



国外经典教材·计算机科学与技术



Optimizing Network Performance with Content Switching

网络性能优化——内容交换技术

Matthew Syme Philip Goldie 著

于涛 罗庆华 谢文亮等 译



清华大学出版社

网络性能优化—— 内容交换技术

Matthew Syme 著
Philip Goldie

于 涛 罗庆华 谢文亮 等 译

清华大学出版社
北 京

内 容 简 介

Simplified Chinese edition copyright © 2004 by PEARSON EDUCATION ASIA LIMITED and TSINGHUA UNIVERSITY PRESS.

Original English language title from Proprietor's edition of the Work.

Original English language title: Optimizing Network Performance with Content Switching, 1st by Matthew Syme & Philip Goldie, Copyright © 2004

EISBN: 0131014684

All Rights Reserved.

Published by arrangement with the original publisher, Pearson Education, Inc., publishing as Prentice Hall Professional Technical Reference.

This edition is authorized for sale only in the People's Republic of China (excluding the Special Administrative Region of Hong Kong and Macao).

本书中文简体翻译版由 Prentice Hall PTR 授权给清华大学出版社在中国境内(不包括中国香港、澳门特别行政区)出版发行。

北京市版权局著作权合同登记号 图字:01-2003-8788

版权所有, 翻印必究。

本书封面贴有 Pearson Education(培生教育出版集团)激光防伪标签, 无标签者不得销售。

图书在版编目(CIP)数据

网络性能优化: 内容交换技术/希姆(Syme, M.), 高弟(Goldie, P.)著;于涛,罗庆华,谢文亮等译. —北京:清华大学出版社, 2004. 7

书名原文: Optimizing Network Performance with Content Switching

ISBN 7-302-08758-X

I. 网… II. ①希… ②高… ③于… ④罗… ⑤谢… III. 计算机网络—程序设计 IV. TP393

中国版本图书馆 CIP 数据核字(2004)第 052473 号

出 版 者: 清华大学出版社

<http://www.tup.com.cn>

社总机: 010-62770175

地 址: 北京清华大学学研大厦

邮 编: 100084

客户服务: 010-62776969

责任编辑: 常晓波

封面设计: 立日新

印 刷 者: 北京四季青印刷厂

装 订 者: 北京鑫海金澳胶印有限公司

发 行 者: 新华书店总店北京发行所

开 本: 185 × 230 印张: 13 字数: 290 千字

版 次: 2004 年 7 月第 1 版 2004 年 7 月第 1 次印刷

书 号: ISBN 7-302-08758-X/TP · 6245

印 数: 1 ~ 4000

定 价: 25.00 元

本书如存在文字不清、漏印以及缺页、倒页、脱页等印装质量问题, 请与清华大学出版社出版部联系调换。联系电话: (010)62770175 - 3103 或 (010)62795704

出版说明

近年来,我国的高等教育特别是计算机学科教育,进行了一系列大的调整 and 改革,急需一批门类齐全、具有国际先进水平的计算机经典教材,以适应当前我国计算机科学的教学需要。通过使用国外先进的经典教材,可以了解并吸收国际先进的教学思想和教学方法,使我国的计算机科学教育能够跟上国际计算机教育发展的步伐,从而培育出更多具有国际水准的计算机专业人才,增强我国计算机产业的核心竞争力。为此,我们从国外知名的出版集团 Pearson 引进这套“国外经典教材·计算机科学与技术”教材。

作为全球最大的图书出版机构, Pearson 在高等教育领域有着不凡的表现,其下属的 Prentice Hall 和 Addison-Wesley 出版社是全球计算机高等教育的龙头出版机构。清华大学出版社与 Pearson 出版集团长期保持着紧密友好的合作关系,这次引进的“国外经典教材·计算机科学与技术”教材大部分出自 Prentice Hall 和 Addison-Wesley。为了组织该套教材的出版,我们在国内聘请了一批知名的专家和教授,成立了一个专门的教材编审委员会。

教材编审委员会的运作从教材的选题阶段即开始启动,各位委员根据国内外高等院校计算机及相关专业的现有课程体系,并结合各个专业的培养方向,从 Pearson 出版的计算机系列教材中精心挑选针对性强的题材,以保证该套教材的优秀性和领先性,避免出现“低质重复引进”或“高质消化不良”的现象。

为了保证出版质量,我们为这套教材配备了一批经验丰富的编辑、排版、校对人员,制定了更加严格的出版流程。本套教材的译者,全部来自于对应专业的高校教师或拥有相关经验的 IT 专家。每本教材的责编在翻译伊始,就定期不间断地与该书的译者进行交流与反馈。为了尽可能地保留与发扬教材原著的精华,在经过翻译、排版和传统的三审三校之后,我们还请编审委员或相关的专家教授对文稿进行审读,以最大程度地弥补和修正在前面一系列加工过程中对教材造成的误差和瑕疵。

由于时间紧迫和受全体制作人员自身能力所限,该套教材在出版过程中很可能还存在一些遗憾,欢迎广大师生来电来信批评指正。同时,也欢迎读者朋友积极向我们推荐各类优秀的国外计算机教材,共同为我国高等院校计算机教育事业贡献力量。

清华大学出版社

《国外经典教材·计算机科学与技术》

编审委员会

主任：

孙家广 清华大学教授、中国工程院院士

委员：

邵志清	华东理工大学教授
吴百锋	复旦大学教授
孟祥旭	山东大学教授
何 晨	上海交通大学教授
龚声蓉	苏州大学教授
沈钧毅	西安交通大学教授
冯少荣	厦门大学教授
阮秋琦	北方交通大学
徐佩霞	中国科学技术大学教授
刘腾红	中南财经政法大学教授
秦小麟	南京航空航天大学教授
刘乐善	华中科技大学教授
冯全源	西南交通大学教授
王成山	天津大学教授
谢希仁	解放军理工大学教授
徐晓飞	哈尔滨工业大学教授
陈 纯	浙江大学教授
周伯生	北京航空航天大学教授

陈道蓄	南京大学教授
王 珊	中国人民大学教授
蒋宗礼	北京工业大学
周立柱	清华大学教授
陈 钟	北京大学教授
曹元大	北京理工大学教授
姚淑珍	北京航空航天大学
钱培德	苏州大学教授
孟祥旭	山东大学教授
李 彤	云南大学
孙吉贵	吉林大学

前 言

目前,关于内容交换及相关应用的技术资料不多。究其原因,我们认为有两点。首先,内容网络是一门极为复杂的学科,涉及许多不同的领域,从交换到缓存乃至内容路由技术,这一点常常让人感到费解。其次,其应用市场还有待开拓。

虽然内容网络确实涵盖了很多技术,但绝大多数都是以现有成熟的技术为基础的。实际应用中可能会用到专用的硬件设备、功能更丰富的软件以及特殊的配置,但这都是在我们已经很熟悉的现有技术基础之上发展起来的。

内容网络纯粹是现有成熟技术应用的延展,我们对之并不缺乏技术和认知。对内容交换机工作原理的深入了解并学习实际应用的方式,这是我们进一步学习的基础。

本书不会面面俱到地讲述整个内容网络,但会介绍内容交换技术,并对相关应用做些讨论。

读者对象和本书目标

本书的目的是为了消除目前存在的关于这项发展中技术的模糊和混乱认识,以简明清晰的方式叙述了基本理论和技术。这是从两层实现方式来完成。

对那些欲对这种新生网络技术有个整体了解的经理人、销售人员及其他各级管理层人员,能够从本书学习到相关主题的介绍性概述。这将使他们对每项应用的优缺点有个全面透彻的了解机会,可以学到为什么安全套接字层(SSL)卸载、防火墙负载平衡、Web 缓存重定向等是技术运用的基础知识,并理解不断革新的技术概念。简言之,本书会让读者对内容交换技术具有基本的认识,并使你在相关的日常工作中做出合理的决策,对从事相关产品营销工作的市场人员也会有很大的帮助。

对于技术人员,本书将全面介绍内容交换技术的深层工作原理,从而为掌握第2层、第3层应用技术打下基础。深入理解了第2层和第3层功能,会让读者更容易理解本书的深层次内容,包括TCP/IP流和会话、延时绑定、URL/URI解析、Cookie持久性、服务器运行监测,以及负载平衡效果等等。所有需要接触这些技术的工程技术人员可以通过本书全面了解应用技术的核心概念和功能,了解到这些技术内幕。

案例研究

本书所举的示例均围绕假想的 Foocorp 公司展开。Foocorp 是一个相当大的公司，从总部到各分公司拥有几千名雇员。雇员需要访问内部应用和系统。Foocorp 公司还通过外部网络一直与其工作伙伴和客户们保持畅通的联系。Foocorp 还建立了在线电子商务平台。

本书各主要章节都将讨论 Foocorp 公司实际面对的效益、设计和问题。

本书图例说明

为了使读者更好地理解图例的含义，现说明如下，可作为学习本书各章节的参考。

用户



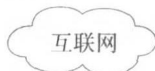
是指某种形式的计算机。不是特指任何一种操作系统，可以代表单独一个或多个用户。通常都有一个 TCP/IP 地址。

R(路由器)



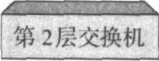
某厂商生产的、工作在第 3 层的路由器，运行某种形式的路由编码。可以是任意路由协议，并且一般和 Internet 保持连接。

Internet(云状图)



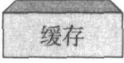
这个图例用来代表某个公共网络或私有网络，可能包含内部网络或外部网络。典型的，可能连有路由器、交换机以及其他互联设备，例如帧中继、ATM 等。

第2层交换机

第2层交换机

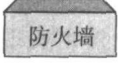
这个图例代表任何一个第2层交换机，可能含有10/100/1000 M以太网端口。这种交换机独立于厂商，实现了高性价比的连接。

缓存

缓存

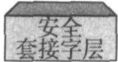
虽然有专用的缓存设备，但也可以在PC平台上运行执行缓存功能的软件。

防火墙

防火墙

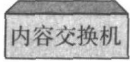
同样，这个图例可以是指作为某个复杂安全系统一部分的专用防火墙，也可能是基于安全因素而实施访问控制列表(ACL)的路由器。虽然第2层协议防火墙并不稀少，但一般是第3层设备。

安全套接字层(SSL 卸载设备)

安全套接字层

某一专门用来处理安全套接字层通信的硬件设备。

CS(内容交换机)

内容交换机

本图例代表工作在第4~7协议层的交换机。本书讨论的所有功能和特性在此类设备中均可用。

服务器



此图例可以代表单一服务器或多个服务器。这些服务器并不指定运行某种操作系统。若可以与基于 IP 协议的网络互联，通常可以达到负载平衡。

技术评论简介

Ray d'Urso 拥有 25 年的语音数据网络工作经验。在此期间，他在伦敦参与了多项大型网络工程建设，凭借他的经验，这些项目都获得了成功。目前，他是某跨国金融机构的技术服务负责人。

Dave Weal 至今已在 Data Networks 公司工作了 15 年。在此期间，他参与设计、实现和维护 X.25 网、ATM、帧中继、令牌环网，当然还有以太网和 TCP/IP 网络。

目前，他是北电网络公司的技术顾问，从技术上和关键网络需求的产品解决方案上为企业的主要客户提供咨询。

Scott Reeves 从 1989 年开始一直从事数据网络方面的工作，从 WAN 到 LAN，拥有较宽的技术面，主要从事系统集成、以太网、令牌环网以及 ATM 终端设备工作。

最近由于北电网络收购了 Alteon Web Systems(内容交换领域的先驱)，他进入了北电网络公司，职务是内容网络产品的欧洲开发先锋。

致谢

写第一本书对每个人来说都是令人心惊胆颤和富有挑战意义的经历，但我们不得不承认本书不是独立完成的。我们感谢技术评论 Ray d'Urso, Dave Weal 和 Scott Reeves，感谢他们宝贵的技术支持和文字工作。还要感谢 Mary Franz 和她在 Prentice Hall 的小组 12 个月以来的热情、鼓励和指导。虽然本书是独立讨论内容交换技术的书籍，但真正的英雄是 Alteon WebSystems 公司和北电网络公司的那些幕后支持者，正是他们编制了本书应用的工具软件。

感谢家人和朋友，他们给予了我很大的支持和鼓励。在这里，向你们说声谢谢。谢谢我深爱的 Fleur，这些日子你受苦受劳了。

Philip Goldie, 2002 年 12 月

我要对那些帮助过我的人说声谢谢。我希望能够给过去几年与我并肩作战的同事们一些回报。感谢 Ray d'Urso 在学习上给予我重大帮助，使我能够更透彻地研究问题。邀请他审阅本书是无尚的荣幸。还要感谢我的家人多年来对我的关爱与支持，特别是 Debra、Jessica 和 Cameron 在那些漫长的夜晚陪伴左右。真诚地对你们说声：谢谢。

Matthew Syme, 2002 年 12 月

小结

为了编写一本讨论新话题、振奋人心的书籍，使得即使经验丰富的专业人员也能通过本书完成复杂的配置或最终解决疑难问题，我们试图确保涉及了技术面，包括内容交换的优势和劣势，但没有讨论针对某个厂商的配置和 CLI 命令。我们的目标是探讨这种技术的强大功能和它所带来的好处。我们热衷于讨论这门技术，但它不能解决所有问题，许多地方都需要多加注意，并更好地设计和规划。总之，内容交换开辟了一个新的领域，使我们能够延伸知识，提高技能，希望还能够获得一些乐趣。我们希望本书成为你的良师益友。

目 录

第 1 章 内容交换技术入门	1
1.1 第 2 层和第 3 层网络的发展	1
1.1.1 更大更快的 Internet	1
1.1.2 更丰富内容的推动力	2
1.1.3 用内容交换解决问题	3
1.1.4 先行者	5
1.1.5 当前市场前景	7
1.2 小结	7
第 2 章 理解第 2 层、第 3 层和第 4 层协议	8
2.1 OSI 7 层模型——层的概念	8
2.1.1 应用层(第 7 层)	8
2.1.2 表示层(第 6 层)	9
2.1.3 会话层(第 5 层)	9
2.1.4 传输层(第 4 层)	9
2.1.5 网络层(第 3 层)	9
2.1.6 数据链路层(第 2 层)	10
2.1.7 物理层(第 1 层)	10
2.1.8 纵览 7 层	10
2.2 不同层之间的交换	12
2.2.1 第 2 层交换	13
2.2.2 第 3 层交换和路由	13
2.3 理解第 4 层协议	14
2.4 传输控制协议(TCP)	15
2.5 用户数据报协议(UDP)	20
2.6 虚拟路由器冗余协议(VRRP)	22
2.6.1 第 2, 3 层冗余	22
2.6.2 VRRP 的组成部分	23
2.6.3 VRRP 寻址	24

2.6.4	VRRP 的运行	24
2.7	小结	25
第3章	理解应用层协议	26
3.1	超文本传输协议 (HTTP)	26
3.1.1	基本 HTTP 页面取回	26
3.1.2	HTTP 方法	27
3.1.3	HTTP 中的 URL	28
3.1.4	HTTP 中的持久连接	28
3.1.5	其他 HTTP 报头	29
3.1.6	Cookie——HTTP 状态管理机制	31
3.1.7	HTTP 参考资料	33
3.2	文件传输协议 (FTP)	33
3.2.1	FTP 基础	34
3.2.2	FTP 主动方式	34
3.2.3	FTP 被动方式	35
3.2.4	FTP 参考资料	37
3.3	实时流协议 (RTSP)	37
3.3.1	RTSP 传送的组成	37
3.3.2	RTSP 操作	39
3.3.3	数据发送的其他选择	40
3.3.4	RTSP 及 RTP 的参考资料	42
3.4	安全套接字层 (SSL)	42
3.4.1	应用的安全需求	43
3.4.2	SSL 在 7 层模型中的位置	44
3.4.3	加密和密码	44
3.4.4	SSL 工作情况	50
3.4.5	SSL 总结	51
3.5	小结	51
第4章	内容交换的概念	52
4.1	虚拟服务和应用重定向	52
4.1.1	虚拟服务和虚拟 IP 地址	52
4.1.2	应用重定向	53
4.2	资源和真实服务器	54

4.3	帧、数据包和会话	54
4.4	深度数据包检测	55
4.5	小结	56
第5章	基本的服务器负载平衡	57
5.1	为什么要使用负载平衡服务器	57
5.2	服务器负载平衡的拓扑	58
5.2.1	第2层(桥接)服务器负载平衡	58
5.2.2	第3层(路由)服务器负载平衡	60
5.2.3	单臂服务器负载平衡	62
5.2.4	实现 SLB 的高度可用性	65
5.3	第4层负载平衡算法	71
5.3.1	最少连接法	72
5.3.2	循环法	72
5.3.3	IP 地址散列法	73
5.3.4	响应时间和服务器代理	74
5.3.5	带宽	74
5.3.6	加权和最大连接数	75
5.4	服务器诊断	75
5.4.1	基于连接的诊断	76
5.4.2	ARP 诊断	76
5.4.3	ICMP 诊断	76
5.4.4	TCP 诊断	76
5.4.5	应用诊断	77
5.4.6	诊断脚本	77
5.5	小结	77
第6章	内容识别服务器负载平衡	80
6.1	什么是第7层服务器负载平衡	80
6.2	为什么要使用第7层服务器负载平衡	80
6.2.1	示例1——所有的服务器并不都是相同的	81
6.2.2	示例2——所有的内容必须是随处可用的	81
6.2.3	示例3——购买服务	81
6.2.4	示例4——将来的验证	81
6.3	处理第7层通信	82

6.3.1	会话的即时绑定与延时绑定	83
6.3.2	使用延时绑定安全机制	84
6.3.3	第7层分析和连接: Keep-Alive 报头	85
6.4	HTTP URL 解析和负载均衡	86
6.5	HTTP 报头负载均衡	89
6.6	HTTP Cookie 负载均衡	94
6.7	FTP 负载均衡	97
6.7.1	主动模式下 FTP 的负载均衡	98
6.7.2	被动模式下 FTP 的负载均衡	100
6.8	第7层 DNS 的负载均衡	102
6.9	RTSP 流媒体的负载均衡	105
6.9.1	仅在第4层上对 RTSP 进行负载均衡	105
6.9.2	第7层 RTSP 负载均衡的应用	107
6.10	小结	110
第7章	持久性、安全性和 Internet	113
7.1	Internet 服务提供商——代理和通信量	114
7.2	IP 地址散列	116
7.3	基于 Cookie 的持久性	117
7.3.1	Cookie	117
7.3.2	Cookie 类型	119
7.3.3	应用 Cookie	119
7.4	SSL	121
7.4.1	SSL 会话 ID 持久性	121
7.4.2	使用 SSL 卸载修复此模型	122
7.5	小结	126
第8章	应用重定向	129
8.1	应用重定向的必要条件	129
8.2	基于 VIP 的 SLB 和应用重定向	130
8.3	Web 缓存重定向(WCR)	131
8.3.1	缓存的原理	131
8.3.2	缓存类型	135
8.3.3	IP 欺骗	140
8.3.4	第4层缓存重定向	141

8.3.5 第7层 WCR	145
8.3.6 供应商问题	149
8.4 安全重定向	150
8.5 路由器/链接负载平衡	151
8.6 小结	154
第9章 防火墙和 VPN 负载平衡	156
9.1 对防火墙和 VPN 交换机进行负载平衡的原因	156
9.2 防火墙简介	157
9.2.1 有状态防火墙	157
9.2.2 防火墙同步技术	158
9.2.3 第二层防火墙	158
9.3 配置防火墙的负载平衡	159
9.3.1 使用 VRRP	159
9.3.2 使用基于软件的方案	160
9.3.3 使用内容交换机	160
9.3.4 创建通道	160
9.3.5 健康检测防火墙	162
9.3.6 通过负载平衡防火墙三明治的通信流	162
9.3.7 网络地址转换防火墙	165
9.4 基于策略的防火墙负载平衡	167
9.4.1 拓扑示例	168
9.4.2 第二层防火墙	170
9.4.3 层式防火墙提供更好的安全性	171
9.4.4 使用内容交换机提供额外保护	171
9.4.5 增加非军事区	172
9.5 VPN 和 MAC 保持	175
9.5.1 VPN 的工作过程	175
9.5.2 负载平衡 VPN	176
9.5.3 失败情况	177
9.6 小结	179
9.6.1 案例研究: 防火墙和 VPN 负载平衡	179
9.6.2 配置防火墙和 VPN 负载平衡	180

第 10 章 内容交换机的体系结构	182
10.1 常见的第二层和第三层体系结构需要考虑的事项	182
10.2 内容交换机彼此不同的原因	183
10.2.1 数据包与会话	183
10.2.2 突破规则	183
10.3 好的内容交换机的组成	184
10.3.1 会话建立与并发会话	184
10.3.2 第 4~7 层处理	185
10.4 不同方法	186
10.4.1 PC 体系结构	186
10.4.2 基于 ASIC 的体系结构	187
10.5 新技术又将是什么	188
10.6 相关的硬件	188
10.7 小结	189