

Shell Scripting

Expert Recipes for Linux*, Bash, and More

SHELL SCRIPTING EXPERT RECIPES FOR LINUX, BASH, AND MORE

INTRODUCTIO	ON	xxix
► PART I	ABOUT THE INGREDIENTS	
CHAPTER 1	The History of Unix, GNU, and Linux	3
CHAPTER 2	Getting Started	
CHAPTER 3	Variables	33
CHAPTER 4	Wildcard Expansion	67
CHAPTER 5	Conditional Execution	83
CHAPTER 6	Flow Control Using Loops	111
CHAPTER 7	Variables Continued	139
CHAPTER 8	Functions and Libraries	161
CHAPTER 9	Arrays	199
CHAPTER 10	Processes	219
CHAPTER 11	Choosing and Using Shells	255
► PART II	RECIPES FOR USING AND EXTENDING SYSTEM T	OOLS
CHAPTER 12	File Manipulation	279
CHAPTER 13	Text Manipulation	315
CHAPTER 14	Tools for Systems Administration	353
► PART III	RECIPES FOR SYSTEMS ADMINISTRATION	
CHAPTER 15	Shell Features	409
CHAPTER 16	Systems Administration	427
CHAPTER 17	Presentation	459
CHAPTER 18	Data Storage and Retrieval	471
CHAPTER 19	Numbers	483
CHAPTER 20	Processes	501
CHAPTER 21	Internationalization	517

► PART IV REFERENCE

APPENDIX	Further Reading	529
GLOSSARY .		. 533
INDEX		. 539

Shell Scripting

EXPERT RECIPES FOR LINUX, BASH, AND MORE

Shell Scripting

EXPERT RECIPES FOR LINUX, BASH, AND MORE

Steve Parker



Shell Scripting: Expert Recipes for Linux, Bash, and More

Published by John Wiley & Sons, Inc. 10475 Crosspoint Boulevard Indianapolis, IN 46256 www.wiley.com

Copyright © 2011 by Steve Parker, Manchester, England

Published by John Wiley & Sons, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-0-470-02448-5 ISBN: 978-1-118-16633-8 (ebk) ISBN: 978-1-118-16632-1 (ebk) ISBN: 978-1-118-16631-4 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or Web site is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or Web site may provide or recommendations it may make. Further, readers should be aware that Internet Web sites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Control Number: 2011932268

Trademarks: Wiley, the Wiley logo, Wrox, the Wrox logo, Programmer to Programmer, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

For my daughters, Bethany and Emily, and my wife, Jackie. Putting up with a professional geek is never easy, particularly when it leads to a career which often means a lot of travel and time spent away from home. Also to God, from whom comes all wisdom, intelligence, and learning. The better we understand the Creation, the better chance we have of understanding the Creator.

For it is written:

"I will destroy the wisdom of the wise; the intelligence of the intelligent I will frustrate. Where is the wise man? Where is the scholar? Where is the philosopher of this age? Has not God made foolish the wisdom of the world?...For the foolishness of God is wiser than man's wisdom, and the weakness of God is stronger than man's strength."

1 Corinthians chapter 1, verses 19, 20, and 25

ABOUT THE AUTHOR

STEVE PARKER is a Unix and Linux consultant with 20 years' experience with Unix, and 15 years' experience with GNU/Linux. He wrote and maintains the online shell scripting tutorial at http://steve-parker.org/sh/sh.shtml.

Steve provides IT consultancy services, and can also provide training courses in shell scripting as well as Unix and Linux technologies. He can be contacted via http://sgpit.com/.

ABOUT THE TECHNICAL EDITOR

JOHN KENNEDY has worked with Linux (and Unix) as a system administrator since 1997. He has worked with Red Hat, SUSE, Debian, Ubuntu, Solaris, and HP-UX. He started bash scripting in 2000 because he felt he was doing too much work and wanted something to do the tedious jobs for him.

Before learning the joys of Linux and Unix, John was in the U.S. Air Force for nine years working as a communications systems operator and spent time in Germany, Texas, and England. Since leaving the military he has lived in Nebraska and Pennsylvania, and is now back in England.

John currently works as an Infrastructure Engineer for a media company based in London. He lives near Oxford with his wife, Michele, and son, Kieran. He has a daughter, Denise, who just finished her university degree in the U.S.

When John is not on his computer, he enjoys watching football (soccer) with his son, spending time with his family, and relaxing.

CREDITS

EXECUTIVE EDITOR

Mary James

PROJECT EDITOR

Christina Haviland

TECHNICAL EDITOR

John Kennedy

PRODUCTION EDITOR

Rebecca Anderson

COPY EDITOR

Nancy Rapoport

EDITORIAL MANAGER

Mary Beth Wakefield

FREELANCER EDITORIAL MANAGER

Rosemarie Graham

ASSOCIATE DIRECTOR OF MARKETING

David Mayhew

BUSINESS MANAGER

Amy Knies

PRODUCTION MANAGER

Tim Tate

VICE PRESIDENT AND EXECUTIVE GROUP

PUBLISHER

Richard Swadley

VICE PRESIDENT AND EXECUTIVE PUBLISHER

Neil Edde

ASSOCIATE PUBLISHER

Jim Minatel

PROJECT COORDINATOR, COVER

Katie Crocker

COMPOSITOR

Jeff Lytle, Happenstance Type-O-Rama

PROOFREADERS

Louise Watson, Word One New York Paul Sagan, Word One New York

INDEXER

Robert Swanson

COVER DESIGNER

Ryan Sneed

COVER IMAGE

© mika makkonen / istockphoto.com

ACKNOWLEDGMENTS

THIS BOOK WOULD NOT HAVE happened without the help (and deadlines) that the people at Wiley gave me. Every step of the process has been a new experience, and Christina Haviland has been a great mentor through each step. John Kennedy has provided feedback and encouragement throughout, and Nancy Rapoport has shown a fantastic eye for detail.

From a personal perspective, I would like to thank all of the people behind Acorn, Sinclair, and other companies in the early 1980s for making affordable computers for kids to learn real programming with. Also the BBC for their foresight in the entire BBC Micro project, the TV programs that they put behind it, and the development work that they pursued. The next generation needs something like the BBC Micro project; not using fancy IDEs to write apps for phones, but working at the bare metal with real systems. The Arduino project deserves credit for promoting this at the hardware level; it is an excellent project, making it easy to hack hardware without having to have a knowledgeable uncle on hand to translate resistor values. The Free Software infrastructure, particularly with more recent injections like the Google Summer of Code, is another ideal breeding ground for this love of hacking to develop afresh for a new (GNU?) generation. The idea of a generation growing up knowing only how to *use* devices, not how to *develop* them, is a disturbing one. The projects mentioned above provide hope for the future.

I also want to thank ICL, where I met Douglas and Capitan, Jit, and Ketan. We tested DRS/NX, and had direct access to userspace and kernel developers. That was a rare treat, and it was where I fell in love with Unix. Also the guys who used to hang out on <code>comp.unix.shell</code> back in the days when Usenet was still readable; you taught us so much, and we must have seemed so naïve (which we were).

What I gained at ICL by being employed by the same company as the kernel and userspace developers became available to everyone with the GNU/Linux operating system. In the course of writing this book, I have been able to quote e-mails written by people that I have never met (and probably will never meet) in the discussion of Unix, Linux, and shell features. Similarly, in a professional context, I have had the honor of chatting online with the key developers of specific Linux kernel features to discuss how they are implemented in different versions of the Linux kernel, none of which would be possible with a different development model. Similarly, Chet Ramey, the maintainer of the bash shell, has responded to emails about implementation details.

From a professional and IT community perspective, I would like to thank Ken Thompson, Dennis Ritchie, Brian Kernighan, Doug McIlroy, David Korn, and Steve Bourne (to name but a few) for C, Unix, and the environment that is so easily taken for granted. The concepts these visionaries came up with have lasted for 40 years and more.

I also thank Dr. Richard M. Stallman for giving the GNU project to the world, for the GPL and the Free Software Foundation, and for dedicating a lifetime to promoting software freedom. The world needs idealists, and Stallman is one of these. It is my belief that Stallman will be proved by history to be right, that it is better to share developments than to hide them. That is the scientific tradition, and it must also be applied to computer science if it is to be treated seriously as a scientific endeavor.

My thanks to all of the GNU programmers and Linux kernel developers for putting all of these Free Software pieces together into a variety of different usable and Free OSs. Also the other Unix developers, who write code under a variety of licenses for various employers.

Finally, my thanks to Bill Joy, Vinod Khosla, Andy Bechtolsheim, and Scott McNealy for Sun Microsystems and the Solaris Operating Environment. Also Jonathan Schwartz for making most of the company's software Open Source (even buying StarDivision in order to release OpenOffice.org) and the contributions that JDS made to the GNOME project, at a time when a lot of the industry didn't understand the model. RIP Sun Microsystems.

CONTENTS

INTRODUCTION	xxix	
PART I: ABOUT THE INGREDIENTS		
CHAPTER 1: THE HISTORY OF UNIX, GNU, AND LINUX	3	
Unix	3	
"Everything Is a File" and Pipelines	5	
BSD	6	
GNU	7	
Linux	11	
Summary	12	
CHAPTER 2: GETTING STARTED	15	
Choosing an OS	15	
GNU/Linux	15	
The BSDs	17	
Proprietary Unix	17	
Microsoft Windows	17	
Choosing an Editor	18	
Graphical Text Editors	18	
Terminal Emulation	21	
Nongraphical Text Editors	22	
Setting Up the Environment	24	
The Shell Profile	24	
Aliases	26	
vim Settings	30	
Summary	31	
CHAPTER 3: VARIABLES	33	
Using Variables	33	
Typing	34	
Assigning Values to Variables	35	
Positional Parameters	39	
Return Codes	42	
Unsetting Variables	45	

Preset and Standard Variables	47
BASH_ENV	47
BASHOPTS	47
SHELLOPTS	48
BASH_COMMAND	50
BASH_SOURCE, FUNCNAME, LINENO, and BASH_LINENO	51
SHELL	55
HOSTNAME and HOSTTYPE	55
Working Directory	55
PIPESTATUS	55
TIMEFORMAT	56
PPID	57
RANDOM	58
REPLY	58
SECONDS	58
BASH_XTRACEFD	59
GLOBIGNORE	60
HOME	62
IFS	62
PATH	63
TMOUT	64
TMPDIR	65
User Identification Variables	65
Summary	66
CHAPTER 4: WILDCARD EXPANSION	67
Filename Expansion (Globbing)	67
Bash Globbing Features	70
Shell Options	71
Regular Expressions and Quoting	75
Overview of Regular Expressions	76
Quoting	77
Summary	81
CHAPTER 5: CONDITIONAL EXECUTION	83
lf/Then	83
Else	85
elif	85
Test ([)	87
Flags for Test	88
File Comparison Tests	95

String Comparison Tests Regular Expression Tests Numerical Tests Combining Tests Case	96 98 101 103 105
Summary	109
CHAPTER 6: FLOW CONTROL USING LOOPS	111
For Loops	111
When to Use for Loops	112
Imaginative Ways of Feeding "for" with Data	112
C-Style for Loops	118
while Loops	119
When to Use while Loops	119
Ways to Use while Loops	119
Nested Loops	125
Breaking and Continuing Loop Execution	126
while with Case	130
until Loops	131
select Loops	133
Summary	137
CHAPTER 7: VARIABLES CONTINUED	139
Using Variables	139
Variable Types	141
Length of Variables	142
Special String Operators	144
Stripping Variable Strings by Length	144
Stripping from the End of the String	146
Stripping Strings with Patterns	147
Searching Strings	151
Using Search and Replace	151
Replacing Patterns	153
Deleting Patterns	153
Changing Case	153
Providing Default Values	153
Indirection	157
Sourcing Variables	158
Summary	159

CHAPTER 8: FUNCTIONS AND LIBRARIES	161
Functions	161
Defining Functions	162
Function Output	162
Writing to a File	164
Redirecting the Output of an Entire Function	167
Functions with Trap	171
Recursive Functions	173
Variable Scope	177
Libraries	181
Creating and Accessing Libraries	183
Library Structures	183
Network Configuration Library	187
Use of Libraries	191
getopts	191
Handling Errors	194
getopts within Functions	195
Summary	197
CHAPTER 9: ARRAYS	199
Assigning Arrays	199
One at a Time	200
All at Once	200
By Index	201
All at Once from a Source	201
Read from Input	203
Accessing Arrays	205
Accessing by Index	205
Length of Arrays	206
Accessing by Variable Index	206
Selecting Items from an Array	209
Displaying the Entire Array	209
Associative Arrays	210
Manipulating Arrays	211
Copying an Array	211
Appending to an Array	213
Deleting from an Array	214
Advanced Techniques	216
Summary	217

CHAPTER 10: PROCESSES	219
The ps Command	219
ps Line Length	220
Parsing the Process Table Accurately	220
killall	223
The /proc pseudo-filesystem	225
prtstat	226
I/O Redirection	227
Appending Output to an Existing File	229
Permissions on Redirections	229
exec	229
Using exec to Replace the Existing Program	230
Using exec to Change Redirection	231
Pipelines	237
Background Processing	237
wait	238
Catching Hangups with nohup	239
Other Features of /proc and /sys	242
Version	242
SysRq	242
/proc/meminfo	245
/proc/cpuinfo	245
/sys	246
/sys/devices/system/node	251
sysctl	253
Summary	254
CHAPTER 11: CHOOSING AND USING SHELLS	255
The Bourne Shell	256
The KornShell	256
The C Shell	256
The Tenex C Shell	257
The Z Shell	257
The Bourne Again Shell	257
The Debian Almquist Shell	258
Dotfiles	258
Interactive Login Shells	259
Interactive Non-Login Shells	260
Non-Interactive Shells	261
Logout Scripts	262

Command Prompts	262
The PS1 Prompt	262
The PS2, PS3, and PS4 Prompts	264
Aliases	265
Timesavers	265
Modifying Behaviors	265
History	266
Recalling Commands	267
Searching History	267
Timestamps	268
Tab Completion	269
ksh	269
tcsh	270
zsh	270
bash	271
Foreground, Background, and Job Control	272
Backgrounding Processes	272
Job Control	273
nohup and disown	275
Summary	276
PART II: RECIPES FOR USING AND EXTENDING SYSTEM TO	noi s
CHAPTER 12: FILE MANIPULATION	279
stat	279
cat	281
Numbering Lines	282
Dealing with Blank Lines	282
Non-Printing Characters	283
cat Backwards is tac	284
Redirection	285
Redirecting Output: The Single Greater-Than Arrow (>)	285
Appending: The Double Greater-Than Arrow (>>)	286
Input Redirection: The Single Less-Than Arrow (<)	288
Here Documents: The Double Less-Than Arrow (<< EOF)	290
dd	292
df	294
mktemp	295
join	297
install	298

grep	300
grep Flags	300
grep Regular Expressions	301
split	303
tee	304
touch	306
find	307
find-exec	310
Summary	313
CHAPTER 13: TEXT MANIPULATION	315
cut	315
echo	316
dial1	316
dial2	319
fmt	320
head and tail	323
Prizes	323
World Cup	324
od	328
paste	331
pr	334
printf	335
shuf	337
Dice Thrower	337
Card Dealer	338
Travel Planner	340
sort	341
Sorting on Keys	342
Sorting Log Files by Date and Time	344
Sorting Human-Readable Numbers	345
tr	346
uniq	350
WC	351
Summary	352
CHAPTER 14: TOOLS FOR SYSTEMS ADMINISTRATION	353
basename	353
date	355
Typical Uses of date	355
More Interesting Uses of date	359

dirname	360
factor	362
identity, groups, and getent	364
logger	367
md5sum	368
mkfifo	370
Master and Minions	371
Reversing the Order	373
Networking	375
telnet	376
netcat	376
ping	378
Scripting ssh and scp	381
OpenSSL	383
nohup	390
seq	391
Integer Sequences	391
Floating Point Sequences	393
sleep	394
timeout	394
Shutdown Script	396
Network Timeout	399
uname	400
uuencode	401
xargs	402
yes	405
Summary	406
PART III: RECIPES FOR SYSTEMS ADMINISTRATION	N
CHAPTER 15: SHELL FEATURES	409
Recipe 15-1: Installing Init Scripts	409
Technologies Used	410
Concepts	410
Potential Pitfalls	410
Structure	410
Recipe	412
Invocation	414
Summary	414
Recipe 15-2: RPM Report	414
Technologies Used	415
Concepts	415

Potential Pitfalls	415
Structure	415
Recipe	417
Invocation	419
Summary	420
Recipe 15-3: Postinstall Scripts	421
Technologies Used	421
Concepts	421
Potential Pitfalls	422
Structure	422
Recipe	423
Invocation	425
Summary	426
CHAPTER 16: SYSTEMS ADMINISTRATION	427
Recipe 16-1: init Scripts	427
Technologies Used	428
Concepts	428
Potential Pitfalls	429
Structure	430
Recipe	431
Invocation	432
Summary	433
Recipe 16-2: CGI Scripts	433
Technologies Used	433
Concepts	434
Potential Pitfalls	434
Structure	435
Recipe	438
Invocation	441
Summary	445
Recipe 16-3: Configuration Files	445
Technologies Used	445
Concepts	445
Potential Pitfalls	446
Structure	446
Recipe	446
Invocation	447
Summary	448
Recipe 16-4: Locks	448
Technologies Used	448
Concepts	448

Potential Pitfalls Structure Recipe Invocation Summary	449 450 453 455 458
CHAPTER 17: PRESENTATION	459
Recipe 17-1: Space Game	459
Technologies Used	459
Concepts	460
Potential Pitfalls	462
Structure	462
Recipe	464
Invocation	469
Summary	470
CHAPTER 18: DATA STORAGE AND RETRIEVAL	471
Recipe 18-1: Parsing HTML	471
Technologies Used	471
Concepts	472
Potential Pitfalls	472
Structure	472
Recipe	473
Invocation	474
Summary	476
Recipe 18-2: CSV Formatting	476
Technologies Used	476
Concepts	476
Potential Pitfalls	477
Structure	477
Recipe	478
Invocation	480
Summary	481
CHAPTER 19: NUMBERS	483
Recipe 19-1: The Fibonacci Sequence	483
Technologies Used	483
Concepts	484

Potential Pitfalls	484
Structure for Method 1	485
Recipe for Method 1	486
Invocation of Method 1	486
Structure for Method 2	487
Recipes for Method 2	488
Invocations of Method 2	489
Structure for Method 3	490
Recipe for Method 3	490
Invocation of Method 3	491
Summary	492
Recipe 19-2: PXE Booting	492
Technologies Used	492
Concepts	493
Potential Pitfalls	493
Structure	493
Recipe	494
Invocation	497
Summary	499
CHAPTER 20: PROCESSES	501
Recipe 20-1: Process Control	501
Technologies Used	501
Concepts	502
Potential Pitfalls	503
Structure	503
Recipe	506
Invocation	511
Summary	516
CHAPTER 21: INTERNATIONALIZATION	517
Recipe 21-1: Internationalization	517
Technologies Used	518
Concepts	518
Potential Pitfalls	519
Structure	520
Recipe	521
Invocation	525
Summary	526

APPENDIX: FURTHER READING	529
Shell Tutorials and Documentation	529
Arrays	530
Tools	530
Unix Flavors	531
Shell Services	531
GLOSSARY	533
INDEX	539