# Optimizing Linux® Performance

## A Hands-On Guide to Linux® Performance Tools

The first comprehensive, expert guide for end-to-end Linux application optimization

Learn to choose the right tools—and use them *together* to solve real problems in real production environments

UYÊN
IỆU

Phillip G. Ezolt

hp invent

# Optimizing Linux® Performance

# Hewlett-Packard® Professional Books

## HP-UX

| | |
|---|---|
| **Cooper/Moore** | **HP-UX 11i Internals** |
| **Fernandez** | **Configuring CDE** |
| **Keenan** | **HP-UX CSE: Official Study Guide and Desk Reference** |
| **Madell** | **Disk and File Management Tasks on HP-UX** |
| **Olker** | **Optimizing NFS Performance** |
| **Poniatowski** | **HP-UX 11i Virtual Partitions** |
| **Poniatowski** | **HP-UX 11i System Administration Handbook and Toolkit, Second Edition** |
| **Poniatowski** | **The HP-UX 11.x System Administration Handbook and Toolkit** |
| **Poniatowski** | **HP-UX 11.x System Administration "How To" Book** |
| **Poniatowski** | **HP-UX 10.x System Administration "How To" Book** |
| **Poniatowski** | **HP-UX System Administration Handbook and Toolkit** |
| **Poniatowski** | **Learning the HP-UX Operating System** |
| **Rehman** | **HP-UX CSA: Official Study Guide and Desk Reference** |
| **Sauers/Ruemmler/Weygant** | **HP-UX 11i Tuning and Performance** |
| **Weygant** | **Clusters for High Availability, Second Edition** |
| **Wong** | **HP-UX 11i Security** |

## UNIX, LINUX

| | |
|---|---|
| **Mosberger/Eranian** | **IA-64 Linux Kernel** |
| **Poniatowski** | **Linux on HP Integrity Servers** |
| **Poniatowski** | **UNIX User's Handbook, Second Edition** |
| **Stone/Symons** | **UNIX Fault Management** |

## COMPUTER ARCHITECTURE

| | |
|---|---|
| **Evans/Trimper** | **Itanium Architecture for Programmers** |
| **Kane** | **PA-RISC 2.0 Architecture** |
| **Markstein** | **IA-64 and Elementary Functions** |

## NETWORKING/COMMUNICATIONS

| | |
|---|---|
| **Blommers** | **Architecting Enterprise Solutions with UNIX Networking** |
| **Blommers** | **OpenView Network Node Manager** |
| **Blommers** | **Practical Planning for Network Growth** |
| **Brans** | **Mobilize Your Enterprise** |
| **Cook** | **Building Enterprise Information Architecture** |
| **Lucke** | **Designing and Implementing Computer Workgroups** |
| **Lund** | **Integrating UNIX and PC Network Operating Systems** |

## SECURITY

| | |
|---|---|
| **Bruce** | **Security in Distributed Computing** |
| **Mao** | **Modern Cryptography: Theory and Practice** |
| **Pearson et al.** | **Trusted Computing Platforms** |
| **Pipkin** | **Halting the Hacker, Second Edition** |
| **Pipkin** | **Information Security** |

## WEB/INTERNET CONCEPTS AND PROGRAMMING

| | |
|---|---|
| **Amor** | **E-business (R)evolution, Second Edition** |
| **Apte/Mehta** | **UDDI** |
| **Chatterjee/Webber** | **Developing Enterprise Web Services: An Architect's Guide** |
| **Kumar** | **J2EE Security for Servlets, EJBs, and Web Services** |

| | |
|---|---|
| **Little/Maron/Pavlik** | Java Transaction Processing |
| **Mowbrey/Werry** | Online Communities |
| **Tapadiya** | .NET Programming |

**OTHER PROGRAMMING**

| | |
|---|---|
| **Blinn** | Portable Shell Programming |
| **Caruso** | Power Programming in HP OpenView |
| **Chaudhri** | Object Databases in Practice |
| **Chew** | The Java/C++ Cross Reference Handbook |
| **Grady** | Practical Software Metrics for Project Management and Process Improvement |
| **Grady** | Software Metrics |
| **Grady** | Successful Software Process Improvement |
| **Lee/Schneider/Schell** | Mobile Applications |
| **Lewis** | The Art and Science of Smalltalk |
| **Lichtenbelt** | Introduction to Volume Rendering |
| **Mellquist** | SNMP++ |
| **Mikkelsen** | Practical Software Configuration Management |
| **Norton** | Thread Time |
| **Tapadiya** | COM+ Programming |
| **Yuan** | Windows 2000 GDI Programming |

**STORAGE**

| | |
|---|---|
| **Thornburgh** | Fibre Channel for Mass Storage |
| **Thornburgh/Schoenborn** | Storage Area Networks |
| **Todman** | Designing Data Warehouses |

**IT/IS**

| | |
|---|---|
| **Anderson** | mySAP Tool Bag for Performance Tuning and Stress Testing |
| **Missbach/Hoffman** | SAP Hardware Solutions |

**IMAGE PROCESSING**

| | |
|---|---|
| **Crane** | A Simplified Approach to Image Processing |
| **Gann** | Desktop Scanners |

# Optimizing Linux® Performance

## A Hands-On Guide to Linux® Performance Tools

Phillip G. Ezolt

www.hp.com/hpbooks

Pearson Education

Upper Saddle River, NJ • Boston• Indianapolis • San Francisco

New York • Toronto • Montreal • London • Munich • Paris • Madrid

Capetown • Sydney • Tokyo • Singapore • Mexico City

Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed with initial capital letters or in all capitals.

The author and publisher have taken care in the preparation of this book, but make no expressed or implied warranty of any kind and assume no responsibility for errors or omissions. No liability is assumed for incidental or consequential damages in connection with or arising out of the use of the information or programs contained herein.

The publisher offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales, which may include electronic versions and/or custom covers and content particular to your business, training goals, marketing focus, and branding interests. For more information, please contact:

U. S. Corporate and Government Sales
(800) 382-3419
corpsales@pearsontechgroup.com

For sales outside the U. S., please contact:

International Sales
international@pearsoned.com

Visit us on the Web: www.phptr.com

Library of Congress Number: 2004117118

Copyright © 2005 Hewlett-Packard Development Company, L.P.

All rights reserved. Printed in the United States of America. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, write to:

Pearson Education, Inc.
Rights and Contracts Department
One Lake Street
Upper Saddle River, NJ 07458

ISBN 0-13-148682-9

Text printed in the United States on recycled paper at RR Donnelley & Sons Company in Crawfordsville, IN

First printing, March 2005

*This book is dedicated to my wife Sarah, (the best in the world), who gave up so many weekends to make this book possible. Thank you, Thank you, Thank you!*

# Table of Contents


| | | |
|---|---|---|
| | Preface | xvii |
| Chapter 1 | Performance Hunting Tips | 1 |
| | 1.1 General Tips | 2 |
| | 1.1.1 Take Copious Notes (Save Everything) | 2 |
| | 1.1.2 Automate Redundant Tasks | 5 |
| | 1.1.3 Choose Low-Overhead Tools If Possible | 5 |
| | 1.1.4 Use Multiple Tools to Understand the Problem | 6 |
| | 1.1.5 Trust Your Tools | 7 |
| | 1.1.6 Use the Experience of Others (Cautiously) | 7 |
| | 1.2 Outline of a Performance Investigation | 8 |
| | 1.2.1 Finding a Metric, Baseline, and Target | 8 |
| | 1.2.2 Track Down the Approximate Problem | 11 |
| | 1.2.3 See Whether the Problem Has Already Been Solved | 11 |
| | 1.2.4 The Case Begins (Start to Investigate) | 13 |
| | 1.2.5 Document, Document, Document | 14 |
| | 1.3 Chapter Summary | 14 |
| Chapter 2 | Performance Tools: System CPU | 15 |
| | 2.1 CPU Performance Statistics | 15 |
| | 2.1.1 Run Queue Statistics | 16 |
| | 2.1.2 Context Switches | 16 |
| | 2.1.3 Interrupts | 17 |
| | 2.1.4 CPU Utilization | 18 |
| | 2.2 Linux Performance Tools: CPU | 18 |
| | 2.2.1 vmstat (Virtual Memory Statistics) | 19 |
| | 2.2.2 top (v. 2.0.x) | 24 |
| | 2.2.3 top (v. 3.x.x) | 31 |
| | 2.2.4 procinfo (Display Info from the /proc File System) | 37 |
| | 2.2.5 gnome-system-monitor | 40 |
| | 2.2.6 mpstat (Multiprocessor Stat) | 42 |
| | 2.2.7 sar (System Activity Reporter) | 45 |
| | 2.2.8 oprofile | 51 |
| | 2.3 Chapter Summary | 60 |

Chapter 3 Performance Tools: System Memory 61

3.1 Memory Performance Statistics 61
3.1.1 Memory Subsystem and Performance 62
3.1.2 Memory Subsystem (Virtual Memory) 62
3.2 Linux Performance Tools: CPU and Memory 65
3.2.1 vmstat (Virtual Memory Statistics) II 65
3.2.2 top (2.x and 3.x) 72
3.2.3 procinfo II 75
3.2.4 gnome-system-monitor (II) 78
3.2.5 free 79
3.2.6 slabtop 82
3.2.7 sar (II) 85
3.2.8 /proc/meminfo 89
3.3 Chapter Summary 92

Chapter 4 Performance Tools: Process-Specific CPU 93

4.1 Process Performance Statistics 93
4.1.1 Kernel Time Versus User Time 94
4.1.2 Library Time Versus Application Time 94
4.1.3 Subdividing Application Time 94
4.2 The Tools 95
4.2.1 time 95
4.2.2 strace 99
4.2.3 ltrace 102
4.2.4 ps (Process Status) 107
4.2.5 ld.so (Dynamic Loader) 109
4.2.6 gprof 112
4.2.7 oprofile (II) 117
4.2.8 Languages: Static (C and C++) Versus Dynamic (Java and Mono) 124
4.3 Chapter Summary 125

Chapter 5 Performance Tools: Process-Specific Memory 127

5.1 Linux Memory Subsystem 127
5.2 Memory Performance Tools 128
5.2.1 ps 129

5.2.2 /proc/<PID> 131
5.2.3 memprof 135
5.2.4 valgrind (cachegrind) 139
5.2.5 kcachegrind 146
5.2.6 oprofile (III) 150
5.2.7 ipcs 157
5.2.8 Dynamic Languages (Java, Mono) 162
5.3 Chapter Summary 162

Chapter 6 Performance Tools: Disk I/O 163

6.1 Introduction to Disk I/O 163
6.2 Disk I/O Performance Tools 164
6.2.1 vmstat (ii) 164
6.2.2 iostat 171
6.2.3 sar 176
6.2.4 lsof (List Open Files) 178
6.3 What's Missing? 181
6.4 Chapter Summary 182

Chapter 7 Performance Tools: Network 183

7.1 Introduction to Network I/O 184
7.1.1 Network Traffic in the Link Layer 185
7.1.2 Protocol-Level Network Traffic 186
7.2 Network Performance Tools 187
7.2.1 mii-tool (Media-Independent Interface Tool) 187
7.2.2 ethtool 188
7.2.3 ifconfig (Interface Configure) 189
7.2.4 ip 191
7.2.5 sar 194
7.2.6 gkrellm 197
7.2.7 iptraf 200
7.2.8 netstat 202
7.2.9 etherape 205
7.3 Chapter Summary 207