

book reviews

ÆLEEN FRISCH, BRAD KNOWLES, AND
SAM STOVER

BASH COOKBOOK: SOLUTIONS AND EXAMPLES FOR BASH USERS

Carl Albing, JP Vossen, and Cameron Newham

O'Reilly, 2007. 622 pp.

ISBN 978-0-596-52678-8

REVIEWED BY ÆLEEN FRISCH

Bash Cookbook is another strong entry in the well-known Cookbook series published by O'Reilly Media. The authors have created a large collection of examples designed to address common tasks and problems as well as to educate readers about running commands and writing shell scripts under bash.

The book consists of a great many relatively short problem/task-plus-solution discussions (with related items loosely gathered into chapters). One of the strengths of the book is that these examples are placed in realistic computing contexts, so, e.g., determining the amount of time between two dates is considered with respect to NTP rather than as a coding exercise in isolation. The most extensive examples focus on text processing, simple parsing, and automating operations on files and directories. Sound best practices advice is integrated into virtually every discussion.

My initial assumption was that this book was about bash shell scripting, based on *bash* in the title and familiarity with other works in the series. However, although items involving bash scripts do comprise somewhat more than half of the book, it also contains a great deal of information that is neither bash-specific nor scripting-related. Much of the book focuses on basic UNIX commands in reasonable detail (most notably `grep`, `find`, `sort`, and `date`), along with related topics such as I/O redirection

and pipes, wildcards and quoting, and search paths. The various items concerned with scripting ultimately cover a comprehensive range of relevant information, including basic script structure, file I/O, user prompting, arguments and variables, control structures and functions, script invocation methods, and security.

Several chapters in this work deserve special mention. The first chapter is an excellent tutorial for absolute beginners with bash and also includes useful information such as how to obtain bash for Windows and a list of Web sites offering free shell access.

Chapter 12 is the jewel at the center of this work. It discusses a series of very well-crafted scripts solving problems of great interest to many people: copying MP3 files to a player, creating a Web photo album for a picture directory (a great first example of generating HTML), and the like. I wish that this chapter had contained many more examples.

The book also provides a comprehensive yet compact reference appendix to all things bash: invocation options, prompt strings (including ANSI color escapes), built-in commands, shell variables, test command operators, arithmetic, and so on. I've made myself a copy that I keep closer than my bookshelf.

In general, titles in the O'Reilly Cookbook series seem to reject systematic organization and take a more exploratory, meandering pathway through their subjects. This book is no exception. I find both the ordering of the chapters and the sequencing of items within chapters very arbitrary. In addition, subjects that seem to this reader to be closely related can be separated by hundreds of pages. This is not surprising, given that the work is designed to be read in a random-access manner, like an encyclopedia. However, this design has the consequence that readers who want to explore certain topics in detail will find themselves jumping around in the text quite a bit. Fortunately, the index is excellent, so finding information is not a problem.

All in all, this is a very useful and well-written book about running UNIX commands and shell scripting in the bash shell. It should appeal to three types of readers. People who have already begun writing bash shell scripts, perhaps after reading a more discursive introductory book, will find a wealth of real-world example fragments and scripts discussed in detail. People who are looking for solutions to specific problems or techniques to specific tasks will find many helpful items within the book. Finally, people who prefer to just jump in

and start trying things can use this book to learn about bash, again as both a command environment and a scripting language. Such people learn better from contextual exploration than from more abstract and linear discussions, and this book is perfect for them. All of the elements of bash scripting are included in the book, although the path connecting them is far from a straight line.

HIGH PERFORMANCE WEB SITES: ESSENTIAL KNOWLEDGE FOR FRONT-END ENGINEERS

Steve Souders

O'Reilly, 2007. 168 pp.

ISBN 10: 0-596-52930-9; ISBN 13: 978-0-596-52930-7

REVIEWED BY BRAD KNOWLES

The author of this book is the Chief Performance Yahoo! at Yahoo!. The subtitle holds a hint of the premise that instead of tuning the back-end systems for maximum performance from the perspective of the people developing on or administrating those machines, we should instead be focusing on tuning the overall system for maximum performance from the perspective of the people using the system. The author shows that Web site performance from the user perspective depends much more on the front-end architecture and on how the overall Web pages are designed and much less on the back-end throughput and how fast they can crank out the HTML. Therefore, by following the 80/20 rule, we should be working on tuning the front end and not the back end.

The author clearly explains his methodology, listing all the tools he uses in his testing, including the tools that he developed to help implement his methodology—with links for everything. In each of the chapters, the author also provides links to sample code that he has on the Web that demonstrates the technique described. Thus, the reader can see firsthand what he is talking about, and how the page is sped up by making the change being highlighted. In all, there are forty-three examples provided, demonstrating how the overall rule being discussed in a given chapter affects various different aspects of the Web page.

The book is organized according to fourteen key steps that can be taken to optimize the front-end performance of a Web site, in order of importance. However, on first glance, the particular order of the rules might not seem to make the most sense. For example, the reader might think that they should be looking at using a Content Delivery Network as one of the last things to optimize their Web site (after all local optimization has been applied). Nevertheless, by the end of Chapter Two, the reader will

be convinced as to why this is the second rule. The reader might still choose to consider CDNs after local optimization, but will at least understand why CDNs are important.

Some of the chapters are very short (just two or three pages); others are longer (ten or fifteen pages). The author doesn't seem to feel the need to make any chapter longer than necessary, which results in a pretty thin book. However, although the book is packed with information, the presentation is light and easy to read. There's a full fourteen-course meal here, but each plate is as small as it can reasonably be, and each serving is already cut up into nice little bite-size chunks. This reviewer read the whole book on a flight lasting less than two hours.

The author also deconstructs the top ten Web sites on the Internet (by volume), including both his own site and major competitors, as well as some others the reader might not have otherwise expected. He is constructive when applying criticism, but he is also refreshingly honest when the competitors do well according to his methodology. Most surprisingly, he publicly applies the same type of criticism to his own site, when it does not perform as well as it could.

Some of the rule names (also used as the chapter titles) would seem to be obvious, but on further explanation the reader comes to understand the full scope of the issue at hand and how this affects the overall user experience and apparent speed of the Web site. Some of the later rules actually relate to and reexpress earlier rules, despite their inclusion in the list. Regardless of some of the apparent obvious names, most of the useful information is actually found within the chapters themselves, so there is little harm in listing the rules:

1. Make fewer HTTP requests.
2. Use a Content Delivery Network.
3. Add a (far-in-the-future) Expires Header.
4. Gzip components.
5. Put stylesheets at the top.
6. Put scripts at the bottom.
7. Avoid CSS expressions.
8. Make Javascript and CSS external (or internal).
9. Reduce DNS lookups.
10. Minimize (or obfuscate) Javascript.
11. Avoid redirects.
12. Remove duplicate scripts.
13. (Eliminate or) configure Etags.
14. Make Ajax cacheable.

The comments in parentheses are recommended alterations by this reviewer. Once the reader has

completed the chapter in question, the reason for the alterations should become clear.

The author provides tips and tricks that make it obvious how some of these things can be done. For example, sites that have lots of dynamic content may think they can't implement rule #3. However, the author shows that by including the version of the object in question within the URL to the object, the front-end engineer can still add a far future "Expires:" header and make sure that the object is cached for as long as possible, while maintaining the dynamism of the site—all that is required is to switch to a different URL for a different version of that object when it gets updated.

In the case of CDNs he not only tells the reader which ones are the most commonly used and which ones are "low-cost" alternatives, he also outlines free solutions that are available. In addition, he mentions some external CDN testing services that can be used to make sure that the reader sees the global perspective on their site, and not just the very distorted picture of how it looks from the high-speed local connectivity the internal workers have from their workstations sitting right next to the servers.

The author also provides browser-specific guidance as to why the reader might want to do certain things in certain ways as opposed to other alternatives.

The one problem with this book is that it is written from the perspective of a group that has complete and total control over every aspect of their megasite, write their own tools, etc. However, most sites on the Internet today are likely to be implemented with Content Management Systems (e.g., Drupal, Joomla!, Mambo), wikis (e.g., WikiMedia, TWiki, MoinMoin), or blogging software (e.g., WordPress, Moveable Type, Bloxsum, LiveJournal), or are hosted at commercial blogging sites (e.g., TypePad, Blogger, LiveJournal, MySpace). Much of the front-end engineering for sites implemented with tools such as these will be encoded into the toolkit itself, and therefore it will be difficult to actually apply these rules.

This is not the fault of the author, but it would be very useful if a companion book were to be produced that took the Yahoo! methodology outlined and showed the reader how to implement as much of that as possible within a variety of popular tools.

Since I'm not really a Web developer or administrator myself, I don't expect to get much more out of it, so my review copy (already well-thumbed) will be handed over to the Webmaster at one of the

open-source projects I help support, and I will be buying several more copies for other Webmasters and Web developers. I've definitely had my perspective on this field, and on performance tuning in general, permanently altered. I only wish someone would buy a few thousand copies of this book and freely distribute them to the key people in the various communities for Web developers, because I believe that everyone on the Internet would benefit from a universal application of these concepts.

ALTERNATE DATA STORAGE FORENSICS

Amber Schroader and Tyler Cohen

Syngress, 2007. 400 pp.

ISBN 978-1-59749-163-1

REVIEWED BY SAM F. STOVER

If you are looking for a cutting-edge book on the forensic procedure for Alternate Data Storage (ADS) devices, this is not the book for you. If you are looking for an introductory look at how ADS devices can be examined, this might be the book for you. Considering the caliber of some of the authors, I have to admit that I was a bit disappointed. I don't do forensics every day, but when I do, I prefer to have a more authoritative reference than this book provides. To be fair, much of my disappointment stems from the chapter on PDA, Blackberry, and iPod Forensic Analysis. I have some degree of experience in this area and was hoping to expand my horizons, and I was really let down. If you are interested in Blackberry hacking in particular, avoid this book altogether: 30 minutes on Google will give you far more than this chapter.

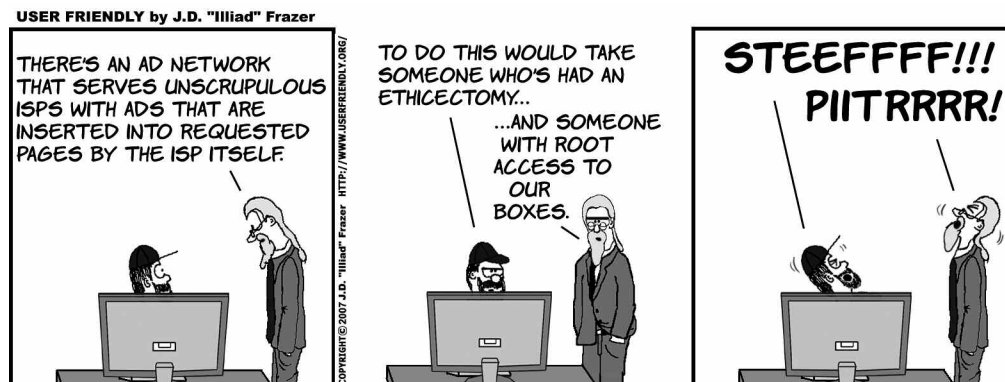
The other chapters are a bit more solid, but there is a lot of introductory text that can be found in plenty of other forensics books. I got the impression that there wasn't enough PDA/handheld-specific material to justify a \$60 book, so the obligatory background filler was used to inflate the book to a final size of approximately 300 pages (not counting the index).

Now that I have all that negativity out of my system, I'd like to focus on the chapters that I did find informative. The first is Chapter 5, which addresses email forensics. As email clients become more and more advanced, extracting the actual data becomes more difficult. It's one thing to grep through someone's mutt or pine mailbox, but another thing entirely to analyze an Outlook PST file. One thing that did confuse me a bit is that the chapter starts out by outlining the exchange between a client and server, but the analysis deals solely with client systems. I was hoping for some tips on analyzing an email server, but unfortunately that was not the

case. The next chapter I liked was Chapter 6, on router forensics. Again, I thought there was a little too much introductory material, but I was happy to see network infrastructure addressed in the book. The final two chapters deal with CD/DVD and MP3 forensics, respectively. The CD/DVD chapter is particularly comprehensive, and it should serve as a great reference. The MP3 chapter is also fairly complete in that it focuses on MP3 players not only as media devices but also as potential platforms for alternative uses, such as running Linux. Good stuff.

In summary, I would say that this book is a little too lean for the price tag, as only four of the chap-

ters really piqued my interest. If you have no real experience in forensics and have a pressing need to analyze an iPod, this is probably a reasonable book to pick up. If you have some forensic background and are interested in nonstandard forensics, I'd recommend perusing the book at your local bookstore before actually ordering a copy. If you are an advanced forensic examiner, I'd wait for the next revision. I think this book has plenty of potential, but it just didn't live up to that potential this time around.



bash cookbook carl albing, jp vossen, and cameron newham copyright 2007 media, inc.. all rights reserved. printed in the united states of america. published. This book will only briefly cover basic and intermediate shell scripting—see Learning the bash Shell by Cameron Newham (O'Reilly) and Classic Shell Scripting by Nelson H.F. Beebe and Arnold Robbins (O'Reilly) for more in-depth coverage. Instead, our goal is to provide solutions to common problems, with a strong focus on the "how to" rather than the theory. The key to mastering any Unix system, especially Linux and Mac OS X, is a thorough knowledge of shell scripting. Scripting is a way to harness and customize the power of any Unix system, and it's an essential skill for any Unix users, including system administrators and professional OS X developers. But beneath this simple promise lies a treacherous ocean of variations in Unix commands and standards. bash Cookbook teaches shell scripting the way Unix masters practice the craft. It presents a variety of recipes and tricks for all levels of shell programmers so that anyone can become a profi bash Cookbook, by Carl Albing, JP Vossen, and Cameron Newham Copyright © 2007 O'Reilly Media, Inc. All rights reserved. Printed in the United States of America. Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472. O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (safari.oreilly.com). For more information, contact our corporate/institutional sales department: (800) 998-9938 or corporate@oreilly.com. Editor: Mike Loukides Production Editor: Laurel R.T. Ruma Copyeditor: Derek