

多协议标记交换：在新型公网中增强路由功能

Juniper 网络公司，爱立信公司，1999 年

概要

不论大型 ISP 还是小型 ISP 都不断面临着使其网络满足快速成长和客户提出的更高的可靠性和多种业务需求的挑战。在 90 年代中期，IP - over - ATM 模型为许多 ISP 提供了实现卓越的性能及流量工程的解决方案。而且，许多运营商发现将 Internet 业务作为 ATM 核心网上所承载的多种业务中的一种进行复用是非常经济的。

最近，Internet 业务的不断增长和光纤领域内波分复用（WDM）技术为将多种业务复用到特定线路上提供一种除 ATM 以外的可选方案。而且，现在的 Internet 骨干网路由器同样可以提供 ATM 交换机具有的更高、更快的带宽。同样重要的是，多协议标计交换为用于分组的流量工程和多业务功能提供了更为简单的机制，并且能够在可扩展性方面提供更多的优势。

IETF 努力地对一些最初在 90 年代中期被建议的专用多层交换方案标准化，进而形成 MPLS 标准。为帮助您进一步了解 MPLS 的重要性和它在 Internet 核心网上的影响，这篇白皮书的第一部分将阐述促使开发和发展这些不同方案的原因，着重分析这些不同的方案所共有的一般功能和设计思路 - 将控制部分与转发部分完全分离并使用标计转发模式。这部分也将讨论普通技术的发展和 IETF 工作组制定 MPLS 的一些情况。

这篇白皮书的下半部分将帮助您理解多层交换技术，着重分析 MPLS 的技术细节。它将讨论 MPLS 工作组的目标及对象，MPLS 组成的核心，一些对 MPLS 的错误概念，MPLS 为 Internet 核心网带来的一些好处，和 MPLS 的一些典型应用。这部分将讨论 MPLS 是怎样为多种业务提供基础，MPLS 如何使 ISP 能够提供传统 IP 路由技术所不能支持的提供新业务的快速支持。

内容

介绍.....	3
多层交换在 Internet 中的发展.....	3
基本组件.....	3
控制与转发部分的分离.....	4
标计交换转发算法.....	4
ISP 转移到 IP - over - ATM 模型.....	6
多层交换可作为 ATM 以外的选择.....	7
各种多层交换方案的相同之处.....	8
各种多层交换方案之间的基本区别.....	9
数据驱动的模式.....	9
控制驱动的模式.....	10
多层交换方案的基本问题.....	11
多协议标记交换.....	11
需求和对象.....	12
对 MPLS 的常见错误概念.....	13
MPLS 的承诺.....	13
MPLS 应用.....	15
流量工程.....	15

服务等级.....	16
虚拟专网.....	17
总结.....	18
参考.....	19

介绍

在过去几年中，在 ISP 们努力与 Internet 的迅猛发展保持同步的过程中，许多新技术被开发和设计以支持他们的需求。这些最先进的技术包括 Internet 骨干网路由器，新的排队和时间安排算法，IPSEC，web 高速缓存业务，免疫力业务，和集成的路由/转发方案。所有这些技术对 Internet 的成功运行和持续增长都起着重要的作用，同时，如果 ISP 希望提供对新型增值客户服务的支持，路由功能性的发展是最基本的。

Internet 核心网中的路由/转发技术的最新发展是多协议标计交换 (MPLS)。MPLS 可以将 IP 路由的控制和第二层交换的简单性无缝地集成起来。而且，因为 MPLS 解决了许多复杂的问题，它为支持先进路由服务的配置提供了基础，这些复杂问题包括：

- MPLS 解决了现在普遍使用的 IP - over - ATM 覆盖模型的可扩展性问题。
- MPLS 明显的减少了网络运行的复杂程度
- MPLS 简化了用于增强传统 IP 路由技术的新型路由能力的推出
- MPLS 提供了一个基于标准的解决方案，从而保证了多个厂商之间的互操作性

IETF 努力地对一些最初在 90 年代中期被建议的专用多层交换方案标准化，进而形成 MPLS 标准。为帮助您全面了解 MPLS 的本质和它在 Internet 中的角色，很有必要去回顾一下促使那些专用的多层交换方案进行开发的因素，和它们最后如何被集成到 MPLS 中去。

Internet 中多层交换的发展

多层交换是将第二层交换和第三层路由集成在一起的一种技术。今天，一些 ISP 网络使用覆盖模型建立的，即逻辑 IP 路由拓扑运行在一个底层的独立的第二层交换拓扑（ATM 或帧中继）之上。第二层技术提供高速链接，同时 IP 路由器被放置于网络的边缘 - 通过第二层虚电路的结网进行互联 - 以提供转发 IP 报文的智能性。这种实现方式的难点在于在两个分离结构间进行映射，需要对分离的拓扑，地址空间，路由协议，信令系统，和资源分配方案进行定义和维护。多层交换解决方案和 MPLS 是 Internet 发展的一部分，它们通过将第二层交换和第三层路由集成到一个完全的解决方案中去来降低复杂性。

基本成分

在开设讨论 Internet 中的多层交换之前，了解所有多层交换方案和 MPLS 所包含的基本成分是非常重要的：

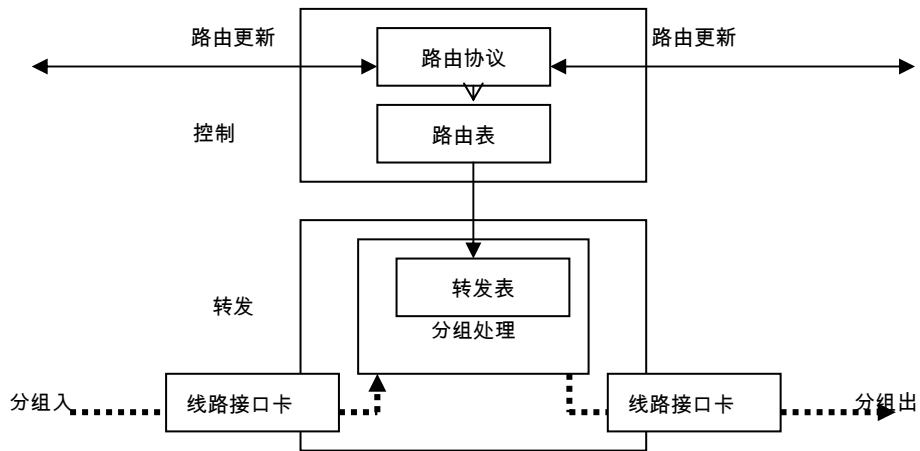
- 控制和转发部分的分离
- 分组交换转发算法

控制和转发部分的分离

所有多层交换技术，包括 MPLS，都由两个分立的基本组成部分 - 控制部分和转发部分。控制部分通过标准的路由协议（OSPF，IS - IS，BGP - 4）同其它路由器交换信息，以建立并维护转发表。当数据包到达时，转发部分查询由控制部分维护的转发表，为每个数据包作出转发决定。特别地，转发部分检查包含在数据包报头里的信息，在转发表中进行匹配查询，并通过系统的交换结构将包从输入接口引导至输出接口。

由于控制部分和分组转发部分的完全分离，可以对每个部分进行独立地开发和修改。唯一的需求是控制部分通过管理包转发表与转发部分继续进行通信。我们将了解到一个非常简单的转发算法的配置，如标计交换，可以为支持新型的增值客户服务提供所需的扩展转发能力。

图 1：路由功能模块：控制及转发



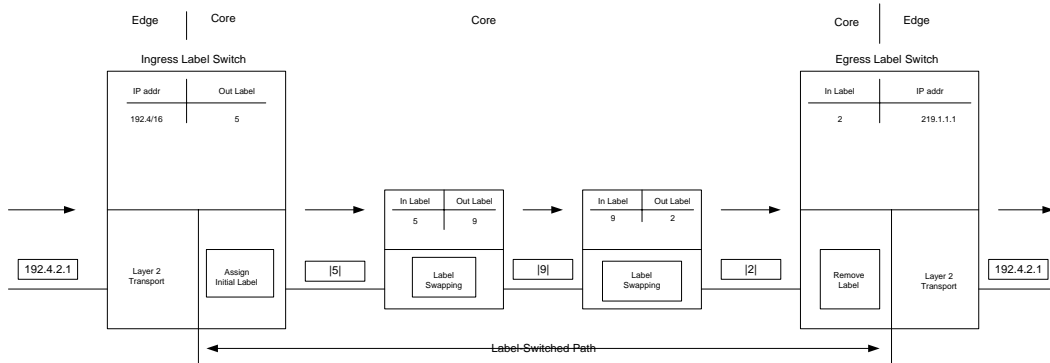
标记交换转发算法

事实上所有的多层交换方案和 MPLS 的转发部分都是基于标记交换转发算法。这个算法与在 ATM 和帧中继交换机中使用的转发数据的算法相同。信令和标记分配是标记交换转发算法运行的基础，但是，我们将不在此白皮书中过多解释。

标记是在包头中包含的短的且具有固定长度的值，它定义了转发等效级别（FEC）。标记与链接标识符类似，如 ATM 的 VPI/VCI 或帧中继的 DLCI，因为其本地链接的特点，标记并不对网络层的报头进行编码，只是将业务映射到特定的 FEC 上去。FEC 是指一系列使用相同路径转发而通过网络的数据包，即使他们的最终目的不相同。例如，对于传统的最长匹配 IP 路由，FEC 的一个例子是，一系列单点传送的数据包，它们的目的地映射到一个给定的 IP 地址前缀。

标记交换转发算法要求分组在网络的入口边缘进行分类，为每个分组分配一个特定的标记。在图 2 中，输入标记交换机接收到一个目的地为 192.4.2.1 的无标记的分组。标记交换机执行最长匹配路由表查询，经分组映射到一个 FEC - 192.4/16。数据标记交换机为该分组分配一个标记（标记值为 5），并将其转发至标记交换路径中的下一跳。

图 2：分组在一条标记交换路径中传输



LSP 与虚电路的功能相似，因为它定义了一条通过网络的入口 - 出口路径，所有分配了特定 FEC 的分组都将通过此路径。LSP 中的第一的标记交换机被称作入口，或起始标记交换机。LSP 中的最后一个标记交换机被称作出口，或终止标记交换机。

在网络的核心部分，标记交换机忽略分组的网络层报头，只是简单的使用标记交换算法对分组进行转发。当分组到达一个交换机时，转发部分使用输入端口号码和标记在其转发表内进行精确匹配查询。当发现匹配时，转发部分从转发表重新获得输出标记，输出接口和下一

跳地址。转发部分则使用输出标记交换（或替换）输入标记，引导分组至输出接口，使其传输到 LSP 的下一跳。

当被标记的分组到达出口标记交换机时，转发部分搜索其转发表。如果下一跳不是标记交换机，处理交换机将去除标记，并按传统的最长匹配 IP 转发对分组进行转发。

标记交换与传统的一跳接一跳网络层路由相比提供了一些重要的操作优势：

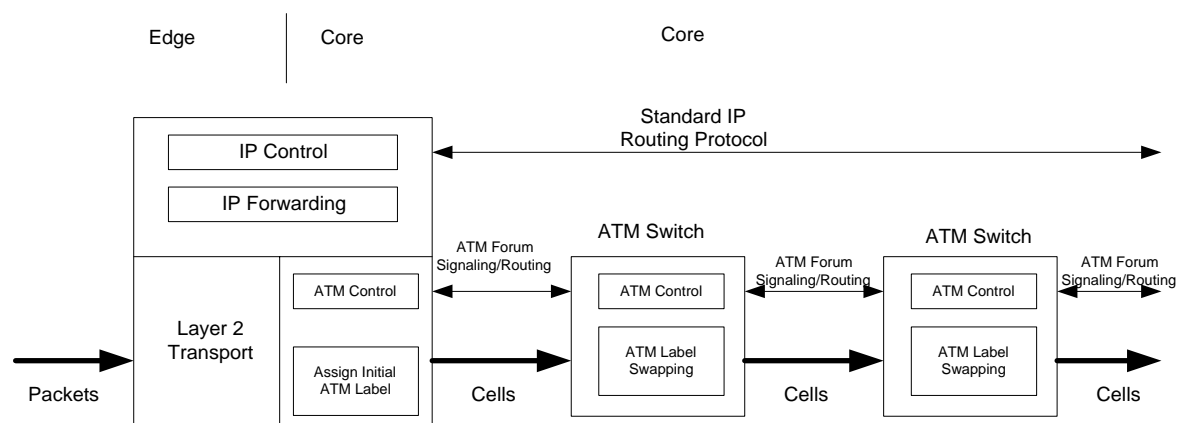
- 标记交换为服务提供商在将分组分配至 FEC 的方法上提供了极大的灵活性。例如，为模仿传统 IP 转发，入口标记交换机可以被配置成基于分组的目的地地址来将分组分配至一个 FEC。但是，分组也可以基于其它一些基于策略的考虑被分配至某个 FEC，如源地址，应用类型，进入标记交换网络的切入点，离开标记交换网络的地址，分组报头所传送的 CoS，或上面提及的任意组合。
- 服务提供商可以制定定制的 LSP 以支持特定的应用需求。LSP 可被设计成使跳转数最小，满足特定的带宽需求，支持精确的性能需求，旁路潜在的阻塞点，引导业务离开由 IGP 选择的缺省路径，或简单地强迫业务通过某特定链接和网络中的某个节点。
- 标记交换算法的最重要的优势是可以将任何类型的用户业务分配至一个 FEC，并将此 FEC 映射到一个满足该 FEC 需求的特定的 LSP 上去。这种基于标记交换转发技术的技术的实施使 ISP 能够对其网络中的业务流进行明确地控制。这种空前级别的控制使网络的运行更为有效，并可以提供更多可预测的业务。

ISP 转移到 IP - over - ATM 模型

90 年代中期，一些 ISP 将其网络从基于路由器的核心网转移到在 ATM 上运行 IP 的覆盖模型上去。ISP 采取这种转移是因为他们需要更大的带宽，更为确定的转发性能，和支持在其网络中发生的爆发性增长的流量工程。IP - over - ATM 模型能够满足这些运行需求的一个主要原因就是 ATM 使用标记交换转发算法。

IP - over - ATM 模型 (见图 3) 在网络中心使用 ATM 的功能，网络核心部分的每个系统都要求具有 ATM 软件控制 (信令和路由) 和硬件转发 (标记交换) 。 IP - over - ATM 模型通过在网络的边缘使用第三层的功能来满足应用需求，依靠核心部分的高速、标记交换 ATM 交换机和 PVC 使网络的吞吐量最大。 IP 路由仅局限于网络的边缘，因为这种模型将基于软件的路由器作为网络低性能的主要原因。

图 3 : IP - over - ATM 模型



90 年代中期，并没有专门为 Internet 骨干网而设计的网络设备，因此，ISP 们只能采用那些已商品化的设备以支持其快速增长的网络。只有 ATM 交换机设备能够提供支持他们直接需求的带宽和转发能力。但是，随着 ISP 继续其爆发式的增长及一些以 Internet 应用为特定目的而制造的设备的出现，继续延用 IP - over - ATM 模型及其内在的扩展性方面的问题变得越来越没有意义。扩展性问题包括 ATMSAR 接口的带宽局限性，20% 的信元税，PVC“ N²” 问题内部网关协议 (IGP) 的压力，和在非 ATM 基础结构上运行的无力。如果您对由于 IP - over - ATM 覆盖模型所引入的挑战不十分熟悉，请参阅 Juniper 网络公司的白皮书：“ *新型公网中的流量工程*” 。

尽管有这些扩展性方面的问题，最大的问题还在于对两个分立网络进行管理的复杂性。IP 和 ATM 基于完全不同的协议体系 (无连接和基于连接) ，它们都有其各自的地址模型，路由协议，信令协议，和资源分配方案。随着 ISP 对 ATM 和标记交换所提供的性能和控制的不断增长，他们发现在一个基于分组的网络环境中证明 IP - over - ATM 模型的复杂性是很

困难的。今天，随着专门为 Internet 核心网设计的高性能 Internet 骨干网路由器的出现，没有什么原因继续坚持那种需要两套分立设备的复杂的覆盖型实现。

作为 IP - over - ATM 模型备选方案的多层交换技术

随着 ISP 继续向 IP - over - ATM 模型转移，一系列的技术、市场及经济方面的原因开始发生影响，需要开发一种为 Internet 核心网设计的新技术。公众开理解 Internet 在为新的全球经济提供基础过程中的重要角色。我们可以很清楚的认识到的，Internet 的市场已经大到需要开发一种专门用于 Internet 骨干网应用的设备的程度了。IP 迅速成为能够战胜 IPX，AppleTalk，OSI 及 SNA 的唯一的协议。

“IP 收敛”的概念为风险投资者及发起人与现有的生产厂商竞争提供的一线机会。为获得成功，发起人必须能够提供 ATM 交换机的性能与价格和 IP 路由器的控制功能，同时还要避免 IP - over - ATM 模型的复杂映射需求。在 1996 年末，许多厂商提出了集成了 ATM 交换和 IP 路由的专有的多层交换解决方案，其中包括：

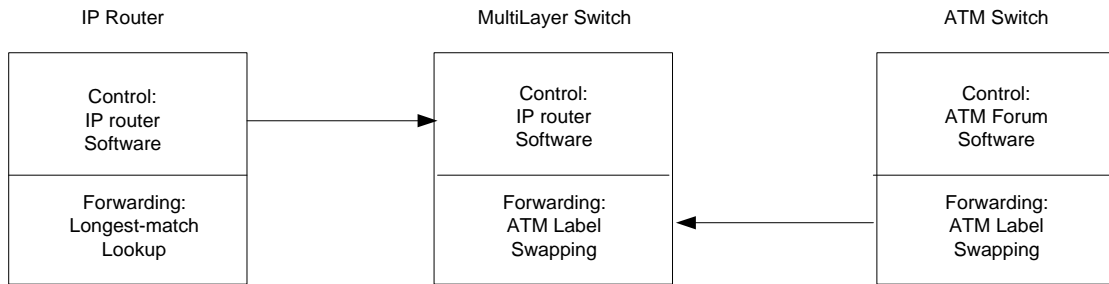
- Ipsilon/Nokia 设计的 IP Switching
- Cisco 开发的 Tag Switching
- IBM 设计的 Aggregate Route-Based IP Switching (ARIS)
- Cascade/Ascend/Lucent 推出的 IP Navigator
- 东芝开发的 Cell Switching Router (CSR)

虽然这些方法有许多共同之处，但是由于它们采用了不同的技术对 IP 路由和 ATM 交换进行集成，因此这些方案并不能互通。但是，在 1997 初期，Internet 界的许多人都对这些方案的简单和简洁的特点留下深刻的印象，他们开始将多层交换作为设计大型 ISP 骨干网的下一个逻辑进化步骤。

多层交换解决方案的共同点

每个多层交换方案都在力图组合 IP 路由和 ATM 交换的最佳特点，同时仍维持以 IP 为中心。这些策略所采取的基本方法是从 IP 路由器获取控制软件，将其与 ATM 交换机的转发性能和标记交换集成，从而建立一个高速而经济的 IP 路由器（见图 4）。

图 4：多层交换机作为快速 IP 路由器



对于控制部分，每个多层交换机都运行标准的 IP 路由软件（OSPF，IS - IS，和 BGP - 4）和专有的标记捆绑机制。路由软件使多层交换机能够交换第三层的网络可达信息。标记捆绑机制将第三层路由映射到标记（即，到 ATM 的 VPI/VCI），并将它们分布到邻居中去以在网络的核心建立 LSP。在系统的核心部分运行路由协议要比只在系统的边缘运行路由协议在增强网络的运行方面获得更多的好处：

- 避免了 IP - over - ATM 模型的“ N^2 ” PVC 扩展问题
- 降低了由于动态减少对等体数量而对内部网关协议产生的压力
- 可以获得实际的核心网络物理拓扑，能够在网络层进行路由

对于转发部分，多层交换机使用传统的 ATM 交换硬件和标记交换在网络的核心部分对信元进行转发（见图 5）。但是，用于为路由器分配标记、在多层交换机中分配标记和建立转发表的 control 部分是通过专有的基于 IP 的协议进行管理的，并不是 ATM 论坛的协议。在网络核心部分的 ATM 标记交换提供许多优势：

- 标记交换通过基于硬件的转发优化网络性能。厂商们确信这将提供新一代的具有极高性能价格比的产品，同时大大缩短推向市场的时间。在个问题只出现在 90 年代中期，

其实这并不是多层交换所具有的最先进的优点，因为技术的进步为建造 Internet 骨干网路由器提供了更好的方法。

- 标记交换实现了外在路由。外在路由即是对业务通过服务提供商网络的节点序列预先进行设置，从而使转发路径不同于通常的使用基于目的路由而计算出的路径。外在路由是 ISP 能够对业务流进行精确的控制，使对流量工程、QoS，和阻止循环的支持成为可能。
- 标记交换提供了比传统基于目的路由更强的扩展控制工具。多层交换所提供的比传统路由机制支持更强的增强转发控制能力表现在它在网络设计过程总所表现出的持久的作用。在此白皮书的后面我们将讨论多层交换是如何简化新一代路由功能的实施的。

因为不包含 ATM 论坛的路由和信令协议，多层交换避免了对两个不同协议体系 - IP 和 ATM 的协调和映射，从而减少了操作的复杂性。虽然多层交换仍继续使用标准的 ATM/VPI/CVI 作为标记，但标记是通过专有的基于 IP 的协议检修分配和分布的，而不是采用标准的 ATM 论坛协议。这样，通过避免提供不同体系间的复杂映射而简化了 IP 路由与 ATM 交换的集成。但是，多层交换解决方案的最主要的局限性在于它们被局限在运行在基于信元的 ATM 基础结构，而与此同时，Internet 正在向基于分组的方向发展。

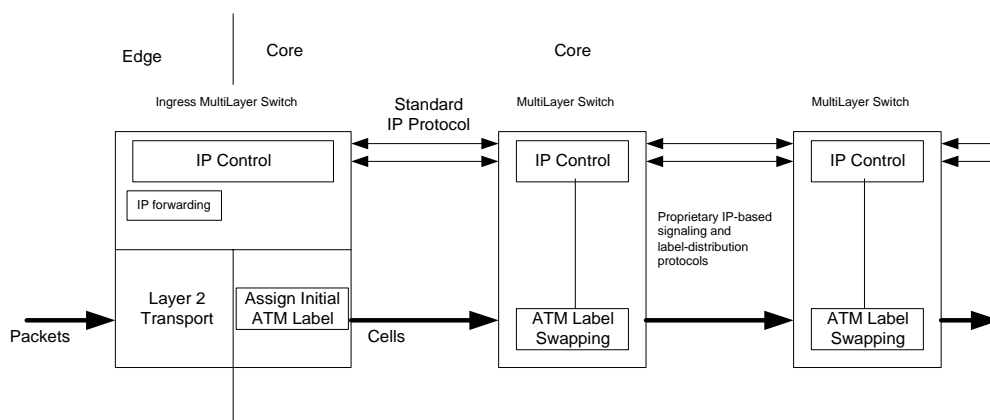


图 5：多层交换解决方案 - - 使 IP 路由和 ATM 交换一体化

各种多层交换解决方案的基本区别

虽然各种多层交换解决方案有许多共同的功能特点，但从根本上，它们是基于两种不同的方法来在建立 LSP 时进行标记捆绑的分配与分布的：

- 数据驱动模型
- 控制驱动模型

数据驱动模型

在数据驱动模型中，数据捆绑是在用户的数据包到达时建立的。数据流是一系列具有相同的源及目的 IP 地址和 TCP 或 UDP 端口数的分组。标记捆绑即可以在多层交换机接收到数据流中的第一个分组时建立，也可以在接收到数据流中一定数量的分组时再建立。接收到数据流中一定数量的分组再建立标记捆绑可以确保数据流足够长以利于标记分配及分布的开销。多层交换解决方案中提供数据驱动方法的有 IP 交换（Ipsilon）和信元交换路由器（东芝）。注意，MPLS 不支持数据交换模型。

数据驱动模型的优势在于标记捆绑只有在业务流使用标记捆绑时才被建立。但是，这些模型在具有大量业务流的大型 ISP 网络中实施时也存在一些局限性：

- 每个多层交换机都必须提供复杂的和高性能的分组分级能力以定义业务流。
- 通常，在识别一个业务流和为该业务流分配一个标记之间有一个延迟。这就意味着每个多层交换机必须在准备阶段支持最长匹配的 IP 转发，对没有被分配到业务流中的分组进行转发而不被丢弃。
- 分派标记捆绑所需的控制业务流的数量直接与业务流的数量成比例。
- 大量相关的短期业务流的出现加重了网络运行的负担。

传统的专家认为数据驱动模型不具备 Internet 核心网应用所需的扩展性。

控制驱动模型

在控制驱动模型中，只有当控制信息到达时才建立标记捆绑。标记参考路由协议业务的普通处理或类似于 RSVP 业务的控制业务来进行分配，或基于静态配置。执行控制驱动模型的多协议交换解决方案包括标记交换（CISCO），IP 导航器（Ascend/Lucent），和 ARIS（IBM）。另外，MPLS 使用控制驱动模型。

在大型 ISP 网络的核心部分配置控制驱动模型有许多好处：

- 标记在用户数据业务到达前便被分配和分布。这就意味着每条出现在 IP 转发表中的路由都已经分配了标记，因此业务到达多层交换机时可以立即通过标记进行交换。
- 因为标记交换路径只与 IP 转发表中的条目有关而与特点业务流的数量无关，因此扩展性要明显优于数据驱动模型。对于大型 ISP 网络中的流量工程，扩展性要更为优秀 - 因为其只与网络中出现的节点数成比例。
- 对于一个稳定的拓扑，标记分配和分布的消耗要比数据驱动模型低，这是因为标记交换路径只有在拓扑发生改变或控制业务到达时才被建立，而并不因为每个“新”的业务流的到达而建立。
- 业务流中的每个分组是通过标记进行交换的，而并不象数据驱动模型那样只在业务流的终止处进行。

多层交换解决方案的基本问题

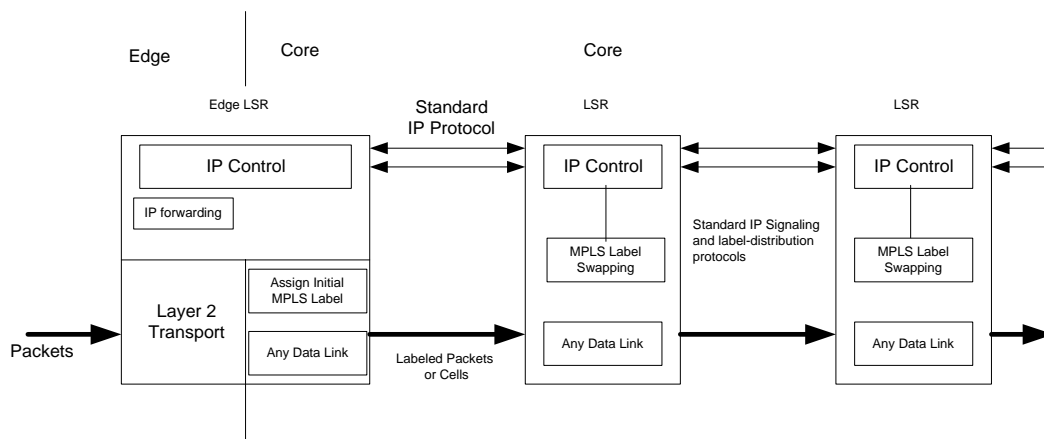
每个多层交换解决方案都包括 IP 控制部分和使用 ATM 标记交换的转发部分。ISP 界所面临的挑战是每个解决方案都是专有的，因此不能进行互操作。同时，因为多层交换解决方案不能工作在混合媒体结构中（帧中继、PPP、SONET，和 LAN），因此其需要 ATM 传输。如果多层交换希望在 ISP 中广泛应用，在需要有一种运行在任何连接层技术上的多厂商标准。在 1997 年初，IETF 建立的 MPLS 工作组以制定一个统一的、可在多层交换中进行互操作的标准。

多协议标记交换

MPLS 是 Internet 多层交换发展中的最新阶段，它是建立在多种专有多层交换解决方案上的基于 IETF 标准的方案。

MPLS 使用控制驱动模型对用于建立标记交换路径 (LSP) 的标记捆绑的分配及分布进行初始化。LSP 本质上是很简单的 (业务以一个方向从起点流向终点)，全双工业务需要两条 LSP，每条 LSP 负责一个方向上的业务。LSP 是通过连接一个或几个标记交换点来建立的，运行一个分组在 MPLS 域内从一个标记交换路由器 (LSR) 转发到另一个标记路由器。LSR 是一个支持基于 MPLS 转发的路由器。

图 6：多协议标记交换



MPLS 控制部分位于 IP 功能的中心，与专有的多层交换解决方案相似 (见图 6)。但是，MPLS 定义了新的基于标准的 IP 信令和标记分配协议以及现有协议的扩展，以支持多个厂商间的互操作性。MPLS 不采用任何 ATM 论坛的信令和路由协议，因此，避免了对两个不同协议体系进行协调的复杂性。这样，MPLS 为基于分组的 Internet 带来了明显的优势。MPLS 转发部分基于标记交换算法。如果第二层技术支持标记区域 (如 ATMVPI/VCI，或帧中继 DLCI 区域)，则原来的标记区域将封装 MPLS 标记。但是，如果第二层技术不支持标记区域，MPLS 标记则通过位于第二层和 IP 报头之间的标准 MPLS 报头 (见图 7) 进行封装。MPLS 运行任何链接层技术承载 MPLS 标记，因此，其可以从 LSP 的标记交换中获得好处。

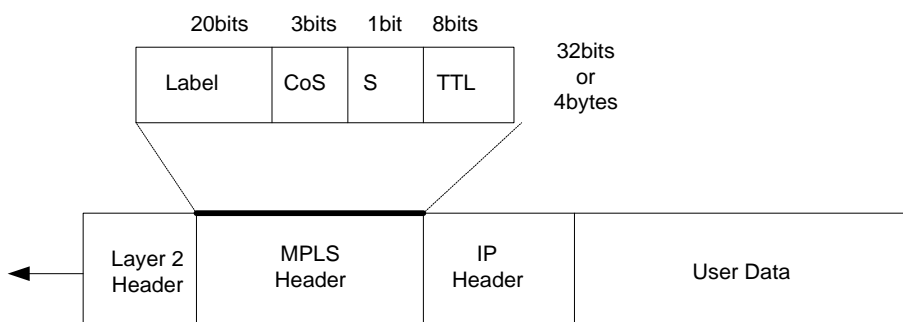


图 7 : MPLS 报头

32 位的 MPLS 报头包括以下区域：

- 承载 MPLS 标记实际值的标记区域 (20 位)。
- CoS 区域 (3 位) ，用于在分组通过网络时施加在分组上的排队和丢弃算法。
- 堆栈区域 (S 区域，1 位) ，用于支持标记堆栈序列。
- TTL 区域 (8 位) ，提供传统的 IPTTL 功能。

需求及目的

MPLS 工作组的目的是标准化一个基本的技术，它综合了具有标记交换的转发部分和具有网络层路由信息的控制部分。为达到这一目的，MPLS 工作组必须提出一个能够满足一系列要求的解决方案，包括：

- MPLS 要能够运行在任何链接层技术之上，并不仅仅局限于 ATM。
- MPLS 核心技术必须能够支持单点和多点传送业务流的转发。
- MPLS 必须能够与 IETF 的集成业务模型兼容，包括 RSVP。
- MPLS 必须能够进行扩展以支持持续发展的 Internet 增长。
- MPLS 至少要能够提供与现有 IP 网络相同的运行，操作和维护的简易性。

对于 MPLS 普遍的错误概念

对于 MPLS 在 Internet 核心网中的作用有许多错误的概念。以下 Internet 业界的人们认为 MPLS 的开发是为允许厂商将 ATM 交换机转移到高性能 Internet 骨干网路由器提供标准。这可能是 90 年代中期的专有多层交换解决方案原来的目的之一，但是，半导体技术的最新发展使仅有 ASIC 的 IP 路由查询引擎的运行速率与 MPLS 或 ATMVPI/VCI 查询引擎相同。尽管 MPLS 能够增强基于处理器相同的转发性能，但提高分组转发性能并不是 MPLS 工作组成立的主要动力。

另一些 Internet 业界的人士认为 MPLS 设计的目的是为了避开传统的最长匹配 IP 路由。这点将永远不会是 MPLS 工作组的目的，因为它们的成员理解 Internet 永远需要传统的第三层路由。

- 防火墙和 ISP 边缘处的包过滤是支持安全性和执行管理策略的基本部分。因为包过滤需要对分组报头进行仔细的检验，对于这些应用仍需要传统的第三层转发。
- 并不是大量的主机系统将执行 MPLS。这意味着主机传输的每个分组仍需要转发到第一跳的第三层设备，那里，分组的报头将在分组被转发之前被检验。然后，第一跳路由器即可以使用传统的最长匹配路由进行转发，也可以对其分配一个标记，沿一条 LSP 进行转发。
- 如果路径中的第三层设备检测 IP 报头并分配一个标记，则标记代表了一个集合的路由，这是因为不可能为整个 Internet 上的每个主机都维持标记捆绑。这意味着，在路径上的某些点上，另一些第三层设备必须检测 IP 报头，为继续转发分组决定更好的间隔。这个路由器可以选择使用传统路由进行转发，或为其分配一个标记，然后沿一条新的标记交换路径对分组进行转发。
- 在目的主机之前的最后一跳，分组必须使用传统的第三层路由进行转发，因为为整个子网上的每个主机都分配一个独立的标记是不实际的。

MPLS 的承诺

留下来的问题是：“为什么 ISP 要在其网络的核心部分配置 MPLS？”MPLS 的最重要的优势在于它允许 ISP 提供现有传统 IP 路由技术所不能支持的新业务。ISP 面临的挑战不仅包括需要提供优异的基础业务，同时也需要提供新型业务以区别于其他竞争对手。MPLS 允许服务提供商控制费用，提供更高等级的基础服务，并提供新的增值客户服务。

图 8 说明了 MPLS 是如何通过支持那些不仅仅需要基于目的路由的应用来提供增强的路由能力的。假设在网络核心部分的路由器提供传统的最长匹配 IP 转发。如果主机 A 和主机 B 都向主机 C 传送一个分组，分组将沿路径 1 通过网络的核心，因为它通过 IGP 计算得到的最短路径。

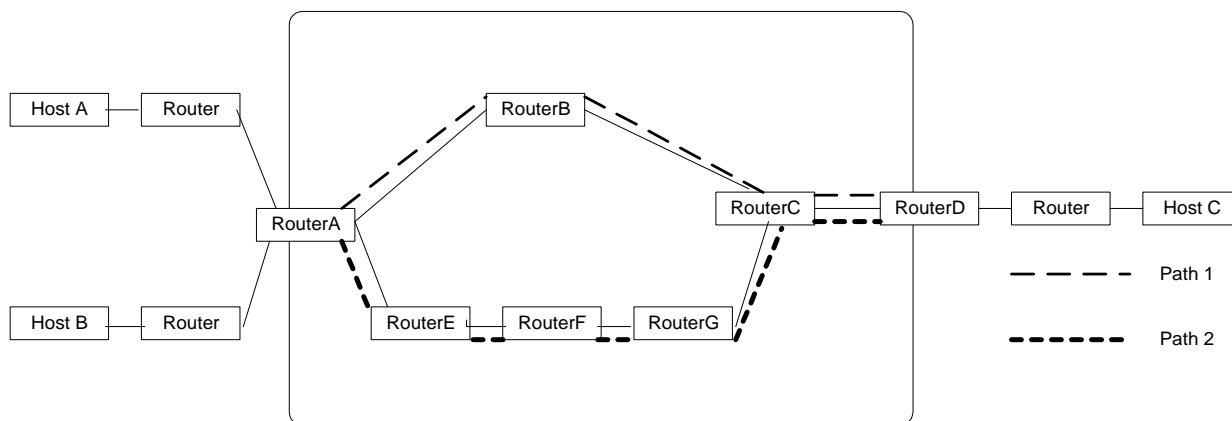


图 8：MPLS 增强路由功能性

假定网络管理员已经对业务参数进行监视，需要对路由器 B 实施策略以控制阻塞。策略将通过将业务负载分散到网络的其它路径来降低路由器 B 处的阻塞。从主机 A 发起、目的为主机 C 的业务将通过 IGP 的最短路径，路径 1。从主机 B 发起、目的为主机 C 的业务将使用另一条路径，路径 2。如果使用传统 IP 路由，将不能实施这种策略，因为路由器 A 处的所有转发都基于分组的目的地址。

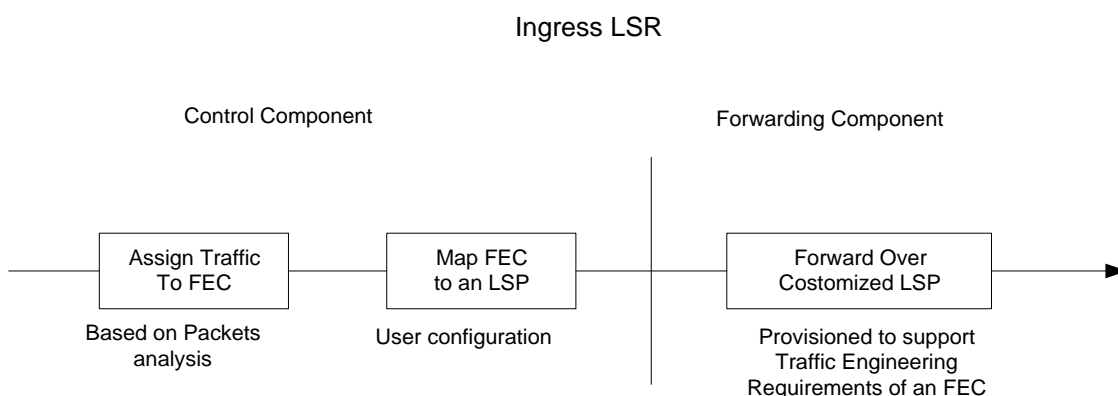
现在，如果网络核心处的路由器执行 LSR 的功能，它将很容易的执行策略以减少 LSRB 处的阻塞。网络管理员配置 LSP1 通过路径 1，配置 LSP2 通过路径 2。最后，网络管理员将所有从主机 A 接收、目的为主机 C 的业务分配到 LSP1，同时 LSRA 将所有从主机 B 接收

到的并将主机 C 作为目的的业务分配到 LSP2。这种将任何 FEC 分配到客户定义的 LSP 上的能力是网络管理员能够在业务流经提供商网络时对其进行精确的控制。

通过仔细的计划，MPLS 为 ISP 提供了一个无与伦比的业务控制级别，使网络能更为有效的运行，支持更多的业务，并能够为支持不停变化的客户需求提供灵活性。您应该注意，余下的部分我们将讨论 MPLS 基于丰富的分组分级能力将业务分配至 FEC 的潜力。最初的 MPLS 实施将提供一套较为有限的分组分级能力，希望其能随着软件执行控制部分的成熟得到发展。

随着 ISP 需要不停的推出新的客户服务，MPLS 转发结构仍能够被保持。新的业务可通过对控制单元进行简单的修改，将分组分配到一个 FEC，然后将每个 FEC 映射到客户定制的 LSP (见图 9)。例如，分组可以基于目的子网和应用类型的组合，源和目的子网的组合，特殊的 QoS 需求，IP 多点传送组，或虚拟专网 (VPN) 标识符被分配到一个 FEC。类似地，网络管理员可以规定 LSP 以满足特定的 FEC 需求 - 使跳转数最小，满足特定的带宽需求，强迫业务通过网络的特定链路及其它。发展路由功能的最后一步是配置起始 LSR，将分配的特定 FEC 的分组放置到满足了 FEC 需求的 LSP 中。

图 9：MPLS 如何增强路由功能性



MPLS 应用

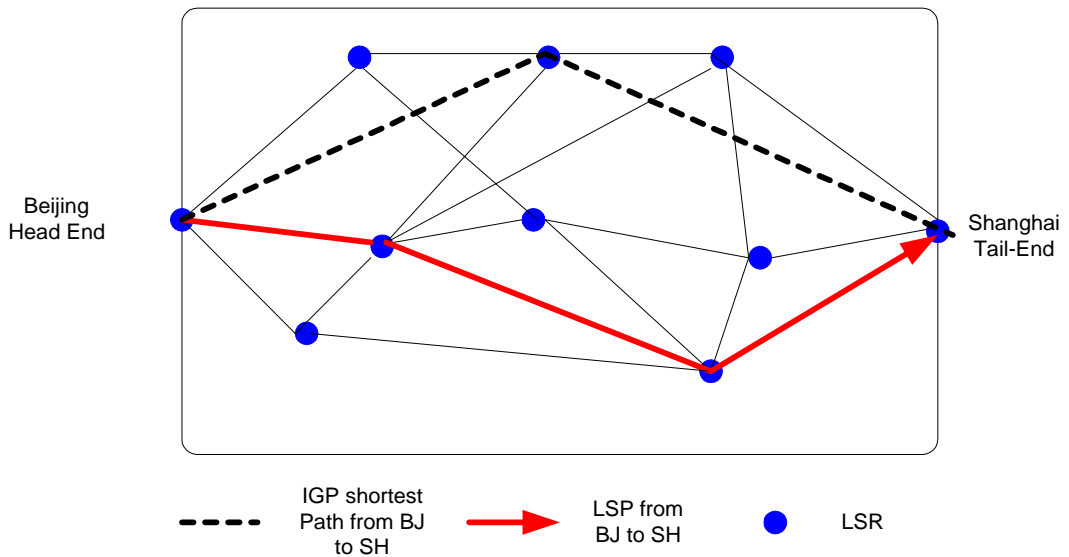
现在，在大型 ISP 网络中对于 MPLS 有三中流行的 应用：

- 流量工程
- 服务等级 (CoS)
- 虚拟专网 (VPN)

流量工程

流量工程是 ISP 能够将业务流从 IGP 计算得到的最短路径转移到网络中潜在的、具有较少阻塞的物理路径上去 (见图 10)。流量工程是 MPLS 的一个主要应用，原因是因为对网络资源需求的快速增长，IP 应用的重要任务属性，和服务提供商市场中不断增加的竞争等。一个成功的流量工程解决方案能够平衡网络中的各种链接、路由器和交换机上的网络汇集业务负载，使这些特定的单元不会被过分使用也不会未被充分利用。这样可以使网络的运行更为有效，并能提供更多的可预测的业务。

图 10：通过服务提供商网络的流量工程 LSP 与 IGP 最短路径比较



MPLS 非常适合于为大型 ISP 网络中的流量工程提供基础，其有以下原因：

- 对于明确路径的支持使网络管理员能够为通过服务提供商网络的 LSP 定义一条精确的物理路径。

- 每条 LSP 的统计参数可用于网络计划和分析工具的输入，用于确定瓶颈和中继线的使用情况并对将来的扩展进行计划。
- 基于约束的路由提供增强的性能，使 LSP 能在建立之前便满足特定的性能需求。
- 基于 MPLS 的解决方案能够运行在基于分组的网络之上，并不局限在 ATM 结构之上。

对于更为详细的 Juniper 网络公司流量工程的介绍，请参考白皮书“*新型公网中的流量工程*”。

服务等级

MPLS 可以为服务提供商开始推出区分服务 (DiffServ) 时提供优势。区分服务模型定义了一系列将业务分配到较少数量的服务等级上去的机制。用户逐渐将 Internet 作为公共传输工具，其包括从传统的文件传输到象语音及视频等对延迟敏感的业务的不同应用。为满足客户需求，ISP 不仅需要流量工程技术，同事也需要业务分级技术。

ISP 一般采用两种方法对基于 MPLS 的服务等级转发进行支持：

- 在业务流经特定的 LSP 时，业务可基于 MPLS 报头中承载的优先级位在每个 LSR 的输出接口出排队。
- ISP 可以在每对边缘 LSR 间提供多条 LSP。每条 LSP 可通过流量工程提供不同的性能和带宽保证。起始 LSR 可在一条 LSP 上设置高优先权，另一条 LSP 设置成中等优先权，将第三条 LSP 设置成最努力业务，第四条 LSP 设置成比最努力业务还要低的优先级业务。

MPLS 为 ISP 在处理不同类型业务时提供了极大的灵活性，使他们能够为不同的客户提供不同业务。优先级为只用于将分组分配到几种服务等级中去。每个服务等级所支持的服务类型是由 ISP 来决定的。

虚拟专网

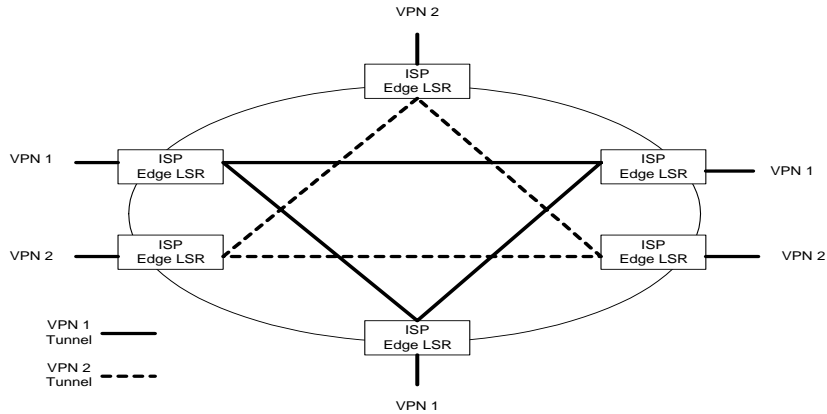
虚拟专网 (VPN) 即在公共 Internet 上模拟专用广域网 (WAN) 的运行。为给客户提供一个可行的 VPN 服务，ISP 必须解决数据保密及支持在 VPN 内使用非唯一的专用 IP 地址。MPLS 能够提供简单而有效的解决方案以解决来自这两方面的挑战，因为 MPLS 的转发决定是基于标记的值，而并不依赖于分钟报头内所包含的目的地址。

VPN 通常由 4 个基本部分组成：

- 防火墙，用来包含每个用户的站点及为 Internet 提供安全接口
- 认证，用于检验每个用户站点之一有效的远端站点交换数据
- 加密，用以保护数据在通过 Internet 时不被检测到或被使用
- 隧道封装，用以提供多协议传输服务及能在 VPN 内使用专有 IP 地址空间

MPLS 允许 ISP 通过提供简单、灵活，而且有力的隧道机制提供 VPN 服务 (见图 11)。VPN 通过提供一组 LSP 来建立 VPN，以在 VPN 内提供不同站点之间的链接。每个 VPN 站点将向 ISP 广播一组本地站点可以到达的报头。ISP 的路由系统通过将标记附着在路由协议更新之上或使用标记分配协议发布这些新型。VPN 识别器允单个路由系统支持具有较迭内部地址空间的多条 VPN。最后，每个输入 LSR 基于包的目的地址和 VPN 关系信息将业务分配到 LSP 中。

图 11：MPLS 方便 VPN 实现



总结

所有 ISP 都面临着对快速增长进行管理挑战。在网络的核心部分，WDM 和更高的光纤可能性，连同客户需求的增长为在专用线路或波长上承载 Internet 业务提供了动机。同时，复用技术，流量工程，和从 IP - over - ATM 模型上获得的性能上的优势必须在多种服务的环境下继续维持。MPLS 在提供 IP - over - ATM 模型所具有流量工程优势的同时还具有其它以下优势：

- 简单的网络设计及运行
- 改进的扩展性能

MPLS 是 Internet 核心多层交换计算的最新发展。它是基于 IETF 标准的实现方法，从多种专有多层交换解决方案中获取了许多优点。MPLS 将转发部分的标记交换和控制部分的 IP 路由、基于标准的 IP 信令及标记分布协议组合在一起。而且，MPLS 可以运行在任何链接层技术之上 - 不仅仅是 ATM 结构 - 从而简化了向基于 SONET/WDM 和 IP/WDM 结构的下一代光 Internet 的转化。

MPLS 的最重要的优势在于它允许 ISP 提供传统 IP 路由技术所不能支持的新型业务。MPLS 通过支持不局限于基于目的转发的技术提供了增强的路由能力。一些可通过 MPLS 实现的新型节约及增值业务包括流量工程，基于 CoS 的转发，及 VPN。通过将控制部分和转发部

分分离，MPLS 提供了发展控制功能的灵活性而不需要改变转发机制，因此，MPLS 可以支持增强的转发能力，以满足 Internet 爆发性增长的需求。

参考

Textbooks

Davie, B., P. Doolan, and Y. Rekhter, *Switching in IP Networks: IP Switching, Tag Switching, and Related Technologies*, Morgan Kaufmann, 1998, ISBN 1-55860-505-3.

Metz, Christopher, *IP Switching: Protocols and Architectures*, McGraw-Hill, New York, 1999, ISBN 0-07-041953-1.

Requests for Comments

RFC 1953, *Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0*, P. Newman, W.L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, May 1996.

RFC 1954, *Transmission of Flow Labelled IPv4 on ATM Data Links Ipsilon Version 1.0*, P. Newman, W.L. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, May 1996.

RFC 2105, *Cisco Systems' Tag Switching Architecture Overview*, Y. Rekhter, B. Davie, D. Katz, E. Rosen, and G. Swallow, February 1997.

RFC 2297, *Ipsilon's General Switch Management Protocol Specification Version 2.0*, P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Ching Liaw, T. Lyon, and G. Minshall, March 1998.

Internet Drafts

Baker, F., and Y. Rekhter, *Tag Switching with RSVP*, draft-baker-tags-rsvp-00.txt, December 1996.

Callon, R., G. Swallow, N. Feldman, A. Viswanathan, P. Doolan, and A. Fredette, *A Framework for Multiprotocol Label Switching*, draft-ietf-mpls-framework-02.txt, November 1997.

Callon, R., A. Viswanathan, and E. Rosen, *Multiprotocol Label Switching Architecture*, draft-ietf-mpls-arch-02.txt, July 1998.

Davie, B., Y. Rekhter, A. Viswanathan, S. Blake, V. Srinivasan, and E. Rosen, *Use of Label Switching With RSVP*, draft-ietf-mpls-rsvp-00.txt, March 1998.

Davie, B., T. Li, Y. Rekhter, and E. Rosen, *Explicit Route Support in MPLS*, draft-davie-mpls-explicit-routes-00.txt, November 1997.

Doolan, P., B. Davie, D. Katz, Y. Rekhter, and E. Rosen, *Tag Distribution Protocol*, draft-doolan-tdp-spec-01.txt, May 1997.

Feldman N., and A. Viswanathan, *ARIS Specification*, draft-feldman-aris-spec-00.txt, March 1997.

Viswanathan, A., N. Feldman, R. Boivie, and R. Woundy, *ARIS: Aggregate Route-Based IP Switching*, draft-viswanathan-aris-overview-00.txt, March 1997.

Next Generation Networks 1998: Conference Tutorials and Sessions

Awduche, Daniel, UUNET Technologies, Inc., "MPLS and Traffic Engineering in Service Provider Networks," Multilayer Switching Symposium, November 6, 1998.

Callon, Ross, Iron Bridge Networks, "Technologies for the Core of the Internet," Conference Tutorial, November 2, 1998.

Downey, Tom, Cisco Systems, Inc., "MPLS for Network Service Providers," Multilayer Switching Symposium, November 6, 1998.

Halpern, Joel, Newbridge Networks, "Technologies for Advanced Internet Services," Conference Tutorial, November 2, 1998.

Heinanen, Juha, Telia Finland, Inc., "MPLS --- ATM Like Functionality for Router Backbones," Multilayer Switching Symposium, November 6, 1998.

McQuillan, John, McQuillan Consulting, "Major Trends in Broadband Networking," Conference Introduction, November 3, 1998.

Metz, Chris, Cisco Systems, Inc., "A Survey of Advanced Internet Protocols," Conference Tutorial, November 2, 1998.

Metz, Chris, Cisco Systems, Inc., "IP Switching in the Enterprise," Multilayer Switching Symposium, November 6, 1998.

Rybczynski Tony, Nortel Multimedia Networks, "Multilayer Switching in Enterprise WANs and VPNs," Multilayer Switching Symposium, November 6, 1998.

Sindhu, Pradeep, Juniper Networks, Inc., "Foundation for the Optical Internet," Next Generation Core Routers Session, November 3, 1998.

UUNET IW-MPLS '98 Conference

Barnes, Bill, UUNET Technologies, Inc., "Traffic Engineering with the Overlay Model," Traffic Engineering Session, November 13, 1998.

Huitema, Christian, Bellcore, "Scaling Issues in MPL Networks," Architecture and System Design Session, November 12, 1998.

St. Arnaud, Bill, Canarie, Inc., "MPLS and Architectural Issues for an Optical Internet," Architecture and System Design Session, November 12, 1998.

Swallow, George, Cisco Systems, Inc., "Traffic Engineering in MPLS Domains," Traffic Engineering Session, November 13, 1998.