

Linux

服务器性能调整

Performance Tuning for Linux Servers

- Web、文件、打印、数据库和应用服务器优化的公认技术
- 性能管理开源工具的实用介绍
- 基于Red Hat Enterprise Linux和Novell SUSE Linux Enterprise Server的通俗示例

Sandra K. Johnson
(美) Gerrit Huizenga 著 韩智文 译
Badari Pulavarty



Linux 服务器性能调整

	Sandra K. Johnson	
(美)	Gerrit Huizenga	著
	Badari Pulavarty	
	韩智文	译

清华大学出版社

北 京

Authorized translation from the English language edition, entitled Performance Tuning for Linux[®] Servers, 9780137136285 by Sandra K. Johnson, Gerrit Huizenga, Badari Pulavarty, published by Pearson Education, Inc, publishing as IBM Press, Copyright © 2005.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by PEARSON EDUCATION ASIA LTD., and TSINGHUA UNIVERSITY PRESS Copyright © 2009.

北京市版权局著作权合同登记号 图字: 01-2009-4898

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

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

Linux 服务器性能调整/(美)约翰逊(Johnson, S. K.), (美)威曾格(Huizenga G.), (美)普拉瓦提(Pulavarty, B.) 著; 韩智文译. —北京: 清华大学出版社, 2009. 9

书名原文: Performance Tuning for Linux[®] Servers

ISBN 978-7-302-20532-6

I. L… II. ①约…②威…③普…④韩… III. Linux 操作系统 IV.TP316.89

中国版本图书馆 CIP 数据核字(2009)第 111077 号

责任编辑: 王 军 于 平

装帧设计: 孔祥丰

责任校对: 成凤进

责任印制: 杨 艳

出版发行: 清华大学出版社

地 址: 北京清华大学学研大厦 A 座

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

邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 刷 者: 清华大学印刷厂

装 订 者: 三河市溧源装订厂

经 销: 全国新华书店

开 本: 185×260 印 张: 27.5 字 数: 669 千字

版 次: 2009 年 9 月第 1 版 印 次: 2009 年 9 月第 1 次印刷

印 数: 1~4000

定 价: 58.00 元

本书如存在文字不清、漏印、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)62770177 转 3103 产品编号: 029688-01

译者序



计算机要解决的基本问题之一是如何在不增添更多硬件能力的情况下使其能够完成更多工作。在当前的全球经济形势下，企业面对更加严峻的成本压力，性能经过优化的服务器系统能够节省运营资金和时间，因此更加凸现了系统性能调优的重要性。Linux 服务器在许多企业中扮演着关键性的角色，这也意味着其性能具有极为重要的影响。当 Linux 作为企业级服务器操作系统使用时，需要监控其性能并在必要时进行调优，以便清除可能为用户带来影响的瓶颈。

性能调优是一门在天平上跳舞的科学和艺术。这是一项非常困难的任务，要求对计算机硬件、操作系统及应用都有深入的理解。在优化 Linux 服务器的系统性能时，有必要将性能调优技术、系统优化要素、合适的工具以及用于性能测量的服务器应用加以合理集成。

近年来，开源社区中对 Linux 系统优化的讨论日渐热烈，关于此类技术的书籍也不断增多。本书阐述了各种对运行于 Linux 系统之上的 Web、文件、打印、数据库和应用服务器加以优化的实用技术；具体描述了 Linux 的性能调优方法、服务器的性能监控和分析工具，以及特定服务器应用的关键调优参数。其目的是解释如何分析和调优 Linux 操作系统，以便为在其上运行的任何应用提供卓越的性能。

同时，本书将操作系统的先进设计理论、系统性能的基准测试方法以及大量的实际测试分析工作加以适当集成。它不流俗于浮浅的信息堆积，而是对于每个讨论主题既给出了详尽的技术背景信息和可调优参数，也交织着实用的配置和测试说明。作者基于多年积累的丰富实践经验提供了大量操作范例，从而有助于读者了解可用的技术并做出明智的调优决策。例如，第 8 章不仅提供了调度器可调优参数的列表，还包含了其描述信息和性能影响。

值得一提的是，本书虽由 IBM 出版社发行，但并非一部仅以 IBM 产品为中心的书籍。本书基于主要的企业级 Linux 平台 Red Hat Enterprise 和 Novell SUSE Linux Enterprise 展开论述，但大部分内容具有普遍适用性，可应用于几乎所有针对 2.6 内核以及部分针对 2.4 内核系列的 Linux 发行版本。书中多数内容只是用于阐释特定技术要点，而不要求读者遵循一种逐步执行过程。

本书由世界领先的 Linux 研发实验室——IBM Linux 技术中心的众多性能专家合著而成，凝结了群体的智慧。这种编撰形式极易导致整部书丧失一致的著作风格。但本书并非如此，它的各章内容结构流畅，风格统一，难以看出多位编著者斧凿的痕迹。本书作者在写作时不仅极好地凝练了资料素材，还将一些高度技术化的复杂概念深入浅出地呈现给读者。例如，第 2 章“内核概述”详细清晰地讨论了 Linux 内核的体系结构。相信本书将会成为长期驻留于读者案头上的参考指南。

译者

2008 年 12 月

作者简介



Sandra K. Johnson 博士是一名拥有 16 年经验的高级技术人员。她的研究领域涉及 cache 一致性协议、存储子系统、并行 I/O 子系统、Java 服务器以及 Linux 领域的工作负荷描述与性能分析。她是 IBM Linux 技术中心的 Linux 性能架构师，目前是 IBM 系统与技术事业部的全球中小型企业部门的首席技术官。

Gerrit Huizenga 是 IBM 俄勒冈州比弗顿 Linux 技术中心的 Linux 基础技术(Linux Base Technologies)的工程师和架构师。**Gerrit** 已经致力于架构、设计和实现操作系统的能力达 20 余年，重点关注性能、可扩展性、标准和安全。在为 IBM Linux 技术中心工作之前，**Gerrit** 是 Sequent Computer Systems 公司的操作系统首席专家。

Badari Pulavarty 是 IBM 俄勒冈州比弗顿 Linux 技术中心的高级工程师；具有 15 年开发 UNIX 操作系统的经验；主要研究领域是 Linux 在 VM、I/O、文件系统和存储子系统方面的性能问题。

贡献人员简介

Vaijayanthimala Anand, 高级软件工程师, 具有多年网络协议/驱动程序开发和体系结构设计经验。过去 3 年里她致力于 Linux 内核性能的研究。获 Houston 大学计算机科学硕士学位。

Steve Best 为 IBM 得克萨斯州奥斯汀 Linux 技术中心服务。他当前正在开发 Linux 项目中的 Journaled File System(JFS)。Steve 在操作系统开发领域完成了广泛的工作, 重点是文件系统、国际化及安全方面。

Edward G. Bradford 博士, 高级工程师。具有 15 年以上的 UNIX 操作系统开发及开发管理经验。目前致力于 Linux 与其他平台的性能比较问题。

Mark Brown, 具有 20 年以上的 UNIX 和 Linux 经验, 包括服务于 GNU C 库指导委员会(GNU C Library Steering Committee)。他的具体研究领域是操作系统 API 和 ABI 规范, 以及 C 运行库的问题。

Mingming Cao, IBM 俄勒冈州比弗顿 Linux 技术中心的 Linux 内核开发人员。她的研究领域包括进程间通信、文件系统和 I/O。

Ruth Forester, IBM 俄勒冈州比弗顿 Linux 技术中心的性能工程师。她从事终端用户和系统设计方面的 UNIX 和数据库性能研究已有 15 年。参与开发了 TPCD 并为 TPC 服务 7 年。她完成的工作涵盖应用程序性能以及一级、二级和三级 cache 的性能影响。

Steven French, IBM Linux 技术中心的高级工程师, 负责 Linux 文件系统设计, 具有 15 年以上的网络软件的设计和开发经验。是 Linux 内核中 CIFS 文件系统的编写者, Samba 团队成员, 存储网络产业联盟(Storage Networking Industry Association)的 CIFS 工作组主席。

Dominique Heger, 致力于研究操作系统性能、性能建模、算法和数据结构, 以及 I/O 扩展性。他曾为 IBM、Hewlett-Packard 以及 Unisys 工作过。获信息系统博士学位。

前言

廉价的硬件、运行的高稳定性以及不懈的技术开发造就了 Linux(开源)操作系统的不断发展。Linux 最常见的用途是充当 Web 服务器。当前, 它可以提供众多企业级服务, 如电子邮件、防火墙、代理、网关、数据库、应用程序、文件服务器、打印等。在更高端的企业服务器领域, Linux 也即将成为一种具有竞争力的操作系统。主要的华尔街公司、电影制作公司以及其他众多企业都在转向使用 Linux 以满足其对企业服务器的需求。

随着服务器的部署日益增多, 服务器操作系统的性能已成为系统整体效率的决定因素, 并将会影响到所有用户。虽然使用更新更快的服务器来替换整个服务器也是一种解决方案, 但更合适的办法常常是只替换或添加必需的组件而使其他组件保持不变。系统性能低下的原因多源于个别硬件子系统中存在着瓶颈、操作系统配置不当或者应用程序缺少性能优化等问题。

提高性能的关键是要理解系统中有哪些可用的软硬件配置选项, 理解性能工具及其用法, 对利用这些工具获取的结果进行分析并对系统进行适当调整, 从而对服务器性能产生正面影响。

本书提供了理解和改善 Linux 服务器性能所需的基本知识和技能。由 Linux 专业人员编写, 并基于大家共同的实际经验, 描述了 Linux 性能监控、评估、测量和分析的方法以及 Linux 服务器的调优方法。还讨论了对基于 Intel 硬件平台以及 Linux 操作系统运行的商业服务器应用程序的性能加以改进和优化的方法。

本书中采用的是简单易懂的渐进式性能调优方法。首先介绍 Linux 内核的有关知识, 然后讨论 Linux 服务器、性能工具和调优技术, 之后介绍 Linux 服务器的性能特征, 并给出稍加改动就可适用于各种情形的调优示例。

本书中的论述主要基于两种 Linux 版本: Red Hat Enterprise Linux 和 Novel SUSE Linux Enterprise Server。另外, 假定服务器基础硬件为基于 Intel 的平台, 这是因为这种硬件平台在 Linux 社区中得到普遍使用。但 Linux 服务器的性能调优机制也很容易移植到其他平台上。服务器的具体应用包括 Web 服务器、文件和打印服务器、数据库服务器、网络管理服务器以及 Web 应用服务器等。本书中涉及到的相关讨论和示例都基于 Linux 2.4、2.5 和 2.6 内核。

本书由 5 部分构成: “Linux 概述”、“性能分析工具”、“系统调优”、“Linux 服务器应用的性能特征”与“调优案例分析”。

第 I 部分“Linux 概述”简述了与 Linux 安装以及 Linux 内核和服务器体系结构相关的问题。

第 II 部分“性能分析工具”详细介绍了与 3 个领域相关的性能工具: 系统性能和监控、

系统跟踪以及基准测试。

第III部分“系统调优”讨论了性能调优的原则与策略，还讨论了调度器、内存、I/O、文件系统以及网络 and 代码的调优机制。

第IV部分“Linux 服务器应用的性能特征”描述了服务器应用程序的性能特征，包括影响其性能的潜在因素。还给出了关于 Web、文件、数据库和应用程序服务器的讨论。

第V部分“调优案例分析”给出了大量示例来阐释如何将前述各部分中的所有工作加以集成，从而实施一个包含各种场景的调优方案。所提供的案例分析包括调度器、I/O、文件系统以及网络和商业工作负荷的调优。

本书的附录中讨论了许多可调优的内核参数，并阐述了使用 `sysctl`、`/proc` 和 `sysfs` 接口来调优 Linux 内核的方法。

目 录

第 I 部分 Linux 概 述

第 1 章 Linux 系统安装的性能问题.....3

1.1 引言.....3

1.2 安装前的规划.....3

1.2.1 选择分区位置.....4

1.2.2 使用多个硬盘驱动器.....4

1.2.3 选择文件系统.....4

1.2.4 转换文件系统.....5

1.2.5 配置 RAID.....6

1.3 Linux 2.6 内核的可配置属性.....8

1.3.1 I/O 提升器.....8

1.3.2 超大 TLB 页面支持.....8

1.4 Linux 日志工具.....9

1.4.1 /var/log/messages 文件.....9

1.4.2 /var/log/XFree86.0.log 文件.....9

1.4.3 日志轮转.....10

1.4.4 日志工具.....10

1.4.5 日志定制.....10

1.5 BSD 与系统 V 的初始化机制.....11

1.5.1 初始化表(/etc/inittab).....11

1.5.2 BSD 的初始化表(Slackware).....12

1.5.3 系统 V 的初始化表(Red Hat).....13

1.6 小结.....14

1.7 参考文献.....14

第 2 章 Linux 内核机制.....15

2.1 引言.....15

2.2 Linux 的发展历程.....15

2.3 Linux 内核体系结构.....15

2.3.1 内核的职能.....16

2.3.2 内核组织结构与模块.....16

2.3.3 内核服务.....16

2.3.4 /proc 文件系统的外部性能视图.....16

2.3.5 内存管理.....17

2.4 进程管理.....18

2.4.1 进程、任务与内核线程.....19

2.4.2 调度与上下文切换.....19

2.5 进程间通信.....20

2.5.1 信号.....20

2.5.2 管道.....20

2.5.3 系统 V 的 IPC 机制.....21

2.6 Linux 对称多处理(SMP)模型.....22

2.6.1 多处理系统类型.....22

2.6.2 同步与数据串行化.....22

2.6.3 锁、锁粒度与锁开销.....22

2.6.4 cache 一致性.....23

2.6.5 处理器亲和度.....23

2.7 文件系统.....23

2.7.1 虚拟文件系统(VFS).....23

2.7.2 ext2fs.....23

2.7.3 LVM 与 RAID.....24

2.7.4 磁盘卷组.....24

2.7.5 设备相关文件.....25

2.7.6 devfs.....25

2.8 Linux 2.6 内核的新特性.....26

2.9 小结.....26

2.10 参考文献.....27

第3章 服务器体系结构	29
3.1 引言	29
3.2 Linux 服务器	29
3.3 多处理器与多处理机制	30
3.3.1 服务器拓扑结构	30
3.3.2 处理器的混用机制	31
3.4 内存	32
3.5 I/O	33
3.6 Linux 企业级服务器	34
3.7 Linux 集群	35
3.7.1 高性能集群	35
3.7.2 高可用集群	36
3.8 服务器系统示例	36
3.8.1 IBM zSeries 大型机	36
3.8.2 刀片服务器	39
3.8.3 NUMA	39
3.9 小结	44

第II部分 性能分析工具

第4章 系统性能监控	47
4.1 引言	47
4.2 Linux 与性能分析的背景知识	47
4.3 CPU 利用率	49
4.3.1 vmstat	51
4.3.2 top 与 gtop 工具	53
4.3.3 sar	54
4.4 内存利用率	56
4.4.1 /proc/meminfo 与 /proc/slabinfo	56
4.4.2 ps	57
4.4.3 vmstat	59
4.5 I/O 利用率	59
4.5.1 iostat	61
4.5.2 sar	62
4.6 网络利用率	63
4.6.1 网络统计信息	64
4.6.2 接口信息	66
4.6.3 TCP/IP 协议统计数据	66
4.6.4 nfsstat	67

4.7 小结	67
4.8 参考文献	67
第5章 系统跟踪工具	69
5.1 引言	69
5.2 系统跟踪机制的需求	69
5.3 top	70
5.4 strace	72
5.5 OProfile	74
5.5.1 opcontrol	75
5.5.2 数据概况描述工具	76
5.6 Performance Inspector	82
5.6.1 Above Idle	83
5.6.2 Per-Thread Time	84
5.6.3 Trace Profiling	86
5.6.4 指令跟踪	89
5.6.5 Java Profiler	91
5.6.6 Java Lock Monitor	103
5.6.7 Performance Inspector 的执行工具	106
5.7 小结	107
5.8 参考文献	107
第6章 工作负荷的性能基准测试	109
6.1 引言	109
6.2 改进工作负荷的基准测试方法	110
6.3 基准测试的类型	111
6.4 微基准测试	111
6.4.1 操作系统基准测试	111
6.4.2 磁盘基准测试	124
6.4.3 网络基准测试	132
6.4.4 应用基准测试	133
6.5 Web 服务器基准测试	137
6.5.1 SPECweb、SPECweb SSL 与 TPC-W	137
6.5.2 SPECjAppServer 与 ECPeef	138
6.5.3 其他应用基准测试	139
6.6 小结	139

第III部分 系统调优

第7章 系统性能原理和策略:

基准测试方法的案例分析 143

7.1 引言 143

7.2 性能评价方法 143

7.2.1 跟踪机制 143

7.2.2 工作负荷特征描述 144

7.2.3 数值分析 144

7.2.4 模拟方法 144

7.3 基准测试案例分析 144

7.4 分析方法 145

7.4.1 软硬件配置 145

7.4.2 运行规则 146

7.4.3 设置目标 146

7.4.4 测量、分析与调优 146

7.4.5 退出策略 147

7.5 基准测试程序 148

7.5.1 基准测试程序介绍 148

7.5.2 性能结果 149

7.6 小结 151

7.7 致谢 152

7.8 参考文献 152

第8章 调度器调优 153

8.1 引言 153

8.2 单处理器系统 153

8.3 对称多处理 154

8.4 非一致内存访问 154

8.5 对称多线程 155

8.6 Linux 2.6 内核调度器 155

8.7 负载平衡 156

8.8 调度器的可调参数 157

8.8.1 *CHILD_PENALTY 157

8.8.2 *CREDIT_LIMIT 157

8.8.3 *EXIT_WEIGHT 157

8.8.4 *INTERACTIVE_DELTA 157

8.8.5 *MAX_SLEEP_AVG 158

8.8.6 *MAX_TIMESLICE 158

8.8.7 *MIN_TIMESLICE 158

8.8.8 *PARENT_PENALTY 158

8.8.9 *PRIO_BONUS_RATIO 158

8.8.10 *STARVATION_LIMIT 158

8.9 小结 159

8.10 参考文献 159

第9章 Linux 虚存的性能问题 161

9.1 引言 161

9.2 内存与地址空间 162

9.2.1 地址空间 162

9.2.2 用户地址空间 162

9.2.3 VM 区域 163

9.2.4 内核地址空间 164

9.3 高端内存支持 165

9.4 分页与交换机制 165

9.4.1 替换策略 166

9.4.2 页面替换与内存平衡 167

9.5 Linux 页表 167

9.6 Linux 2.6 内核中的新特性 169

9.6.1 rmap 与 objrmap 169

9.6.2 大型页面的支持 169

9.6.3 页面分配与替换 170

9.6.4 Slab 分配器 171

9.6.5 VM 的可调参数 171

9.6.6 CPU 调度器 173

9.7 小结 175

9.8 参考文献 175

第10章 I/O 子系统的性能问题 177

10.1 引言 177

10.2 I/O 调度与块 I/O(BIO)层 177

10.2.1 Linux 2.6 内核的 I/O

调度器 178

10.2.2 Linux 2.4 内核的 I/O

调度器 179

10.2.3 Linux 2.6 内核的最终期限

I/O 调度器 179

10.2.4 调度器的可调参数 180

10.2.5	Linux 2.6 内核的 Anticipatory I/O 调度器	181
10.2.6	Linux 2.6 内核的 CFQ 调度器	183
10.2.7	Linux 2.6 内核的 noop I/O 调度器	183
10.2.8	I/O 调度器的性能含义	183
10.3	批量读写请求	184
10.4	读预测启发算法	184
10.5	影响性能的 I/O 部件	185
10.6	I/O 设备寻址	186
10.7	小结	186
10.8	参考文献	186
第 11 章	文件系统调优	189
11.1	引言	189
11.2	文件系统基本知识	189
11.2.1	文件系统的实现考虑	189
11.2.2	创建优化的文件系统	190
11.2.3	文件系统基本术语	190
11.3	日志型文件系统	192
11.3.1	文件系统的故障处理机制	193
11.3.2	事务机制	193
11.3.3	选项	194
11.4	影响文件系统性能的 磁盘因素	195
11.4.1	物理磁盘存储器的 考虑事项	195
11.4.2	磁盘的性能特征	196
11.4.3	传输速率	196
11.4.4	hdparm	197
11.5	文件系统的碎片整理	199
11.6	文件同步	199
11.6.1	确保数据或文件完整性	200
11.6.2	使用函数调用	200
11.6.3	使用文件描述符	200
11.6.4	sync/fsync 的性能含义	201
11.7	bdflush 参数	201
11.8	异步输入与输出	202
11.8.1	使用异步 I/O	202
11.8.2	异步 I/O 函数	202
11.9	原始磁盘 I/O	203
11.9.1	在 Linux 上设置原始 I/O	203
11.9.2	示例	203
11.10	Ext2 与 Ext3 文件系统	204
11.10.1	Ext2 的组织结构	204
11.10.2	Ext2 文件系统 的块分配方式	206
11.10.3	创建 Ext2 文件系统	206
11.10.4	Ext2 文件系统的 Ext3 扩展机制	207
11.10.5	Ext3 的内核配置支持	207
11.10.6	Ext3 文件系统调优方法	208
11.10.7	创建 Ext3 分区	208
11.10.8	Ext2 与 Ext3 的转换	209
11.10.9	建立外部日志	209
11.10.10	Ext2/Ext3 工具	209
11.11	ReiserFS	210
11.11.1	ReiserFS 的内核 配置支持	211
11.11.2	ReiserFS 文件系统 调优方法	211
11.11.3	创建 ReiserFS 文件系统	211
11.11.4	建立外部日志	212
11.11.5	挂接 ReiserFS 文件系统	212
11.11.6	挂接选项	213
11.11.7	调优 ReiserFS	213
11.11.8	ReiserFS 文件系统 实用工具	214
11.12	日志式文件系统	214
11.12.1	JFS 的内核配置支持	215
11.12.2	JFS 文件系统的 调优方法	215
11.12.3	创建 JFS 文件系统	216
11.12.4	建立外部日志	216
11.12.5	挂接文件系统	216
11.12.6	挂接选项	216
11.12.7	调优 JFS	217

11.12.8	JFS 文件系统实用工具	217	13.4.2	semmsns	239
11.13	下一代文件系统	217	13.4.3	semmsl	239
11.13.1	XFS 的内核配置支持	218	13.4.4	semopm	239
11.13.2	XFS 文件系统的 调优方法	218	13.4.5	semvmx	239
11.13.3	创建 XFS 文件系统	219	13.4.6	未用的信号量参数	239
11.13.4	容量饱和时的文件 系统行为	219	13.5	消息队列参数	240
11.13.5	建立外部日志	219	13.5.1	msgmni	240
11.13.6	挂接文件系统	220	13.5.2	msgmax	240
11.13.7	挂接选项	220	13.5.3	msgmnb	241
11.13.8	调优 XFS	220	13.5.4	未用的消息队列参数	241
11.13.9	XFS 文件系统实用工具	221	13.6	共享内存段参数	241
11.14	小结	222	13.6.1	shmmni	242
11.15	参考文献	223	13.6.2	shmmax	242
第 12 章	网络调优	225	13.6.3	shmmmin	242
12.1	引言	225	13.6.4	shmall	242
12.2	网络协议栈	225	13.7	可配置 IPC 参数的动态 修改机制	242
12.3	内核参数调优机制	226	13.7.1	使用/proc	242
12.4	内核自动调优机制	226	13.7.2	使用 sysctl	243
12.5	核心内核参数	227	13.8	IPC 参数的静态配置	243
12.6	TCP/IPv4 协议内核参数	228	13.9	管道机制	243
12.6.1	TCP 缓冲区与内存管理	228	13.10	小结	244
12.6.2	TCP 选项	230	第 14 章	代码调优	245
12.6.3	TCP 连接管理	231	14.1	引言	245
12.6.4	TCP 连接保持管理	232	14.2	一般原则	245
12.6.5	IP 端口范围	233	14.3	应用的概要描述工具	246
12.7	小结	233	14.4	编译器选项调优	246
12.8	参考文献	233	14.4.1	性能调优的基本步骤	247
第 13 章	进程间通信	235	14.4.2	编译器优化的问题	247
13.1	引言	235	14.5	代码调优	247
13.2	进程间通信的定义	235	14.6	算法设计调优	248
13.3	Linux SysV 的 IPC 资源 与 ipcs 命令	236	14.6.1	问题与解决方案的可能性	248
13.3.1	ipcs 命令	236	14.6.2	问题描述	248
13.3.2	IPC 标识符及其限制	237	14.6.3	程序	249
13.4	信号量参数	238	14.6.4	设计代码	251
13.4.1	semmsni	238	14.6.5	服务器	251
			14.6.6	计时	252
			14.6.7	Socket	253

14.6.8	线程	256
14.6.9	同步	259
14.6.10	文件 I/O	262
14.6.11	客户端	264
14.6.12	代码讨论	266
14.6.13	编译选项	266
14.6.14	链接库	267
14.7	小结	268

第IV部分 Linux 服务器应用的性能特征

第 15 章	Web 服务器的性能调优	271
15.1	引言	271
15.2	HTTP 请求与响应	272
15.3	Web 服务器的网络行为	273
15.4	Web 服务器事务	275
15.5	Web 服务器模型	275
15.6	Web 服务器的调优方法	276
15.6.1	Web 服务器的通用 调优机制	277
15.6.2	Apache 服务器的 调优机制	277
15.6.3	Flash 与事件驱动的其他 服务器的调优机制	278
15.6.4	Tux 的调优机制	278
15.6.5	Web 服务器的性能 评价工具	278
15.7	小结	279
15.8	参考文献	279

第 16 章	文件与打印服务器的 性能调优	281
16.1	引言	281
16.2	专用网络存储服务器的类型	282
16.3	网络存储性能的优化方法	282
16.3.1	确定远程存储的数据	282
16.3.2	SAN 与网络文件 系统/NAS	282
16.3.3	网络文件系统协议	283

16.3.4	客户与服务器的实现选择	286
16.3.5	Linux 客户端调优的 关键概念	287
16.3.6	Linux 文件服务器的 调优机制	290
16.3.7	性能测量	292
16.3.8	改进容量规划的负荷 测量方法	292
16.3.9	打印服务器的性能因素	293
16.4	参考文献	293

第 17 章	数据库服务器的性能调优	295
17.1	引言	295
17.2	数据库体系结构	295
17.3	数据库的性能调优领域	296
17.3.1	I/O 调优	296
17.3.2	队列长度与响应时间	296
17.3.3	负载均衡	297
17.3.4	全局内存	298
17.3.5	日志设备	299
17.3.6	典型的数据库工作负荷	300
17.4	进程管理	300
17.5	内存管理	301
17.6	I/O 管理	302
17.6.1	回弹缓冲区	302
17.6.2	原始 I/O	302
17.6.3	向量 I/O	303
17.6.4	异步 I/O	303
17.6.5	直接 I/O	303
17.6.6	块 I/O	304
17.6.7	I/O 请求锁	304
17.7	小结	304

第 18 章	应用服务器的性能调优	305
18.1	引言	305
18.2	应用服务器概念	305
18.3	Java、J2EE 与应用服务器	305
18.4	应用服务器的性能特征	307
18.4.1	应用服务器的特征	308
18.4.2	Linux 上的应用服务器	308

18.5	性能与高可用性的改进	316
18.5.1	SMP 扩展性	316
18.5.2	集群	317
18.5.3	拓扑结构	319
18.5.4	性能调优	322
18.6	小结	324
18.7	参考文献	325

第 V 部分 调优案例分析

第 19 章 Linux 2.6 内核 I/O 调度器

调优案例分析

19.1	引言	329
19.2	基准测试环境与工作 负荷概况	330
19.3	I/O 调度器与性能	331
19.4	单 CPU 单磁盘配置	331
19.5	8 路 RAID-5 配置	332
19.6	16 路 RAID-0 配置	335
19.7	AS 串行读性能	335
19.8	AS 与 Deadline 的性能	336
19.9	CFQ 性能	337
19.10	小结	339
19.11	参考文献	339

第 20 章 文件系统调优实例分析

20.1	引言	341
20.2	文件布局分析	341
20.2.1	Ext2/Ext3 文件系统布局	341
20.2.2	日志文件系统(JFS)布局	347
20.2.3	ReiserFS 文件系统布局	349
20.2.4	XFS 文件系统布局	351
20.3	文件系统的调优机制	353
20.3.1	Ext3 的外部日志调优选项	353
20.3.2	ReiserFS 的外部日志 调优选项	357
20.3.3	JFS 的外部日志调优选项	358
20.3.4	XFS 的调优选项	359
20.4	I/O 的测量方法	361

20.4.1	iostat	361
20.4.2	iostat 与 sar 工具	362
20.5	小结	367
20.6	参考文献	367

第 21 章 Linux 系统网络性能

实例分析

21.1	引言	369
21.2	实例分析中使用的基准测试	369
21.2.1	NetBench	370
21.2.2	Netperf3	370
21.2.3	VolanoMark	370
21.2.4	SPECWeb99	370
21.3	Linux 2.4 与 2.6 内核中 的增强机制	371
21.3.1	SendFile	371
21.3.2	TCP 分段卸载	373
21.3.3	网络负荷中的进程与 IRQ 亲合性	375
21.3.4	NAPI	376
21.3.5	TCP 卸载引擎	378
21.4	示例分析	379
21.4.1	NetBench	380
21.4.2	Netperf3(千兆位以太网 调优实例分析)	381
21.4.3	VolanoMark	383
21.4.4	SPECWeb99	384
21.5	小结	386
21.6	参考文献	387

第 22 章 商用负载调优实例分析

22.1	引言	389
22.2	商用负载调优概述	389
22.3	J2EE 的标准商用负载模型	390
22.4	商用负载模型实例： 股票交易	390
22.4.1	系统配置	391
22.4.2	Trade3 的详细信息	392
22.4.3	性能分析方法	395
22.5	性能分析试验	397

22.5.1	Web 服务器	398
22.5.2	数据库后端程序	399
22.5.3	Java 虚拟机	404
22.5.4	应用服务器	406
22.5.5	超线程	408
22.6	小结	409
22.7	参考文献	410
附录 A	内核参数调优	411
A.1	引言	411
A.2	sysctl 接口	411
A.3	procfs 接口	412
A.4	sysfs(只针对 Linux 2.6 内核)	413
A.5	通用内核参数	413
A.5.1	共享内存	413
A.5.2	进程	414
A.5.3	信号	414
A.5.4	概要分析/调试	414
A.5.5	系统	414
A.6	虚存参数	415
A.7	文件系统参数	416
A.8	网络内核参数	417
A.8.1	网络核心参数	418
A.8.2	ICMP	418
A.8.3	邻居	419
A.8.4	配置	419
A.8.5	路由	420
A.8.6	TCP	420
A.8.7	IP 分段	421