ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ

**Федеральное государственное
бюджетное образовательное учреждение высшего образования
«САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М. А. БОНЧ-БРУЕВИЧА»**

**А. Б. Гольдштейн, А. В. Никитин,
А. А. Шкрыль**

ТРАНСПОРТНЫЕ СЕТИ IP/MPLS

ТЕХНОЛОГИЯ И ПРОТОКОЛЫ

Учебное пособие

*Разработано в рамках договора между ОАО «Ростелеком» и СПбГУТ
на выполнение научно-исследовательских работ
по разработке учебно-методических комплексов с интерактивным обучением
по дисциплинам базовой кафедры «Инновационные технологии телекоммуникаций»
ОАО «Ростелеком» в СПбГУТ*

СПб ГУТ)))

**САНКТ-ПЕТЕРБУРГ
2016**

УДК 621.395.74
ББК 32.882
Г63

Рецензенты:
заведующий кафедрой автоматической электросвязи,
профессор *А. П. Пшеничников* (МТУСИ),
доктор технических наук, профессор *Н. А. Соколов* (ЛО ЦНИИС)

Утверждено редакционно-издательским советом СПбГУТ
в качестве учебного пособия

Гольдштейн, А. Б.

Г63 Транспортные сети IP/MPLS. Технология и протоколы : учебное пособие / А. Б. Гольдштейн, А. В. Никитин, А. А. Шкрыль ; СПбГУТ. – СПб., 2016. – 80 с.

ISBN 978-5-89-160-129-1

Написано в соответствии с рабочими программами дисциплин «Транспортные сети NGN» и «Инфокоммуникационные протоколы». Приведен теоретический материал по технологиям пакетных сетей связи, рассмотрены вопросы работы протоколов маршрутизации, механизмов и протоколов сигнализации в IP/MPLS, вопросы QoS и Traffic Engineering, а также развития MPLS, T-MPLS, MPLS-TP в дальнейших направлениях эволюции технологии.

Предназначено для студентов, обучающихся по направлениям подготовки 11.03.02 «Инфокоммуникационные технологии и системы связи», 09.03.01 «Информатика и вычислительная техника».

УДК 621.395. 74
ББК 32.882

ISBN 978-5-89-160-129-1

Предисловие

Инфокоммуникационные технологии проникли в нашу жизнь настолько глубоко, что мыслить дальнейшее развитие человечества вне контекста средств и технологий связи просто невозможно.

Национальная телекоммуникационная компания ПАО «Ростелеком» является не только ключевым игроком отечественной отрасли связи, но и образцом внедрения инновационных решений. Обеспечение бесперебойной работы крупнейших по своей протяженности сетей связи, расположившихся от Крайнего Севера до Ирана, от Лондона до Токио, является сложнейшей задачей, лежащей на наших сотрудниках. Подготовка специалистов, способных не только поддержать существующие процессы и досконально изучить используемые технологии, но и найти силы для поиска новых революционных решений – вот это и есть высшая цель, стоящая перед создателями актуального учебного пособия.

Объединяя лучшее от практики и теории, авторы постарались максимально компактно и ярко изложить механизмы, протоколы и требования современных сетей IP/MPLS.

Руководство ПАО «Ростелеком» выражает признательность всему коллективу авторов за проведенную работу и желает всем слушателям укрепить свои знания и достойно встретить все вызовы профессиональной карьеры.

Вице-президент – Директор макрорегионального
филиала «Северо-Запад» ПАО «Ростелеком»
А. В. Балаценко

Обеспечение занятий теоретическим материалом, способным отразить актуальную картину в области транспортных сетей связи, является важным вопросом в подходе к построению эффективного обучения. Учебное пособие «Транспортные сети IP/MPLS. Технология и протоколы» является кратким отражением мирового опыта эксплуатации технологии IP/MPLS и объясняет фундаментальные положения пакетных технологий сетей связи, успешно используемых в построении операторских сетей.

Бурное развитие инфокоммуникационной отрасли, охватывающей все новые отрасли экономики, порождает новые требования к транспортным сетям связи. Большая часть методических пособий по технологиям IP/MPLS устарела с точки зрения анализа таких требований. Появление новых стандартов мобильной связи, развитие облачных услуг, сетей центров обработки данных и постоянный рост объемов трафика – все вместе требует актуализации понимания технологии IP/MPLS. Данное пособие полностью соответствует предъявляемым требованиям и опирается на последние нормативные документы отрасли.

1. ВВЕДЕНИЕ В СОВРЕМЕННЫЕ СЕТЕВЫЕ ТРАНСПОРТНЫЕ ТЕХНОЛОГИИ

Первое десятилетие XXI в. ознаменовалось всесторонним проникновением технологий информатизации в повседневную жизнедеятельность человечества. Стремительное увеличение решаемых прикладных задач, подключенных устройств и объемов передаваемых данных в корне изменило пути развития сетевых технологий, услуг связи и информационных систем управления сетями связи.

Традиционные услуги связи на современном этапе развития инфокоммуникаций больше не являются локомотивом развития отрасли и составляют процентное меньшинство по прибыльности и популярности среди абонентов. Услуги фиксированного доступа заменяются мобильными, а передача голоса заменяется передачей видеоконтента. Растут скорости доступа и объемы передаваемой информации, как и требования к качеству передаваемой информации. Возникает большое количество услуг, реализация которых происходит в реальном времени и требует от оператора новых архитектурных решений.

Если раньше оператор являлся монополистом в области создания услуг, то сейчас он стремительно уступает место специализированным компаниям, создающим различные программные приложения для обмена широким спектром типов информации (сообщения, фото, видео, игры и т. д.). Успех так называемых социальных сетей (facebook, vk.com, Myspace), операторов потокового видеовещания (NetFlix), служб forsqaure, Instagram и других показывают востребованные пользователями услуги и решения в масштабах многомиллионных аудиторий. Очевидно, что оператор самостоятельно не может разрабатывать конкурентные решения в области онлайн-сервисов.

Смещение оператора в сторону известной «битовой» трубы заставляет оператора переосмысливать свою роль, корректировать стратегию и планировать развитие в соответствии с потребностью рынка.

Современные решения в области транспортных технологий предназначены для удовлетворения потребности операторов фиксированной и сотовой связи в интеграции своих различных сетей (для передачи голоса, видео, трафика сигнализации и данных) в единую магистраль. Таким образом, оператор должен решить проблему снижения расходов на построение и эксплуатацию этих магистралей, а также подготовиться к разворачиванию сложных инфокоммуникационных сетей и сетей следующего поколения.

Важнейшие механизмы, используемые при построении транспортных сетей, ориентированы на интеллектуальное управление потоками трафика, обеспечение требований SLA и QoS и дальнейшее развитие, а также на интеллектуальное управление всеми ресурсами транспортных сетей.

Динамичное подключение центров обработки данных и высокопроизводительных локальных сетей по всему миру к транспортным сетям постоянно увеличивает требования к надежности и автоматизации процесса подключения, обеспечения надежности и мониторинга.

Ярким примером инновационного развития транспортных сетей связи является опыт ОАО «Ростелеком».

В 2006 г. компания ОАО «Ростелеком» (далее «Ростелеком») начала строить федеральную сеть передачи данных на основе современных пакетных технологий. В 2007 г. сеть передачи данных IP/MPLS Ростелеком была принята в эксплуатацию. Сегодня протяженность сети составляет более 40 тыс. км. Она состоит из 10 опорных и свыше 100 региональных узлов и свыше 350 точек доступа на всей территории РФ.

Сеть IP/MPLS Ростелеком является высокоскоростной сетью передачи данных на основе коммутации пакетов и функционирует поверх первичной транспортной сети, построенной на основе ВОЛС SDH и DWDM. Предназначена она для конвергенции услуг по передаче видео, речи и данных и призвана обеспечить возможность построения инфраструктуры интеллектуальных сетей следующего поколения. В настоящее время сеть передачи данных IP/MPLS Ростелеком присутствует на зарубежных точках обмена трафиком, имеет сеть собственных региональных Дата-центров в Москве, Екатеринбурге, Красноярске и Хабаровске, предоставляет услуги L2 VPN и VPLS.

Сеть имеет динамическую маршрутизацию, поддерживает протоколы IPv4, IPv6, MPLS Fast Reroute и обеспечивает передачу в режиме реального времени различных типов трафика для мультимедиа, голоса, данных, видео, сети Интернет, с предоставлением услуг виртуальных частных сетей и различных классов обслуживания.

Высокоскоростная IP-магистраль построена на базе ресурсов собственной первичной сети по технологии MPLS (Multi-protocol Label Switching), обеспечивающей конвергенцию услуг по передаче видео, речи и данных.

Возможности сети:

- покрытие всех регионов РФ: 24 узла уровня ядра и 110 региональных распределяющих узлов;
- свыше 350 точек доступа на всей территории РФ;
- пропускная способность сети IP MPLS – 7,8 Тбит/с;
- зарубежный сегмент в Европе и Азии емкостью 1,5 Тбит/с;
- присутствие на зарубежных узлах доступа: в Стокгольме, Лондоне, Гонконге, Франкфурте и Амстердаме;

- присутствие в российских и международных точках обмена интернет-трафиком;
- самое современное в России на сегодняшний день оборудование (магистральные маршрутизаторы Juniper T1600 производительностью до 1,6 Тбит/с; пограничные маршрутизаторы Juniper MX960, Juniper MX480, Juniper M320, Juniper M40);
- наличие сертификата соответствия сети требованиям информационной безопасности ФСТЭК;
- сеть собственных Дата-центров в Москве, Казани, Екатеринбурге, Новосибирске, Хабаровске.

2. ИСТОРИЧЕСКОЕ РАЗВИТИЕ ТЕХНОЛОГИЙ ТРАНСПОРТНЫХ СЕТЕЙ СВЯЗИ

Объединение маршрутизации и коммутации в одном устройстве. Появление технологии IP-switching.

Технология *многопротокольной коммутации по меткам MPLS* явилась результатом слияния нескольких сходных технологий, которые были изобретены в середине 1990-х гг. Наиболее известная из них (хотя и не первая, увидевшая широкий свет) была названа ее изобретателями – компанией Ipsilon – *IP Switching.* Ранее компания Toshiba уже описала похожий механизм – *Cell Switching Router* (CSR), а вскоре были обнародованы сведения и о некоторых других технологиях, среди которых отметим *Tag Switching* (Cisco Systems) и *ARIS* (IBM). Эти механизмы имеют ряд общих черт. Все они используют для пересылки пакетов простой метод замены меток и разработанную для Интернета структуру управления, т. е. IP-адреса и стандартные протоколы маршрутизации, например OSPF и BGP.

С начала 1990-х годов мощными двигателями развития технологии коммутации по меткам были проблемы обеспечения совместимости протоколов IP и ATM. Поэтому на историю MPLS значительное влияние оказала сложившаяся тогда ситуация вокруг проблемы маршрутизации пакетов IP по сетям ATM. Попытки развивать в этом направлении стандарты протоколов ATM предпринимались еще раньше – в конце 1980-х гг. Уже тогда неоправданно преувеличенные перспективы технологии ATM обусловили начало разработки механизма переноса IP-пакетов по сетям ATM. Этой проблемой занялись сразу несколько рабочих групп в составе комитета IETF, а на рубеже 1993–1994 гг. ими были опубликованы два важных документа серии RFC.

Первый стандарт, посвященный IP-поверх-ATM (IPoATM), описан в RFC 1483 и касается простой проблемы: каким образом инкапсулировать IP-дейтаграммы (и пакеты других протоколов) в канал ATM. Второй стандарт, описанный в RFC 1577, определяет классический механизм передачи IP-пакетов по сети ATM и протокол преобразования адресов ATMARP. Классический механизм предполагает, что маршрутизаторы пакетов IP и хосты, находящиеся в одной и той же подсети (т. е. сетевые и подсетевые составляющие их адресов одинаковы), могут взаимодействовать через эту подсеть. Если они находятся в разных подсетях, то для пересылки пакета от подсети отправителя к подсети получателя необходим еще один или несколько маршрутизаторов. При определении классической модели IP-поверх-ATM было признано, что IP-устройства могут подключаться к общей сети ATM (например, к крупной сети ATM, принадле-

жащей оператору сети общего пользования) и при этом находиться в подсетях, которые в эту общую сеть не входят. Таким образом, была предложена идея логической подсети IP (LIS), которая состоит из совокупности хостов и маршрутизаторов IP-пакетов, подключенных к общей сети ATM и имеющих общий адрес с IP-сетью и с подсетью.

Документ RFC 1577 специфицирует только взаимодействие внутри LIS и предполагает, что для передачи пакета из одной LIS в другую он должен проходить через маршрутизатор, подключенный к обеим LIS. Прежде всего, в связи с классической моделью, нужно отметить следующее. Она подразумевает, что два IP-устройства, подключенных к одной и той же сети ATM, но через разные подсети LIS, не смогут использовать для обмена IP-дейтаграммами единый виртуальный канал VC, проходящий через сеть ATM. Вместо этого они будут вынуждены передавать пакеты через отдельный маршрутизатор. Такой механизм многим специалистам казался тогда непривлекательным, особенно в связи с тем, что на тот период времени производительность серийно выпускавшихся коммутаторов ATM значительно превышала производительность маршрутизаторов. Несмотря на то что возможным вариантом решения этой проблемы могло бы стать введение правила, согласно которому сеть ATM должна представлять собой одну LIS, такой вариант часто трудно реализовать по организационным причинам. Например, маловероятно, что две совершенно разные организации, подключенные к одной сети ATM общего пользования, захотят использовать для своих IP-адресов общее адресное пространство. Далее, после создания общей концепции LIS в RFC 1577 определен ключевой механизм управления организацией связи между двумя IP-устройствами, входящими в одну подсеть LIS, – протокол ATMARP. Используется протокол преобразования адресов ARP в традиционной локальной сети, позволяющий IP-устройствам получать адресную информацию, которая необходима для организации связи, например адреса оборудования локальной сети Ethernet. Аналогичным образом ATMARP позволяет двум IP-устройствам узнать ATM-адреса друг друга. В протокол ATMARP было введено новое понятие ARP-сервера подсети LIS, который преобразует IP-адреса в адреса оборудования сети ATM для данной LIS. Каждое IP-устройство подсети LIS регистрируется на ARP-сервере и получает его адрес в сети ATM и его IP-адрес. После этого любое устройство LIS может запросить у ARP-сервера преобразование IP-адреса в адрес сети ATM, а уже снабженное адресом сети ATM устройство сможет создать виртуальный канал к этому адресу, используя сигнализацию ATM, и затем передать свои данные.

Решение проблемы создания виртуального канала ATM к IP-устройству другой LIS, отсутствующее в RFC 1577, взяла на себя рабочая группа

ROLC в составе IETF. Предложенный ею механизм – *протокол Next Hop Resolution Protocol (NHRP)*. Протокол NHRP позволяет IP-устройству одной логической подсети IP узнать ATM-адрес другого IP-устройства, с которым ему нужно установить связь, с помощью специального *сервера следующей пересылки (Next Hop Server)* и организовать виртуальный канал связи с этим устройством, используя сигнализацию ATM.

Однако во всех этих предшествовавших MPLS работах не подвергался сомнению базовый принцип: маршрутизаторы выполняют функции маршрутизации, а коммутаторы выполняют функции коммутации, и устройства этих двух типов всегда функционируют порознь. При этом имеются в виду не только IP-маршрутизаторы и ATM-коммутаторы, но и само правило разделения функций уровней 3 и 2 между различными технологиями и устройствами.

Компания *Toshiba* практически впервые подвергала сомнению этот принцип и в 1994 г. анонсировала *маршрутизатор коммутации ячеек CSR (Cell Switching Router)*. В архитектуре CSR реализована идея управления коммутационным полем ATM-коммутатора с помощью протоколов IP, а не протоколов сигнализации сети ATM типа Q.2931. Подобный подход, будучи доведенным до логического завершения, смог бы свести на нет необходимость использования практически всей сигнализации ATM и всех функций мэппинга между IP и ATM. Этот подход позволяет совместно использовать традиционные коммутаторы ATM и оборудование CSR. Например, CSR могут обеспечить взаимосвязь между подсетями LIS, устраняя необходимость в протоколе NHRP. Проект CSR был представлен на обсуждение рабочей группы комитета IETF в 1994 г., а немного позже, в начале 1995 г., – на технической сессии BOF комитета IETF, однако интерес к этой проблеме тогда был довольно низким.

Компании Ipsilon (позднее была приобретена Nokia) благодаря более полным техническим спецификациям *IP Switching* и наличию готового продукта IP switch обычно приписывается честь создания первой по-настоящему формализованной концепции MPLS-подобной коммутации по меткам в сетях IP, получившей значительно большее признание, нежели технология CSR. Сам IP Switch состоял из ATM-коммутатора и контроллера IP-коммутатора, который выполнял функции управления. Контроллер IP-коммутатора фактически являлся отдельным устройством, содержащим функциональные объекты маршрутизации и пересылки данных.

Среди рассматриваемых в этом разделе технологий в пользу технологии IP Switching можно привести существенный аргумент: IP Switching позволяет устройству, обладающему функциональными возможностями ATM-коммутатора, выполнять также и работу маршрутизатора, а мэппингу между IP и ATM – вообще не использовать управляющие протоколы

ATM. В мае 1996 г. вышел документ RFC 1953 «Ipsilon Flow Management Protocol Specification for Ipv4. Version 1.0», а в августе того же года – RFC 1987 «Ipsilon's General Switch Management Protocol Specification. Version 1.1». Эти публикации позволили компании Ipsilon официально заявить, что их технология является открытой, поскольку использованные в ней базовые протоколы общедоступны. Коммутация по меткам в IP Switching фактически основывалась на классификации потоков по таким параметрам, как IP-адрес и номер порта отправителя, IP-адрес и номер порта получателя, тип протокола. Потоки классифицировались как устойчивые (пересылка файлов по протоколу FTP, трафик HTTP, Telnet и др.) и как кратковременные (обращения к системе имен доменов DNS, сообщения протокола SNMP и протокола сетевого времени NTP). *Протокол Ipsilon Flow Management Protocol (IFMP)* относил трафик к тому или иному классу и помогал создавать виртуальный канал, требующийся для пересылки трафика этого класса, а для того чтобы конфигурировать коммутационное поле ATM-коммутатора, использовался *протокол Generic Switch Management Protocol (GSMP)*, который позволял добавлением внешнего контроллера превратить практически каждый коммутатор ATM в коммутатор пакетов IP. Механизм IP Switching был чрезвычайно важным шагом, поскольку предлагал жизнеспособную парадигму замены меток, а также вводил метод классификации IP-трафика.

Вскоре компания *Cisco Systems* анонсировала свой вариант технологии коммутации по меткам под названием *коммутация по тегам (Tag Switching)*, которая существенно отошла от двух рассмотренных выше технологий IP Switching и CSR. В частности, для создания таблиц пересылки в коммутаторе она не опиралась на поток трафика данных и к тому же была специфицирована для ряда технологий уровня 2, отличных от ATM. Таким образом, технология Tag Switching оказалась намного ближе к окончательной концепции MPLS, чем механизм IP Switching. Более того технология MPLS в значительной степени вышла из механизма Tag Switching. Сам так называемый *тег*, т. е. фиксированное количество битов, используемых для адресации, во многом аналогичен метке MPLS. Механизм Tag Switching предназначался для совместной работы с рядом протоколов нижних уровней и включал в себя *протокол распределения тегов (Tag Distribution Protocol, TDP)*. Как и в MPLS, механизм Tag Switching поддерживал образование стека тегов. Новые маршрутизаторы не только осуществляли более быстрый поиск адреса, но и могли обслуживать вызовы по-разному, в зависимости от требуемого качества обслуживания (при передаче речи, видео и изображений). Кроме того, все маршрутизаторы производства Cisco, в которых был реализован механизм Tag Switching, были позже модернизированы и смогли поддерживать MPLS.

Как и Ipsilon, Cisco Systems выпустила RFC, в котором была описана предлагаемая технология. Однако в отличие от Ipsilon компания Cisco объявила о своем намерении провести стандартизацию технологии Tag Switching через IETF. В связи с этим было выпущено большое число проектов интернет-стандартов, описывающих разные аспекты технологии Tag Switching, включая функционирование в сети ATM, с протоколами PPP и каналами 802.3, поддержку многоадресной маршрутизации, а также функций резервирования ресурсов с помощью протокола RSVP.

Практически сразу же после того, как Cisco опубликовала информацию о технологии Tag Switching и объявила об ее предполагаемой стандартизации в IETF, от корпорации IBM поступили проекты интернет-стандартов, в которых предлагалась другая технология коммутации по меткам – *Aggregate Route-based IP Switching (ARIS).* Механизм ARIS предназначался для использования с ATM- и FR-коммутаторами, а также с устройствами коммутации на уровне 2 в локальных сетях. Устройство, в котором был реализован механизм ARIS, получило название *ARIS Integrated Switch Router (ISR)*. Технология ARIS имеет больше общих черт с технологией Tag Switching, нежели с другими уже упоминавшимися технологиями, – в обеих для создания таблиц пересылки используется трафик управляющей информации, а не трафик данных, – но при этом технология ARIS имеет некоторые существенные отличия от Tag Switching. Основное отличие состоит в том, что ARIS основан на маршрутах, а не на потоках. Маршруты в домене ARIS строятся на базе выходного узла. Конфигурируются выходные узлы домена ARIS, а затем от них распространяются маршруты в сторону входных узлов. Выходной узел может быть задан рядом идентификаторов: префиксом получателя протокола IPv4, IP-адресом выходного маршрутизатора, идентификатором маршрутизатора OSPF или идентификатором пары многоадресной передачи. Маршруты устанавливаются независимо от потоков пакетов. Многие из идей технологии ARIS перешли в окончательный стандарт MPLS.

Еще одной предшествовавшей MPLS технологией является *IP Navigator*, предложенная компанией Cascade. Cascade была затем куплена компанией Ascend, которая в свою очередь стала частью Lucent Technologies. В технологии IP Navigator были использованы многие идеи коммутации в IP-сетях, разработанные ранее компаниями Toshiba, Ipsilon, Cisco и IBM.

После публикации первой серии проектов стандартов Tag Switching 9–13 декабря 1996 г. в Сан-Диего, Калифорния, состоялась рекордная за всю историю IETF по посещаемости сессия BOF, на которой Cisco Systems, IBM и Toshiba провели презентации своих технологий. Такой интерес, а также тот факт, что столь много ведущих компаний разработали

во многом близкие технические предложения для решения проблемы, позволили сделать очевидный вывод о необходимости создать для стандартизации механизма коммутации по меткам специальную группу. В апреле 1997 г. в Мемфисе, Теннеси, состоялось первое заседание этой рабочей группы MPLS WG. Само название *Multiprotocol Label Switching* было принято в первую очередь по той, уже упомянутой нами причине, что названия *IP Switching* и *Tag Switching* ассоциировались с продуктами, выпускаемыми конкретными компаниями, и требовался нейтральный термин (рис. 2.1).

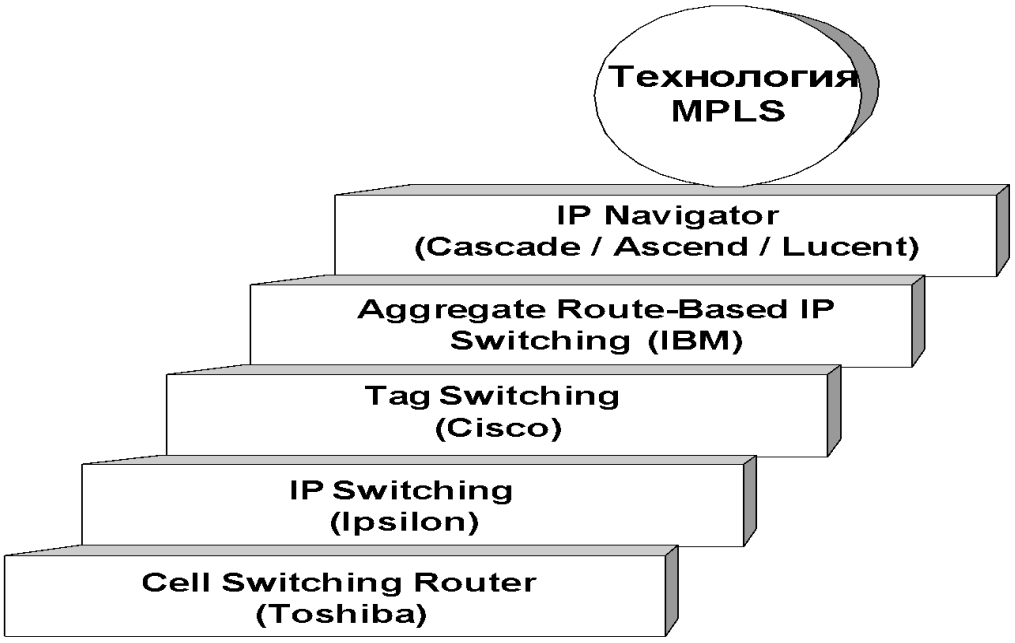

Рис. 2.1. Эволюция технологий

Архитектура MPLS специфицирована в документе RFC 3031 «Multiprotocol Label Switching Architecture» (табл. 2.1). Сегодня вопросами MPLS продолжают заниматься рабочие группы IETF (Routing Area Working Group – рабочая группа по маршрутизации; MPLS Working Group – рабочая группа по MPLS) и в ATM Forum (Traffic Management Working Group – рабочая группа по управлению трафиком; ATM-IP Collaboration Working Group – рабочая группа по совместной работе сетей ATM и IP).

Таблица 2.1

Документы серии RFC, разработанные комитетом IETF по MPLS

RFC	Описание
RFC 2702	Requirements for Traffic Engineering over MPLS – определяет возможности управления трафиком в сети MPLS и алгоритмы эффективных и надежных сетевых операций в домене MPLS. Эти алгоритмы могут использоваться для оптимизации использования сетевых ресурсов и для улучшения рабочих характеристик, связанных с передаваемым трафиком
RFC 3031	MPLS Architecture – специфицирует архитектуру многопротокольной коммутации по меткам MPLS
RFC 3032	MPLS Label Stack Encoding – специфицирует кодирование стека меток, а также правила и процедуры обработки разных полей стека меток, которые использует маршрутизатор LSR для передачи снабженных метками пакетов по звеньям данных протокола двухточечной связи PPP, звеньям данных локальной вычислительной сети и, возможно, по другим звеньям данных
RFC 3033	The Assignment of the Information Field and Protocol Identifier in the Q.2941 Generic Identifier and Q.2957 User-to-User Signaling for the IP – специфицирует назначение информационного поля и идентификатора протокола в общем идентификаторе Q.2941 и сигнализации «пользователь-пользователь» по Q.2957 для IP
RFC 3034	Use of Label Switching on Frame Relay Networks Specification – определяет модель и типовые механизмы использования MPLS в сетях Frame Relay. Расширяет и уточняет составляющие архитектуры MPLS и протокола распределения меток LDP в плане их использования в сетях Frame Relay
RFC 3035	MPLS using LDP and ATM Virtual Channel (VC) Switching – специфицирует процедуры, используемые при распределении меток к/от маршрутизаторов ATM-LSR, когда эти метки представляют классы эквивалентности пересылки (FEC), для которых алгоритмами маршрутизации сетевого уровня определены маршруты «по участкам». Специфицирует также инкапсуляцию MPLS, которая используется при передаче снабженных метками пакетов к/от маршрутизаторов ATM-LSR
RFC 3036	LDP Specification – определяет набор процедур протокола LDP, посредством которого LSR распределяют метки для пересылки пакетов MPLS
RFC 3037	LDP Applicability – описывает применимость протокола LDP
RFC 3038	Virtual Channel ID (VCID) Notification over ATM link for LDP – специфицирует процедуры обмена значениями идентификатора виртуального канала VCID между смежными маршрутизаторами ATM-LSR
RFC 3107	Carrying Label Information in BGP-4 – специфицирует способ, посредством которого информация о привязке метки к FEC для определенного маршрута вкладывается в то же сообщение протокола BGP, которое используется для рассылки информации о самом маршруте. Когда для выбора определенного маршрута используется протокол BGP, он может также использоваться для передачи метки MPLS, которая назначена для этого маршрута

3. ТЕХНОЛОГИЯ КОММУТАЦИИ ПО МЕТКАМ

MPLS может рассматриваться как совокупность технологий, которые, работая совместно, обеспечивают доставку пакетов от отправителя к получателю контролируемым, эффективным и предсказуемым способом. В MPLS для пересылки пакетов на уровне 2 используются коммутируемые по меткам тракты LSP, которые были организованы с помощью протоколов маршрутизации и сигнализации уровня 3.

В дополнение к приведенным в табл. 3.1 рассмотрим еще некоторые базовые понятия, важные для понимания принципов работы технологии MPLS, такие как *пересылка, коммутация* и *маршрутизация*.

Таблица 3.1

Основные понятия

Понятие	Пояснение
FEC (Forwarding Equivalence Class) – класс эквивалентности пересылки	Множество пакетов, которые пересылаются одинаково, например с целью обеспечить заданное QoS
Label – метка	Короткий идентификатор фиксированной длины, определяющий принадлежность пакета тому или иному FEC
Label swapping – замена меток	Замена метки принятого узлом сети MPLS пакета новой меткой, связанной с тем же FEC, при пересылке этого пакета к нижестоящему узлу
LER (MPLS edge router – пограничный узел сети MPLS)	Пограничный узел сети MPLS, который соединяет домен MPLS с узлом, находящимся вне этого домена
Loop detection – выявление закольцованных маршрутов	Метод, позволяющий обнаружить, что пакет прошел через узел более одного раза
Loop prevention – предотвращение образования закольцованных маршрутов	Метод выявления и устранения закольцованных маршрутов
LSP (Label Switched Path) – коммутируемый по меткам тракт	Путь через один или более LSR, по которому следуют пакеты одного и того же FEC
LSP-ER (explicitly routed LSP) – LSP с явно заданным маршрутом	Тракт LSP, который установлен способом, отличным от традиционной маршрутизации пакетов IP
LSR (Label Switching Router) – маршрутизатор коммутации по меткам	Маршрутизатор, способный пересылать пакеты по технологии MPLS
MPLS domain – домен MPLS	Совокупность узлов MPLS, между которыми существуют непрерывные LSP
MPLS egress node – выходной узел сети MPLS	Последний MPLS-узел в LSP, направляющий исходный пакет к адресату, который находится вне MPLS-сети
MPLS ingress node – входной узел сети MPLS	Первый MPLS-узел в LSP, принимающий исходный пакет и помещающий в него метку MPLS

Маршрутизация – это выбор маршрута или того его элемента, который ведет к ближайшему узлу, входящему в этот маршрут, как правило, функция уровня 3 модели OSI. Очень важно правильно воспринять это понятие, потому что технология MPLS дополняет его общепринятую трактовку и «вклинивается» между сетевым уровнем 3 и уровнем звена данных 2. Маршрутизация в традиционном смысле, без MPLS, представляет собой процесс определения следующего участка, по которому должен пойти пакет в направлении получателя, путем анализа заголовка сетевого уровня. Процесс маршрутизации в каждом маршрутизаторе использует различные протоколы и алгоритмы маршрутизации для отыскания маршрутов и создания таблицы пересылки, которая используется уже в плоскости пересылки данных.

Коммутация – это выбор исходящего порта в соответствии с результатом маршрутизации и создание связи между входящим и выбранным исходящим портами, т. е. создание внутри узла условий (можно сказать, внутриузлового пути) для отправки пакета по уже выбранному маршруту. Как правило, это функция уровня 2 модели OSI. В традиционном смысле *коммутатор* представляет собой устройство, которое принимает пакеты во входных портах, анализирует информацию заголовка уровня 2 (звена данных), использует свои внутренние таблицы коммутации, чтобы создать условия для отправки пакетов к надлежащим выходным портам. Обычно коммутаторы работают быстрее маршрутизаторов, но имеют меньше функциональных возможностей. Добавление в коммутатор функций MPLS превращает его в LSR.

Пересылка – это использование созданных посредством коммутации условий (внутриузлового пути) для того, чтобы передать пакет из входящего порта по упомянутому маршруту через выбранный при коммутации исходящий порт.

Это несколько упрощенные определения, но они являются хорошей отправной точкой для обсуждения в следующей главе того, что же такое метки MPLS, как именно они распределяются и как по ним производится коммутация.

4. КЛАССЫ ЭКВИВАЛЕНТНОСТИ ПЕРЕСЫЛКИ FEC

Рабочими группами, упомянутыми в предыдущем разделе, определены три основных элемента технологии MPLS: **FEC** – Forwarding Equivalency Class – класс эквивалентности пересылки, **LSR** – Label Switching Router – маршрутизатор коммутации по меткам и **LSP** – Label Switched Path – коммутируемый по меткам тракт. Начнем с классов эквивалентности пересылки.

При традиционной транспортировке пакета через сеть с использованием в уровне 3 протокола, не предусматривающего создания виртуальных соединений, каждый маршрутизатор на пути следования пакета самостоятельно принимает решение о том, к какому маршрутизатору переслать этот пакет дальше (способ транспортировки *hop-by-hop*). Иначе говоря, в каждом маршрутизаторе на пути следования пакета анализируется его заголовок и выполняется алгоритм сетевого уровня. Здесь и далее используется английское слово *hop* – прыжок, скачок, – а в терминах маршрутизации – «одна пересылка». Под пересылкой пакета понимается его передача к ближайшему маршрутизатору из тех, что расположены на возможном пути следования этого пакета, т. е. слово «пересылка» используется как эквивалент английского слова *forwarding*.

В заголовке пакета содержится гораздо больше информации, чем нужно для того, чтобы выбрать следующий маршрутизатор. Этот выбор можно организовать проще – путем выполнения двух функций. Одна из них состоит в разделении всего множества прибывающих пакетов на классы, которые называются *классами эквивалентности пересылки FECs (Forwarding Equivalence Classes)*. Вторая ставит в соответствие каждому FEC определенное «направление» пересылки (слово «направление» написано в кавычках потому, что в сети используется режим hop-by-hop, и разные пакеты одного и того же FEC могут пересылаться к разным маршрутизаторам, т. е. физические направления пересылки могут быть разными). С точки зрения выбора следующего маршрутизатора все пакеты, принадлежащие одному FEC, неразличимы.

Идея классов эквивалентности более универсальна, чем MPLS. При традиционной IP-маршрутизации тот или иной маршрутизатор тоже может считать, что два пакета принадлежат одному и тому же условному классу эквивалентности, если в его таблицах маршрутизации используется некий адресный префикс, идентифицирующий направление, в котором предполагаемые маршруты транспортировки этих двух пакетов совпадают наиболее долго. По мере продвижения пакета по сети каждый следующий маршрутизатор анализирует его заголовок и приписывает этот пакет к такому из собственных (определенных только в данном маршрутизаторе) классов эквивалентности, который соответствует тому же направлению.

При использовании многопротокольной коммутации по меткам MPLS пакет приписывается к определенному классу FEC только один раз, когда он попадает в сеть. Этому FEC присваивается *метка* – идентификатор фиксированной длины, передаваемый вместе с пакетом, когда тот пересылается к следующему маршрутизатору. Благодаря этому в остальных маршрутизаторах заголовок сетевого уровня не анализируется. *Метка*, установленная пограничным маршрутизатором при входе пакета в MPLS-сеть, используется как указатель входа таблицы, которая определяет очередной маршрутизатор для пересылки к нему пакета, а также новую метку для FEC, к которому относится пакет.

Таким образом, *класс эквивалентности пересылки FEC* является формой представления группы пакетов с одинаковыми требованиями к направлению их передачи, т. е. все пакеты в такой группе обрабатываются в маршрутизаторе одинаково и одинаково следуют к пункту назначения. Примером FEC могут служить все IP-пакеты с адресами пунктов назначения, соответствующими некоторому префиксу, например 212.18.6. Возможны также FEC на основе префикса адреса и еще какого-нибудь поля IP-заголовка (например, тип обслуживания – ToS). Каждый маршрутизатор сети MPLS создает таблицу, с помощью которой определяет, каким образом должен пересылаться пакет. Эта таблица, которая называется *информационной базой меток LIB*, содержит используемое множество меток и для каждой из них – *привязку «FEC-метка»*. Метки, используемые маршрутизатором LSR при привязке «FEC-метка», подразделяются на следующие категории:

- *на платформенной основе,* когда значения меток уникальны по всему тракту LSP; метки выбираются из общего пула меток, и никакие две метки, распределяемые по разным интерфейсам, не имеют одинаковых значений;
- *на интерфейсной основе,* когда значения меток связаны с интерфейсами: для каждого интерфейса определяется отдельный пул меток, из которого для этого интерфейса и выбираются метки. При этом метки, назначаемые для разных интерфейсов, могут быть одинаковыми.

Метод пересылки пакетов на основе привязки «FEC-метка», принятый в MPLS, имеет ряд преимуществ перед методами, основанными на анализе заголовка блоков сетевого уровня. В частности, пересылку по методу MPLS могут выполнять маршрутизаторы, которые способны читать и заменять метки, но при этом либо вообще не способны анализировать заголовки блоков сетевого уровня, либо не способны делать это достаточно быстро.

Так или иначе, действия маршрутизатора LSR зависят от значения метки, которую он принимает от предшествующего LSR. Фактически действия, выполняемые LSR, специфицированы в *Next Hop Level Forwarding Entry (NHLFE)*, указывающем следующий участок, операцию, которая должна

быть выполнена со стеком меток, и кодирование, которое следует использовать для стека в исходящем тракте. Выполняемая со стеком операция может состоять в том, что LSR должен изменить метку на вершине стека. Эта операция может потребовать, чтобы LSR просто вытолкнул верхнюю метку из стека, или вытолкнул и заменил ее новой, или просто поместил новую метку над той, которая до этого была верхней, ничего не выталкивая и не заменяя. Следующим участком для обрабатываемого пакета с метками может оказаться и тот же самый LSR. В этом случае LSR выталкивает верхнюю метку стека и пересылает пакет самому себе. В этот момент пакет может иметь еще одну метку, которую следует анализировать, или оказаться без меток, т. е. исходным пакетом IP. В последнем случае пакет пересылается в соответствии со стандартной маршрутизацией IP.

Если маршрутизатор обнаруживает, что он оказался предпоследним LSR в тракте, то он должен удалить весь стек и передать пакет в последний LSR. Благодаря этому минимизируется объем обработки, которую должен выполнить последний LSR. То, каким образом LSR определяет, что он в данном тракте предпоследний, является задачей распределения меток и используемого для этого протокола распределения меток. Но прежде поясним, что же представляет собой этот тракт.

5. КОММУТИРУЕМЫЕ ПО МЕТКАМ ТРАКТЫ LSP

Путь следования потока пакетов в сети MPLS определяется тем FEC, который установлен для этого потока во входном LSR. Такой путь носит название *коммутируемого по меткам тракта LSP* (*Label-Switched Path)* и идентифицируется последовательностью меток в LSR, расположенных на пути следования потока от отправителя к получателю.

LSP организуются либо перед передачей данных (*с управлением программой*), либо при обнаружении определенного потока данных.

Метки в LSP назначаются с помощью *протокола распределения меток LDP (Label Distribution Protocol)*, рассматриваемого в разделе 6, причем существуют разные способы такого распределения на основе данных вспомогательных протоколов, в частности протокола RSVP-TE. Подготавливают процесс распределения меток протоколы маршрутизации, такие как OSPF, IS-IS или BGP. С помощью этих протоколов маршрутизации создается «древо» сети, на которое «развешиваются» метки.

Главная задача распределения меток – это организация и обслуживание трактов LSP, в том числе определение каждой привязки «FEC-метка» в каждом LSR тракта LSP. Маршрутизатор LSR использует протокол распределения меток, чтобы информировать о привязке «FEC-метка» вышестоящий LSR. Нижестоящий LSR может непосредственно сообщать о привязке «метка-FEC» вышестоящему LSR, что называется привязкой *по инициативе нижестоящего* (*unsolicited downstream*). Кроме того, возможно извещение о привязке, передаваемое *нижестоящим по требованию* (*downstream on demand*), когда вышестоящий LSR запрашивает привязку у нижестоящего LSR. Организуемый LSP всегда является односторонним. Трафик обратного направления идет по другому LSP.

Технология MPLS поддерживает следующие два варианта создания LSP:

● *последовательная маршрутизация по участкам маршрута (hop-by-hop routing)* – каждый LSR самостоятельно выбирает следующий участок маршрута для данного FEC. Эта методология сходна с той, что применяется сейчас в IP-сетях. LSR использует имеющиеся протоколы маршрутизации, такие, например, как OSPF;

● *явная маршрутизация (ER)* – сходна с методом маршрутизации со стороны отправителя. Входной LSR (т. е. LSR, от которого исходит поток данных в сети MPLS) специфицирует цепочку узлов, через которые проходит ER-LSP. Специфицированный тракт может оказаться не оптимальным. Вдоль тракта могут резервироваться ресурсы для обеспечения заданного QoS трафика данных. Это облегчает оптимальное распределение трафика по всей сети и позволяет предоставлять дифференцированное обслуживание потокам трафика разных классов, сформированных на основе принятых правил и методов управления сетью.

Рассмотрим логически завершенный (и в определенном смысле автономный) домен сети MPLS, изображенный на рис. 5.1. Завершенность этого домена выражается в том, что он имеет вполне определенную замкнутую границу, вдоль которой размещено четыре так называемых *пограничных узла MPLS* (MPLS edge nodes или, как их еще иногда называют, LER – Label Edge Router), обозначенных на рис. 2.1 как LSR1, LSR5, LSR6, LSR7. Помимо этих узлов, внутри домена сети MPLS (когда это не вызывает двоякого толкования, мы будем для удобства называть его просто MPLS-сетью) имеется множество маршрутизаторов, каждый из которых имеет с остальными маршрутизаторами (в том числе и с пограничными узлами) либо прямые, либо коммутируемые связи. В последнем случае коммутация, необходимая для создания такой связи, производится другими маршрутизаторами из этого множества, которые не обязательно являются пограничными узлами MPLS и могут не иметь функций LSR. Более того, некоторые коммутируемые связи между LSR могут проходить через подсети, встроенные в рассматриваемую MPLS-сеть, но не содержащие в себе функции MPLS. Они, разумеется, не показаны в примере на рис. 5.1 где изображены только три внутренних маршрутизатора LSR2, LSR3 и LSR4.

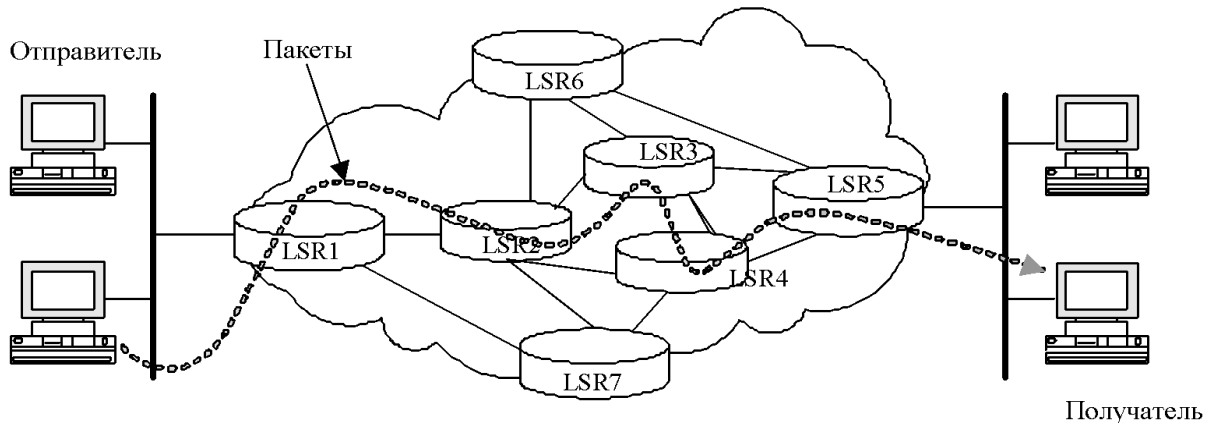

Рис. 5.1. Пример домена MPLS-сети

Напомним, что на рис. 5.1 изображен лишь упрощенный домен MPLS-сети. Пакеты, поступающие в него, могут приходить как непосредственно от отправителей (что показано на рис. 5.1), так и из смежной сети, которая может быть MPLS-сетью более высокого уровня (т. е. содержать в себе рассматриваемый домен). Эти пакеты принимаются пограничным узлом MPLS (в данном случае LSR1), который является по отношению к этим пакетам *входным MPLS-узлом*. Пакеты, направляемые сетью в другую смежную сеть, передаются туда другим пограничным узлом, который является по отношению к этим пакетам *выходным MPLS-узлом* (в данном случае LSR5). В общем случае все пакеты, транспортируемые через MPLS-сеть от входного MPLS-узла *LSR1* к выходному MPLS-узлу *LSR*5, принадлежат

одному FEC и следуют по одному и тому же виртуальному *коммутируемому по меткам тракту LSP*, который может проходить через несколько LSR и маршрутизаторов без функций LSR.

Таким образом, в MPLS-сети имеются маршрутизаторы двух типов: пограничные LSR и транзитные LSR. Пограничные маршрутизаторы LSR в ряде случаев включают в себя шлюзы интерфейсов сетей разных видов (например, Frame Relay, ATM или Ethernet) и пересылают их трафик в MPLS-сеть после организации трактов LSP, а также распределяют трафик обратного направления при выходе его из MPLS-сети. К этому следует добавить, что любой MPLS-совместимый маршрутизатор должен быть способен принимать в любом своем интерфейсе пакет со вставленной меткой, отыскивать ее в таблице коммутации, вставлять новую метку в соответствующем формате и затем отправлять пакет через другой интерфейс. Иными словами, пограничный LSR может коммутировать пакет с меткой от любого интерфейса к любому другому интерфейсу с заменой метки. Такой подход гораздо гибче, чем в случае ATM, так как он не ограничен исключительно каналами передачи ячеек. Пограничные маршрутизаторы выполняют основную роль в процессе назначения и удаления меток, когда трафик поступает в MPLS-сеть или выходит из нее.

При этом полезно отметить, что любой транзитный LSR способен принимать пакеты без меток, т. е. с обычными IP-заголовками. Довольно часто встречающееся в литературе утверждение, что внутри домена MPLS пакеты между транзитными LSR маршрутизируются только по меткам, не совсем верно. Для обычного MPLS-трафика это действительно так, но служебные сообщения передаются с использованием IP-заголовков.

К выходному узлу LSR5 (рис. 5.1) поступают потоки пакетов от нескольких входных узлов (от LSR1, LSR6 и LSR7). В промежуточных маршрутизаторах некоторые из этих потоков могут «сливаться», т. е. объединяться в один общий поток пакетов, которые приобретают в этой точке слияния общий FEC. Таким образом, множество трактов LSP, идущих к одному выходному узлу, образует ветвящееся дерево, корень которого находится в этом выходном узле.

Каждый из четырех пограничных узлов выполняет, в общем случае, функции и входного, и выходного узла, так что в изображенной на рис. 5.1 MPLS-сети существует четыре дерева такого рода, которые вместе содержат $4 \cdot (4 - 1) = 12$ трактов LSP. Ясно, что через один промежуточный маршрутизатор LSR может проходить несколько LSP, в том числе LSP, принадлежащих разным деревьям. Если учесть, что физическая топология сети отличается от топологии виртуальной сети LSP (и еще раз вспомнить про режим hop-by-hop), то станет ясно, что на практике могут возникать случаи «закольцовывания» путей прохождения пакетов, и, следовательно, в MPLS-сетях нужно предусматривать меры обнаружения и/или предотвращения таких случаев.

6. ПРОТОКОЛ LDP

В этом разделе речь пойдет о том, каким образом производится распределение меток по всем LSR сети MPLS с использованием протокола LDP (Label Distribution Protocol).

В спецификации LDP к настоящему моменту установлены два типа элементов, с помощью которых может определяться FEC:

● *Address Prefix* – адресный префикс любой длины от нуля до полного адреса;

● *Host Address* – полный адрес хоста.

Решения о назначении меток могут основываться на критериях пересылки, таких как одноадресная маршрутизация к получателю, оптимизация распределения трафика в сети, многоадресная рассылка, виртуальная частная сеть VPN, механизмы обеспечения качества обслуживания QoS и др. Спецификация же протокола LDP определяет правила, по которым устанавливается соответствие между входным пакетом и его LSP.

Для распределения меток могут использоваться разные методы:

● *метод на основе топологии (topology-based method)* – использует стандартную обработку протоколов маршрутизации (например, OSPF и BGP)

● *метод на основе запросов (request-based method)* – использует обработку управляющего протокола на основе запросов (например, протокола RSVP);

● *метод на основе трафика (traffic-based method)* – запускает процедуру присвоения и распределения меток при получении пакета.

Методы на основе топологии и запросов являются примерами привязки меток к FEC, управляемой от программы, а метод на основе трафика является примером привязки, управляемой данными. Во всех этих случаях архитектурой MPLS предусматривается, что назначение метки, т. е. ее привязку к определенному FEC, производит LSR, который является выходным пограничным маршрутизатором для пакетов этого FEC. По-английски такой LSR называется *downstream LSR*, т. е. расположенный «ниже по течению», мы будем называть его *нижним* или *нижестоящим LSR*, а расположенный «выше по течению» – *upstream LSR* – будем называть *верхним* или *вышестоящим LSR*. Таким образом, назначение меток всегда производится *снизу*, т. е. в сторону, противоположную направлению трафика. Нижний LSR информирует соседние верхние LSR о том, какие метки он привязал к каждому FEC поступающих к нему пакетов. Этот процесс и называется *распределением меток*, а обеспечивает его *протокол распределения меток LDP* (*Label Distribution Protocol*).

Подчеркнем, что архитектура MPLS не требует обязательного применения LDP. Для распределения меток могут применяться модификации существующих протоколов маршрутизации, позволяющие использовать их для передачи информации о метках, например протокол BGP.

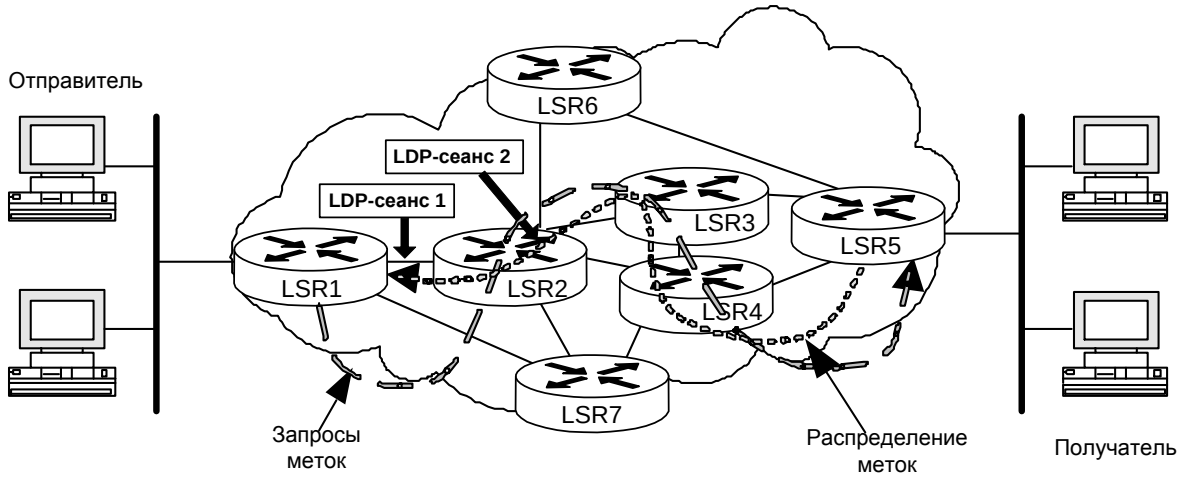

Рис. 6.1. Пример фрагмента MPLS-сети

Протокол RSVP также имеет расширения, обеспечивающие поддержку обмена метками с уведомлением. Однако протокол распределения меток LDP был признан комитетом IETF наиболее удачным и, что еще важнее, хорошо специфицирован им. Кроме того, определено расширение базового протокола LDP для поддержки явной маршрутизации с учетом обеспечения качества обслуживания QoS и управления трафиком TE – *протокол LDP с учетом ограничивающих условий CR-LDP (Constraint-Based LDP)*. Ко всему прочему, LDP устанавливает надежные транспортные соединения со смежными маршрутизаторами LSR по протоколу TCP, причем в случае, если между двумя LSR надо одновременно установить несколько LDP-сеансов, используется единственное TCP-соединение.

Имеются следующие схемы обмена метками:

- LDP – преобразует в метки IP-адреса получателя при одноадресной передаче;
- RSVP и CR-LDP – используются для оптимизации распределения трафика в сети и для резервирования ресурсов;
- BGP – работает с внешними метками VPN.

Все эти схемы рассматриваются далее.

7. ОСНОВЫ ПРОТОКОЛА LDP

Протокол LDP разработан комитетом IETF в рабочей группе по MPLS с целью специфицировать процедуры распределения меток MPLS. В связи с тем что протокол LDP тесно взаимодействует с протоколами внутренней маршрутизации IGP, его часто называют *протоколом пересылки по участкам*. Протокол распределения меток LDP представляет собой *набор процедур и сообщений, при помощи которых LSR организуют тракты коммутации по меткам (LSP), обмениваясь информацией о привязке меток к FEC с «соседними» LSR, а также поддерживают и прекращают LSP-сеансы.*

Речь идет именно об *обмене* информацией, поскольку протокол предусматривает двусторонний диалог взаимодействующих LSR, являющихся в данном контексте *одноранговыми узлами LDP*. Этот диалог называется *LDP-сеансом*, в ходе которого каждый из взаимодействующих LSR получает сведения о привязке меток к FEC в другом LSR. В протоколе определен также механизм передачи уведомлений и обнаружения в LSP закольцованных маршрутов.

Итак, процедуры протокола LDP позволяют LSR, выполняющему этот протокол, создавать тракты LSP. Конечной точкой такого тракта является либо смежный LSR, имеющий прямую связь с данным LSR, либо выходной LSR, доступный этому LSR через некоторое количество транзитных LSR. Процессы обнаружения конечных точек этих двух типов называются соответственно процессом *базового обнаружения* и процессом *расширенного обнаружения*. LDP создает двустороннюю связь двух смежных LSR, которые становятся одноранговыми узлами LDP и взаимодействуют друг с другом посредством LDP-сеанса. При обмене между LSR информацией, связанной с привязкой «метка-FEC», используются четыре категории сообщений LDP:

- *сообщения обнаружения (discovery messages),* используемые для того, чтобы объявить и поддерживать присутствие LSR в сети;
- *сеансовые сообщения (session messages),* предназначенные для создания, поддержки и прекращения LDP-сеансов между LSR;
- *сообщения-объявления (advertisement messages),* используемые для создания, изменения и отмены привязки метки к FEC;
- *уведомляющие сообщения (notification messages),* содержащие вспомогательную информацию и информацию об ошибках.

Хотя «раздает» метки всегда нижний LSR, инициатором их распределения не обязательно должен быть он. Процесс может инициировать и верхний LSR, направив к нижнему LSR соответствующий запрос. Такой режим называется *downstream on-demand*. В той или иной сети может использоваться распределение меток либо только по запросам сверху, либо только по инициативе нижнего LSR *(unsolicited downstream)*, либо и то и другое вместе.

Распределение меток может быть *независимым* или *упорядоченным*. В первом случае LSR может уведомить вышестоящий LSR о привязке метки к FEC до того, как получит информацию о привязке от нижестоящего LSR. Во втором случае высылать подобное уведомление «наверх» разрешается только после получения таких сведений «снизу».

Заметим, что нижний LSR распределяет метки не только по тем верхним LSR, которые имеют с ним прямые связи. Протокол распределения меток может быть использован и для диалога двух LSR, между которыми существует лишь коммутируемая связь, однако результат распределения в этом случае зависит от того, в каком из двух режимов, либеральном или консервативном, работает верхний LSR.

Консервативный режим распределения меток. В этом режиме сообщения-объявления о привязке «метка-FEC», получаемые от несмежных LSR, не принимаются и отбрасываются. LSR привязывает метку к FEC только в том случае, если он является выходным маршрутизатором или если он получил сообщение о привязке от смежного с ним LSR. Такой режим позволяет LSR обслуживать меньшее число меток.

Либеральный режим распределения меток. В этом режиме привязка метки, выданная тем нижним LSR, с которым нет прямой связи, запоминается и используется. Такой режим удобен тем, что при реконфигурации сети соответствие между меткой и FEC сохраняется, даже если связь с LSR, определившим это соответствие, стала не коммутируемой, а прямой. Недостаток либерального режима состоит в том, что в верхнем LSR приходится хранить и обрабатывать заметно больше информации о соответствии между метками и FEC.

Как уже говорилось, метка всегда локальна, т. е. обозначает некоторый FEC только для пары маршрутизаторов, между которыми имеется прямая или коммутируемая связь, и используется при пересылке пакетов этого FEC от того из маршрутизаторов данной пары, который является в ней верхним (*upstream LSR*), к тому, который является нижним (*downstream LSR*). Для пересылки пакетов того же FEC к следующему маршрутизатору используется другая метка, идентифицирующая этот FEC для новой пары маршрутизаторов, в которой маршрутизатор, бывший в предыдущей паре нижним, приобретает статус верхнего, а статус нижнего получает второй маршрутизатор этой новой пары. Отсюда ясно, что каждый маршрутизатор MPLS-сети должен хранить соответствие между входящими и исходящими метками для всех FEC, которыми он оперирует.

Одной из важнейших функций протокола LDP является обнаружение петель. Для этой цели можно использовать два поля в сообщениях *Label Request* и *Label mapping* (рассматриваемых в разделах 6 и 7), а именно *Path Vector* и *Hop Count*.

Поле *Path Vector* содержит список *LSR-идентификаторов* (первые 4 октета LDP-идентификатора), принадлежащих тем LSR, через которые прошло содержащее его сообщение. Если LSR получает сообщение и обнаруживает в поле *Path Vector* свой собственный LSR-идентификатор, он «понимает», что возникла петля. В протоколе LDP предусматривается возможность задать максимально допустимое значение поля *Path Length,* по достижении которого тоже принимается решение о возникновении петли.

Второй вариант – поле *Hop Count,* которое содержит счетчик LSR, пройденных сообщением LSR. Каждый пройденный LSR увеличивает его значение на единицу. Маршрутизатор, обнаруживший, что счетчик достиг максимально допустимой величины, обрабатывает сообщение, как сделавшее петлю.

8. ПРОТОКОЛ CR-LDP

Протокол LDP может следовать только таблицам маршрутизации IP. Чтобы преодолеть это ограничение, было предложено расширение LDP, названное CR-LDP.

Протокол CR-LDP является вариантом LDP, в котором определены механизмы создания и поддержания трактов LSP с явно заданной трассой. Для создания тракта CR-LSP используется больше информации, чем можно получить от традиционных протоколов внутренней маршрутизации. CR-LDP используется для таких приложений MPLS, как TE и QoS, где требуется дополнительная информация о маршрутах. В этом протоколе запрос метки может не следовать слепо вдоль дерева маршрутизации для данного адресата, так как предусмотрена возможность точно сообщить, как он должен следовать, включив в сообщение явно заданный маршрут. При этом программное обеспечение CR-LDP не использует для маршрутизации запроса таблицы пересылки, а маршрутизирует его в соответствии с содержащимися в сообщении инструкциями.

Протокол CR-LDP не поддерживает динамического вычисления явно задаваемых маршрутов, поэтому сведения о динамическом резервировании пропускной способности должны включаться в вещательную информацию протоколов OSPF или IS-IS, или в извещения о состоянии каналов LSA. Используя эти механизмы, CR-LDP может занимать и резервировать пропускную способность. Доступная пропускная способность изменяется в соответствии с запросом, и ее новое значение рассылается другим узлам с помощью расширений протоколов OSPF и IS-IS, которые будут рассмотрены ниже. Новые маршруты можно вычислить затем с учетом принятых ограничений, используя модифицированный алгоритм Дейкстры. В результате протокол CR-LDP получает в свое распоряжение явный маршрут для организации LSP. Тракт создается посредством запроса метки, содержащего динамически вычисляемый явный маршрут.

Протокол CR-LDP имеет также другие, новые по сравнению с базовой версией LDP функциональные возможности:

– явная маршрутизация с точно определенными и свободными маршрутами, при которой маршрут задается в виде последовательности групп узлов. В случае если в группе указано более одного маршрутизатора, при создании явного маршрута возможна определенная гибкость;

– спецификация параметров трафика (например, пиковая скорость передачи, гарантированная скорость передачи и допустимая вариация задержки);

– закрепление маршрута (*route pinning*), которое может использоваться в тех случаях, когда изменять трассу LSP нежелательно, например в сегментах со свободной маршрутизацией, когда в этом сегменте становится доступным лучший маршрут;

– механизм приоритетного вытеснения LSP с помощью системы приоритетов создания и удержания. Ранжируются существующие тракты LSP (приоритет удержания) и новые тракты LSP (приоритет создания) с тем, чтобы определить, может ли новый LSP вытеснять существующий LSP. Для приоритетов предложен диапазон значений от 0 (высший приоритет) до 7 (низший приоритет);

– новые коды состояний LSR;

– LSPID – уникальный идентификатор тракта CR-LSP в сети;

– классы (цвета) сетевых ресурсов, назначаемые администратором сети.

В протоколе CR-LDP не появилось новых сообщений, но модифицированы некоторые сообщения базовой версии (например, *Label Request, Label Mapping*), а некоторые остались без изменений (*Label Release, Label Withdraw, Label Abort Request*).

Добавлен ряд новых TLV, передаваемых в модифицированных сообщениях или в качестве необязательных параметров в других сообщениях:

TLV явного маршрута (*Explicit route TLV*) – специфицирует трассу создаваемого тракта LSP, содержит одно или более полей *Explicit route hop TLV*;

TLV сетевых узлов явного маршрута (*Explicit route hop TLV*) – серия полей переменной длины, в каждом из которых указывается адрес маршрутизатора(ов);

TLV параметров трафика (*Traffic parameters TLV*) – описывает параметры трафика;

TLV вытеснения (*Preemption TLV*) – содержит данные о приоритетах создания и удержания;

LSPID TLV – содержит уникальный идентификатор LSP, состоящий из идентификатора входного LSR (или его IP-адреса) и локально уникального идентификатора LSP для этого LSR;

поле класса (цвета) ресурса (*Resource class (color) TLV*) – определяет, какие типы каналов передачи приемлемы для LSP. Выражается 32-битовым кодом;

TLV закрепления маршрута (*Route pinning TLV*) – указывает на наличие или отсутствие запроса закрепления маршрута.

8.1. Аспекты безопасности LDP

После краткого описания CR-LDP уделим внимание вопросам информационной безопасности, связанным с применением протокола LDP.

8.1.1. *Несанкционированные действия*

Имеется два типа взаимодействий по протоколу LDP, которые могут стать объектом атаки со стороны нарушителей.

Сообщения обнаружения, переносимые по транспортному протоколу UDP. LSR, связанные друг с другом на уровне звена данных, обмениваются

базовыми сообщениями Hello. Угроза несанкционированных сообщений Hello может быть снижена путем приема базовых сообщений Hello только через интерфейсы, прямо связанные с теми LSR, которым можно доверять, и/или за счет того, что будут игнорироваться базовые сообщения Hello, не адресованные «Ко всем маршрутизаторам» в группе многоадресной рассылки данной подсети. LSR, которые не связаны прямым звеном данных, могут извещать о своей готовности установить LDP-сеанс расширенными сообщениями Hello (Extended Hello). LSR может снизить риск появления несанкционированных расширенных сообщений Hello путем их фильтрации и приема только тех из них, которые исходят от источников, включенных в список разрешенных пользователей. Однако процедура фильтрации отнимает много ресурсов LSR.

Информация о сеансе, переносимая по транспортному протоколу TCP. Протокол LDP может использовать в качестве опции сигнатуру TCP Message Digest 5 (MD5). Это предотвращает введение имитированных сегментов TCP в LDP-сеансов. Использование сигнатуры TCP MD5 в протоколе LDP аналогично ее использованию в протоколе BGP (специфицировано в RFC 2385 «Protection of BGP Sessions via the TCP MD5 Signature Option»). Сам алгоритм MD5 представлен в RFC 1321 «The MD5 Message-Digest Algorithm». Однако LDP может использовать и любую другую методику, обеспечивающую защиту TCP-сообщений. Следовательно, когда будет специфицирована методика, более сильная, чем MD5, протокол LDP можно будет относительно просто модернизировать.

8.1.2. *Конфиденциальность*

Протокол LDP не предусматривает никакого специального механизма, обеспечивающего конфиденциальность распределения меток, но сам механизм, гарантирующий аутентичность и сохранность сообщений протокола LDP, дает уровень защиты не хуже того, который дают протоколы маршрутизации.

Известны высказывания, что для снижения угрозы несанкционированных действий распределение меток требует конфиденциальности. Однако такая конфиденциальность не защитит от имитации меток, поскольку пакеты данных переносят метки в явном, незашифрованном виде. Более того, такого рода атаки возможны и без знания злоумышленником конкретной привязки метки к FEC. Для предотвращения имитации меток необходима гарантия того, что пакеты получают метки от надежных и достоверных LSR, и что эти метки надлежащим образом контролируются маршрутизаторами, которые их назначают.

8.1.3. *Отказ в обслуживании*

LDP имеет две потенциальные мишени для атак, приводящих к отказам в обслуживании.

Известный (well-known) UDP-порт для процедур обнаружения. Администратор LSR может принимать меры для устранения угроз передачи несанкционированных базовых сообщений Hello, приводящих к отказам в обслуживании, гарантируя, что LSR непосредственно подключен только к надежным узлам, от которых нельзя ожидать инициирования подобных атак. Если существует надежная и достоверная область MPLS, то маршрутизаторы, расположенные на границе этой области, могут использоваться для защиты внутренних LSR от фатальных атак.

Известный TCP-порт для организации LDP-сеанса. Подобно другим управляющим протоколам, которые в качестве транспортного протокола используют TCP, протокол LDP может стать мишенью атак, приводящих к отказу в обслуживании, например SYN-атак. В этом отношении LDP уязвим в той же степени, что и другие управляющие протоколы, использующие протокол TCP. Применение на границе области MPLS списков доступа (списков разрешенных пользователей) способом, аналогичным предложенному выше для расширенных сообщений Hello, позволяет защитить внутреннюю область MPLS от атак из-за пределов области.

8.2. Сигнализация LDP

В завершение рассмотрим некоторые сценарии распределения меток с помощью протокола LDP.

Как уже отмечалось, маршрутизаторы LSR могут использовать либо независимый, либо упорядоченный способ распределения меток. При независимой раздаче меток LSR может объявлять метки одноранговым объектам LDP в любой момент, когда пожелает, т. е. может передать сообщение Label Mapping в вышестоящий LSR до получения сообщения Label Mapping от нижестоящего LSR. При упорядоченной раздаче LSR может объявлять метку вышестоящим LSR только тогда, когда он имеет привязку «FEC-метка» для следующего участка, или когда сам является выходным маршрутизатором MPLS.

Если мы комбинируем режим «нижестоящим-по-требованию» с упорядоченным способом раздачи меток, то их распределение будет происходить следующим образом. Входной LSR передает сообщение Label Request в нижестоящий LSR, и это сообщение будет продвигаться по сети до выходного LSR. Выходной LSR отвечает соседнему вышестоящему LSR сообщением Label Mapping. Этот вышестоящий LSR запоминает привязку «FEC-метка» для своего выходного интерфейса, создает новую привязку «FEC-метка» для своего входного интерфейса и информирует об этой привязке следующий вышестоящий LSR с помощью собственного сообщения Label Mapping. Процесс продолжается до тех пор, пока входной LSR не получит сообщение Label Mapping в ответ на свой первоначальный запрос.

LDP является протоколом прикладного уровня, использующим в качестве транспортных протоколов для связи с другими LSR, которые могут стать его одноранговыми узлами, как протокол UDP, так и протокол TCP. В частности, транспортный протокол UDP используется LSR для передачи сообщений обнаружения, информирующих другие LSR о том, что данный LSR является потенциальным кандидатом на роль однорангового узла LDP, а транспортный протокол TCP используется для передачи всех других сообщений, в частности сообщений *Session, Advertisement* и *Notification*, которым требуется механизм надежной и своевременной их доставки, что и обеспечивает TCP.

UDP-портом по умолчанию для сообщений *Hello* протокола LDP является порт 646. TCP-портом по умолчанию для всех других сообщений протокола LDP также является порт 646.

Прежде всего протокол LDP должен обнаружить маршрутизаторы LSR, с которыми возможна организация LDP-сеансов. Для этой цели в LDP предусмотрено два механизма: базовый и расширенный. Базовый механизм обеспечивает обнаружение LSR путем периодической передачи сообщений *Hello* к UDP-порту 646 по IP-адресу многоадресной рассылки «Все маршрутизаторы подсети» (224.0.0.2). Расширенный механизм используется для обнаружения LSR, не имеющих прямой связи (на уровне звена данных) с данным маршрутизатором. При этом пакеты *Hello* передаются по IP-адресу запрашиваемого LSR.

В сообщениях *Hello* передается LDP-идентификатор *пространства меток*, которое LSR, передающий эти сообщения, намерен использовать на следующем этапе процесса распределения меток – при создании соединения между маршрутизаторами по протоколу TCP, – а также вспомогательная информация. Сведения о смежных LSR, полученные посредством сообщений *Hello*, действительны в течение времени, заданного в этих сообщениях. По истечении заданного времени они могут быть аннулированы, о чем инициатор должен известить встречный LSR уведомляющим сообщением *Notification Message*, если только таймер не установлен на бесконечность, или если до срабатывания таймера не получено другое сообщение *Hello*.

После обмена сообщениями *Hello*, одноранговые узлы устанавливают TCP-соединение через порты 646. Сеанс TCP-связи инициируется тем из двух LSR, чей *транспортный адрес* (идентификатор пространства меток) короче. LSR, инициировавший TCP-соединение, является *ведущим (master)* маршрутизатором сеанса.

После того как TCP-соединение создано, оба LSR по инициативе ведущего маршрутизатора обмениваются сообщениями *Initialization*. Этими сообщениями они информируют друг друга о версии протокола LDP, о желательной дисциплине распределения меток, о диапазонах значений меток для ATM (VPI/VCI) и Frame Relay (DLCI), о значении таймера KeepAlive, а также о других параметрах *TRV*.

Этап инициирования сеанса связи завершается обменом между LSR сообщениями *KeepAlive*. Таким образом, поверх TCP-соединения организуется LDP-сеанс. При наличии неразрешенных конфликтов и проблем с совместимостью LSR передает уведомление об ошибке, и LDP-сеанс разрушается. В этом случае ведущий LSR может предпринять, но не ранее чем через 15 с, повторную попытку установить связь, причем при нескольких отказах подряд выдержка времени каждый раз увеличивается, достигая максимально 2 мин.

Установив LDP-сеанс, каждый LSR ведет мониторинг прихода сообщений LDP в рамках этого сеанса. Если в течение времени, установленного таймером KeepAlive, от смежного узла не получено никакой информации, LSR закрывает LDP-сессию, разрушая транспортное соединение и оповещая об этом встречный маршрутизатор сообщением *Shutdown*. Каждый раз, когда от корреспондента принимается сообщение, таймер KeepAlive сбрасывается. Таким образом, оба участника соединения должны обеспечивать передачу любого протокольного сообщения внутри интервала времени, заданного таймером. Если маршрутизатор не имеет полезной информации, он должен передавать сообщение *KeepAlive*.

Итак, LDP-сеанс организован, и теперь LSR может запросить метку, передав сообщение *Label Request*. Это сообщение предусматривает ограничение максимального числа пересылок для предотвращения бесконечной циркуляции запроса по сети при возникновении закольцованных маршрутов. Обязательным параметром сообщения *Label Request* является параметр FEC TLV, указывающий, для какого FEC запрашивается метка. В состав сообщения *Label Request* может также входить рассмотренный выше вектор тракта (Path Vector), представляющий собой перечень всех маршрутизаторов, через которые прошло это сообщение.

Если в это время какой-либо узел столкнется с проблемами, связанными с совместимостью, с отсутствием ресурсов и т. д., вышестоящему маршрутизатору, запрашивающему метку, будет передано соответствующее уведомление, и оно будет пересылаться дальше по цепочке, пока не дойдет до входного пограничного маршрутизатора. Все соединения между маршрутизаторами этого, еще не созданного тракта будут разрушены. Если сообщение *Label Request* в конце концов благополучно дойдет до выходного маршрутизатора (при упорядоченном режиме), в выходном маршрутизаторе будет генерироваться сообщение *Label Mapping*, содержащее метку, которая имеет локальное значение на участке между выходным и соседним с ним вышестоящим маршрутизатором. Если на всех следующих далее вышестоящих маршрутизаторах успешно пройдет привязка меток к FEC, то после обработки во входном пограничном маршрутизаторе сообщения *Label Mapping*, полученного от соседнего с ним нижестоящего маршрутизатора, тракт LSP будет создан.

9. ТЕХНОЛОГИЯ MPLS VPN. ЧАСТНЫЕ СЕТИ И ТРАНСПОРТНЫЕ СЕТИ

Общие предпосылки VPN традиционно иллюстрируются следующими типовыми примерами. Представим гипотетическую сеть предприятия современной глобализованной экономики. Офисы разбросаны по всему миру, сотни или тысячи сотрудников работают в командировках, сотни или тысячи сотрудников работают дома – и все это необходимо объединить в единую сеть. Причем не просто объединить, а организовать доступ к информационным ресурсам, разделить полномочия, обеспечить надежность и безопасность. Естественно, можно проложить каналы, установить маршрутизаторы и устройства доступа, т. е. организовать свою собственную частную сеть связи. Предприятие, имеющее такую сеть, не зависит от операторов сети общего пользования, решает проблемы безопасности, доступа к услугам, может строить все что угодно и на чем угодно.

Сеть MPLS VPN делится на две области: IP-сети клиентов и магистраль провайдера. Классическая конструкция MPLS L3VPN состоит из следующих компонентов: граничные маршрутизаторы провайдера LER, обращенные к клиентскому оборудованию CE, соединены между собой маршрутизаторами LSR в домене MPLS.

Маршрутизаторы, расположенные на входе или выходе MPLS-сети называются LER (англ. Label Edge Router – граничный маршрутизатор меток). LER на входе в MPLS-сеть добавляют метку MPLS к пакету данных, а LER на выходе из MPLS-сети удаляет метку MPLS из пакета данных.

Маршрутизаторы, выполняющие маршрутизацию пакетов данных, основываясь только на значении метки, называются LSR (англ. Label Switching Router – коммутирующий метки машрутизатор).

LDP (англ. Label Distribution Protocol) – протокол, который используется LSR-маршрутизаторами для обмена информацией о метках. Предоставляет возможность маршрутизаторам обнаруживать друг друга и устанавливать взаимодействие.

LSP (англ. Label Switch Path) – последовательность LSR, которые коммутируют помеченный пакет через сеть MPLS. Фактически это маршрут следования пакетов через сеть MPLS.

Инфраструктура MPLS L3VPN предполагает обеспечение изоляции распределенных клиентских IP-сетей в рамках VPN, т. е. обеспечивается только обмен пакетами между IP-сетями одной VPN.

Каждая VPN логически связана с одним или более комплексов маршрутизации и пересылки (VRF). VRF определяет членство в VPN подсети за узлом CE, подключенного к PE. Интерфейсы PE маршрутизаторов, обращенные к CE, логически связаны с индивидуальными VRF.

VPN (англ. Virtual Private Network – виртуальная частная сеть) – обобщенное название технологий, позволяющих обеспечить одно или несколько сетевых соединений (логическую сеть) поверх другой сети. В зависимости от применяемых протоколов и назначения VPN может обеспечивать соединения трех видов: узел-узел, узел-сеть и сеть-сеть.

VRF – это объект, в состав которого входят:

- множество интерфейсов PE, к которым подключаются CE из одного VPN;
- VRF-таблица.

Атрибуты и правила распространения маршрутной информации.

VRF локален для каждого устройства PE. Понятие VPN распространяется на всю сеть. Таким образом, VPN не равно VRF. VRF – это, скорее, описание VPN в рамках одного устройства PE.

Определяются два ключевых параметра в VRF: RD (англ. Route Distinguisher) и RT (англ. Route Target).

RD – уникальный для каждого VRF идентификатор (он может быть только один), который превращает маршрут IPv4 в уникальный для всей сети провайдера маршрут VPNv4. Фактически именно он и позволяет использовать пересекающуюся адресацию у клиентов (да и сети оператора тоже).

RT – параметр, который добавляется в специальное поле при рассылке маршрутов внутри сети оператора по MP-BGP и который определяет, в таблицу маршрутизации какого или каких VRF этот маршрут попадет.

MP-BGP (англ. MultiProtocol BGP) – это расширение для BGP, которое добавляет возможность работы с маршрутами VPNv4 и необходимо для создания MPLS VPN.

10. ПРЕИМУЩЕСТВА ОРГАНИЗАЦИИ VPN НА БАЗЕ MPLS

Основными преимуществами организации VPN на базе MPLS можно назвать: масштабируемость, пересечение адресных пространств, узлов, подключенных в различные VPN, изолирование трафика VPN друг от друга на втором уровне модели OSI.

Масштабируемость достигается за счет того, что подключение нового узла в существующий VPN производится только перенастройкой одного PE, к которому подключается данный узел.

В различных VPN адресные пространства могут пересекаться, что может быть чрезвычайно полезным в случае, если оператору необходимо предоставить VPN нескольким клиентам, использующим одинаковое приватное адресное пространство.

Устройства P(LSR) при коммутации анализируют только внешнюю метку, определяющую LSP между PE, и не анализируют заголовок IP-пакета, тогда справедливо говорить о том, что эти устройства выполняют функции коммутации на втором уровне модели OSI. Устройства PE также разделяют маршрутную информацию, таблицы маршрутизации, интерфейсы, направленные в сторону устройств CE, между VRF. Тем самым процессы маршрутизации разных VPN полностью разделяются, и обеспечивается разделение трафика от разных VPN на втором уровне модели OSI.

Все сети VPN, рассматриваемые в начале раздела, обладают общими функциональными возможностями: масштабируемостью, наличием средств управления, защитой и обработкой частных адресов

Масштабируемость – крайне полезное для VPN свойство, поскольку эти сети часто нуждаются в расширении с ростом бизнеса предприятия. Сети VPN должны быть *управляемыми*, чтобы их можно было конфигурировать и контролировать в соответствии с быстро меняющимися бизнес-процессами предприятия. Важным может оказаться также управление учетом предоставляемого обслуживания для целей начисления оплаты и других аспектов. *Защита* является свойством, которое сохраняет букву P в аббревиатуре VPN и становится ключевым элементом в сетевых коммуникациях в современных условиях угроз похищения информации и атак на интеллектуальную собственность. Поскольку большинство VPN в настоящее время ориентируются на IP, возможность *обработки частных адресов*, чтобы гарантировать уникальность IP-адреса, также важна.

Рассмотренное в предыдущем разделе понятие LSP-туннелей составляет важный аспект MPLS VPN, так как прилагательное «частный» в VPN на базе MPLS относится к физическому разделению трафика между LSP-туннелями. Это весьма похоже на способ виртуальных каналов, которые

устанавливаются для VPN-приложений ATM и FR. В настоящее время в области сетей MPLS VPN существуют два главных направления: *BGP/ MPLS VPN* и *VPN с виртуальными маршрутизаторами на базе IP.*

Ядро сети строится на базовых маршрутизаторах MPLS, называемых *внутренними маршрутизаторами провайдера P* и взаимодействует с пользователем VPN не напрямую, а посредством соединения между *граничным устройством маршрутизации заказчика CE (Customer Edge router)* и *граничным устройством маршрутизации провайдера PE (Provider Edge router)*. CE могут быть статически подсоединены к PE провайдера через закрепленные каналы или могут использовать коммутируемые линии связи. При этом для подключения CE к PE могут задействоваться каналы любого типа, например, ATM, FR, Ethernet, PPP, а также такие механизмы туннелирования, как упоминавшиеся IPSec, L2TP или *GRE* (*Generic Route Encapsulation*), рис. 10.1.

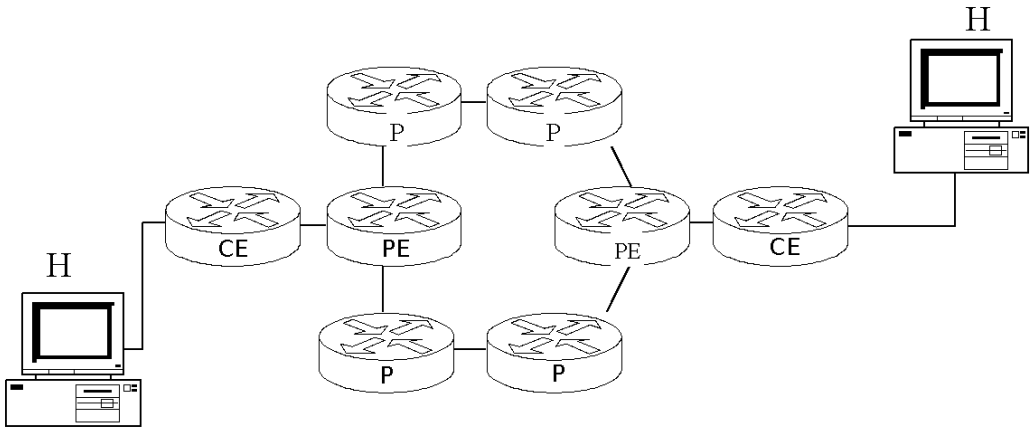

Рис. 10.1. Эталонная модель MPLS VPN

Оба метода MPLS VPN сходны в функциональности создаваемой провайдером услуги VPN. В одном методе протокол BGP используется для создания специальных расширенных адресов при передаче пакетов через ядро MPLS, а в другом – VR хранят отдельные таблицы путей MPLS для каждой VPN. Фактическая же реализация этих двух методов совершенно различна, а выбор метода – решение провайдера услуг VPN на основе возможностей оборудования, ситуации с взаимодействием сетей и других факторов. Создана новая рабочая группа IETF, названная *Provider-Provisioned VPNs (PPVPNs)*, которая разрабатывает структуру и соответствующие спецификации для этих двух типов сетей VPN.

11. СЕТИ MPLS/BGP VPN

Модель MPLS/BGP VPN (рис. 11.1) базируется на расширениях протокола маршрутизации внешнего шлюза BGP, называемых многопротокольными расширениями BGP и касающихся специальных расширенных адресов. Эти адреса используются для обмена информацией о доступности между маршрутизаторами PE только между членами одной и той же VPN. Каждый маршрутизатор PE в MPLS/BGP VPN поддерживает отдельную *таблицу маршрутизации VRF (VPN Routing and Forwarding table)*. Каждая VRF содержит все маршруты для одного пользователя одной VPN, что обеспечивает эффективную изоляцию сетей. Данная модель позволяет использовать перекрытие частных IP-адресаций разными предприятиями.

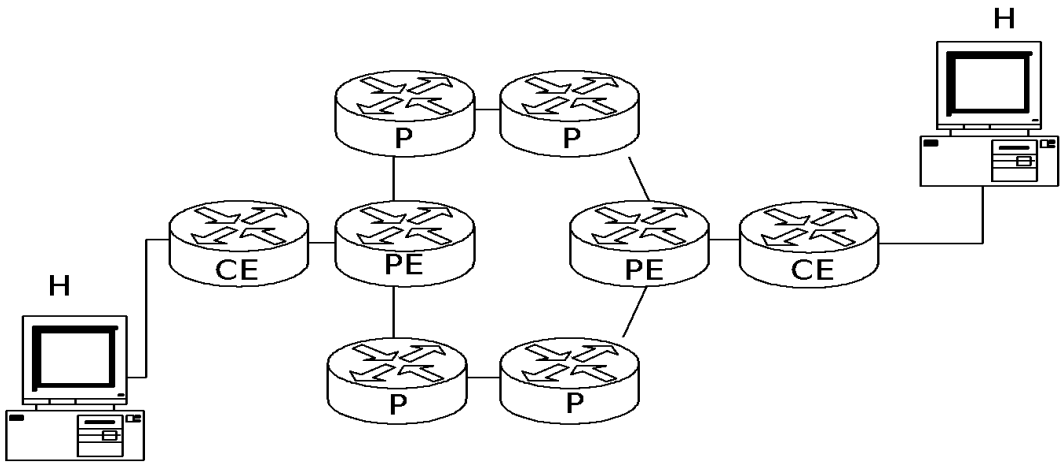

Рис. 11.1. Модель сети MPLS/BGP VPN

Основная цель этого типа реализации MPLS VPN – позволить провайдеру обеспечить требуемую заказчику конфигурацию VPN. Заказчиком в данном случае может быть предприятие, группа предприятий, которым нужна сеть Экстранет, другой сервис-провайдер или даже другой провайдер VPN, который может использовать это MPLS-приложение с целью построить сеть VPN своим собственным клиентам. Модель VPN позволит заказчикам использовать сетевые услуги масштабируемо и гибко, а провайдеру – упростить менеджмент, защиту и обслуживание VPN. Спецификация MPLS/BGP VPN впервые была опубликована как RFC 2547 «BGP/ MPLS VPNs», а позже пересмотрена и усовершенствована рабочей группой.

12. ВИРТУАЛЬНАЯ СЕТЬ НА БАЗЕ MPLS IP (MPLS VPN)

Другая популярная модель MPLS VPN связана с использованием маршрутизаторов VR (рис. 12.1). Маршрутизатор VR является результатом логического разделения физического маршрутизатора на несколько логических (виртуальных) маршрутизаторов. Каждый VR поддерживает свою собственную независимую таблицу для каждой VPN. Ясно, что при таком решении для информации о доступности маршрутизатора BGP не требуется.

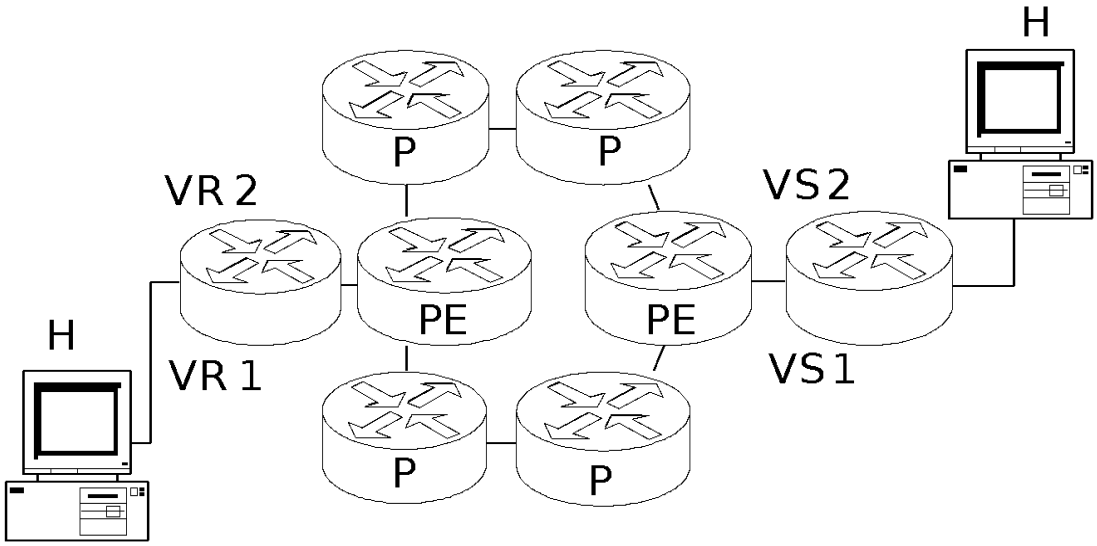

Рис. 12.1. Модель VPN на базе MPLS IP

Концепция VR подразумевает разделение физического маршрутизатора на несколько маршрутизаторов VR. Каждый VR использует существующие механизмы и инструменты для конфигурирования, обеспечения, эксплуатации, начисления платы и техобслуживания. Маршрутизаторы VR могут совместно использовать физические ресурсы маршрутизации или использовать отдельные объекты. В этой модели каждый VR имеет свой собственный объект маршрутизации для распределения информации о доступности VPN среди маршрутизаторов VR, участвующих в данной VPN. При этом VPN может использовать любой протокол маршрутизации, а для получения необходимой доступности VPN не требуется никакого специального расширения к протоколу маршрутизации. Такие конфигурации VPN могут быть очень гибкими в отношении того, как маршрутизаторы VR соединены в опорной сети, эта архитектура может допускать различные сценарии развертывания опорной сети: провайдер VPN MPLS может быть владельцем опорной сети, или же провайдер услуги VPN может получать необходимые услуги опорной сети от одного или нескольких других провайдеров MPLS.

Организация MPLS VPN

В общем случае у клиента может быть несколько территориально обособленных IP-сетей, каждая из которых, в свою очередь, может включать несколько подсетей, связанных маршрутизаторами. Такие территориально изолированные сетевые элементы корпоративной сети принято называть *сайтами*. Принадлежащие одному клиенту сайты обмениваются IP-пакетами через сеть провайдера и образуют виртуальную частную сеть этого клиента. Каждый сайт имеет один или несколько граничных пользовательских маршрутизаторов CE, соединенных с одним или более граничными провайдерскими маршрутизаторами PE посредством каналов PPP, ATM, Ethernet, Frame Relay и т. п.

CE-маршрутизаторы различных сайтов не обмениваются маршрутной информацией непосредственно и даже могут не знать друг о друге. Адресные пространства подсетей, входящих в состав VPN, могут перекрываться, т. е. уникальность адресов должна соблюдаться только в пределах конкретной подсети. Этого удается добиться преобразованием IP-адреса в VPN-IP-адрес и использованием протокола MP-BGP для работы с этими адресами.

Каждый PE-маршрутизатор должен поддерживать столько таблиц маршрутизации, сколько сайтов пользователей к нему подсоединено, т. е. на одном физическом маршрутизаторе организуется несколько виртуальных. При этом маршрутная информация, касающаяся конкретной VPN, содержится только в PE-маршрутизаторах, к которым подсоединены сайты данной VPN. Таким образом, решается проблема масштабирования, неизбежно возникающая в случае наличия этой информации во всех маршрутизаторах сети оператора. Считается, что CE-маршрутизатор относится к одному сайту, но сам сайт может принадлежать к нескольким VPN. К PE-маршрутизатору может быть подключено несколько CE-маршрутизаторов, находящихся в разных сайтах и даже относящихся к разным VPN.

13. МАРШРУТИЗАЦИЯ MPLS VPN

13.1. Таблицы маршрутизации в PE-маршрутизаторах

Как уже говорилось, PE-маршрутизаторы поддерживают на каждый подключенный сайт по одной, ассоциированной с ним таблице маршрутизации. Рассмотрим сеть, в которой существуют три PE-маршрутизатора: PE1, PE2 и PE3. Каждый из них соединяется с одним из CE-маршрутизаторов: CE1, CE2 и CE3 соответственно. При этом CE1, CE2 и CE3 принадлежат сайтам, входящим в одну VPN. В таком случае PE1 использует протокол MP-BGP, чтобы передать PE2 и PE3 сведения о маршрутах, которые он получил от CE1. В свою очередь, PE2 и PE3 вносят эти маршруты в свои таблицы маршрутизации, ассоциированные с сайтами, в которых находятся CE2 и CE3. Если сайт принадлежит к нескольким VPN, то соответствующая ему таблица маршрутизации в PE может содержать маршруты всех этих VPN.

Если PE-маршрутизатор получил от сайта пакет с адресом, которого нет в ассоциированной таблице маршрутизации, то либо он отбросит данный пакет, либо, если оператор предоставляет услуги доступа в Internet через эту VPN, произойдет обращение к таблице маршрутизации Internet.

Для обеспечения требуемой изоляции одной VPN от другой важно, чтобы ни один из маршрутизаторов, составляющих магистральную сеть MPLS (Backbone Router), не принимал пакеты с метками от маршрутизаторов, не относящихся к данной сети, за исключением случая, когда верхняя метка стека была уже распространена маршрутизатором, входящим в магистральную сеть MPLS. При этом обнаруживается, что использование этой метки приведет к тому, что пакет покинет сеть до того, как будут обработаны остальные метки стека и проанализирован IP-заголовок.

Ассоциированные таблицы маршрутизации в PE используются только для пакетов, полученных от непосредственно соединенных с PE сайтов, и, как результат, может существовать несколько различных маршрутов к одной системе, причем маршрут будет определяться сайтом, из которого пакет попал в магистральную сеть.

В некоторых случаях сайт клиента может быть разделен им на несколько виртуальных сайтов, например с использованием VLAN. Такие виртуальные сайты могут быть членами различных VPN, а PE должен поддерживать различные таблицы маршрутизации для каждого виртуального сайта, как это описано в RFC 2917.

13.2. Маршрутная информация по BGP

PE-маршрутизаторы используют BGP для передачи друг другу маршрутной информации. BGP-спикер может определить и распространить только один маршрут для каждого конкретного адресного префикса, но

любая VPN может иметь свое собственное адресное пространство, и один и тот же адрес может повторно использоваться в разных VPN. Для устранения этого противоречия было специфицировано новое *семейство адресов VPN-IPv4 Address Family*.

Адрес VPN-IPv4 имеет длину 12 байтов, первые восемь из которых занимает префикс, называемый *различителем маршрутов RD (Route Distinguisher)*, а оставшиеся 4 байта содержат IPv4-адрес. Даже если в двух VPN будут совпадающие IPv4-адреса, то PE-маршрутизатор преобразует их в уникальные VPN-IPv4-адреса. Таким образом, решена задача установления различных маршрутов к устройствам, имеющим один и тот же IP-адрес, но принадлежащим различным VPN.

Сам RD не является информативным и служит лишь для целей создания нескольких отдельных маршрутов по общему IPv4-адресному префиксу. Также RD может быть использован для создания нескольких различных маршрутов к одному и тому же сетевому устройству путем создания двух отличных VPN-IPv4-адресов, имеющих общую IPv4-часть.

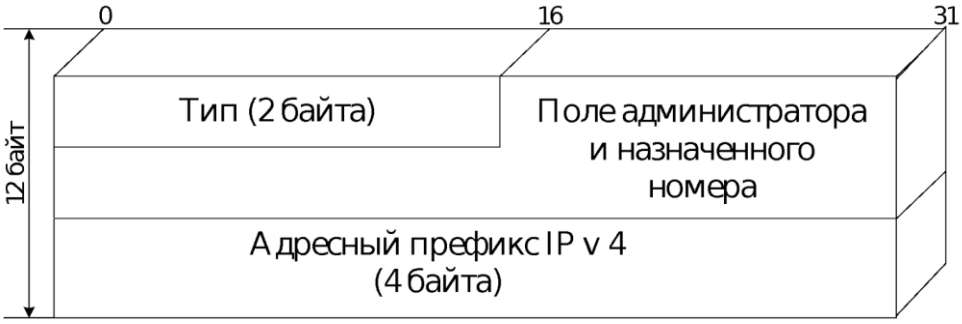

Рис. 13.1. Кодирование Route Distinguisher

Различители маршрутов RD имеют структуру, позволяющую каждому провайдеру администрировать свое собственное «номерное пространство», не конфликтуя при этом с RD, назначенными другими провайдерами. RD состоит из 2-байтового поля типа, поля администратора (Administrator) и поля назначенного номера (Assigned Number). Причем значение поля типа определяет длину обоих последующих полей, а также семантику поля администратора. Поле администратора определяет авторство номера, а поле назначенного номера содержит соответственно его значение. Данная структура различителя маршрутов полностью игнорируется протоколом BGP, когда он сравнивает два таких адресных префикса. Она имеет смысл только для провайдера услуг. Если же поля администратора и назначенного номера заполнены нулями, то адрес имеет то же значение, что и обыкновенный IPv4-адрес, и они будут сравнимы для BGP.

Конкретная ассоциированная таблица маршрутизации будет иметь только один VPN-IPv4-маршрут для любого заданного IPv4-префикса.

Когда адрес назначения ставится в соответствие с VPN-IPv4-маршрутом, то соотносится только IPv4-часть. PE-маршрутизатор должен ассоциировать маршруты, ведущие к конкретному CE с определенным RD. При этом PE может быть сконфигурирован так, чтобы ассоциировать все маршруты, ведущие к одному CE, с одним RD или с разными RD.

13.3. Распространение маршрутной информации

13.3.1. *Атрибут целевой VPN*

Каждой ассоциированной таблице маршрутизации в PE-маршрутизаторах присваивается один или несколько *атрибутов целевой VPN (Target VPN)*. Когда PE-маршрутизатором создается маршрут VPN-IPv4, он ассоциируется с одним или более атрибутами Target VPN. Они переносятся протоколом BGP как атрибуты маршрута. Каждый маршрут, ассоциированный с Target VPN «T», должен быть объявлен всем PE-маршрутизаторам, имеющим таблицу маршрутизации, ассоциированную с Target VPN «T». Когда такой маршрут принимается PE-маршрутизатором, он имеет право быть включенным в любую из ассоциированных таблиц маршрутизации, имеющих атрибут Target VPN «T».

Существует набор атрибутов Target VPN, которыми PE-маршрутизатор снабжает маршруты, полученные от сайта «S». Одновременно с этим есть набор атрибутов Target VPN, который PE-маршрутизатор использует для принятия решения о праве маршрута, полученного от другого PE, быть включенным в ассоциированную с сайтом «S» таблицу маршрутизации. Эти два набора атрибутов независимы и могут быть различными.

Функции, выполняемые атрибутом Target VPN, схожи с функциями BGP Communities Attribute, однако формат последнего не удовлетворяет требованиям, так как имеет только 2-байтовое номерное пространство.

Когда BGP-спикер получает 2 маршрута для одного VPN-IPv4-адресного префикса, он выбирает один согласно традиционным правилам предпочтения маршрутов в протоколе BGP.

Следует отметить, что маршрут может иметь только один RD, но несколько атрибутов Target VPN. Если существует один маршрут с несколькими атрибутами, в BGP масштабируемость улучшается по сравнению с вариантом, когда создается несколько отдельных маршрутов. Всегда остается возможность убрать атрибут Target VPN, создав большее количество маршрутов, но ухудшив характеристику масштабируемости.

Существует несколько возможных подходов, реализуемых PE, для определения атрибутов Target VPN, которые необходимо ассоциировать с данным маршрутом. PE может быть сконфигурирован так, чтобы ассоциировать все маршруты, ведущие к конкретному сайту с определенным

атрибутом Target VPN. Или ассоциировать часть маршрутов, ведущих к данному сайту, с одним Target VPN, а другую часть – с отличным Target VPN. Также возможен вариант, когда CE-маршрутизатор, отправляя маршрутную информацию к PE-маршрутизатору, задает для каждого маршрута один или несколько атрибутов Target VPN. Последний вариант передает пользователю контролирующие механизмы VPN. При этом желательно оставить за PE возможность исключать любые Target VPN, запрещенные согласно конфигурации PE, а также добавлять определенные Target VPN, являющиеся для какого-либо PE обязательными.

13.3.2. *Атрибут VPN-источник*

Маршрут VPN-IPv4 может быть опционально ассоциирован с атрибутом *VPN-источника (VPN of Origin)*. Этот атрибут однозначно идентифицирует группу сайтов и соответствующий маршрут, объявленный одним из маршрутизаторов, находящихся в этих сайтах. В качестве одного из применений данного атрибута может быть идентификация предприятия, владеющего сайтом, к которому ведет маршрут, или сети Intranet, к которой он принадлежит. Точнее было бы назвать этот атрибут источником маршрута, так как определяемая им группа сайтов не обязательно составляет VPN.

13.3.3. *Атрибут сайт-источник*

Этот атрибут *Site of Origin* уникальным образом идентифицирует сайт, от которого PE-маршрутизатор получил информацию о данном маршруте. Все маршруты, полученные от конкретного сайта, должны быть ассоциированы с конкретным значением этого атрибута, даже если сайт имеет несколько соединений с одним PE или соединен с несколькими PE. Для разных сайтов должны использоваться разные атрибуты Site of Origin.

14. ПЕРЕДАЧА МАРШРУТНОЙ ИНФОРМАЦИИ МЕЖДУ PE

Если два сайта принадлежат к одной автономной системе AS, то их PE-маршрутизаторы могут обмениваться маршрутной информацией по IBGP непосредственно друг с другом или, когда имеется большое их количество, через *отражатель маршрутов*. Отражатели маршрутов позволяют сообщать полученные от маршрутизаторов IBGP маршруты другим маршрутизаторам IBGP, т. е. «отражать» их, уменьшая тем самым количество связей внутри AS.

Если же сайты находятся в разных AS (например, если они подсоединены к разным провайдерам), то PE-маршрутизатор должен будет передать маршрутную информацию по IBGP граничному маршрутизатору ASBR или отражателю маршрутов, клиентом которого является ASBR. Затем ASBR должен будет по EBGP передать маршруты ASBR, находящемуся в другой AS. Таким образом, становится возможным подключение сайтов к различным провайдерам, но между последними должно быть предварительно заключено доверительное соглашение, так как маршруты из глобальной сети Интернет по умолчанию не принимаются.

Когда PE-маршрутизатор распространяет маршруты по BGP, он использует свой собственный адрес в качестве параметра BGP_Next_Hop. Также он назначает и распространяет метку MPLS: фактически PE распространяет не VPN-IPv4-маршруты, а «помеченные» VPN-IPv4-маршруты, но на текущий момент документы IETF, специфицирующие перенос меток MPLS протоколом BGP-4, находятся в статусе *draft*. Когда PE обрабатывает полученный пакет, имеющий такую метку наверху стека, он извлекает стек и отправляет пакет сайту, к которому ведет маршрут. Обычно это означает, что он пошлет пакет CE-маршрутизатору, объявившему ему данный маршрут. Метка также может определять инкапсуляцию трафика звена данных.

В большинстве случаев метка, присвоенная PE-отправителем, обеспечит доставку пакета прямо к CE, и PE-получателю не придется анализировать адрес назначения. Однако также возможно присвоение метки, однозначно идентифицирующей конкретную таблицу маршрутизации, к которой PE-получателю необходимо обратиться для дальнейшей пересылки пакета. За счет того, что при построении MPLS-VPN допускается использование любых стандартных архитектур на основе отражателей маршрутов, а также за счет того, что PE-маршрутизаторы принимают только такие объявления о маршрутах, которые относятся к подключенным к ним VPN, а P-маршрутизаторы, т. е. маршрутизаторы MPLS-ядра, вообще не принимают сообщений о VPN-IPv4-маршрутах, достигается исключительная масштабируемость MPLS-VPN-решений.

14.1. Пересылка данных по магистральной сети

Маршрутизаторы магистральной сети не имеют сведений о маршрутах к сайтам VPN, а пересылка пакетов осуществляется посредством технологии MPLS с двухуровневым стеком меток. PE-маршрутизаторы (а также ASBR, распространяющие адреса VPN-IPv4) должны добавлять свои адресные префиксы в маршрутные таблицы IGP магистральной сети. Это позволяет MPLS на каждом узле магистральной сети назначить метку соответствующему маршруту к каждому из PE.

Получив пакет от одного из CE, PE-маршрутизатор обращается к некоторой ассоциированной таблице маршрутизации. Если пакет предназначается CE-узлу, подсоединенному непосредственно к рассматриваемому PE, он немедленно отсылается этому маршрутизатору. В противном случае анализируется поле BGP_Next_Hop, а также находится соответствующая ему метка и помещается наверх стека меток. Затем PE ищет IGP-маршрут к BGP_Next_Hop и метку, соответствующую следующей пересылке IGP. Эта метка также добавляется наверх стека, и пакет отсылается на адрес следующей пересылки IPG (в случае если адреса следующего узла для IGP- и BGP-маршрутов совпадают, может назначаться только первая метка). После пересылки по MPLS-сети PE-маршрутизатор удаляет метку и отсылает пакет CE-получателю, анализирующему только заголовок IP-пакета.

Следует обратить внимание, что в силу двухуровневой модели стека меток маршрутизаторы магистральной сети могут даже не иметь маршрутов до CE, а только до PE-маршрутизаторов.

14.2. Передача маршрутной информации между CE и PE

PE-маршрутизатор, подсоединенный к конкретной VPN, должен знать, как адреса этой VPN распределены по ее сайтам. Если CE-устройство является рабочей станцией (*host*) или коммутатором (*switch*), этот набор адресов будет конфигурироваться на обслуживающем его PE-маршрутизаторе. Если же CE-устройство – маршрутизатор, то существует несколько возможных способов получения PE-маршрутизатором данного набора адресов.

PE преобразует эти адреса в VPN-IPv4-адреса, используя различитель маршрутов *RD*. Далее PE использует эти VPN-IPv4-адреса в качестве информации на входе BGP.

Возможность применения той или иной техники обмена маршрутной информацией между PE и CE зависит от того, принадлежит ли CE к так называемой «транзитной VPN» или нет. *Транзитная VPN* – это VPN, содержащая маршрутизатор, принимающий маршрутную информацию от маршрутизаторов (не принадлежащих данной VPN, но также не являющихся PE-маршрутизаторами) и передающий эту информацию PE-маршрутиза-

тору. В противном случае VPN является *тупиковой VPN (stub VPN)*, называемой иногда *шлейфной VPN*. Большая часть VPN, включая практически все корпоративные сети предприятий, попадает в разряд stub VPN.

А теперь рассмотрим непосредственно способы обмена маршрутной информацией. Статическая маршрутизация, т. е. назначаемая администратором сети, может использоваться только в *stub VPN*.

Маршрутизаторы PE и CE могут быть одноранговыми узлами протокола RIP. Тогда CE сможет передавать PE-информацию о доступных адресных префиксах, используя RIP.

Точно так же PE- и CE-маршрутизаторы могут быть одноранговыми узлами протокола OSPF. В таком случае необходимо, чтобы сайт являлся отдельной зоной OSPF, CE был ABR в этой зоне, а PE – ABR в другой зоне. Этот способ может использоваться только в *stub VPN*.

PE- и CE-маршрутизаторы могут быть одноранговыми узлами протокола BGP, и CE может использовать BGP (в частности, EBGP), чтобы объявлять PE о доступных адресных префиксах. Данная техника пригодна как для транзитных, так и для *stub* VPN и с чисто технической точки зрения является наилучшим решением, так как в отличие от IGP-техники, она не требует наличия в PE модулей поддержки нескольких протоколов маршрутизации для общения с разными CE. Это и понятно, так как BGP изначально создан для обеспечения обмена маршрутной информацией между независимо администрируемыми системами. Использование BGP также позволяет CE легко передавать PE атрибуты маршрутов.

С другой стороны, работа с протоколом BGP может не поддерживаться администратором CE, за исключением, конечно, случаев, когда пользователь услуг VPN сам является интернет-провайдером.

Если сайт находится не в транзитной VPN, то он может не иметь уникального *номера автономной системы ASN*, и этот номер может выбираться из частного плана нумерации ASN. Петли маршрутов при этом предотвращаются с помощью атрибута *Site of Origin*, описанного выше. Если группа сайтов образует транзитную VPN, то удобно представить ее как *BGP-конфедерацию* для того, чтобы внутренняя структура VPN была скрыта от любого маршрутизатора, не входящего в состав VPN. В этом случае каждый сайт в VPN должен будет иметь два BGP-соединения к магистральной сети, одно является внутренним, а второе – внешним относительно BGP-конфедерации. В стандартные внутриконфедерационные процедуры следует внести некоторые изменения для того, чтобы учесть возможность существования в магистральной сети и в сайтах различных сетевых политик.

Прежде чем распространять VPN-IPv4-маршруты, полученные от сайта, PE должен присвоить маршрутам определенные *атрибуты Site of Origin, VPN of Origin и Target VPN*.

Теперь обсудим распространение маршрутной информации от PE к CE. Разумеется, CE-устройство в этом случае должно быть маршрутизатором. В общем случае PE может объявлять CE любой маршрут, находящийся в его таблице маршрутизации, ассоциированной с сайтом, содержащим рассматриваемый CE. При этом существует одно ограничение: маршрут, имеющий атрибут *Site of Origin*, никогда не должен объявляться CE-маршрутизаторам, принадлежащим сайту, идентифицируемому этим атрибутом. Для передачи маршрутной информации от PE к CE будут использоваться те же процедуры, что и для передачи от CE к PE.

В большинстве случаев будет, однако, удобно назначить PE маршрут по умолчанию к CE, а в некоторых случаях также и от CE к PE.

14.3. Поддержка CE-маршрутизатором MPLS

Если CE поддерживает MPLS и хочет импортировать себе все маршруты, относящиеся к его VPN, PE может распространить ему метки для каждого из этих маршрутов. Когда PE будет принимать от CE пакеты с подобными метками, он должен заменять их соответствующими метками, полученными по BGP, и добавлять наверх стека метку для следующей пересылки BGP (*BGP Next Hop*).

15. ИНЖИНИРИНГ ТРАФИКА В ТРАНСПОРТНЫХ СЕТЯХ. MPLS-TE

Задача маппинга (*англ. Mapping* – наложение, проецирование) потоков трафика данных на существующие сетевые топологии называется инжинирингом трафика (TE). TE-механизмы обеспечивают возможности по переносу потоков трафика с маршрутов, рассчитанных IGP-протоколами на маршруты, которые обеспечивают большую пропускную способность и являются менее загруженными.

Инжиниринг трафика представляет собой мониторинг и моделирование трафика, а также управление трафиком с тем, чтобы обеспечить нужное качество его обслуживания путем рационального использования сетевых ресурсов за счет сбалансированной их загрузки. В отечественной литературе совокупность механизмов, обеспечивающих выполнение перечисленных функций, называется также *перераспределением* трафика, *конструированием* трафика, *оптимизацией* трафика, *проектированием* трафика, *управлением* трафиком. Более точным является название *управление разнотипным трафиком*, так как оно подчеркивает связь рассматриваемых здесь механизмов с задачей обеспечить разное качество обслуживания (QoS) трафика разных типов, но в данном учебном пособии будем использовать наиболее распространенную прямую кальку с английского эквивалента – *Traffic Engineering (TE)*.

Хотя методы TE появились раньше технологии MPLS, именно эта технология очень хорошо подходит для управления трафиком, может предоставить большую часть функций TE по относительно низкой (в сравнении с конкурирующими решениями) цене. Не менее важно, что в MPLS можно автоматизировать функции управления трафиком. Для всего этого нужно только одно – протоколы сигнализации MPLS должны уметь переносить информацию, которая необходима для работы механизмов управления трафиком, находящихся на прикладном уровне, а также «прокладывать» тракты LSP по явно заданным маршрутам. При этом в MPLS есть возможность достичь дополнительной гибкости, так как маршрут можно задать не только строго, но и не строго, т. е. группа узлов может быть задана как «абстрактный узел», внутри которого существует известная свобода выбора маршрута.

Инжиниринг трафика в MPLS основан на управлении наборами атрибутов, значения которых учитываются при выборе маршрутов для создаваемых в MPLS-сети LSP и LSP-туннелей.

Основными компонентами подсистемы TE являются:

- *пользовательский интерфейс,* через который администратор сети может управлять политикой TE;

- *IGP-компонент,* распространяющий информацию о топологии сети и сведения о состоянии сетевых ресурсов;
- *маршрутизация на основе ограничений* – модуль, который производит расчет маршрута в сети MPLS на основе информации, получаемой от пользовательского интерфейса и IGP-компонента;
- *компонент сигнализации* для создания и поддержки LSP (или LSP-туннеля), для управления LSP (LSP-туннелем) и для резервирования сетевых ресурсов;
- *компонент пересылки данных,* в качестве которого выступает сама сеть MPLS.

Напомним, что TE – это механизм оптимизации сети, и потому сетевые узлы наряду с вычислением маршрутов на основе ограничений должны уметь рассчитывать традиционные маршруты согласно алгоритму SPF.

В процессе эксплуатации сети MPLS расширенные маршрутные IGP-протоколы, такие как OSPF-TE и IS-IS-TE, распространяют в сети следующую информацию о топологии сети и состоянии ресурсов:

- максимальная пропускная способность звена;
- максимальная пропускная способность, доступная для резервирования звена;
- резервированная для звена пропускная способность;
- текущее использование пропускной способности;
- «цвет» ресурса.

Поскольку состояние сетевых ресурсов изменяется намного чаще, чем топология сети, эти расширенные версии протоколов OSPF и IS-IS создают в сети более интенсивный служебный трафик, нежели базовые протоколы, однако это оправданные затраты.

При маршрутизации, основанной на ограничениях (CSPF-маршрутизация), модуль маршрутизации вычисляет маршрут в сети, используя информацию, хранящуюся в TED. Расчет может быть *тактическим* или *стратегическим*.

Тактический расчет используется при возникновении аварий или перегрузок на маршруте, и вместо маршрута, рассчитанного, например, IGP-протоколом, будет автоматически создан обходный TE-LSP.

Стратегический расчет может быть разделен на *online*- и *offline-маршрутизацию*. При стратегической online-маршрутизации все маршруты TE-LSP (или TE-туннелей) вычисляются между граничными узлами MPLS-домена в соответствии с заданными ограничениями, и производится соответствующее резервирование ресурсов, причем вычисления выполняют сами узлы.

Стратегическая offline-маршрутизация отличается от online-маршрутизации только тем, что для вычисления всех маршрутов используется

отдельный сервер, имеющий общее видение сети и ее ресурсов. Таким образом, при offline-маршрутизации можно добиться наиболее эффективного использования сетевых ресурсов, так как в сети не будет возникать противоречивых запросов резервирования, а при вычислении маршрута будут учитываться и остальные TE-LSP (TE-туннели).

Но online-маршрутизация быстрее адаптируется к изменениям, происходящим в сети, поэтому на практике эти два подхода часто используются совместно: TE-LSP рассчитывает offline-сервер, а online-маршрутизация запускается лишь тогда, когда рабочие характеристики TE-LSP перестают удовлетворять предъявляемым к ним требованиям, или происходит заметное изменение состояния сети. Кроме того, процессы offline- и online-маршрутизации могут запускаться с разным периодом, причем первый из них, имеющий больший период, вычисляет маршруты, а второй производит их коррекцию.

Вычисленные таким образом маршруты для TE-LSP позволяют организовать в MPLS-сети сами эти тракты. Модуль маршрутизации передает в сигнальный модуль данные о последовательности *абстрактных узлов*. При создании каждого TE-LSP происходит обмен протокольными сообщениями, в процессе которого вдоль маршрута распределяются метки и резервируются сетевые ресурсы. Могут также учитываться приоритетность создаваемых LSP, производиться вытеснение низкоприоритетного трафика и обрабатываться ситуации соперничества за ресурсы.

После того как все TE-LSP созданы, подсистема TE продолжает эффективно их поддерживать, используя свои дополнительные возможности. Здесь нельзя не упомянуть о такой функции, как *быстрая ремаршрутизация FRR (Fast ReRoute)*. Поскольку с появлением мощных и производительных маршрутизаторов возможности MPLS, упрощающие процесс маршрутизации, постепенно отходят в тень, основными преимуществами этой технологии становятся, во-первых, гибкое управление прохождением трафика (собственно, основная задача TE) и, во-вторых, именно FRR.

Важность быстрой ремаршрутизации обусловлена тем, что для оператора весьма опасны потери при выходе из строя звена. Разумеется, о такой ситуации будет информирован оконечный маршрутизатор, он предпримет попытку создать новый LSP (LSP-туннель) в обход поврежденного участка сети, но из-за задержек, возникающих при передаче сигнальных сообщений к оконечному узлу, в процессе расчета нового маршрута могут произойти ощутимые потери данных в неисправном звене. FRR обеспечивает защиту от этих потерь, ремаршрутизируя трафик, проходящий по LSP, в обход поврежденного звена. При этом решение о ремаршрутизации принимается узлом, непосредственно соединенным с неисправным звеном. Такая локальная ремаршрутизация позволяет предотвратить дальнейшую потерю пакетов и выиграть время для того, чтобы информировать оконечный узел и создать новый LSP.

16. СРАВНЕНИЕ ПРОТОКОЛОВ CR-LDP И RSVP-TE

16.1. Сравнение функциональных возможностей

Следует сразу же отметить, что оба рассматриваемых протокола отвечают требованиям документа RFC 2702 к сигнализации MPLS. Однако для выполнения этих требований CR-LDP и RSVP-TE используют различные механизмы, хотя и схожего в реализации ряда функций очень много. В этом разделе рассмотрены основные функции исследуемых протоколов и варианты их реализации в каждом из них.

Возможность присвоения и переноса параметров трафика и QoS в RSVP-TE реализуется посредством передачи в сообщениях непрозрачных данных, предназначенных для подсистемы управления трафиком. CR-LDP может определять *правила для пограничных узлов (edge rules)* и рекомендации для *промежуточных пересылок (per hop behaviors)*, базирующиеся на скорости передачи данных, на полосе пропускания звена и на весах, присвоенных этим параметрам.

О неисправности протокол RSVP-TE извещает сообщением об ошибке, но момент его отправки зависит от установленных значений таймеров для обновления состояния. CR-LDP же пользуется для извещения о неисправности возможностями транспортного протокола TCP.

Восстановление после неисправности протокол RSVP-TE обеспечивает ремаршрутизацией туннеля в соответствии с принципом *«make-before-brake»*, когда сначала создается новый LSP, затем в него переводится трафик, и лишь после этого разрушается неисправный тракт. CR-LDP позволяет задать политику обработки неисправности в каждом из узлов, через которые проходит LSP.

Обнаружение закольцованных маршрутов требуется лишь для не строго заданных маршрутов и в RSVP-TE производится при помощи объекта RRO. Протокол CR-LDP использует для этого традиционный для LDP объект Path_Vector_TLV. Оба этих объекта могут также применяться для определения того, какой маршрут использует LSP.

Для управления трактами в обоих протоколах применяется идентификация каждого LSP идентификатором LSP ID. При этом RSVP-TE может идентифицировать и туннель (Tunnel ID), позволяя перевести его из одного LSP в другой.

Оба протокола поддерживают функцию вытеснения высокоприоритетным трафиком низкоприоритетного. В обоих протоколах это достигается присвоением приоритетов удержания и захвата ресурса.

Разумеется, оба протокола могут определять для тракта явно заданный маршрут, причем оба они могут работать с абстрактными узлами и со строго и не строго заданными маршрутами.

16.2. Сравнение технических характеристик

Ключевыми различиями между протоколами CR-LDP и RSVP являются используемые тем и другим транспортные протоколы и то, в каком направлении – прямом или обратном – производится резервирование ресурсов. Из различия по этим двум признакам вытекают и многие другие различия этих протоколов. В чем сходны и чем различаются протоколы CR-LDP и RSVP с поддержкой LSP-туннелей, показывает табл. 16.1.

Таблица 16.1

Сравнение протоколов CR-LDP и RSVP-TE

Параметы	CR-LDP	RSVP-TE
Используемый транспортный протокол	TCP	Исходный IP
Надежность операторского класса	Нет	Да
Поддержка трафика «много точек – точка»	Да	Да
Поддержка вещательной рассылки	Нет	Нет
Поддержка слияния LSP	Да	Да
Явная маршрутизация	Со строгими и с нестрогими участками маршрута	Со строгими и с нестрогими участками маршрута
Ремаршрутизация LSP	Да	Да
Закрепление маршрута	Да	Да, путем записи маршрута
Вытеснение потоков в LSP	Да, на основе приоритета	Да, на основе приоритета
Средства безопасности	Да	Да
Защита LSP	Да	Да
Состояние LSP	Жесткое	Нежесткое
Регенерация состояния LSP	Не требуется	Периодическая, по участкам
Резервирование совместно используемых ресурсов	Нет	Да
Обмен параметрами трафика	Да	Да
Управление трафиком	В прямом направлении	В обратном направлении
Авторизация пользователей	Неявная	Явная
Индикация протокола уровня 3	Нет	Да
Ограничения в зависимости от класса ресурса	Да	Нет

Наиболее очевидным различием протоколов CR-LDP и RSVP, показанным в таблице, является то, какой *транспортный протокол* используется для передачи запросов меток. RSVP использует для этого протокол IP, не ориентированный на соединение. CR-LDP использует транспортный протокол UDP, но только для обнаружения одноранговых маршрутизаторов

сети MPLS; для передачи протокольных сообщений он использует ориентированные на соединение TCP-сеансы. В RSVP требуется, чтобы все принятые пакеты IP, переносящие сообщения протокола RSVP, доставлялись к модулю протокола без ссылки на фактический IP-адрес получателя, содержащийся в пакете. Эта особенность может потребовать внесения незначительных изменений в реализацию IP. Имеются два негативных аспекта использования протоколом CR-LDP транспорта TCP. Во-первых, реализованный в TCP механизм предотвращения перегрузки может заметно тормозить передачу информации между маршрутизаторами. Во-вторых, когда между двумя LSR имеется только одно TCP-соединение, TCP принудительно вводит дисциплину FIFO обслуживания очередей сообщений, при которой для критически важного сообщения нет возможности прийти к адресату раньше менее важного сообщения, которое было отправлено первым. Кроме того, когда теряется какой-либо пакет, все сообщения, следующие за этим пакетом, задерживаются до тех пор, пока не будет успешно выполнена его повторная передача.

Протокол CR-LDP наследует все функции *обеспечения безопасности*, которые есть у протокола TCP. К сожалению, TCP уязвим со стороны атак, имеющих цель нарушить обслуживание, при которых рабочие характеристики TCP-сеанса могут серьезно пострадать в результате несанкционированного доступа к сети, что отрицательно повлияет на работу протокола CR-LDP. Адресатом сообщений Path протокола RSVP является выходной LSR, а не промежуточные LSR. Это значит, что *IPSec* (серия разработанных комитетом IETF проектов стандартов обеспечения аутентификации и защиты передаваемых по протоколу IP пакетов путем шифрования) не может использоваться, поскольку промежуточные LSR будут не в состоянии получить доступ к информации, содержащейся в сообщениях Path. Но RSVP имеет свои собственные механизмы аутентификации и авторизации пользователей, позволяющие проверять полномочия каждого отправителя сообщений и не допускать несанкционированного или злонамеренного резервирования ресурсов. Аналогичные возможности могли бы быть специфицированы и для протокола CR-LDP, но ориентированный на соединение характер TCP-сеанса делает это требование менее актуальным, так как TCP может использовать какой-нибудь стандартный криптографический алгоритм.

IP-трафик *«точка – много точек»* не поддерживается протоколами CR-LDP и RSVP-TE. Но базовый протокол RSVP первоначально разрабатывался с учетом возможности резервировать ресурсы для деревьев многоадресной рассылки по протоколу IP, так что его проще расширить для поддержки многоадресного трафика. Поддержка многоадресной рассылки в настоящее время не определена ни для одного из существующих

протоколов распределения меток, но 28 января 2004 г. в рабочей группе MPLS IETF был предложен *draft*-документ, который, если он будет принят, должен определить требования к расширению протокола RSVP-TE для поддержки режима «точка – много точек». Так что у протокола RSVP-TE может в ближайшем будущем появиться функция многоадресной рассылки.

Еще во время спецификации протокола RSVP возникли сомнения относительно возможности его применения в крупных сетях из-за недостаточной *масштабируемости*. Эти сомнения были вызваны тем фактом, что RSVP резервирует ресурсы для индивидуальных «микропотоков», т. е. для потоков таких данных, которые соответствуют, как правило, одному пользовательскому приложению, функционирующему на паре рабочих станций. Число «микропотоков», проходящих через один маршрутизатор в крупной IP-сети, измеряется миллионами, что может предъявить серьезные требования к объему памяти и к производительности маршрутизаторов. Это обстоятельство иногда приводит к утверждению о слабой масштабируемости RSVP. Но в действительности мы говорим об отсутствии хорошего масштабирования RSVP только тогда, когда базовый RSVP используется, чтобы резервировать ресурсы для «микропотоков» индивидуальных пользовательских приложений, а когда протокол RSVP-TE используется для создания явно заданных трактов LSP, такое «микрорезервирование» не делается. Доминирующим фактором, определяющим масштабируемость протокола создания LSP, является число создаваемых трактов LSP, которое, разумеется, от протокола не зависит.

Все ориентированные на соединение протоколы требуют *хранения данных* о состояниях соединений как на оконечных LSR, так и на промежуточных. При использовании протокола RSVP-TE требования во многом схожи во всей сети, потому что информация о состоянии должна храниться и периодически обновляться в каждом LSR. В эту информацию включаются параметры трафика, данные о резервированных ресурсах и об явно заданных маршрутах. Объем этой информации составляет порядка 500 байтов на один LSP.

Протокол CR-LDP требует, чтобы входной и выходной LSR хранили сходные объемы информации о состоянии, включая параметры трафика и данные об явно заданных маршрутах. Суммарный объем информации о состоянии, требуемый для функционирования протокола CR-LDP, составляет на оконечных узлах тоже порядка 500 байт. В промежуточных LSR возможно снизить объем хранимой информации примерно до 200 байт за счет отказа от функции модификации LSP, т. е. от возможности ремаршрутизации LSP или от учета изменений ресурсов. Следует отметить, что буферы пересылки данных, необходимые для обеспечения гарантированного QoS при передаче по LSP, будут иметь намного больший размер, чем

объем памяти, нужный для хранения данных о состоянии. Таким образом, различие между протоколами RSVP и CR-LDP в сети MPLS, не требующей поддержки модификации LSP, менее существенно.

Под *надежностью операторского класса* понимается коэффициент готовности в «пять девяток», т. е. 99,999 %. Высокая готовность достигается за счет своевременного обнаружения и устранения неисправностей без какого-либо (или, в крайнем случае, только минимального) нарушения обслуживания. Вопросы обеспечения живучести LSP при возникновении программных или аппаратных отказов относятся к реализации оборудования, и эти вопросы обязан рассматривать и решать каждый поставщик сетевого оборудования MPLS.

В связи с тем что протокол RSVP использует транспортный механизм без установления соединения, он хорошо приспособлен к системе, которая должна быть устойчивой к аппаратным отказам или сбоям программного обеспечения. Сведения о любых управляющих действиях, теряющиеся во время аварийного переключения на резервную систему, могут быть восстановлены посредством встроенного в протокол RSVP-TE механизма регенерации состояния.

С другой стороны, протокол CR-LDP предполагает надежную доставку управляющих сообщений и поэтому он более чувствителен к аварийным переключениям на резерв. Кроме этого, протокол TCP очень сложно сделать отказоустойчивым, и поэтому аварийное переключение на резервный стек протоколов TCP приводит к потере TCP-соединений. Такая ситуация интерпретируется протоколом CR-LDP как отказ всех соответствующих LSP, которые надо создавать заново от входного LSR.

Таким образом, изначально протокол RSVP может обеспечивать лучшие решения для сетей MPLS с высоким уровнем готовности. Ситуацию с CR-LDP исправляет документ RFC 3479, специфицирующий механизмы отказоустойчивости для этого протокола. После внедрения этих расширений CR-LDP приблизится по характеристикам обеспечения высокой готовности к протоколу RSVP.

С надежностью непосредственно связано *обнаружение отказов в звеньях и в маршрутизаторах.* Если два LSR непосредственно связаны двухточечным каналом, например каналом ATM, неисправность в LSP можно, как правило, обнаружить путем мониторинга состояния интерфейсов LSP. Например, если в канале ATM пропадает сигнал, то и протокол CR-LDP, и протокол RSVP могут использовать уведомление об отказе в интерфейсе для обнаружения отказа LSP. Если два LSR соединены друг с другом через совместно используемую среду передачи, например Ethernet, или соединены друг с другом не непосредственно, а через облако WAN, то они не обязательно получают уведомление об отказе звена из канального

оборудования. В таких случаях задача обнаружения отказов LSP перекладывается на средства, имеющиеся в протоколах сигнализации. Протокол CR-LDP использует обмен сообщениями Hello и Keepalive для того, чтобы удостовериться, что смежный LSR и звено продолжают оставаться активными. Несмотря на то что протокол TCP имеет встроенную систему поддержания соединения, она, как правило, слишком медленно, с точки зрения нужд LSP сети MPLS, реагирует на отказы в звене и в маршрутизаторе. В протоколе RSVP периодически передаваемые для регенерации состояния сообщения Path и Resv образуют фоновый трафик, который указывает на то, что звено продолжает оставаться работоспособным. Однако для сведения к минимуму этого трафика в относительно стабильной сети для таймера регенерации может быть установлено достаточно большое значение. В протоколе RSVP-TE для подтверждения активности и работоспособности звена и смежного LSR может использоваться обмен сообщениями Hello. Таким образом, методика и быстрота обнаружения отказов в обоих сравниваемых протоколах схожи.

Организация лямбда-сетей поднимает значительную совокупность проблем, возникающих при реализации технологии MPLS в оптической сети. Всеми преимуществами спектральной коммутации можно воспользоваться только в том случае, если коммутация LSP выполняется аппаратными средствами без помощи ПО. Число волн, однако, слишком мало по сравнению с вероятным числом LSP. Кроме того, возможности одной волны значительно превышают обычные требования одного LSP, поэтому организация взаимно-однозначного соответствия между ними была бы непозволительно расточительной тратой сетевых ресурсов. Рабочей группой IETF по вопросам MPLS выпущены документы RFC, содержащие необходимые дополнения как для протокола RSVP-TE, так и для CR-LDP, обеспечивающие их работоспособность в лямбда-сетях в рамках концепции GMPLS, о чем будет сказано в следующем разделе.

Важно отметить, что в контексте *управления трафиком* протоколы CR-LDP и RSVP-TE выполняют резервирование ресурсов на разных стадиях процесса создания LSP.

Протокол CR-LDP переносит полную информацию о параметрах трафика в сообщении запроса метки Label Request. Это позволяет каждому маршрутизатору сети MPLS управлять трафиком при создании LSP. Параметры трафика могут согласовываться по мере продвижения процесса создания LSP, а окончательные значения параметров передаются в обратном направлении в сообщении назначения метки Label Mapping, что обеспечивает контроль доступа и резервирование ресурсов в каждом LSR в реальном времени. Этот подход позволяет не создавать LSP по маршруту, который не имеет в данный момент времени достаточных ресурсов.

Протокол RSVP-TE переносит в сообщении Path набор параметров трафика в виде спецификации потока данных отправителя Tspec. Эта спецификация описывает данные, которые будут передаваться по LSP. Транзитные LSR могут анализировать эту информацию и на ее основе принимать решения о маршрутизации. Однако только тогда, когда сообщение Path дойдет до выходного LSR, спецификация Tspec будет преобразована в спецификацию Flowspec, которая переносится в обратном направлении сообщением Resv и в которой даются детали резервирования ресурсов, требующихся для LSP. Это означает, что резервирование не происходит до тех пор, пока сообщение Resv не пройдет через сеть, а результатом может стать то, что на выбранном маршруте создать LSP не удастся из-за нехватки ресурсов.

Протокол RSVP-TE содержит необязательную функцию Adspec, посредством которой можно сообщить о доступных ресурсах в сообщении Path. Это позволяет выходному LSR узнать, какие ресурсы доступны, и в соответствии с этим модифицировать спецификацию Flowspec, переносимую в сообщении Resv. К сожалению, эта функция не только требует, чтобы ее поддерживали все LSR на маршруте, но также имеет тот очевидный недостаток, что ресурсы, сведения о которых несет сообщение Path, к тому времени, когда будет получено сообщение Resv, уже могут оказаться занятыми другим LSP.

Частичное решение этой проблемы для LSR, использующих протокол RSVP, лежит в области реализации: можно обеспечить предварительное резервирование ресурсов при обработке сообщения Path. Это резервирование будет лишь приблизительным, поскольку оно опирается на спецификацию потока данных отправителя Tspec, а не на спецификацию Flowspec, но, тем не менее, оно может значительно облегчить положение.

Протокол CR-LDP предлагает несколько более жесткий подход к управлению трафиком, особенно в сетях, испытывающих высокую нагрузку.

Протокол RSVP-TE позволяет переносить в сообщениях Path и Resv объект *авторизации пользователей* с непрозрачным содержимым. Эта информация используется при обработке сообщений для контроля доступа в соответствии с установленными правилами. Это позволяет тесно привязать протокол RSVP-TE, поддерживающий распределение меток, к протоколам авторизации и надзора, например в соответствии с RFC 2749 – к COPS (Common Open Policy Service).

В отличие от этого протокол CR-LDP позволяет переносить в неявном виде только информацию авторизации в форме адресов получателя и класса административного ресурса в параметрах трафика.

Различие между протоколами существует и в *индикации протоколов уровня 3*. Хотя LSP может переносить любые данные, бывают случаи, когда транзитному или выходному маршрутизатору сети MPLS может

потребоваться информация о протоколе уровня 3. Если транзитный LSR не в состоянии доставить пакет (например, из-за отказа ресурса), он может передать в обратном направлении специфическое для протокола уровня 3 сообщение об ошибке, уведомляющее отправителя данных о возникшей проблеме. Чтобы этот механизм работал, LSR, который выявляет ошибку, должен знать, какой именно протокол уровня 3 используется. Информация о протоколе уровня 3 может также помочь выходному маршрутизатору пересылать пакеты данных.

RSVP-TE идентифицирует один единственный протокол передачи полезной нагрузки на этапе создания LSP, но CR-LDP и этого делать не в состоянии. Даже RSVP не может помочь, когда в один LSP направляется трафик более чем одного протокола.

Тракты LSP, созданные с выбором оптимального маршрута в сети, могут управляться на их входе и контролироваться на их выходе с помощью *административной базы данных MIB* сети MPLS. Эта MIB в настоящее время разрабатывается, и в рабочей группе на эту тему существует ряд проектов *draft*. В связи с весьма пессимистическими планами работы над CR-LDP эти проекты, скорее всего, будут ориентированы на поддержку RSVP-TE, хотя на ранней стадии разработки проектов MIB предпочтение отдавалось как раз CR-LDP.

16.3. Служебный трафик в RSVP-TE и CR-LDP

Оба протокола создают потоки служебных сообщений для раздачи меток, передачи сквозного запроса и сквозного ответа на запрос. Протокол RSVP ориентируется на «нежесткое состояние» маршрутов. Это значит, что протокол должен периодически обновлять состояние каждого LSP между смежными узлами, что позволяет автоматически учитывать изменения в дереве маршрутизации. В качестве транспортного средства протокол RSVP использует IP-дейтаграммы, так что управляющие сообщения могут теряться, и смежный узел, не получив соответствующего уведомления, может прекратить обслуживание. Регенерация состояния LSP гарантирует, что оно надлежащим образом синхронизировано между смежными узлами. Нагрузка на сеть от таких периодических обновлений зависит от чувствительности к отказам, которая регулируется выбором значения таймера регенерации (обновления). Сообщение Path протокола RSVP будет иметь длину порядка 128 байт, увеличиваясь на 16 байт в каждом маршрутизаторе, если используется фиксированный маршрут. Сообщение Resv будет иметь длину порядка 100 байт. При 10 000 LSP в межузловом звене и периоде регенерации 30 с на регенерацию будет потребляться свыше 600 кбит/с пропускной способности звена. Много это или нет – зависит от характеристик звена и от того, какую это составляет долю от передаваемого по нему трафика.

Протокол CR-LDP не требует от LSR периодической регенерации каждого LSP после его создания. Это достигается благодаря тому, что в качестве транспортного средства для передачи сигнальных сообщений протокола CR-LDP используется протокол TCP. Протокол CR-LDP может обеспечивать надежную доставку сообщений Label Request и Label Mapping. Использование транспортного протокола TCP в тракте сигнализации никакой служебной информации в этот тракт не вносит, а только добавляет 20 байт к длине каждого сигнального сообщения. Для поддержания связности со смежными узлами протокол CR-LDP использует сообщения Hello, с помощью которых он удостоверяется, что смежные узлы продолжают оставаться активными, и сообщения KeepAlive для мониторинга TCP-соединений. Периодический обмен этими относительно короткими сообщениями ведется для всего звена, а не для каждого из многочисленных LSP, по нему проходящих, и потому эти сообщения практически не оказывают влияния на пропускную способность звена. Таким образом, протокол CR-LDP, в принципе, вносит меньшую нагрузку в сеть, чем протокол RSVP. Но IETF специфицировал документ RFC 2961, в который вошли идеи уменьшения числа сообщений регенерации, требуемых протоколом RSVP. Рассмотрим подробнее решение этой проблемы, предложенное и детализированное в рабочих группах MPLS и RSVP в составе комитета IETF.

С учетом взаимосвязи между надежной доставкой информации и непроизводительными затратами на процедуры регенерации, одним из шагов для снижения объема обрабатываемой информации регенерации могло бы стать добавление в протокол RSVP механизма надежной доставки сообщений. Это было сделано путем определения пары объектов протокола RSVP, а именно объектов MESSAGE_ID и MESSAGE_ID_ACK. Эти объекты обеспечивают надежную доставку сообщений следующим образом. Предположим, что узел LSR1 имеет сообщение протокола RSVP, представляющее новую часть состояния, например новое сообщение PATH, которое он должен передать к соседнему узлу LSR2. Узел LSR1 создает локально уникальный идентификатор нового состояния, помещает этот идентификатор в объект MESSAGE_ID и добавляет его к сообщению протокола RSVP, установив в объекте флаг «требуется подтверждение». Получив сообщение, LSR2 подтверждает его получение, передавая к LSR1 сообщение, которое содержит объект MESSAGE_ID_ACK с тем же идентификатором. Если LSR2 уже имеет какое-либо сообщение, подлежащее передаче к LSR1, он просто включает объект MESSAGE_ID_ACK в это сообщение, в противном случае ему надо будет передать специально созданное сообщение подтверждения ACK. Этот простой механизм может следующим образом обеспечить решение проблемы, связанной с большими объемами информации регенерации. Узел, поддерживающий RSVP,

может использовать очень короткий таймер регенерации для передаваемых им сообщений, приводящих к установлению нового состояния. При получении от соседнего узла подтверждения того, что новое состояние установлено, он может переключиться на очень длинный таймер регенерации. Таким образом, объем информации регенерации снижается без ущерба для своевременности резервирования.

При этом такой механизм отличается от «жесткого состояния», поскольку сообщения регенерации все равно должны периодически передаваться. Следовательно, сохраняются присущие нежесткому состоянию свойства «самовосстановления», включая интерактивную обработку отказов. Кроме того, этот механизм обладает обратной совместимостью, т. е. не создает никаких проблем, если в одних узлах он реализован, а в других – нет.

Дальнейшее снижение объема информации регенерации достигается с помощью механизма, который называется *сокращенной* или *ускоренной регенерацией (summary refresh).* Суть его в том, что после того, как идентификатор сообщения уже привязан к сообщению протокола RSVP, для регенерации состояния нет необходимости передавать снова и снова полное сообщение, как это обычно делается. Узел может просто передавать сам идентификатор сообщения. Кроме того, возможна упаковка довольно большого числа идентификаторов сообщений в одно сообщение, так что одно-единственное сообщение протокола RSVP может обновлять целый ряд состояний. Такая возможность полезна, поскольку часто требуются значительные вычислительные затраты только на получение сообщения в процессор маршрутизатора, вне зависимости от того, как велико это сообщение, или от того, каковы затраты на обработку его содержимого.

Для поддержки механизма сокращенной регенерации было определено новое сообщение протокола RSVP – сообщение SREFRESH. При адресации к определенному устройству это сообщение содержит просто перечень ранее переданных идентификаторов сообщений. Узел, получающий такое сообщение, ведет себя так, как если бы он получил набор сообщений регенерации, каждое из которых идентично сообщению, в котором был первоначально передан объект MESSAGE_ID, т. е. он сбрасывает таймер регенерации.

Одна из возможных проблем – получение идентификатора MESSAGE_ID, который не распознается. Это может произойти, если следующий участок маршрута изменился, но передающий узел это изменение не обнаружил. Для выхода из такой ситуации принимающий узел, не распознавший идентификатор сообщения, может передать отрицательное подтверждение MESSAGE_ID_NAK. Передающий узел, принявший такое подтверждение, должен передать полное сообщение протокола RSVP, к которому относился MESSAGE_ID.

В результате всех этих усовершенствований были решены проблемы, связанные с масштабируемостью RSVP при его использовании для MPLS. Появилась возможность многократно производить резервирование ресурсов и трактов LSP при незначительном объеме трафика регенерации.

Проблема снижения объема передаваемой служебной информации в RSVP-TE была рассмотрена нами столь подробно потому, что обычно одним из аргументов противников RSVP-TE был именно трафик обновления состояний. Теперь же видно, что после реализации в протоколе RSVP механизма, сокращающего объем информации регенерации, различие между RSVP-TE и CR-LDP по этому признаку становится несущественным.

16.4. Сравнение ремаршрутизации в RSVP-TE и CR-LDP

Рассмотрим вопрос изменения маршрута для LSP после получения уведомления об отказе или при изменении топологии сети. Предварительное программирование резервных (альтернативных) маршрутов для тракта LSP называется *защитой LSP*. Явно заданный LSP может быть ремаршрутизирован только входным LSR – отправителем данных. Следовательно, об отказе в некоторой точке LSP должен быть информирован входной LSR, и при этом постепенно разрушается весь LSP. Однако нестрого специфицированный участок LSP с явным маршрутом и любая часть LSP, маршрут для которого задавался по участкам, могут быть ремаршрутизированы, если обнаружен отказ звена или смежного маршрутизатора (*локальное восстановление*), или если стал доступен лучший маршрут, или если ресурсы LSP требуются для создания нового LSP с более высоким приоритетом (*приоритетное вытеснение*). Выше уже отдельно рассматривался случай, когда сеть поддерживает функцию Fast ReRoute.

Ремаршрутизацию, управляемую входным узлом, поддерживает как протокол CR-LDP, так и протокол RSVP-TE, хотя имеются незначительные различия в том, как это делается.

Маршрутизатор LSR, использующий протокол RSVP-TE, может определить новый маршрут, как только станет доступным и/или потребуется альтернативный маршрут, просто путем обновления LSP, в результате чего выбирается другой следующий маршрутизатор. Старый тракт при этом не используется и разрушается после срабатывания таймера, поскольку сообщения регенерации по нему больше не передаются. Ясно, что при таком способе непроизводительно тратятся ресурсы старого LSP.

Этого можно избежать, передав по явно заданному маршруту сообщение *PathTear* или *ResvTear*. При этом активизируется *механизм с включением нового тракта до разрыва старого (make-before-break)*, т. е. механизм, при котором старый тракт продолжает использоваться (и обновляться) все

время, пока создается новый тракт, а после его создания LSR, производящий ремаршрутизацию, выполняет переключение на новый тракт и разрушает старый тракт. Эта методика позволяет избежать двойного резервирования ресурсов как в протоколе CR-LDP (использованием значения *modify* флага действия в сообщении Label Request), так и в протоколе RSVP-TE (использованием фильтров при стиле резервирования Shared-Explicit).

В нестабильных сетях интенсивный служебный трафик, обеспечивающий ремаршрутизацию нестрого специфицированных участков LSP на промежуточных LSR при появлении лучших маршрутов, может приводить к возникновению перегрузок. Для того чтобы это предотвратить, нестрого специфицированный участок маршрута может закрепляться следующим образом.

В протоколе CR-LDP это делается просто путем пометки нестрогого участка явно заданного маршрута как закрепленного. Это означает, что как только маршрут будет определен, к нему будут относиться, как к строго специфицированному маршруту, который изменяться не может.

В протоколе RSVP закрепление требует некоторой дополнительной обработки. Пусть первоначальный маршрут специфицирован с нестрогим участком. Чтобы информировать входной LSR о выбранном маршруте, в сообщениях Path и Resv используется объект Record Route (RRO). Входной LSR может затем использовать эту информацию для повторной передачи сообщения Path, которое будет содержать строго специфицированный явный маршрут.

И RSVP-TE, и CR-LDP используют гибкий подход к ремаршрутизации LSP и к применению механизма с включением нового LSP до разрыва старого. Протокол CR-LDP опирается на внесенное в его спецификацию добавление, позволяющее поддерживать включение нового до разрыва старого, а протокол RSVP-TE требует дополнительного обмена сообщениями для закрепления маршрута.

Стоит отметить один из недостатков протокола CR-LDP, связанный с использованием протокола TCP для LDP-сессий. При работе CR-LDP имеют место дополнительные затраты времени при определении отношения смежности между двумя LSR. Перед тем как маршрутизаторы смогут инициировать LDP-сеанс, они должны пройти процедуру вхождения в связь по протоколу TCP. Это дает преимущество протоколу RSVP-TE, который не требует установления соединения перед началом процедуры распределения меток. Можно сказать, что протокол RSVP имеет «облегченные отношения смежности», которые позволяют определять новые взаимосвязи между соседними маршрутизаторами быстро, по мере необходимости. И это важно для выполнения быстрой ремаршрутизации.

Модификация LSP, например, при изменении параметров трафика в LSP, является операцией, эквивалентной ремаршрутизации, хотя при модифика-

ции LSP изменение маршрута не является обязательным. Следовательно, эта функция всегда присутствует в протоколе RSVP-TE и будет присутствовать в протоколе CR-LDP при условии, что реализация протокола поддерживает значение modify флага действия в сообщениях Label Request (дело в том, что ранние реализации протокола модификацию LSP не поддерживают).

Защита LSP заключается в программировании резервных путей для тракта LSP с автоматическим переключением на резервный тракт при отказе основного тракта. Несмотря на то что с концептуальной точки зрения эта функция также аналогична ремаршрутизации, защита LSP обычно рассматривается как намного более важная операция, целью которой является оперативное переключение на новый тракт с минимально возможным прерыванием передачи данных по LSP (как правило, расчетная длительность прерывания должна быть менее 50 мс). В обоих сравниваемых протоколах могут поддерживаться несколько уровней защиты LSP.

Простейшим видом защиты LSP является попытка входного или транзитного LSR выполнить ремаршрутизацию LSP немедленно по получении уведомления об отказе. Такая возможность существует в обоих протоколах, однако из-за необходимости передавать разнообразные сигнальные сообщения аварийное переключение на новый маршрут происходит относительно медленно (обычно это занимает как минимум несколько секунд). Такая невысокая скорость переключения неприемлема для IP-телефонии и некоторых других приложений реального времени.

Намного более быстрой защиты LSP можно достичь, если звено между двумя LSR защищено схемой защиты на уровне 2. Такая защита прозрачна для LSP и может применяться с любым из двух сравниваемых протоколов. Впрочем, реализация защиты на уровне 2 может оказаться дорогим делом, и сама защита ограничена участком LSP между двумя соседними маршрутизаторами. К тому же, защита звена не обеспечивает защиты от отказа отдельных LSR.

17. РАЗВИТИЕ MPLS. GMPLS. ЦЕНТРАЛИЗОВАННОЕ УПРАВЛЕНИЕ ОПТИЧЕСКИМИ СЕГМЕНТАМИ ТРАНСПОРТНЫХ СЕТЕЙ СВЯЗИ

Эта идея о возможности распространить парадигму замены меток MPLS на новые оптические технологии была впервые представлена как спецификация для управления световыми путями, или оптическим следами. Каждый *узел OXC (Optical Cross-Connect)* взаимодействует подобно MPLS LSR с плоскостью управления. Между соседними OXC устанавливается отдельный канал управления IP. Световые пути становятся путями LSP, а выбор лямбд и портов кросс-соединения является процессом, аналогичным назначению меток и связок метка-FEC.

Таким образом, MPLS эволюционирует к GMPLS путем расширения (генерализации) понятия метки на различные коммутационные приложения: временное мультиплексирование TDM, частотное мультиплексирование FDM и пространственное мультиплексирование SDM. Метки больше не являются исключительно дополнительными полями в заголовке пакета сетевого уровня, как это описано в разделе 2, они могут быть также оптическими лямбдами и т. п.

Существуют четыре класса путей, которые можно создавать с помощью сигнализации GMPLS:

- статистически мультиплексированные пути – обычные пакеты MPLS, использующие промежуточный заголовок;
- пути TDM – каждый временной канал является меткой;
- пути FDM – каждая электромагнитная частота (т. е. длина световой волны) является меткой;
- пути SDM – меткой является позиция, например местоположение волокна в пучке.

Технология GMPLS специфицирована в документе RFC 3471, описывающем саму технологию, а также в RFC 3472, посвященном сигнализации CR-LDP, и RFC 3473, описывающем протокол RSVP-TE. Все три документа имеют схожую структуру, но если в RFC 3471 специфицируются только информационные поля сообщений, то в двух других документах эти поля вдобавок обрамляются служебной информацией конкретного протокола.

Это позволяет добавить еще одну трактовку слова *Generalized* в название технологии. Речь идет о том, что GMPLS расширяет и число возможных типов оборудования, входящего в состав MPLS-сети. Универсальность новой архитектуры заключается в том, что она может включать в себя LSR, неспособные анализировать заголовки пакетов, но осуществляющие маршрутизацию, основываясь на временных интервалах, длинах волн или физических портах. Таким образом, все LSR, а точнее – интерфейсы на LSR, могут быть подразделены на следующие классы:

• интерфейсы *Packet-Switch Capable (PSC)*, способные различать границы пакетов и ячеек и осуществлять маршрутизацию, основываясь на содержании их заголовков, например, обычные LSR или ATM-коммутаторы;

• интерфейсы *Time-Division Multiplex Capable (TDM)*, через которые маршрутизируют данные, основываясь на их временных интервалах, например интерфейсы SDH или входы коммутационного поля цифровой АТС;

• интерфейсы *Lambda Switch Capable (LSC)*, через которые осуществляется маршрутизация на основе длины волны, т. е. интерфейсы *оптических лямбда-коммутаторов*;

• интерфейсы *Fiber-Switch Capable (FSC)*, через которые маршрутизируются данные, основываясь на их реальной физической среде переноса, например интерфейсы оптического коммутатора, работающего на уровне одного или нескольких оптических волокон.

Отметим, что PSC-интерфейсы ничем не отличаются от используемых в традиционной MPLS-сети, в связи с чем ниже будем различать *PSC-* и *не-PSC-интерфейсы*, так как именно наличие последних является причиной появления большинства опций GMPLS.

Уже обсуждавшийся выше принцип вложенных LSP позволяет масштабировать систему, формируя иерархию маршрутизации. Наверху иерархии расположатся FSC-интерфейсы, за ними – LSC, потом TDM и наконец PSC. Таким образом, LSP, который начинается и заканчивается на PSC-интерфейсе, может (вместе с несколькими другими LSP) быть помещен в LSP, начинающийся и заканчивающийся на TDM-интерфейсе. Также LSP уровня TDM и других уровней могут быть вложены в LSP, иерархически расположенные выше.

По причине использования в GMPLS интерфейсов, отличных от PSC, появилась необходимость установления двунаправленных LSP. В частности, они используются для предотвращения конфликтной ситуации захвата ресурсов, возникающей при установлении различных LSP по отдельным сигнальным сессиям; а также для упрощения процедур восстановления после сбоев на не-PSC оборудовании. Также двунаправленные LSP имеют преимущества меньшего времени задержки при их установлении и меньшего количества необходимых для этой операции сообщений.

В технологии GMPLS формализуется возможное разделение каналов сигнализации и каналов данных, что важно для поддержки технологий, где сигнальный трафик не может посылаться вместе с пользовательской информацией. Также GMPLS предусматривает возможность расширения протоколов сигнализации специфическими параметрами для поддержки конкретных технологий.

GMPLS определяет еще несколько усовершенствований, которые необходимы для выполнения работы MPLS в оптических сетях, в число

которых входят связывание каналов, ненумерованные каналы и новый базирующийся на IP *протокол LMP (Link-Management Protocol)*. Связывание каналов представляет собой агрегатирование атрибутов каналов более чем одного параллельного канала в единый пучок каналов. Выигрыш от такого связывания состоит в уменьшении величины базы данных состояний каналов и улучшении некоторых важных характеристик масштабирования. Ненумерованные каналы поддерживают каналы, которые не сконфигурированы IP-адресами. Использование альтернативной идентификации каналов упрощает многие задачи управления каналами. Вновь предложенный ненумерованный тег канала является кортежем «идентификатор маршрутизатора/номер канала». LMP – дополнительный протокол управления, который обусловлен особыми оптическими требованиями мониторинга и управления между двумя соседними оптическими узлами. LMP обеспечивает верификацию связности каналов, корреляцию свойств каналов, управление управляющими каналами и локализацию неисправностей.

18. TP-MPLS и T-MPLS

T-MPLS (*Transport Multiprotocol Label Switching – транспортная многопротокольная коммутация по меткам*) представляет собой технологию, разработнанную специально для применения в пакетных траспортных сетях операторов связи. Как следует из названия, T-MPLS основана на технологии MPLS. Работа над набором стандартов T-MPLS была начата в феврале 2006 г. Предполагалось создание ряда стандартов, описывающих построение пакетной сети, ориентированной на соединения. Реализация T-MPLS в рамках разработанного ряда стандартов представляется как упрощенная технология MPLS. Упрощение достигается за счет отказа от функциональности, не связанной с ориентированной на соединения пакетной сетью (упрощение самой технологии MPLS), и за счет добавления механизмов, критичных для функционирования транспортной сети (упрощение управления и обслуживания пакетной сети).

Работа над развитием стандартов T-MPLS была прекращена ITU-T в декабре 2008 г. В качестве последующего развития идей, заложенных в T-MPLS, предлагается новая технологией MPLS-TP.

В T-MPLS используются те же самые принципы построения архитектуры сети, что и в технологиях SDH или OTN. Выбор архитектуры построения технологии T-MPLS, близкой к технологиям SDH и OTN обусловлен тем, что операторы связи имеют разработанные процессы управления и рабочие процедуры для развития и обслуживания сетей связи, основанных на этих технологиях.

При разработке стандартов T-MPLS из технологии MPLS были удалены некоторые функции, относящиеся к сетевому уровню. Это позволило на всем участке сети T-MPLS между границами гарантировать целостность OA&M пакетов, используемых для мониторинга и защитного переключения туннелей.

Отличием T-MPLS от MPLS является отсутствие в стандарте T-MPLS следующих функций:

- удаление MPLS метки на предпоследнем узле сети (PHP, *Penultimate Hop Popping*);
- объединение LSP, когда весь трафик, посылаемый в одном направлении с одного узла, может использовать одну и ту же MPLS метку;
- множественные маршруты равной стоимости (ECMP, *Equal Cost Multiple Path*);
- однонаправленные LSP. В стандарте T-MPLS используются только двунаправленные LSP.

При удалении MPLS-метки на предпоследнем узле в обычной MPLS-сети последнее соединение в маршруте LSP основано на протоколе IP, а

не на MPLS. Это позволяет снизить нагрузку на маршрутизатор, но контроль целостности OA&M-пакетов до последнего узла становится более сложным и даже невозможным, так как значение MPLS-метки теряется. Поэтому данная функция была удалена из стандарта T-MPLS, так как она противоречит ориентированной на соединения архитектуре.

Объединение LSP также противоречит ориентированной на соединения архитектуре, так как при этом затрудняется определение источника трафика, возникают проблемы с OA&M-пакетами и мониторингом производительности.

Функция множественных маршрутов равной стоимости (ECMP) также была удалена из стандарта T-MPLS, так как позволяла трафику внутри одного LSP передаваться поверх нескольких путей. Это усложняло обработку IP-заголовков, MPLS-меток и механизмы OA&M и мониторинга производительности.

Причина отказа от однонаправленных LSP в стандарте T-MPLS заключается в особенности использования T-MPLS как технологии транспортной сети оператора. Транспортные сети обслуживают двунаправленные соединения, поэтому T-MPLS работает с парами прямых и обратных LSP, следующих через одни и те же узлы и соединения. Двунаправленные LSP позволяют использовать новые механизмы защиты трафика, основанные на использовании центральной системы сетевого управления, отличные от используемой в MPLS быстрой перемаршрутизации (FRR – *Fast Reroute*).

MPLS-TP, или MPLS Transport Profile, разрабатывается совместными усилиями ITU-T и IETF. Изначально поставщики телекоммуникационного оборудования предложили облегченную версию MPLS, содержащую только ту часть MPLS, которая необходима для создания туннеля, ориентированного на установление соединения. Ставилась задача, с одной стороны, упростить IP/MPLS, а с другой – добавить необходимые для транспортных сетей функции и сделать технологию IP/MPLS независимой от сигнальных протоколов IP-маршрутизации (отделить плоскости управления и продвижения данных от предоставляемых сетью сетевых служб более высоких уровней). Это предложение уже рассмотрено ITU, оно получило название T-MPLS. Эта технология должна была упростить структуру и управление по сравнению с полной версией MPLS, что обещало привести и к удешевлению оборудования. Однако проблема совместимости нового стандарта с существующими устройствами MPLS заставляет ITU вернуть в T-MPLS исходные функции MPLS. В апреле 2008 г. создается объединенная рабочая группа с IETF для проработки вопросов совместимости. Результат: в июне того же года происходит официальный отказ от T-MPLS в пользу нового протокола MPLS-TP. Предполагается, что MPLS-TP будет основы-

ваться на тех же самых архитектурных принципах разделения на сетевые уровни, которые используются в давно представленных транспортных технологиях подобно SDH, SONET и OTN. Технология будет заимствовать от MPLS функции создания соединений на основе коммутации пакетов с добавлением к ним механизмов, которые обеспечивают поддержку критической транспортной функциональности. Фактор широкой распространенности MPLS-технологии играет в пользу этого подхода, однако скептики опасаются, что в результате MPLS-TP может стать такой же сложной, громоздкой и дорогой, как и исходная технология MPLS.

19. ДАЛЬНЕЙШАЯ ЭВОЛЮЦИЯ ТРАНСПОРТНЫХ СЕТЕЙ СВЯЗИ

Длительное время телекоммуникационная отрасль обсуждала пользу от создания унифицированного контроля в транспортных сетях. Так было принято решение двигаться эволюционно, развивая достаточно успешный MPLS. GMPLS был разработан как набор MPLS-протоколов управления с постепенным расширением их на специфику работы сетей с коммутацией каналов. Специалисты технических групп и научных лабораторий планировали развивать протоколы сигнализации и распределения меток, чтобы постепенно перейти к интеллектуальному, автоматизированному и унифицированному уровню управления различными сетями с разными принципами передачи данных (пакеты, тайм-слоты, длины волн и т. д.).

GMPLS не стал технологией-революцией и широкого распространения в глобальной сети не получил, несмотря на научно-исследовательскую работу и обсуждения ведущими организациями: ITU, IETF и OIF.

Функционирование транспортных технологий семейства MPLS/GMPLS опирается на уже работающую сеть с маршрутизацией OSPF, IS-IS или EIGRP. Данные протоколы сетевой маршрутизации не обладают высокоуровневыми и централизованными механизмами управления. При возникновении отказа на сети, подключения новых узлов или настройки ограничения распространения маршрутной информации между узлами, возникают проблемы со стабильностью маршрутизации и быстрым восстановлением сети. Например, обеспечить быстрое подключение услуги MPLS VPN для нового сайта через сеть OSPF – не простая задача. Сквозная автоматизация процесса подключения таких услуг требует изменения архитектуры сети и правил взаимодействия с OSS/BSS-инфраструктурой.

За последние несколько лет ряд научно-технических исследований был посвящен проблеме создания единого уровня управления в транспортных сетях. Так, в исследовании Стэнфордского университета, посвященном унификации уровня управления транспортными сетями, выносится тезис о том, что протоколы управления MPLS (сигнализация и распределение меток) для создания единого уровня управления транспортной сетью не подходят. Дальнейшее развитие этих протоколов и их адаптация не является эффективной, так как путь усложнения приводит к обратному эффекту: увеличение функциональности протоколов сигнализации одновременно усложняет работу сетевых узлов, снижает гибкость и масштабируемость сети и ее устойчивость к сбоям.

Необходимо разработать гибкое, программируемое и зонтичное решение для реализации единого уровня управления.

Внедрение нового решения управления может вызвать дополнительные сложности в области взаимодействия с существующей сетью и сетевыми элементами. Интерфейсы взаимодействия между единым уровнем управления и инфраструктурой оператора, ровно, как и сетью, должны быть открытыми. Иначе оператор рискует быть зависимым от функционала, предоставляемого одним вендором.

Создание подхода единого управления сетью, максимально безболезненного с точки зрения внедрения и максимально гибкого с точки зрения функционала, есть критически важная задача отрасли связи.

Единственным кандидатом для создания такой архитектуры с единым уровнем управления является подход программно-конфигурируемых сетей.

Как уже отмечалось, оригинальные исследования по созданию концепции SDN проводились в университетах Stanford и UC Berkley с середины 2000-х гг., тогда сам термин стал проникать в обиход. Исследовательское понимание SDN самими создателями концепции можно выразить, процитировав Мартина Касадо, одного из главных идеологов подхода: «SDN – это сетевая архитектура, в которой уровень управления отделен от уровня передачи данных и сетевые устройства обеспечивают большую функциональность по сравнению с традиционными решениями».

Таким образом, SDN ставит своей задачей создать такую архитектуру сети, при которой сеть открыта к интеллектуальному, автоматизированному и унифицированному развитию и управлению.

Сейчас существуют два независимых подхода к реализации SDN. Один из них развивается консорциумом Open Network Foundation, другой развивается IETF. Подходы получили названия OpenFlow (OF) и PCE соответственно.

ONF определяет OF как интерфейс между уровнями управления и передачи данных. Весь потенциал SDN должен быть раскрыть благодаря протоколу, по сути переносящему таблицы (правила) действий с уровня управления на уровень передачи данных. OF не говорит, как организовать уровень управления, как принимать решения для создания тех самых правил и таблиц управления. OF открывает интерфейс между двумя функциональными уровнями в архитектуре SDN.

Интересно отметить, что IETF позиционирует PCE как эволюционное развитие сетевого управления, способного привести ко всем благам SDN. В своих докладах, посвященных сравнению с другими подходами, IETF отмечает, что OpenFlow открывает лишь возможности для создания программируемого уровня управления для коммутации потоков данных.

Попробуем понять, в чем суть подхода, предлагаемого IETF. PCE – это элемент архитектуры сети SDN, ответственный за вычисление маршрутов LSP. Элемент может быть реализован как роутер, часть OSS системы или как виртуальный элемент, поднятый в облачной среде.

Принцип работы подхода IETF заключается в том, что сетевой транспортный узел при прокладке маршрута обращается к PCE, тот, зная всю сетевую топологию и статус сети, вычисляет оптимальный маршрут. Вычислив маршрут PCE, возвращает узлу (Explicit Route Object) данные для LSP, которые передаются дальше по сети с помощью сигнального протокола RSVP.

На рис. 19.1 видно, что подходы по своей функциональной архитектуре реализуют главные принципы SDN: разделение уровней и функций передачи данных от управления.

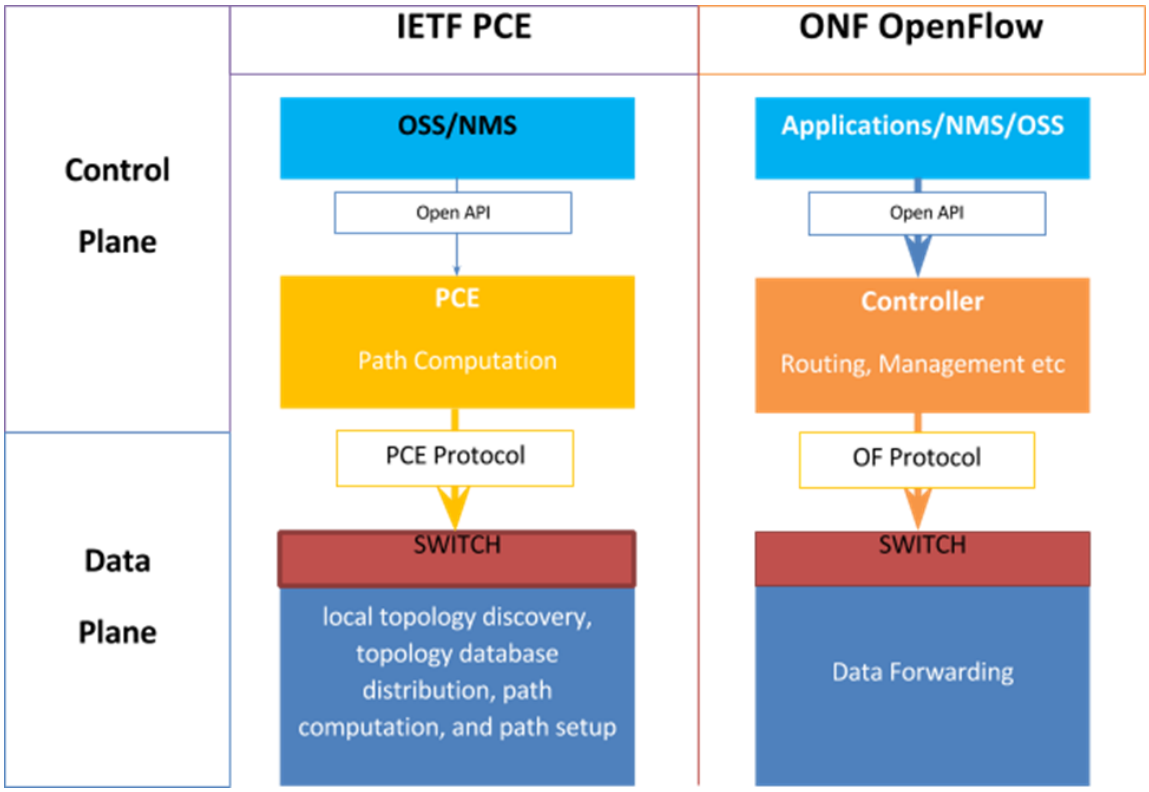

Рис. 19.1. Подходы IETF и OF

Очевидное отличие в том, что OF забирает все функции управления и задачи принятия решения с сетевых устройств и переносит на уровень управления. В то время как PCE отставляет часть функций принятия решений самим сетевым устройствам.

Управление в GMPLS распределено по сети и реализуется на каждом сетевом элементе независимо. Рассмотрим основные модули, которые формируют подсистему управления. Модуль HAL, указанный на рис. 19.2, использует SNMP для конфигурации сетевых элементов. С помощью OSPF-TE маршрутизатор GMPLS собирает данные о соединениях на сети и их атрибуты для базы данных traffic engineering (TE Database). С помощью модуля Path Computation реализуется расчет маршрута LSP, а RSVP-TE реализует резервирование ресурсов для LSP.

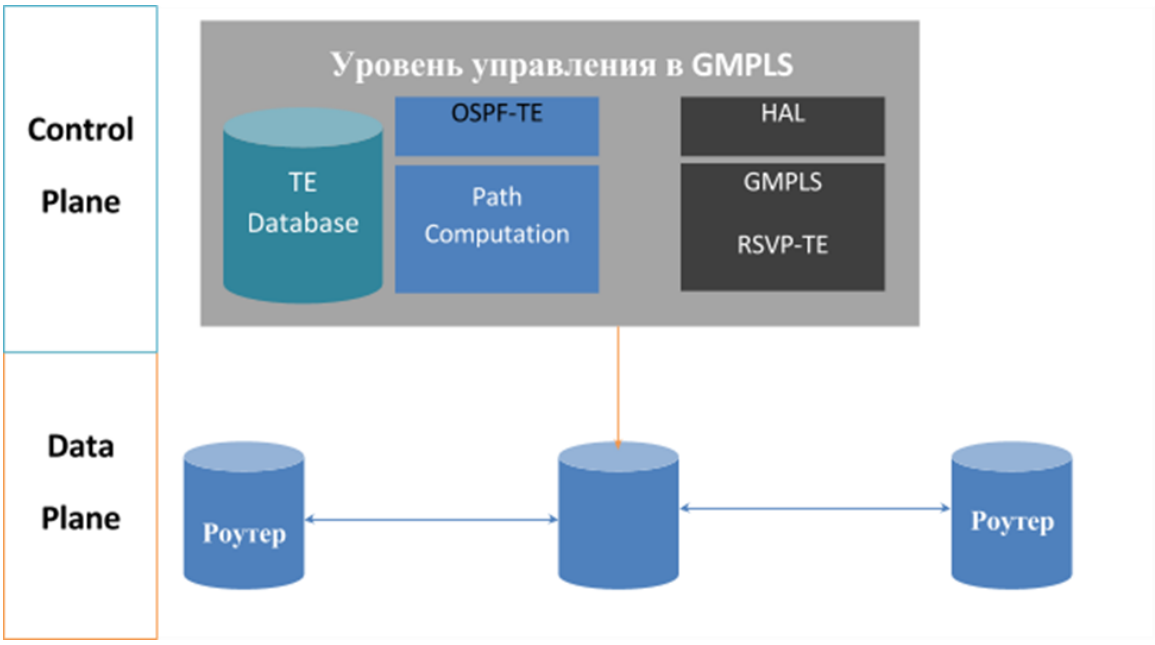

Рис. 19.2. Управление в GMPLS

Реализация единого уровня управления через протокол Open Flow будет иметь следующий вид (рис. 19.3). Все функции принятия решений по управлению вынесены на уровень централизованного контроллера и связаны с уровнем передачи данных через интерфейс OpenFlow. Уточним, что функции по управлению могут быть не только представлены в самом контроллере, но могут быть вынесены в отдельное приложение или компонент OSS.

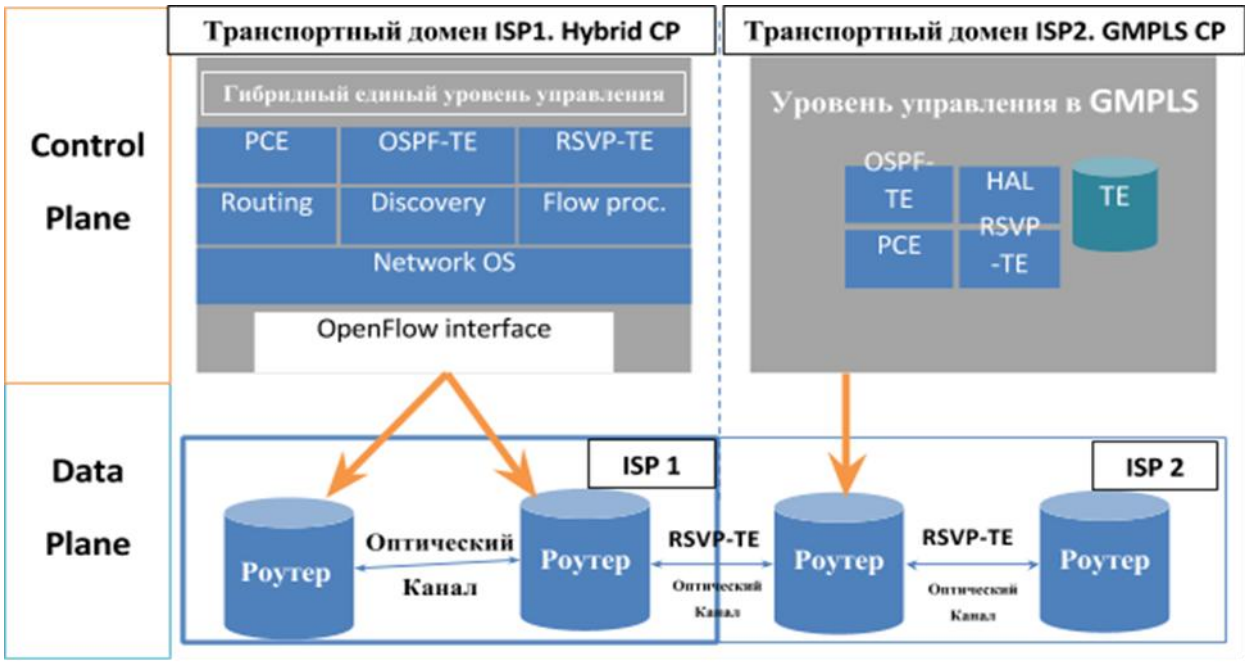

Рис. 19.3. Гибридная схема

Таким образом, OpenFlow разделяет уровень передачи данных от централизованного контроля. Функции *Discovery* собирают информацию о топологии сети, ее статусе, а также о связности сетевых элементов. Функция *Routing* выполняет задачи вычисления маршрутов для потоков данных. Отметим, что *Routing* в OpenFlow отличен от привычной маршрутизации пакетов: если маршрутизация пакетов оперирует заголовками полей протокола IP, то OpenFlow расширяет возможность анализа заголовков протоколов от канального до транспортного уровня, обеспечивая возможность управления потоками.

Flow processor – создает новые записи о потоках проходящих данных, формирует для них правила, которые дальше будут переданы через OpenFlow на сетевые элементы уровня передачи данных.

Сделаем промежуточный вывод о том, что GMPLS не подходит для создания единого уровня управления по причинам сложности и закрытости уровня управления и локализации интеллекта на сетевых элементах. Возможности развития новых функций, а следовательно, и услуг практически отсутствуют.

OpenFlow открывает новые возможности для создания приложений управления сетью и появления нового функционала. Вопрос заключается в том, получится ли создать приложения SDN для управления крупными транспортными сетями? Успех на рынке решений SDN пока завоевывают решения в области управления IP/Ethernet сетями кампусов, центров обработки данных и прочими локальными сетями.

Созданием коммерческого решения в области транспортных сетей пока не может похвастаться никто. Дополнительная сложность заключается в том, что так или иначе в игру за обладание такими технологиями подключаются ведущие разработчики сетевых решений. Однако их конечная цель – окупить свои инвестиции и ослабить позиции конкурентов в борьбе за долю рынка. Так, возможно, здесь оканчивается «открытость» SDN и начинается проприетарность OpenFlow.

Рассмотрим компромиссный вариант, обсуждаемый частью научно-инженерного комьюнити SDN. Попробуем сделать эволюционный вариант управления с помощью революционной архитектуры. Почему нельзя взять лучшее от OF и PCE?

Одной из ключевых особенностей PCE является возможность работать в разных доменах. Допустим у нас есть два провайдера один из них смог поставить решение OF для создания единого управления своим транспортным участком. А его партнер по транспортной сети ограниченно использует GMPLS без унифицированного управления. Очевидно, что здесь теряется автономность управления и создания LSP, так как используются разные подходы к управлению уровнем передачи данных.

Выходом является гибридное решение, позволяющее сохранить возможности OF и разрешить работать протоколам PCE и RSVP. В этом случае проблема «обратной совместимости» технологий пропадет сама собой.

Объединив функции от PCE и OpenFlow на уровне управления и сохранив возможность взаимодействия с GMPLS, мы получаем вариант более «жизнеспособный», чем узкоспециализированные варианты.

Именно объединение технологий, работающих на уровне Northbound, является основным ключом к дальнейшему пониманию того, как оператор сможет развивать свои системы поддержки эксплуатации в эру программно-конфигурируемых сетей.

Литература

Гольдштейн, А. Б. Технология и протоколы MPLS / А. Б. Гольдштейн, Б. С. Гольдштейн. – СПб. : BHV, 2005. – ISBN 5-8206-0126-2.

Содержание

Гольдштейн Александр Борисович
Никитин Алексей Владимирович
Шкрыль Александр Александрович

ТРАНСПОРТНЫЕ СЕТИ IP/MPLS

Технология и протоколы

Учебное пособие

Редактор *И. И. Щенсняк*
Компьютерная верстка *Н. А. Ефремовой*

План издания 2016 г., п. 31 а

Подписано к печати 19.04.2016
Объем 5,0 усл.-печ. л. Тираж 30 экз. Заказ 638
Редакционно-издательский отдел СПбГУТ
191186 СПб., наб. р. Мойки, 61

Отпечатано в СПбГУТ