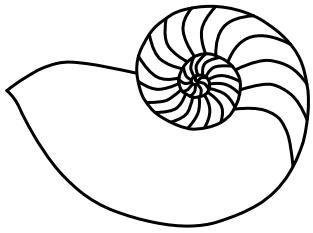
Vol. 15 No. 6 February 2003
$2.50



MUUGlines

The Manitoba UNIX User Group Newsletter

February 11, 2003: Wireless Network Setup and Security

Wireless networks are becoming a popular alternative for home users and small businesses. A small LAN can be installed more quickly (and sometimes more cheaply) with standards-compliant wireless equipment than with traditional Ethernet wiring. Unfortunately, many users value the flexibility of wireless networks without recognizing the security risks involved. Adam Thompson will discuss and demonstrate connecting a Red Hat 8.0 system to an existing ethernet infrastructure and, time permitting, configuring WEP (Wired Equivalent Privacy).

Meetings are held at the IBM offices at 400 Ellice Ave. (between Edmonton and Kennedy). When you arrive, you will have to sign in at the reception desk, and then wait for someone to take you (in groups) to the meeting room. Please try to arrive by about 7:15 PM, so the meeting can start promptly at 7:30 PM. Don't be late, or you may not get in.

Limited parking is available for free on the street, or in a lot across Ellice from IBM, for $1.00 for the evening. Indoor parking is also available nearby, at Portage Place, for $2.00 for the evening.

Review: Perl & LWP

Perl & LWP
Sean M. Burke
O'Reilly, 2002
242 pp
US$34.95, C$54.95
Reviewed by Sean Walberg

The only disappointing thing about this book is its title. 'Perl & LWP' conjures up images of simply grabbing a web page, and maybe posting to a form through Perl. That isn't much fodder for a book, as a cursory glance at the man pages will tell you how to do that. However, one must dig into the by-line of the book to learn what this book is really about, namely 'Fetching Web pages, Parsing HTML, Writing Spiders, and More'.

Unless you've done this stuff before, it's hard to appreciate how difficult it really is. The act of getting data from a web server is trivial. The book covers it quickly, writes a function, and then largely ignores the issue. The trick is what to do with the data once you've got it. That's the hard stuff, unless you've got this book by your side.

The book starts off with a quick look at what web automation is, why you'd use it, and where it's appropriate. Thirty-five pages later, you know how to GET and POST forms several different ways using the LWP (Library for WWW in Perl) class library. Then it's into a discussion of URLs, and then forms. That's five chapters, and 85 pages. It's also where the book goes from simply 'good', to 'great'.

Many people will use regular expressions to parse the data from here, and chapter six talks about ways to do that. Basically, it involves a lot of looking at the source, and making some rough guesses such as 'look for a sequence of characters, followed by some spaces, then comes our data.' The author shows iterative techniques to arrive at the solution. He then also shows shortcomings of this technique.

To overcome the problem with regular expressions, token parsing is introduced. Again, good ways to use the technique are demonstrated, and the limitations presented. To overcome these limitations, trees are introduced. Though three different methods are used to solve much the same problem, the relative advantages and disadvantages are made clear.

This book is full of examples. Every example is from a real live site, from grabbing a book's rankings from Amazon.com, to checking out license plates from the California DMV, to pulling out news items from BBC news. His code is exceptionally clear, there are very few Perl tricks used, meaning that you don't need to be a Perl expert to make use of it. Unlike many Perl books I have, the focus is on providing easy to understand code rather than efficient and terse code. Furthermore, the author isn't scared to show you the things that don't work

alongside the things that do work.

Perhaps the most impressive thing I found about Perl & LWP was that it embodied the Perl motto: There is more than one way to do it. Regular expressions are shown as a quick and dirty way to solve a problem. Tokens are introduced to overcome the limitations of regular expressions. Cases where tokens fall apart are solved with trees. Some examples are even redone to show how the improvements happen.

Particularly interesting was the chapter on writing web spiders. Not only is it a cool application, it pulls together techniques learned from most of the previous chapters. The design and build of the spider is laid out here; you follow the author's thoughts as it's put together.

If you're working with Perl to retrieve web pages without this book, you're either the author, or wasting a lot of your time. I gained more time-saving techniques reading this on the bus than I have in several years of hacking around with LWP. The time it saves is more than worth the price. For a book on programming, it's surprisingly easy to read, even bordering on fun – just what is the author's fascination with pie? I heartily recommend 'Perl & LWP' to anyone looking to automate even the smallest of web tasks using Perl.

Free High-Quality Fonts!

In a move to fill the last remaining gap in font display capabilities for open source software, the GNOME Foundation has reached an agreement with Bitstream to release fonts under an open source license.

A total of 10 serif, sans serif and monospaced fonts will be released under a special open source license, after they have been adapted to meet the requirements for technical use. The set of 10 fonts will bring advanced font capabilities to all developers and users of free and open source software.

"The donation of these fonts to the free software community is the final piece that will give full functionality to projects like Freetype, XFT2 and X Render extensions in the Xfree86 project, Pango, KDE and Trolltech's QT Rendering," said Jim Gettys, GNOME Foundation board member.

"These fonts will be available to all developers and users, giving GNOME and other open source programs a great look right out of the box that has been lacking until now."

The Bitstream Vera™ fonts will be available for free copying and redistribution and can be modified as long as the font name is changed. The fonts cannot be packaged by themselves for sale, but can be sold with any software. The GNOME Foundation will incorporate the fonts into future GNOME releases, giving end-users of all levels, as well as GNOME developers, the advanced display capabilities they offer.

"Bitstream has had a long relationship with the developer community," explained Bob Thomas, director of product management, Bitstream. "We're very happy to contribute these fonts to the open source community and we hope that developers will generate new fonts and continue to enhance the experience of end-users at all levels."

Samples of the Bitstream fonts

"GNOME was created to give free software users a friendly, powerful desktop environment,"

said Miguel de Icaza, GNOME Foundation president and CTO of Ximian Inc. "This agreement with Bitstream will enhance the experience and graphics capabilities of corporate, enterprise, educational, governmental and individual GNOME users and will give powerful tools to open source developers everywhere."

"Bitstream's generous donation is a tangible measure of the importance and growth of open source," said Russ Nelson, vice president of the Open source Initiative. "I praise the GNOME Foundation's efforts to bring this agreement into being."

The agreement between Bitstream and the GNOME Foundation will bring invaluable tools to open source developers and users worldwide.

FreeBSD 5.0 (nearly)

From the FreeBSD Release Engineering Team

FreeBSD 5.0 marks the first new major version of FreeBSD in over two years. Besides a number of new features, it also contains a number of major developments in the underlying system architecture. Along with these advances, however, comes a system that incorporates a tremendous amount of new and not-widely-tested code. Compared to the existing line of 4.*X* releases, 5.0 may have regressions in areas of stability, performance, and occasionally functionality.

For these reasons, the Release Engineering Team <re@FreeBSD.org> does not encourage users to blindly update from older FreeBSD releases to 5.0. Specifically, for more conservative users, we recommend running 4.*X* releases (such as 4.7-RELEASE or the upcoming 4.8-RELEASE) for the near-term future. We feel that such users are probably best served by upgrading to 5.*X* only after a 5-STABLE development branch has been created; this may be around the time of 5.1-RELEASE or 5.2-RELEASE.

FreeBSD employs a model of development that relies on multiple development branches within the source code repository. The main branch is called "CURRENT", and is referred to in the CVS repository with the HEAD tag. New features are committed first to this branch; although this means that CURRENT is the first to see new functionality, it also means that it occasionally suffers from breakages as new features are added and debugged.

Most FreeBSD releases are made from one of several "STABLE" branches. Features are only added to these branches after some amount of testing in CURRENT. At the moment, only one STABLE branch is under active development; this branch is referred to as "4-STABLE", and all of the FreeBSD 4.*X* releases were based on it. This branch has the tag RELENG_4 in the CVS repository.

FreeBSD 5.0 will be based on the CURRENT branch. This will be the first release from this branch in over two years (the last was FreeBSD 4.0, in March 2000).

At some point after the release of FreeBSD 5.0, a "5-STABLE" branch will be created in the FreeBSD CVS repository with the branch tag RELENG_5. The past two stable branches (3-STABLE and 4-STABLE) were created immediately after their respective "dot-oh" releases (3.0 and 4.0, respectively). In hindsight, this practice did not give sufficient time for either CURRENT or the new STABLE branches to stabilize after the new branches were created.

Therefore, the release engineering team will only create the 5-STABLE branch in the CVS repository after they have found a relatively stable state to use as its basis. It is likely that there will be multiple releases in the 5.*X* series before this happens; we estimate that the 5-STABLE branch will be created sometime after 5.1-RELEASE or 5.2-RELEASE.

New Features

A large attraction of FreeBSD 5.0 is a number of new features. These new features and functionality generally involve large architectural changes that were not feasible to port back to the FreeBSD 4-STABLE development branch. (By contrast, many self-contained enhancements, such as new device drivers or userland utilities, have already been ported.) A brief, but not exhaustive list includes:

- SMPng: The "next generation" support for SMP machines (work in progress). There is now partial support for multiple processors to be running in the kernel at the same time.
- KSE: Kernel Scheduled Entities allow a single process to have multiple kernel-level threads, similar to Scheduler Activations.
- New architectures: Support for the sparc64 and ia64 (e.g. Itanium) architectures, in addition to the i386, pc98 (a Japanese standard), and Alpha.
- GCC: The compiler toolchain is now based on

GCC 3.2.1, rather than GCC 2.95.*X*.
- MAC: Support for extensible, loadable Mandatory Access Control policies.
- GEOM: A flexible framework for transformations of disk I/O requests. An experimental disk encryption facility has been developed based on GEOM.
- FFS: The FFS filesystem now supports background
- fsck(8) operations (for faster crash recovery) and filesystem snapshots.
- UFS2: A new UFS2 on-disk format has been added, which supports extended per-file attributes and larger file sizes.
- Cardbus: Support for Cardbus devices.

A more comprehensive list of new features can be found in the release notes for FreeBSD 5.0 (http://www.freebsd.org/releases/5.0R/relnotes.html).

IBM's Help For Developers

Responding to strong developer demand, IBM updated the Speed Start Your Linux App offering for 2003. The updated offering includes a new website with free technical support and a new IBM Software Evaluation Kit (SEK) for Linux, a set of four CDs that include no-charge evaluation copies of IBM's most popular software offerings: Eclipse-based WebSphere Studio tools, WebSphere Application Server and Web Services Toolkit, DB2® Database, Lotus Domino collaboration software, and Tivoli Management Software.

The new SEK includes Tivoli software and includes support for the new UnitedLinux V1.0 server distribution, in addition to Red Hat Advanced Server. Developers can receive the new SEK via CDs or as a download from the newly updated developerWorks Speed Start site (http://www.ibm.com/developerworks/speed), IBM's resource for open, standards-based development. In 2002, more than 345,000 kits were distributed to developers creating new Linux-based applications for IBM software with over 34,000 developers registering at the Speed Start site for free technical support for Linux and the IBM software for Linux and developing over 4,200 new Linux applications using IBM Software.

For The True Geek

To thoroughly understand what makes Linux tick and why it's so efficient, one needs to delve deep into the heart of the operating system—into the Linux kernel itself. The kernel is Linux—in the case of the Linux operating system, it's the only bit of software to which the term "Linux" applies. The kernel handles all the requests or completed I/O operations and determines which programs will share its processing time, and in what order. Responsible for the sophisticated memory management of the whole system, the Linux kernel is the force behind the legendary Linux efficiency.

"Understanding the Linux Kernel, Second Edition" by Daniel P. Bovet and Marco Cesati (O'Reilly, US $49.95) takes readers on a guided tour through the most significant data structures, many algorithms, and programming tricks used in the kernel. Probing beyond the superficial features, the authors offer valuable insights to people who want to know how things really work inside their machine. Relevant segments of code are dissected and discussed line by line. The book covers more than just the functioning of the code, it explains the theoretical underpinnings of why Linux does things the way it does.

The new edition of "Understanding the Linux Kernel" has been updated to cover version 2.4 of the kernel, which is quite different from version 2.2: the virtual memory system is entirely new, support for multiprocessor systems is improved, and whole new classes of hardware devices have been added.

Sending Us E-Mail?

Due to the amount of e-mail MUUG receives, we've set up an auto-reply to give you immediate feedback, and redirect some of the e-mail to the appropriate places. Why not look at http://www.muug.mb.ca/about.html#contacts first?

Share Your Thoughts

E-mail us with your comments on the newsletter, whether it's criticisms or commendations, and continue to send in articles or ideas for same. If you have a How-To or other idea, and aren't ready to give a presentation at MUUG, an article is a great alternative! If you can write better than the editor, that's terrific; if you can't, submit it anyway and we'll get it into shape for publication. We know that many of you have some great ideas and lots of knowledge. Why not share? Mailto: editor@muug.mb.ca.