

路由器新技术丛书

路由器

原理与设计

◆白建军 卢泽新 编著

人民邮电出版社
POSTS & TELECOMMUNICATIONS PRESS

路由器新技术丛书

路由器原理与设计

白建军 卢泽新 编著

人民邮电出版社

图书在版编目(CIP)数据

路由器原理与设计/白建军, 卢泽新编著. —北京: 人民邮电出版社, 2002.7

(路由器新技术丛书)

ISBN 7-115-10208-2

I. 路... II. ①白...②卢... III. 计算机网络—路由选择 IV. TN915.05

中国版本图书馆 CIP 数据核字(2002)第 030401 号

内 容 提 要

本书从路由器的基本原理出发, 结合工程实践和国际上路由器技术的发展现状与趋势, 深入浅出地介绍了高性能路由器设计的一系列关键技术和实现难点。主要内容包括: 高性能路由器的体系结构、高速接口、系统软件、路由体系及路由协议栈等方面的设计与实现技术, 同时对路由表问题、高级交换技术以及路由器的安全体系和 QoS 作了深入介绍。最后还给出了当前国际上一些主流高端路由器产品的主要技术指标。

本书内容详尽, 是一本当前少有的详细介绍路由器设计技术的书籍, 既适合于具备一定网络基础知识的网络工程师、高等院校网络专业师生阅读, 也可供网络数据产品研发及相关人员参考。

路由器新技术丛书

路由器原理与设计

-
- ◆ 编 著 白建军 卢泽新
责任编辑 梁 凝

 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
读者热线 010-67180876
北京汉魂图文设计有限公司制作
北京顺义振华印刷厂印刷
新华书店总店北京发行所经销

 - ◆ 开本: 787×1092 1/16
印张: 13.5
字数: 323 千字 2002 年 7 月第 1 版
印数: 1-5 000 册 2002 年 7 月北京第 1 次印刷

ISBN 7-115-10208-2/TN · 1857

定价: 24.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

丛书前言

伴随着世界经济格局在世纪之交的大调整和新技术革命带来的更大的挑战，网络经济，特别是与 Internet 相关的经济，已经成为知识经济时代的主要模式。互联网络成为加快和改变人类信息与知识流通的巨大推动力，信息高速公路的出现更是消除了地域和时间上的差异，成为各种经济行为的基础。同时，随着 Internet 网络技术的飞速发展，各种形式的网络应用深入到千家万户的日常生活中，人们对主要的网络互连设备的性能、网络安全性以及稳定性的期望越来越高。

作为 IP 网络的核心，路由器技术，特别是高性能路由器技术已经成为当前网络领域研究的热点和重点，越来越多的研究机构和商业团体开始重视路由器技术的发展。

在我国，虽然路由器技术起步较晚，但是也已经取得了可喜的成绩。具有自主知识产权的高端路由器于 2001 年初的问世，不仅标志着我们的网络领域已经赶上了国际上的最高水平（某些指标甚至已经是国际领先水平），也极大地激起了网络科研工作者和业界的自豪感与使命感，业界人士已经开始广泛关注路由器技术。为了满足广大网络工作者以及在校学生的需求，受同仁委托，我们组织编写了这套“路由器新技术丛书”。

本丛书以路由器技术为主线，涉及到 Internet 路由方方面面的知识和技能，包括路由器的原理与设计、路由器的测试、路由协议和 MPLS 技术等众多领域。丛书共有以下 5 本。

1. 《Internet 路由结构分析》
2. 《路由器原理与设计》
3. 《高端路由器测试》
4. 《新型的骨干网路由平台——MPLS》
5. 《组播路由协议设计及应用》

本套丛书的作者大都是在网络领域从事研究与工程实践的工作人员，他们对路由器以及相关技术有深入的研究和理解，广大读者可以从他们的实践中汲取很多其他途径很难获取的经验和知识。

相信本套丛书会为那些想深入学习和理解高级网络技术和广大技术人员、工程人员以及相关专业的在校学生带来难得的惊喜！

作者
2002 年 1 月

9501 X 10

前 言

随着 Internet 的迅猛发展，一个专门为解决 Internet 骨干网运营商所面临的特殊问题的网络设备——路由器随之出现。

传统的 IP 网络在传输实时业务时的主要瓶颈在于路由器速度太慢，时延和时延抖动太大，不能满足 QoS 要求，因此研制核心路由器的任务显得非常重要。传统路由器采用软件实现路由识别、计算和包转发，但即使是昂贵的高档路由器的包转发速度也远远低于网络交换机的速度。自 1997 年下半年以来，一些公司开始陆续推出采用硬件专用电路(ASIC)进行路由识别、计算和转发的新型 Internet 骨干路由器。

骨干路由器不仅能够集合带宽及业务的吞吐量，而且还要具备丰富的软件功能及控制特性。为能同时实现增加带宽及丰富的软件功能，新型路由系统的设计者必须在设计中采用以前只有在交换机中才具有的转发功能。但是，相对于为固定长度数据包提供线速性能的直接交换而言，实现应用最长匹配查询的不定长度数据包线速路由与转发要复杂得多。特别是，这种数据包的处理已不能通过基于微处理器或微处理器辅助的方式来实现，而必须使用一种基于随 Internet 环境变化而改变其路由策略的专用集成电路芯片（ASIC）的方式完成。这种新型的 Internet 骨干路由器所面临的系统上的挑战也是非常复杂的。它们不仅要能适应现有的 OC-12 及 OC-48 速率的骨干网和相关的 Intra-POP（Intra Point of Presence，内部层现网络）结构，而且还要能够支持具有 OC-48 速率和 10Gbit/s 以太网 Intra-POP 结构的 OC-192 速率的骨干网络。另外，新的路由系统必须要加速 Internet 从“尽力服务”模式向最基础的可靠性服务的方向转变。Internet 用户希望能够像使用公共电话网络那样，在任何需要通信的时候，听到“拨号音”，便可得到高质量的服务。

本书以高端路由器技术为主，介绍路由器的原理与设计，全书共分 10 章。第 1 章从路由器技术基础开始谈起，介绍了路由器的基本组成、工作机制、路由器的分类以及高端路由器的关键组成。在第 2 章中，介绍了路由器技术的发展与当前主要的路由器实现标准。第 3 章概括阐述了路由器实现过程中的一些关键技术。从第 4 章开始具体介绍实现细节，包括报文交换体系结构、路由器软件系统、高速硬件接口设计技术以及路由协议的设计与实现技术。作为一个独立部分，第 8 章着重阐述了 Internet 骨干网的路由表问题，这不仅仅是因为随着 Internet 的急速增长，路由表容量对路由器路由转发的性能影响越来越严重，而且也因为路由表的查找与维护也是路由器设计过程中的一个难点。最后两章概要介绍了高级交换技术、路由器安全以及 QoS 技术，这些技术必将在未来的高端路由器中成为主流，也是当前网络领域的研究重点。由于路由器技术，特别是高端路由器技术涉及的领域繁多，比如路由器中实现组播以及光路由等新技术，因此在一本书里不可能给出完整的介绍和阐述。我们只希望本书能够为网络技术的研究与普及作出微薄贡献，其他未能涉及的领域请参阅相关书籍和网络资源。

本书参考了大量的网络资源以及学术论文，还有很多信息来源于论坛，国际上一些网络

设备公司的技术支持中心、一些国际性标准化中心组织的公开文档。另外，作者在国家“863-300”核心路由器项目的测试和验收期间，与 Agilent 中国公司的梁勇高级工程师、信息产业部电信传输研究所的卢彧、袁琦、武静、魏亮、吴江等同志、大唐电信集团的沈祺工程师、巨龙公司的骆璐以及网通集团的吕东芳等工程师的讨论，为本书积累了大量的素材，在此对他们一一表示感谢。感谢王宝生老师、朱培栋博士、孙志刚博士、彭伟博士为作者的工作和学习给予的指导和关心。

本书由白建军组织编写，第 6 章由张锐执笔，其余各章由白建军执笔，最后由卢泽新教授审阅全书并定稿。王乐春、朱珂、张小哲、张明杰等博士生也为本书提出了很多建议。书中大部分图表都由张锐绘制。

由于水平所限，加上时间仓促，书中欠妥乃至错误之处在所难免，恳请广大读者批评指正。意见、建议和批评请发至电子邮件：jianjun_bai@163.net，作者将不胜感激。

作者
2002 年 3 月

目 录

| | |
|-------------------------------------|----|
| 第 1 章 路由器技术概述 | 1 |
| 1.1 引言 | 1 |
| 1.2 路由器的组成及其功能 | 2 |
| 1.2.1 路由器的基本组成 | 3 |
| 1.2.2 路由器的硬件组成及初始化过程 | 5 |
| 1.2.3 路由器的功能 | 8 |
| 1.2.4 路由器与网桥的比较 | 9 |
| 1.3 路由器分类 | 9 |
| 1.3.1 核心路由器 | 10 |
| 1.3.2 企业级路由器 | 11 |
| 1.3.3 接入路由器 | 11 |
| 1.4 路由器的关键性能指标 | 12 |
| 1.4.1 吞吐量 | 12 |
| 1.4.2 背板能力 | 13 |
| 1.4.3 丢包率 | 13 |
| 1.4.4 时延以及时延抖动 | 13 |
| 1.4.5 突发量能力 | 13 |
| 1.4.6 路由表容量 | 13 |
| 1.4.7 服务质量 | 13 |
| 1.4.8 网管能力 | 14 |
| 1.4.9 可靠性和可用性 | 14 |
| 1.5 核心路由器的组成 | 15 |
| 1.5.1 高速接口子系统 | 16 |
| 1.5.2 报文转发子系统 | 16 |
| 1.5.3 交换开关子系统 | 16 |
| 1.5.4 软件子系统 | 17 |
| 第 2 章 路由器技术的发展与相应的国际标准 | 19 |
| 2.1 路由器的发展 | 19 |
| 2.1.1 传统路由器 | 19 |
| 2.1.2 现代路由器技术 | 20 |
| 2.1.3 T 比特路由器 | 20 |
| 2.1.4 未来光交叉连接路由器 | 21 |
| 2.1.5 主动网络路由器 | 22 |
| 2.1.6 未来的发展 | 22 |

| | | |
|--------------|-----------------------------|-----------|
| 2.2 | 路由器设计标准 | 24 |
| 2.2.1 | 链路层实现要求 | 24 |
| 2.2.2 | 网络层 IP 协议实现要求 | 26 |
| 2.2.3 | 网络层 ICMP 协议实现要求 | 34 |
| 第 3 章 | 路由器关键技术 | 38 |
| 3.1 | 体系结构及调度技术 | 38 |
| 3.1.1 | 第一代单总线单 CPU 结构路由器 | 38 |
| 3.1.2 | 第二代单总线多 CPU 结构路由器 | 39 |
| 3.1.3 | 第三代交叉开关体系结构的 GSR 路由器 | 41 |
| 3.1.4 | 第四代多级交换路由器 | 42 |
| 3.2 | 高速硬件接口 | 44 |
| 3.3 | 系统软件 | 44 |
| 3.4 | 路由协议 | 44 |
| 3.5 | 高层交换与 MPLS | 46 |
| 3.6 | QoS | 46 |
| 3.7 | 安全 | 47 |
| 第 4 章 | 路由器交换结构设计 | 50 |
| 4.1 | 交换体系结构 | 50 |
| 4.1.1 | 进程交换 | 51 |
| 4.1.2 | 快速交换 | 52 |
| 4.1.3 | 最优交换 | 55 |
| 4.1.4 | 快速转发 CEF | 56 |
| 4.2 | 早期的报文交换体系结构 | 57 |
| 4.2.1 | 共享总线交换 | 57 |
| 4.2.2 | 共享存储交换 | 58 |
| 4.2.3 | 基于存储片的交换 | 61 |
| 4.3 | 高性能交换 | 63 |
| 4.3.1 | Crossbar 交换 | 64 |
| 4.3.2 | Cisco 12000 系列路由器交换结构 | 65 |
| 第 5 章 | 高端路由器系统软件设计 | 68 |
| 5.1 | 软件体系结构 | 68 |
| 5.1.1 | 概述 | 68 |
| 5.1.2 | 路由器系统软件的组成 | 70 |
| 5.2 | 路由器系统软件的实现方法 | 78 |
| 5.2.1 | 对实时系统的要求 | 79 |
| 5.2.2 | VxWorks | 80 |
| 5.2.3 | JUNOS 软件系统 | 82 |
| 第 6 章 | 高速接口设计技术 | 84 |
| 6.1 | 以太网接口设计 | 84 |

| | | |
|--------------|-------------------------|------------|
| 6.1.1 | 以太网的基本原理 | 84 |
| 6.1.2 | 具体接口设计 | 88 |
| 6.2 | POS 接口设计 | 91 |
| 6.2.1 | POS 技术概述 | 91 |
| 6.2.2 | POS Line Card 组成及数据处理过程 | 93 |
| 第 7 章 | 路由协议的设计与实现 | 99 |
| 7.1 | 路由结构 | 99 |
| 7.2 | RIP 协议设计 | 101 |
| 7.2.1 | RIP 协议简介 | 101 |
| 7.2.2 | RIP 协议的实现 | 106 |
| 7.3 | OSPF 协议设计 | 109 |
| 7.3.1 | OSPF 协议简介 | 109 |
| 7.3.2 | OSPF 的关键特征 | 110 |
| 7.3.3 | OSPF 协议的实现 | 114 |
| 7.4 | BGP 协议设计 | 126 |
| 7.4.1 | BGP 协议简介 | 126 |
| 7.4.2 | 协议操作机制 | 131 |
| 7.4.3 | BGP 协议的实现 | 133 |
| 7.4.4 | BGP 协议路由策略 | 148 |
| 第 8 章 | 路由表问题 | 153 |
| 8.1 | 概述 | 153 |
| 8.2 | Internet BGP 路由表的分析 | 154 |
| 8.2.1 | BGP 路由表的增长 | 154 |
| 8.2.2 | 路由表引发的一系列问题 | 156 |
| 8.2.3 | 地址耗尽 | 157 |
| 8.2.4 | 其他一些畸形现象 | 158 |
| 8.2.5 | 结论 | 158 |
| 8.3 | 高效的路由查找算法 | 159 |
| 8.3.1 | IP 地址组成变化及对路由查找的影响 | 159 |
| 8.3.2 | 常用的路由表查找算法 | 160 |
| 8.4 | 一种基于压缩树的路由查找算法 | 162 |
| 8.4.1 | 压缩树算法 | 163 |
| 8.4.2 | 实现压缩树算法的改进 | 167 |
| 第 9 章 | 高级交换技术 | 169 |
| 9.1 | 概述 | 169 |
| 9.2 | MPLS 交换 | 170 |
| 9.2.1 | MPLS 体系结构 | 170 |
| 9.2.2 | 转发粒度和等效前传类 | 173 |
| 9.3 | 第三层交换 | 174 |

| | | |
|---------------|-----------------------------------|------------|
| 9.3.1 | 出现的原因 | 174 |
| 9.3.2 | 主要技术 | 175 |
| 9.3.3 | Ipsilon 公司的 IP switching | 177 |
| 第 10 章 | 路由器 QoS 与安全 | 179 |
| 10.1 | 路由器 QoS | 179 |
| 10.1.1 | 拥塞管理 | 180 |
| 10.1.2 | 拥塞避免 | 183 |
| 10.2 | 安全 | 184 |
| 10.2.1 | 路由器的安全威胁 | 184 |
| 10.2.2 | 提高路由器安全性的主要技术 | 185 |
| 附录 A | 国际上主流高端路由器性能指标 | 190 |
| A.1 | Cisco 12416 Internet 路由器 | 190 |
| A.2 | Juniper 公司 M160 骨干路由器主要性能指标 | 195 |
| 附录 B | 参考的 RFC 标准列表 | 198 |
| 参考文献 | | 203 |

第 1 章 路由器技术概述

互联网络是一个被网络界广泛接受的术语，它表示相互连接的许多网络的集合。每一个独立的网络都有它自己的网络号，此网络号在特定的互联网络中必须是惟一的。路由器的基本作用是根据从网络协议获悉的有关信息，控制通过互联网络的通信量。

在一个有几百台、甚至几千台计算机连在一起的互联网络中，必须有一些约定的方式供这些设备相互访问和通信。随着网络规模的增大，让每一台计算机记住互联网络上其他所有计算机的地址是不切实际的，因此必须有一些机制来减少每台计算机为实现与其他所有计算机通信而维护的信息量。已使用的机制是将一个互联网络分成许多独立但互相连接的网络，这些网络本身可能又被分为许多子网(见图 1-1)。连接这些独立网络的任务可以交给被称为路由器的专用计算机来完成。使用这种方法，网络上的计算机只需记住互联网络中的独立网络的地址，而不需记住网络上的每一台计算机的地址。

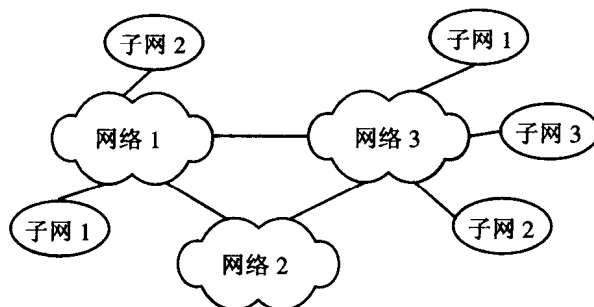


图 1-1 网络和子网的连接

1.1 引 言

通信网络是由一些系统和节点组成的集合，这些系统和节点负责传输连接在通信网络上的用户之间的信息。在一个网络中主要定义两种系统：端系统和中间系统。端系统是支持端用户应用或者服务的设备，中间系统是连接多个网络并允许这些网络的端系统相互之间进行通信的设备。

简单来说，路由器就是一个中间系统，它主要用来连接两个或者多个网络，这些网络可能是同构的也可能是异构的。一个路由器简单到可以只是一台有两块或更多网络接口卡的微机，从接收接口进来的数据报，经过处理，转发到适当的发送接口。路由器是负责在网络层

对 IP 数据包进行转发的主要设备，它还负责对 IP 数据包进行灵活的路由选择，把数据逐段向目的地转发。它掩盖了下层网络的细节，使各类网络都在 IP 上达到统一。

我们通过描述互联网上的计算机如何相互寻址来引入路由器。说明路由器的作用的最好的类比是邮局服务系统。当邮寄一封信时，需要提供公寓号码、街区名称和号码、城镇和省名。在计算机术语中，发送信息时，需要提供应用端口号、主机号、子网号和网络号。

核心的概念是当邮局接收到发往另一个城镇的信件时，邮政人员首先将它发送到目的城镇所在的分局。从那里，这封信被交给负责特定街区的某个邮递员。最终，这封信被投递到目的地。

计算机网络也采用相似的过程。发往互连网络的信息首先被送到与目的网络相连的路由器。路由器实际上起着这个网络的分发中心的作用，它把信息送到目的子网。最后此信息被送到目的主机的目的端口上。

图 1-2 给出了一个简单的互连网络的系统，其中由路由器将不同的网络连接起来。图中的网络 2、4 上有主机，而网络 5、6、7 上没有。网络 5、6、7 仅仅是与局域网或广域网上的路由器相连。在此网络中，主机 1 和 2 必须配置成同样的网络号(本例中为 4)。除此之外，与同一网络相连的路由器接口必须配置成同样的网络号。

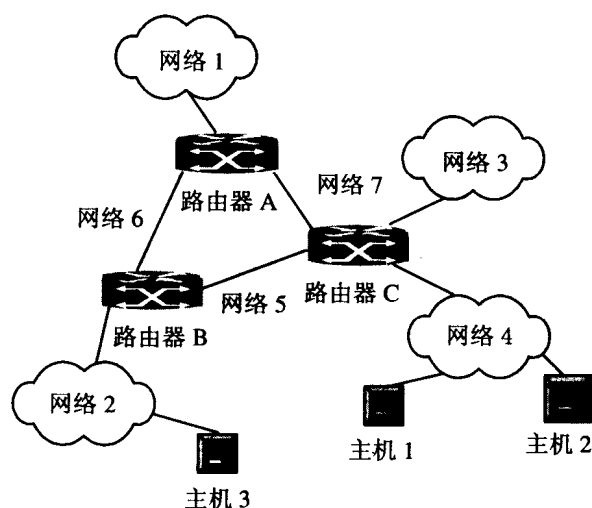


图 1-2 一个简单的网络连接

从本质上说，路由器的作用是将报文分组从一个网络路由到另一个网络。这句话暗含着两个含义。第一，一个路由器的多个接口不能被配置成相同的网络号(但是，通过使用子网掩码，可以在同一路由器的不同接口上配置相同的网络号，但是有不同的子网号)；第二，由于每条路由带有一个目的网络号，故在缺少目的网络号的情况下路由器不能转发路由报文。

1.2 路由器的组成及其功能

路由器工作在 OSI 参考模型的网络层(如图 1-3 所示)，完成不同网络之间的数据存储、

分组和转发。它可以根据报文来传输数据，完成网络层路由和转发任务。由于在两个不同网络的网络层之间按报文传输数据时，需要改变两个不同类型网络报文中的第二层地址，即决定在网络之间数据传输时的路由去向，所以称之为“路由器”。

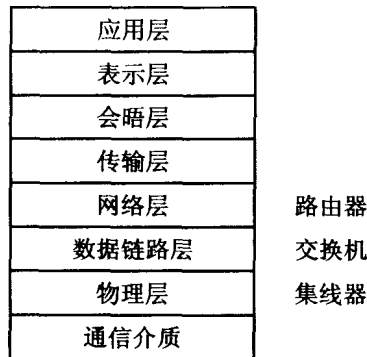


图 1-3 路由器在 OSI 参考模型中的位置

路由器的基本用途是连接多个逻辑上分开的网络，必须具有判断网络地址和选择路径的功能，能够在多个网络互联环境中建立灵活的连接，并可用完全不同的数据分组和介质访问方法连接各种子网。路由器只接受源站或其他路由器的信息，属于网络层的一种互联设备，它不关心各子网使用的硬件设备，但要求运行与网络层协议相一致的软件。

虽然路由器可以支持多种协议(例如 TCP/IP、IPX/SPX、AppleTalk 等协议)，但是在我国绝大多数路由器运行 TCP/IP 协议。本书假定路由器主要运行 TCP/IP 协议簇(实际情况也是如此)。目前网络层使用的协议为 IPv4，因为这是最流行的网络层协议，它所涉及的概念与其他网络层协议是类似的。

1.2.1 路由器的基本组成

路由器通常连接两个或多个由 IP 子网或点到点协议标识的逻辑端口，至少拥有 1 个物理端口。路由器根据收到的数据包中的网络层地址和路由器内部维护的路由表决定输出口以及下一跳地址，并且重写链路层数据包头实现转发数据包。路由器通常动态维护路由表来反映当前的网络拓扑，通过与网络上其他路由器交换路由和链路信息来维护路由表。

我们先以一个路由器模型为基础，介绍路由器的一般组成，后面还会有更详细的描述。一个典型的路由器主要由四部分组成：输入端口、输出端口、交换网络和路由处理器。如图 1-4 所示。

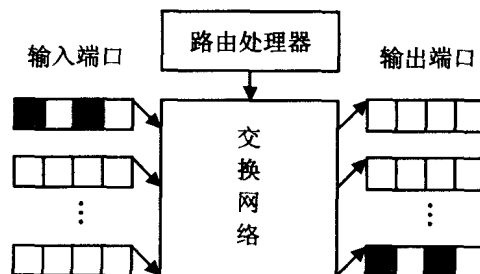


图 1-4 路由器基本组成

1. 输入端口

输入端口在线卡上，是物理链路的连接点也是报文的接收点。输入端口的设计遵守物理链路设计标准，完成的功能主要包括：

(1) 数据链路层帧的拆封。

(2) 在一些路由器的设计中，输入端口拥有一个转发表(是主处理器路由表的一个简单映射)，它可以直接在转发表中执行路由查找并把数据报送往输出端口，从而减主处理器的交换负担。早期的路由器都是简单地把所有接收到的数据报由输入接口送给主处理器进行处理，以至于主处理器的处理能力成为整个路由器性能的瓶颈。

(3) 为了提供 QoS，输入端口也可能根据预先指定的策略把接收的报文进行分类。

(4) 有时，输入端口还要运行数据链路层协议(如 SLIP、PPP 等)或者网络层的协议(如 PPTP)。

(5) 把一些特定的控制数据报(利用 RIP、OSPF、BGP 等路由信息数据报)送给路由处理器，由主处理器更新路由表。

2. 输出端口

输出端口的主要功能包括排队和缓冲管理，在交换网络以比接口速度更快的速率传送数据报时，这种功能显得尤其重要。输出端口使用复杂的调度算法(例如：RED 算法、WFQ 算法等)实现 QoS 功能，另外输出端口还要执行与输入端口类似的功能——数据报的封装和支持物理链路协议。

3. 交换网络

交换网络完成输入端口和输出端口之间的互联。交换网络有从简单的 Crossbar 交换到复杂的 Perfect shuffle、Clos 交换等许多实现方法，现在常用的主要有三种：总线交换、共享存储交换和互连网络交换。

最简单的交换方式是总线交换：当一个数据报到达输入端口时，通过一个共享总线直接被送往输出端口而不用路由处理器的干预，一旦发现总线忙，数据报就会在输入端口排队。这种方式的不足在于因为总线是共享的，路由器的交换能力受限于共享总线的带宽。

共享存储交换使用一个共享的存储器完成报文的交换，它没有独立的路由处理器，路由器的主 CPU 本身完成路由转发。当输入端口接收到数据报时，它首先中断主 CPU，将接收到的数据报从输入端口缓冲送到共享存储器，处理器经过路由表的查找为该数据报选择合适的输出端口，最后将该数据报送给输出端口。Cisco 8500 系列路由器使用的就是这种共享存储的交换方式。这种方法的不足是交换速度受限于访存速度。

互连网络交换把输入端口和输出端口用多条可用的交换路径连接起来，从而消除了共享总线带宽所带来的限制。一种典型的互连网络交换是 Crossbar。如图 1-5 中所示，Crossbar 通过交换网络把 N 个输入接口和 N 个输出接口连接起来，Crossbar 可以被看作一个由 $N \times N$ 个交叉点连接成的 $2N$ 条总线，当一个交叉点处于“可用”状态时，它所连接的两条总线之间就可以进行数据传输。Cisco 12000 系列路由器就是用这种互联方式进行交换，达到 60Gbit/s 的交换能力。

4. 路由处理器

路由处理器运行系统软件和各种路由协议，实现维护路由表和计算转发表等功能。其部分功能既可以用软件实现，也可以用硬件实现。对路由处理器的软件与硬件分工的优化也是

影响路由器性能的主要因素。

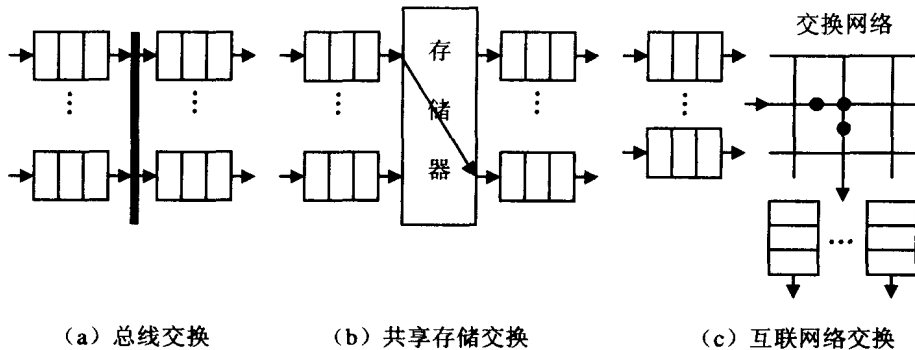


图 1-5 典型交换网络

可见，只要具有了以上四个基本要素，就可以实现路由器的基本功能。如何把这四个基本要素合理地结合起来，使路由器具有高性能的同时也具有低的价格一直是路由器体系结构开发所关注的问题。随着新硬件的不断上市，硬件成本也不断下降，原来是高不可攀的体系结构实现技术，现在已经有了实现的可能。同时由于网络的作用越来越受到重视，用户也愿意支付较高的投资来消除目前网络的瓶颈，从而大幅度提高整个网络的性能。

1.2.2 路由器的硬件组成及初始化过程

1. 硬件组成

目前市场上有大量各种类型的路由器产品。尽管这些产品在处理能力和所支持的接口数上有所不同，但它们都使用一些核心的硬件部件。图 1-6 展示了第一代基于总线的路由器的关键部件，其中 CPU(或微处理器)的类型、ROM 与 RAM 的大小以及 I/O 端口的数目和介质转换器，根据产品的不同会有些相应的变化，但每一个路由器都有图中所示的各个硬件部件。通过分析各硬件的功能，就能从整体上了解路由器的工作方式及其所提供的功能。

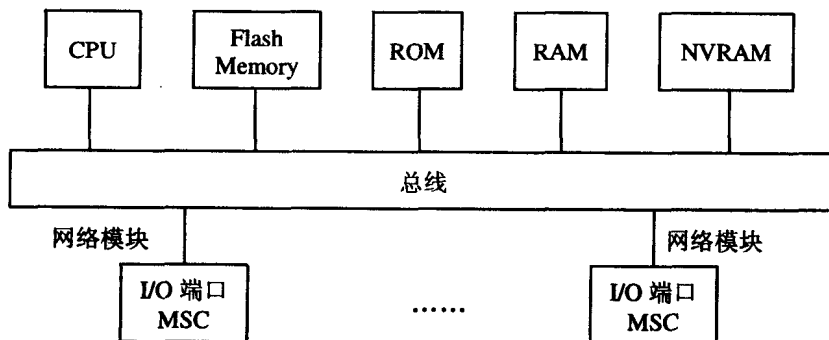


图 1-6 路由器的关键部件

(1) 中央处理单元

中央处理单元(Central Processing Unit, CPU)，或者称为微处理器，负责执行组成路由器操作系统(OS)的命令和通过控制台(Console)和 Telnet 连接输入的命令。因此，CPU 的处理能

力与路由器的处理能力直接相关。

(2) 闪存

闪存(Flash Memory)是一种可擦写、可编程的 ROM。在许多路由器上,闪存作为一种选择性的硬部件,负责保存操作系统的映像(Image)和路由器的微代码(Micro Code)。因为修改闪存无需更换和移动芯片,在需要定期修改存储内容的情况下,其代价低,且使用方便。只要空间允许,用户可以在闪存中存储多个操作系统的映像。这项功能对于测试新的映像十分有用。路由器的闪存还能通过使用普通文件传输协议(Trivial File Transfer Protocol, TFTP)将操作系统的映像加载到另一个路由器上。

(3) 只读存储器

只读存储器(Read Only Memory, ROM)中所包含的代码执行加电检测,这一点与许多微机所执行的加电自检(Power On Self-Test, POST)相同。ROM 中的启动程序还负责加载操作系统软件。尽管许多路由器需要软件升级时,只能通过更换路由器系统板上的 ROM 芯片才能做到,但另一些路由器可能使用不同类型的存储方式来保存操作系统。

(4) 随机存取存储器

随机存取存储器(Random Access Memory, RAM)用来保存路由表,进行报文缓存等。在许多流量流向一个通用接口时,报文可能不能直接输出到该接口,此时 RAM 可以提供报文排队所需的存储空间。在设备操作期间, RAM 还能提供保存路由器配置文件所需的存储空间。当 RAM 为缓存 ARP 信息提供空间时,它能减少 ARP 的流量,并增强多个网络连接到路由器时的传输能力。路由器关电时, RAM 中的内容被清除。

(5) 非易失的 RAM

非易失的 RAM(Nonvolatile RAM, NVRAM)在路由器关电时,仍能保持其内容。通过在 NVRAM 中保存配置文件的一个副本,路由器在电源故障时可以快速地恢复。使用 NVRAM 后,路由器就不再需要硬盘或软盘来保存其配置文件。因此,路由器中没有移动部件可以延长各部件的使用寿命。计算机系统的许多故障都是因为对移动部件的拆装,如硬盘故障等引起的。

因为路由器没有硬盘和软盘,所以通常的方法是将配置文件存储在一台微机(Personal Computer, PC)上,这样,用户可以使用文本编辑器方便地进行修改。配置文件可以通过网络使用 TFTP 直接加载到 NVRAM 中。当通过路由器来加载配置时,路由器相当于一个客户机,而文件所在的 PC 相当于一台服务器。这意味着用户要在 PC 上安装 TFTP 服务器软件来操作 PC,以便将文件在计算机上进行移入或移出。

(6) 输入/输出端口和特定介质转换器

输入/输出端口(Input/Output Port, I/O 端口)是报文进出路由器的连接装置。每一个 I/O 端口都连到一个特定介质转换器(Media-Specific Converter, MSC)上, MSC 提供物理接口到特定类型介质,如以太网、令牌环网、RS-232 连接或 POS(Packet Over SONET)接口等。

根据路由器的数据流,当路由器从网络上接收到数据后,报文的第 2 层报文头被丢掉,并进入 RAM 中。此时, CPU 检查路由表,以决定报文的输出端口及报文的封装方式。这种处理过程称为进程式交换模式(Process Switching Mode)。在进程式交换模式中,每一个报文都必须被 CPU 处理, CPU 要查询路由表并决定往何处发送报文。

Cisco 路由器还有一种交换模式称为快速交换(Fast Switching)。在快速交换模式中,路由

器维护着一个内存缓冲区，其中包含目的 IP 地址与下一跳的接口信息。路由器通过存储以前从路由表中获取的信息而创建这样的缓存。去往特定目标的第一个报文需要 CPU 查询路由表。一旦得到路由信息，在决定去往特定目标的下一跳时，该信息就已经存储在快速交换缓存中。当新的发往该目标的报文到达时，就不再需要查询路由表。这种机制使得路由器交换报文的速度提高，并且路由器 CPU 的负载也会相应降低。

快速交换的变种存在于特殊的硬件结构中——它只在高端路由器模型中才存在。但快速交换的变种与快速交换两者的原理在本质上是一样的，对于所有的交换模式都需要一个包含从目的地址到接口映射的缓存。

还有一种交换模式称为网络流交换(Net-flow Switching)，这种方式不仅缓存目的 IP 地址，还缓存源 IP 地址和上层的 TCP 或 UDP 端口等。

对于快速交换还有一些要点需要指出。首先，对路由表或 ARP 缓存的任何修改都会导致快速交换缓存中的内容被强制清除。该动作一般发生在拓扑结构发生改变时，此时快速交换缓存要重建。而且，快速交换缓存中的路由项会随着路由表中内容的改变而改变。快速交换缓存中的路由项与路由表中的相应项要能够匹配。

例如，如果路由器有一个到网络 10.1.1.0/24 的路由，路由器将会缓存目标 10.1.1.0/24。如果路由器只有一个到网络 10.1.1.0/16 的路由，路由器将会缓存目标 10.1.0.0/16。如果路由表中没有网络或子网的路由项，路由器就使用缺省路由，并使用缺省主网络屏蔽码，同时将缓存目的 10.0.0.0/8。这种方式只工作在对特定目标只有一个路由的情况下。如果有多个代价相同的路径，且无缺省路径，路由器就会缓存整个 32 位的目标。

例如，若目的 IP 地址为 10.1.1.1，而路由器中有两条到 10.1.1.0/24 网络的路由，则系统将缓存 10.1.1.1/32，并将其匹配成第一跳。下一个去往 10.1.1.0/24 网络的路由，如 10.1.1.2/32，将会缓存并匹配成第二跳。如果还有第三条代价相同的路径，则 10.1.1.0/24 网络上的下一个目标将成为第三跳，等等。注意，上述的情形只有在没有缺省路由的情况下才会发生。如果路由器要使用缺省路由来发送报文，则将缓存主网络号，而不是整个 32 位地址。

本质上，路由器使用一种称为轮询(round-robin)的方法缓冲每一个目标的下一跳地址，这意味着路由器的负载基于每一个目的站点进行了平衡。也就是说，因为快速缓存中包含了目的节点与接口的映射，一旦缓存中缓冲了某个表项，其他去往相同目的地的报文就可以使用缓存中的表项内容。路由器不会在快速交换缓存中缓冲同一目的结点的多个接口。

进程交换模式下，路由器基于每个报文进行负载平衡。因为没有快速交换缓存，每一个报文都要经过查询才能发送到后续的接口。如果有多条路径，这种方式就能使得网络的负载更加平衡，但路由器的 CPU 负载却增加了，从而降低了路由器传递报文的速度。

2. 路由器初始化过程

打开路由器的电源，路由器执行一系列预定义的操作。路由器所执行的附加操作还依赖于具体设备以及对设备进行的配置。要了解路由器的初始化过程，首先要分析设备加电时所发生的主要事件。

图 1-7 以流程图方式说明路由器初始化过程所执行的主要功能。路由器加电时，首先执行一系列诊断性测试，以验证处理器、存储器和接口电路能否正常工作。由于该过程是在加