

;login:

THE MAGAZINE OF USENIX & SAGE

July 2001 • Volume 26 • Number 4

inside:

BOOK REVIEWS:

UNIX ADMINISTRATION: A BEGINNER'S GUIDE

by Steve Shaw

Reviewed by Paul Guglielmino

DATA MUNGING WITH PERL

by David Cross

Reviewed by William Annis

DNS AND BIND, 4TH EDITION

by Paul Albitz and Cricket Liu

Reviewed by Rik Farrow

USENIX & SAGE

The Advanced Computing Systems Association &
The System Administrators Guild

book reviews

LINUX ADMINISTRATION: A BEGINNER'S GUIDE

STEVE SHAH

New York: McGraw-Hill, 2001. Pp. 643 +
CD-ROM. ISBN 0072131365.

REVIEWED BY PAUL GUGLIELMINO
<paulg@ccs.neu.edu>

[Paul Guglielmino is a UNIX administrator in Boston.]

Steve Shah's *Linux Administration: A Beginner's Guide* is a well written and comprehensive book on Linux administration. The book is divided into seven sections: installing Linux as a server, single-host administration, Internet services, intranet services, advanced Linux networking, and two appendixes. Some specific areas dealt with are software installation, the bootup and shutdown process, Samba, DHCP, and backups. Internet services covered are SMTP with Sendmail, DNS, FTP, and Web services with Apache. This book is written for beginners, but there is a section about some of the more advanced features of the Linux kernel. This includes, among other things, IP masquerading, IP chains, packet filtering and the /proc file system. Unfortunately, in trying to cover as much ground as he has, Shah can't delve into the details of any one subject, but I think he does a good job of giving a clear overview of each topic without overwhelming new users with details.

Even though the book does not have "RedHat" in the title, it is obviously centered on the RedHat distribution. There are several references to RedHat in the text; more specifically, the installation chapter only describes the installation process for RedHat. In fact, other distributions are only mentioned in a few sentences in the opening chapter. The book comes with a CD, which has a watered down RedHat 7 distribution.

Although very good overall, the book could stand some improvement in a few places. My biggest complaint is that the security chapter could have been devel-

oped more. There was no mention of TCP wrappers or the importance of good passwords. (On the plus side, there was a whole chapter dedicated to SSH.) The installation chapter could have said more about dual booting Linux and Windows, since this is an issue that could potentially scare away new users of Linux. I was also a little disappointed in the chapter devoted to the Linux kernel. Since this book is about Linux administration, I would expect this chapter to be one of the most thorough. Although Shah does explain many of the different options that are available in the kernel and then goes on to show you how to compile your own, he could have described the history of the kernel and the kernel module system in more detail.

In addition, the primary discussion of LILO occurs in an earlier chapter, but it would have been more useful to cover LILO in the context of the Linux kernel. Since the book is RedHat-centric, showing how to install the kernel via RPMs would have been helpful. Another small complaint is that there is a small chapter on POP but no mention of IMAP anywhere.

The introduction states that the reader should be a "strong user in Windows." The best part of the book may come in the first chapter in a section titled "The major differences between Windows 2000 and Linux." Here there are discussions of the separation of the GUI from the kernel, single and multi-user philosophies, the Windows registry versus text-based configuration files, and Active Directory vs. NIS. Other than that great section and a couple pages of blueprints about the boot and shutdown processes, there is not much reason to be well versed in Windows to read this book. I liked that the software installation chapter explained both how to use RPM and how to compile packages from the source. RPM is a great tool, but it's important for a system administrator to understand what her tools do and how

they do it. The SMTP chapter was a very good introduction to Sendmail, which can be a very complex program to administer. The chapter was Sendmail-centric, but that doesn't matter much because it is by far the most common MTA on the Internet. In each of the "Internet" chapters, Shah has included a mechanics section that could be considered a very condensed version of the relevant RFC. It is important for troubleshooting to know not only what software does but also how it does it. Each chapter has a nice set of summary bullets, and most chapters include a small list of useful links or book references to learn more.

I would definitely recommend this book to those new to the world of Linux and UNIX. However, if you have some experience in UNIX, but are new to Linux, I would advise checking into other Linux books on the market or just looking to the Linux Documentation Project on the Web for help.

DATA MUNGING WITH PERL

DAVID CROSS

Greenwich, CT: Manning Publications Co.,
2001. Pp. 283. ISBN 1-930110-00-6.

REVIEWED BY WILLIAM ANNIS
<annis@biostat.wisc.edu>

With entire rows of bookstores full of books on learning CGI programming with Perl in 10 easy lessons, it's nice to see one highlighting Perl's data processing capabilities more generally.

The book is divided into three parts with several chapters apiece. The first part is an overview. After defining "munging" and introducing his CD collection file – which will show up in examples throughout the rest of the book – Cross discusses Perl itself then goes on in Chapter 2 to discuss important issues like separating parsing from munging, choosing the right data structure for your problem, and approaching audit trails and data validation.

Chapter 3, “Useful Perl Idioms,” gives an excellent overview of sorting in Perl, including a description of the Orcish Maneuver and a very nice explanation of not only how the Schwartzian Transform works but also why you’d use it. Tools for debugging and benchmarking are introduced in this chapter, as is the DBI interface. Chapter 4 goes over important string manipulation functions and regular expressions.

Part 2, “Data Munging,” is the heart of the book. Starting with simple transformations of ASCII data to other formats, Chapter 5 ends with a discussion of several CPAN libraries for manipulating and formatting numeric data. Chapter 6, “Record-oriented Data,” starts with a three-page digression on the `while (<FILE>) {}` idiom, which seems a bit strange coming so late in the book. It goes on to various types of line-oriented records and introduces useful idioms using the Perl special variables which apply to records (`$`, `$/`, `$`, etc.). Data caching, CSV, and multi-line records are touched on, and Chapter 6 ends with a good overview of date and time manipulation, including parsing dates, using `POSIX::strftime`, `Date::Calc`, and `Date::Manip`. Chapter 7, with a discussion of various sorts of fixed-width and binary data formats, ends Part 2.

Chapter 8 starts off Part 3, “Simple Data Parsing,” by introducing more complex data files and metadata. There is a brief section showing how regular expressions are not sufficient to parse HTML, followed by an introduction to parsing terminology. Chapter 9, “HTML,” builds a few tools for summarizing and manipulating HTML using the CPAN modules `HTML::Parser`, `HTML::LinkExtor`, and `HTML::TokeParser`. The chapter ends with an example for getting a weather report from Yahoo!. Chapter 10, “XML,” starts off with a quick introduction to XML and makes a clear distinction between valid and well-formed XML. The chapter is only concerned really with

well-formed XML, so there is no discussion of DTDs. After discussing several modules for XML parsing – `XML::Parser`, `XML::DOM`, and `XML::RSS` – Cross ends with a tool for turning XML markup of documentation into POD, HTML, and plain text. Chapter 11, “Building Your Own Parsers,” gives a nice description of `Parse::RecDescent`. After introducing parser-writing using Windows INI files as an example, the chapter ends with a parser for Cross’s CD collection file.

The emphasis on using existing tools from CPAN is a strong feature of this book. Appendix A gives brief documentation on the modules used in the main text. Finally, Appendix B is a very dense and brief overview of Perl itself. This, with the Perl boosterism early in the book, is a bit odd. No one unfamiliar with Perl is going to be able to use this book as a recipe book, but for experienced Perl programmers it is an excellent overview of Perl’s many data manipulation capabilities.

DNS AND BIND, 4TH EDITION

PAUL ALBITZ AND CRICKET LIU

O’Reilly & Associates, Sebastopol, CA, 2001.
Pp. 622. ISBN 0-596-00158-4

REVIEWED BY RIK FARROW

rik@spirit.com

I have the second edition of this excellent book, but remembered a couple of small problems I had when I first examined *DNS and BIND* back in 1997. For example, I only dimly understood the meaning of “canonicalization,” and I really wanted Albitz and Liu to explain the term in their book so I wouldn’t have to ask someone and appear clueless. The second edition just assumes that you know what the canonical name means, but the fourth edition actually defines it on page 8. Thank you, guys.

Of course, that’s not the biggest change in a book that has grown by 200 pages in two editions. A lot of the book has been revised “for the first time,” and that has

made the new edition easier to understand. The explanations are a bit longer and are certainly clearer than before.

The other big difference is that the fourth edition includes both version 8 and version 9 features, something my old edition misses entirely. So, you can learn about how to configure BIND to handle IPv6 addresses, and to use transaction signatures. If you are using version 9, you can even use `DNSSEC`, a method that will quadruple the size of your zone files while adding security that only other participating BIND 9 servers will appreciate. Mind you, I do think that DNS does need better security, and anxiously await the outcome of the ongoing experiments with `DNSSEC` in The Netherlands, Germany, and Sweden. Evi Nemeth has a new chapter in her `sysadmin` book about `DNSSEC`, as well as a version of that information in the last special edition of *login*. Even so, it helped having yet another explanation of a complicated topic.

Overall, I liked the changes and improvements to the fourth edition and can recommend it to anyone who uses BIND. I think the authors have done a great job in improving this book and deserve to be congratulated.