

IP网络技术丛书

多协议标签交换 技术与应用

(美) Bruce Davie Yakov Rekhter 著

罗志祥 朱志实 黄本雄 李非 译



机械工业出版社
China Machine Press

本书介绍多协议标签交换(MPLS)技术的来龙去脉、体系结构、应用领域,内容包括算法和IP控制协议的使用、Ipsilon公司IP交换、Cisco公司的标志交换、MPLS的核心协议(LDP)、QoS的主要方法,以及使用MPLS建立VPN的方法等,内容丰富、图文并茂,是MPLS的完整参考手册,可供网络设计与开发人员参考。

Bruce Davie & Yakov Rekhter: MPLS: Technology and Applications.

Copyright © 2000 by Academic Press.

Translation Copyright © 2001 by China Machine Press. All rights reserved.

本书中文简体字版由美国Morgan Kaufmann公司授权机械工业出版社独家出版。未经出版者书面许可,不得以任何方式复制或抄袭本书内容。

版权所有,侵权必究。

本书版权登记号:图字:01-2000-2810

图书在版编目(CIP)数据

多协议标签交换技术与应用/(美)大卫(Davie, B.), (美)瑞格特(Rekhter, Y.)
著;罗志祥等译. - 北京:机械工业出版社, 2001.1

(IP网络技术丛书)

书名原文: MPLS: Technology and Applications

ISBN 7-111-08666-X

I. 多… II. ①大… ②瑞… ③罗… III. 宽带通信系统-计算机通信网-通信技术
IV. TN915

中国版本图书馆CIP数据核字(2000)第87891号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)

责任编辑:王云

北京市密云县印刷厂印刷·新华书店北京发行所发行

2001年1月第1版第1次印刷

787mm×1092mm 1/16·12.5印张

印数:0 001-5 000册

定价:28.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换

前 言

作为因特网中的技术革新飞速发展的缩影，“网络时代”的概念到目前为止已经广为人知了。多协议标签交换技术——MPLS是网络时代发展最迅速的技术。两年前，当我们写作关于标签交换技术的第一本书的时候，我们只描述了MPLS未来的可能方向，因为那时候MPLS还处于定义和标准化的早期阶段。我们甚至不敢将MPLS的术语置于封面上，因为担心很少有人知道MPLS是什么。

MPLS现在已经成为因特网中的一个根本性的重要技术，并且现在出版关于这个主题的书显然是适时的（如果没有迟到的话）。一些最大的因特网服务提供商已经在它们的网络中采用MPLS来解决一些问题，如流量工程，并在ATM骨干网络上提供IP服务。已经出现了在MPLS基础之上的虚拟专用网络（VPN），并且这些服务的第一批用户已经热心地采用了这个新技术。现在大多数高端路由器可以无缝支持MPLS，并论证了其互操作性能。

尽管MPLS取得了这些成功，还是有人对MPLS不太了解。有些人甚至错误地认为标签交换的主要吸引力是它可以有效地实现并因此提高了转发性能；还有些人则把转发性能的提高归功于高速的IP查找算法和硬件的成功，从而争论说MPLS根本并没有提供任何价值。在本书中，我们正确地解释了MPLS是什么以及MPLS是如何工作的，并且详细介绍了MPLS的一些优点。虽然这些优点并没有暗示应当在这个世界的每个网络中都采用MPLS，但这些绝对重要性引起了绝大多数的大型网络操作者以及那些必须从服务提供商处购买网络服务的人员的深思。

我们力图使本书深入浅出并且为大众接受。我们假定绝大部分的读者已对IP路由具备一定程度的了解，但是我们还是在适当的地方提供了关于该领域的一些背景知识。更高层次的读者将可以跳过那些对路由的标准进行讨论的部分。我们在最后两章中讨论了一些相当高深的路由功能，对路由没有很强的背景知识的读者可不必阅读这些内容。同样，本书关于服务质量（QoS）的讨论，对于那些寻找关于这个主题的全面指导的读者来说可能太简短了，然而某些更富有经验的读者可跳过关于这个主题的概述。

我们希望本书对于服务提供商和企业这两种环境下的网络设计者和工程师来说很有价值，网络设计者和工程师需要充分了解MPLS，以便决定在他们的网络中MPLS是否是一个可能的候选方案。本书将提供大量关于标签交换与其他技术（比如，传统的路由和第二层交换）相比的知识，并针对另外一个技术来权衡不同的选项。这将使得你对MPLS将在你的网络中扮演什么样的角色（如果可能的话）作出一个有见地的判断。这本书也将为开发MPLS产品的工程师提供有关此领域的一些有用的知识。这里的信息将使得对那些可以从因特网草案中找到的详细的技术资料、请求注释文档（RFC）和由各个MPLS厂家提供的文献的理解更加容易。在那些

标准还没有完成的领域，你也许会产生顾虑：是采用现在建议的方法，还是等将来标准定下来再说。

关于本书的结构

本书的章节分成三部分：

- 介绍与概述：第1章和第2章。
- 两个先前技术的细节：第3章和第4章。
- MPLS协议与应用：第5章到第8章。

在第1章中我们论述了发明通常的标签交换以及特别的MPLS的必要性。我们介绍了标签交换技术旨在解决的许多问题及其简要历史。第2章描述了标签交换整个领域的总体结构问题。在MPLS和它之前的一些标签交换技术之间存在某些基本的相似之处，例如转发算法和IP控制协议的使用。另外，任何一种标签交换方法的设计者都必须作出一些相同的、关键结构的选择，例如，在控制驱动和数据驱动的标签分配之间进行选择。我们论述了这些选择的分歧。

第3章和第4章对作为MPLS产物先导的两种最重要的技术提供了详细的描述。这其中的第一种技术是Ipsilon公司的IP交换，它是使得标签交换在网络工业中出名的最大的原因。第二种技术是Cisco公司的标志交换，MPLS的许多基本的概念起源于标志交换。通过详细分析这两种技术，我们看到存在许多设计选择，并使形成MPLS设计的一些因素更清晰。

在第5章中我们通过论述MPLS体系结构的基础以及一些核心协议，特别是标签分布协议（LDP）来开始对MPLS的详细描述。我们在第6章描述了MPLS在支持服务质量中的角色，并解释了因特网中的两种主要的QoS方法（集成服务和差分服务）在一个MPLS网络中如何支持。第7章解释了MPLS对基于约束路由的应用，并论述了可能被用于这个应用中的协议范围。

最后，第8章描述了使用MPLS建立VPN的一种方法。本书中，我们讨论了不同方法以及各种各样的设计判断。我们尽可能地在分析中保持客观的、公正的态度。因为我们参与了某些协议的设计和标准化进程，所以对某些设计选择的偏爱是不可避免的。然而，我们已经尽力客观地介绍不同的方法，避免“过度吹嘘”MPLS，使得你可以得出自己的结论。

目 录

前言

第1章 绪论1

1.1 MPLS的出现原因2

1.1.1 因特网的增长与演化2

1.1.2 价格与性能3

1.1.3 ATM上的IP集成5

1.1.4 扩展路由功能9

1.2 MPLS的历史回顾11

1.2.1 ATM上的IP11

1.2.2 东芝公司的信元交换路由器13

1.2.3 IP交换13

1.2.4 标志交换14

1.2.5 IBM公司的ARIS14

1.2.6 多协议标签交换工作组14

1.3 小结15

1.4 参考信息16

第2章 基本概念17

2.1 网络层路由的功能器件17

2.1.1 转发等价类18

2.1.2 提供一致的路由19

2.2 标签交换的转发功能器件19

2.2.1 什么是标签19

2.2.2 标签交换的转发表20

2.2.3 在分组中携带标签20

2.2.4 标签交换的转发算法21

2.2.5 单一的转发算法22

2.2.6 转发粒度23

2.2.7 多协议：上层多协议与下层多协议23

2.2.8 标签交换转发功能器件：小结24

2.3 标签交换的控制功能器件24

2.3.1 本地绑定与远程绑定25

2.3.2 上游绑定与下游绑定26

2.3.3 “自由的”标签27

2.3.4 建立与撤消标签绑定：控制驱动与数据驱动标签绑定27

2.3.5 分布标签绑定信息：选项是什么33

2.3.6 组播考虑35

2.3.7 处理路由瞬态36

2.4 边缘设备37

2.5 标签交换和网络层寻址与路由的关系38

2.6 小结38

2.7 参考信息39

第3章 IP交换40

3.1 IP交换概述41

3.2 Ipsilon流管理协议45

3.2.1 IFMP的邻接协议46

3.2.2 IFMP的重定向协议46

3.2.3 重定向流的封装49

3.2.4 IFMP与安全52

3.2.5 IFMP和TTL52

3.3 通用交换机管理协议55

3.3.1 GSMP邻接协议56

3.3.2 GSMP连接管理协议57

3.4 实现59

3.5 小结59

3.6 参考信息60

第4章 标志交换61

4.1 概述61

4.1.1 支持基于目的地的路由62

4.1.2 通过路由知识体系结构来提高路由可扩展性67

4.1.3 组播71

4.1.4 使用标志交换的RSVP74

4.1.5 显式路由75

4.2 ATM上的标志交换	76	第7章 基于约束的路由	124
4.2.1 携带标志信息	76	7.1 概述	124
4.2.2 基于目的地的转发	77	7.2 基于约束路由的功能器件	126
4.3 非ATM链路上的标志封装	79	7.2.1 约束的最短路径优先	127
4.4 处理标志故障	81	7.2.2 MPLS作为转发机制	130
4.5 在路由瞬态过程期间处理转发环路	81	7.2.3 RSVP扩展	131
4.6 标志分布协议	83	7.2.4 CR-LDP	133
4.7 小结	84	7.2.5 OSPF与IS-IS扩展	135
4.8 参考信息	85	7.2.6 CR-LDP与RSVP的比较	136
第5章 MPLS核心协议	86	7.3 应用到流量工程	138
5.1 工作组起源与章程	86	7.3.1 问题的描述	138
5.2 MPLS体系结构	87	7.3.2 用ATM或者帧中继解决流量工程	139
5.2.1 有序的控制对独立的控制	88	7.3.3 为什么普通的IP路由是不够的	140
5.2.2 环路检测与预防	90	7.3.4 使用MPLS的基于约束的路由解决 流量工程	142
5.3 封装	93	7.4 应用于快速重路由	146
5.4 标签分布	95	7.4.1 普通IP路由的路由收敛	146
5.4.1 标签分布协议(LDP)	95	7.4.2 使用基于约束路由的快速重路由	148
5.4.2 利用BGP的标签分布	98	7.5 应用到QoS	151
5.5 ATM问题	99	7.5.1 QoS与路由之间的关系	151
5.6 组播	102	7.5.2 保证带宽的LSP	152
5.7 小结	103	7.6 小结	152
5.8 参考信息	104	7.7 参考信息	153
第6章 服务质量	105	第8章 虚拟专用网络	154
6.1 集成服务与RSVP	105	8.1 什么是VPN	154
6.1.1 集成服务概述	106	8.2 叠加模型	155
6.1.2 MPLS支持RSVP	109	8.3 对等模型	159
6.1.3 RSVP与可扩展性	111	8.4 路由信息的约束分布	160
6.2 差分服务	113	8.5 多个转发表	163
6.2.1 差分服务概述	113	8.6 VPN-IP地址	164
6.2.2 MPLS支持差分服务	115	8.7 MPLS作为转发机制	166
6.3 明确的拥塞通告	119	8.8 可扩展性	168
6.3.1 ECN概述	119	8.9 安全	169
6.3.2 MPLS支持ECN	120	8.10 QoS支持	170
6.4 小结	122	8.11 高级话题	174
6.5 参考信息	122		

8.11.1 因特网服务提供商作为用户	174	8.12 小结	181
8.11.2 BGP/MPLS VPN服务提供商作为 用户	177	8.13 参考信息	182
8.11.3 多个提供商的操作	178	结束语	183
		术语表	185

第1章 绪 论

因特网长期以来为技术的革新提供了肥沃的土壤。随着新的网络技术雨后春笋般地涌现，有些技术设法引起了人们特别的关注。多协议标签交换——MPLS——就是这样一种技术。本书对MPLS提供了一个全面并且均衡的论述，并且回答关于MPLS技术的一些基本的问题。第一个问题，MPLS是什么？第二个问题，它是怎样工作的？最后一个问题，它提供了什么好处？在回答最后这个问题的时候，我们希望表明，MPLS提供的好处可以证明它在网络工业界得到重视是应当的。

MPLS是从90年代中期发明的一些相似的技术演化而来的。这些技术中最出名的（虽然不是最先公之于众的）是IP交换，由它的发明者在开办Ipsilon公司时给予了这个称号。东芝公司在此之前也描述了一种类似的方案，并在公司的信元交换路由器（CSR）上实现了这个方案。一些其他的方法不久也公布了，著名的是Cisco公司的标志交换以及IBM公司的基于聚合路由的IP交换（ARIS）。所有这些方法都存在着一些共同的特点。它们都是利用简单的标签交换技术来转发数据的。而且它们都使用因特网协议集的控制，也就是说，它们都使用IP地址和标准的因特网路由协议，如OSPF和BGP。然而，不同的方法在它们的目的以及它们的实现细节上有很大的差异。

对于所有这些举动，毫不令人惊讶的是因特网工程任务组（IETF）特许设立了一个工作组来对吸收这些思想的共同方法进行标准化。IETF不期望选择的名称隐含对任何一个公司产品的认可，它为这个工作组挑选了一个中性的（稍微有些麻烦的）名称：多协议标签交换（MPLS）。MPLS技术目前正处于成为一个工业标准的过程中。

本书提供了对MPLS技术的详细描述和分析，以及这本书着手解决的一组问题。本书也描述了MPLS协议的体系结构问题和细节，探索了已经制定的设计抉择，并且讨论了这些选择的得与失。本书还描述了由MPLS的一些先行者发展起来的许多关键的概念。这有助于解释隐藏于MPLS之后的一些设计决策，并有助于阐明MPLS必须着手解决的广泛需求。最后，我们讨论了MPLS的一些主要的应用，并且着手解决了当使用MPLS适合和不适合时的一些问题。

我们在这个早期阶段就规范一些术语是很重要的。我们已经使用过很多次的一个词是“转发”。我们用这个词指代交换机和路由器这两者对分组执行的共同的操作：交换机和路由器在一个输入端口接收分组，通过检验此分组中的某些域来确定此分组需要到哪里去，并将它发送到合适的输出端口。

在本书中描述的一组方法将全部被称为标签交换技术。标签只不过是一个相对较短的、固

定长度的被用于转发分组的标识符。标签值对于一个单一的链路（更准确地说，一个单一的数据链路层子网）来说通常是本地有效的，因而没有全球的意义。它们也是无结构的，也就是说，它们不是由明显的功能器件组成的。一个标签交换设备在转发分组到下一跳之前通常用某个新的值替换这个分组中的标签。正是由于这个原因，我们称此转发算法为标签交换。在标签基础之上的转发判决使用精确匹配算法来决定发送分组到哪里去；我们将在第2章中详细地描述这个算法。并且，根据我们的定义，一个我们称为标签交换路由器（LSR）的标签交换设备运行标准的IP控制协议（比如路由协议、RSVP等等）来确定转发分组到哪里去。

我们在本书中论述了几个MPLS的先行者。根据上面的定义，它们全部都是标签交换技术，并且每一个都有其自己的术语。关于LSR，不同的厂商有不同的术语如信元交换路由器（CSR）、IP交换、标志交换路由器（TSR）或者标志交换，以及集成式交换路由器（ISR）。其他还有一些我们还没有囊括进来的方法，因为它们与这四种方法的差别不够大而还需要增加一些讨论，或者因为它们不是标签交换的方法。特别是，我们还没有囊括基于交换机的路由器，这些路由器的外部行为像一个传统的路由器，但是它们内部是以一个交换机构为中心建立的。这样的设备是一些硬件块，但是它们却解决了一组不同于LSR的问题。这种路由器的功能主要集中于性能问题，而性能问题仅仅只是标签交换着手解决的问题中的一个很小的部分。

在我们考虑这些方法中的任意一个的细节之前，考虑引起我们对这个方面关注的一系列因素将会是很有帮助的。

1.1 MPLS的出现原因

许多因素导致了标签交换技术的发展。通常假定的仅仅是一个因素——对快速、廉价的路由器的需求。这对于一两个MPLS的先行者来说可能是对的，但是，就如我们将在下面的章节中看到的，标签交换领域作为一个整体不仅仅只是由速度的要求来驱动的。在下面的讨论中，每一个因素只不过在标签交换方法的一个子集的发展中可能是重要的，但是，所有这些因素组合起来就导致了该领域发展到这一步，即在可预见的将来，MPLS似乎肯定是网络世界中的一部分。

1.1.1 因特网的增长与演化

如今谈论因特网的“爆炸性的”或者“指数的”增长是陈词滥调、老生常谈，但是事实仍然是因特网已经经历了显著的增长。因特网以几乎任何一种的尺度正在明显地扩大，并且这种增长已经造成了很多技术上的挑战。标签交换在某种程度是对这些挑战的一个响应。

因特网在用户的数目以及用户对带宽的需求这两方面的增长，已经对因特网服务提供商（ISP）的网络提出了越来越多的需求。为了满足对带宽不断增长的要求，ISP需要更高性能的交流与路由产品。我们将在下一节中讨论性能在推动标签交换的努力中的角色。

除了变得更快以外,网络还需要处理增加的节点数、路由表中更多的路由、通过一个给定点的更多的流,等等。一般说来,网络提供商需要考虑可扩展性,我们可以宽松地定义可扩展性使其具有在所有这些尺度上扩展网络而不会碰到不可克服的问题的能力。标签交换部分是由可扩展性的需求推动的,我们将在1.1.3小节中进一步地讨论这个问题。

或许标签交换身后最重要的,而在网络团体中不被看好的一个确切的推动因素,是因特网的路由功能和IP网络的路由功能进化发展的需要。因特网的增长正持续不断地对路由协议提出新的需求,并且还有一个对新的路由功能的不断增长的需求——既要处理网络增长本身又要满足不断增长的用户群体的发展的要求。

在过去,路由功能是非常难于进化发展的,其原因部分在于IP网络中路由和转发之间的紧密耦合。其中一个例子是采用无类的域间路由(CIDR)的过程。CIDR的效果就是说以前是8,16或者24比特长的IP网络前缀,现在的长度可以是任意的。这就大大地增强了在因特网中可以分配的地址的有效性,并且也有助于选址和路由信息的聚合有更大的可扩展性。然而,进行这个改变也需要实际应用中的所有IP路由器的转发算法有所改变,因为前缀现在可以是任意长度的。这些算法对一个路由器的性能来说是至关重要的,并且可以在硬件或非常相配的软件上实现。对转发算法所作的改动是一个典型、昂贵且耗时的提议。

标签交换的吸引力之一在于转发算法是固定的,并且新的控制不需要作任何改动就可以被采用。就如我们将在下一章中要看到的一样,多种多样的控制模块可以用于控制标签交换的过程,并且这些控制模块正好全部都使用相同的转发算法。因此,确实可行的做法是把转发算法植入硬件或者调整一次快速路径软件,而不必考虑每次需要一个新的路由功能时都将需要重新最优化。这对于缩短在IP网络中开发和使用新的路由功能所花费的时间来说具有重大的潜在优势。我们相信这是标签交换的最大的好处,因而标签交换可能成为下一代路由体系结构的基础。

在这一点上,有些读者可能想知道第6版的IP(IPv6)发生了什么。它不是被假定为下一代IP网络的基础吗?真实情况是,IPv6服务于一个非常特定的目的:扩展IP地址空间,因而更多的IP节点可以被唯一地编址。IPv6实际上对为第4版的IP(IPv4)发展起来的路由体系结构没有作任何改变。其他的新功能,比如资源预留、安全等等,虽然经常与IPv6同时讨论,但是在很大程度上它们是独立于IPv6的。

有趣的是,标签交换有潜力来简化IPv6的使用,因为它将不需要改变转发算法。如果给定携带IP v6地址的路由协议的可利用性,那么在本书中描述的所有标签交换技术就都能够和IP v6一起运行。

1.1.2 价格与性能

在基于因特网协议集的任何网络中,不管它是全球因特网的一部分还是一个专有的互连网

络[⊖]，关键的器件之一是路由器。一个路由器最基本的任务是转发IP分组（或者数据报）穿越网络。就如我们将要在下一章中更详细地看到的一样，转发IP数据报是一个相当复杂的操作。此外，路由器除了转发分组之外还经常执行广泛的功能，比如在一个网络的不同部分之间过滤分组的流。的确，路由器对于很多应用来说，最重要的特点不是它可以多快地转发分组，而是它可以提供一组多么丰富的功能。

另外一个重要的网络器件是交换机。然而路由器是第三层的设备（它们转发IP分组，而IP是7层协议模型中的第三层协议），交换机则是第二层的设备——它们转发第二层协议的分组。与路由器相比，交换机倾向于更简单一些，它不能提供一组同样丰富的功能，并且它们通常只支持很有限的协议和接口类型。与此相反，路由器通常支持许多的协议以及非常广泛的接口类型和速度。交换机的转发算法总是非常简单。某些类型的交换机，特别是ATM交换机和帧中继交换机，都使用在标签交换基础之上的转发算法。

如果让路由器和交换机执行不同复杂度的任务，那么路由器和交换机将倾向于显示出不同的性能/价格特点就毫不令人惊讶了。首先，我们将定义我们所指的性能是什么意思。刻画一个交换机或者一个路由器的性能可能是非常复杂的，因为这依赖于非常多的因素，比如呈现在一个设备的输入端口前的精确的业务流模式。然而，提出一个合理的且有代表性的性能参数通常是可能的，此参数要么根据设备在输入端口和输出端口之间能够转发的每秒分组数而定，要么根据设备的总带容量而定。比如，如果一个交换机有10个输入端口，每一个输入端口能够以150MB/s（ 150×10^6 比特/秒）的速度接受数据，并且交换机能够同时从所有这些输入端口交换数据，那么我们就可以说交换机总的容量是1.5GB/s（ $10 \times 150 \times 10^6 = 1.5 \times 10^9$ B/s）。

当我们着眼于交换机和路由器的性能/价格特点时，我们经常发现交换机的表现领先。根据这一点我们可以说，对于一个给定的性能等级，路由器的价格要比等效的交换机的价格高。反过来说，对于一个给定的代价，交换机比路由器提供更高的性能等级。当你考虑到路由器必须转发分组并执行各种各样其他的服务而交换机在本质上除了转发分组外几乎不提供其他服务这一点时，上述观点则不会令人惊讶了。由于将在下一章中讨论的一些原因，这种性能差别因为路由器实际的转发操作比交换机更加复杂这一事实而恶化。

与此观测结论相关的是这样一个事实，即在任意给定的时间内最高级的性能通常在交换机中被发现而不是在路由器中被发现。因而，例如具有处理总容量为10GB/s能力的交换机在相似容量的路由器出现之前就被较好地利用了。

关于观测到的交换机和路由器之间的价格/性能差别是根本上的，还是由于不同的设计目标、市场策略以及大量的其他因素导致的，或者仅仅只是历史遗留的产物的争论还存在着相当大的空间。然而，当各种各样的标签交换策略被提出来时，及时地在这一点上争论观测到的在价格/性能上的差别是非常困难的。

⊖ 在本书中，我们遵循惯例使用因特网（拼写为大写的I）这个术语来特指大多数读者都熟悉并且众所周知的全球网络，而我们使用互连网络这个通用术语来表示基于因特网协议集的任何一网络。

这就直接导致了让所有的标签交换方法发展的关键动机之一。如果你能够制造一个看起来象一个交换机使用硬件的设备来完成路由器最重要的工作——转发IP分组——会怎么样呢？你可能会拥有一个具有交换机的性能/价格特点的，而又具有路由器功能的产品。当然，你可能不会得到所有的更深层的路由器的特点，但是在很多场合中，那些特点并不是真正必需的。由于因特网的成功，而使IP是路由器需要处理的唯一一个协议的情况增加，这一点已经显得更具有吸引力。在一个交换机的价格/性能的水平上提供IP转发的期望，激发了在本书中描述的大量工作。

1.1.3 ATM上的IP集成

推动在本书中描述的这样一些工作的另外一个因素是期望集成IP和ATM（异步转移模式）。ATM交换机在二十世纪八十年代末期开始出现在市场上，并且许诺提供优于早期网络技术的巨大的性能改进。然而，随着ATM网络标准的演化，由主体（body）例如国际电信联盟（ITU；电信设备主要的标准设置主体，以前被称为CCITT）和ATM论坛所推动，ATM获得了一个与IP体系结构存在很大差别的体系结构模型。特别的是，尽管IP是基于数据报或者是基于无连接的数据传递模型，而ATM是基于面向连接的数据传递模型或者是虚电路模型。IP和ATM也具有完全独立的选址策略以及许许多多其他的差别，包括组播通信和资源分配的不同模型。这些不同的体系结构模型提出了一个重大的挑战，标签交换的努力在某种程度是对此挑战的一个响应。

在ATM发展的早期阶段，ATM将“主宰世界”的情况似乎是可能的（至少对某些观测家而言），也就是说，ATM将成为占支配地位的网络技术。许多提倡ATM的中坚分子都设想基于天然ATM的应用的发展，以及被设计来利用ATM的特点的新的协议堆栈。然而，许多ATM网络的一个主要功能将是转发IP数据报，这一事实很快就变得明显起来。这很大部分是因特网成功的结果，也是基于TCP/IP的应用——比如网页浏览器——已经牢固占领市场这一事实的结果。因此，ATM和IP标准的主体面临这样一个问题，即如何把IP体系结构“映射”到ATM网络上。

使得这个问题更加紧迫的一个因素是ATM和商业上可获得的路由器之间的性能上的差别。ATM交换机相对较高的性能以及ATM被设计为在广域网路上运行的这一事实，使得ATM成为建立一个互连网络的骨干网络的一种有吸引力的技术。现今的因特网相当大的部分用ATM交换机扩建，而在ATM交换机的外围是速度相对较低的路由器。这样的一种网络设计的例子如图1-1所示。

在这里示意的网络类型，通常被认为是使用了叠加模型。其思想是在ATM网络之上叠加IP网络。ATM网络提供了一个高速连通的核心，而由ATM虚电路互连的一组路由器组成的IP网络，则提供转发IP数据报的智能，而这是网络的主要工作。

将IP映射到ATM上的问题由许多标准组织来解决，主要是由ATM论坛和IETF来解决。问题的复杂度在某种程度上可以仅仅通过计算处理这个问题的各方面的不同的工作组数目来估

计。这些工作组包括：

- ATM上的IP（IPATM）工作组，该工作组定义了IP数据报携带在ATM适配层的PDU中时的封装，并为IP地址映射到ATM地址定义了一个地址解析协议（ATMARP），这在后来被扩展用于处理组播。
- 大的公用数据网络上的IP（IPLPDN）和稍后的大的云上的路由（ROLC）工作组，这两个工作组定义了下一跳解析协议（NHRP），使得分离很远的主机和路由器能够穿越一个ATM网络建立直接的虚电路。
- LAN仿真工作组（LANE），该工作组定义了一些程序，使得一个ATM网络看起来更像一个多路访问LAN。
- ATM上的多协议（MPOA）工作组，该工作组联合并扩展了许多其他工作组的工作来支持多个网络层协议（与仅仅只支持IP相反）。
- 特定链路层上的集成服务（ISSLL）工作组，该工作组正在定义将IP的资源预留模型映射到ATM（与其他链路层一起）的资源预留模型上的程序。

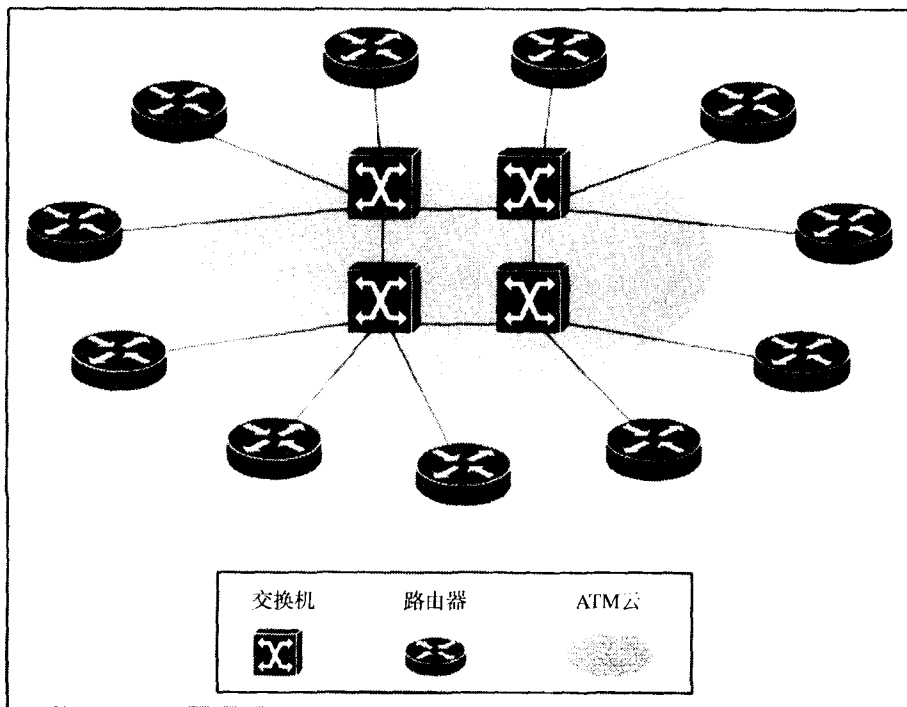


图1-1 一个ATM上的IP网络、路由器连接到一个ATM交换机的“云”

正如这个快速的综述所建议的一样，IP和ATM之间的映射涉及到相当大的复杂性。上述工作组的大多数已经定义了某种服务（ATMARP、MARS、NHRP、BUS以及一些待命名的）来

处理映射功能中的一个方面，连同一起的还有与服务器交互所必需的协议。一个一直存在的问题就是如何处理服务器表现出来的单点失效的这一事实，并因而期望设置冗余的服务器，这就当然需要同步协议来使得它们彼此保持一致。

该复杂性来自于何处呢？从本质上来说，它全部来源于一个事实，即因特网协议和ATM协议是在彼此互不关心的情况下发展的，并且碰巧在不同的位置上结束。这种实现方案已经促使许多人想知道ATM交换机是否可以和一组不同于ATM论坛和ITU定义的协议——与因特网体系结构更加一致并因此而消除过多的复杂映射的需要的一组协议一起使用。本书中描述的一些标签交换的工作事实上就企图定义这样一组协议，该组协议能够以这样一种方法控制ATM交换机，使得ATM交换机可以自然地转发IP分组，而不需要在IP和ATM之间加入半打提供映射的服务器的帮助。

正如我们将在后续章节中看到的一样，使用标签交换转发分组的各种各样的方案全部都废除了叠加模型及其带来的复杂性。所有的提案都没有采用两个不同的协议体系结构，诸如不同的寻址、选路协议、资源分配方案等等，而是都使得IP控制协议能够直接在ATM硬件上运行。ATM交换机仍然使用标签交换技术转发分组，但是ATM交换机用于建立转发表以及分配资源的机制则全部都是由IP控制协议来驱动的。从控制的观点来看，ATM交换机已经有效地变成了IP路由器，因此免除了IP和ATM控制模型之间映射的需要。

可扩展性问题

有一个很少有人知道的但是很重要的扩展问题，也就是只要一个IP网络叠加建立在一个第二层网状网络上时，就如可以利用帧中继骨干网以及ATM骨干网实现的一样，这个问题就会产生。对于“扩展的问题”，我们指的是网络中的某些复杂程度比网络中的节点数增长得更快，因此在某点上任意扩大网络是非常难于实现的。

考虑图1-2中所示的网络。如果位于中间的ATM网络在所有的路由器之间提供高速的连通，那么一个完全的虚电路网状网络就必须互连所有的路由器。任何的减少将意味着在某些路由器对之间存在着一个额外的路由器跳，这种情况将导致位于路径中间的这个额外的路由器成为一个瓶颈。然而，一个完全的虚电路网状网络意味着所有的路由器实际上都彼此相互直接连接。在图1-2中，为了避免使图形混乱，我们仅仅只显示了在其中一个路由器上发起的VC（虚电路）。实际上，一个完全的网状网络将包含 $n(n-1)/2$ 个VC。也就是说在本例子中为55个VC。

在这里为了评估这个问题，就必须对路由协议是如何工作的有所了解。一个路由器通常被配置为与每一个直接与它连接的邻机都有一个邻接。邻接使得路由器能够明了谁与它直接相连并明了任何一个链路是否已经断开，并且邻接被用来与那些邻机交换路由信息。

当路由器由一个完全的虚电路网状网络连接时，任何一个路由器具有的直接的邻机数目等于围绕它的路由器的数目减1（它自己）。在路由器中间存在ATM交换机的这一事实并不妨碍这

些路由器在网络层看上去似乎是直接相连的表现——交换机在这一层被有效地隐形了。于是，在图1-2中，任何一个路由器具有的邻接的数目都是10——连接到此云的其他路由器的数目。因此，在该图中的线不仅表示VC，而且还表示路由邻接。

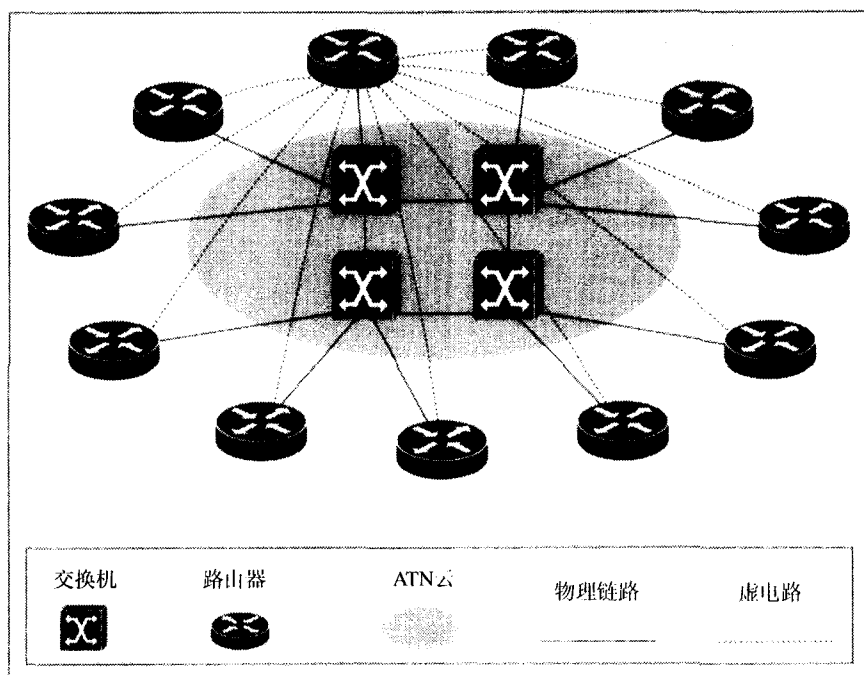


图1-2 在一个ATM网络上的路由器之间的虚电路

可以表明的是，当网络的核心面对拓扑的改变时，在这样一种网络中被传输的路由信息的总量差不多可以是 n 的4次方，其中 n 是围绕此核心的路由器的个数（这个结果的推导太棘手所以不在这里列出）。因为随着 n 的增加而使信息总量增加得如此之快，所以它可以达到这样一点，即仅路由业务流就可以使路由器超载，导致路由器的性能非常差。因此，我们推断这样一种设计从扩展性观点来看是不好的。

一些方法可以被采用来缓解这个扩展的问题。开始的一点是消除VC的完全的网状网络。正如我们观察到的一样，这意味着现在某些路由器对之间的路径中将包含一个额外的路由器跳，如果这个中间的路由器的性能足以处理额外的业务流，那么这种方法可能是可以接受的。做为选择的是，下一跳解析协议（NHRP）允许路由器建立VC，在此VC上路由器可以发送数据而不需要在此VC上建立路由邻接。这种方法有它自己的一套问题，包括运行许多的NHRP服务器的需要，以及引入持久的转发环路的可能性。NHRP也仅仅只适合于单播业务流；而没有为组播业务流定义NHRP。

解决扩展问题的另外一个方法涉及到标签交换。首先，让我们回想一下，根据定义，一个

LSR运行IP控制协议，包括IP路由协议，并且LSR有可能使用没有改动的ATM硬件来实现。因此，不需要改变图1-2中的物理拓扑或者设备，通过在ATM交换机上运行IP路由协议，我们能够大大地减少任何一个路由器邻接的数目。这种情况如图1-3所示。

因为ATM交换机有能力运行IP路由协议，所以中间邻机的概念现在已经从一条VC另一端的上的路由器改变为在物理链路另外一个端点上的设备——路由器或者LSR。因此，任何一个路由器具有的邻接的最大个数被大大地减少了，并且此数目再也不会随着网络的大小而增加，这就导致了一种更加易于扩展的设计。注意到在图1-2中的曾经有十个邻接的路由器，现在仅仅只有一个邻接，并且每个LSR（过去曾经是ATM交换机）都不超过五个邻接——和每一个直接相连的路由器有一个邻接，不管此LSR是一个传统的路由器，还是一个作为LSR运行的ATM交换机。

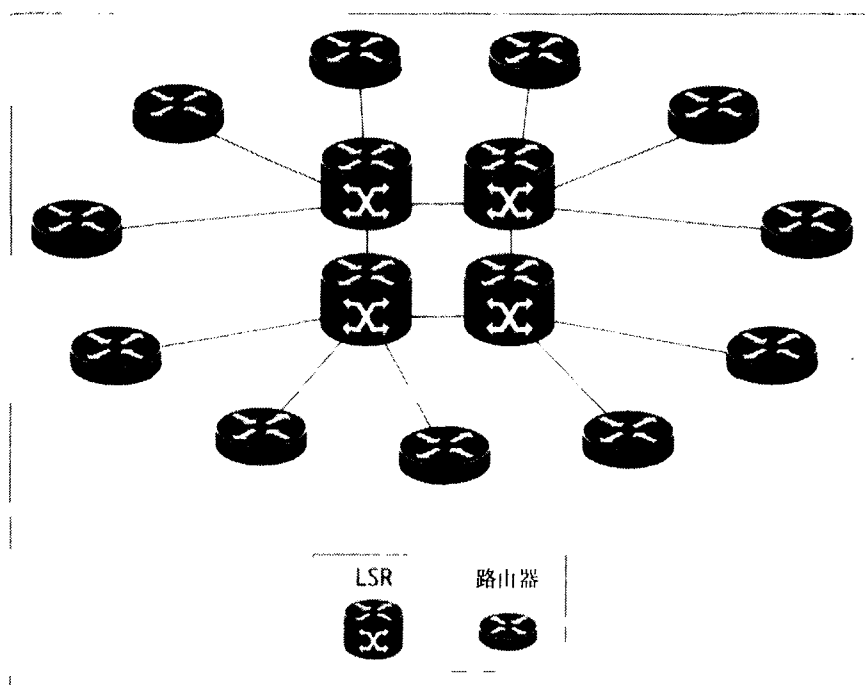


图1-3 使用标签交换技术消除邻接的完全网状网络。每一个路由器或者LSR与它的最近的邻机都有一个邻接。与图1-2不同的是，在这个图中，每一个物理链路上只有一个邻接

这不是使用标签交换来提高可扩展性的唯一情况。我们将在4.1.2节中考虑另外一种不是针对ATM的情况。

1.1.4 扩展路由功能

标签交换不仅仅只是使ATM硬件表现得更象一个IP路由器，强调这一点是很重要的。它也

不只是关于制造更快更廉价的IP路由器。标签交换也能提供新的功能，而由现有的IP路由技术来提供此功能是很不容易的。提供新的路由性能的能力已经成为标签交换发展的另外一个幕后的推动因素。在此小节中我们将讨论的问题略微不同于1.1.1小节中路由演化的描述。这里的一个问题是标签交换使得传统的路由不能够支持的路由能力成为可能。

标签交换许诺扩展传统的IP路由能力，在这方面最著名的例子是基于目的地的路由。实际上在当今的IP网络内的所有路由都是基于目的地的；也就是说，关于转发分组到哪里去的判决仅仅只是在目的地地址的基础之上作出的。从原则上来说，当决定转发一个分组到何处去时，IP分组头中的其他域（例如源地址和服务类型比特）可以被使用，但是各种各样的设计考虑已经导致路由器的设计几乎总是单独地在目的地地址的基础上转发分组。与此相反，依赖诸如帧中继和ATM的网络技术的标签交换技术能够提供不同的功能。

为了了解标签交换如何能够提供新的功能，考虑图1-4中所示的网络。这是传统的路由很难提供的一种功能类型的一个典型例子。网络有时被称为“鱼形图”，因为网络在断面上象一条鱼。考虑一种情况，在这种情况下路由器B是一个传统的路由器，它仅仅只使用IP目的地地址来转发分组。当来自路由器B的其中一个邻机的分组到达路由器B时，转发判决除了受目的地地址的影响外不受其他任何因素的影响。现在假定我们用路由器B来实现一个简单的策略：“来自路由器A且将要到路由器F去的分组应当通过路由器D，而所有其他的其目的地是路由器F的分组应当通过路由器E”。如果仅仅只考虑目的地的转发机制很明显不能够实现这种策略。

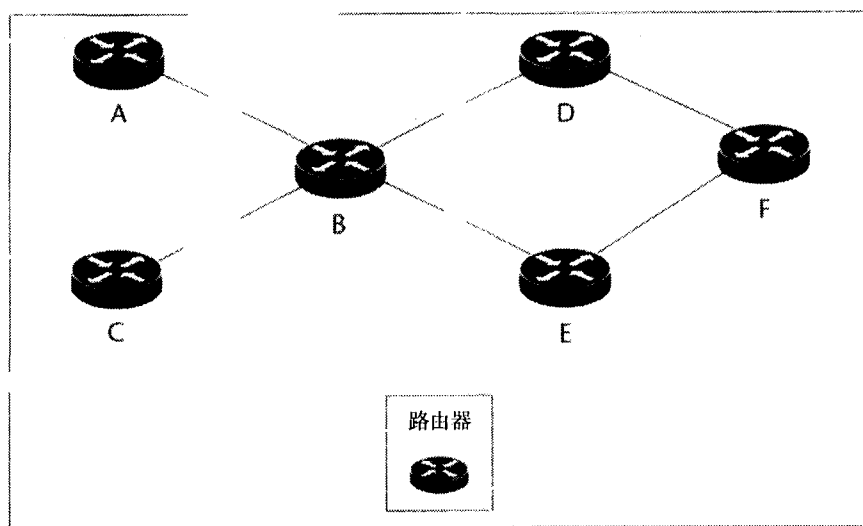


图1-4 标签交换式路径

与此相反，如果路由器B在标签交换的基础上转发分组，那么就很容易在路由器B上实现这样一种策略。主要原因是路由器A和路由器C没有义务为所有将前往路由器F的分组使用相同的标签。假定路由器A对目的地是路由器F的分组使用一个值为5的标签，而路由器C使用一个值