

А.Б. Гольдштейн, Б.С. Гольдштейн

МРПС

ТЕХНОЛОГИЯ И

ПРОТОКОЛЫ

А. Б. Гольдштейн, Б. С. Гольдштейн

Технология и протоколы MPLS

Санкт-Петербург
«БХВ-Петербург»
2014

УДК 621.39
004.724
ББК 32.973.202
32.81
С59

Гольдштейн А. Б., Гольдштейн Б. С.
Технология и протоколы MPLS
СПб.: БХВ-Петербург, 2014. — 304 с.: ил.

ISBN 5-8206-0126-2

Технология MPLS представляет собой синтез всего самого лучшего из технологий уровня 2 (ATM, Frame Relay, Ethernet) и маршрутизации уровня 3 пакетных сетей. Изучение рассмотренных в книге протоколов MPLS позволит читателю оценить, является ли MPLS именно тем инструментом, с помощью которого сегодняшний хаос неуправляемой передачи пакетов по IP-сети может быть превращен в стройный, эффективно функционирующий механизм для сети связи следующего поколения NGN. На этот вопрос читатель сможет ответить сам, прочитав книгу, а также узнать про классы эквивалентности FEC, метки, протоколы LDP, CR-LDP, RSVP, RSVP-TE, OSPF, BGP-4, IS-IS, трафик-инжиниринг, GMPLS и многое другое.

Для технических специалистов, занятых разработкой и эксплуатацией сетей связи, студентов старших курсов и аспирантов соответствующих специальностей, для всех, кто интересуется современными инфокоммуникациями.

Научно-техническое издание

ISBN 5-8206-0126-2

© Гольдштейн А. Б., Гольдштейн Б. С., 2005

Содержание

Предисловие	9
Глава 1. Основы MPLS	13
1.1. Технология MPLS	13
1.2. Предыстория MPLS	15
1.3. Классы эквивалентности пересылки FEC	22
1.4. Коммутируемые по меткам тракты LSP	24
1.5. Основные понятия	28
Глава 2. Метки и функционирование MPLS	30
2.1. Коммутация по меткам	30
2.2. Структура метки	34
2.3. Стек меток MPLS	36
2.4. Инкапсуляция меток	38
2.5. Таблицы пересылки	43
2.6. Привязка «метка-FEC»	45
2.7. Режимы операций с метками	48
2.8. Фиксированные значения метки	52
Глава 3. Протокол LDP	54
3.1. Классы эквивалентности пересылки и LDP	54
3.2. Основы протокола LDP	56
3.3. Формат и параметры сообщений LDP	59
3.3.1. Блоки данных протокола LDP	59
3.3.2. Схема Type-Length-Value	60
3.3.3. Параметры TLV	61
3.3.4. Формат сообщений LDP	66
3.4. Сообщения LDP	68
3.4.1. Уведомляющее сообщение Notification Message	68
3.4.2. Приветственное сообщение Hello	69
3.4.3. Иницизирующее сообщение Initialization	71
3.4.4. Сообщение KeepAlive	73
3.4.5. Адресное сообщение Address	73
3.4.6. Сообщение отмены адреса Address Withdraw	73

3.4.7.	Сообщение привязки метки Label Mapping	74
3.4.8.	Сообщение запроса метки Label Request	75
3.4.9.	Сообщение отмены запроса метки Label Abort Request	76
3.4.10.	Сообщение отмены привязки метки Label Withdraw	77
3.4.11.	Сообщение освобождения метки Label Release	78
3.4.12.	Дополнительные сообщения и TLV	78
3.5.	Протокол CR-LDP	79
3.6.	Аспекты безопасности LDP	81
3.6.1.	Несанкционированные действия	81
3.6.2.	Конфиденциальность	82
3.6.3.	Отказ в обслуживании	82
3.7.	Сигнализация LDP	83

Глава 4. Протокол RSVP для MPLS 88

4.1.	Стили резервирования ресурсов для MPLS	88
4.2.	Основы протокола RSVP	92
4.3.	Роль RSVP и RSVP-TE в MPLS	95
4.4.	Расширение RSVP-TE	98
4.5.	Ремаршрутизация TE-туннелей	99
4.6.	Форматы RSVP-TE	101
4.6.1.	Сообщения создания LSP	101
4.6.2.	Объект LABEL	102
4.6.3.	Объект LABEL_REQUEST	102
4.6.4.	Объект EXPLICIT_ROUTE (ERO)	104
4.6.5.	Объект RECORD_ROUTE (RRO)	106
4.6.6.	Объект SESSION	109
4.6.7.	Объект SENDER_TEMPLATE	110
4.6.8.	Объект FILTER_SPEC	111
4.6.9.	Объект SESSION_ATTRIBUTE	111
4.7.	Расширение сообщения Hello	113
4.8.	Управление трафиком в MPLS	115

Глава 5. Протокол OSPF 116

5.1.	Протоколы OSPF и RIP	116
5.2.	Метрики OSPF	119
5.3.	Алгоритм Дейкстры	120
5.4.	Области OSPF	122
5.5.	Структура OSPF-пакета	125

5.6.	Типы пакетов OSPF.....	126
5.6.1.	Пакеты-приветствия	126
5.6.2.	Пакет описания базы данных OSPF	129
5.6.3.	Запросы сведений о состоянии каналов.....	130
5.6.4.	Корректировка сведений о состоянии каналов.....	131
5.6.5.	Подтверждение извещений о состоянии каналов	132
5.7.	Извещения LSA	132
5.8.	Базы данных OSPF.....	135
5.8.1.	База данных о смежности.....	135
5.8.2.	Топологическая карта сети	136
5.8.3.	Таблица маршрутизации	136
5.9.	Принципы работы OSPF.....	137
5.10.	SDL-диаграмма поведения маршрутизатора OSPF.....	139
Глава 6.	Протокол IS-IS.....	144
6.1.	Еще раз о маршрутизации по состоянию каналов.....	144
6.2.	Проблема flooding в протоколе IS-IS.....	146
6.3.	Метрики IS-IS	148
6.4.	Адресация IS-IS.....	148
6.5.	Маршрутизация IS-IS	149
6.6.	Пакеты IS-IS	153
6.6.1.	Пакеты-приветствия Hello	154
6.6.2.	Пакеты состояния каналов LSP.....	157
6.6.3.	Пакеты порядкового номера SNP	159
Глава 7.	Протокол маршрутизации BGP.....	161
7.1.	Использование протокола BGP в MPLS.....	161
7.2.	Алгоритм Беллмана-Форда	163
7.3.	Нумерация автономных систем в BGP	164
7.4.	Маршрутизаторы BGP	165
7.5.	Протокол EBGP.....	167
7.6.	Протокол IBGP.....	168
7.7.	Конфедерации BGP	169
7.8.	Карты маршрутов	170
7.9.	Метрики маршрутов	171
7.10.	База данных маршрутизации.....	172

7.11. Сообщения BGP	172
7.11.1. Общий заголовок	172
7.11.2. Запрос соединения OPEN	173
7.11.3. Сообщение об обновлении UPDATE.....	174
7.11.4. Уведомление NOTIFICATION	175
7.11.5. Сообщение подтверждения связи Keepalive.....	176
7.12. Синхронизация BGP	177
7.13. Многопротокольные расширения BGP	177

Глава 8. Виртуальные частные сети и туннели 182

8.1. Виртуальные частные сети VPN	182
8.2. Туннелирование в MPLS	185
8.3. Виртуальные частные MPLS-сети	190
8.3.1. Общие предпосылки MPLS-VPN	190
8.3.2. Сети MPLS/BGP-VPN.....	192
8.3.3. Виртуальная сеть на базе IP/MPLS	192
8.3.4. Организация MPLS-VPN	193
8.4. Маршрутизация MPLS-VPN	193
8.4.1. Таблицы маршрутизации в PE-маршрутизаторах	193
8.4.2. Распространение маршрутной информации по протоколу BGP.....	194
8.5. Распространение маршрутной информации	196
8.5.1. Атрибут целевой VPN	196
8.5.2. Атрибут VPN-источник.....	197
8.5.3. Атрибут сайт-источник	197
8.5.4. Передача маршрутной информации между PE	197
8.6. Пересылка данных по магистральной сети	198
8.7. Передача маршрутной информации между CE и PE.....	199
8.8. Поддержка MPLS маршрутизатором CE	201
8.8.1. Виртуальные сайты	201
8.8.2. VPN Интернет-провайдера.....	201
8.9. Стандартизация технологии MPLS-VPN.....	201
8.10. Сценарии организации VPN на основе туннелей MPLS	202

Глава 9. Инжиниринг трафика 210

9.1. Концепция инжиниринга трафика в MPLS.....	210
9.2. Протоколы сигнализации	214

9.3.	Атрибуты потоков трафика и сетевых ресурсов	215
9.3.1.	Атрибуты объединенных потоков трафика	216
9.3.2.	Атрибуты сетевых ресурсов	218
9.4.	Маршрутизация на основе ограничений	218
9.5.	Механизмы TE в MPLS	219
9.6.	Сравнение протоколов CR-LDP и RSVP-TE.....	224
9.6.1.	Сравнение функциональных возможностей	224
9.6.2.	Сравнение технических характеристик.....	225
9.6.3.	Служебный трафик в RSVP-TE и CR-LDP.....	231
9.6.4.	Сравнение ремаршрутизации в RSVP-TE и CR-LDP	234
9.6.5.	Сравнение протоколов по их внедрению	237
Глава 10. Эволюция к GMPLS.....		239
10.1.	MPLambS и GMPLS	239
10.2.	Метки в GMPLS.....	244
10.2.1.	Запрос универсальной метки	244
10.2.2.	Универсальная метка	247
10.2.3.	Коммутация диапазонов волн	248
10.2.4.	Предлагаемая метка	249
10.2.5.	Набор меток.....	249
10.3.	Двунаправленные LSP	250
10.4.	Уведомления и сообщения об ошибках	253
10.4.1.	Запрос уведомления	253
10.4.2.	Уведомляющее сообщение	254
10.4.3.	Разрушение процесса пересылки сообщением об ошибке PathErr	254
10.5.	Метки для явно заданного маршрута.....	255
10.6.	Информация о защите.....	256
10.7.	Информация об административном статусе.....	257
10.8.	Разделение общего тракта управления	258
10.8.1.	Идентификация интерфейсов	258
10.8.2.	Обработка ошибок	260
10.9.	Форматы сообщений.....	262
10.10.	Что дальше?	263
Глава 11. Камо Грядеши?		265
11.1.	Внедрение и перспективы MPLS.....	265
11.2.	MIB и MPLS.....	266

11.3. Оборудование MPLS.....	269
11.4. Тестирование MPLS.....	272
11.5. VoMPLS	275
11.6. «Все» через MPLS.....	277
11.7. Многоадресность.....	279
11.8. DiffServ-aware MPLS-TE	280
11.8.1. Объединение технологий	280
11.8.2. DiffServ в MPLS.....	281
11.8.3. Class of Type — СТ	284
11.8.4. Вычисление пути.....	285
11.8.5. Сигнализация тракта.....	286
11.8.6. Модели назначения полосы пропускания	287
11.9. MPLS и QoS	290
Литература	293
Глоссарий.....	298
Предметный указатель.....	302

Предисловие

«Когда я назначаю кого-то на лидирующую позицию, то имею девяносто девять недовольных и одного неблагодарного», говорил Людовик XIV. Существование, как минимум, девяносто девяти претендентов на позицию главного механизма обеспечения качества обслуживания QoS (Quality of Service) в сетях связи следующего поколения NGN (Next Generation Network) обусловлено лавинообразным ростом числа пользователей IP-сетей и соответствующим увеличением мультимедийного трафика. Изначально же IP-технологии были ориентированы на передачу данных простейших приложений, для чего было вполне достаточно программных маршрутизаторов. По мере появления новых мультимедийных приложений типа IP-телефонии [20], требующих более высоких скоростей передачи и поддержки более широкой полосы пропускания, возникла потребность в создании технологий и устройств, обеспечивающих возможность быстрой коммутации на уровне 2 (уровне звена данных) и уровне 3 (сетевом уровне) аппаратными средствами. Появились устройства коммутации на уровне 2, решающие проблему узких мест коммутации в среде LAN, а также новые маршрутизаторы, улучшающие ситуацию с маршрутизацией на уровне 3 путем перевода в быстродействующую аппаратную реализацию процедур просмотра маршрутных таблиц для пересылки пакетов. Достигли своего апогея наиболее перспективные технологии 90-х годов Frame Relay и ATM. Появились разнообразные средства обеспечения качества обслуживания DiffServ и IntServ, рассматриваемый в этой книге протокол резервирования RSVP и многое другое.

И все же, все эти девяносто девять претендентов на наиболее эффективное решение проблемы сетевого QoS при передаче мультимедийной информации в реальном времени с учетом таких показателей, как задержка, дрожание фазы, перегрузка и т.п. фактически уже уступили занявшей лидирующую позицию многопротокольной коммутации по меткам MPLS (MultiProtocol Label Switching).

MPLS является весьма изящным и универсальным решением проблем QoS, стоящих перед сегодняшними пакетными сетями, решением, которое обеспечивает скорость передачи, масштабируемость, оптимизацию распределения трафика и эффективную маршрутизацию (на основе показателей QoS) в пакетных сетях IP, ATM и Frame Relay. Лидерство MPLS обусловлено, по мнению авторов, удачно выбранной позицией, позволяющей оптимальным образом отображать сквозной трафик третьего уровня от исходящего сетевого узла (маршрутизатора) к входящему узлу в трафик между соседними узлами на втором уровне сетевой иерархии. Т.о., MPLS, являясь гибридом уровней 2 и 3 семиуровневой модели OSI, собрала вместе лучшее из двух миров: уровня 2 и уровня 3, мира ATM и мира IP.

Некоторая «неблагодарность» в ответ на столь значительное внимание к технологии MPLS проявилась в возникшей путанице относительно того, что такое MPLS, к какому уровню OSI она относится, что эта технология может и для чего она предназначена. Такая путаница обусловлена, в частности, тем, что технология MPLS не использует какой-либо единый формат для транспорта пакетов, а базируется на нескольких протоколах (и соответствующих им стандартах), каждый из которых решает отдельную проблему. Попытка распутать этот клубок и каким-то образом структурировать описания протоколов MPLS как раз и предпринята в этой книге и, прежде всего, в ее первой главе.

Вторая глава посвящена собственно меткам. Сами по себе, метки не являются чем-то новым. Технологии X.25, ATM, Frame Relay и, до некоторой степени, TDM используют инкапсуляцию с помощью меток в течение многих лет. Хороший пример такой инкапсуляции в TDM при передаче сообщений протокола DSS1 ISDN через универсальный интерфейс сети доступа V5.2 был рассмотрен в одной из предыдущих книг этой серии, посвященной протоколам сети доступа [21]. Но только в MPLS эта концепция реализуется в общем виде, т.е. метки не привязаны к какому-либо конкретному протоколу второго уровня. Метка MPLS представляет собой число, уникальным образом идентифицирующее некоторую совокупность передаваемых пакетов. Метка имеет только локальное значение, т. е. по мере следования пакетов вдоль маршрута ее необходимо изменять. О самих метках, о распределении меток, о коммутации по меткам, о соответствующих протоколах сигнализации и маршрутизации, о методах инжиниринга трафика, о VPN, о туннелях MPLS и о многом другом рассказывается в этой книге.

Книга построена таким образом, что в ней можно условно выделить три части. В первых главах представлено описание технологии MPLS и рассматриваются:

- принципы MPLS, сильные и слабые стороны этой технологии,
- краткая история MPLS, основы архитектуры,
- классы эквивалентности пересылки FEC и коммутируемые по меткам тракты LSP,
- структура метки, методы распределения меток и коммутации по меткам.

Задача этих глав — ввести ключевые определения и термины, относящиеся к MPLS, и объяснить технологии, в результате слияния которых и была создана MPLS. Сюда входит рассмотрение технологий маршрутизации и коммутации, ранних исследований в области коммутации ячеек и тегов, а также других важных технологий, которые являются ключом к пониманию эволюции технологии MPLS.

Задача следующей, наиболее объемистой части книги — объяснить базовую технологию и протоколы MPLS, включая протоколы сигнализации, механизмы распределения меток и коммутации по меткам, и содержит следующие главы:

- протокол LDP и распределение меток,
- протокол RSVP и идеи инжиниринга трафика,
- протокол OSPF и метрики маршрутизации,
- протокол BGP и пограничные шлюзы,
- протокол IS-IS и внутризонавая маршрутизация.

Уже простое перечисление глав показывает, что эта книга могла бы быть издана в серии «Телекоммуникационные протоколы», как первоначально и планировалось. Но книга, которая получилась и которую вы держите в руках, существенно отличается от вышедших в серии «Телекоммуникационные протоколы» книг, в первую очередь, тем, что отнюдь не является справочником. Это вполне объяснимо: справочники по телекоммуникационным протоколам написаны для гораздо более «зрелых» технологий ОКС7 [22,24], V5 [23], R1.5 [25]. В данной же книге представлен не столько справочник, сколько, в определенном смысле, путеводитель по технологии MPLS, чрезвычайно молодой, но, тем не менее, уже заслужившей место в «Энциклопедии современных инфокоммуникаций».

Это изменение подхода к книге обусловило и то, что в третьей ее части обсуждаются направления текущих и будущих работ в области технологии и протоколов MPLS, роль MPLS в конвергенции сетей и услуг связи, перспективные вопросы MPLS:

- инжиниринг трафика в MPLS,
- туннели MPLS и виртуальные частные сети VPN,
- MPLambdaS и GMPLS,
- перспективы технологии MPLS.

Заключительные главы книги также требуют краткого предварительного комментария. Первоначально MPLS рассматривалась как технология, которая сумела бы радикально улучшить маршрутизацию IP-пакетов на уровне 3 модели OSI. До самого последнего времени главным был тот аргумент, что благодаря применению MPLS значительно возрастает производительность сети, т.к. анализ коротких меток фиксированной длины выполняется намного быстрее, чем анализ длинных IP-заголовков сетевого уровня при традиционной маршрутизации IP-пакетов. Однако этот аргумент уже сегодня становится менее актуальным в связи с появлением аппаратных решений на интегральных схемах прикладной ориентации ASIC, на программируемых в процессе эксплуатации матрицах FPGA и т.п. Анализ заголовков IP-пакетов в самом ближайшем будущем не будет являться узким местом для производительности терабитовых (а вскоре — и петабитовых!) маршрутизаторов традиционных IP-сетей.

С другой стороны, на первый план вышла проблематика качества обслуживания QoS, возрастает интерес к новым приложениям, таким как использование MPLS в оптических сетях коммутации, инжиниринг трафика TE, речь поверх MPLS, виртуальные частные сети VPN, что и обусловило как подбор материала, так и структуру третьей части книги.

Список литературы весьма обширен и содержит 125 наименований. И даже в этот список вошло далеко не все, что уже написано про MPLS, но поскольку перечисленные источники сами содержат литературные ссылки, наш перечень будет нетрудно расширить. Представляются также полезными предметный указатель и список использованных в книге аббревиатур, составляющий основу глоссария.

Авторы пользуются случаем выразить свою признательность профессору Г.Г. Яновскому, впервые обратившему внимание авторов на эту проблематику, причем тогда, когда вышеприведенные предложения о перспективности MPLS были отнюдь не так очевидны как сегодня. Кроме того, книга была бы заметно хуже, если бы не научное редактирование, выполненное В.А. Соколовым. Весьма полезными были и стимулирующие дискуссии с коллегами из ЛОНИИС, СПбГУТ, ГК Экран, НТЦ Протей, Уралсвязьинформ, МТУ-информ/Комстар, Ленсвязь, ЮТК, АДЭ и др. Значительная помощь в работе над книгой была оказана студентами старших курсов СПбГУТ им. проф. М.А. Бонч-Бруевича — А. Атциком и его коллегами, материалы дипломных проектов которых приведены в качестве некоторых сценариев и примеров в ряде глав книги.

Все это, безусловно, содействовало решению главной задачи, которую ставили перед собой авторы: показать, что многопротокольная коммутация по меткам — это именно тот инструмент, с помощью которого сегодняшний хаос неуправляемой передачи пакетов по IP-сетям может быть превращен в стройный, эффективно функционирующий механизм для NGN.

Глава 1

Основы MPLS

1.1. Технология MPLS

Технология *многопротокольной коммутации по меткам MPLS* явилась результатом слияния нескольких сходных технологий, которые были изобретены в середине 1990-х годов. Наиболее известная из них (хотя и не первая, увидевшая широкий свет) была названа ее изобретателями — компанией Ipsilon — *IP Switching*. Панее 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 Интернет структуру управления, т.е. IP-адреса и стандартные протоколы маршрутизации, например OSPF и BGP. Мы кратко рассмотрим эти технологии в следующем параграфе.

Возрастающий интерес к коммутации с использованием меток привел к созданию специальной рабочей группы *IETF*, целью которой было выработать на основе вышеупомянутых механизмов общий стандарт. Не желая давать группе название, которое соответствовало бы продукту какой-либо одной компании, IETF остановила свой выбор на нейтральном (правда, слегка громоздком) названии: *многопротокольная коммутация по меткам*.

Можно встретить другие переводы словосочетания *label switching*: «коммутация меток», «коммутация на базе (или, что то же самое, — на основе) меток». Первый из этих переводов просто неверен — ведь коммутируются не метки, а пакеты. Второй несколько искажает действительное положение вещей — базу (или основу) коммутации составляют далеко не только метки.

Перед пересылкой принятого пакета на следующий участок маршрута устройство коммутации по меткам обычно вставляет в этот пакет, некоторую новую метку вместо той, которая в нем содержалась. Эта операция называется *заменой меток (label swapping)*. Для того чтобы определить, куда пересылать пакеты,

устройство коммутации по меткам, которое называется *маршрутизатором LSR (Label Switching Router)*, использует стандартные протоколы управления IP-сетью.

Многопротокольной (Multi-Protocol) коммутацией MPLS называется потому, что ее средства применимы к *любому* протоколу сетевого уровня, т.е. MPLS — это своего рода инкапсулирующий протокол, способный транспортировать информацию множества протоколов низших уровней модели OSI, как это показано на рис. 1.1.

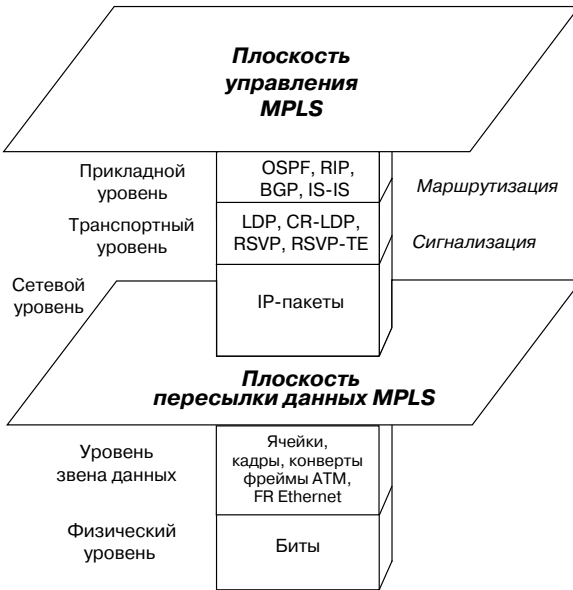


Рис. 1.1. Плоскости MPLS

Напомним, что первый, *физический уровень* (physical layer) содержит функции, обеспечивающие использование физической среды для двусторонней *передачи битов* (с такой достоверностью, какую обеспечивает эта среда) по прямому тракту, связывающему два узла сети. Второй уровень — *уровень звена данных* (data link layer) — содержит функции, обеспечивающие формирование в этом тракте надежного логического звена связи, по которому происходит двусторонний обмен *информационными блоками* между названными узлами; при этом путем обнаружения и исправления ошибок гарантируется заданная достоверность передачи. Третий, *сетевой уровень* содержит функции, обеспечивающие транспортировку информационных блоков от отправителя к получателю через несколько узлов сети по подходящему маршруту транспортировки, который составляется из звеньев второго уровня. Общая идея протоколов всех уровней (кроме физического) состоит в том, что информационный блок каждого уровня содержит заголовок и информационное поле, и в том, что блок протокола вышележащего

уровня помещается в информационное поле блока протокола расположенного сразу под ним нижележащего уровня.

Представленная на рис. 1.1 плоскость пересылки данных MPLS не образует полноценного уровня, она «вклинивается» в сетях IP, ATM или Frame Relay между 2-м и 3-м уровнями модели OSI, оставаясь независимой от этих уровней. Можно сказать, что одновременное функционирование MPLS на сетевом уровне и на уровне звена данных приводит к образованию так называемого *уровня 2.5*, где, собственно, и выполняется коммутация по меткам.

1.2. Предыстория MPLS

Теперь, когда, более или менее, перечислены мотивы, побудившие создателей различных технологий коммутации по меткам взяться в свое время за эту работу, имеет смысл обратиться к последовательности событий, которые привели, в конечном итоге, к сегодняшней MPLS.

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 годов. 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 1993/94 годов 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, принадлежащей 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 принадлежащих разным подсетям 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 чисто 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 определения 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. Дальнейшая история соответствует высказыванию, которое приписывают Альберту Эйнштейну: «Все давно знают, что то-то и то-то совершенно невозможно. Но вот находится невежда, который этого не знает, и он-то и совершает открытие».

В роли такого «невежды» выступила компания *Toshiba*, практически впервые подвергнув сомнению этот принцип и в 1994 году анонсировавшая *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 году, а немного позже, в начале 1995 года, — на технической сессии BOF комитета IETF, однако интерес к этой проблеме тогда был довольно низким.

Компании Ipsilon (в настоящее время — подразделение фирмы 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 параграфе 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 года 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, по мнению авторов, 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. Кроме более быстрого поиска адреса, новые маршрутизаторы могли обслуживать вызовы по-разному в зависимости от требуемого

качества обслуживания (при передаче речи, видео и изображений). К тому же, 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.

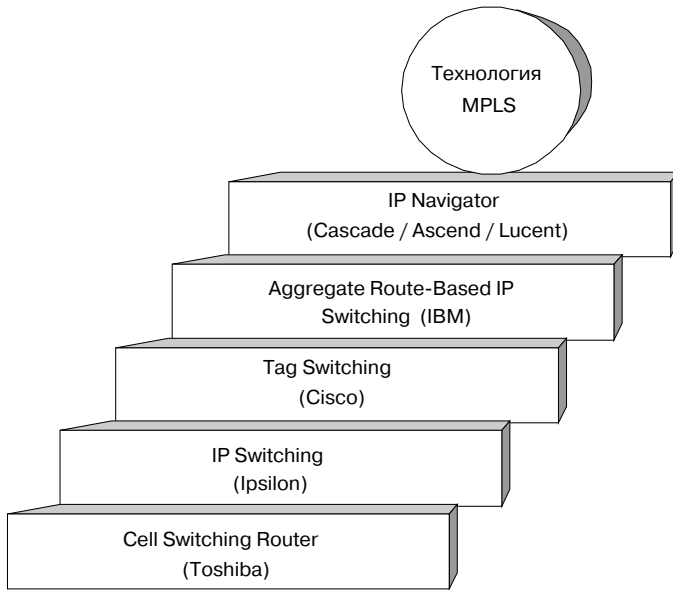


Рис. 1.2. Эволюция технологии

После публикации первой серии проектов стандартов Tag Switching 9-13 декабря 1996 года 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 года в Мемфисе, Тенниси, 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. Появившиеся вслед за этим документы RFC по технологии и протоколам MPLS приведены в табл. 1.1.

Таблица 1.1. 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 применяться 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

Архитектура MPLS специфицирована в документе RFC 3031 «Multiprotocol Label Switching Architecture». Сегодня вопросами MPLS продолжают заниматься рабочие группы IETF (Routing Area Working Group — рабочая группа по маршрутизации — и MPLS Working Group — рабочая группа по MPLS) и в ATM Forum (Traffic Management Working Group — рабочая группа по управлению трафиком — и ATM-IP Collaboration Working Group — рабочая группа по совместной работе сетей ATM и IP). Основные идеи и результаты работы этих групп уже упоминались в предыдущем параграфе и будут рассмотрены в книге далее.

1.3. Классы эквивалентности пересылки FEC

Рабочими группами, упомянутыми в предыдущем параграфе, 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 протокола уровня 3, 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, то есть 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 же 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 создаст 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.

Понятия «метка» и «база данных LIB» будут рассмотрены более подробно в главе 2, а сейчас важно отметить, что значение метки, как правило, изменяется по мере продвижения пакета по сети.

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)*, который указывает следующий участок, операцию, которая должна быть выполнена со стеком меток (стек меток подробнее будет также рассмотрен в главе 2) 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.

1.4. Коммутируемые по меткам тракты LSP

При рассмотрении классов FEC в параграфе 1.3 отмечалось, что 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 от программы*), либо при обнаружении определенного потока данных (*управляемые данными*).

Метки в LSP назначаются с помощью *протокола распределения меток LDP (Label Distribution Protocol)*, рассматриваемого в главе 3, причем существуют разные способы такого распределения на основе данных вспомогательных протоколов, в частности, рассмат-

риваемого в главе 4 протокола RSVP-TE. Подготавливают процесс распределения меток протоколы маршрутизации, такие как OSPF, IS-IS или BGP, рассматриваемые в главах 5, 6 и 7, соответственно. С помощью этих протоколов маршрутизации создается «дерево» 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 LSP 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 запрашивают 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 заверченный (и, в определенном смысле, автономный) домен сети MPLS, изображенный на рис. 1.3. 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. 1.3 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. Они, разумеется, не показаны в примере на рис. 1.3, где изображены только три внутренних маршрутизатора LSR2, LSR3 и LSR4.

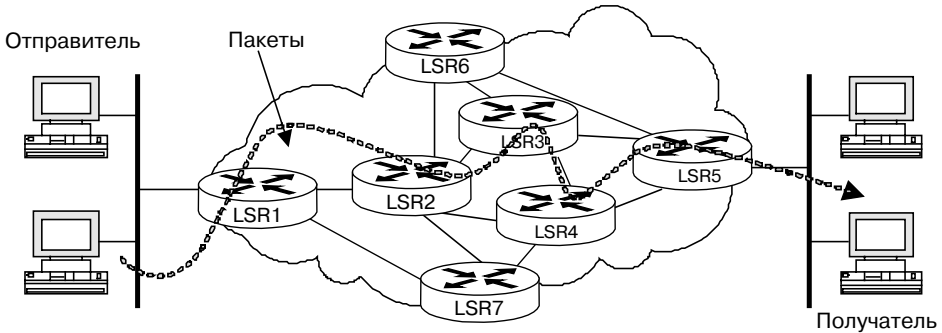


Рис. 1.3. Пример домена MPLS-сети

Напомним, что на рис. 1.3 изображен лишь упрощенный домен MPLS-сети. Пакеты, поступающие в него, могут приходиться как непосредственно от отправителей (что показано на рис. 1.3), так и из смежной сети, которая может быть MPLS-сетью более высокого уровня (то есть 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 LSR5, 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. Пграничные 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 при-

нимать 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 обсуждению разделения пакетов с IP-заголовками и пакетов с метками в сети MPLS мы вернемся в главе 11, где попробуем проанализировать сегодняшние решения производителей для сетей MPLS.

К выходному узлу LSR5 (рис. 1.3) поступают потоки пакетов от нескольких входных узлов (от LSR1, LSR6 и LSR7). В промежуточных маршрутизаторах некоторые из этих потоков могут «сливаться», то есть 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 рисунке MPLS-сети существует четыре дерева такого рода, которые вместе содержат $4 \times (4 - 1) = 12$ трактов LSP. Ясно, что через один промежуточный маршрутизатор LSR может проходить несколько LSP, в том числе, LSP, принадлежащих разным деревьям. Если учесть, к тому же, 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. В главе 3, посвященной протоколу LDP, достаточно внимательно рассматриваются средства борьбы с такого рода петлями. К рис. 1.3 мы еще вернемся в следующих главах книги, а сейчас, в заключительном параграфе главы, попробуем свести воедино основные рассмотренные в ней понятия.

1.5. Основные понятия

Итак, 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 используются рассмотренные выше коммутируемые по меткам тракты LSP, которые были организованы с помощью рассматриваемых в следующих главах протоколов маршрутизации и сигнализации уровня 3. Основные специальные термины MPLS сведены в табл. 1.2.

Таблица 1.2. Основные термины MPLS

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	Приходящий через один или более LSR тракт, по которому следуют пакеты одного и того же FEC
ER — LSP — (explicitly routed LSP) — LSP с явно заданным маршрутом	Тракт LSP, который организован 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

В дополнение к приведенным в таблице рассмотрим еще некоторые базовые понятия, уже упоминавшиеся выше и весьма важные для понимания принципов работы технологии MPLS, такие как *пересылка, коммутация и маршрутизация*.

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, как это показано на рис. 1.1.

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.