
P. Traina, "BGP-4 Protocol Analysis," RFC 1774, March 1995.
P. Traina, "Experience with the BGP-4 Protocol," RFC 1773, March 1995.
Y. Rekhter and T. Li, "A Border Gateway Protocol 4 (BGP-4)," RFC 1771, March 1995.
D. Haskin, "A BGP/IDRP Route Server Alternative to a Full Mesh Routing," RFC 1863, October 1995.
J.Hawkinson and T. Bates, "Guidelines for Creation, Selection, and Registration of an Autonomous System (AS)," RFC 1930, March 1996.
P. Traina. "Autonomous System Confederations for BGP," RFC 1965, June 1996.
T. Bates and R. Chandra, "BGP Route Reflection An Alternative to Full Mesh IBGP," RFC 1966, June 1996.
E. Chen and T. Bates, "An Application of the BGP Community Attribute in Multi-home Routing," RFC 1998, August 1996.
R. Chandra, P. Traina and T. Li, "BGP Communities Attribute," RFC 1997, August 1996.
J. Stewart,!. Bates, R. Chandra and E.Chen, "Using a Dedicated AS for Sites Homed to a Single Provider," RFC 2270, January 1998.
T. Bates, R. Chandra, D. Katz and Y. Rekhter, "Multiprotocol Extensions for BGP-4," RFC 2283, February 1998.
C. Alaettinoglu, T. Bates, E. Gerich, D. Karrenberg, D. Meyer, M. Terpstra, C. Villamizar,

"Routing Policy Specification Language (RSPL)," RFC 2280, January 1998.
A. HefTernan, "Protection of BGP Sessions via the TCP MD5 Signature Option," RFC 2385, August 1998.
C. Villamizar, R. Chandra and R. Govindan, "BGP Route Flap Damping," RFC 2439, November 1998.
E. Chen and J. Stewart, "A Framework for Inter-Domain Route Aggregation," RFC 2519, February 1999.
P. Marques and F. Dupont, "Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing," RFC 2545, March 1999.
C. Alaettinoglu, T. Bates, E. Gerich, D. Karrenberg, D. Meyer, M. Terpstra, C. Villamizar, "Routing Policy Specification Language (RPSL)," RFC 2622, June 1999.
C. Alaettinoglu, D. Meyer, C. Orange, M. Prior, J. Schmitz, "Using RPSL in Practice," RFC 2650, August 1999.
C. Alaettinoglu, D. Meyer, S. Murphy, C. Villamizar, "Routing Policy System Security," RFC 2725, December 1999.
B. Aiken, B. Carpenter, I. Foster, C. Lynch, J. Mambretti, R. Moore, J. Strassner, B. Teitelbaum, "Network Policy and Services: A Report of a Workshop on Middleware." RFC 2768, February 2000.
T. Bates, R. Chandra and E. Chen, "BGP Route Reflection An Alternative to Full Mesh IBGP" RFC 2796. April 2000.
R. Chandra and J. Scudder, "Capabilities Advertisement with BGP-4," RFC 2842, May 2000.

Фильтр исходящих маршрутов (Outbound Route Filter— ORF) представляет собой новую функцию в протоколе BGP, которая используется для минимизаиции количества обновлений, посылаемых на соседний узел. В этом приложении мы приводим некоторые сведения о применении фильтра исходящих BGP-маршрутов в IOS версий 12.0ST и 12.1.

Основная идея заключается в выставлении локально сконфигурированного фильтра BGP-префиксов для работы с удаленным узлом, что позволяет последнему использовать его как еще один фильтр исходящих маршрутов. В результате этого достигаются следующие преимущества.

- Уменьшается количество посланных префиксов, что, в свою очередь, приводит к снижению общего числа обновлений маршрутов, а следовательно снижает нагрузку на полосу пропускания.
- Локальный маршрутизатор тратит меньше ресурсов на обработку обновлений маршрутов, чем раньше. Благодаря этому снижается потребление ресурсов памяти, обрабатывается меньше атрибутов и создается меньший кэш.

Приложение B. Фильтр исходящих BGP-маршрутов (Outbound Route Filter — ORF)

В этом приложении мы рассматриваем лишь фильтр исходящих маршрутов. Список входящих префиксов и фильтр на его основе могут быть организованы на другом узле с учетом выходных правил маршрутизации.

Фильтр исходящих маршрутов представляет собой новую функцию в IOS и описан в ней кодом 130. Эта функция отражает самые последние разработки, которые включают в себя все поддерживаемые типы ORF с возможностью фильтрации по передаче, приему или по обеим направлениям. Это означает, что между маршрутизаторами, использующими эти функции, существует определенная степень совместимости. Старые механизмы управления при применении ORF могут работать только на прием.

Локальный узел, который объявляет о поддержке фильтра ORF в режиме передачи, будет помещать свой список входящих префиксов только в том случае, если от удаленного узла он получит подтверждение о возможности его работы с фильтром ORF в режиме приема. Удаленный узел будет ожидать приема запроса ROUTE-REFRESH или запроса от ORF с меткой IMMEDIATE и лишь затем посылать обновление мар-. шрутов. Обращаем ваше внимание, что все эти операции делаются для каждого семейства адресов (address family — AF) в сообщениях об обновлениях, которыми обмениваются узлы, в зависимости от объявленной функции BGP ORF.

Когда необходимо использовать BGP ORF

Фильтр BGP ORF позволяет BGP-спикеру установить собственный фильтр на основе списка входящих префиксов, поступающих с удаленного узла. Этот фильтр можно использовать для уменьшения нежелательных обновлений маршрутов.

Так, например, эта функция может использоваться для адресации принимаемых ("нежелательных") полных маршрутов от клиентов, подключенных к нескольким провайдерам. В этом случае клиенту необходимо лишь разрешить использование этой функции, что позволит его провайдеру управлять фильтрацией маршрутов и избежать распространения нежелательных обновлений маршрутов по сети.

Конфигурация

В последующих разделах мы рассмотрим основные этапы конфигурации фильтра BOP ORF. Фильтр ORF асимметричен и на разных концах соединения конфигурируется независимо.

Для конфигурации фильтра BGP ORF нужно предпринять следующие шаги.

1. Разрешить функцию BGP ORF *в режиме передачи (send-mode).*
2. Разрешить функцию BGP ORF *в режиме приема (receive-mode).*
3. Убедиться в обратной совместимости со старыми системами фильтрации.

Разрешение BGP ORF в режиме передачи

Команда, разрешающая работу BGP ORF в режиме передачи, имеет такой вид:

```
[no] neighbor x.x.x.x capability orf prefix list send
```

По этой команде локальный маршрутизатор конфигурируется для объявления работы с функцией ORF (значение 128) в режиме передачи (значение 2) в течение установления сеанса связи с указанным удаленным узлом. Эта команда может использоваться как для отдельных BGP-узлов, так и для узлов, которые являются членами группы или даже для группы узлов в целом.

По умолчанию эта функция отключена, поэтому при конфигурации маршрутизатора ее нужно включить.

Эта функция доступна для всех семейств адресов.

Разрешение BGP ORF в режиме приема

Команда, разрешающая работу BGP ORF в режиме приема, выглядит так:

```
[no] neighbor x.x.x.x capability orf prefix list receive
```

По этой команде локальный маршрутизатор конфигурируется для объявления работы с функцией ORF с BGP ORF (значение 128) в режиме приема (значение 1) в течение установления сеанса связи с указанным удаленным узлом. Эта команда может использоваться только с отдельными узлами или с группами взаимодействующих узлов. Она не может применяться по отношению к узлам, которые являются членами группы узлов.

По умолчанию эта функция отключена, поэтому при конфигурации маршрутизатора ее нужно включить.

Эта функция поддерживается для всех семейств адресов (в том числе для уникальной адресации IPv4, групповой адресации IPv4 и т.д.).

Обеспечение обратной совместимости со старыми системами фильтрации

Команда, позволяющая убедиться в обратной совместимости со старыми системами фильтрации, выглядит так:

```
[no] neighbor x.x.x.x  capability prefix-filter
```

Если существовал какой-либо фильтр, то новое программное обеспечение разрешает его скрытное использование и будет объявлять его как новый фильт BGP ORF, работающий в режиме приема. Обратите внимание, что он будет использоваться теперь в новом формате NVGEN. Это означает, что удаленный узел должен быть модернизирован для работы с ORF.

Если старый фильтр сконфигурирован примерно так:

```
[no] neighbor  x.x.x.x  send prefix-filter,
```

то новое программное обеспечение воспримет его как скрытый и объявит как новую функцию BGP ORF в режиме передачи. Обратите внимание, что он будет использоваться теперь в новом формате NVGEN. Это означает, что удаленный узел следует модернизировать для работы с ORF.

Команды EXEC

Передача списка префиксов и прием обновления маршрутов от соседнего узла

Синтаксис команды для передачи списка префиксов и приема обновления маршрутов от соседнего узла следующий (обратите внимание, что имеется возможность указания любого семейства адресов):

```
clear  ip bgp   x.x.x.x  in prefix-filter
clear  ip bgp  x.x.x.x vrf  foo  in prefix-filter
clear  ip bgp  x.x.x.x  ipv4  multicast  in prefix-filter
```

Когда список входящих префиксов изменяется (или просто удаляется), эта команда может использоваться для передачи нового списка префиксов и, следовательно, для приема обновления маршрутов от соседнего узла уже на основе нового списка префиксов.

Ключевое слово **prefix-filter** игнорируется, если от соседнего узла не получено подтверждение о работе с фильтром BGP ORF или если локальный спикер не разрешил соседнему узлу работу с ORF в режиме передачи.

Без ключевого слова **prefix-filter** команда **clear ip bgp *x.x.x.x* in** будет просто выполнять обычное обновление маршрутов от соседнего узла. Она не передает текущий фильтр списка входящих префиксов соседнему узлу. Эта команда полезна лишь при использовании входных правил маршрутизации, отличных от фильтра спика префиксов, таких как карта маршрутов.

Отображение списка префиксов, полученных от соседнего узла

Синтаксис команды для отображения списка префиксов, принятых от соседнего узла, следующий:

```
show ip bgp neighbor  x.x.x.x received prefix-filter
show ip bgp vpnv4 vrf foo neighbor x.x.x.x received prefix-filter
show ip bgp ipv4 multicast neighbor x.x.x.x received prefix-filter
```

Отображение изменений в таблице BGP-маршрутов на соседнем узле

С помощью команды **show ip bgp neighbor *x.x.x.x*** отображается следующая информация:

```
AF-dependent capabilities:
     Outbound Route Filter (ORF) type (128) Prefix-list:
        Send-mode: advertised, received
     Receive-mode: advertised, received
Outbound Route Filter (ORF): sent, received (2 entries)
First update is deferred until ORF or ROUTE-REFRESH is received
Scheduled to senr: the Prefix-list filter test
```

Заключительное замечание

Даже после того как BGP-спикер выдвигает на удаленный конец свой список входящих префиксов, он продолжает использовать локальный список входящих префиксов для фильтрации принимаемых обновлений маршрутов.

Приложение Г. Мультипротокольные расширения BGP (Multiprotocol BGP — MBGP)

Вместе с поддержкой функций БGP (BGP Capabilities – BGPCAP[1] в мультипротокольном расширении для протокола BGP (Multiprotocol BGP – MBGP[2]) описываются и обратно совместимые дополнения для протокола BGP-4[3], которые позволяют ему передавать маршрутную информацию для нескольких протоколов сетевого уровня, (например, IPv6, IPX и т. п.). Отдельные протоколы сетевого уровня идентифицируются семейством адресов (address family — AF), как описано в RFC 1700[4]. Эти расширения позволяют набору BGP-узлов обмениваться информацией о взаимной доступности для различных семейств адресов (например, IPv4, IPv6, IPX и т. п.), а также *у* подсемейств адресов (уникальные или групповые IP-адреса).

Примечание

Полный переход от старого интерфейса командной строки к новому, ориентированному на семейства адресов, пока еще не полностью реализован в оборудовании компании ^ Cisco. Поэтому если вы встретитесь с какими-либо противоречиями, проконсультируйтесь с персоналом из отдела технической поддержки компании Cisco или обратитесь к документации для вашей версии Gisco-IOS.

Причины перехода к новому интерфейсу командной строки

Существование нескольких семейств адресов обусловило развитие нескольких параллельных топологий, которые в итоге привели бы к полной взаимной несовместимости. Кроме того, наборы правил маршрутизации, задаваемые в определенном сеансе могут видоизменяться в зависимости от семейства адресов (address family— AF). Начальный интерфейс командной строки (Command Line Interface — CLI) был разработан для обслуживания только одного AF (IPv4) и был связан лишь с обработкой двух подсемейств адресов уникальныХ'И групповых адресов (unicast and multicast). В результате использования этого интерфейса'для поддержки дополнительных AP усложнились выражения, описывающие правила маршрутизации, а также повысились затраты на обслуживание этих AF.

Тогда был предложен новый подход — четкое отделение основных параметров сеанса работы маршрутизатора от специфических параметров, присущих определенному AF. Кроме отдельных аспектов управления, новый подход обеспечивает и другие преимущества.

- Входные и выходные правила маршрутизации могут отличаться *для* различных AF.

- BGP-маршрутизатор может быть настроен как отражатель маршрутов для одного AF или сразу нескольких AF.
- Не требуется дополнительных усилий для конфигурации и поддержки "ванильного" BGP (IPv4 с уникальными адресами).
- Префиксы могут независимо поступать из любых источников (путем преобразования, выражений **aggregate-address** и выражений **network)** внутри каждого AF.
- Группы взаимодействующих узлов будут обслуживаться внутри соответствую щего AF, так как это касается генерации сообщений UPDATE.

Новая модель более гибкая при использовании нескольких семейств адресов. На основе приведенных выше преимуществ субкоманды **router bgp** можно разделить на три группы,

1. **Глобальные команды BGP** — влияют на работу протокола BGP во всем маршрутизаторе. Например, это такие команды, как **bgp deterministic-med и bgp cluster-id.**
2. **Команды для идентификации соседних узлов/групп узлов** — описывают соседние узлы или группы взаимодействующих узлов, которые по умолчанию доступны согласно таблице маршрутов, путем определения параметров сеанса работы с ними. Примерами таких команд могут служить **команды neighbor 1.2.3.4 remote-as *as* и neighbor 1.2.3.4 ebgp-multihop /7/.**
3. **Команды для семейств адресов** — к этому типу могут быть отнесены два набора команд:
    - **глобальные команды семейств адресов** — это те команды, которые не зависят от конфигурации соседних узлов и влияют на работу по протоколу BGP определенного AF. Префиксы могут поступать от различных источников (с использованием преобразования, выражений **network или aggregate-address)** для **AF,** относящегося к этой категории. Примеры таких команд: **network 1.2.0.0 mask 255.255.0.0, redistribute dvmrp и bgp scan-time.**
    - **специфические команды семейств адресов для соседних узлов/группы узлов -** Эти команды позволяют настроить правила работы с соседними узлом (узлами) или в группе взаимодействующих узлов с помощью списков преобразования, списков префиксов или карт маршрутов. Соседние узлы можно также конфигурировать как клиентов отражателя маршрутов или как дополнительных членов группы узлов. При этом соседние узлы должны быть явным образом "активированы" в целях разрешения обмена префиксами в рамках MBGP. Подобными командами являются **neighbor 1.2.3.4 filter-list in, neighbor peergroupl route map foo in и neighbor 1.2.3.4 activate.**

Организация групп команд в новой конфигурации

Для группы команд 1 отсутствует неоднозначность, так что они могут задаваться в качестве глобальных параметров BGP. Эти команды задаются при конфигурации узлов только один раз.

Отсутствует неоднозначность и для команд группы 2, которые следуют за командами группы 1 с глобальными параметрами BGP. Соседние BGP-узлы, за исключением случая виртуальных частных сетей (Virtual Private Networks — VPN), описываются при конфигурации лишь один раз.

В режим конфигурации с использованием команды **router** bgp *autonomous-system* был введен новый дополнительный режим для команд AF.

При использовании нового AF-ориентированного интерфейса командной строки конфигурация представляет собой примерно следующее:

```
router bgp  autonomous-system
```

```
 address-family   afi   [sub-afi]
        redistribute protocol
        neighbor  1.2.3.4   activate
...
 exit-address-fainily
exit
```

Соседний узел по отношению к другому узлу может применять несколько различных правил маршрутизации (например, выражения route-map или prefix-list) одно для каждого AF.

Конфигурация BGP на основе IPv4 с уникальными адресами позволяет, чтобы наборы правил маршрутизации для соседних узлов были сконфигурированы лишь на основе команд группы 2. Это очень напоминает ситуацию, которую мы имеем сегодня (старый CLI). Хотя рекомендуется явная конфигурация, команда address-family ipv4 unicast также является однозначной.

Возможно сконфигурировать правила маршрутизации BGP для IPv4 с уникальными адресами в режиме address-family. Задав команду, мы получим результат выполнения команды address-family ipv4 unicast, глобальные IPv4 с уникальными адресами (глобальные по AF) и команды формирования правил маршрутизации (специфичные для AF для соседних узлов/групп узлов), разрешенные в этом режиме.

В режиме address-family глобальные для группы AF команды задаются в первую очередь. Все эти команды являются глобальными по отношению к AF.

За глобальными командами AF задаются специфичные для AF команды для соседних узлов/групп узлов. Эти команды реализуют наборы правил маршрутизации для соседних узлов в определенном AF.

Прежде чем для соседнего узла будут описаны какие-либо правила в соответствии с заданным AF, его следует "активировать" для данного AF. Синтаксис команды для активации соседнего узла в определенном AF таков:

```
router bgp autonomous-system
  address-family   afi   [sub-afi]
neighbor   1.2.3.4   activate
...
  exit-address-family
exit
```

Новая структура конфигурации представляет собой следующее:

```
router bgp  1
no synchronization                      ! Global to BGP
bgp deterministic-mod                   ! Global to BGP
bgp bestpath med confed                 ! Global to BGP
neighbor ebgp peer-group                ! Peer group defn., global to BGP
neighbor 1.1.1.1 remote-as 2            ! Neighbor defn., global to BGP
neighbor 2.2.2.2 remote-as 1            ! Neighbor defn., global to BGP
neighbor 3.3.3.3 remote-as 3            ! Neighbor defn., global to BGP
!
address-family ipv4 unicast             ! Address-family IPv4—uriicast
bgp scan-time 45                        ! Global  to IPv4-unicast
aggregate-address 50.0.0.0 255.255.0.0  ! Global  to  IPv4-unicast
neighbor ebgp activate                  ! Activate neighbor  for  IPv4-unic
neighbor ebgp route-map ucast-out out   ! Peergroup IPv4-unicast policy
neighbor 1.1.1.1 activate               ! Activate neighbor  for IPv4-unic
neighbor 1.1.1.1 peer-group ebgp        ! Neighbor membership -  IPv4-unic
neighbor 1.1.1.1 route-map ucast-in in  ! Neighbor IPv4-unicast policy
neighbor 2.2.2.2 activate               ! Activate neighbor for  IPv4-unic
neighbor 2.2.2.2 route-reflector-client ! RR client -  IPv4-unicast
neighbor 3.3.3.3 activate               ! Activate neighbor  for  IPv4-unic
 neighbor 3.3.3.3 peer-group ebgp        ! Neighbor membership -  IPv4-unic
```

```
no auto-summary                          ! Disable  IPv4-unicast auto
summarization
exit-address-famlly                      ! Exit AF sub-mode
!
address-family ipv4 multicast            ! AF sub-mode
network 100.0.0.0 mask 255.255.0.0       ! Global to IPv4-multicast
redistribute dvmrp route-map redist-map  ! Global  to  IPv4-multicast
neighbor ebgp activate                   ! Activate neighbor for IPv4-mult
neighbor 1.1.1.1 peer-group ebgp         ! Neighbor membership -  IPv4-mult
neighbor 1.1.1.1 route-map mcast-in      ! Neighbor  IPv4-multicast polic
neighbor 3.3.3.3 peer-group ebgp           ! Neighbor membership        IPv4-
mult
exit-address-family                      ! Exit AF sub-mode
exit
```

В последующих разделах приводится более подробная информация о командах. В каждом разделе приводятся два стиля конфигурации:

- **Старый стиль** — Используемый в настоящее время способ конфигурации BGP 12.0S.
- **Стиль AF** - Новый способ конфигурирования BGP с применением режима **address-family.**

В приведенных далее разделах рассматриваются основные этапы конфигурации — активация узла, объявление сети, групп взаимодействующих узлов, создание карт маршрутов, преобразование маршрутов, отражение маршрутов и объединение (агрегация).

activate

Команда **activate** лишь недавно введена в Cisco IOS в качестве расширении AF CLI. Она применяется для разрешения (или активации) поддержки указанного AF на BGP-yaie. В этом разделе будет рассмотрен старый метод разрешения использования AF с соседним узлом, а затем — этапы конфигурации этой же ф>ᵀнкции с использованием нового метода.

Старый стиль

В версии 12.OS команды **activate** не было. Соседние узлы активировались для работы с IPv4 по BGP автоматически. Однако при разрешении использования MBGP в команде **neighbor** требуется указать ключевое слово nlri:

```
Router (config-router)#neighbor 1.2.3.4  remote-as  10 nlri unicast multicast
```

Если ключевое слово **nlri** не указано, то маршрутизатор будет обмениваться только префиксами IPv4. Если же оно было укзано только с параметром **multicast**, то обмен будет проводиться только групповыми **NLRI** по IPv4. Таким образом, вы получаете возможность активировать только обмен уникальными, групповыми адресами или теми и другими одновременно.

Стиль AF

Чтобы разрешить использование определенного AF между соседними узлами, на

BGP-маршрутизаторе нужно задать команду **activate** или перейти в режим конфигурации **address-family.** При этом соседние BGP-узлы автоматически активируются для работы с **IPv4.** Для всех остальных **AF** соседние узлы следует активировать явным образом. Чтобы запретить соседнему узлу работать с AF, задайте команду активации с ключевым словом **no.** Ниже приводится синтаксис команды, которая разрешает узлу с адресом 1.2.3.4 поддержку групповых адресов IPv4:

```
Router (config-router)#address-family ipv4 multicast
Router (config-router-af)#neighbor 1.2.3.4 activate
```

network

Теперь команду **network** можно задавать в соответствии с определенным AF в порядке аннексирования сети в указанном AF. В этом разделе мы рассмотрим старый и новый стили конфигурирования маршрутизатора с помощью команды **network.**

Старый стиль

Старый стиль применения команды **network** предполагает объявление сети в BOP. В команде может дополнительно указываться расширение **nlri,** которое объявляется как поддерживающее работу с уникальными, групповыми или и теми, и другими адресами. Отсутствие в команде ключевого слова **nlri** автоматически означает, что поддерживается только работа с уникальными адресами IPv4. Например:

```
Router (config-router)#bgp 10
Router (config-router)#network 2.2.2.0 mask 255.255.255.0
Router (config-router)#network 3.0.0.0 mask 255.0.0.0 nlri multicast
Router (config-router)#network 1.1.0.0 mask 255.255.0.0 nlri unicast
multicast
```

Стиль AF

Согласно этому стилю, наличие режима **address-family** устраняет необходимость использования ключевого слова **nlri.** Чтобы объявить сеть через IPv4, команда должна быть указана на маршрутизаторе в режиме конфигурации BOP. Чтобы сеть была объявлена как групповая NLRI, команду **network** следует указывать в режиме AF с параметрами **ipv4 multicast.**

Приведенные ниже команды объявляют сеть всем соседним узлам семейства адресов IPv4:

```
Router (config-router)#network 1.1.0.0 mask 255.255.0.0
Router (config-router}#network 2.2.2.0 mask 255.255.255.0
```

Чтобы объявить сеть семейства групповых адресов IPv4, нужно определить сеть для использования с AF групповых адресов IPv4:

```
Router (config-router)#address-family ipv4 multicast
Router (config-router-af)#network 1.1.0.0 mask 255.255.0.0
Router (config-router)#network 2.2.2.0 mask 255.255.255.0
```

Таким образом, сети могут независимо объявляться как сети с уникальными, групповыми или теми и другими адресами IPv4 NLRI одновременно.

Группы взаимодействующих узлов

Группы взаимодействующих узлов также теперь можно конфигурировать для специфических AF, обеспечивая вместе с тем большую гибкость по сравнению со старым стилем конфигурирования. В этом разделе мы обсудим и старые, и новые методы конфигурации групп взаимодействующих узлов.

Старый стиль

Группа взаимодействующих узлов описывалась на маршрутизаторе в режиме конфигурации BGP. Для разрешения обмена групповыми префиксами внутри группы использовалось ключевое слово nlri. С помощью ключевого слова nlri можно было указывать уникальные и групповые адреса. Если же это слово не указывалось, то поддерживались только уникальные адреса IPv4.

Члены группы узлов автоматически поддерживали уникальные иди групповые адреса — в соответствии с тем, что было задано для группы. Приведенный ниже синтаксис команды показывает, как конфигурировались группы взаимодействующих узлов старым методом:

```
Router {config-router)#bgp 10
Pouter (config-router)#neighbor external peer-group nlri unicast
multicast
Router (config-rcuter)#neighbor 1.2.3.4 remote-as 20
Router (config-router)#neighbor 1.2.3.4 peer-group external
```

Стиль AF

Группы узлов (или их члены) описываются на маршрутизаторе в режиме конфигурирования BOP. Но поскольку имеются дополнительные режимы конфигурирования для различных AF, отпадает необходимость в использовании ключевого слова nlri- Группа узлов должна теперь активироваться в AF групповых адресов IPv4, что позволит проводить обмен групповыми префиксами IPv4. Как и с командой neighbor, группа взаимо-дсисптощнх узлов и ее члены по умолчанию активируются только для поддержки \нн-кольных адресов IPv4. Однако это действие можно перекрыть, задав команду activate с ключевым словом по. Следующий синтаксис команд показывает, каким образом конфигурируется группа узлов с использованием нового AF-ориентированного CLI:

```
Router (config-router}#bgp 10
Router (config-router)#neighbor external peer-group
Router (config-router)#neighbor 1.2.3.4 remote-as 20
Router (config-router)#neighbor 1.2.3.4 peer-group external

Router (config-router)#address-family ipv4 multicast
Router (config-router-af)#neighbor external activate outer
Router (config-router-af)#neighbor 1.2.3.4 peer-group external
```

Карты маршрутов

Старый метод, конфигурирования карт маршрутов требовал, чтобы правила лля все отдельных AF описывались внутри одной карты маршрутов. Новый AF-ориентированный стиль конфигурации позволяет создавать индивидуальные карты маршрутов для каждого семейства адресов. В этом разделе мы остановимся на старом и новом методах описания и

применения карт маршрутов относительно различных AF.

Старый стиль

В старом стиле одна карта маршрутов использовалась для описания правил маршрутизации для всех AF. Затем эта карта маршрутов применялась либо ко входящему, либо к исходящему трафику отдельного узла или группы узлов. Правила маршрутизации для двух AF, которые могли передаваться в сеансе BGP одновременно (уникальные и групповые адреса IPv4), были представлены в одной карте маршрутов путем указания ключевого выражения match nlri в последовательности карты маршрутов. Выражение match nlri в карте маршрутов семантически выглядело так:

```
match nlri multicast          !   Matches only IPv4 multicast
match nlrimulticast unicast   !   Matches both IPv4 unicast and
multicast
match nlri unicast            !   Matches only IPv4 unicast
match nlri                    !   [Unspecified; matches only IPv4
unicast
```

В приведенном ниже примере показано, как сконфигурировать BGP, чтобы от узла 1.1.1.1 принимались любые групповые маршруты, соответствующие списку разрешенных узлов 1:

```
router bgp 109
 neighbor 1.1.1.1 remote-as 1 nlri unicast multicast
 neighbor 1.1.1.1 route-map filter-mcast in
!
route-map filter-mcast permit 10
 match nlri multicast
 match ip address 1
```

Стиль AF

Одной из важнейших причин для перехода от старого стиля конфигурации к AF-ориентированному стилю было выражение правил маршрутизации с применением ключевого выражения **match nlri,** которое включалось в последовательность карты маршрутов и становилось практически неуправляемым, особенно при описании сложных и специфических правил для различных AF. Благодаря поддержке в новой 1OS большего числа AF, стало очевидным, что необходимо разработать более масштабное решение для поддержки всех правил маршрутизации. Применение одной карты маршрутов для выражения всех правил маршрутизации было признано нецелесообразным и очень сложным.

Представление нового режима анализа для каждого отдельного AF способствовало введению нового способа конфигурирования правил маршрутизации — теперь уже на основе заданных AF (т.е. для каждого AF создавалась отдельная карта маршрутов). Но на основе AF можно создавать не только карты маршрутов, но и осуществлять разного рода фильтрацию, т.е. создавать списки префиксов, списки преобразования, списки разрешения доступа на основе атрибута AS_PATH и т.д. Теперь, с введением нового стиля конфигурирования правил маршрутизации, уже не требуется ключевое слово **nlri,** и при анализе карты маршрутов выражения **match nlri** игнорируются.

Согласно AF-стилю, правила маршрутизации теперь можно описать следующим образом:

```
router  bgp 109
neighbor 1.1.1.1   remote-as   1
neighbor 1.1.1.1   route-map   filter-ucast   in
!
```

```
address-family   ipv4   multicast
neighbor   1.1.1.1   activate
neighbor   1.1.1.1   route-map   filter-mcast   in
!
route-map filter-mcast permit 10
match ip address 1
!
route-map filter-mcast permit 10
 match ip address 2
```

Преобразование

Преобразование -- это процесс импортирования маршрутов из одного протокола маршрутизации в другой. Когда маршруты преобразуются в BGP-маршруты посредством команды **redistribute,** следует создать таблицу BGP-маршрутов (Adj-RIB-In), в которую и будут импортироваться маршруты. Эта таблица BGP-маршрутов может содержать как уникальные адреса, так и групповые.

Старый стиль

Старый стиль задания таблицы, в которую должны были помещаться маршруты, предполагал применение выражения **set nlri** в карте маршрутов. При выполнении преобразования это выражение в карте маршрутов использовалось следующим образом:

- **set nlri multicast**-- преобразование соответствующих префиксов в групповую таблицу;
- **set nlri unicast multicast** -- преобразование соответствующих префиксов в таблицу с уникальными адресами и в таблицу с групповыми адресами;
- **set nlri unicast** — преобразование соответствующих префиксов только в таблицу с уникальными адресами;
- **set nlri**-- преобразование с неявным указанием (в этом случае выполнялось преобразование в таблицу с уникальными адресами).

Ниже на примере показано преобразование в BGP. Здесь, как видите, все "связанные" префиксы, соответствующие списку разрешения доступа I в таблице маршрутов, импортируются как групповой NLRI:

```
router bgp 109
redistribute connected route-map mbgp-source-map
!
route-map mbgp-source-map
 match ip address 1
 set nlri multicast
```

Стиль AF

Команда **redistribute,** заданная в режиме AF, определяла таблицу, в которую помещались префиксы после преобразования. Например, если выражение **redistribute** задавалось в режиме address-family **ipv4 multicast,** то преобразованные префиксы помешались в таблицу как групповые NLRI IPv4. Следовательно, предыдущая конфигурация преобразования, согласно старому стилю, трансформировалась в следующую:

```
router 109
```

```
 address-family ipv4 multicast
 redistribute connected route-map mbgp-source-map
!
route-map mbgp-source-map
 match ip address 1
```

Отметим, что с вводом выражения redistribute в режиме AF, отпадает необходимость в использовании ключевой цепочки set **nlri** и далее при анализе карты маршрутов последняя будет игнорироваться.

Отражатель маршрутов

Теперь конфигурация отражателя маршрутов также может выполняться на основе семейства адресов, что позволяет делать это более гибко и с меньшими затратами по сравнению со старым "централизованным" методом. В этом разделе мы обсуждаем и старый, и новый стили конфигурации отражателя маршрутов.

Старый стиль

Согласно старому стилю свойства клиентов отражателя маршрутов описывались глобально, и вся конфигурация затем применялась ко всем AF, поддерживаемым его клиентами. Отражатель маршрутов "знал", что отражение маршрутов от и к клиентам должно осуществляться по команде route-reflector-client, где указывался определенный узел или группа узлов IBGP. Далее приведен пример, где узел IBGP с адресом 1.1.1.1 становится клиентом отражателя маршрутов и для уникальных, и для групповых префиксов IPv4:

```
router bgp  109
neighbor  1.1.1.1   remote-as   109  nlri  unicast multicast
neighbor   1.1.1.1   route-reflector-client
```

Стиль AF

Может ли узел или группа узлов быть клиентами отражателя маршрутов, зависит, согласно стилю AF, от семейства адресов. В таком случае конфигурация их как клиентов проводится в режиме AF. Другими словами, то, что узел является клиентом отражателя маршрутов, поддерживающего уникальную адресацию по IPv4, совершенно не означает, что этот узел является одновременно и клиентом отражателя маршрутов для групповых адресов IPv4. Теперь это необходимо указывать явно при конфигурации клиента в режиме AF как поддерживающего работу с групповыми адресами IPv4. Таким образом, предыдущая конфигурация, когда узел 1.1.1.1 становился клиентом для работы по уникальным и групповым адресам, теперь будет иметь такой вид:

```
router bgp  109
 neighbor   1.1.1.1   remote-as   109
 neighbor   1.1.1.1   route-reflector-client
!
 address-family  ipv4  multicast
 neighbor   1.1.1.1   activate
 neighbor   1.1.1.1  route-reflector-client
```

Новый режим дает оператору определенную гибкость при конфигурации маршрутизатора в качестве отражателя маршрутов только для определенных AF. Таким образом, топология отражателей маршрутов для различных AF может варьироваться.

Объединение

Конфигурация объединения одного или нескольких AF была всегда довольно сложной задачей, особенно, если возникала необходимость в обеспечении определенных правил маршрутизации. В этом разделе мы приводим некоторые сведения о старых и новых методах конфигурации объединения маршрутов.

Старый стиль

В старом стиле объединение маршрутов с групповыми адресами конфигурировалось точно так же, как и объединение маршрутов с уникальными адресами -- с помощью команды **aggregate-address.** Команда **aggregate-address** затем была дополнена, что позволило с помощью ключевого слова **nlri** указывать, к каким адресам (уникальным или групповым) она относится. Ниже приведен пример генерации объединенного маршрута в таблице групповых BGP-адресов:

```
router bgp 109
aggregate-address 174.0.0.0 255.0.0.0 as-set nlri multicast
```

Параметры **nlri,** которые можно было использовать в команде **aggregate-address, - unicast, multicast и unicast multicast.** С их помощью в таблице BGP-маршр/тов генерировались объединенные маршруты с уникальными, групповыми или теми и другими адресами. В отсутствие ключевого слова **nlri** объединенные маршруты генерировались только с уникальными адресами.

Стиль AF

С появлением режимов AF для каждого семейства адресов исчезла необходимость использования ключевого слова **nlri** в команде **aggregate-address.** Режим AF, в котором теперь указывался объединенный маршрут, определял в какую таблицу будет помещаться сгенерированный объединенный префикс. Таким образом, в стиле AF объединенный маршрут описывался следующим образом:

```
router bgp 109
!
          address-family ipv4 multicast
          aggregate-address 174.0.0.0 255.0.0.0 as-set
```

Список команд BGP

В табл. Г-1 приводится список команд протокола BGP и категория, к которым они относятся.

Таблица Г-1. *Команды и субкоманды BGP, сгруппированные по категориям*

Команда/Субкоманда	Категория
address-family ipv4 unicast	Может один раз применяться в режиме router bgp
address-family ipv4 multicast	Может один раз применяться в режиме router bgp

aggregate-address	Согласно AF
auto-summary	Согласно AF
bgp ahvays-compare-med	Глобальная для BGP
bestpath	Глобальная для BGP
clicnt-io-cliem	Согласно AF
cluster-id	Глобальная для BGP
confederation	Глобальная для BGP
dampening	Согласно AF
default	Глобальная для BGP
determministic-med	Глобальная для BGP
fast-external-fallover	Глобальная для BGP
log-neighbor-changes	Глобальная для BGP
redistribute-internal	Глобальная для BGP
router-id	Глобальная для BGP
scan-time	Согласно AF
default- metric	Глобальная для BGP
distance	Согласно AF
maximum-paths	Согласно AF
neighbor	Применяется по согласованию между соседними узлами
activate	Согласно AF
advertisement-interval	Глобальная для соседнего узла (на время сеанса)
default-originate	Согласно AF (правила маршрутизации)
description	Глобальная для соседнего узла
distribute-list	Согласно AF (правила маршрутизации)
ebgp-multihop	Глобальная для соседнего узла (на время сеанса)
filter-list	Согласно AF (правила маршрутизации)
local-as	Глобальная для соседнего узла (на время сеанса)
maximum-prefix	Согласно AF (правила маршрутизации)
next-hop-self	Глобальная для соседнего узла (на время сеанса)
password	Глобальная для соседнего узла (на время сеанса)
peer-group	Согласно AF (правила маршрутизации)
prefix-list	Согласно AF (правила маршрутизации)
remote-as	Глобальная для соседнего узла (на время сеанса)
remove- private- AS	Глобальная для соседнего узла (на время сеанса)
route-map	Согласно AF (правила маршрутизации)
route-reflector-clicnt	Согласно AF
send- community	Согласно AF (правила маршрутизации)
shutdown	Глобальная для соседнего узла (на время сеанса)
soft-reconfi?unilion	Согласно AF Справила маршрутизации)
Команда/Субкоманда	Категория
timers	Глобальная для соседнего узла (на время сеанса)
update-source	Глобальная для соседнего узла (на время сеанса)
version	Глобальная для соседнего узла (на время сеанса)
weight	Согласно AF (правила маршрутизации)
network	Согласно AF

redistribute	Согласно AF
syncronization	Согласно AF
table-map	Согласно AF
timers	Глобальная для BGP

Переход к использованию стиля AF

Для того чтобы совершить плавный переход от одного стиля к другому, в новой версии была сохранена поддержка некоторых команд версии 12.OS (которые позволяют использовать ключевое слово **nlri).** Вот эти команды:

- **neighbor**
- **network**
- **aggregate**
- **set nlri и match nlri** в картах маршрутов

Единственное предостережение: команды старого стиля могут использоваться до те.х пор, пока не требуется активизировать новые функции. В этом случае команды BGP старого стиля следует транслировать в команды нового стиля.

Чтобы перейти к использованию нового набора команд, в режиме конфигурации маршрутизатора нужно ввести команду **bgp up grade-cli:**

```
Router(config-router)#bgp  upgrade-cli
```

Старая конфигурация будет преобразована в новую. Далее, как обычно, нужно выполнить **wr** mem, чтобы сохранить новую конфигурацию. (Замечание: сама команда **bgp upgrade-cli** в конфигурации не отображается).

Ссылки

[1] RFC 2842, "Capabilities Advertisement with BGP-4," www.isi.edu/in-notes/rfc2842.txt
[2] RFC 2283, "Multiprotocol Extensions for BGP-4," www.isi.edu/in-notes/rfc2283.txt
[3] RFC 1771, "A Border Gateway Protocol 4 (BGP-4)," www.isi.edu/in-notes/rfcl771.txi
[4] RFC 1700, "Assigned Numbers," www.isi.edu/in-notes/rfcl700.txt

Принципы маршрутизации в Internet

2-е издание

- **Откройте для себя функции, атрибуты и области применения протокола BGP-4 — стандартного протокола междоменной маршрутизации — на практических примерах**
- **Изучите структуру сети Internet и узнайте, как выбирать провайдера на основе оценки его возможностей с точки зрения организации маршрутизации и способа подключения к Internet**
- **Овладейте методами распределения адресного пространства, включая бесклассовую междоменную маршрутизацию (Classless Interdomain Routing — CIDR) — наиболее перспективного подхода для сдерживания лавинообразного роста сети Internet**
- **Разработайте оптимальные правила маршрутизации для своей сети с учетом требований резервирования, распределения трафика, симметричности и стабильности**
- **Изучите, как интегрировать ваши внешние и внутренние домены маршрутизации и управлять сложными быстрорастущими автономными системами**

Эта книга является частью серии ***Cisco Press Core Series****, которая ориентирована на профессионалов по работе с сетями, и содержит ценную информацию по созданию оптимальных сетей, о новых технологиях и о том, как сделать карьеру в этой области*

ВИЛЬЯМС
www.williamspublishing.com

Cisco Systems
Cisco Press
ciscopress.com

В этой книге рассматриваются все тонкости организации междоменной маршрутизации в сетях с использованием протокола граничного шлюза версии 4 (Border Gateway Protocol Version 4 — BGP-4), который сегодня является стандартом для междоменной маршрутизации. Здесь вы найдете всю необходимую информацию для принятия обоснованных решений при подключении сети вашей организации к Internet.

Приобретенные знания позволят вам провести интеграцию вашей сети в Internet и создать собственные крупномасштабные автономные системы. Вы научитесь управлять внутренними протоколами маршрутизации с помощью BGP-4, разрабатывать качественные и стабильные сети, настраивать с помощью программного обеспечения Cisco IOS™ необходимые правила маршрутизации, а также узнаете о распространенных в сети Internet приемах и способах маршрутизации. Используя эту книгу как справочник, вы сможете на уровне эксперта управлять маршрутной информацией.

Эта книга является самым полным справочником по маршрутизации в сетях на базе TCP/IP и в сети Internet.

Автор книги, **Сэм Хелеби** (Sam Halabi), является одним из лучших экспертов на рынке провайдеров Internet. Сэм Хелеби, недавно назначенный вице-президентом компании Marketing at an IP networking startup, до этого несколько лет возглавлял отдел IP-маркетинга в компании Cisco Systems. Сэм Хелеби является экспертом по сложным протоколам маршрутизации и специализируется на разработке и организации крупномасштабных IP-сетей.

Принимая активное участие в развитии рынка вычислительных сетей, Хелеби является членом технологических форумов по разработке и внедрению новых технологий Optical Internetworking Forum и MPLS Forum.

Категория:
Internet — Сети

Предмет рассмотрения:
Глобальные сети и протокол BGP-4

ISBN 5-8459-0188-X

9 785845 901880 01058