

高性能路由器 设计与实现

Gaoxingneng Luyouqi
Sheji yu Shixian

朱培栋 孙志刚 张晓哲 陈一骄 编著



國防工業出版社
National Defense Industry Press

高性能路由器 设计与实现

朱培栋 孙志刚 张晓哲 陈一骄 编著



国防工业出版社

· 北京 ·

内 容 简 介

本书系统介绍高性能路由器的体系结构、设计和实现的关键技术及其最新进展,揭示路由器结构演化规律和发展趋势,深入探讨高性能路由器设计方法和实现的核心技术。

高性能路由器的设计和实现包括系统结构、硬件系统、软件系统和技术体制等多个方面,涵盖高速交换、高速接口、高性能转发引擎、网络处理器技术及路由协议、组播、MPLS、IPv6、安全等各主要功能的实现。本书结合近年来网络技术创新、新型业务发展和网络结构演化等对高性能路由器的研制提出的新要求,系统介绍高性能路由器体系结构和实现技术的最新发展,重点揭示高可扩展交换结构、支持多业务灵活处理的高性能转发引擎和新型路由器软件系统实现的关键技术,同时对集群路由器、可重构路由器和虚拟路由器等新型路由器的结构设计 with 系统实现进行了全面的探讨。

本书适合于系统学习和掌握路由器技术、新型网络技术的大专院校师生阅读,同时可供从事网络设备研发和大型网络运维的工程技术人员参考。

图书在版编目(CIP)数据

高性能路由器设计与实现/朱培栋等编著. —北京:国防工业出版社,2013. 8

ISBN 978-7-118-08966-0

I. ①高... II. ①朱... III. ①计算机网络—路由选择 IV. ①TN915.05

中国版本图书馆 CIP 数据核字(2013)第 179230 号

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 710×1000 1/16 印张 14 字数 263 千字

2013 年 8 月第 1 版第 1 次印刷 印数 1—3000 册 定价 46.00 元

(本书如有印装错误,我社负责调换)

国防书店: (010)88540777

发行邮购: (010)88540776

发行传真: (010)88540755

发行业务: (010)88540717

前 言

因特网规模的迅速扩展,网络流量的激增,以及新型网络应用的不断出现,对路由器等网络设备提出新的要求。高性能路由器要具有强大的处理能力和良好的扩展性,具有非常高的稳定性、可靠性和安全性,同时具备支持多协议、多类型端口、多种安全认证的灵活性。高性能路由器的设计和实现面临着非常多的挑战。

2005 年我们编著出版了《高性能路由器》一书。从 1999 年起,在国家科技部和国家自然科学基金委员会等部门的支持下,我们承担了多项高性能网络设备研制任务和骨干网络路由系统相关的研究项目,并取得一些重要成果,例如,“银河玉衡核心路由器”“新一代 IPv6 高性能路由器”“安全高速路由器”“域间路由监测系统”。可以说,2005 年版的《高性能路由器》是 5 年多高性能设备研发技术攻关和域间路由系统理论研究的结晶。

近 10 年来,网络规模持续增长,网络新业务不断涌现,互联网的体系结构依然未停止演化的步伐,从而对作为网络基础设施的高性能路由器的研制提出了新的要求。这些变化反映在高性能路由器的设计和实现上,主要表现在以下方面。在路由器的部件上,主要体现在 40Gb/s 和 100Gb/s 接口的广泛应用及未来更高速率接口的研发,转发引擎的多业务灵活处理需求,以及高可扩展多级交换结构的成熟运用。在路由器体系结构上,主要体现在三种新型路由器的研发需求:集群路由器、可重构路由器、虚拟路由器。集群路由器从 10 年前的原型研制,发展到现在的适当规模化应用;网络技术的创新和新型网络业务的发展,对可重构路由器的研制提出迫切要求;云计算环境的深化应用促进了虚拟路由器的研发。

在国家 973 课题“新一代互联网路由交换理论”,国家 863 专项“新一代高可信网络”的“开放式构件化通用路由器研制”课题,以及国家自然科学基金“新一代互联网路由系统协同机理研究”“基于互联网自然特性的域间路由系统基本理论和关键技术研究”“新一代互联网社会网络感知模型与能力设计”等项目的支持下,我们结合未来互联网的发展需求以及现实应用和部署要求,对新型路由器的设计和实现技术及域间路由系统的行为和机理进行了比较系统全面的研究,同时还基于可重构路由器技术研发了 NetMagic 实验平台,希望本书的出

版能够反映这一领域的最新发展。

2005年版的《高性能路由器》从系统结构、硬件系统、软件系统和技术体制四个方面考察高性能路由器,涵盖了高速交换、高速接口、高性能转发引擎、网络处理器技术及路由协议、组播、MPLS、IPv6、安全等各主要功能的实现。新版的《高性能路由器设计与实现》保留了原书最核心的内容并新增对三种新型路由器的关键设计与实现技术。内容的具体安排如下:第1章,高性能路由器概述和最新发展;第2~3章,选择最关键的、技术含量最丰富的高速交换技术和高性能转发引擎深入探讨高性能路由器设计和实现的具体技术;第4章,高性能路由器软件系统设计及其最新发展;第5章,路由协议的并行实现,揭示的是集群路由器实现的关键技术;第6章,可重构路由器设计原理及开放式构件化可重构路由器的设计方案;第7章,虚拟路由器及其对网络虚拟化的支持。

感谢蔡开裕教授、夏卓群博士、杨成同学对书稿提出的修改建议和编排工作,感谢徐明教授、卢泽新研究员、王宝生教授等对高性能路由器研制工作和域间路由系统研发工作的大力支持,感谢吕高峰博士、李韬博士和汤庆新同学等提供了相关的技术资料,感谢“新一代高可信网络”项目组的同行,包括解放军信息工程大学汪斌强教授、兰巨龙教授及清华大学的徐恪教授、徐明伟教授等。

本书的编写参考了国内外大量的资料,主要的在参考文献中已经列出。但是由于资料众多,逐一列举又恐拖沓、影响本书的可读性,特向文献的原作者及相关媒体表示感谢。高性能路由器及其相关技术的发展很快,书中涉及的某些关键技术的研究还刚刚起步,成熟的成果还比较少,难免有不准确、不完善之处,敬请读者和业内专家不吝赐教。

目 录

第 1 章 高性能路由器概述	1
1.1 高性能路由器的兴起与发展	1
1.1.1 高性能路由器的基本特征	1
1.1.2 高性能路由器产品演化历程	2
1.1.3 高性能路由器发展的应用需求	3
1.1.4 高性能路由器使能技术	5
1.2 高性能路由器的组成	6
1.2.1 高性能路由器的基本组成	6
1.2.2 高性能路由器的结构特点	7
1.2.3 高性能路由器软件系统	10
1.2.4 高性能路由器信息模型	11
1.3 高性能路由器新型体系结构	14
1.3.1 集群路由器	14
1.3.2 开放架构路由器	16
1.3.3 可编程路由器	19
1.3.4 构件化可重构路由器	19
1.3.5 虚拟路由器	20
1.4 高性能路由器的扩展性设计	23
1.4.1 高性能路由器实现的扩展性	24
1.4.2 并行技术的应用	25
1.5 高性能路由器可靠性设计	26
1.5.1 路由器的可靠性要求	26
1.5.2 平稳重启和无中断转发	27
1.5.3 硬件系统的冗余设计	28
1.5.4 软件系统的可靠性	29
1.5.5 在线升级能力	29
1.5.6 对网络可靠性的支持	29
1.6 高性能路由器服务的灵活性	30

1.6.1	服务分离	30
1.6.2	高性能的服务	30
1.6.3	控制平面的调节能力	30
1.6.4	接口的灵活性	30
1.6.5	转发引擎的灵活性	31
1.6.6	服务质量支持	31
1.7	高性能路由器服务的能耗设计	31
第2章	高速交换技术	34
2.1	基本交换结构	34
2.1.1	基本交换结构	34
2.1.2	共享内存结构	34
2.1.3	Crossbar 交换结构	35
2.2	Crossbar 工作模式与调度算法	37
2.2.1	Crossbar 交换结构的分类	37
2.2.2	Crossbar 调度算法类型	41
2.2.3	单播调度算法	42
2.2.4	组播调度算法	45
2.2.5	支持服务质量的调度算法	48
2.2.6	交换背板的设计	51
2.3	高可扩展交换结构	51
2.3.1	Crossbar 的局限性	51
2.3.2	多级交换结构	52
2.3.3	动态多级交换网络	53
2.3.4	静态多级交换网络	58
2.3.5	并行分组交换网络	60
第3章	高性能转发引擎	62
3.1	转发引擎的功能与结构	62
3.1.1	多协议转发功能	62
3.1.2	服务质量控制	63
3.1.3	安全策略控制	64
3.1.4	转发引擎的结构	64
3.2	路由查找	65
3.2.1	最长前缀匹配	65
3.2.2	IP 路由查找面临的挑战	66

3.2.3	高性能路由器的路由表查找	68
3.2.4	硬件直接查表算法示例	69
3.2.5	TCAM 算法示例	70
3.3	转发引擎流量管理系统	72
3.3.1	流量管理系统的组成	72
3.3.2	报文分类	73
3.3.3	流量测量	73
3.3.4	流量管制	75
3.3.5	流量整形	76
3.3.6	报文调度	76
3.3.7	缓冲区管理	77
3.4	报文分类	79
3.4.1	报文分类技术	79
3.4.2	报文分类面临的挑战	80
3.4.3	典型的分类算法	81
3.5	排队与调度	81
3.5.1	排队	81
3.5.2	RED 算法	83
3.5.3	报文调度	88
3.5.4	基本调度算法	90
3.6	Cisco 快速转发技术	95
3.6.1	基于 Cache 的转发	96
3.6.2	Cisco 快速转发	96
3.6.3	Cisco 分布式快速转发	97
第 4 章	高性能路由器软件系统	99
4.1	路由器软件系统结构的演化	99
4.1.1	传统路由器软件结构的局限性	99
4.1.2	高性能路由器软件系统基本结构	100
4.2	高性能路由器软件系统的功能组成	101
4.2.1	软件组成	101
4.2.2	操作系统	102
4.2.3	路由协议和信令	106
4.2.4	用户界面	106
4.3	Cisco IOS 软件系统	107

4.3.1	IOS 概述	107
4.3.2	Cisco IOS 结构	107
4.3.3	Cisco IOS XR	109
4.4	高性能路由器软件系统的性能设计	110
4.4.1	路由系统的扩展性	110
4.4.2	软件系统性能设计	110
4.4.3	路由协议及其扩展性	112
4.4.4	提高路由系统扩展性的途径	114
4.5	高性能路由器软件系统的可靠性设计	115
4.5.1	实现软件系统可靠性的基本途径	115
4.5.2	稳定性设计	116
4.5.3	软件系统可靠性结构	116
第 5 章	路由协议并行实现	120
5.1	BGP 并行处理方法	120
5.1.1	域间路由性能问题	120
5.1.2	BGP 协议并行处理方法	121
5.2	BGP 并行处理模型	124
5.2.1	BGP 协议实体	125
5.2.2	C - BGP 协议	126
5.3	BGP 协议的并行路由计算算法	127
5.3.1	算法选择依据	127
5.3.2	BGP 邻居会话的划分算法	128
5.3.3	内部路由视图一致性维护	129
5.4	BGP 并行处理模型性能评价	133
5.4.1	BGP 实体的功能分析	133
5.4.2	BGP 实体的性能分析	134
5.5	多路由协议并行处理结构的转发表同步机制	135
5.5.1	相关工作	136
5.5.2	AREF 路由同步机制	138
5.5.3	AREF 同步机制性能分析	140
5.6	小结	144
第 6 章	可重构路由器的设计	145
6.1	可重构路由器概念与相关技术	145
6.1.1	研发需求	145

6.1.2	相关工作	146
6.1.3	可重构计算系统	147
6.1.4	构件化技术	150
6.1.5	软件定义网络(SDN)	151
6.2	开放可重构路由器体系结构	153
6.2.1	开放可重构路由器系统组成	153
6.2.2	开放可重构路由器平台模型	154
6.2.3	控制组件的构件支撑技术	155
6.2.4	硬件平台可重构的支撑技术	157
6.3	开放可重构路由器硬件设计	158
6.4	开放可重构路由器软件设计	160
6.4.1	可重构路由器软件系统总体设计	160
6.4.2	可重构控制实体软件结构设计	161
6.5	构件运行平台及构件设计	163
6.5.1	可重构路由器构件总体运行环境	163
6.5.2	硬件构件设计	164
6.5.3	软件构件设计	171
6.5.4	可重构路由器转发构件设计及其实例	173
6.6	基于 NetMagic 的硬件开发	180
6.6.1	NetMagic 模型	180
6.6.2	NetMagic 的硬件开发模型	181
6.6.3	NetMagic 硬件开发的关键技术	184
6.6.4	NetMagic 的硬件参考设计——NM - Probe	185
6.7	小结	187
第 7 章	虚拟路由器的设计	189
7.1	虚拟路由器研发现状	189
7.1.1	网络虚拟化的概念	189
7.1.2	网络虚拟化技术的研究与试验	190
7.1.3	路由器厂商对网络设备虚拟化的支持	191
7.1.4	网络设备的虚拟化与软件定义网络	192
7.2	虚拟路由器实现模式及其比较	193
7.2.1	虚拟路由器实现模式	193
7.2.2	虚拟路由器实现模式的比较	193
7.2.3	虚拟路由器实现模式的选择	195

7.2.4 虚拟路由器实现示例	195
7.3 软硬件结合的虚拟路由器设计	196
7.3.1 网络虚拟化设备支撑体系	196
7.3.2 软硬件结合的二维虚拟化路由器结构模型	197
7.3.3 网络资源的虚拟化技术	199
7.3.4 虚拟网络资源的统一管理和调度技术	201
7.3.5 控制平面的软件虚拟化技术	203
7.3.6 数据平面的虚拟化技术	204
7.4 小结	205
附录 缩略语	207
参考文献	212

第1章 高性能路由器概述

扩展性、可用性和服务的灵活性是高性能路由器设计的主要目标。高性能路由器的设计与实现的关键技术涉及:体系结构设计,交换网络,高速接口,高速转发引擎,网络处理器应用,软件系统结构,路由协议实现,IPv6/MPLS/组播技术支持,安全功能,对新型应用、网络管理和流量工程的支持,以及与光网络的互操作等。本章首先考察高性能路由器设计的主要目标,并以此作为主线和指导原则,在后续章节分别探讨高性能路由器的设计与实现的关键技术。

1.1 高性能路由器的兴起与发展

1.1.1 高性能路由器的基本特征

顾名思义,高性能路由器应该是性能指标高的路由器。路由器的高性能是一个相对的概念。在2000年前,系统吞吐量在数 Gb/s ($1\text{G} = 1000\text{M}$) 的路由器可称为高速路由器,现在普遍认为吞吐量达到数百 Gb/s、数 Tb/s ($1\text{T} = 1000\text{G}$) 或更高速率的路由器才能称得上高性能路由器。

路由器的高性能不只是体现在吞吐量上,还要有比较高的协议处理速度和比较强的安全处理能力。性能不只是体现在一个绝对数值,而更强调高扩展能力。高性能路由器除了要具有强大的计算和处理能力外,还要具有高稳定、高可靠、高安全等特点,而 IPv6 技术的应用、光通信技术的迅速普及 MPLS 技术成为主流,要求新一代高性能路由器不仅具有更大容量的交换网络,同时要具备支持多协议、支持高密度多种端口、支持多种安全认证的能力。

在硬件体系结构上,高性能路由器在关键的 IP 业务流程处理上采用了可编程的、专为 IP 网络设计的网络处理器技术。它通过若干微处理器和一些硬件协处理器并行处理,通过软件来控制处理流程。对于一些复杂的操作(如内存操作、路由表查找算法、QoS 的拥塞控制算法、流量调度算法等)采用硬件协处理器来提高处理性能。

就目前而言,高性能路由器一般应该具有以下特征:

(1) 系统具有足够高的数据包处理能力以达到数百 Mp/s (每秒百万个包) 的转发能力;

(2) 分布式可扩展的大容量矩阵式交换结构;

- (3) 高密度、多端口和扩展性;
- (4) 采用网络处理器技术实现 IP 报文处理和转发;
- (5) 全面的冗余设计,高可靠性;
- (6) 应具有有效的 QoS 手段来满足客户在不同场合对不同服务质量的要求;
- (7) 能支持多种现行协议标准和支持未来协议发展的能力如 IPv6 和 MPLS 等。

1.1.2 高性能路由器产品演化历程

以下是不同时期商用路由器交换容量的大致量级:1992 年约 2Gb/s,1995 年约 10Gb/s,1998 年约 40Gb/s,2001 年约 160Gb/s,2003 年约 640Gb/s,2004 年 Cisco 推出 CRS-1 支持 Tb/s。下面对目前几类高性能路由器产品及其主要特征作一介绍。

1. Gb/s 路由器

典型配置:160 Gb/s 以上交换容量,接口速度达 OC-192c (10Gb/s)以上;较高的接口密度,例如 8 端口的 10G 以太网接口板;较多深层输出队列;大型转发表;可靠性增强;使用最新网络处理器。例如 Cisco 公司的 12000 系列和 XR 12000 系列中的 12416 吞吐量达到 320Gb/s。Juniper 的 T640 路由器承载 40 个 10G 接口、8 个 40G 接口,吞吐量可达 640Gb/s。华为 NetEngine 80E 核心路由器交换容量 640Gb/s,接口容量 320Gb/s,转发性能 400Mp/s。

2. Tb/s 路由器产品

典型配置:Tb(Terabit,或称太比特,其中 1T = 1000G)交换容量,往往采用多机箱结构或采用分布式配置;大量的 OC-768c 或 40/100GE 接口和更大的接口密度;支持 DWDM;更小的报文转发时延。

典型的 Tb/s 路由器主要有:早期的 Avici System 公司的 TSR 路由器,Pluris 公司的 TeraPlex20。Cisco 公司的 CRS-3 系列多机柜路由器,1152 插槽,吞吐量可达 322Tb/s。Juniper 公司的 T4000 路由器吞吐量 4Tb/s,TX Matrix Plus 通过互联多个 T4000 机柜,吞吐量可达 16Tb/s。华为 NE5000 集群系统采用 16 拖 64 结构,最多可达 200Tb/s。阿尔卡特朗讯 7950 XRS 支持高达 32Tb/s 的交换容量,整机可支持 160 个 100GE 端口。

3. 高性能路由器的国家标准

2009 年工信部发布了“路由器设备技术要求——核心路由器”(YD/T 1097—2009),以取代 2001 年发布的“路由器设备技术规范——高端路由器”(YD/T 1097—2001)。在功能要求中,增加了 MPLS、QoS、IPsec 及 VPN 等实现要求;新增 10GE、10G POS 接口要求,删除 E3 接口。在性能指标中,系统双向交换容量从 20Gb/s 改为 60Gb/s,删除吞吐量和丢包率指标要求,要求 IP 包转

发延迟均应小于 1ms;删除设备处理能力的要求,增加对标记交换吞吐量和转发延迟指标;路由表容量仍然是 250000 条路由,但是支持的 BGP 对等由 50 个变为 500 个,IGP 邻居由 50 个变为 1000 个。增加了安全性和 IPv6 要求,专门发布“路由器设备安全技术要求——高端路由器”(YD/T 1359—2005)和“IPv6 网络设备技术要求——支持 IPv6 的核心路由器”(YD/T 1454—2006)。

“路由器设备技术要求——核心路由器”(YD/T 1097—2009)是对在骨干网络部署的路由器的最基本要求,实际上 2005 年前后生产的路由器已经能够满足这些指标。相关部门也在组织制定“路由器设备技术要求——集群路由器”国家标准,主要对集群路由器形态、分布处理能力、可靠性、4B 自治系统号的支持等提出要求,要求具有更高的吞吐量(例如,单机框系统双向交换容量应大于 640Gb/s,多框级联形态下双向容量应大于 1.28Tb/s),更大的路由表和更强的协议处理能力等。

1.1.3 高性能路由器发展的应用需求

1. 下一代网络对路由器发展的要求

为了满足网络市场需求,下一代网络应该具有以下特点:无线网络的移动性,公共电话网络(PSTN)的稳定性,因特网的广泛可达性,专线的安全性和服务隔离,IP 和 MPLS 支持集成的数据、话音和视频服务的灵活性,光网络的大容量,公共基础设施运行的高效性。

许多服务提供商认为下一代网络应该建立在新的基础之上,这个新基础包括以下三个方面:

(1) 基础设施聚合:将多种特殊的服务网络聚合成以 MPLS 为核心的公共网络;

(2) 服务聚合:数据、话音和视频服务聚合到灵活的 IP/MPLS 网络;

(3) 网络简化:简化网络尤其是 POP(Point Of Presence, 汇接点)的结构,去除复杂的分层和冗余,构造高扩展、高可靠、可管理的 IP/MPLS 网络。

2. 应用速度要求

按照斯坦福大学 Nick McKeown 教授的分析,从 1990 到 2004 年,互联网流量大致每 12 个月增长 1 倍,而路由器的容量每 18 个月才增长 1 倍。从 2005 到 2010 年,互联网的流量每年大约增长 0.5 倍;Cisco 预测 2011—2016 年全球互联网流量将增长 3 倍,到 2016 年达到 110EB(1E = 1000P, 1P = 1000T),每年平均增长约 29%。虽然互联网流量的增长速度有所放缓,但是对路由器吞吐量的要求却不断提高。

3. 软件性能要求

高性能路由器和传统路由器对协议软件性能的要求不同。最初的路由器采用软件转发,要求软件系统具有非常高的转发性能。高性能路由器采用硬件转

发,软件系统主要负责控制平面的工作。随着网络规模的扩展,路由表容量、路由更新速度、支持网络邻居数目、支持的 LSP(标记交换路径)数目、IPv6 隧道数目等成为考察路由器软件系统性能的主要指标。

4. 简化网络结构与管理

运营商通过网络汇接点(Point Of Presence,POP)与用户互连,POP 可能由多台路由器组成,如图 1-1 所示。

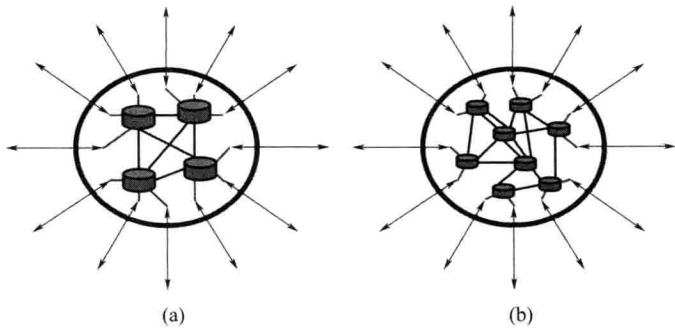


图 1-1 POP 的组成

(a)大型路由器构造的 POP;(b)众多小型路由器构造的 POP。

有的运营商网络设施比较复杂,会有数百乃至几千个 POP 分布在多处,有的分多个层次,互连线路更是数以千计。例如,在 2008 年中国电信 163 网的省间链路总带宽就达到 6Tb/s,部分 POP 节点交换容量超过 1Tb/s,每个 POP 节点超过 10 台路由器,汇聚了 100 多条 10Gb/s 链路。简化单个 POP,会给整个网络的管理带来收益。单个 POP 的简化,还有可能将地理上分离的多个 POP 合并,从而进一步简化整个网络。

采用端口直接互连多台 Gb/s 路由器,可以使整个 POP 的性能达到 Tb/s 级。但是这种方法需要独立管理各个路由器,不但管理复杂,而且端口浪费非常严重,费用高。以每个路由器 8 个端口为例,如果 4 台路由器全互连,可用端口只有 20 个,而用于 POP 内机群(Cluster)互连的端口就有 12 个。如果作为一个 Tb/s 路由器,采用专用交换网络,从内部互联各个接口,可以显著降低 Tb/s POP 的构造成本,简化网络拓扑与系统管理。

图 1-2 是一个传统 POP 的结构,路由器已经按照核心层、汇聚层和边缘层组织。

这种结构的缺点是:

- (1) 可用性是通过路由器和链路的冗余来实现的,代价大。
- (2) POP 的扩展能力有限,需要较多的插槽用于 POP 内部互连,链路升级往往要求路由器更新换代。
- (3) 管理的网元多,监测和维护的设备多,配置难度大,路由器发生故障的

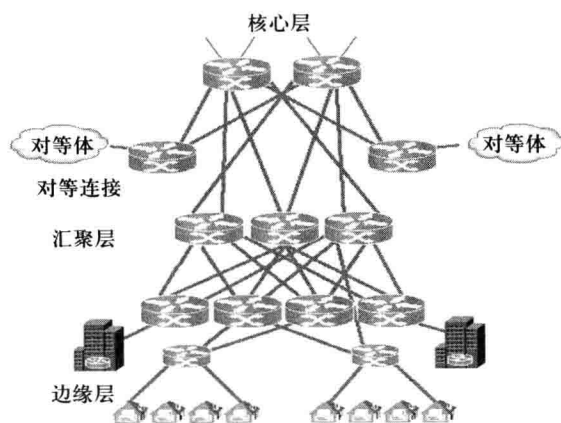


图 1-2 典型的 POP 结构

可能性大。

(4) 路由复杂,配置的路由协议邻居多,增加了路由收敛时间。

(5) 不同层次路由器服务能力的差异使 POP 支持安全、QoS、组播等服务的能力受到影响。

只有采用新型的路由系统,简化 POP 结构,才能避免以上问题。

1.1.4 高性能路由器使能技术

1. 光传输技术

近年来随着光通信技术和微电子技术的发展,密集波分复用(Dense Wavelength Division Multiplexing, DWDM)得到飞速进步,网络容量不再受限于传输链路。早在 2005 年 1.6Tb/s (160 * 10Gb/s) 的波分复用系统进入商用阶段,7.04 Tb/s(176 * 40Gb/s) 的系统已完成长距离传输试验。阿尔卡特朗讯推出了 400Gb/s 的光业务引擎(Photonic Service Engine, PSE),不但直接支持(200/400) Gb/s 速率,而且将 100Gb/s 的传送距离提升到 3000km。阿尔卡特朗讯采用 80nm 谱宽的光放大器实现了 1022 个波长复用,2013 年 5 月华为发布了 40Tb/s 的波分复用系统。中国电信已经开始集中采购 100Gb/s DWDM 系统。

2. 光网络技术

自动交换光网络(Automatically Switched Optical Network, ASON)通过信令系统具有完整的光网路由和控制功能,支持 IP 路由设备动态向光网申请端口资源,被认为是光传输网发展的主要方向。光网的自动交换和动态配置主要有两种模型:客户—服务器模型(也称层叠模型)和对等模型。ITU-T 支持客户—服务器模型,拟定了相关的建议;IETF 采用对等模型,提出通用多协议标记交换(Generic Multiple Protocol Label Switching, GMPLS)协议框架和信令规范。光网

络技术的发展对路由器的设计和实现提出了新的要求。

3. 新型设计与制造工艺

早在 2005 年基于 SiGe 和 InP 材料工艺生产 OC-768 前端电路,具有 CMOS 所无法比拟的高频特性。Cisco 的 100Gb/s 接口采用硅光子技术,将光与电子元件组合至独立的微芯片中以提升芯片之间的连接速度,并且降低了能耗。

4. 高速交换网络芯片

Agere Systems 公司在 2002 年就推出的交换开关 PI-40 芯片组最高支持 OC-768c 速率。Broadcom 的 BCM88650 芯片支持 100Gb/s 接口。VITESSE、Internet Machines 公司都提供从网络处理器到交换模块的套片及整体解决方案,相互之间可无缝连接进行互操作。

5. 高速分组存储技术

在 2002 年,并行分组缓冲技术(Parallel Packet Buffering, PPB)使用 DRAM 结构完全能够满足 OC-192 线速所需的带宽。快速分组缓冲技术(Fast Packet Buffering, FPB)以及基于乱序执行的流队列管理技术等都可以支持 OC-192 以上速率接口。

6. 网络处理器技术

网络处理器(Network Processing Unit, NPU 或 Network Processor, NP),从 2000 年初到现在被许多网络设备制造商选作新一代高端路由器设备的核心处理器,而在这段时间里,能够开发出成熟的 NPU 芯片的公司也从开始的两三个迅速增加到了十几个,而且 NPU 的处理能力也从 2.5Gb/s 扩展到 10Gb/s 和 40Gb/s。从 2010 年 Xelerated 公司就开始推出 100Gb/s 网络处理器,Broadcom、Marvell 等目前成为 100Gb/s 网络处理器的主力厂商,阿尔卡特朗讯在 2011 年推出了 400Gb/s 网络处理器 FP3。网络处理器以其杰出的包处理性能及可编程性已经成为构成路由转发引擎不可替代的部分。

7. 新型网络体制和协议

MPLS 技术和新型 QoS 机制的出现,以及软件定义网络(Software Defined Network, SDN)的发展,为高速路由器的设计和应用提供了更好的支持并提出新的需求。

1.2 高性能路由器的组成

本节介绍高性能路由器的基本组成、结构特点和信息模型。

1.2.1 高性能路由器的基本组成

如图 1-3 所示,路由器在逻辑上可以抽象为 4 个部分,即输入端口、交换结构、输出端口和路由处理器。