高等学校研究生系列教材

Internet 技术及其实现

Internet Technology and Implementation

胡越明

高等教育出版社
HIGHER EDUCATION PRESS

高等学校研究生系列教材

Internet 技术及其实现

胡越明

高等教育出版社

内 容 提 要

本书主要讲授有关 Internet 的最新技术，包括交换技术；用于多媒体业务流的各种 QoS 技术；虚拟专用网络技术，包括隧道技术和安全 IP 协议；网络安全技术，包括防火墙技术和入侵检测技术；尚处于研究初期阶段的主动网技术；网络路由和交换设备的组成原理；Intel 公司的网络处理器及其外围芯片，并以 Intel 公司的网络处理器开发为例，介绍 Internet 网络设备软件开发方法，使读者能具体了解实现 Internet 新技术的方法。

本书的特点是介绍了 Internet 技术的研究前沿，适合作为研究生教材或计算机和通信专业高年级本科生的选修课教材，也可作为从事计算机网络和 Internet 技术的研究和开发人员的参考书。

图书在版编目(CIP)数据

Internet 技术及其实现/胡越明.-北京:高等教育出版社,2003.9

ISBN 7-04-013310-5

Ⅰ.Ⅰ… Ⅱ.胡… Ⅲ.因特网 Ⅳ.TP393.4

中国版本图书馆 CIP 数据核字(2003)第 072527 号

出版发行	高等教育出版社	购书热线	010-64054588
社　　址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010-82028899		http://www.hep.com.cn
经　　销	新华书店北京发行所		
印　　刷	潮河印业有限公司		
开　　本	787×1092　1/16	版　　次	2003 年 9 月第 1 版
印　　张	22.5	印　　次	2003 年 9 月第 1 次印刷
字　　数	470 000	定　　价	28.00 元

编 者 的 话

Internet 的迅速推广应用使其不断面临新的挑战。面对这些挑战，在 Internet 中形成了许多新技术并不断得到发展。这些新技术主要包括各种交换技术、多媒体信息传输技术、虚拟专用通路技术、网络安全技术和主动网技术等。这些技术构成了未来 Internet 的核心技术。为了掌握 Internet 的未来，首先要掌握 Internet 的核心技术。

目前，能够系统介绍 Internet 技术发展现状和趋势的教材比较缺乏，国内高校计算机专业开设的课程也没有较完整地介绍这些 Internet 新技术。在国际各著名大学中，介绍 Internet 新技术的课程也在建设之中。为了使我国在这一技术领域的教学跟上国际水平，作者编写了这本教材。本教材力图反映在 Internet 核心技术领域的最新发展，尽量用通俗的语言向读者介绍这一领域涉及到的概念和方法，避免复杂的公式推导。网络技术的一个特点是大量涌现的新名词、新概念，本书编写过程中力图有序地引出这些名词概念，力图在每个新名词术语第一次出现时就对其做出较明确的解释，使读者能较顺利地掌握这一领域的技术基础。网络技术的另一个特点是复杂的协议，本书主要介绍这些协议的基本思想，而不是罗列其全部内容。Internet 的各种协议都以 RFC 文件的形式描述，可以从网上下载。

本书第一章介绍交换技术，包括 IP 交换、标记交换和 MPLS；第二章介绍用于多媒体业务流的各种 QoS 技术；第三章介绍虚拟专用网络技术，包括隧道技术和安全 IP 协议；第四章介绍网络安全技术，包括防火墙技术和入侵检测技术；第五章介绍还处在研究初期阶段的主动网技术；第六章介绍网络路由和交换设备的组成原理，以及 Intel 公司的网络处理器及其外围芯片；第七章以 Intel 公司的网络处理器开发为例，介绍 Internet 网络设备软件开发方法，使读者能具体了解实现 Internet 新技术的方法。

本书的特点是介绍了 Internet 技术的研究前沿，适合作为研究生教材或计算机和通信专业高年级本科生的选修课教材，也可作为从事计算机网络和 Internet 技术的研究和开发人员的参考书。本书的内容建立在本科生计算机网络课程的基础之上，要求读者具备 TCP/IP 协议的基础知识和数据加密的基础知识。

本教材是作者在上海交通大学硕士研究生教学实践的基础上编写完成的。本书的编写以及上海交通大学 Internet 技术课程的建设得到了 Intel 公司在技术、资金和设备方面的支持，在此向 Intel 公司表示感谢。本书的审稿工作由复旦大学计算机科学系高传善教授完成，在此也表示感谢。

由于 Internet 技术发展迅速，作者的知识能力有限，本书难免存在不足。广大读者如能提出批评和建议，作者在此先表示衷心感谢。

胡越明

2003 年 8 月

目　录

第一章 交换技术

Internet 的迅速发展使得它已经渗透到各个网络通信的应用领域，包括各种文字、数据、语音和视频通信领域，成为一种极具竞争力的网络通信技术。Internet 业务量的增长需要更多的带宽与容量，而传统路由器采用的路由机制根本无法满足这一需求。为此，引入了在电信网络中广泛采用的交换技术。本章介绍交换技术在 Internet 中的发展和应用。

1.1 路由与交换的概念

路由和交换是两种不同的实现网络信息传输的技术。前者是构成传统 Internet 的基础，后者是现代通信技术发展的成果。

1.1.1 路由技术

Internet 采用 TCP/IP 协议。Internet 协议（IP）是一个网络层的协议，它规定数据以分组的形式在网络上传输，实现数据分组在全球范围内的传递。为此，IP 协议定义了网络层的分组格式、IP 地址编址方案（包含在分组中）；定义了对分组转发的机制，即路由器和路由协议；定义了报告分组中的错误的机制，即 ICMP 协议。

Internet 是由路由器构成的网络。IP 协议是一个无连接的通信协议，在网络结点间进行通信之前不需要建立相互连接的关系。一个结点向另一个结点发送数据时只需要在分组中给出接收结点的 IP 地址，路由器将数据分组转发到指定的目的结点。IP 地址是包含在分组中的全球惟一的地址编码，可以由网管员手动分配给主机，也可以由动态主机配置协议（DHCP）动态分配。

在 Internet 中，把数据分组从一个网络转发到另一个网络的过程称为分组转发，或 IP 转发。网络中的**业务流**指的是，网络上从一个特定的源发送到一个目的地的、具有先后关系的分组序列。当分组到达一个路由器时，路由器从这个分组头中提取出目的 IP 地址，与路由表中的表项进行比较。**路由表**是根据网络可达性信息计算得到的一张清单列表，其中列出了各个 IP 地址的转发出口，用于将分组传送到其目的地。分组传送的路由取决于目的结点的位置和网络的拓扑结构。路由表中包括一个目的网络的列表和与之对应的通路中的**下一跳**（Next Hop，即下一个结点）地址。路由器根据分组的目的地址从路由表中查找转发路由中下一跳的端口，并将分组从这个端口发送出去。路由器是一种能够把 IP 分组（或

称 IP 包）从一个网络转发到另一个网络的设备。路由器的每一个网络接口都分别连接到一个互不相同的网络上。路由器之间不断地传递信息，运行路由协议，为业务流寻找通往目的地的路线。所有路由器共同维护一个网络信息数据库，记录各个网络的方位。这个路由数据库根据路由协议建立，路由协议使得路由器之间能够交换有关网络拓扑和可达性的信息。

在转发方式上，IP 路由是无状态的。每个路由器独立地对 IP 分组进行转发，转发每一个分组都要重新查找路由表，转发完成后不保留与该分组有关的任何信息；分组出错时也只是简单地将分组丢弃。因此，IP 技术具有简单灵活的特点，特别适合于少量数据通信的场合，这是它在早期能得到广泛应用的一个重要原因。此外，IP 分组是变长的，可以方便地在其中增加附加信息；IP 路由表是单向的，双向传输可以通过不同的路径；路由表是在路由协议的控制之下动态改变的，每个路由器独立自主地进行路由选择；IP 网络既可以实现点到多点的组播，也可以实现多点到多点的组播。

IP 分组的头信息如图 1-1 所示。其中版本字段的值为 4，表示 IPv4；IHL 表示分组头长度；服务类型用于提供不同特征的服务级别；总长度字段标识整个 IP 分组的字节数，包括分组头；标识字段表示分组在流中的编号，用于分组片断数据报的重组；标志字段用于控制分组的分割；段偏移字段用于该分组的数据在整个数据报中的相对位置；生存时间（TTL）用于进行跳计数，以避免分组在网络中循环；头校验和用于传输中的纠错。IP 分组的头部包含了足够的用于路由选择的信息。

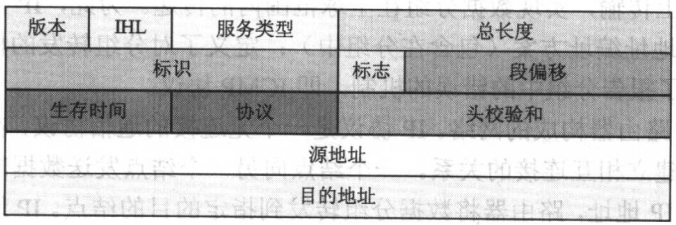

图 1-1　IP 分组头的信息结构

每个 IP 分组通过一个链路层的分组进行传输。IP 分组的大小应当适合于在数据链路上的传输。一条传输链路所能支持的最大分组大小称为该链路的最大传输单元 **MTU**（Maximum Transmission Unit），IP 分组应小于或等于 MTU。当大于链路 MTU 的分组到达时，分组就被分割成较小的分组，被分割的分组在到达目的地后再重新装配成原来的分组。IP 分组头中的标志字段和段偏移字段用于分组的重新装配。

在无连接的网络层协议中，数据分组从一个路由器传递到另一个路由器，每个路由器都独立地对该分组作出转发的决定，即每个路由器对分组的头信息进行分析，运行一个网络路由算法；每个路由器根据分析和算法独立地选择一条出口通路作为分组的下一跳链路。

在IP路由系统中，整个Internet被表示成大量**自治系统**（**AS**）的集合，每个自治系统是某个网络建设和管理实体所维护的网络系统，如大型企业网络或者它们的集合。在一个AS之内的路由方式由域内路由协议确定，跨越AS的路由则由域间路由协议确定。域内路由协议用于在同一个AS的各个路由器之间计算分组转发通路，又称为内部网关协议IGP，如RIP、OSPF；域间路由协议用于在不同AS之间计算分组传输需要经过哪些AS，又称为外部网关协议EGP，如BGP。在域内路由协议中，确定分组在域内各个结点之间的传递路径，路径的计算以一个特定的衡量指标的最优化为基础。在RIP中，这个指标是跳数，即分组转发过程中经过的路由器数量。在有多条路径可供选择时，路由协议使用Bellman-Ford算法选择跳数最少的路径。在OSPF中，使用将路径的管理衡量尺度最小化的Dijkstra最短路径优先（SPF）算法来计算路径。路径的管理衡量尺度为沿此路径的所有链路上管理衡量尺度的总和。域间路由协议用于确定分组的传递所经过的AS序列。BGP协议不使用域内路由的距离向量算法或者链路状态算法，它是一个通路向量协议。由于各自治系统可以有各自不同的路由衡量尺度，BGP不能简单地根据各个自治系统的衡量尺度来计算路由。BGP不传递路由衡量尺度的信息，只传递路由通路信息，如通向目标AS所经过的一系列AS。BGP的路径信息只是指定AS，而不是指定AS中的路由器，因为每个AS可以自主地决定由哪个路由器来处理域间的路由。

路由协议必须满足以下要求：

(1) 动态发现网络的拓扑结构。建立最短通路转发树，能处理关于外部网络的信息，这些信息可能采用与本地网络不同的衡量标准。对网络的拓扑结构变化作出反应，更新最短通路树。

(2) 可伸缩性。即网络扩展性，减少路由器之间的路由信息传递，减少计算路由表所消耗的计算资源，在网络规模增大时能够支持网络路由。

(3) 避免环路。当计算分组转发路径时，能够防止出现环路。

(4) 协议可扩展性。路由协议在不改变其基本操作和兼容性的情况下能够加入新功能。

典型的路由器要对每一个分组执行基于目的地址的路由表查询，以及TTL递减、分组头校验和介质转换等操作。在典型的路由器中，这些操作是基于软件实现的。有效的业务流转发不仅依赖于路由器自身的性能，而且依赖于最佳的路由选择。

网络中的每一个路由器都包含控制功能器件和转发功能器件。控制功能器件由协议组成，提供路由信息的交换以及路由表的更新操作。转发功能器件由算法或程序组成，对分组进行转发决策。转发功能包括单播、带服务分类的单播以及组播。

在路由器中，到达的数据分组被归结到有限的子集中，路由器以相同的方式处理每个子集中的数据分组，这种子集称为**转发等价类**（**FEC**）。转发等价类中目标地址的数量称为该FEC的**粒度**（Granularity）。在一个转发等价类中的数据分组尽管可能具有不同的网络层属性，但从转发的角度它们是等价的，在路由表中用同一个表目描述。转发等价类将

不同的业务流汇聚起来，形成了较粗的转发粒度，减少了路由表的长度。但是转发粒度太粗会使得网络缺少灵活性，因为它不能区分不同类型的业务流。业务流的汇聚可以大大缩短路由表的长度，减少路由信息的交换和路由表查找时间，提高路由协议的可伸缩性。

早期的 Internet 中把 IP 地址划分成 A 类、B 类和 C 类，IP 地址的类型可以通过高 3 位进行判定。IP 地址的其余部分划分为网络地址和主机地址，A 类、B 类和 C 类网络中的主机地址字段分别为 24 位、16 位和 8 位，分别对应于 4 百万、64 万和 256 个主机的网络。这种地址格式不便于网络的划分，因为绝大部分网络的规模不完全符合这种规模类型，从而在 IP 地址分配时将大量浪费地址资源。为此提出了**无类域间路由（CIDR）**方式，使用 IP 地址（X）加前缀（Y）的方式表示各种规模的网络以及转发等价类。路由器对每一对 X/Y 值创建一个 32 位的屏蔽字，其最高 Y 位的值为 1，其余位为 0。将这个屏蔽字作用于分组的 IP 地址，将未屏蔽的部分与路由表中的 X 比较，如果结果相等（匹配），则将该分组转发到路由表输出字段指定的网段，继续查表。如果路由表中没有匹配的表项，则丢弃该分组。这种 IP 地址加前缀的表示方法有利于更加合理地分配 IP 地址。CIDR 取消了地址分类的分配方式，可以任意设定网络号和主机地址字段的边界，即根据网络规模的需要可以重新定义地址屏蔽字。

CIDR 能够支持路由的汇聚。在基于目的路由的无类域间路由网络中，一个转发等价类与一个地址前缀相关联，这样一来，一个路由表项就可以代表大量传统分类路由所表示的地址范围，从而可以限制路由表的增长，加速路由表查询的速度。通过使用单播路由协议（如 OSPF、RIP、BGP）提供的信息，路由器可以在 FEC 和它们的下一跳之间建立映射，并使用这种映射进行分组的转发。在基于目的路由的网络中，路由选择是两种功能的组合：第一种功能是将各种可能的分组划分成一组转发等价类 FEC，第二种功能将每一种 FEC 映射为对下一跳结点的选择。从转发决策的角度来看，划分到同一 FEC 的各种分组是不可区分的。同一个 FEC 中的分组都沿着相同的通路转发。

在无类域间路由的 IP 转发中，一个 IP 地址可以在路由表中找到多个匹配的表项，这些表项分别对应包含该目的地址的不同颗粒度的等价类。路由器在进行分组转发时采用最长前缀匹配的原则。如果在路由器的路由表中有一个地址前缀 Y，而且 Y 是每个分组的目标地址的最长匹配，则这个路由器通常将这两个分组放在同一个 FEC 中。所谓目标地址的最长匹配就是在匹配的目标地址中选择一个 Y 值最大的。因为具有较大 Y 值的网址前缀的路由信息所描述的 IP 目的地址的集合较小，或者说目标地址更加详细具体，等价类的颗粒度更小，所以最长地址匹配的路由方式能够将分组传送到更具体的目标网络或结点。当分组在网络上传输时，每一跳的路由器都要重新检查这个分组，将其归入某一个 FEC 中。

在 IP 协议之上有 UDP 协议和 TCP 协议等。UDP 协议是一个简单的传输层协议，它在 IP 协议之上增加了复用/解复用功能以及一些简单的纠错功能。为了实现复用/解复用功能，它加入了源端口号和目标端口号，为了实现纠错功能，它加入了长度和校验和的信息。端口

号与 IP 地址一起，可以将分组转发给指定主机上的指定应用程序进程。在 UDP 的发送和接收过程中没有使用握手机制，它是一种无连接的通信机制。在 UDP 通信机制中，主机上无需保存通信的状态信息，分组头的信息量较少，分组发送的速率不受约束，没有拥塞控制机制。建立在 UDP 上的应用类型有网络文件系统、多媒体流、IP 电话、网管协议等。

TCP 协议是建立在 IP 协议之上的传输层协议，它建立起面向连接的双向传输通路，因而能提供可靠的、按序传输数据的服务。TCP 协议提供复用/解复用和纠错功能。在传输数据之前，双方必须先进行握手，即通过交换分组建立数据传输的参数。TCP 协议只在通信的双方主机上运行，不需要在中间结点上运行。TCP 的连接是一种点到点的连接，不能建立点到多点的连接关系。

TCP 是一种面向流的协议，它将数据看成一个字节流，发送方不需要考虑发送业务的长度。TCP 能够将发送业务分成小段进行传输，并能够对丢失的某一段进行重传，能够对发送过程中打乱了顺序的各段进行重新排序。每一小段的大小通常根据 IP 分组的大小进行选择，以避免分片。每一小段数据封装在一个 IP 分组中，并加上 TCP 头，构成 TCP 段。TCP 协议将数据看作一个无结构的有序数据字节流。TCP 头中的段号字段指出段中的一个字节在字节流中的序号，接收方收到数据段后返回等待接收的字节号。

TCP 协议能够对业务流提供流量控制，以避免分组在网络中发生拥塞现象。即当网络畅通时增加数据分组的流量，当网络不畅通（拥塞）时降低分组的流量。它采用可变发送窗口，通过调节窗口的大小来调节数据流量。发送窗口指定了可发送的分组序列范围，这个范围的初始大小在建立连接时由双方商定，并在业务流的传输过程中进行动态调整。判断网络拥塞的依据是分组传输在网络中的传输时间过长，或者分组被丢弃。建立在 TCP 协议之上的应用类型通常是要求可靠传输的业务流，如电子邮件、远程终端、Web 浏览、文件传输等。

IP 网络通常采用边缘-核心模型，将路由器分成边缘路由器和核心路由器两类。边缘路由器将各种不同的局域网络与网络核心部分连接起来，接收进入网络的用户信息，处理各种网络协议。为此，边缘路由器应当具有多种不同网络的接口，支持多种不同的网络协议。核心路由器以提供高传输带宽为主，在其他核心路由器以及边缘路由器之间提供分组的转发服务。核心路由器与其他路由器的连接接口可以是某一种统一的网络链路接口，可以采用硬件的实现方式以提高分组转发速度，可以只支持少量的网络协议。

当网络的流量迅速增长之后，传统的路由技术暴露出以下两个主要问题：

(1) 路由器的处理能力低，速度和吞吐量较低，用户接入速度太低。主要原因是数据分组的存储和处理过程复杂，路由表太长导致查询时间长。虽然可以通过网络设备的升级来解决，但是受到它的高成本限制。

(2) IPv4 协议对实时业务、灵活的路由机制、流量控制和安全性的支持不够。当用户数增加时，网络的某些区域会发生拥塞，无法保证可靠的服务质量。

1.1.2　交换技术

分组交换方式是源自通信技术中的一种方式，它对每个输入的分组进行缓存直到该分组输入完毕，然后根据分组头查找转发表，对输入的分组头进行转换，构成输出分组，再输出这个分组。采用这种交换方式的有 X.25、帧中继和 ATM。分组交换方式的基本过程是一种异步时分复用的过程，由交换设备在输入和输出之间建立对应关系。

ATM 是一种不同于时分交换模式的新型信息传输模式。它采用统计复用方式，不再把时隙固定地分配给某一个业务流。它采用 53 字节的固定长度的信元，其中 5 个字节为信元头，其余 48 字节为承载的数据。只要网络链路中时隙有空闲，就可以传输任何一个业务流的信元，业务流之间根据其信元头的标志进行区分。这种技术可以动态分配网络的带宽资源，适合于传输速率可变的业务流。ATM 不需要在每一个中间结点为每个信元确定路由。ATM 的信息交换方式是面向连接的和有状态的。在 ATM 网络中，对每一个业务流都要预先建立起端到端的连接关系。在建立连接时，每一个 ATM 交换机都为该连接分配一个标识符，该标识符对应了输入输出的端口。ATM 交换机通过一个**转发表**建立一个连接及其与输入输出端口的对应关系。

ATM 支持永久、半永久和动态的虚电路（VC）以及虚通道（VP），可为不同的业务流类型提供不同的传输服务质量要求。VC 是两个或多个终端系统之间的传输信元和专用逻辑信道。VP 是一组 VC，可以是永久的（PVP）或者交换的（SVP）。每个 VC 都属于一个 VP。在网络中，VP 是进行大量 VC 交换的一种高效手段。信元头中包含 8 位或者 12 位的虚通道标识符 VPI 和 16 位的虚电路标识符 VCI。此外还有净荷类型指示（PTI）、信元丢弃优先级（CLP）和信头差错检验（**HEC**）字段等。

在 ATM 网络中，连接关系的建立通过人工配置手段或者信令协议建立，反映在各个交换机结点中的转发表的对应表项中。在信元传递经过的每一个结点上，需要为建立这种 VP、VC 虚连接分配对应的标识符 VPI 和 VCI。同时还要为虚连接进行网络资源的分配和其他必要的配置。通过这种方式，ATM 就可以对业务流参数提供严格的保证。此外，ATM 信元是固定长度的；ATM 连接是双向的，而且是静态的；ATM 网络只能实现点到多点的组播，不能实现多点到多点的组播。在这些方面，ATM 与 IP 有着很大的区别。VCI 是每个结点为一条虚电路分配的标识，只具有局部意义，它作为在这条虚电路上传输的信元的标识。每个交换机为一条 VC 确定其本地的 VCI 值，并且保证与它直接相连的链路具有惟一的 VCI 值。这样，VCI 在每一段链路上具有惟一的值。当信元在一条链路上通过时，每经过一段链路，VCI 的值都在改变。

在 ATM 中，包含在信元中的路由信息是虚电路的标识符，而不是完整的目标地址，这样可使得信元传输路径的表示更为简短，也更适合于信元的交换。ATM 交换的过程是：当交换机在输入端口上发现输入信元具有标识符 m 时，就查找转发表中对应于该输入标识

的表项，得出其输出端口 j 和输出标识符 n，然后把该信元标以 n 标识符并送到输出端口 j，下一个交换机就使用这个新的标识符。这里的标识就是信元中的 VCI。ATM 的转发表必须预先建立，建立转发表的过程就是建立连接的过程。ATM 的这种交换方式如图 1-2 所示，图中到达交换机 1 号端口的信元标识为 A，从转发表中查出：从 1 号端口进来的标识为 A 的信元应向 3 号端口转发，并加上标识 C。

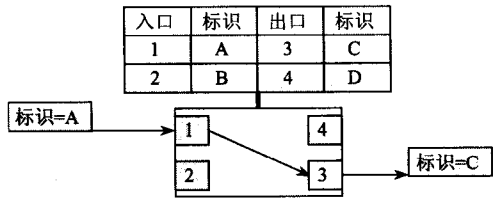

图 1-2 ATM 的交换方式

在这种交换方式下，交换机需要建立一个由输入端口号、输入标识符、输出端口号和输出标识符构成的转发表。建立连接的过程就是在转发表中建立相应的转发项的过程。当一个信元到达交换机时，其输入端口号和标识符就被用于在转发表中搜索匹配的项，然后根据查到的输出端口号和输出标识符进行转发。这个转发过程可以用硬件快速完成。在一个信元的传输过程中，每经过一个交换机，它的标识符就被改变一次，一个标识符只在两个交换机之间使用，所以只具有局部的意义。标识的这种局部性使得网络中可以存在相同的标识，增加了网络中可供使用的 VCI 的值的数量，或者说减少了 VCI 的长度，从而保证了网络的可扩展性。

交换机的优势在于成本和性能上的因素。与局域网结合的交换技术可以比用传统的 LAN 技术低得多的成本来提供高得多的性能。交换技术的第一个特点是工作在协议栈的第二层以下，分组的交换对于网络层及以上的协议是透明的。交换机转发信元的过程比路由转发分组的过程简单得多，因此可以由硬件完成。交换机把一个输入端口上接收到的数据分组发送到一个输出端口，这个过程可以不需要处理器的干预，具有比网络层处理方式更高的性能、更低的延迟和更低的成本。集成电路技术的发展将可以把更多的网络处理功能集成到一个廉价的芯片中。这为交换技术带来广阔的发展前景。

将转发机制与路由建立机制分离是交换技术优势的一个关键。转发机制是一种简单的机制。随着网络的扩大，路由机制必然越来越复杂，路由表的查找必然越来越复杂。在路由器中，对每一个分组都实施路由机制，使得它的效率大大降低。而交换技术中仅对一个业务流进行一次性路由建立的过程，对于分组的转发则采用简单高效的转发机制，提高了转发的速度。

ATM 的这种连接建立机制是一种信令机制。除了信令机制外，连接通路的建立还可以通过地址学习、生成树、广播和发现、链路状态路由等机制建立。在信令机制下，在将数

据分组进行转发之前，发送方必须先通过网络向接收方发送一条建立连接的请求消息（信令）。这条消息建立 ATM 连接的一套参数，这些参数指定了连接的特性，包括信息传输速率、可接受的最大延迟等。当数据发送方收到接收方的一个应答并确认存在一条有足够容量的通道的消息（信令）时，才开始发送数据。建立连接之后，数据分组都沿着建立的交换通路到达接收方，不再需要计算路由，而且能够保证通路有足够的带宽和一定的时间延迟进行数据的传输。

ATM 技术由于是从通信技术发展而来，具有高速度、低延时和延迟时间均匀等传统路由器网络所不具备的性能，已经得到了广泛的应用。ATM 可以承载任何业务流，包括数据、语音和图像等。ATM 将这些信息分成信元，通过预先建立的通路进行传输。ATM 在几个方面与 IP 路由技术不同。ATM 交换机中，信息的转发速度比传统的路由器快得多，因为它的转发表比路由表小得多，而且由硬件实现。ATM 技术的问题是技术复杂，连接方式不够灵活。由于 Internet 应用主要以突发性业务为主，协议特征与 ATM 有很大区别，所以 ATM 无法直接支持 Internet 的应用。为此，人们提出了将 IP 技术与 ATM 技术相结合的思想，如把 IP 协议建立在 ATM 的基础上，代替 ATM 交换机中的信令和路由协议，形成采用 IP 协议的交换机。这种交换机执行 IP 路由协议，进行传统的逐级跳方式的 IP 分组转发。当检测到一个大数据量的业务流时，为其分配一个虚通道和虚信道（VPI/VCI）来进行传输。这种技术包括东芝公司的 CSR、Ipsilon 的 IP 交换、Cisco 的标记交换、IBM 的聚合的基于路由的 IP 交换（ARIS）。

1.1.3　IP 技术与 ATM 技术的结合

在 Internet 中，IP 协议有着不可动摇的地位；而在远程通信领域，ATM 等交换技术则具有技术优势。因此，如何将这两种技术相互结合，形成高效的能满足各种服务要求的网络，成为人们探索研究的一个课题。

在 Internet 中，使用以 ATM 为代表的交换技术的优点是：

(1) 在 ATM 交换技术中，数据传输在第二层进行，从路由器的角度看，整个 ATM 交换网络是一个单跳网络，只有边缘设备进行网络层协议处理，可提高路由处理效率。使用 ATM 交换结构的 IP 网络将会因为路由器跳数的减少而获得较低的成本，网络的复杂性也随之降低。

(2) ATM 信元交换特性以及硬件交换能力使得 ATM 交换机具有极高的端口速率和较低的网络传输时延，可以为 IP 骨干网络的数据传输提供可靠的保证。

(3) 降低设备成本。交换机能比路由器提供更多的端口，同时使每个端口的费用更低。

(4) 交换设备中可以根据业务的需要为不同的数据流分配不同的转发优先级、带宽等服务质量参数，传输质量更好，并且可以定义多个虚连接。

(5) 交换机具有比路由器更高的可靠性。

(6) 简化网络的管理。因为交换机只是链路层的设备，不占用 IP 地址，网络拓扑结构的配置也十分灵活。

IP 技术与交换技术的结合方式有**叠加模型**和**对等模型**两种。叠加模型是在交换网络之上叠加 IP 网络，将 IP 建立在交换协议之上，两层协议完全独立，它们之间通过一系列中间协议实现互通。叠加模型的网络由运行 IP 路由协议、具有 IP 地址的 IP 设备和运行交换协议的信令及路由协议的交换机组成。这种网络可以利用和运行大部分原有的标准协议，具有独立寻址的特点，运行独立的路由协议。交换网络提供一个高速连通的核心网络，网络边缘采用路由器连接用户的网络，由交换技术中的虚电路互联的一组路由器组成的 IP 网络提供 IP 数据包的转发功能。在叠加模型的开发方面，ATM 论坛和 IETF 组成了一些工作组，开展了一系列的研究开发。这些工作组包括：

(1) ATM 上的 IP（IPATM 或 IPOA）工作组。该工作组定义 IP 数据包携带在 ATM 适配层的 PDU 的封装方式，定义 ATM 地址到 IP 地址的解析协议（ATMARP）。ATM 上的经典 IP 最初在 RFC1577 中描述，在 RFC2225 中作了修改。RFC1932 是 IPOA 的框架文档，RFC2331 描述了支持 IPOA 的 ATM 信令。此外，RFC1483 描述了如何将 IP 分组封装在 ATM 信元中。这种方式需要使用将逻辑子网的 IP 地址转换到 ATM 地址的 ARP 服务器，不同逻辑子网的通信只能通过路由器。

(2) 大型公用数据网络上的 IP（IPLPDN）工作组和大型云上路由（ROLC）工作组。该工作组定义下一跳解析协议（NHRP），使得相隔很远的主机和路由器能够穿越 ATM 网络建立直接的虚电路（RFC2332）。为了支持组播，还定义了组播地址解析服务器（MARS）和组播服务器（MCS），有关文档是 RFC2022。

(3) LAN 仿真（LANE）工作组。这是 ATM 论坛的一个工作组，该工作组定义在 ATM 上仿真 LAN 的程序。LANE 仿真了一个物理 ATM 网络之上的典型的广播式媒体 LAN 的功能和行为，从而可以支持 IP 等网络协议。LANE 提供的功能包括：数据封装和传输、地址解析以及组播的管理。但是，LANE 只能支持较小规模的网络。

(4) ATM 上的多协议（MPOA）工作组。这是 ATM 论坛的工作组，该工作组联合并扩展其他工作组的工作来支持多个网络层协议，而不是只支持 IP 协议。MPOA 是 LANE 和 NHRP 的结合，对 LANE 进行了改造，采用 NHRP 协议和流检测机制以构造不同 LANE 之间的捷径。RFC2022 描述了 MPOA 协议。

(5) 特定链路层上的集成服务（ISSLL）工作组。该工作组定义将 IP 的资源预留模型映射到 ATM 的资源预留模型上的程序。

在这种叠加模型中，共享同一地址前缀的 IP 主机和路由器组合成一个逻辑 IP 子网，物理上与同一个 ATM 网络相联。同一子网的 IP 主机之间可通过 ATM VC 直接通信，不同子网间通过 NHRP 提供 ATM VC 连接来实现 IP 主机间在 ATM 网络上直接通信的能力。ATM 到 IP 地址的转换通过 ATMARP 协议进行。叠加模型将 IP 建立在 ATM 之上，两层