

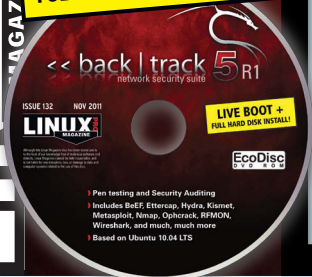
FREE DVD

back | track 5 R1
network security suite

REAL-WORLD HPC
BUILD A CLUSTER
STEP BY STEP

LINUX
MAGAZINE
PRO

LIVE BOOT +
FULL HARD DISK INSTALL!



LINUX

MAGAZINE

NOVEMBER 2011

Build a cluster with open source tools

Real-World HPC

OpenOffice tricks: Send a doc by email and get a PDF in reply

FASTER BOOT WITH
SYSTEMD SERVICE
MANAGER

GPT Partitioning:
WHY IS THE GUID PARTITION
TABLE REPLACING THE MBR?



Scripting with Squirrel


This lean little interpreter combines the best of Lua with an easy and familiar syntax

ZONEMINDER

Turn an ordinary webcam into a professional security video tool

BLUESPICE

MediaWiki add-on for business environments

Issue 132	US\$ 12.99
November 2011	CAN\$ 13.99
	
0 74470 58049 2	1 1

SUPERMICRO®

TwinBlade™

The Evolution of Blade Technology

10 "Double Density" Blades (20 DP Nodes) in 7U

- 20 DP (Intel® 5600 and AMD G34 Processor Series) nodes in 7U Enclosure
- Unparalleled 0.35U per DP Node
- 128 GB ECC Registered DDR3 per DP Node
- 94% Efficiency, N+ 1 Redundant 2500W Power Supplies
- Dual 40Gb InfiniBand, 10 GbE and 8Gb FCoE Switches
- Dual 1/10 Gb Layer 2/3 Ethernet Switches
- Dual Hot-Plug 2.5" SATA/SSD HDDs per DP Node
- Centralized Remote Management (IPMI 2.0, KVM-over-IP)

The Cost Advantages:

\$400+ Energy Cost Savings Per Node†

Compared to 1U Server:

- **Blade Only Per Node Saves 33%***
- **Plus GbE Switch Per Node Saves 33%***
- **Plus QDR IB Switch Per Node Saves 50%***

†Compared with typical 1U server over 3 year deployment.

*Based on average industry price.

It's In Our DNA!

**Selected 2010
Best Blade
Based Solution**
--Blade Systems
Insight



OfficeBlade™
10 DP Blades in 7U
(50dB) Optimized for
SMB & Personal
Supercomputing



Storage Blade
10 DP Blades in 7U
60 DP Blades with
360 SAS2 HDDs
in 42U Rack



DatacenterBlade™
14 DP Blades in 7U
PCI-E 2.0 Expansion
1008 Processor Cores
per 42U Rack



4-Way SuperBlade®
10 MP Blades in 7U
2880 Processor Cores
per 42U Rack



AMAX
1-800-800-6328
www.amax.com

Arrow Electronics
1-888-427-2250
www.arrowna.com

ASI
1-800-2000-ASI
www.asipartner.com

Bell Micro
1-800-232-9920
www.bellmicro.com

Ingram Micro
1-800-456-8000
www.ingrammicro.com

MA LABS
1-408-941-0808
www.malabs.com

Synnex
1-800-756-5974
www.synnex.com

Tech Data
1-800-237-8931
www.techdata.com

SED International
800-444-8962
www.sedonline.com

©Super Micro Computer, Inc. Specifications subject to change without notice. Intel, the Intel logo, Xeon, and Xeon inside, are trademarks or registered trademarks of Intel Corporation in the US and other countries. AMD, the AMD logo, Opteron, and Magny Cours, are trademarks or registered of AMD. All other brands and names are the property of their respective owners.

www.supermicro.com

HP PHONE HOME

Dear Linux Pro Reader,

As you will see if you read our news, Hewlett Packard, which everyone calls HP, was big this month. The computer giant announced some sweeping changes, including sweeping changes to undo the sweeping changes they announced a few months ago. Sweeping changes are always part of the scene in an economic downturn, but I can't help thinking it is better to figure out what you want to do and sweep in one direction, rather than trying to sweep several ways at once.

As a cultural icon, HP predates the personal computer era. When I was a teenager, no one had a smartphone, but all the geeky kids had calculators, and HP calculators were the coolest and geekiest calculators of all. With their elegant design and distinctive reverse polish notation, they washed coolness back onto the owner. At once simple, yet inscrutable to the uninitiated, they were the iPhone of their era.

But that was many years ago. HP's mystique didn't make the transition into the computer era. Their printers had a solid reputation for a while, but HP was one of many computer vendors to wake up one day and discover that the PC hardware market just doesn't lend itself to mystical brand names. In today's world, people buy PCs the way they buy cans of corn and beans at the grocery – a vague memory of a logo lingers until around dinnertime, but no one has any real expectation that one particular logo leads to a significantly different experience.

That's why I was interested in HP's announcement last February that it was working toward including its webOS mobile operating system in its new generation of personal computers. What a game changer. Of course, I'm partial to Linux, but to be honest, Android is looking (and acting) less like Linux every day. From a standpoint of the PC chess game, it was intriguing to see if HP could build up a highly distinctive and recognizable hardware/software combo platform that could compete with Apple/iOS, then actually extend that platform from the smartphone space back into the PC/tablet realm.

It is hard to believe HP would cash out this strategy after only six months and one failed TouchPad line. In fact, now it looks like they might be leaving the personal electronics business completely. Of course, they've had some big transitions, with a CEO leaving unexpectedly and another one arriving. Could it be because their new leader, Leo Apotheker, is consolidating his authority and turning the company toward the kind of behind-the-scenes business integration and support he used to preside over at SAP? Perhaps, or maybe, despite my best efforts to extract secret meaning in their moves, the real problem is just a bad

economy. Still, I can't help thinking these grandiose, drastic actions hither and thither, aiming for seemingly arbitrary and discontinuous targets, is an indication of the brash, "take charge" mentality we have come to value in business, in which upheaval is regarded as equivalent to innovation.

When the board of directors says, "What? Your TouchPad didn't make money? What are you going to do about it?" it is increasingly difficult for the CEO to say, "Not much. I'll just make some minor corrections and we'll keep doing what we've been doing."

The Wall Street-style model of uber-compensated, itinerant CEOs doesn't lend itself to the kind of single-minded determination it takes to navigate through hard times and stay on course into uncharted business conditions.

Up in Redmond, Microsoft's founder sits in the chairman's office. In Mountain View, a Google founder calls the shots, and over in Cupertino, until a few weeks ago, Apple's iconic creator remained at the reins. It is worth noting that, during HP's heyday as the maker of the coolest, most mystical calculators in the dorm, Bill Hewlett was still at the helm.

I bet they wish they could call him back.

Joe

Joe Casad,
Editor in Chief





LINUX PRO MAGAZINE

news

8 NEWS

• Is Linux really a threat to Microsoft?

9 HP UNCERTAINTY

• The TouchPad goes down. What about webOS?

10 KERNEL.ORG HACKED

• Latest GIMP.
• PC growth estimates.

11 ARM DEVICES

• HP buys Autonomy.

12 TECH TOOLS

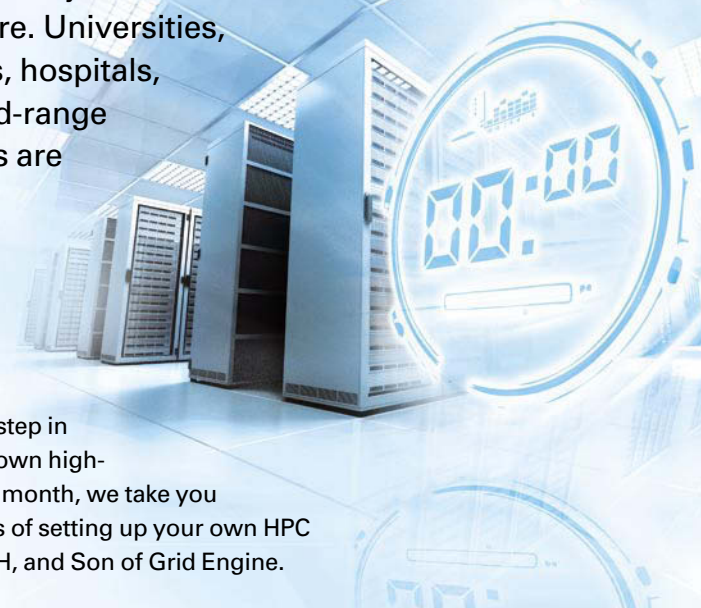
• SunGard Recover2Cloud.
• A10 SoftAX ADC.
• Cloud Cruiser for OpenStack.

Real World HPC

High performance isn't just for research institutions anymore. Universities, hosting companies, hospitals, and even some mid-range corporate networks are taking up the HPC challenge.

14^{HPC}

Buying the hardware is only the first step in the path to building your own high-performance cluster. This month, we take you through through the steps of setting up your own HPC system with Kickstart, SSH, and Son of Grid Engine.



SERVICE

3 Comment

6 DVD

96 Featured Events

97 Call for Papers

98 Preview

LINUX PRO MAGAZINE (ISSN 1752-9050) is published monthly by Linux New Media USA, LLC, 616 Kentucky St., Lawrence, KS, 66044, USA. Periodicals Postage paid at Lawrence, KS and additional mailing offices. Ride-Along Enclosed. POSTMASTER: Please send address changes to Linux Pro Magazine, 616 Kentucky St., Lawrence, KS 66044, USA.

Community Notebook

89 Doghouse: LinuxCon NA

maddog takes a trip to LinuxCon North America and reminisces on 20 years of Linux .

90 Kernel News

Kernel developer Zack Brown reports on rebooting Kernelhoops.org and weighs in on long-term kernel plans.

92 Projects

If you've tried Drupal and WordPress and you're ready for something new, post up with the Joomla and concrete5 content management systems.



HIGHLIGHTS

14 HPC
Follow these steps to set up your own high-performance cluster.

32 SQUIRREL
If you like Lua, you'll love the Squirrel scripting language.

38 GPT
Find out why this partition scheme is replacing the MBR.

74 ZONEMINDER
Let this helpful video monitor keep an eye on your stuff.

FEATURES

54 Mallard
Build better and simpler topic-based help documents.



58 systemd
Speed up the boot process and manage system services.

28 BlueSpice
This innovative add-on brings the MediaWiki collaboration engine to enterprise users.

32 Squirrel
Gather nuts with this lean and lively scripting engine.

38 GPT Partitions
The GUID partition table replaces the legacy Master Boot Record partitioning scheme.

43 PDF Converter
Send a doc by email and receive a PDF as a reply.

46 Security Lessons: Passwords
Tips for proper password storage.

48 Perl: VirtualBox Snapshots
Testing CPN modules in a pristine environment with VirtualBox snapshots.

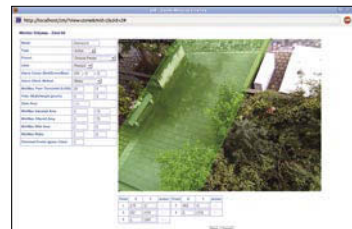
57 Charly: OpenNetAdmin
Track and manage network objects with OpenNetAdmin.

66 Ask Klaus!
Klaus Knopper answers your Linux questions. This month: Dynamic Doldrums and KVM.

LINUXUSER

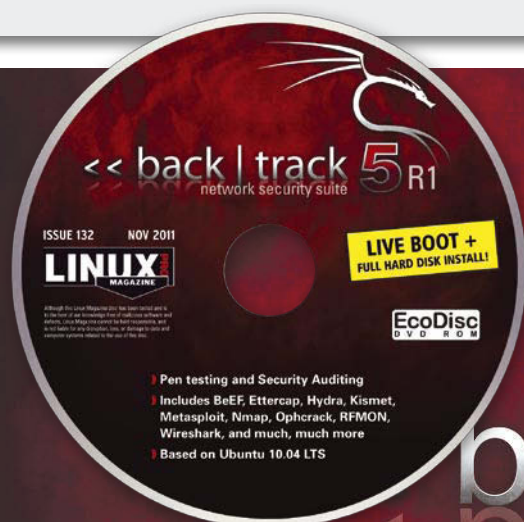
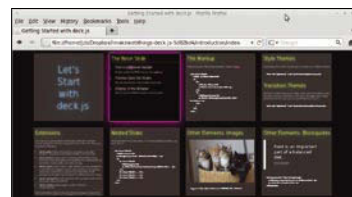
70 Workspace: phTagr
Publish and share your images on the web.

74 Zoneminder
Turn your webcam into a security camera.



82 Command Line: Compiling
You can go a long way with package managers, but sooner or later, you'll need to compile some source code.

86 deck.js
This nifty presentation tool is so simple – if you know a little HTML.



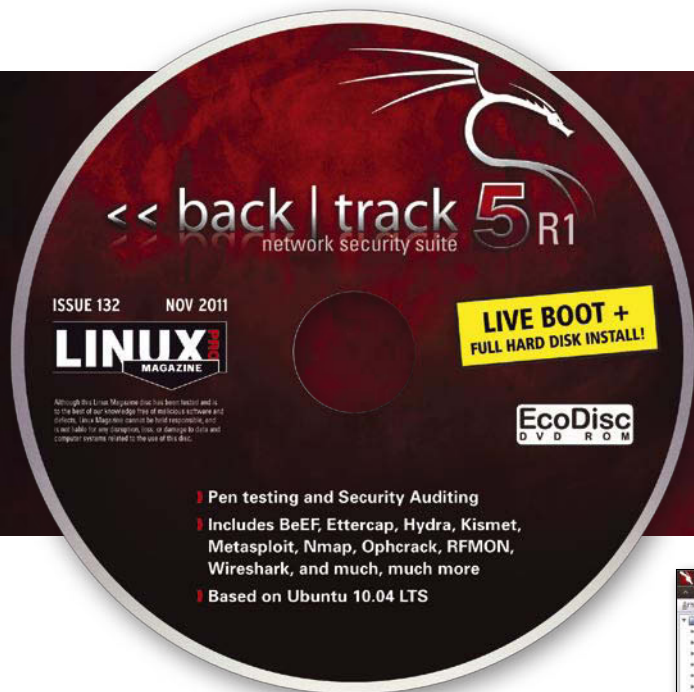
Includes:

- Pen testing and security auditing tools
- BeEF, Ettercap, Hydra, Kismet, Metasploit, Nmap, Ophcrack, RFMON, Wireshark, and much, much more
- Based on Ubuntu 10.04 LTS

SEE P6 FOR FULL DETAILS

back | track 5

On the DVD



- Pen testing and Security Auditing
- Includes BeEF, Ettercap, Hydra, Kismet, Metasploit, Nmap, Ophcrack, RFMON, Wireshark, and much, much more
- Based on Ubuntu 10.04 LTS

Backtrack 5 R1

This issue's DVD comes with the BackTrack 5 R1 [1][2][3] pen test distribution. BackTrack provides a great collection of pen testing and security auditing tools. You can boot into BackTrack Live from the DVD or install BackTrack permanently on your hard disk.

BackTrack 5 R1, which is based on Ubuntu 10.04 LTS, includes more than 120 bug fixes. 70 applications have received updates, and 30 new applications appear for the first time.

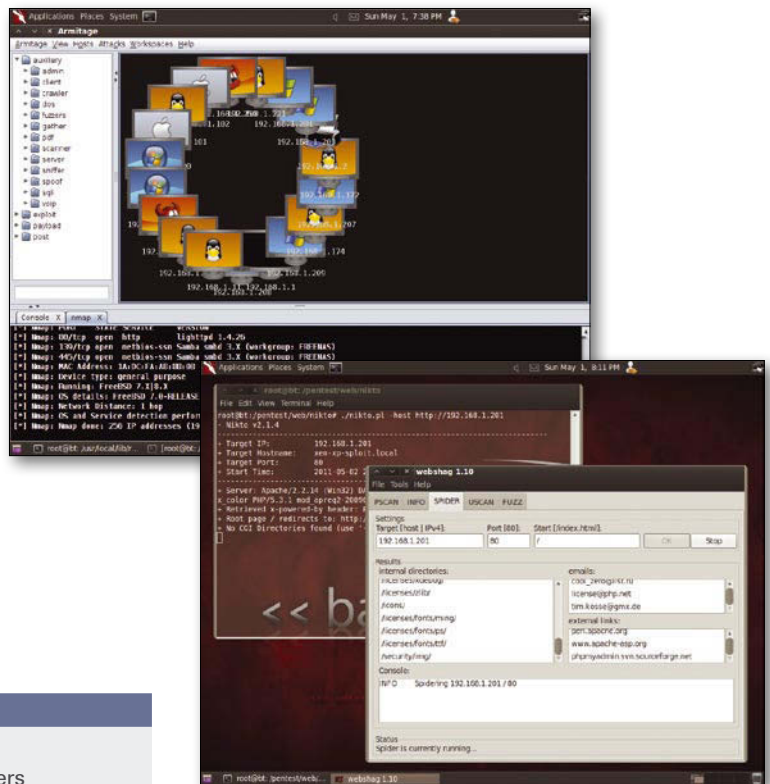


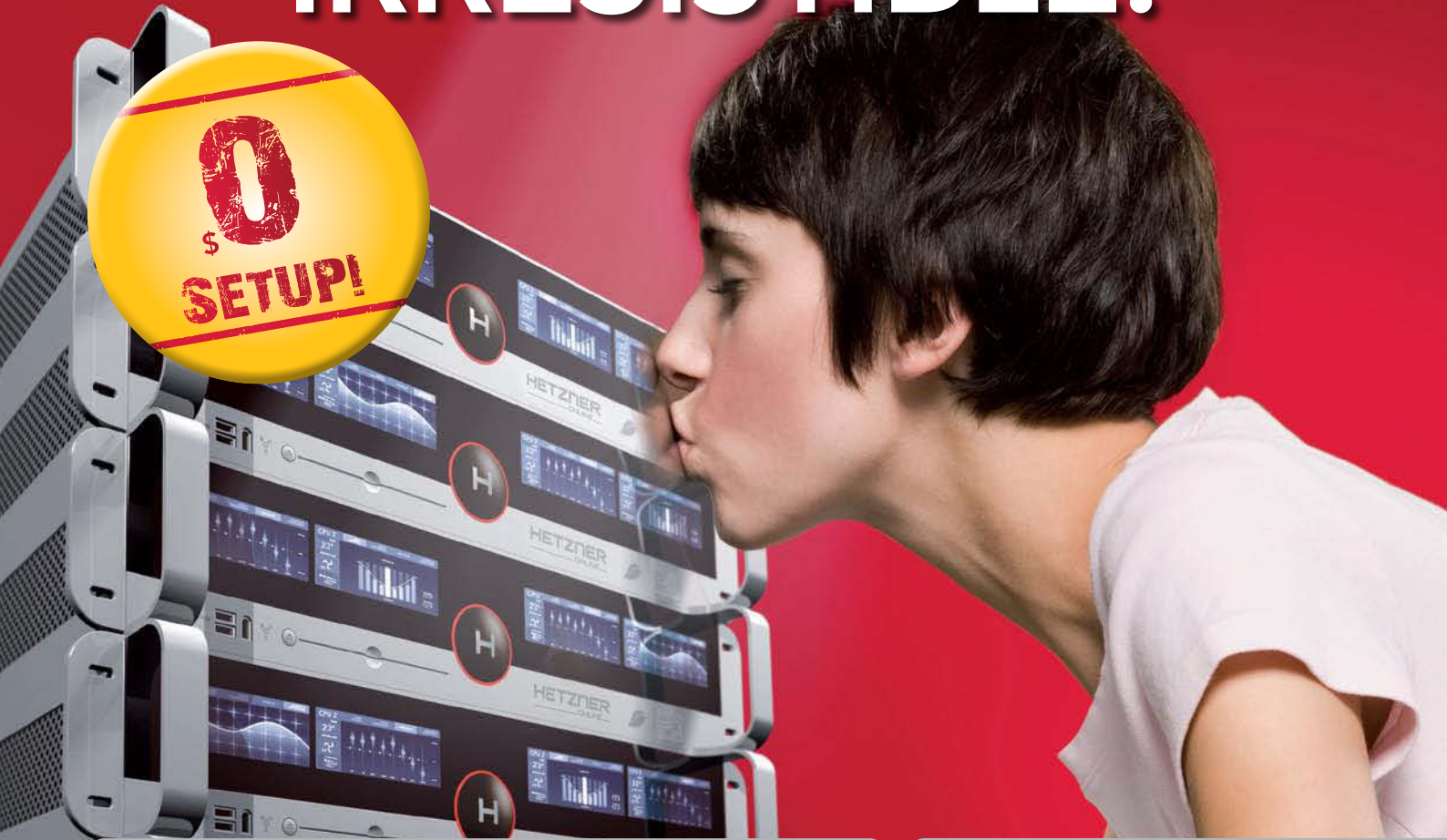
TABLE 1: Backtrack Applications

Application	Function
BeEF	Browser exploitation
Cisco OCS	Scanning for Cisco routers
Ettercap	Man-in-the-middle attacks on LAN
Hydra	Password cracking
Kismet	Packet sniffer
Metasploit	Pen testing
Nmap	Network scanning
Ophcrack	Password cracking
RFMON	Wireless traffic monitor
Quyp	Terminal emulator
Wireshark	Packet analyzer

ADDITIONAL RESOURCES

- [1] BackTrack Website:
<http://www.backtrack-linux.org/>
- [2] BackTrack Wiki:
<http://www.backtrack-linux.org/wiki/>
- [3] BackTrack Forums:
<http://www.backtrack-linux.org/forums/>

SIMPLY IRRESISTIBLE!



HETZNER vSERVER VQ7

- Single Core CPU
- 512 MB RAM
- 20 GB HDD
- Linux operating system
- 1 TB traffic*
- Full root access
- IPv6 subnet (/64)
- Rescue system
- Image Installer
- Domain Registration Robot
- No minimum contract
- Setup Fee \$ 0

monthly \$

10

HETZNER vSERVER VQ12

- Single Core CPU
- 1024 MB RAM
- 40 GB HDD
- Linux operating system
- 2 TB traffic*
- Full root access
- IPv6 subnet (/64)
- Rescue system
- Image Installer
- Domain Registration Robot
- No minimum contract
- Setup Fee \$ 0

monthly \$

16

HETZNER vSERVER VQ19

- Single Core CPU
- 2048 MB RAM
- 80 GB HDD
- Linux operating system
- 4 TB traffic*
- Full root access
- IPv6 subnet (/64)
- Rescue system
- Image Installer
- Domain Registration Robot
- No minimum contract
- Setup Fee \$ 0

monthly \$

24

* If the limit is exceeded we will permanently restrict the connection speed to 10 MBit/s. 100 MBit/s speed can be optionally restored by committing to pay \$ 8 per additional TB used.



Hetzner Online actively supports the protection of the environment by using 100% renewable energy. Take the step with us towards a cleaner future.

WWW.HETZNER.INFO

All prices exclude VAT and are subject to the terms and conditions of Hetzner Online AG. Prices are subject to change. All rights reserved by the respective manufacturers.

news

UPDATES ON TECHNOLOGIES, TRENDS, AND TOOLS

Joe "Zonker" Brockmeier

This Month

Threats

As you could probably guess, a Microsoft exec's August 16 tweet, stating that Linux is "not a threat to the Windows desktop any longer" caused a colorful range of emotions in the open source blog circles. Many saw this as a slight – an insult to the honor of Linux after so many people have worked for so long to get the Linux desktop into the big leagues.

At least it is good that Microsoft is backing away from fearful battleground terms like "threat." Businesses talk about "threats" and "dangers" all the time, but that ought to be something they reserve for other businesses. The word "threat" never did fit very well with the world of free software, where software products co-exist all the time without threatening each other. Maybe IBM is a threat for Microsoft. Maybe Novell/Suse was a threat (until it went away). Maybe even Canonical is a threat – but not Linux itself, because it is free to anyone who wants to use it, including Microsoft.

Microsoft saying Linux is a threat seems a little like a soft drink company saying water is a threat – like they are publicly accusing their competitors (or "enemies," to continue with the martial metaphor) of maliciously spreading the dangerous rumor that people could be drinking water for free rather than paying money for Lucky Star Cola.

To mark this passing of the Linux threat to Microsoft, it is only fitting to look back and remember what one of the top Redmond rangers said about Linux back in 2001. Microsoft CEO Steve Ballmer famously stated that "Linux is a cancer that attaches itself in an intellectual property sense to everything it touches."

Hey, so maybe Linux isn't so bad after all, eh Microsoft? Like maybe you overstated it just a teeny teeny tiny bit back in 2001? Like maybe Linux isn't a cancer? Like maybe you aren't so credible, are you now...

THIS MONTH'S NEWS

- 09 HP Uncertainty**
 - The TouchPad goes down. What about webOS?
 - News Bites:
 - New RHEV due out
 - New X.org server released
- 10 Kernel.org Hacked**
 - Latest GIMP
 - PC growth estimates cut
- 11 ARM Devices**
 - HP buys Autonomy
- 12 Tech Tools**
 - SunGard Recover2Cloud
 - A10 SoftAX ADC
 - Cloud Cruiser for OpenStack

MORE NEWS ONLINE

- Caché Database offers high availability
<http://www.admin-magazine.com/News/Cache-Database-Now-Offers-High-Availability>
- ASP.Net security vulnerability patched
<http://www.admin-magazine.com/News/Microsoft-Patches-The-ASP.Net-Security-Vulnerability>
- Vyatta certified for IPv6
<http://www.admin-magazine.com/News/Router-distribution-Vyatta-is-now-certified-for-IPv6>
- New beta of Opennode software
<http://www.admin-magazine.com/News/New-Beta-of-Opennode-Virtualization>
- Debconf videos now online
<http://www.admin-magazine.com/News/Videos-of-Debconf-2010-online>

Future Uncertain for webOS

Just two months after it became available, HP announced in a press release that it would discontinue its TouchPad tablets and smartphone devices that run webOS, the Linux-based operating system it inherited in the \$1.2 billion purchase of Palm in April 2010. Sluggish sales and complaints from retailers about unsold inventories led HP to cut prices on the tablets several times, eventually slashing them to US\$ 99 and US\$ 149 for the 16GB and 32GB models. That move set off a buying frenzy for the reduced-price tablets at online merchants and big box retailers, some of whom had already shipped their inventories back to HP. The lack of inventory and sudden demand caused prices for the tablets to rise in some aftermarket online auctions to as much as \$300, and the sudden interest even led HP to start a Twitter feed to keep people updated on future TouchPad availability. With the unexpected demand and their manufacturing partners reportedly stuck with parts for a large number of the devices, HP announced on its blog that it would produce a limited number of TouchPads in its fourth fiscal quarter, which ends October 31.

In addition to price, TouchPad critics disliked the hardware and lack of apps, but webOS seems to have been reasonably well regarded, even by HP. In the week before the announcement, an HP spokesman was quoted in a *Wall Street Journal* article saying HP had plans to try to license webOS for use in household appliances and automobiles. And while the TouchPad may be toast, HP said it will continue to explore options to optimize the value of webOS software going forward, which has led to speculation that firms such as Amazon, HTC, Facebook, or Samsung may have designs on acquiring it. Samsung CEO Choi Gee Sung was later quoted in Berlin as saying his company would “never” make such an acquisition, and that “...it’s not right that acquiring an operating system is becoming a fashion.” Samsung has its own mobile OS, Bada, but uses Android in its Galaxy line of tablets.

It seems clear that the patents behind webOS are valuable, but their ownership appears complex, which may complicate any HP effort to sell or license them, and HP has announced no concrete plans on either front. The firm has, however, apparently decided to split management of the former Palm software assets, including webOS, from their hardware brethren. The site www.precentral.net published information from two leaked memos to employees of HP’s webOS Global Business Unit that revealed plans to move management of the software to the HP Office of Strategy and Technology, while leaving the hardware under its Personal Systems Group.

In the wake of the TouchPad fire sale, at least two different groups announced plans to get Google’s Android operating system running on the TouchPad hardware, and one group, CyanogenMod, aims to make the TouchPad bootable by multiple operating systems, including webOS. The group has released a video showing an alpha version of its Android variant booting up on a TouchPad. Another group, which dubbed its effort TouchDroid, posted plans on a RootzWiki forum to develop a port of Android 2.3 (Gingerbread) on TouchPad hardware and has also launched a Touchdroid wiki for developers who may be interested.

Adding another wrinkle to the saga, Microsoft is reportedly offering phones, development tools, and training on their Windows phone platform to webOS developers who have had applications added to the webOS arsenal, and as many as 1,000 developers have reportedly contacted the firm about the offer.

NEWS BITES

RHEV 3.0 Will Close Curtain on Windows

Red Hat announced that it has redesigned its Red Hat Enterprise Virtualization (RHEV) management system as a Java application running on the JBoss Enterprise Application Platform on Red Hat Enterprise Linux (RHEL), which means that RHEV 3.0 customers will no longer have to run it on Microsoft’s Windows platform. RHEV 3.0 is expected to be available by the end of the year. The new version will still be Windows-compatible and will continue to support Active Directory, and the zero-downtime upgrade will be available at no charge to customers. RHEV 3.0 will also feature a new version of the KVM hypervisor based on RHEL 6, and it will use the open source PostgreSQL database.

Notable enhancements include a portal that enables power users to provision virtual machines, set template definitions, and administer their own systems, new local storage features, and a reporting engine to help track usage trends and provide reports. For more information, see <http://www.redhat.com/about/news/prarchive/2011/Red-Hat-Enterprise-Virtualization-3-0-Beta-Now-Available>

X.org Releases Server 1.11

The X.org project, home of a major open source implementation of the X Window System, announced the release of X.org Server 1.11, a new release that incorporates six months of improvements, highlighted by a number of bug fixes.

This is not a major release, and a X.org plans to address a number of issues in subsequent releases – including a video driver ABI break, the RandR extensions, and the touch features for X Input 2.1. Still, it’s great to see the fixes in 1.11. The X.org Server 1.12, which is expected in about six months, also could include new support for smooth scrolling. For more information, visit http://www.phoronix.com/scan.php?page=news_item&px=OTg0MA

Kernel.org hacked; Linux Kernel Moved to Github

According to a message sent to Kernel.org users by John “Warthog9” Hawley, Chief Kernel.org Administrator, servers that host, maintain, and distribute the Linux kernel code suffered a Trojan-based security breach “no later than August 12” that was discovered August 29. According to the message, “a trojan startup file was added to rc3.d” and “Files belonging to ssh (openssh, openssh-server and openssh-clients) were modified and running live. These have been uninstalled and removed, all processes were killed and known good copies were reinstalled.” Users were urged to change their password and SSH keys, and the site was shut down temporarily for maintenance. The organization later confirmed in a blog post that the attacker gained root access, modified software, and recovered passwords of developers who used several servers on the system.

Site maintainers at Kernel.org posted a statement to the site’s homepage shortly after news of Hawley’s message broke that said, “We believe they may have gained this access via a compromised user credential; how they managed to exploit that to root access is currently unknown and is being investigated. Probably the most important point in the message, and one of the most interesting things about the case, is the knowledge that the security features in use in the system, which controls how Linux code is developed and tracked, limits how much havoc a rogue entity can wreak. The statement explained that for each of the nearly 40,000 files in the Linux kernel, a cryptographically secure SHA-1 hash is calculated to uniquely define the exact contents of that file, and it is not possible to change the old versions without the change being noticed. Further, because each hash is ar-

chived on so many different systems around the world, it is possible to verify the validity of Linux components before implementing them.

With the Kernel.org site down, Linux creator Linus Torvalds decided to temporarily move the latest Linux kernel code to Github, a hosted Git repository. Git is a distributed software management program originally created by Torvalds for Linux Kernel Development. “Since I did a github account for my divelog thing, why not see how well it holds up to me just putting my whole kernel repo there, too? So, while kernel.org is down for the count, let’s just see how github does,” Torvalds wrote in announcing the latest release candidate, 3.1RC5. Torvalds also wrote that he will move the code back to Kernel.org when it comes back and will use the GitHub repository as a mirror. Visit Github at <https://github.com/>

Latest GIMP features single-window mode

The spry folk that manage the free graphics application GIMP (GNU Image Manipulation Program) have released version 2.7.3, which includes the much anticipated “single-window mode,” an optional view that unites all open windows and toolbars in a single interface and makes GIMP more competitive with various popular commercial graphics apps.

The single window mode does not replace the multi-window mode to which users have become so accustomed, but it does offer a nice change of pace for those who like things in a single view. The multi-window mode has also been enhanced with a number of improvements that went into the single window mode, including support for horizontal docking, which allows you to dock multiple panels side by side. Many other interface enhancements are included here, such as the ability to add multiple items to the image window and create new images from them, sharing brush modifying paint options among other paint tools, and memory of column widths in multi-column dock windows. Also welcome are the ability to use ALT + number and ALT + Tab shortcuts to navigate between images in both single window and multi-window mode, and a fix to make the rectangle tool’s guides draw correctly, along with lots of bug fixes, translation updates, and additional tool tips.

Unless you are waiting for the next stable release, 2.8, you can download GIMP at <http://www.gimp.org/downloads/>. Details on the 2.7.3 release are at <http://developer.gimp.org/NEWS>

Gartner Cuts PC Shipment Growth Estimates

In a press release on its website, IT research and advisory firm Gartner, Inc. announced lower growth estimates for worldwide PC shipments in 2011 and 2012, cutting their 2011 estimate from 9.3% to 3.8%. For 2012, the estimate was reduced from 12.8 to 10.9%. In all, Gartner expects 364 million units to ship in 2011 and 400 million to move in 2012. In dropping the estimates, Gartner cited unfavorable economic conditions, excess PC inventory, disappointing back-to-school PC sales numbers, users repairing and extending the lives of devices instead of replacing them, and the rise of mobile devices. “Media tablets have dramatically changed the dynamic of the PC market and HP’s decision to rethink its PC strategy simply highlights the pressure that PC vendors are under to adapt to the new dynamic or abandon the market,” said George Shiffler, research director at Gartner. Ranjit Atwal, also a research director at Gartner, stated, “More worrisome for the long term is that Generation Y has an altogether different view of client devices than older generations and are not buying PCs as their first, or necessarily main, device. For older buyers, today’s PCs are not a particularly compelling product, so they continue to extend lifetimes, as PC shops and IT departments repair rather than replace these systems.” The analysts expect sales to pick up in the second half of 2012 as economies improve and “new mobile PC form factors enter the market.” Gartner’s PC shipment numbers do not include media tablets, which are included in separate estimate. For more information, visit <http://www.gartner.com/it/page.jsp?id=1786014>

Android and Ubuntu images for ARM Devices

If you are hunting for a Linux or Android build that is optimized for small-form-factor, low-power, multimedia-ready boards for mobile and embedded applications that run chips using the



ARM architecture, look no further than the Linaro website. The Linaro team announced the release of the latest versions of its open source offerings for Android and Ubuntu 11.04 (Natty Narwhal) on devices using ARM Cortex-based processors. The latest release, 11.08, is the third one distributed on the team's new monthly cycle, and it is built

on version 3.0.3 of the Linux kernel, includes the merge of 3.1-rc1 core ARM updates, and uses the GCC 4.6 tool-chain. Changes in this release include improved graphics, enhanced multimedia features, better power management, and the default file system has been changed to Ext4.

The Android 2.3.5 package is available in a fully supported release for the PandaBoard, which is a low-power, single-board computer based on Texas Instruments'

OMAP4430 processor (ARM Cortex-A9) with a community supported development platform, and its predecessor, the BeagleBoard (ARM Cortex-A8). Unsupported preview builds are available for boards using Freescale's i.MX53 processor (ARM Cortex-A8 core), Samsung's Origen (ARM Cortex-A9), and the Snowball from ST-Ericsson (ARM Cortex A-9).

The Ubuntu package is available in a fully supported release for the PandaBoard and in preview builds for the i.MX53 and Origen boards.

Founded in June, 2010 to develop standards-based open source Linux tools, kernel, and middleware software for consumer electronics using processors based on the ARM architecture, the not-for-profit software engineering foundation includes ARM Holdings, Freescale, IBM, Samsung, ST-Ericsson, Texas Instruments and Partner Program members Canonical, Collabora, Genesi, Mentor Graphics, and Thundersoft. The group's goal is to make Linux a player in the TV, tablet, smartphone, and netbook markets dominated by Apple, Symbian, Blackberry, and Microsoft. Linaro plans to support tools on distributions including Android, LiMo, MeeGo, Ubuntu, and perhaps, the recently deprecated webOS. For more information, visit <http://www.linaro.org/>

HP Buys Autonomy, May Exit PC Biz

In a strategic move aimed at leaving behind the low profit margins of the PC hardware business, HP announced that its board of directors has "authorized the evaluation of strategic alternatives" for its personal computer sales business, alternatives that include spinning off its Personal Systems Group (PSG) into a separate company, or even selling the business outright. At press time, no determination had been made about whether to sell or spin off the unit, but HP PSG continues to release products, having just announced the availability of seven new all-in-one machines for the US market and a new Pavilion DM1 notebook.

HP PSG is the world's largest PC vendor, with US\$ 40 billion in revenues in 2010, and a 16.6% market share of desktops, netbooks, notebooks and pads, according to a first quarter 2011 survey by Canalsys. Acer and Dell are second and third in the same survey, at 12.8 and 11.3 percent.

Not content with simply getting out of the hardware business, HP also announced an approximately US\$ 11 billion deal to purchase English enterprise search software provider Autonomy for US\$ 42.11 a share, a 64% premium over the stock's closing price the day before the deal was announced. HP will own Autonomy through a wholly owned subsidiary, HP SPV, and expects it to strengthen its enterprise software business and cloud-based offerings to help it better compete in those segments with global providers such as IBM, which sold its PC division to Lenovo in 2005. Read more at <http://www.hp.com/hpinfo/newsroom/press/2011/news2011.html>

 The advertisement for the Ada Initiative features a dark background with a stylized, glowing red and white graphic of a hand or a flame-like shape. The text is arranged as follows:

- On the left, vertically oriented: www.adainitiative.org
- Top right: "support the" in white, followed by "Ada" in large white letters, and "initiative" in red letters below it.
- Bottom right: "Shape the future." in red, followed by "Increase the participation of women in open technology and culture." in white.

TECH TOOLS

SunGard Announces Cloud Recovery Solution

SunGard announced the availability of Recover2Cloud, a new suite of recovery services for cloud platforms. The new offering is “backed by guaranteed service levels,” meaning SunGard takes responsibility for recovering wayward data, performs round-the-clock monitoring, runs environment testing and troubleshooting, and performs the recovery, should one be necessary. “Leveraging a cloud platform enables us to expand the recovery options we offer to our customers, both as standalone services and as part of tiered recovery solutions. Based on

what we’re hearing from customers, the managed recovery aspect we are building into many of our services is expected to help deliver significant value. It is an area where in-house knowledge is often missing, and where SunGard has significant experience,” said Larry Coble, senior vice president and general manager of recovery services at SunGard Availability Services, in the press release.

The new suite offers a variety of service levels for recovery time and recovery point. Recover2Cloud for Server Replica-



tion uses asynchronous server-based replication and is appropriate for

“virtual, physical, and cloud-based applications” that might need a sub-four-hour recovery window. Recover2Cloud for Vaulting offers recovery of applications within 24 hours and restoration of data from an online “vault.” A related service to help customers leverage storage virtualization technology is due for release later in the year. For more information, see <http://www.sungardas.com>.

A10 Networks Adds to SoftAX Line

Networking and security solution provider A10 Networks announced the rollout of two new additions to its line of SoftAX virtual machines (VMs) for VMware: a basic 200Mbps offering and a high-performance 8Gbps version. The SoftAX is the only virtual Application Delivery Controller (ADC) to be certified by both Microsoft (for Microsoft Office Communications Server (OCS), Microsoft Lync, and Microsoft Exchange 2010) and VMware (“VMready” for vSphere compatibility). “We are committed to ensuring the SoftAX, along with our flagship AX Series hardware appliances, meets the exact standards of leading applica-

tion vendors, guaranteeing the best experience for our customers,” said Lee Chen, Founder and CEO for A10 Networks, in a press release.



The high-end 8Gbps SoftAX release is intended for data centers and cloud providers requiring a high-performance, high-bandwidth solution that can accommodate unexpected surges in demand. The 8Gbps SoftAX joins an existing 4Gbps offering in A10’s line of “production - High-performance Editions,” while the 200Mbps edition becomes the entry point in A10s “Production-Entry Level/Lab Editions,” which also include a 1Gbps version. “The new 8Gbps SoftAX is the best performing virtual appliance available in the industry, and we are pushing the boundaries in this nascent yet rapidly growing market,” said Chen. The 200Mbps SoftAX is designed for clients who want SoftAX ADC features on a lower budget and can be upgraded on a “pay-as-you-grow” basis.

For more information, visit <http://www.a10networks.com/products/axseries-softax.php>

Cloud Cruiser Debuts Cost Management Tool for OpenStack

Cloud computing cost management solution provider Cloud Cruiser announced the availability of a new cost management solution for OpenStack that includes workflow tools with accounting and management features for optimizing costs in cloud environments. OpenStack is an Apache-licensed FOSS cloud computing project with more than 100 members, including AMD, Citrix, Dell, Intel, and others. Announced as the first of its type for

OpenStack, the solution targets server providers and enterprise OpenStack users. “By offering the industry’s first cost management solution for OpenStack, users will now have the capability to gain control of costs and visibility of their cloud-based computing environments with new tools to measure, manage, and optimize capex and opex costs,” said

Dave Zabrowski, Cloud Cruiser Founder and CEO, in the press release.

The new offering promises the ability to track and limit IT costs in a variety of cloud environments supporting internal, private, and public applications and allows users to “see, charge, control, and own” cloud processing costs right down to the individual user. For more information, see <http://www.cloudcruiser.com>.



SUPERMICR®

Double Side IT

The Evolution of Storage

- Up to 88 Hot-Swappable HDDs in 4U (2.5" HDDs)
- High Performance 6 Gb/s SAS 2.0 Expanders, RAID Cards, and Serverboards
- 2.5" and 3.5" HDD Support
- Gold/Platinum Level (93%/94%) high-efficiency power supplies
- 100% Cooling Redundancy
- Platinum/Gold Level Redundant Power Supplies (94%+ Efficiency)
- Up to 7 expansion LP slot or 4 FH Slot for UIO System Board.
- DP and UP Serverboards Compatibility
- 2U, 3U, and 4U form factors
- Ideal for enterprise data centers and high performance applications
- Top Performance with High Density, Serviceability, Reliability, and Scalability



SC216A/E1/E2/E16/E26
24 Hot-Swap HDDs in 2U

SC936/A/E1/E16/E26
16 Hot-Swap HDDs in 3U



SC836/A/E1/E2/E16/E26
and **SC836-JBOD**
16 Hot-Swap HDDs in 3U



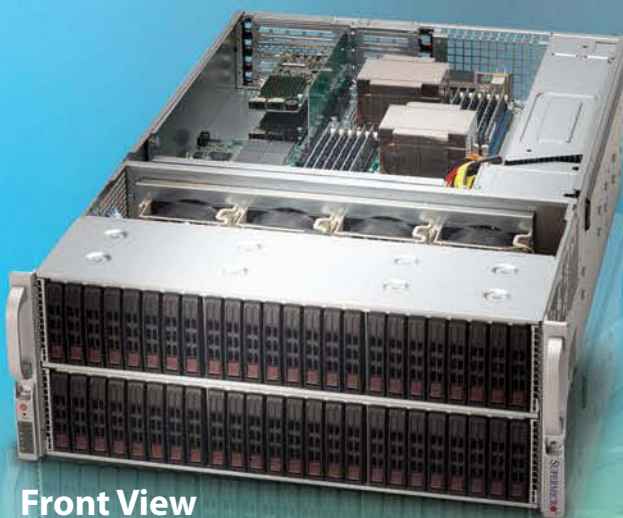
Front View

SC847
(24) 3.5" SATA/SAS
Hot-Swap HDDs in 4U

Rear View

SC847-JBOD
45 3.5" SATA/SAS Hot-Swap HDDs in 4U
22 TB Disk Capacity Per U

It's In Our DNA.



Front View
SC417E16/E26
Double Sided Storage
(72) 2.5" Hot-Swap HDDs in 4U



Rear View
SC417-JBOD
Double Sided Storage
(88) 2.5" Hot-Swap HDDs in 4U



AMAX
1-800-800-6328
www.amax.com

Arrow Electronics
1-888-427-2250
www.arrownaep.com

ASI
1-800-2000-ASI
www.asipartner.com

Avnet
1-800-409-1483
www.avnet.com

Ingram Micro
1-800-456-8000
www.ingrammicro.com

MA LABS
1-408-941-0808
www.malabs.com

Synnex
1-800-756-5974
www.synnex.com

Tech Data
1-800-237-8931
www.techdata.com

SED International
800-444-8962
www.sedonline.com

©Super Micro Computer, Inc. Specifications subject to change without notice. Intel, the Intel logo, Xeon, and Xeon inside, are trademarks or registered trademarks of Intel Corporation in the US and other countries. All other brands and names are the property of their respective owners.

www.supermicro.com



Real-world HPC step by step

Up Close

Setting up a real-world HPC cluster with Kickstart, SSH, Son of Grid Engine, and other free tools. *By Gavin W. Burris*

High-Performance Computing (HPC) clusters are characterized by many cores and processors, lots of memory, high-speed networking, and large data stores – all shared across many rack-mounted servers. User programs that run on a cluster are called jobs, and they are typically managed through a queueing system for optimal utilization of all available resources. An HPC cluster is made of many separate servers, called nodes, possibly filling an entire data center with dozens of power-hungry racks. HPC typically involves simulation of numerical models or analysis of data from scientific instrumentation. At the core of HPC is manageable hardware and systems software wrangled by systems programmers, which allow researchers to devote their energies to their code. This article describes a simple software stack that could be the starting point for your own HPC cluster.

What does an HPC cluster look like? A typical cluster for a research group might contain a rack full of 1U servers (Figure 1). The cluster might be as small as four nodes, or it could fill an entire rack with equipment. (See the box titled “Hardware Specs” for more on the hardware configuration used for this article.)

Before You Begin

Before delving into the setup steps, I’ll start with a look at some planning and purchasing issues. Being a systems programmer with eyes on the code, I am often surprised at the number of real-world logistical issues in need of consideration.

Before you order the hardware, you’d better give some thought to a few basic questions. Why do you need a cluster? What are your peers doing? Is a cluster the correct solution? Have system requirements been collected and base specifications decided? Start with an initial quote from a choice vendor. From there, move on to obtain at least two other comparable competing vendor quotes. Even with comparable quotes in-hand, go the extra mile. Ask if you can run a representative job on proposed hardware in a vendor lab for benchmarking. What are the application bottlenecks, and can the budget be applied to maximize those metrics?

Before issuing a purchase order, consider the space and logistical issues. Will the hardware be assembled, cabled, and verified by the vendor? On- or off-site? Have you reserved an adequate rack footprint to house the equipment? A cluster generates excessive heat, so server room footprint placement is key. The space must have adequate cooling capacity. Are you interleaving the hottest racks in the server room? With a space decided, move on to scheduling an electrician and securing budget codes for billing. Don’t forget to put in all requests for required networking, inventory, and card/key access for primary systems administrators. No matter how much planning you do, you’re bound to be surprised by something.

Be sure to consider the backup scheme and include it in your budgeting process. Backup can double storage costs and will blow your budget if it is left for last. Always make sure you have the correct rack configuration, correct voltage PDUs, and correct plugs, with only needed systems on UPS. Make sure you acquire all software licensing and installation media ahead of time. If possible, request vendor off-site assembly and

AUTHOR

Gavin W. Burris is a Senior Systems Programmer with the University of Pennsylvania School of Arts and Sciences. In his position with the Information Security and Unix Systems group, he works with HPC clusters, server room management, research computing desktops, web content management, and large data storage systems. He maintains a blog at: <http://idolinux.blogspot.com>



integration. The vendor must provide a properly trained, equipped, and insured delivery service for *inside delivery* to the designated footprint with a lift-gate truck and debris removal. A rack of equipment can literally

weigh a ton. If the rack falls off the delivery truck, you do not want to be the one responsible for or underneath it. And there will be *NO* unscheduled deliveries directly to the server room.

The Configuration

Once you have made all the necessary plans, chosen a preferred vendor, and had the system delivered, it's time for the on-site software installation and configuration. For something as complicated as an HPC cluster, all configurations will vary, but the example configuration described in this article will give you an idea of the kind of choices you'll face if you try this on your own.

For this article, I assume you have some familiarity with managing and configuring Linux systems. Most of the commands mentioned in this article are documented in a relevant man page, or you can find additional information through the Red Hat Enterprise Linux [1] or CentOS [2] documentation. In particular, the Red Hat Enterprise Linux Installation Guide [3] and the section on Kickstart installations [4] are important reading.

The basic steps for getting your HPC cluster up and running are as follows:

1. Create the admin node and configure it to act as an installation server for the compute nodes in the cluster. This includes configuring the system to receive PXE client connections as well as setting it up to support automated Kickstart installations.
2. Boot the compute nodes one by one, connecting to the admin server and launching the installation.

3. When all your nodes are all up and running, install a job queue system to get them working together as a high performance cluster.

At this point, I'll assume you have a full rack of equipment, neatly wired (Figure 2) and plugged into power, housed in a hosting facility. Let's call the whole system *Fiji*, a short and pleasant name to type. Before you go to the server room, the first step is to download and burn the CentOS install DVD media for 64-bit [5], currently version 5.6.

Once you have the media in-hand, begin by installing the admin node, which will act as a kickstart installation server to automate the install of all the compute nodes that fill out the rest of the cluster. Even though this article is targeted toward installing a cluster on enterprise-grade rack hardware, there is no reason one cannot follow along at home with a few beige boxes and

HARDWARE SPECS

- * 42U rack
- * 1U KVM
- * 3000VA UPS 208 Volt
- * Dell PowerConnect 6248 gigE switch
- * 2x Dell PowerEdge R610 for head/login and storage/admin nodes
- * 2x Dell PowerVault MD1000 for 12TB of storage
- * 29x Dell PowerEdge R410 compute nodes
- * 29x quad data rate HBAs and 1x switch InfiniBand fabric interconnect from QLogic



Figure 1: A new cluster still on the pallet.



Figure 2: Cluster cabling can get out of hand quickly, so be neat.

an 10Mb hub, as long as the servers are capable of booting from the network with PXE [6]. PXE (which stands for Preboot eXecution Environment”) lets you install and configure the compute nodes without having to stop and boot each one from a CD.

When deploying a multi-user cluster, where users log in via SSH, VNC or NX, it is also a good idea to have a separate head node for the interactive logins. For a small to medium-sized cluster, it is okay to combine what is often called the *storage node* with the admin node. In this article, I assume the cluster consists of one full rack of equipment, but you can extend the techniques described here to replicate this process *N* times for a larger cluster.

The admin/storage node and the head/login node both have multiple Ethernet interfaces. The first interface (`eth0`) connects to the private, internal gigabit switch and 192.168.1.0 network, where all internal communication between nodes and management interfaces will occur. The second Ethernet interface (`eth1`) will connect to the public network for remote user and admin login. This is important to note because, in this case, the baseboard management controller (BMC/DRAC) boards, which are controllable via IPMI [7], are configured to share the first Ethernet port. This means you might see *two* different MAC addresses and *two* different IPs on the `eth0` port; you want both facing internally.

Begin by hooking up the keyboard-video-mouse (KVM) switch to the admin node and powering on. Check BIOS settings, set a password, and choose the boot order. Then configure the BMC, setting an admin username and password and configuring an internal IP address for the device.

Everything should be redundant on the admin/storage and head/login nodes, including the root drives, which will hold

the OS. Configure a RAID 1 mirror for the internal root drive through the RAID BIOS.

At this point, you also need to configure any large external data storage enclosures. Our lab has 10TB, which we configure as RAID 6 plus one hot spare. I recommend letting storage fully initialize before continuing. After full RAID initialization, power off the node and unplug the external RAID enclosure before the OS install because the OS install partitioner can have difficulty with very big disks.

Setting Up the Admin Node

The first step is to set up the admin node that will also act as a storage and Kickstart installation server for the cluster. To get started installing the admin node, insert the CentOS installation DVD into the system that will act as the admin node. Boot the installer and select mostly all defaults. The only steps that differ from the defaults are setting a bootloader password, configuring the two network interfaces, and switching SELinux to permissive mode. In this example, I will call the admin/storage node `fijistor` with internal IP 192.168.1.2. For the package selection, do not change the defaults. Note that the *cluster* packages in the CentOS anaconda installer do not refer to HPC clustering in this context.

Once the installer completes, shut down the system, hook the external RAID back up, remove the “boot from DVD” option from the BIOS and power on. Admire your fresh login screen. Note that you can accomplish most of the rest of this build via remote login through SSH. The only time you need to be physically in the server room is when you are messing with the BIOS and powering the nodes for the first time (Figure 3).

Because SSH is the main entry to this system, I need to lock down `sshd`. Install and test a personal SSH key for the root user, then limit admin node logins to admin group-only and key-only. You only need to add a few lines to the `sshd` configuration file `/etc/ssh/sshd_config`:

```
PermitRootLogin without-password
AllowGroups fijiaadmin
```

Then, add the `admin` group and restart `sshd`:

```
# groupadd fijiaadmin
# usermod -G fijiaadmin root
# service sshd restart
```

Generate an SSH key for the `fijistor` root account, which you can then provide to all compute nodes:

```
# ssh-keygen # just hit enter, no passphrase
# cp /root/.ssh/id_rsa.pub /root/.ssh/authorized_keys.fijistor
# chmod -R go- /root/.ssh
```

If you don't plan to access the admin node through the console GUI, you can disable the graphical login by commenting out the relevant line in the `/etc/inittab` file. This will free up some memory:

```
# Run xdm in runlevel 5
#x:5:respawn:/etc/X11/prefdm -nodaemon
```


ANYONE INTERESTED IN SAVING MONEY?

Looks like these guys are comfortable overpaying
for enterprise storage. Are You?

“Hewlett-Packard Co. agreed to buy 3Par Inc. for \$2.35 billion” — *Bloomberg.com*

“EMC to Buy Isilon Systems Inc. for \$2.25 Billion” — *Wall Street Journal*

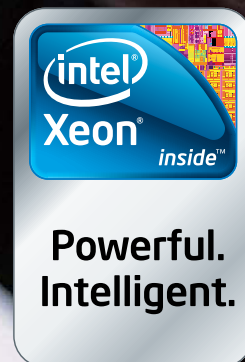
“Dell to Buy Compellent for \$960 Million” — *CNBC*

So what “benefit” will you see by this spending spree, other than higher costs?

The AberSAN Z-Series scalable unified storage platform, featuring the Intel® Xeon® processor 5600 series, brings the simplicity of network attached storage (NAS) to the SAN environment by utilizing the innovative ZFS file system. The AberSAN Z20 is easily found starting under \$20,000.

Who gives you the best bang for the buck?

	3Par InServ F200	Compellent Storage Center Series 30	Isilon NL-Series	Aberdeen AberSAN Z20
Storage Scale-Out	✓	✓	✓	✓
Thin Provisioning	✓	✓	✓	✓
HA Clustering	✓	✓	✓	✓
VMware® Ready Certified	✓	✓	✓	✓
Async / Synchronous Replication	✓	✓	✓	✓
iSCSI / Fibre Channel Target	✓	✓	iSCSI Only	✓
Unlimited Snapshots	X	✓	✓	✓
Native Unified Storage: NFS, CIFS	X	X	✓	✓
Virtualized SAN	X	X	X	✓
Deduplication	X	X	X	✓
Native File System	none	none	OneFS	ZFS 128-bit
RAID Level Support	5 and 6	5 and 6	Up to N+4	5, 6 and Z
Raw Array Capacity (max)	128TB	1280TB	2304TB	Unlimited
Warranty	3 Years	5 Years	3 Years	5 Years
Online Configurator with Pricing	Not Available	Not Available	Not Available	Available



Above specific configurations obtained from the respective websites on Feb. 1, 2011. Intel, Intel Logo, Intel Inside, Intel Inside Logo, Pentium, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. All trademarks are the property of their respective owners. All rights reserved. For terms and conditions, please see www.aberdeenninc.com/abpoly/abterms.htm. Ip007

888-297-7406
www.aberdeenninc.com/1p007



Now perform a first full-system package update – installing the development tools – and reboot into the new kernel:

```
# yum -y update
# yum -y install @development-tools
# reboot
```

It is important to keep system logs, so don't forget to extend the retention time for logfiles by changing a few settings in the `/etc/logrotate.conf` file:

```
# keep 4 weeks worth of backlogs
#rotate 4
rotate 999
```

Enable network forwarding on the admin node, so that all the exec nodes can resolve the public services when needed. To set up IP forwarding on the admin node, add the following to the `/etc/sysctl.conf` file:

```
net.ipv4.ip_forward = 1
```

The next step is to tweak the iptables firewall settings, or nothing will get through the firewall. In this case, the speed at which SSH may be attempted is limited, to curtail SSH brute forcing. All traffic from the internal cluster network is allowed. Dropped packets are logged for debug purposes – a setting you can always comment out and disable if you need to later. Finally, all outbound traffic from the internal network is put

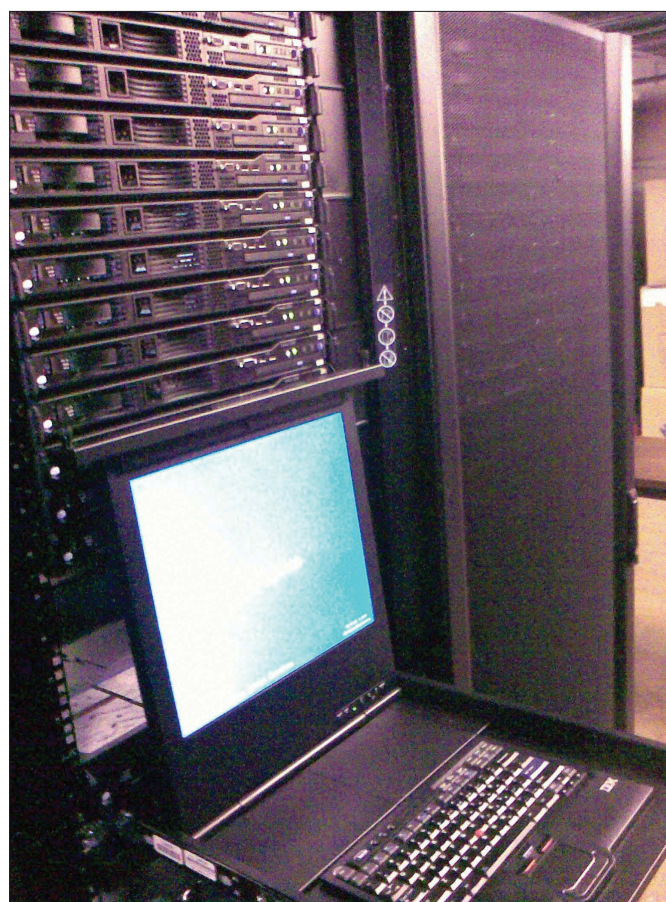


Figure 3: Good systems programming minimizes the time spent at the KVM.

through NAT. See Listing 1 for some critical settings for the `/etc/sysconfig/iptables` file.

Now reload the kernel `sysctl` parameters and restart the firewall:

LISTING 1: `/etc/sysconfig/iptables`

```
01 *filter
02 :INPUT ACCEPT [0:0]
03 :FORWARD ACCEPT [0:0]
04 :OUTPUT ACCEPT [0:0]
05 :RH-Firewall-1-INPUT - [0:0]
06 -A INPUT -j RH-Firewall-1-INPUT
07 -A FORWARD -j RH-Firewall-1-INPUT
08 -A RH-Firewall-1-INPUT -i lo -j ACCEPT
09 -A RH-Firewall-1-INPUT -p icmp --icmp-type any -j ACCEPT
10 -A RH-Firewall-1-INPUT -p 50 -j ACCEPT
11 -A RH-Firewall-1-INPUT -p 51 -j ACCEPT
12 -A RH-Firewall-1-INPUT -p udp --dport 5353 -d 224.0.0.251
    -j ACCEPT
13 #-A RH-Firewall-1-INPUT -p udp -m udp --dport 631
    -j ACCEPT
14 #-A RH-Firewall-1-INPUT -p tcp -m tcp --dport 631
    -j ACCEPT
15 -A RH-Firewall-1-INPUT -m state
    --state ESTABLISHED,RELATED -j ACCEPT
16 ### begin ssh
17 #-A RH-Firewall-1-INPUT -m state --state NEW -m tcp
    -p tcp --dport 22 -j ACCEPT
18 -A RH-Firewall-1-INPUT -p tcp --dport 22 --syn
    -s 192.168.1.0/255.255.255.0 -j ACCEPT
19 -A RH-Firewall-1-INPUT -p tcp --dport 22 --syn -m limit
    --limit 6/m --limit-burst 5 -j ACCEPT
20 ### end ssh
21 ### begin cluster
22 -A RH-Firewall-1-INPUT -i eth0
    -s 192.168.1.0/255.255.255.0 -j ACCEPT
23 -A RH-Firewall-1-INPUT -i eth0 -d 255.255.255.255
    -j ACCEPT
24 -A RH-Firewall-1-INPUT -i eth0 -p udp --dport 67:68
    -j ACCEPT
25 -A RH-Firewall-1-INPUT -i eth0 -p tcp --dport 67:68
    -j ACCEPT
26 ### end cluster
27 ### begin log
28 -A RH-Firewall-1-INPUT -m limit --limit 10/second -j LOG
29 ### end log
30 -A RH-Firewall-1-INPUT -j REJECT
    --reject-with icmp-host-prohibited
31 COMMIT
32 ### begin nat
33 *nat
34 :PREROUTING ACCEPT [0:0]
35 :POSTROUTING ACCEPT [0:0]
36 :OUTPUT ACCEPT [0:0]
37 -A POSTROUTING -o eth1 -j MASQUERADE
38 #-A PREROUTING -i eth1 -p tcp --dport 2222 -j DNAT
    --to 192.168.1.200:22
39 COMMIT
40 ### end nat
```



```
# service iptables restart
```

The next step in the process is to format the external 10TB RAID storage on `/dev/sdx`, obviously substituting in the correct device for `sdx`. If your configuration will not include any additional external storage, you can skip this step. Notice that you can overwrite the pre-existing partition information on the device and then use LVM directly (Listing 2). I also specify the metadata size to account for 128KB alignment. The `pvs` com-

LISTING 2: Configuring Partitions

```
01 # cat /proc/partitions
02 # dd if=/dev/urandom of=/dev/sdx bs=512 count=64
03 # pvcreate --metadatasize 250k /dev/sdx
04 # pvs -o pe_start
05 # vgcreate RaidVolGroup00 /dev/sdx
06 # lvcreate --extents 100%VG
    --name RaidLogVol100 RaidVolGroup00
07 # mkfs -t ext3 -E stride=32 -m 0 -O
    dir_index,filetype,has_journal,
    sparse_super /dev/RaidVolGroup00/RaidLogVol100
08 # echo "/dev/RaidVolGroup00/RaidLogVol100 /data0 ext3
    noatime 0 0" >>/etc/fstab
09 # mkdir /data0 ; mount /data0 ; df -h
```

LISTING 3: Setting Up the Install System

```
01 # mkdir -p /data0/repo/CentOS/5.6/iso/x86_64
02 # ln -s /data0/repo /repo
03 # cd /repo/CentOS
04 # ln -s 5.6 5
05 # wget http://mirrors.gigenet.com/centos/
    RPM-GPG-KEY-CentOS-5
06 # cd /repo/CentOS/5.6/iso/x86_64
07 # cat /dev/dvd > CentOS-5.6-x86_64-bin-DVD-1of2.iso
08 # wget http://mirror.nic.uoregon.edu/centos/5.6/isos/
    x86_64/shalsum.txt
09 # shalsum -c shalsum.txt
10 # mount -o loop CentOS-5.6-x86_64-bin-DVD-1of2.iso /mnt
11 # mkdir -p /repo/CentOS/5.6/os/x86_64
12 # rsync -avP /mnt/CentOS /mnt/repodata
    /repo/CentOS/5.6/os/
13 # mkdir -p /repo/CentOS/5.6/updates/x86_64
14 # rsync --exclude='debug' --exclude='*debug info*'
    --exclude='repoview' \
15 --exclude='headers' -irtCO --delete-excluded --delete \
16 rsync://rsync.gtlib.gatech.edu/centos/5.6/updates/x86_64
    /repo/CentOS/5.6/updates/
17 # mkdir -p /repo/epel/5/x86_64
18 # cd /repo/epel
19 # wget http://download.fedora.redhat.com/pub/epel/
    RPM-GPG-KEY-EPEL
20 # rsync --exclude='debug' --exclude='*debuginfo*'
    --exclude='repoview' --exclude='headers' \
21 -irtCO --delete-excluded
    --delete rsync://archive.linux.duke.edu/
    fedora-epel/5/x86_64 /repo/epel/5/
22 # mv /etc/yum.repos.d/* /usr/src/
23 # cat /dev/null >/etc/yum.repos.d/CentOS-Base.repo
24 # cat /dev/null >/etc/yum.repos.d/CentOS-Media.repo
```

mand will reveal that requesting 250KB actually results in a 256KB setting.

Creating a File Repository

With storage ready, it is time to build a file repository, which will feed all system Yum updates and Kickstart-based node installs. Create a directory structure with the base installer packages from the CentOS install DVD and then pull all the latest updates from a local rsync file mirror (Listing 3). I strongly urge you to find your own local mirror [8] that provides rsync. The commands in Listing 3 also pull the EPEL repository [9] for some additional packages.

Now create a repo configuration for yum in the `/etc/yum.repos.d/fiji.repo` file (Listing 4). Then create a script to update the file repo (Listing 5):

Now that I have a file repository for all RPMs, I can share it out via NFS, along with the `/data0` and `/usr/global/` directories. On this cluster, user home directories happen to be in the

LISTING 4: /etc/yum.repos.d/fiji.repo

```
01 ### CentOS base from installation media
02 [base]
03 name=CentOS-$releasever - Base
04 #mirrorlist=http://mirrorlist.centos.org/?release=$releas
    ever&arch=$basearch&repo=os
05 #baseurl=http://mirror.centos.org/centos/$releasever/
    os/$basearch/
06 baseurl=file:///repo/CentOS/$releasever/os/$basearch/
07 gpgcheck=1
08 gpgkey=http://mirror.centos.org/centos/
    RPM-GPG-KEY-CentOS-5
09 protect=1
10 ### CentOS updates via rsync mirror
11 [update]
12 name=CentOS-$releasever - Updates
13 #mirrorlist=http://mirrorlist.centos.org/?release=$releas
    ever&arch=$basearch&repo=updates
14 #baseurl=http://mirror.centos.org/centos/$releasever/
    updates/$basearch/
15 baseurl=file:///repo/CentOS/$releasever/
    updates/$basearch/
16 gpgcheck=1
17 gpgkey=http://mirror.centos.org/centos/
    RPM-GPG-KEY-CentOS-5
18 protect=1
19 ### Extra Packages for Enterprise Linux (EPEL)
20 [epel]
21 name=Extra Packages for Enterprise Linux 5 - $basearch
22 #baseurl=http://download.fedora.redhat.com/pub/
    epel/5/$basearch
23 #mirrorlist=http://mirrors.fedoraproject.org/mirrorlist?
    repo=epel-5&arch=$basearch
24 baseurl=file:///repo/epel/$releasever/$basearch
25 enabled=1
26 protect=0
27 failovermethod=priority
28 gpgcheck=1
29 gpgkey=http://download.fedora.redhat.com/pub/epel/
    RPM-GPG-KEY-EPEL
```



/data0/home/ directory. I will use /usr/global/ just like /usr/local/ except it is mounted on all nodes via NFS. The script global.sh (Listing 6) is symlinked on all nodes to /etc/profile.d/global.sh to set user environment variables.

The NFS /etc/exports file describes which directories are exported to remote hosts (see Listing 7). When the configuration is complete, start the NFS server:

```
# chkconfig nfs on
# service nfs start
```

Enable remote syslog logging from nodes by adding the following options in the /etc/sysconfig/syslog file:

```
SYSLOGD_OPTIONS="-m 0 -r -s fiji.baz.edu"
```

Then restart syslog:

```
# service syslog restart
```

All node hostnames and IPs for the cluster should be listed in the /etc/hosts file (Listing 8). Don't forget to enable network time synchronization for consistent logging:

LISTING 5: /root/bin/update_repo.sh

```
01 #!/bin/bash
02 OPTS='-vrtCO --delete --delete-excluded --exclude=i386*
--exclude=debug \
03 --exclude=*debuginfo* --exclude=repoview
--exclude=headers'
04 rsync $OPTS rsync://rsync.gtlib.gatech.edu/centos/5.6/
updates/x86_64 /repo/CentOS/5.6/updates/
05 rsync $OPTS rsync://archive.linux.duke.edu/fedora-epel/5/
x86_64 /repo/epel/5/
```

LISTING 6: /usr/global/etc/profile.d/global.sh

```
01 # Grid Engine
02 export SGE_ROOT=/usr/global/sge
03 . /usr/global/sge/default/common/settings.sh
04 alias rsh='ssh'
05 alias qstat='qstat -u "*"
06
07 # Intel compilers
08 . /usr/global/intel/Compiler/11.1/064/bin/iccvars.sh
intel64
09 . /usr/global/intel/Compiler/11.1/064/bin/fortvars.sh
intel64
10 . /usr/global/intel/Compiler/11.1/064/mkl/tools/
environment/mklvars64.sh
11 export INTEL_LICENSE_FILE=/usr/global/intel/
licenses:$INTEL_LICENSE_FILE
```

LISTING 7: /etc/exports

```
01 /data0 fiji(rw,async,no_root_squash) 192.168.1.0
/255.255.255.0(rw,async,no_root_squash)
02 /usr/global 192.168.1.0/255.255.255.0(
rw,async,no_root_squash)
03 /kickstart 192.168.1.0/255.255.255.0(ro)
04 /repo 192.168.1.0/255.255.255.0(ro)
```

```
# ntpdate -u -b -s 1.centos.pool.ntp.org
# hwclock --utc --systohc
# chkconfig ntpd on ; service ntpd start
```

Starting Kickstart

The next step is to install all the services needed for a Kickstart installation server (Listing 9) and build a file structure for network booting via PXE, TFTP, and NFS. This step will allow a new compute node to boot via the network in a manner similar to booting from a CD or DVD.

First, edit the PXE menu in the /tftpboot/pxelinux.cfg/default file (Listing 10). Next, generate your own password to replace the stub in Listing 10. SHA-1 encrypted passwords start with \$4\$:

```
# shalpass password
$4$gS+7mITP$y3s1L4Z+5Udp2v1ZHChNXd8lhAg$
```

Copy in all the files that will sync to all nodes:

```
# cp /etc/hosts /kickstart/fiji/etc/
# cp /root/.ssh/id_rsa.pub /kickstart/fiji/authorized_keys
# cp /etc/yum.repos.d/*.repo /kickstart/fiji/etc/yum.repos.d/
```

LISTING 8: /etc/hosts

```
01 127.0.0.1 localhost.localdomain localhost
02 ::1 localhost6.localdomain6 localhost6
03 192.168.1.1 fiji.baz.edu fiji
04 192.168.1.2 fijiistor.baz.edu fijiistor
05 192.168.100.9 ib
06 192.168.1.101 node01
07 192.168.1.102 node02
08 192.168.1.103 node03
09 192.168.1.104 node04
10 192.168.1.200 fiji-bmc
11 192.168.1.201 node01-bmc
12 192.168.1.202 node02-bmc
13 192.168.1.203 node03-bmc
14 192.168.1.204 node04-bmc
```

LISTING 9: Setting Up Kickstart

```
01 # yum install dhcp xinetd tftp tftp-server syslinux
02 # mkdir -p /usr/global/tftpboot ;
ln -s /usr/global/tftpboot /tftpboot
03 # mkdir -p /tftpboot/pxelinux.cfg
/tftpboot/images/centos/x86_64/5.6
04 # cd /tftpboot/images/centos/x86_64/ ; ln -s 5.6 5
05 # rsync -avP /mnt/isolinux/initrd.img
/mnt/isolinux/vmlinuz
/tftpboot/images/centos/x86_64/5.6/
06 # cd /usr/lib/syslinux
07 # rsync -avP chain.c32 mboot.c32 memdisk menu.c32
pxelinux.0 /tftpboot/
08 # mkdir -p /usr/global/kickstart ;
ln -s /usr/global/kickstart /kickstart ; cd /kickstart
09 # mkdir -p /kickstart/fiji/etc ; cd /kickstart/fiji/etc
10 # mkdir -p rc.d/init.d profile.d ssh yum/pluginconf.d
yum.repos.d
11 # touch rescue.cfg ks-fiji.cfg ; ln -s ks-fiji.cfg ks.cfg
```



COLLABORATION
EDUCATION



OPPORTUNITY
VANCOUVER



CANADA

AUGUST 17-19, 2011



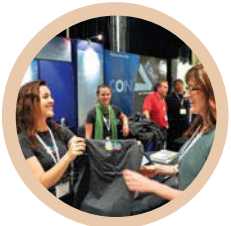
LINUXCON

PRAGUE



CZECH REPUBLIC

OCTOBER 26-28, 2011



THE PREMIERE

LINUX CONFERENCE



HOSTED BY



LEARN MORE AT EVENTS.LINUXFOUNDATION.ORG

Then you can proceed to the next step in the process, which is to edit the `/kickstart/fiji/etc/ntp.conf` file and change the server to the admin node:

```
#server 0.centos.pool.ntp.org
#server 1.centos.pool.ntp.org
#server 2.centos.pool.ntp.org
server 192.168.1.2
```

When you get to booting the compute nodes, you will be able to paste each MAC addresses into the `/etc/dhcpd.conf` file (Listing 11) and then restart the `dhcpd` process. The node will then boot off the network with the configured IP address. Booting each of the compute nodes one at a time and pasting each address into the `/etc/dhcpd.conf` file is a bit of a tedious process, which surely could be automated with more advanced management techniques.

Set the disabled option in `/etc/xinetd.d/tftp` to no (as shown in Listing 12). Now load the newly configured services:

```
# service xinetd restart
# chkconfig dhcpd on
# service dhcpd restart
```

With a complete PXE file structure in place, I can move on to the major step of preparing Kickstart configuration files. Kickstart configuration files are flat text files that specify all of the options the installer can accept, along with all of the pre- and post-install scripting you care to add. Listing 13 is a very sim-

LISTING 10: `/tftpboot/pxelinux.cfg/default`

```
01 DEFAULT menu.c32
02 PROMPT 0
03 TIMEOUT 100
04 ONTIMEOUT local
05 NOESCAPE 1
06 ALLOWOPTIONS 0
07 MENU TITLE Fiji Cluster PXE Menu
08
09 LABEL local
10 MENU LABEL Boot local hard drive
11 LOCALBOOT 0
12
13 LABEL centos
14 MENU LABEL CentOS 5 Fiji Node Install
15 KERNEL images/centos/x86_64/5/vmlinuz
16 APPEND ks=nfs:192.168.1.2:/kickstart/ks.cfg
    initrd=images/centos/x86_64/5/initrd.img
    ramdisk_size=100000 ksdevice=eth0 ip=dhcp
17
18 LABEL rescue
19 MENU PASSWD $4$XXXXXX
20 MENU LABEL CentOS 5 Rescue
21 KERNEL images/centos/x86_64/5/vmlinuz
22 APPEND initrd=images/centos/x86_64/5/initrd.img
    ramdisk_size=10000 text
    rescue ks=nfs:192.168.1.2:/kickstart/rescue.cfg
```

ple Kickstart config file that only boots into the rescue mode of the installer. This file is good to have ready to debug the inevitable failed node.

A full compute node Kickstart file `/kickstart/ks.cfg` is available for download at the *Linux Magazine* website [10]. I recommend pasting in this beginner kickstart configuration file and continually test installing the first compute node. You can edit the Kickstart file again and again, making small adjustments with each step, then completely re-install the first node until the installation is perfected. After the first node is correct, move on to power up each additional node in the cluster.

Don't forget to generate and replace the dummy password stubs with your own password. Encrypted passwords in MD5 format start with `1`:

LISTING 11: `/etc/dhcpd.conf`

```
01 ddns-update-style interim;
02 ignore client-updates;
03 option option-128 code 128 = string;
04 option option-129 code 129 = text;
05
06 subnet 192.168.1.0 netmask 255.255.255.0 {
07 option routers                192.168.1.2;
08 option subnet-mask            255.255.255.0;
09 option nis-domain              "fiji.baz.edu";
10 option domain-name            "fiji.baz.edu";
11 option domain-name-servers    123.123.123.123;
12 option time-offset            -18000; # Eastern
13 option ntp-servers            192.168.1.2;
14 default-lease-time 21600;
15 max-lease-time 43200;
16
17 allow booting;
18 allow bootp;
19 next-server 192.168.1.2;
20 filename "/pxelinux.0";
21
22 host node01 {
23 hardware ethernet 00:11:22:33:44:a0;
24 fixed-address 192.168.1.101;
25 }
26
27 host node01-bmc {
28 hardware ethernet 00:11:22:33:44:8d;
29 fixed-address 192.168.1.201;
30 }
31
32 host node02 {
33 hardware ethernet 00:11:22:33:44:83;
34 fixed-address 192.168.1.102;
35 }
36
37 host node02-bmc {
38 hardware ethernet 00:11:22:33:44:61;
39 fixed-address 192.168.1.202;
40 }
```

```
$ grub-md5-crypt
Password:
Retype password:
$1$d6oPa/$iUemCR50qSyvGSVtX9NrX1
```

With the admin node installed and the Kickstart configuration set up and ready for action, the next step is to install and configure all compute nodes that will execute the jobs submitted to the cluster.

Setting Up the Compute Nodes

Once the whole infrastructure is in place, you are ready to start installing compute nodes. The exact steps will vary depending on your configuration, but the key is to be methodical and minimize the time spent at each node.

Please note that the head/login node, having hostname *fiji*, can be installed manually or automated in the same manner as the compute nodes, with the additional step of configuring *eth1* as a public-facing network interface. A sample compute node installation procedure might include the following steps:

1. Cable the KVM switch and hit the power button.
2. F12 for PXE boot.
3. Ctrl + S to get the system Ethernet MAC address.
4. Ctrl + E to get the BMC Ethernet MAC address
5. Set LAN parameters: IPv4 IP Address Source: DHCP
6. Set LAN user configuration: enter and confirm password
7. Hit Esc; save changes and exit.
8. Add MAC addresses to *fijistor's* */etc/dhcpd.conf*, then enter *service dhcpd restart*.
9. Hit Enter on Fiji Cluster Node Install PXE menu.
10. Wait for a package dependency check.
11. Go to the next node.

LISTING 12: */etc/xinetd.d/tftpd*

```
01 service tftpd
02 {
03 socket_type          = dgram
04 protocol             = udp
05 wait                = yes
06 user                 = root
07 server               = /usr/sbin/in.tftpd
08 server_args          = -s -v /tftpboot
09 disable              = no
10 per_source           = 11
11 cps                  = 100 2
12 flags                = IPv4
13 }
```

LISTING 13: */kickstart/rescue.cfg*

```
01 lang en_US
02 keyboard us
03 mouse none
04 nfs --server=192.168.1.2
    --dir=/repo/CentOS/5/iso/x86_64
05 network --bootproto=dhcp
```

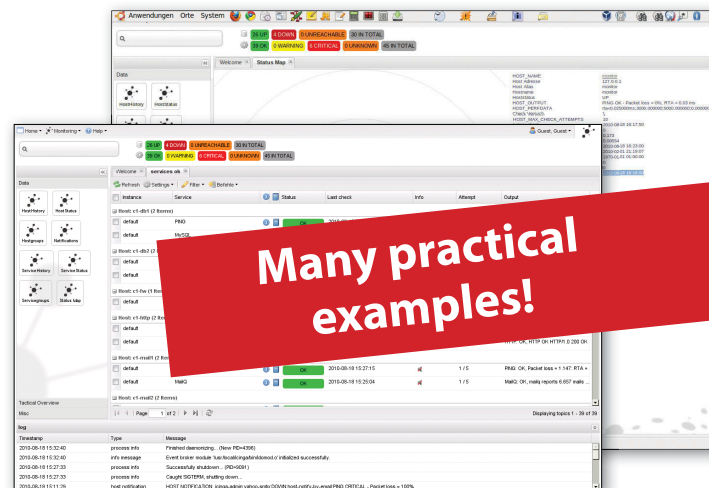
Online Training at Linux Magazine Academy

Monitoring with Nagios

Getting your IT under control the easy way:



- Web front-end
- Monitoring Windows/Linux/Unix
- Structuring the configuration
- Monitoring SNMP-components
- Nagvis, Grapher V2, and ND02DB add-ons



Many practical examples!

You might want to use a management tool like IPMI to control the DRAC/BMC, remotely power-cycling locked nodes:

```
# yum -y install OpenIPMI-tools
# ipmitool -H 192.168.1.2XX -U root -P PASSWORD -I lanplus chassis status
# ipmitool -H 192.168.1.2XX -U root -P PASSWORD -I lanplus chassis power cycle
```

With a cluster of Linux servers, it is desirable to run the same commands across all the compute nodes. You can run these common commands with the use of a simple for loop script, or you can use a cluster-enabled shell like Dancer's Shell [11] or ClusterSSH [12]. To execute the commands, create a file called `/etc/machines.list` comprising a list of all nodes, one per line, then use the simple `/root/bin/ssh_loop.sh` script (Listing 14). You must keep user authentication in sync across all nodes. Although you have better ways to do this, for the sake of brevity, I will do a simple sync of auth files from the head/

LISTING 14: /root/bin/ssh_loop.sh

```
01 #!/bin/sh
02 for I in `grep -v "\#" /etc/machines.list`; do
03 echo -n "$I " ; ssh $I "$@"
04 done
```

LISTING 15: /root/bin/update_etc.sh

```
01 rsync -a fiji:/etc/passwd /root/etc/
02 rsync -a fiji:/etc/shadow /root/etc/
03 rsync -a fiji:/etc/group /root/etc/
04 rsync -a /root/etc/passwd /root/etc/shadow
   /root/etc/group /etc/
05 rsync -a /root/etc/passwd /root/etc/shadow
   /root/etc/group /kickstart/fiji/etc/
06 for NODE in `cat /etc/machines.list | grep -v "\#"`; do
   rsync -a /kickstart/fiji/etc $NODE:/
done
```

LISTING 16: Setting up Son of Grid Engine

```
01 # mkdir /usr/global/sge-6.2u5_PLUS_3-26-11
02 # ln -s /usr/global/sge-6.2u5_PLUS_3-26-11 /usr/global/sge
03 # export SGE_ROOT=/usr/global/sge
04 # cd /usr/global/sge
05 # adduser -u 186 sgeadmin
06 # mkdir src ; cd src
07 # wget http://bioteam.net/dag/
   gridengine-courtesy-binaries/sge-6.2u5_
   PLUS_3-26-11-common.tar.gz
08 # wget http://bioteam.net/dag/
   gridengine-courtesy-binaries/sge-6.2u5_
   PLUS_3-26-11-bin-lx26-amd64.tar.gz
09 # cd /usr/global/sge
10 # tar xzvf src/sge-6.2u5_PLUS_3-26-11-common.tar.gz
11 # tar xzvf src/
   sge-6.2u5_PLUS_3-26-11-bin-lx26-amd64.tar.gz
12 # chown -R sgeadmin.sgeadmin .
```

login node to all other nodes with the script `/root/bin/update_etc.sh` (Listing 15).

Getting on the Grid

The final step is to enable a job queue to manage the workload. I will use Son of Grid Engine [13] from the BioTeam cour-

LISTING 17: Example SGE cluster.conf

```
01 SGE_ROOT="/usr/global/sge"
02 SGE_QMASTER_PORT="6444"
03 SGE_EXECD_PORT="6445"
04 SGE_ENABLE_SMP="false"
05 SGE_ENABLE_ST="true"
06 SGE_CLUSTER_NAME="p6444"
07 SGE_JMX_PORT="6666"
08 SGE_JMX_SSL="false"
09 SGE_JMX_SSL_CLIENT="false"
10 SGE_JMX_SSL_KEYSTORE="/tmp"
11 SGE_JMX_SSL_KEYSTORE_PW="/tmp"
12 SGE_JVM_LIB_PATH="/tmp"
13 SGE_ADDITIONAL_JVM_ARGS="-Xmx256m"
14 CELL_NAME="default"
15 ADMIN_USER="sgeadmin"
16 QMASTER_SPOOL_DIR="/usr/global/sge/default/spool/qmaster"
17 EXECD_SPOOL_DIR="/usr/global/sge/default/spool"
18 GID_RANGE="20000-20100"
19 SPOOLING_METHOD="classic"
20 DB_SPOOLING_SERVER=""
21 DB_SPOOLING_DIR="spooldb"
22 PAR_EXECD_INST_COUNT="20"
23 ADMIN_HOST_LIST="fijistor"
24 SUBMIT_HOST_LIST="fiji"
25 EXEC_HOST_LIST="node01 node02 node03 node04"
26 EXECD_SPOOL_DIR_LOCAL=""
27 HOSTNAME_RESOLVING="true"
28 SHELL_NAME="ssh"
29 COPY_COMMAND="scp"
30 DEFAULT_DOMAIN="none"
31 ADMIN_MAIL="none"
32 ADD_TO_RC="true"
33 SET_FILE_PERMS="true"
34 RESCHEDULE_JOBS="wait"
35 SCHEDD_CONF="1"
36 SHADOW_HOST=""
37 EXEC_HOST_LIST_RM=""
38 REMOVE_RC="true"
39 WINDOWS_SUPPORT="false"
40 WIN_ADMIN_NAME="Administrator"
41 WIN_DOMAIN_ACCESS="false"
42 CSP_RECREATE="true"
43 CSP_COPY_CERTS="false"
44 CSP_COUNTRY_CODE="DE"
45 CSP_STATE="Germany"
46 CSP_LOCATION="Building"
47 CSP_ORGA="Organisation"
48 CSP_ORGA_UNIT="Organisation_unit"
49 CSP_MAIL_ADDRESS="name@yourdomain.com"
```


Save the Date!

LISA'11

December 4–9, 2011, Boston, MA

25TH LARGE INSTALLATION
SYSTEM ADMINISTRATION
CONFERENCE

SPONSORED BY

usenix

IN COOPERATION
WITH LOPSA and SNIA



Come to LISA '11 for **training** and **face time** with experts in the **sysadmin** community.

The theme for LISA '11 is "DevOps: New Challenges, Proven Values."

LISA '11 will feature:

6 days of training on topics including:

- Virtualization
- Security
- Configuration management
- And more!

Plus a 3-day Technical Program:

- Invited Talks
- Paper Presentations
- Guru Is In Sessions
- Practice and Experience Reports
- Vendor Exhibition
- Workshops
- Posters and WiPs

Find out more at www.usenix.org/lisa11/lp

tesy binaries [14]. Son of Grid Engine is a community-based project that evolved from Sun's Grid Engine project. When Oracle bought Sun and phased out the free version of Grid Engine, the Son of Grid Engine developers stepped in to maintain a free version. For more information on working with Son of Grid Engine, see the how-to documents at the project website [15].

Remember that the user home directories and `/usr/global/` are exported to the entire cluster via NFS. Users will have SSH keys without passwords for remote shell commands from the head node to all compute nodes.

To get started with Son of Grid Engine, follow the steps in Listing 16.

You can interactively install the master host on the admin node, with all settings per default except the following: install as `sgeadmin` user, set network ports with environment, `sge_qmaster` port 6444, `sge_execd` port 6445, cell name default, cluster name `p6444`, say no to `pkgadd` and yes to verify permissions, say no to JMX MBean server, select classic spooling method, GID range 20000-20100, enter each hostname, and set no `shadow` host.

```
# cd /usr/global/sge/  
# ./install_qmaster
```

Then we set the user environment:

```
./usr/global/sge/default/common/settings.sh
```

The SGE logs are found at: `/usr/global/sge/default/spool/qmaster/messages`, `/tmp/qmaster_messages` (during `qmaster` startup), `/usr/global/sge/default/spool/HOSTNAME/messages`, and `/tmp/execd_messages` (during `execd` startup).

You need to install SGE on all compute nodes. This step is most easily done with an automated `install` config. Listing 17 shows an example of an `SGE cluster.conf` file, stripped of comments.

To set up an automated installation:

```
# cd /usr/global/sge/  
# cp util/install_modules/inst_template.conf cluster.conf  
# vim cluster.conf  
# ./inst_sge -x -auto /usr/global/sge/cluster.conf  
automated install log  
sge_root/spool/install_HOSTNAME_TIMESTAMP.log
```

After a successful install run, verify that Son of Grid Engine is working and check the configuration with:

```
$ ps ax | grep sge  
$ qconf -sconf
```

A good first step is to submit a simple job test to the cluster as a normal user:

```
# adduser testuser  
# update_etc.sh  
# su - testuser  
$ qsub /usr/global/sge/examples/jobs/simple.sh  
$ qstat
```

Conclusion

Now that you have a functional super computer cluster with job and resource management, what next? Should you install Blender [16] and start rendering a feature-length movie? Will you install MPI and explore parallel computing [17]? Maybe you can run the Linpack benchmark and compete in the Top 500 [18]? Well, that all depends on your chosen discipline, but I imagine that your group of researchers will have enough work to keep the system busy. ■■■

INFO

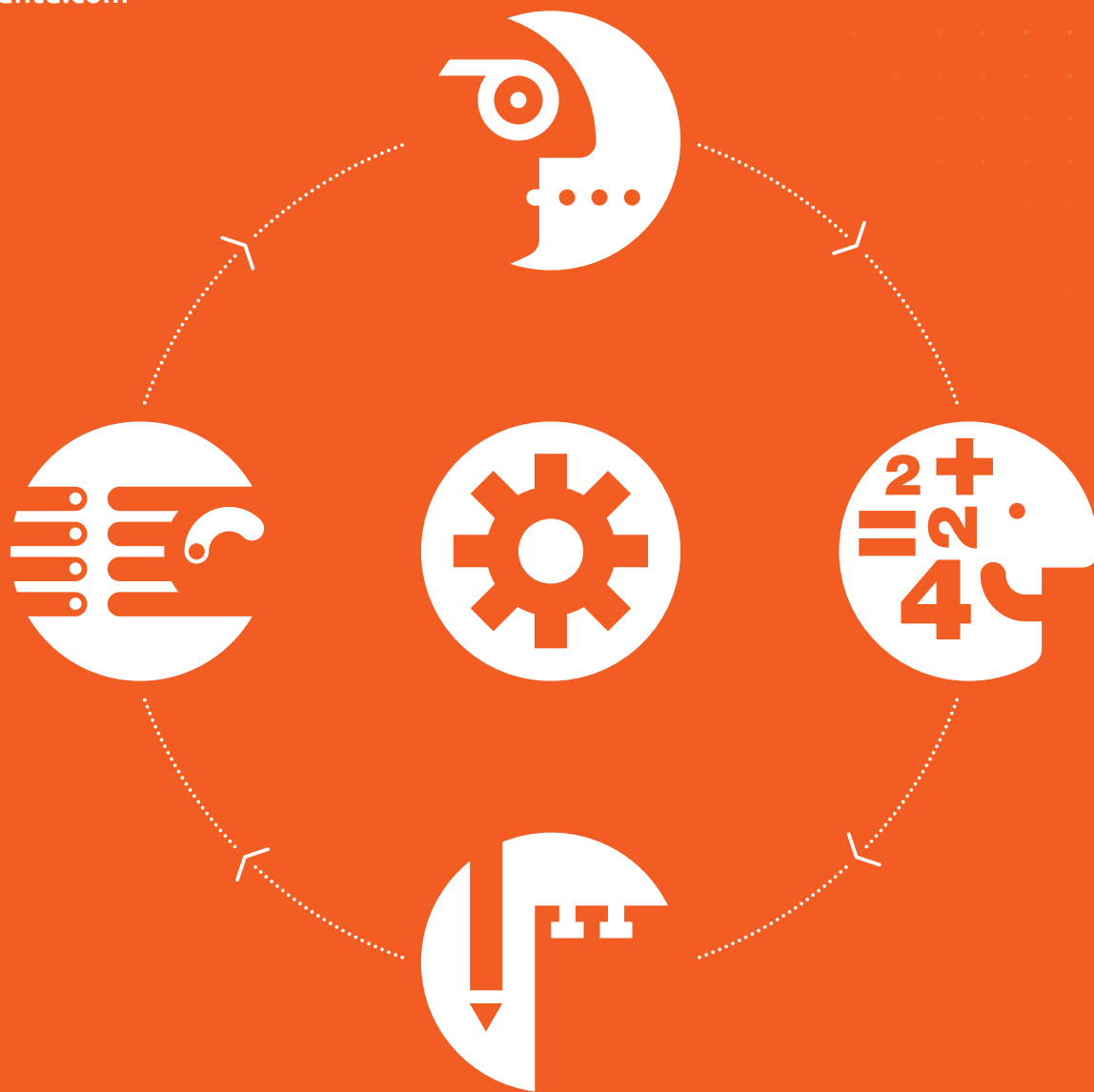
- [1] Red Hat Enterprise Linux documentation: http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/index.html
- [2] CentOS-5 documentation: <http://www.centos.org/docs/5/>
- [3] RHEL install guide: http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5/html/Installation_Guide/index.html
- [4] Kickstart installations: http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5/html/Installation_Guide/ch-kickstart2.html
- [5] CentOS ISO mirrors: http://isoredirect.centos.org/centos/5/isos/x86_64/
- [6] Preboot execution environment: http://en.wikipedia.org/wiki/Preboot_Execution_Environment
- [7] Intelligent platform management interface: http://en.wikipedia.org/wiki/Intelligent_Platform_Management_Interface
- [8] CentOS North American mirrors: <http://www.centos.org/modules/tinycontent/index.php?id=30>
- [9] Fedora EPEL public active mirrors: <http://mirrors.fedoraproject.org/publiclist/EPEL/>
- [10] Code for this article: <http://www.linux-magazine.com/Resources/Article-Code>
- [11] Dancer's shell: <http://www.netfort.gr.jp/~dancer/software/dsh.html.en>
- [12] ClusterSSH: <http://clusterssh.sourceforge.net/>
- [13] Son of Grid Engine: <https://arc.liv.ac.uk/trac/SGE>
- [14] BioTeam grid engine courtesy binaries: <http://bioteam.net/dag/gridengine-courtesy-binaries/>
- [15] Grid engine how-to: <http://arc.liv.ac.uk/SGE/howto/howto.html>
- [16] Blender: <http://www.blender.org/>
- [17] Introduction to parallel computing: https://computing.lln.gov/tutorials/parallel_comp/
- [18] The Linpack benchmark: <http://www.top500.org/project/linpack>



UBUNTU DEVELOPER SUMMIT

31 October - 4 November 2011

uds.ubuntu.com

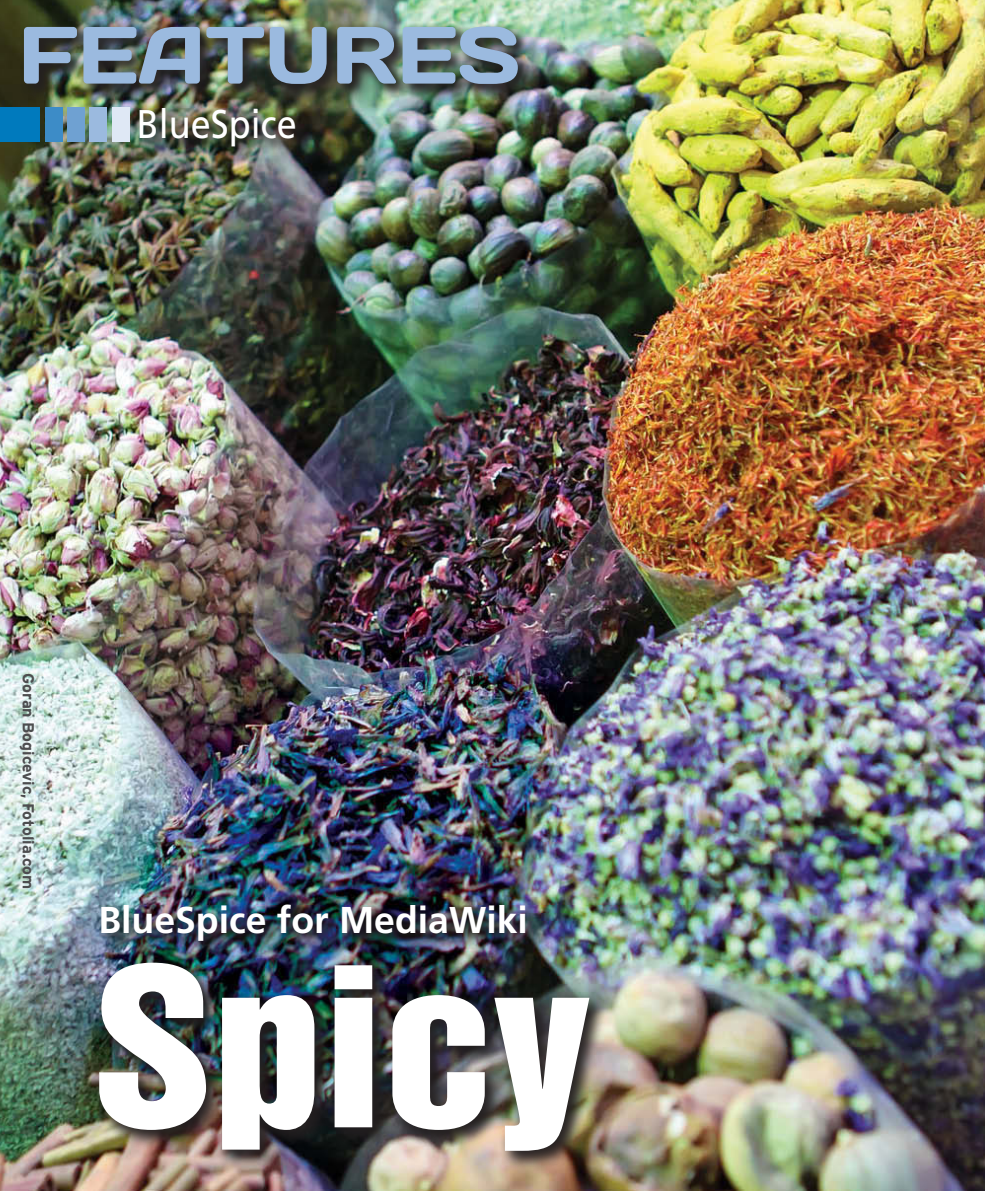


Join us at UDS Orlando, the event that brings together developers, engineers, Ubuntu community members and more. Concentrating on four primary focal points; **ideas**, **discussion**, **planning** and **collaboration**.

Be part of the discussion and planning around the next release of Ubuntu.

This is an opportunity not to be missed!





BlueSpice for MediaWiki

Spicy

A German company has prepared a tasty recipe for the wiki community. Take one slightly jaded MediaWiki, mix in a pinch of secret blue powder, add some fresh modules, and gently heat to the ideal Tomcat temperature – dinner is served! *By Mela Eckenfels*

MediaWiki [1] is a pioneering wiki engine. Everybody uses the software, everybody complains about it, but nobody can really do without it. Additionally, it has no serious competitors [2]. However, major corporations have very little patience with the quirks of this open source project. What they need is a state-of-the-art wiki with a state-of-the-art configuration interface – a kind of business MediaWiki for corporate users.

A company from Regensburg, Germany, by the name of Hallo Welt Medienwerkstatt GmbH [3] has set out to provide just this software. The stable version of BlueSpice for MediaWiki [4] was released July 4, 2011. Among other

things, it includes a WYSIWYG editor, with an improved search function, page templates, and PDF exports as additional features. But does the software, released under the GPL, really make life easier for wiki administrators and users?

Hunger

First, wikis facilitate the process of creating websites and make HTML and its cohorts more accessible to less technically accomplished people. On closer inspection, MediaWiki, with its powerful variables and includes and its inconsistent markup, actually makes life more difficult for many wiki users. Wiki engines only appear to abstract the design from the content of a page.

Second, wikis are perfect for large knowledge bases, but you need a good search function to find all of this know-how. If you look at the referenced installation, Wikipedia, you don't realize how bad the MediaWiki search function really is. But, if your wiki does not have thousands of industrious helpers prepared to push the articles into the right categories, the only real workaround is to use external search engines to find content in your own wiki.

These problems are just two that administrators and users of large wiki installations are confronted with every day. They hunger for better and more feature-rich wiki offerings. Hallo Welt comes to the rescue with a collection of utensils and the various ingredients required to put something new on the table. The company has been working on developing the open source wiki BlueSpice since 2007, and the project was originally initiated by IBM, who wanted to deploy MediaWiki but was unable to live with its downsides.

The Hallo Welt developers began by adapting the basics of Wikipedia to reflect corporate needs. After various corporations tested the business wiki for four years, the team finally dared to publish its results in 2011. A demo [5] of the latest release invites users to experiment, and the software itself can be downloaded from SourceForge [6]. Two versions are up for grabs: Basic and Fully Featured, but I'll come back to that subject later.

The Recipe

BlueSpice is a MediaWiki extension that bundles additional features the developers consider useful. The first thing you'll notice is the improved graphics. The sidebar on the left has a completely new look and feel (Figure 1), and besides the standard *Navigation* button, it now features two other items, *Focus* (your own pages and postings) and *Admin* (preferences).

On the right-hand side, you can click the asterisk to unfold a widget bar with important features. This is where you will find templates, other users, tools, and the PDF export feature if this is installed.

The developers chose the Tiny MCE [7] JavaScript editor libraries for their WYSIWYG editor. Media management

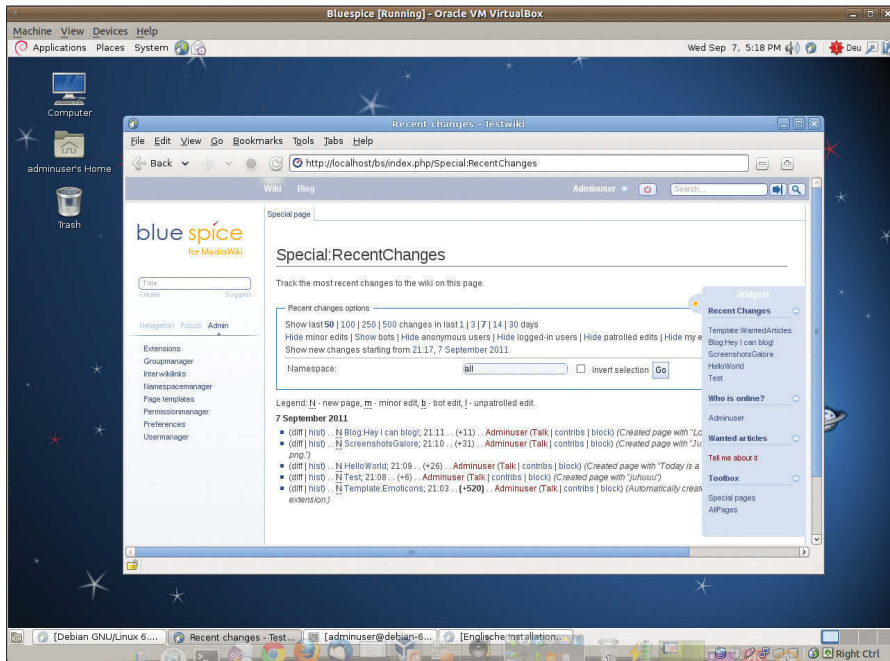


Figure 1: The improved BlueSpice interface is easy to navigate.

was also totally revamped and is now far more user friendly. Users are given a convenient approach to managing the modules grouped and recommended by BlueSpice (currently between 40 and 50 in number) in the user interface (*Admin | Add-ons*). Color coding of the development status also paints a clearer picture. Yellow modules are beta, whereas stable versions are blue.

The advanced search function is only available in the Fully Featured version (see the next section). BlueSpice relies on Apache Solr [8] for this, which is an open source search server that uses the Lucene search Java libraries [9] – also courtesy of the Apache project.

Two Flavors

As I mentioned previously, two versions of the business wiki are available for downloading. The Basic and Fully Featured variants differ in terms of ingredients. The former is designed for users with standard hosting packages, although you do need access to the command line, because BlueSpice currently relies on having a MediaWiki installation in place. But, according to Hallo Welt, a “no worries” package is planned.

Server side, the Basic version expects an up-to-date LAMP or WAMP system (Apache 2, MySQL 5, and PHP 5.2.x or 5.3.x.). PHP 5.3.1 will not work because of a bug; in fact, it won’t even support an unmodified MediaWiki. The addi-

tional PHP package requirements are gd, json, curl, tidy, dom, openssl, mcrypt, and soap.

As of this writing, BlueSpice will run on MediaWiki versions 1.15.x and 1.16.x. BlueSpice does not support the current 1.17.x version with its newly implemented JavaScript environment, although the developers are working on a solution.

MediaWiki installation from the distribution repositories caused some issues in my tests. For example, BlueSpice refused to cooperate with the current Ubuntu MediaWiki package. The best approach is to download the MediaWiki sources [10] and install them below the web server’s document root.

The Fully Featured version has the same requirements with respect to the installation environment. Additionally, Lucene search and the XHTML renderer, for the PDF export feature, require a Java application server. The developers recommend Apache Tomcat [11].

Lightly Salted

Good developers can install and configure software in their sleep. However, it’s far more difficult to explain a complex program to novices. The makers of BlueSpice currently have little experience with third-party installations outside of their own wiki walls. Thus, after unpacking the sources directly in the root directory of the MediaWiki installa-

tion, users will search in vain for an INSTALL or README file. All you get is the release notes and some licensing information – and neither of these files tells you anything about, or points the way to, the installation guide.

The blog [4] and the archive on SourceForge [6] do not provide anything useful either, at first sight. However, the latter does take you to a couple of installation guides in PDF format in the beta version directories. Finally, the tag cloud in the blog puts users on the right track and takes them to the documentation wiki [12], which really would benefit from a more prominent linking. Armed with the right guide, users can embark on the installation [13].

The first step is to open the `installcheck.php` file in a browser of your choice to see whether all the modules and PHP settings are in place with the right versions. The next step is to copy the `bluespice-skin` directory into the `skins` directory. Following the guide from the documentation wiki, you then only need to add a couple of lines to the `LocalSettings.php` file. This is easily done with copy and paste, but automating this step might be preferable from the user’s point of view.

The version history in the guide indicates that the vendor revises and extends the guide more or less on a daily basis. One error that escaped quality assurance, however, is the instruction telling you to open `bluespice-core/config/setup.php` – this file doesn’t exist and just returns a 404 error.

Instead, you need to run `bluespice-core/setup/index.php` to start the configuration (Figure 2). After doing so, add the MySQL credentials for the MediaWiki installation, complete the path in `Core::BlueSpiceScriptPath`, be sure to enter *en* for English next to the `Core::LanguageCode` variable, and keep all the other defaults. Then, you can follow the remaining steps in the guide to complete the setup and wait for the main wiki page to appear in its new BlueSpice clothes.

Tasty

BlueSpice doesn’t just wipe out the MediaWiki core but carefully extends it with a selection of flavorings. In other words, any users you created in MediaWiki will still be around in BlueSpice,

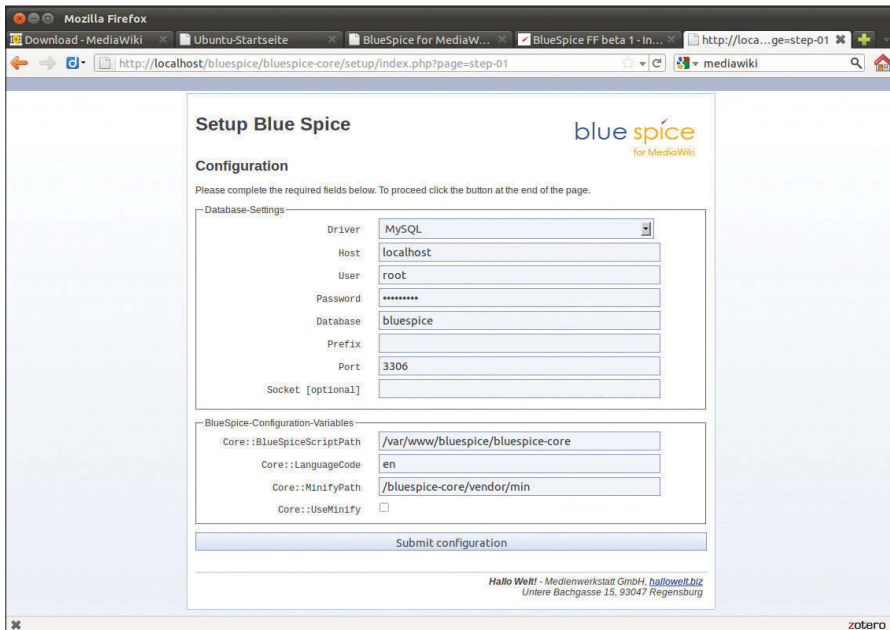


Figure 2: The BlueSpice configuration page.

and you can use your MediaWiki admin account to access the business wiki's administrative interface. The interface is clearly structured: The navigation bar on the left gives you everything you need to manage users, groups, and permissions in the *Admin* area. Here you also will find the settings for add-ons, interwiki links, the namespace manager, and page templates.

If you need proof of the pudding, check out how well BlueSpice's user management compares with its counterpart in an unmodified MediaWiki installation. Whereas the latter spreads tasks over multiple sub-pages, BlueSpice keeps most management tasks in a single interface (Figure 3). And, you can assign users to groups here, rather than needing to go to group management. Although separating items isn't necessarily a bad thing, hard-working admins will always prefer an interface that gives them as much access to as many features as possible at one time.

Users can use the general settings to configure both MediaWiki and BlueSpice. The business wiki doesn't achieve its full potential here and is not as complete as you might like. Annoying management tasks such as cleaning URLs are all manual in MediaWiki, and BlueSpice doesn't simplify them – at least not yet. On a brighter note, the product scores points when it comes to add-ons. Brief explanations save administrators guesswork when it comes to the

roles and capabilities of the individual modules.

Side Dishes

Interwiki links are abbreviations that the wiki software expands to full-fledged links. This feature removes the need for users to insert lengthy URLs into their articles. If the URL for a link changes, you can just modify the address in the interwiki list, and all the links automatically point to the new URL. In Wikipedia, [[en:Spice]] links to the English

version of an article. The en automatically expands to the online encyclopedia's English language namespace.

If multiple departments in a company work with their own wikis, or maintain a multilingual wiki, it is a good idea for the administrator to add intuitive abbreviations to the interwiki list. MediaWiki makes this task as difficult as possible: Administrators have to modify the list directly in the database. In BlueSpice, you can forget the SQL statements and handle the whole thing in the web interface.

Namespace management lets you divide the wiki into topical areas. For example, you could define a *Kansas*: namespace for an office in Kansas or adopt a project-related approach, as in *ProjectA*:, *ProjectB*:, and so on. Although a wiki can have only one *ToDo* page, namespaces let you maintain to-do lists in, say, *ProjectA:ToDo* and *ProjectB:ToDo*. Privilege management allows the administrator to restrict access of a specific group to specific areas of the wiki, making it possible to restrict write access for the *Trainees* group to the *Trainees*: namespace while still granting read access for the rest of the wiki. Unfortunately, the privilege management front end isn't exactly perfect; in fact, its cluttered appearance tends to put people off.

However, the WYSIWYG editor, which generates clean wiki markup, more than compensates for this drawback. The in-

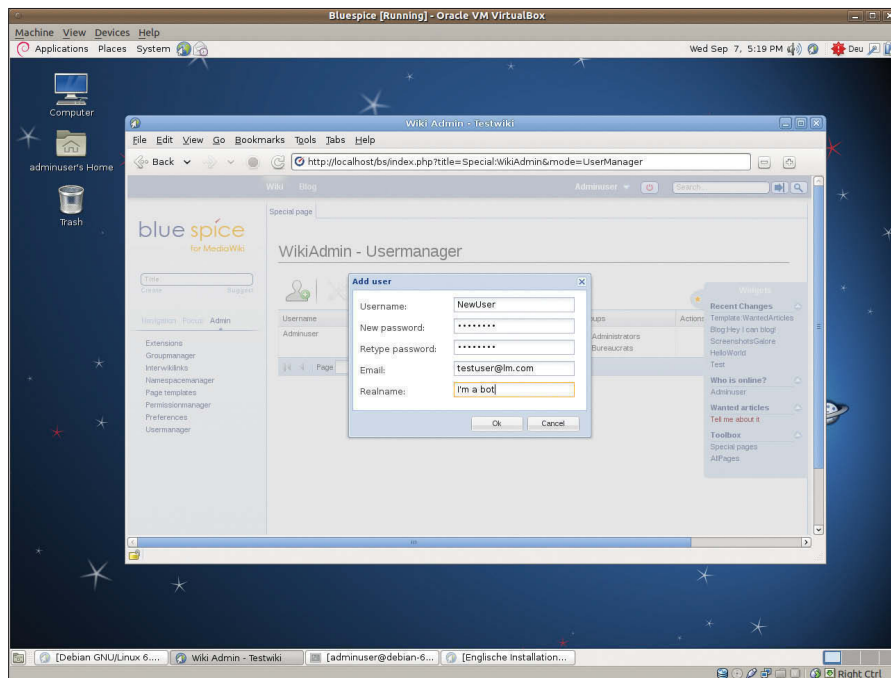


Figure 3: MediaWiki spreads individual user management tasks over multiple pages, whereas BlueSpice combines many tasks in a single dialog.

dividual buttons take users to formatting options and file uploads. Here, you can add images and other files to your article and assign categories and create links. The upload tool is impressive and practical compared with its MediaWiki counterpart, giving easy access to, and helping you search for, files you have already uploaded.


Three functions not a part of the MediaWiki standard ensure good flow of communication between wiki users. One is the widgets bar I referred to previously, which you can open and close by clicking on the asterisk and which shows the currently logged in users. A shout-box is available for chatting while you work and for adding comments, meaningful or not. For more important announcements, a *Blog* item in the menu at the top lets you integrate an online diary. BlueSpice is thus not just suitable for internal communications, it is also a platform for publicly accessible articles.

Easily Digestible

BlueSpice adds verve to the somewhat cumbersome MediaWiki interface and

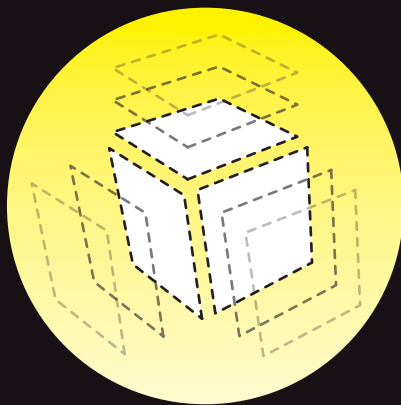
shows a number of approaches to making wiki management a more easily digestible process. Of course, some rough edges remain, and the Hallo Welt team could definitely add a touch more seasoning here.

Above all, the business wiki is excellent for documentation and communication within closed user groups. Even if other application scenarios are conceiv-

able, whether or not they make sense right now is a totally different question. The universality of BlueSpice could be given another boost if the developers remove more MediaWiki quirks and, above all, continue to work on the essential administrative interface. When this happens, the Hallo Welt chefs will be well on their way to earning their first stars. 

INFO

- [1] MediaWiki: <http://www.mediawiki.org/wiki/MediaWiki>
- [2] "Weaving the Web" by Mela Eckens, *Linux Magazine*, December 2010, pg. 38
- [3] Hallo Welt: <http://hallowelt.biz/en/home/>
- [4] BlueSpice for MediaWiki: <http://blog.blue-spice.org>
- [5] BlueSpice demo: <http://demo.blue-spice.org> (in German)
- [6] Download page for both BlueSpice versions: <http://sourceforge.net/projects/bluespice/files/>
- [7] Tiny MCE editor: <http://tinymce.moxiecode.com>
- [8] Apache Solr: <http://lucene.apache.org/solr>
- [9] Apache Lucene: <http://lucene.apache.org/java/docs>
- [10] MediaWiki download: <http://www.mediawiki.org/wiki/Download>
- [11] Apache Tomcat: <http://tomcat.apache.org>
- [12] BlueSpice helpdesk: <http://hilfe.blue-spice.org> (in German)
- [13] BlueSpice installation manual: http://hilfe.blue-spice.org/index.php/Englische_Installationsanleitung_Stable_1.0



Develop.



Deploy.



Scale.

Full root access on your own virtual server for as little as \$19.95/mo

Multiple Linux distributions to choose from • Web-based deployment • Five geographically diverse data centers • Dedicated IP address • Premium bandwidth providers • 4 core SMP Xen instances • Out of band console access • Private back-end network for clustering • IP fail-over support for high availability • Easily upgrade or add additional Linodes • Free managed DNS

For more information visit www.linode.com or call us at 609-593-7103



linode.com



Introduction to the Squirrel scripting language

Speedy Squirrel

If you are looking for a scripting language with a lean interpreter that is easy to integrate with your applications, your search will probably take you to Lua. But a trusted competitor by the name of Squirrel has been hiding out in Lua's shadow for many years. *By Tim Schürmann*

More than eight years ago, Alberto Demichelis was working for the popular German game developer Crytek and tackling his first game, Far Cry, the program logic of which was to be formulated in the Lua scripting language. When he tried to integrate the results, Demichelis discovered an issue with automated memory management (garbage collection). After trying some workarounds, he finally took the plunge and wrote his own scripting language. He wanted its interpreter to be like Lua's – small, lean, and easy to integrate with any application. The results were Squirrel [1], whose interpreter currently weighs in at 442KB.

Old Friends

Squirrel is an imperative, object-oriented scripting language whose syntax looks like a mix of C, C++, Java, JavaScript, and Python. Squirrel copied tables from Lua, using the flexible data structure for nearly everything. And, it has a garbage collection feature that autonomously handles memory management.

The bottom line is that Squirrel combines an easy and familiar syntax with the benefits of Lua. The reference imple-

mentation of the interpreters was placed under the MIT license in version 3.0. The source code is thus not just open source; the interpreter can be used free of charge for commercial projects.

Squirrel is currently mainly used for computer games, the best known commercial games being Left 4 Dead 2, Final Fantasy Crystal Chronicles: My Life as a King, and Portal 2 [2]. The developers of the open source Transport Tycoon Deluxe clone, OpenTTD, wrote the artificial intelligence of the computerized players in Squirrel (Figure 1).

If you use the integrated developer environment Code::Blocks, you also will have encountered Squirrel. Squirrel scripts enhance the functional scope [3] (Figure 2) and is thus also recommended for applications beyond gaming – as the Squirrel Shell goes to prove, you can even use it as a Bash replacement [4].

Manual Work

Unfortunately, Squirrel is not included in the repositories of most popular distributions. This typically gives users no alternative but to grab the reference implementation from the Squirrel website and build the interpreter themselves [1]. The source code archive is available in two

flavors: a standalone variant, which runs Squirrel scripts directly at the command line; and a static library, which you can tack onto your C++ programs. As a bonus, users get the Squirrel Standard Library with some useful Squirrel functions, a handful of sample programs, and a PDF language reference.

To create the interpreter, you need a C++ compiler and then you can run `make` in the directory created by unpacking the archive. This puts the standalone interpreter `sq` in the `bin` directory. You can then pass the Squirrel script you want to run to the interpreter:

```
sq myscript.nut
```

To accelerate execution of the script, the interpreter first converts it to bytecode. You can do this explicitly up front:



Figure 1: Computerized opponents in the OpenTTD use AI courtesy of Squirrel scripts.

Come with us, under the stars...



ballarat 2012
linux.conf.au
16 - 20 January 2012

Come out from the buildings, away from the lights and noise.

Come where the bush meets the city.

Come where history runs hand in hand with high technology.

Come with us, to a different Australia.

Come, join us under the stars.

Come to Ballarat 2012.

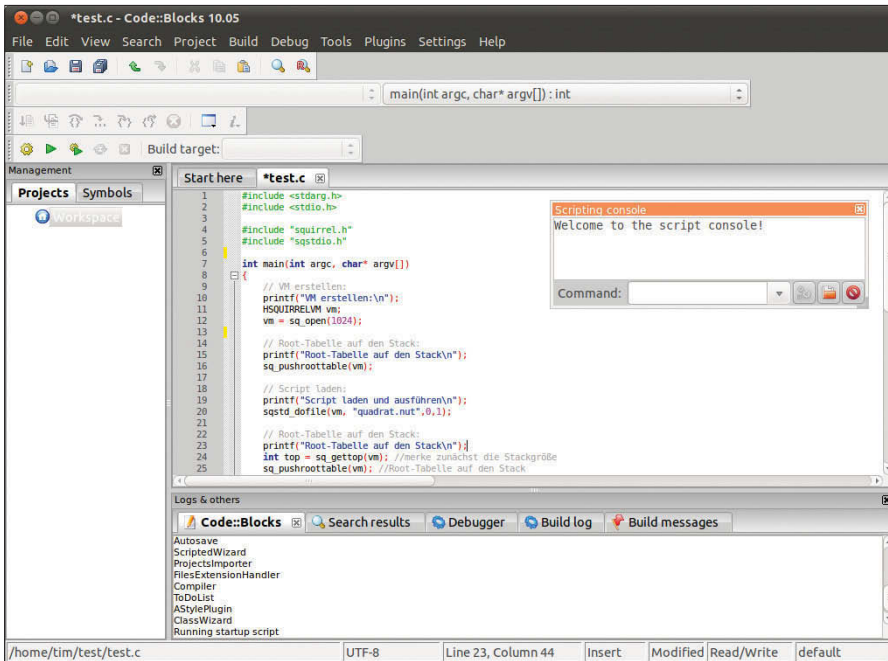


Figure 2: The Code::Blocks developer environment can be extended using Squirrel scripts.

```
sq -c myscript.nut -o myscript.cnut
```

Next, pass in the resulting myscript.cnut:

```
sq myscript.cnut
```

If you call sq without any parameters, this launches the interpreter in an interactive operating mode where you can type Squirrel commands directly (Figure 3). However, to quit the prompt, you will have to press Ctrl + C.

Hello World

The obligatory Hello World in Squirrel is a one-liner:

```
print("Hello World!")
```

The dynamically typed language makes it easy to use variables that you can populate with arbitrary content:

```
local a = 3.14;
local a = "Tux loves Agnes\n";
local a = null;
```

Strings can be dressed with the escape characters you probably know from C. The null keyword empties the variable and is thus similar to, say, the NULL pointer in C, or nil in Lua. Semicolons complete an expression. If an expression is the only thing in a line, as in the example, you can leave out the semicolon. Squirrel is case-sensitive; comments are

enclosed between /* and */, or you can use // for a single-line comment.

Keeping Control

The control structures, if, while, and for work as in C or C++. Listing 1 uses them to compute the largest common denominator of two numbers. Additionally, Squirrel has a practical foreach:

```
local a=[1,2,3,4,5,6]
foreach(number in a)
    print("number="+number+"\n");
```

Listing 1 also shows how Squirrel programmers define a function: The function keyword is followed by the function name and the parameters in parentheses; return returns the results of the computation. In Listing 1, b is set to 10. The interpreter will use this if the second argument for a call is missing. Squirrel even allows a variable number of parameters:

```
function foo(x, ...)
{
    local a = x+1;
    local b = vargv[0];
    /* ... */
}
```

You can pass in any number of arguments to this function and use the vargv array to

access them in the main function. Functions are first class objects in Squirrel: You can put them in variables or pass them in as arguments to other functions.

To save typing, Squirrel versions 3.0 and later use lambda expressions – that is, functions without names. They are really useful for sorting and searching:

```
local numbers = [9,3,2,4,7];
numbers.sort(@(a,b) a <=> b);
```

Alternatively, you could do this:

```
numbers.sort(function(a,b)
    { return a <=> b; } );
```

The comparative operator <=> is also new in Squirrel v3.0. It returns 0 if the two values are equal. sort() is a function of any array, which is really neat.

Tables

Besides arrays in C or C++ style, shown in the following code snippet, Squirrel also supports tables.

LISTING 1: LCD

```
01 // largest common denominator of a
    and b
02 function lcd(a, b=10)
03 {
04     if (b == 0) return a;
05     else {
06         if (a > b) return lcd(a-b, b);
07         else return lcd(a, b-a);
08     }
09 }
10
11 // compute a couple of lcds
12 local x=1;
13 while(x < 10)
14 {
15     print (lcd(x, 27)+"\n");
16     x = x+1;
17 }
```



Figure 3: The @ in front of the string maintains the text between the quotes character for character.

```
local colors = ["red", "green", "blue"];
```

Other languages will call this data structure a dictionary or an associative data field. Like a phone book, it stores any information you pass to it under a unique key.

For example, you can store the value `Tim` under the `Name` keyword. In Squirrel, programmers can save numbers and complete functions as payload data – not just text – and you can mix things as needed:

```
local test=
{
  color="red"
  b=function(x) { return x*x; }
}
```

Squirrel refers to each key/value pair as a slot. Access to stored information is possible using dot notation:

```
print(test.color);
print(test.b(2));
```

You can change an existing slot with a simple assignment:

```
test.b = 20;
```

To add a new slot, use the `<-` operator:

```
test.c <- 20;
```

As of Squirrel 3.0, tables can also use JSON notation:

```
local color = {
  "name": "blue",
  "colornumber": 173,
  "variants": ["metallic", "matt"]
}
```

Squirrel stores global variables, functions and tables in a global, or root, table. At the end of Listing 1, slots for the `lcD()` function and the variable `x` are added to the table. The preceding colon lets you access one of these slots in the root table:

```
::lcD(1,2);
```

In real life, programmers often need to browse lists. In Squirrel, you can use generators to help you do this. These special functions work just like a bubble

gum machine that spits out the next ball of bubble gum whenever you kick it.

Give It to Me!

To create a generator, you simply need to tag the return value in a normal function with `yield`:

```
function vendor(n)
{
  for(local i=0; i<n; i=i+1)
    yield i;
  return null;
}
```

When you call this function, it returns the stopped generator (i.e., the bubble gum vending machine in its original state):

```
local intake=vendor(10);
```

The `resume` command triggers the generator:

```
local x=resume intake;
```

The `vendor()` now runs precisely until the `yield` command occurs, and `yield` returns the results of the expression that follows it – in this example, the content of `i` and thus a `1`.

After this, the `vendor()` function stops again. If you wake it up with `resume`, it will run until the next `yield`, and this would return `2` in this example. Using this approach, the loop

```
while((x=resume intake)!=null) {
  print(x+"\n");
}
```

would pick up all the numbers from the generator from `1` to `n`.

Stop and Go

In addition to generators, Squirrel has threads, which Squirrel calls functions, that stop in midstream and can be woken later. In contrast to generators, thread functions in the interpreter have their own stack, their own root table, and their own error handler.

To convert a function into a thread, Squirrel programmers use `suspend()` to define the points where it should go to sleep:

```
function saytwo()
{
```

```
::print("First\n");
::suspend();
::print("Second\n");
}
```

Here, `newthread()` creates a thread object from the function

```
local onethread = ::newthread(saytwo);
```

and starts the thread using `call()`:

```
onethread.call();
```

It keeps on running up to the first `suspend()`. To wake it up again, do:

```
onethread.wakeup();
```

You can assign a parent table to any table. If you then query a non-existent key in the table, the Squirrel interpreter automatically delegates the request to the parent table.

In Listing 2, you can see a simple example of this: because `_colorname` doesn't exist in the `Mixer` table, the interpreter automatically asks the parent table, `Color`, which kindly returns the value of `_colorname`.

Classic

In contrast to Lua, Squirrel has some object orientation, as you can see in Listing 3. It starts by defining the `Dot` class, from which the `Rectangle` class is then derived. `Rectangle` overwrites the `constructor()` function in this process.

The `base` keyword lets it access the basic class function. `this` ensures that the variables in this class are used, rather than the interpreter creating new,

LISTING 2: Delegation

```
01 Mixer <- {
02 }
03
04 Mixer.Content <- function()
05 {
06   ::print(_colorname);
07 }
08
09 local color = {
10   _colorname="red"
11 }
12 Color.setdelegate(Mixer)
13
14 Color.Content();
```

LISTING 3: Inheritance

```

01 class Dot {
02     x = 0;
03     y = 0;
04
05     function constructor(x1,y1)
06     {
07         this.x=x1; this.y=y1;
08     }
09 }
10
11
12 class Rectangle extends Dot {
13     height = 0;
14     width = 0;
15
16     function constructor (x,y, h,b)
17     {
18         this.height=h; this.width=b;
19         base.constructor(x,y);
20     }
21     function Output()
22     {
23         ::print(x+"," +y+"," +height+","
24             +width+ " \n");
25     }

```

local variables. You can now create an object in the finished class:

```
local onedot = Dot(1,2);
```

Squirrel automatically passes the two arguments to the Dot class's constructor() function and then returns a finished object. You can use dot notation to access its variables and functions:

```
onedot.output();
```

Internally, Squirrel handles classes like tables. The Dot class could also be written as shown in Listing 4.

There is a funny side effect here: You can retrospectively use arrow notation to add functions or variables to a class:

```

Dot.add <- function(c,d)
{
    this.x=this.x+c;
    this.y=this.y+d;
}

```

C++ programmers are familiar with a more common abbreviation for this:

```

function Punkt::add(c,d)
{

```

LISTING 4: Alternative Class

```

01 Dot <- class {
02     x = 0;
03     y = 0;
04
05     function constructor(x1,y1)
06     {
07         this.x=x1; this.y=y1;
08     }
09 }

```

```

    this.x=this.x+c;
    this.y=this.y+d;
}

```

The use of tables comes at the price of all functions and variables being public, but you can store classes in variables, just like functions, or pass them into a function as parameters.

Like Java 5, Squirrel supports annotations; programmers can thus append metadata to classes and functions. This information, which Squirrel refers to as attributes (Listing 5), can then be read and used by a documentation system or a developer environment.

Totally Overloaded

The following notation would be an elegant approach to adding two Dots quickly:

```

local p1 = Dot(1,2);
local p2 = Dot(4,7);
local p3 = p1 + p2;

```

To be able to use this, Squirrel coders simply need to implement the _add() function in the Dot class (Listing 6).

The _add() function belongs to the group of meta methods. If you overwrite

LISTING 5: Attributes

```

01 class Test </ test = "An attribute
    of the class" />{
02     </ test = "Attribute of the
        function"/>
        function PrintTesty()
03     {
04         ...
05     }
06 }

```

them or implement them yourself, you change the behavior of one of Squirrel's built-in functions or, in this case, an operator. In the example shown here, Squirrel knows how to apply the plus sign to Dots.

Conclusions

Squirrel doesn't deserve to lead the life of a wallflower in the shadow of Lua. It is a tried-and-true language and is much easier to learn for object-oriented developers. The fact that the scripting language is mainly used in computer games should not put you off. On the contrary, these programs typically need a very lean and fast interpreter.

On the downside, the documentation is pretty terse. It is mainly a language reference, although it does include many small examples, is easy to read, and is very comprehensive. For more information, you can check out the somewhat chaotic-looking wiki [5], and you can go to the forum if you have any open issues [6]. ■■■

LISTING 6: Overloading

```

01 class Dot
02 {
03     ...
04     function _add(otherdot)
05     {
06         return ::Dot(x+otherdot.x,
07             y+otherdot.y);
08     }

```

INFO

- [1] Squirrel website: <http://squirrel-lang.org/>
- [2] Programs that use Squirrel: <http://wiki.squirrel-lang.org/default.aspx/SquirrelWiki/ProjectsUsingSquirrel.html>
- [3] Squirrel in Code::Blocks: http://wiki.codeblocks.org/index.php?title=Scripting_Code::Blocks
- [4] Squirrel shell: <http://squirrelsh.sourceforge.net/>
- [5] Squirrel wiki: <http://wiki.squirrel-lang.org/default.aspx/SquirrelWiki/SquirrelWiki.html>
- [6] Squirrel forum: <http://forum.squirrel-lang.org/>
- [7] Listings for this article: <http://www.linuxpromagazine.com/Resources/Article-Code>

Too Swamped to Surf?

Our ADMIN Online website offers additional news and technical articles you won't see in our print magazines.

Subscribe today to our free ADMIN Update newsletter and receive:

- Helpful updates on our best online features
- Timely discounts and special bonuses available only to newsletter readers
- Deep knowledge of the new IT

ADMIN
Network & Security



© Yuri Arcurs, fotolia.com

www.admin-magazine.com/newsletter

Replacing the master boot record with the GUID partition table

Table Dance



Large disks use the GUID partition table in place of the legacy master boot record. The usual Linux fdisk programs can't handle the new partitioning scheme, but you can work around it if you have the right tools. *By Hans-Peter Merkel*

The history of hard disks has continually been to hit limits. First, it was the BIOS, then it was the controllers [1], and even the operating system. Maximum numbers of sectors and heads, partition limits and numbers, boot areas – the list is endless.

In the 1980s, MS-DOS 3.2 forced users to split their expensive storage into 32MB partitions. A hard disk of that time weighing in at 85MB would have needed at least three partitions (Figure 1).

Two modern Seagate ST 33000651 AS disks, kindly provided to *Linux Magazine* by the manufacturer for the tests in this article, weigh in at around 3TB each,

AUTHOR

Hans-Peter Merkel has been an active member of the open source community for many years, focusing on forensics.

He trains the staff of crime prevention authorities in Europe, Asia, and Africa and is the founder and chair of FreiOSS and Linux4Afrika.



which exceeds the current 2TB critical limit for PCs by the master boot record (MBR) partition table (Figure 2) [2].

The legacy MBR partitioning schema, CHS (cylinder-head-sector), reached the end of the road many moons ago at 8GB (a maximum of 1,024 cylinders x 255 heads x 63 sectors). Since then, the BIOS and operating system have assumed 254 heads, 63 sectors, and 1,023 cylinders and have used the 4 bytes in the partition table to define the position and size of the sectors (Logical Block Addressing, or LBA).

The PC world has fared well with this trick thus far, but the times, they are a-changing (or already have), because (232 – 1 sectors) x 512-byte sector size results in a maximum size of 2.2TB for a partition and 4TB for a disk [3] [4].

In other words, if you want to use the entire space provided by a state-of-the-art 3TB disk – as a file server or mail server, for example – you are going to need to make some major changes. Welcome to the world of the GUID partition table!

The GUID partition table (GPT) is the successor to the MBR partition table and part of the EFI (Extensible Firmware In-



Figure 1: Old and new side by side: a physically large hard disk, with just 85MB capacity (background), and a state-of-the-art 3TB SATA hard disk by Seagate (foreground).

terface) standard [5]. The EFI specification – the vendor-independent version is known as Unified EFI (UEFI) – replaces and standardizes those parts of the BIOS that are responsible for booting and processing pre-boot programs.

The GPT uses LBA-like modern MBR tables, but with 64 bits, thus making a hard disk of up to 8ZB (zetabytes, or 8 billion TB, if you prefer) addressable for a disk with a 512-byte sector size. Additionally, GPT can manage up to 128 partitions (Figure 3) [6].

GPT Hands-On

The Linux kernel supports the current crop of large disks and GPTs as of version 2.6.25, even though Linus Torvalds decries EFI as “Intel brain damage” [7]. Our lab system addresses the Seagate as `/dev/sdh`; Listing 1 shows the output from the `dmesg` command.

If you try to partition a GPT with the popular `fdisk` or `cdisk` tools, you won't get very far; but, the graphical GParted program supports GPTs (Figure 4). Here, I will use the tool to create a single large data partition on the first disk.

In GParted, you first select *Device | Create Partition Table*; this leads to a dialog window with a huge warning that you will lose all the data on the disk. In the same window, click on the triangle beside *Advanced*, and in the drop-down

box beside *Select new partition table type*, choose *gpt*.

Selecting Partition | New clears the way for you to create a new primary partition with `ext4` as the filesystem. After you click *Add*, a message lets you know that the operation could take some time, depending on the number and size of operations.

On the test system, this translated to a 10-minute coffee break. At the end, however, I had 3TB of disk space – minus a substantial management overhead of 44.05GB (Listing 2).

Listing 3 shows the 512-byte MBR, which GPT uses and correctly populates for reasons of compatibility. The `0x55AA` in line 5 points to the end of the MBR. The second sector of the GPT contains the header information for GUID partitioning (lines 7 through 12). Partition entries show up in position `0x400`, comprising:

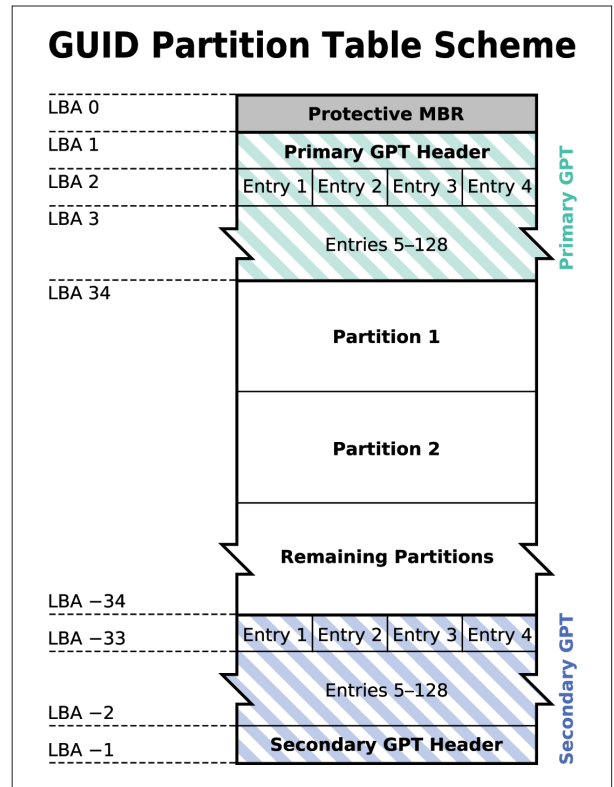


Figure 3: GPT structure. A 100-byte header follows the standard MBR format for compatibility with legacy operating systems. The partition table uses 32 blocks to store the data from a maximum of 128 partitions of 128 bytes. (CC-by-SA-2.5)

- Partition type (16 bytes)
- Partition GUID (16 bytes)
- Start of the partition (8 bytes)
- End of the partition (8 bytes)

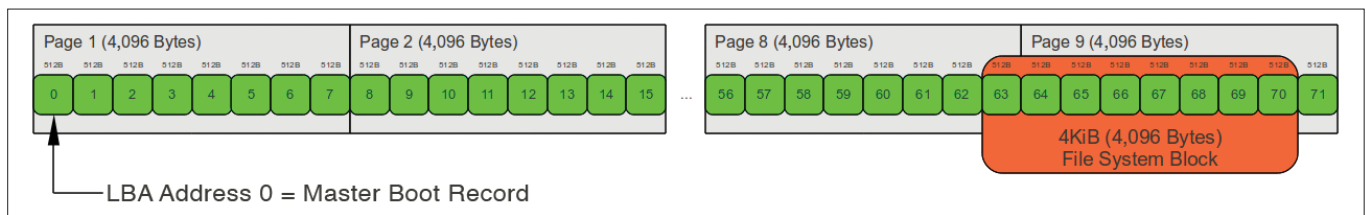


Figure 2: Master boot record structure. The partition table with four entries only has 4 bytes per sector number.

LISTING 1: Kernel Detects 3TB Disks

```
01 [782663.410960] ata6: SATA link up 6.0 Gbps (SStatus 133 SControl 300)
02 [782663.445923] ata6.00: ATA-8: ST33000651AS, CC44, max UDMA/133
03 [782663.445931] ata6.00: 5860533168 sectors, multi 0: LBA48 NCQ (depth 31/32)
04 [782663.446811] ata6.00: configured for UDMA/133
05 [782663.446830] ata6: EH complete
06 [782663.447076] scsi 7:0:0:0: Direct-Access ATA ST33000651AS CC44 PQ: 0 ANSI: 5
07 [782663.447507] sd 7:0:0:0: Attached scsi generic sg9 type 0
08 [782663.447585] sd 7:0:0:0: [sdh] 5860533168 512-byte logical blocks: (3.00 TB/2.72 TiB)
09 [782663.447780] sd 7:0:0:0: [sdh] Write Protect is off
10 [782663.447788] sd 7:0:0:0: [sdh] Mode Sense: 00 3a 00 00
11 [782663.447928] sd 7:0:0:0: [sdh] Write cache: enabled, read cache: enabled, doesn't support DPO or FUA
12 [782663.448846] sdh: unknown partition table
13 [782663.474623] sd 7:0:0:0: [sdh] Attached SCSI disk
```

- Attributes (8 bytes)
- Partition name (72 bytes)

The GUID (Globally Unique Identifier) handles naming in the GPT and is a unique, 16-byte encoding of the partition type. The test system now has a data partition thanks to GParted. The matching GUID is EBD0A0A2-B9E5-4433-87C0-68B6B72699C7; the implementation in line 14 rotates a couple of items in the ID (this looks like a Big Endian-Little Endian thing). Incidentally, EFI stores a copy of the GPT at the end of the data medium as a backup. Thus, you can restore a damaged primary GPT using this copy. If you need to convert a storage medium

from GPT back to MBR partitioning, you also need to delete the second GPT; otherwise, the system will continue to identify the disk as being GPT partitioned.

Other GPT-Capable Tools

Server admins who work at the command line will not find GParted very useful. The choice of GPT-capable fdisk ersatz programs is fairly low right now.

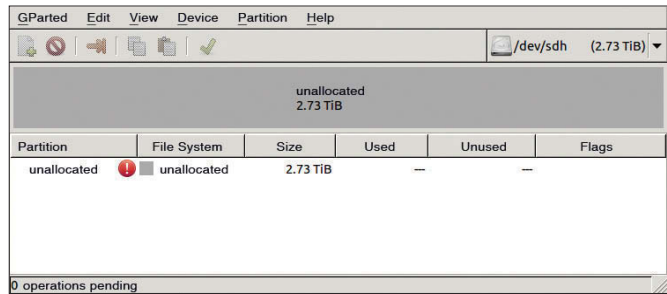


Figure 4: GParted detects an unpartitioned 3TB disk without problem.

The most interesting of them is the GPT fdisk project [8], which comprises the `gdisk`, `sgdisk`, and `fixparts` programs. A current version of Ubuntu will only retrieve `gdisk` from its repository; thus, it makes more sense to pick up all of the packages directly from the openSUSE Build Service [9].

Once you've installed the toolset, you can check and modify the disk. Entering

```
gdisk /dev/sdh
```

returns the following for a new disk:

```
Partition table scan:
  MBR: not present
  BSD: not present
  APM: not present
  GPT: not present
```

In the text-mode menu, choosing `p` prints the partition table (Listing 4), whereas choosing `o` first outputs a security prompt and then creates a GPT table. The `n` option lets you create a new disk-sized data partition, as I did with GParted earlier. `Gdisk` also gives you access to GUID options that GParted doesn't support.

I used the default value of `0700` for Linux/Windows data and then pressed `w` for "Write table to disk and exit." In the test lab, I used `mkfs.ext4` to format the partition. At the command line, the whole process took about 3 minutes less than with GParted. The GPT created was identical to the one GParted made, as a test with `dd` and `xxd` revealed.

From MBR to GPT ...

`Gdisk` can convert MBR-partitioned hard disks to GPT, which removes the need to back up and restore existing partition content. Choosing `gdisk` option `r` sends you to the recovery and transformation menu. For this test, I used an older 200GB disk with an MBR partition. After

LISTING 2: fdisk -lu After Partitioning

```
01 Disk /dev/sdh: 3000 GB, 3000590369280 bytes
02 255 heads, 63 sectors/track, 364801 cylinders, total 5860528065 sectors
03 Units = sectors of 1 * 512 = 512 bytes
04
05 Device Boot Start End Blocks Id System
06 /dev/sdh1 2048 5860532223 2930272033 83 Linux
07 Warning: Partition 1 does not end on cylinder boundary.
```

LISTING 3: dd if=/dev/sdh | xxd (Excerpt)

```
01 [...]
02 00001c0: 0100 eefe ffff 0100 0000 ffff ffff 0000 .....
03 00001d0: 0000 0000 0000 0000 0000 0000 0000 0000 .....
04 00001e0: 0000 0000 0000 0000 0000 0000 0000 0000 .....
05 00001f0: 0000 0000 0000 0000 0000 0000 0000 55aa .....U.
06 [...]
07 0000200: 4546 4920 5041 5254 0000 0100 5c00 0000 EFI PART....\...
08 0000210: 0e0f 49ee 0000 0000 0100 0000 0000 0000 ..I.....
09 0000220: afa3 505d 0100 0000 2200 0000 0000 0000 ..P]....".....
10 0000230: 8ea3 505d 0100 0000 cbb9 6e3f 0765 a746 ..P].....n?.e.F
11 0000240: a7c2 70a1 3db8 25c4 0200 0000 0000 0000 ..p.=.%.....
12 0000250: 8000 0000 8000 0000 6afb 3a17 0000 0000 .....j.:.....
13 [...]
14 0000400: a2a0 d0eb e5b9 3344 87c0 68b6 b726 99c7 .....3D..h.&..
15 0000410: 6903 b666 b91c 7646 a8dd e9d1 d7c6 bc5a i..f..vF.....Z
16 0000420: 0008 0000 0000 0000 ff9f 505d 0100 0000 .....P]....
```

LISTING 4: Partition Information from gdisk

```
01 Disk /dev/sdh: 5860533168 sectors, 2.7 TiB
02 Logical sector size: 512 bytes
03 Disk identifier (GUID): 0AE667A2-818F-4670-A328-500021C76A73
04 Partition table holds up to 128 entries
05 First usable sector is 34, last usable sector is 5860533134
06 Partitions will be aligned on 2048-sector boundaries
07 Use 'l' on the experts' menu to adjust alignment
08 Last sector (2048-5860533134, default =5860533134) or {+}size{KMGTP}:
09 Current type is 'Linux/Windows data'
10 Hex code or GUID (L to show codes, Enter =0700):
```


converting, `gdisk /dev/sde` returned the following results:

```
Partition table scan:
[...]
Found valid GPT with protective MBR; using GPT.
```

Listing 5 shows the output from the `mmls` forensics tool from the sleuthkit package [10]. Line 2 confirms that the conversion was successful. The partition entry in Listing 6 is interesting; the data type is available both in the GUID and in a text entry in lines 4 through 6.

... and Back Again

Gdisk can also convert from GPT to MBR. In some cases, this won't work, so caution is advised. After reconverting the disk I had just GPT'd, `gdisk` returned the output in Listing 7. At this point, the `sgdisk` program from the GPT `fdisk` package is handy for low-level repairs.

Cluster Disks

Normal hard disks store data in sectors of 512 bytes. Operating systems and their

filesystems group multiple sectors to form a cluster; typically, the number of sectors grouped will be eight, thus resulting in a length of 4,096 bytes. To achieve better performance, large disks

now use 4,096-byte cluster sectors internally, and the vendors often refer to these disks as "4k" disks. The Seagate disks that were used during our lab tests were 4k disks.

LISTING 5: `mmls /e3v/sde`

```
01 GUID Partition Table (EFI)
02 Offset Sector: 0
03 Units are in 512-byte sectors
04
05 Slot      Start      End        Length     Description
06 00: Meta  0000000000 0000000000 0000000001 Safety Table
07 01: ----- 0000000000 0000000062 0000000063 Unallocated
08 02: Meta  0000000001 0000000001 0000000001 GPT Header
09 03: Meta  0000000002 0000000033 0000000032 Partition Table
10 04: 00    0000000063 0390716864 0390716802 Linux/Windows data
11 05: ----- 0390716865 0390721967 0000005103 Unallocated
```

LISTING 6: `dd if=/dev/sde | xxd (Excerpt)`

```
01 0000400: a2a0 d0eb e5b9 3344 87c0 68b6 b726 99c7  ....3D..h.&..
02 0000410: 42c4 6d88 c9f9 c84c a33d a0a6 ceb3 2bf3  B.m...L.=...+
03 0000420: 3f00 0000 0000 0000 c0dd 4917 0000 0000  ?.....I.....
04 0000430: 0000 0000 0000 0000 4c00 6900 6e00 7500  .....L.i.n.u.
05 0000440: 7800 2f00 5700 6900 6e00 6400 6f00 7700  x./W.i.n.d.o.w.
06 0000450: 7300 2000 6400 6100 7400 6100 0000 0000  s. .d.a.t.a.....
```

DON'T MISS A SINGLE ISSUE!

The first print magazine created specifically for Ubuntu users! Ease into Ubuntu with the helpful Discovery Guide, or advance your skills with in-depth technical articles, HOW-TOs, reviews, tutorials, and much, much more.

SUBSCRIBE NOW!

4 issues per year for only
£ 24.90 / EUR 29.90 / US\$ 39.95

- ✓ Don't miss a single issue!
- ✓ Huge savings – Save more than 35% off the cover price!
- ✓ Free DVD – Each issue includes a Free DVD!



shop.linuxnewmedia.com

To retain external compatibility for the BIOS, the disk's internal logic will convert the clusters back to 512-byte sectors. To avoid this accommodation of management unit sizes between the operating system and the disk affecting the speed, the disk's and the filesystem's 4KB clusters should match perfectly regardless of whether you use MBR or GPT partitions.

If this is not the case, the results can be disastrous: If you have an offset of 2,048 bytes and the operating system wants to write a cluster to disk, the disk has to write out two half-clusters physically, thus occupying two whole clusters. SSD disks are worst hit by this alignment issue because overlapping write cycles will trigger a read-modify-write cycle on the flash storage.

Start in Sector 63 or 2,048

The problem of operating systems and storage not synchronizing their management units is more the exception than the rule. Microsoft is to blame for this situation because MS-DOS partitioning tools up to Windows XP start the partition immediately after the MBR in sector 63.

Disks partitioned with this method are always misaligned. If the first partition were to start in sector 64, integral groups of 4,096 blocks would be possible. Fdisk on Linux imitates this behavior for compatibility reasons.

Microsoft finally put an end to the problem in Windows Vista and now lets

BOOTING FROM GPT DISKS

Standard server systems, mainframes, and more recent Apple computers comply with the EFI specification and boot from a GPT-initialized hard disk. In contrast, desktop mainboards for Intel and AMD CPUs currently more or less exclusively boot via their MBR BIOS routines.

Linux users are fortunate, in that GRUB 2 has a solution for this issue. The tool's bootloader serves an EFI environment to the computer, and Linux can thus boot from a GPT disk without problem. But, be careful: Early versions of GRUB do not support this feature!

the first partition start in sector 2,048 instead of 63. Depending on the version of fdisk you use on Linux, you might need to disable DOS compatibility mode with the use of a command-line option.

Thankfully, the tools that can create a GPT all use 2,048 as the start sector by default.

This behavior is exactly what I noticed during the tests for this article: Line 6 of Listing 2 shows 2048 as the start sector, which means alignment is perfect on the first 3TB disk. You have to be careful with disks that have been used before. Line 10 of Listing 5 reveals that a GPT can also start in sector 63 rather than 2,048.

You might recall that I converted the second disk from an MBR schema to GPT; this did not change the partition boundary in sector 63.

Whether this problem really is as bad in real life as it seems to be in theory is a subject of much debate, because modern disks use vendor-specific and secret mapping techniques and operating systems and filesystems use a variety of buffers and optimizations. That said, putting the partition boundary on a 4KB

border is a good thing, and administrators should support it.

Conclusions

The era of GPT has definitely started. Modern disks need these new tables to create partitions beyond the 2.2TB limit. Thus, administrators should steer clear of fdisk for the time being, whereas tools like GParted will help plug the gap.

Desktop mainboards have no problem booting Linux from a GPT disk with the GRUB 2 bootloader. In terms of performance, you will need to keep your eyes open: Most state-of-the-art storage uses 4KB clusters internally, instead of 512-byte sectors, which is something you won't notice with LBA; however, this could affect partition performance if worst came to worst. ■■■

INFO

- [1] ST506 interface from 1982: <http://en.wikipedia.org/wiki/ST506>
- [2] MBR: http://en.wikipedia.org/wiki/Master_Boot_Record
- [3] MBR partition table: http://en.wikipedia.org/wiki/Partition_table
- [4] FAQ: Drive Partition Limits: http://www.uefi.org/learning_center/UEFI_MBR_Limits_v2.pdf
- [5] EFI and UEFI: <http://developer.intel.com/technology/efi/>; <http://www.uefi.org/home/>
- [6] GPT: http://en.wikipedia.org/wiki/GUID_Partition_Table
- [7] "Linus On The Extensible Firmware Interface": <http://kerneltrap.org/node/6884>
- [8] GPT fdisk: <http://www.rodsbooks.com/gdisk/>
- [9] GPT fdisk from the openSUSE Build Service: <http://build.opensuse.org/package/show?package=gptfdisk&project=home%3Asrs5694>
- [10] Sleuthkit: <http://sleuthkit.org>

LISTING 7: gdisk /dev/sde

```
01 GPT fdisk (gdisk) version 0.7.1
02
03 Partition table scan:
04  MBR: MBR only
05  BSD: not present
06  APM: not present
07  GPT: not present
08
09 *****
10 Found invalid GPT and valid MBR; converting MBR to GPT format.
11 THIS OPERATION IS POTENTIALLY DESTRUCTIVE! Exit by typing 'q' if
12 you don't want to convert your MBR partitions to GPT format!
13 *****
14
15 Exact type match not found for type code A400; assigning type code for
16 'Linux/Windows data'
```



Email-based PDF conversion service

PDF Machine

Send off a handful of files by email and receive PDF documents as a reply with Open/LibreOffice, Postfix, and a bit of scripting. *By Olaf Hübner*

Many Linux programs create PDF documents from various formats. But what happens if your staff needs to convert multiple documents to PDF in bulk? The method also has to be cheap, simple, fast, and available with a GUI without too much user interaction.

Initial research reveals various potential solutions, but converting multiple documents at the same time always involves a fair amount of user input, and the use of virtual PDF printer drivers is too slow when you have a large number of files.

Office staff are familiar with email, and it's something they use every day for many purposes. This basic thinking is behind a Berlin-based company's idea to combine a mail server and Open/Libre-

Office to create an in-house solution. The user just needs to mail the documents they want to convert as attachments to the server, and they receive the results – PDF documents – in the return mail.

Practical Sequence

A separate mail address is needed to accept the conversion jobs; in this example, I'll use `pdf@domain.tld`. On the conversion side, you need Postfix. The following line in `/etc/aliases` tells the free mail server to pass the message to the conversion script for processing:

```
pdf: z
    "|/usr/local/bin/pdf-konverter.sh"
```

A call to `postalias /etc/aliases` compiles the configuration for Postfix. The `pdf-konverter.sh` script first unpacks the attachments. It then passes them on to `Unoconv` [1], a utility for communication with Open/LibreOffice. The free office suite handles the conversion process before the script finally bundles the files back into attachments and returns them to the sender.

The maximum size of an email is defined by the `message_size_limit` parameter in the Postfix configuration (Listing 1). Because Office is running as a daemon, the conversion process itself is very fast: as the figures show, a 20MB email with 19 attachments just takes about a minute to convert and return – and the test mail went through an antivirus scanner on the inward and outward trip. Users receive a single small office file back in just a couple of seconds.

System Requirements

The people in Berlin used the 64-bit version of Debian Squeeze for their setup. The solution will run just as well on Ubuntu 10.10. Because it only uses open source components, it should work on other distributions, too.

Besides Postfix as the mail server, you also need packers and unpackers for the attachments. `ripMIME` is the tool of choice for unpacking the attachments; it grabs the MIME-encoded attachments from an email and stores them as single files in a selected folder. An alternative tool for this purpose is `Uuview` [2]. Having said this, `ripMIME` was a far faster unpacker in our lab. `ripMIME` is currently available in a source code version only [3]; you can build using `make && make install`. The command

```
chown nobody:nogroup z
    /usr/local/bin/ripmime
```

assigns the correct permissions to the binary. In contrast, the `uuencode` program is readily available in the Debian `sharutils` package. It encodes the converted files into the response mail.

Office Server Control

You will definitely need the `Unoconv` command-line tool to run Open/LibreOf-

LISTING 1: /etc/postfix/main.cf

```
01 myhostname = pdfserver.domain.tld
02 mydomain = domain.tld
03 mydestination = $mydomain
04 myorigin = $mydomain
05 alias_maps = hash:/etc/aliases
06 alias_database = hash:/etc/aliases
07 # optional smarthost
08 # relayhost = mail.domain.tld
09 mynetworks = 127.0.0.0/8
10 message_size_limit = 5120000
```

AUTHOR

Olaf Hübner works as a systems administrator in Berlin, Germany. He started gathering Unix experience in 1990 and had an initial encounter with Linux in 1993, after installing from floppy disks.



office in server mode. The tool uses the free office suites' UNO bindings to read supported formats and output them in the chosen conversion format. The Unoconv homepage [1] features an impressive list of supported import and export formats.

To launch LibreOffice as a daemon, use

```
unoconv --listener
```

which will listen on TCP port 2002 by default, although you can change this

using the `--port` command-line option. Other big advantages of listener mode are the speed at which documents are converted and the simple syntax. More information on starting LibreOffice in daemon mode and about the API can be found online at the LibreOffice help pages [4] and OpenOffice wiki [5].

There is a downside: if you have a standard installation of Debian Squeeze and launch unoconv in listener mode, you will see an error message from the Python UNO bridge. The workaround here is to install the LibreOffice package from backports instead of OpenOffice. Because unoconv cooperates closely with the version of the office suite, you need the package from the Debian unstable branch [6].

After becoming root, enter the following commands to create an init script to

LISTING 2: /etc/init.d/unoconvd

```
01 #!/bin/sh
02 ### BEGIN INIT INFO
03 # Provides: unoconvd
04 # Required-Start: $network
05 # Required-Stop: $network
06 # Default-Start: 2 3 5
07 # Default-Stop:
08 # Description: unoconvd - Convert
    to PDF
09 ### END INIT INFO
10 case "$1" in
11     start)
12         /usr/bin/unoconv --listener &
13         ;;
14     stop)
15         killall soffice.bin
16         ;;
17     restart)
18         killall soffice.bin
19         sleep 1
20         /usr/bin/unoconv --listener &
21         ;;
22     esac
```

LISTING 3: pdf-konverter.sh (curtailed)

```
01 #!/bin/sh
02
03 [...]
04 # Extract file attachments as single files
05 /usr/local/bin/ripmime -i $PDFMAIL -d $PDF_PATH
    --overwrite --syslog --extended-errors
06
07 # list supported files
08 /bin/ls $PDF_PATH/*. * \
09 | /bin/egrep -i
    '\.(do[ct]|doc[mx]|rtf|pp[ts]|xl[st]|csv|od[stgdp]|ott|
    htm|sd[wcd]|sxc|jpg|pts|bmp|png|tif|gif)$' \
10 > $PDF_PATH/files_to_convert.txt
11
12 # compose mail header and body
13 echo "From: $PDF_FROM" >>$EMAIL_RETURN
14 echo "To: $SENDER" >>$EMAIL_RETURN
15 echo "Subject: $SUBJECT" >>$EMAIL_RETURN
16 echo "MIME-Version: 1.0" >>$EMAIL_RETURN
17 echo "Content-Type: multipart/mixed;
    boundary=\"$$BOUNDARY\"" >>$EMAIL_RETURN
18 echo "--$BOUNDARY" >>$EMAIL_RETURN
19 echo "Content-Type: text/plain; charset=\"$$CHARSET\";
    format=flowed" >>$EMAIL_RETURN
20 echo "Content-Transfer-Encoding: quoted-printable"
    >>$EMAIL_RETURN
21 /bin/cat $PDF_PATH/body_return.txt >>$EMAIL_RETURN
22
23 # line-by-line processing of file list
24 while read FILE_NAME
25 do
26     /usr/bin/unoconv -f pdf "$FILE_NAME"
27 done < $PDF_PATH/files_to_convert.txt
28
29 # encode file names as ISO-8859-15, remove special
30 # characters and base64-encode PDF files
31 for ATTACHMENT in $PDF_PATH/*.pdf
32 do
33     FILENAME=$(basename "${ATTACHMENT}")
34     (echo "${ATTACHMENT}" ; echo $FILENAME) >>
    $PDF_PATH/pdf_filenames_and_basenames.txt
35     ( echo -n "\"=${CHARSET}?" ; \
36         echo -n $FILENAME | /usr/bin/perl
    -pe 'use MIME::QuotedPrint;
    $_=MIME::QuotedPrint::encode($_);' \
37         | /bin/sed 's/./=2E/g'
    | /bin/sed 's/ /=20/g'
    | /bin/sed 's/=//g'; \
38     echo "=?\" \
39     ) | tr -d '\n' >$PDF_PATH/filenames_codiert.txt
40     FILENAME_CODIERT=$(/bin/cat
    $PDF_PATH/filenames_codiert.txt )
41     echo "--$BOUNDARY" >>$EMAIL_RETURN
42     echo "Content-Type: application/pdf;" >>$EMAIL_RETURN
43     echo " name=$FILENAME_CODIERT" >>$EMAIL_RETURN
44     echo "Content-Transfer-Encoding: base64"
    >>$EMAIL_RETURN
45     echo "Content-Disposition: attachment;"
    >>$EMAIL_RETURN
46     echo " filename=$FILENAME_CODIERT" >>$EMAIL_RETURN
47     echo "" >>$EMAIL_RETURN
48     # strip first and last line
49     /bin/cat "${ATTACHMENT}"
    | uuencode --base64 $BOUNDARY
    | /bin/sed '1,1d'
    | /bin/sed '$d' >>$EMAIL_RETURN
50 done
51 # end of mail message
52 echo "--$BOUNDARY--" >>$EMAIL_RETURN
53
54 # send mail
55 /bin/cat $EMAIL_RETURN | /usr/sbin/sendmail -t
    -f pdf@domain.tld
56 exit 0
```

control the office server and fill it with the content shown in Listing 2:

```
touch /etc/init.d/unoconvd
chmod 755 /etc/init.d/unoconvd
```

After doing this, the insserv tool makes sure the service is launched at the standard runlevel:

```
insserv -d unoconvd
```

If you don't have insserv, you can set up the service with `update-rc.d` or `chkconfig`, or you can set it up manually. Once you have launched the daemon, whether you use RC or do so manually, as in `/etc/init.d/unoconv start`, you should find `unoconv` and `soffice.bin` in the process list.

The script that accepts the email, converts it, and returns it to the sender will feel quite at home in `/usr/local/bin/pdf-konverter.sh`. It should be owned by `nobody:nogroup`. Listing 3 shows the curtailed content.

If you are unhappy with `unoconv` and Open/LibreOffice, you can change the lines in the converter script to one of many alternatives: `Antiword` [7] (first converts to Postscript and then using `ps2pdf` to PDF), `Pandoc` [8], `JOD` and `PyOD-Converter` [9], and `LOPDFConverter` [10] are your choices.

Future

This working solution basically lets you run a fast PDF conversion server. Because `unoconv` version 0.4 or newer offers special export filters, you could add another account to `/etc/aliases` and write a second, modified script (or parameter evaluation part) for a targeted approach to changing export options [11]. Before you deploy the PDF converter in a production environment, pay some attention to security: running `unoconv` as a defined user, creating cronjobs to delete the temporary paths, creating logfiles, and hardening Postfix. ■■■

INFO

- [1] Unoconv: <http://dag.wieers.com/home-made/unoconv>
- [2] Uudeview: <http://www.fpx.de/fp/Software/UUDeview>
- [3] ripMIME:
<http://www.pldaniels.com/ripmime/ripmime-1.4.0.9.tar.gz>
- [4] Starting LibreOffice with parameters: http://help.libreoffice.org/Common/Starting_the_Software_With_Parameters
- [5] OpenOffice export to the PDF file format:
http://wiki.services.openoffice.org/wiki/Documentation/OOo3_User_Guides/Getting_Started/Exporting_to_PDF
- [6] `unoconv-unstable`:
<http://packages.debian.org/sid/all/unoconv/download>
- [7] `Antiword`: <http://freshmeat.net/projects/antiword>
- [8] `Pandoc`: <http://johnmacfarlane.net/pandoc>
- [9] `JOD` and `PyODConverter`:
<http://artofsolving.com/opensource>
- [10] `LOPDFConverter`:
<http://sourceforge.net/projects/lopdfconverter>
- [11] OpenOffice API for PDF export: http://wiki.services.openoffice.org/wiki/API/Tutorials/PDF_export
- [12] Listings for this article:
<http://www.linux-magazine.com/Resources/Article-Code>

Online Training with Linux Magazine Academy Preparing for the LPIC exam - the easy way:

Equivalent to CompTIA Linux+
powered by LPI

GET YOUR LINUX KNOW-HOW CERTIFIED WITH LPIC:

Professional training
for the exams
LPI 101 and 102



- hardware settings
- package management
- partitioning and filesystems
- shell environment
- automate system administration
- network configuration
- security administration tasks
- troubleshooting

20%
off for current
subscribers

For more information and to order:
academy.linux-magazine.com/LPIC

Storing your passwords properly

Security Blanket

High-performance graphics cards and proper storage can help keep your passwords secure. *By Kurt Seifried*

Computer performance has come a long way in the past few years. Moore's law has continued moving forward; right now, you can easily find a high-end CPU with 6 to 12 cores for between a few hundred and a thousand dollars.

But what if you want to throw a couple hundred cores at a problem? You could buy a rack or two of equipment, or you could just spend a few hundred dollars on a video card. In fact, graphics cards are so good at certain types of computation that both NVidia and AMD now make specialized cards (the NVidia Tesla series and the AMD FireStream series) that have a ton of cores, several gigabytes of memory, and extremely fast interconnects.

So, assuming you have a decent graphics card (or are willing to buy one), what can you do with a few hundred cores? The most obvious answer is encryption, which is embarrassingly easy to parallelize and works very well on GPU-based computing systems.

SSL Support

The good news is that enabling SSL on most web servers won't take more than

KURT SEIFRIED

Kurt Seifried is an Information Security Consultant specializing in Linux and networks since 1996. He often wonders how it is that technology works on a large scale but often fails on a small scale.

a few percentage points of CPU power. The bad news is if you're building a front-end load balancer capable of providing SSL termination for, say, a few thousand clients, you're probably going to need to buy a specialized SSL acceleration card.

Or, you could use a graphics card to provide more cores to handle key setup and data [1]. However, little software is available to provide support for this approach outside of a master's thesis by Urmas Rosenberg (University of Tartu in Estonia) [2] and some associated code that provides AES-128 block support for OpenSSL on CUDA (NVidia-based) cards.

Unfortunately, this is more theoretical than practical right now, "The general result is left as an exercise to the reader." In general, I think generic CPUs are getting so many cores that the need to throw a few hundred cores at this problem will basically become less important.

Password Cracking

Rainbow tables provide a very practical application of GPU-based encryption computation for something useful. The general idea is that people are terrible at choosing passwords. Passwords like

"password" or a pet's name – even with some additional characters – simply do not provide a lot of entropy (randomness). Even if the password is salted, you can still do pre-computation to cut down on the search time when you do want to crack a password. The idea is to do a lot of work upfront so that later on you can quickly retrieve passwords.

Some Numbers with AES-256

If you are using AES-256 to hash and store your passwords, you're doing it wrong (more on this in the next section). But, a lot of people are using AES-256 or SHA-256, and it makes for some impressive numbers. Suppose you want to pre-compute the hash values for all valid characters on a US-English keyboard (26 letters, 10 numbers, 11 other character keys for a total of 94 characters) up to a password length of eight characters (so, $94^1 + 94^2 + \dots + 94^7 + 94^8$ possible passwords). Storing the input and the AES-256 value (ignoring indexing requirements) for this would result in about 1,400,000TB of data.



Although hard drive prices have dropped, they haven't dropped enough to make this possible. But you can cheat by precomputing a chain of values, starting with a value of 1, for example, hashing that, then hashing that result and repeating until you've done it 200,000 times.

You then store the end value of this chain along with the starting value, which means about 14TB of storage (plus indexing requirements, etc.) – or, less than US\$ 1,000 in today's hard drives. When you have a hashed password that you want to crack, you compare it to the stored values, if it matches, you win; if not, you hash the hashed password and look for that value.

In the worst case scenario, you have to search around 200,000 times, but eventually you hash the hashed password and end up with a value for which you have a valid chain. Assuming you have hashed the password and searched 50,000 times until you found a match, you then take the starting value of that chain and hash it 149,999 times to get a value. This value, when hashed, will match the password that you're trying to crack.

Why are MD5, AES-256, and SHA-256 such bad choices for storing your hashed passwords? Because they are very fast algorithms, especially on modern hardware, and especially on modern GPUs. The PostgreSQL project has posted some numbers [3] – on a 1.5GHz Pentium 4, you can do 2,345,086 MD5 hashes per second. On a modern GPU, this increases to hundreds of millions per second. However, you don't actually need to make your own rainbow tables, you can download them (and the software used to create them) from a number of free sources [4] [5]. Two main methods are used to defend against rainbow tables: The first is salting, and the second is using encryption and hash functions, like bcrypt, designed for password storage.

Salting Passwords

Salting passwords primarily defends against situa-

tions in which the attacker obtains the encrypted or hashed password (e.g., by stealing the `/etc/shadow` file or downloading the database) to brute force it. The salt ensures that a pre-computed brute-force attack will take longer because each password must be encrypted with all possible salt values first.

Some significant weaknesses are inherent with salting, however. The first is that it won't really help against badly constructed passwords: Attackers can easily brute-force a list of the million most common passwords even if proper salting is used.

The second issue is that most systems don't care about the password, they care whether the encrypted or hashed value of the password matches the system entry. Because hash functions like AES-256 only provide 2^{256} possible unique outputs, collisions are obviously possible. Ultimately, the attacker wants to find a data string that will encrypt or hash to the same value as the stored one. And, by brute-forcing chains of possible values, chances are the attacker can find a value that works.

Password Storage

The ugly truth is that most encryption and hash functions are designed to provide encryption and hashing, not secure password storage. However, the bcrypt tool [6] is designed just for password storage. Basically, it uses the Blowfish encryption algorithm to hash data but introduces a work function that determines how much work it will take to hash the data.

By setting a large value for the work function, you can make bcrypt take an arbitrarily large amount of CPU time (say 0.1 seconds on a modern system) to encrypt the password. This can obviously affect system performance (e.g., if you have 10 users logging in every second, all the CPU time would be consumed by bcrypt).

The advantage of bcrypt is that, as time goes on, you can increase the work function, defeating attackers in the future (assuming they didn't steal the password file 10 years ago). You might think a similar outcome could be achieved by using multiple rounds of MD5 or AES, but that would actually make the system easier to attack. For example, using 1,000 rounds of MD5 gives

the attacker 1,000 possible values, which when MD5'ed 1,000 times will result in a password stored in their rainbow table. So, please don't do that.

A Note on Software

Of course, none of this matters if the software you are using doesn't work properly. A perfect example is the release of PHP version 5.3.7, which contained a critical security bug. A small change (to avoid a warning) was made to the `crypt()` function when using the MD5 algorithm (the default).

The result of this small change was that, instead of passing back the salt value and the password, `crypt()` concatenated the password to the salt (essentially making a large salt value with no password). Thus, if anyone else tries to login, only the salt values will be compared. The values will of course match, thus allowing the bad guy in. So, I guess the moral of the story is to run your unit tests when you make changes to cryptographic functions.

Conclusion

Brute-forcing older algorithms is definitely possible now (DES and 3DES already fell to brute-force attacks several years ago). The latest algorithms like AES and SHA are good, but, ironically, one of their biggest strengths, their speed, also works against them. So, choosing something slower like bcrypt might be a good idea.

Finally, it doesn't matter how you encrypt passwords if you allow users to pick weak passwords (especially words listed in dictionaries). You might want to download a few dictionaries and check against them when a user attempts to set or change a password. ■■■

INFO

- [1] Accelerating SSL with GPUs: <http://www.ndsl.kaist.edu/papers/comm022t.pdf>
- [2] OpenSSL-GPU: <http://labs.sasslantia.ee/openssl-gpu/>
- [3] pgcrypto: <http://www.postgresql.org/docs/8.3/static/pgcrypto.html>
- [4] Free Rainbow Tables: <http://www.freerainbowtables.com/>
- [5] RainbowCrack Project: <http://project-rainbowcrack.com/>
- [6] bcrypt: <http://bcrypt.sourceforge.net/>

Perl script cleans up virtual machines

Freshen Up

A VirtualBox installation creates snapshots of virtual machines in next to no time at the command line; it also helps protect your privacy while browsing and sends Perl modules through a smoke test. *By Mike Schilli*



If you are continually trying out new Linux distributions, you will probably be able to use a virtualizer like VMware, KVM, Xen, or VirtualBox blindfolded. But programmers who need to develop code for different Linux versions or, heaven forbid, Windows versions, also appreciate locally available virtual machines for quick tests.

The VirtualBox [2] virtualization package is easily controlled through its attractive GUI (Figure 1) and is easily downloaded as an Ubuntu package under the GPLv2 license (Figure 2).

Users quickly can create a handful of guest systems using installation CD/DVDs or ISO files and boot the guests in separate windows at the press of a button. Mouse control on the guest system needs some getting used to; once the guest system has grabbed the mouse

focus, it doesn't let go. The mouse pointer won't move outside of the guest system window borders, ex-

cept for some guest systems that allow a shared focus. To leave the virtual system focus, you need to press a predefined key (the default is the right Alt key); this takes the mouse pointer back to your desktop.

Freezing and Thawing

Efficient snapshot technology lets you freeze a virtual machine's state and re-

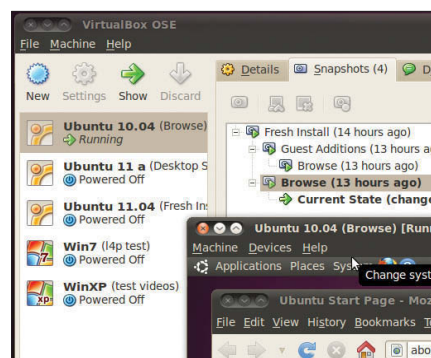


Figure 1: VirtualBox with some Ubuntu and Windows versions as guest systems.

<input checked="" type="checkbox"/>	virtualbox-ose
<input type="checkbox"/>	libpurelibc-dev
<input type="checkbox"/>	virtualbox-ose-dbg
<input checked="" type="checkbox"/>	virtualbox-ose-qt
<input type="checkbox"/>	libpurelibc1
<input checked="" type="checkbox"/>	virtualbox-ose-dkms
<input checked="" type="checkbox"/>	virtualbox-ose-guest-dkms

Figure 2: Synaptic listing of Ubuntu packages that install VirtualBox.

MIKE SCHILLI

Mike Schilli works as a software engineer with Yahoo! in Sunnyvale, California. He can be contacted at mschilli@perlmeister.com. Mike's homepage can be found at <http://perlmeister.com>.

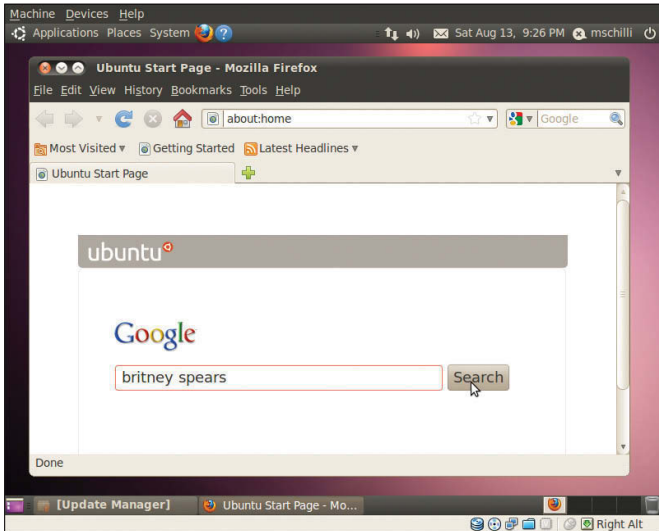


Figure 3: The browser will have forgotten this embarrassing search term the next time you fire up the virtual machine.

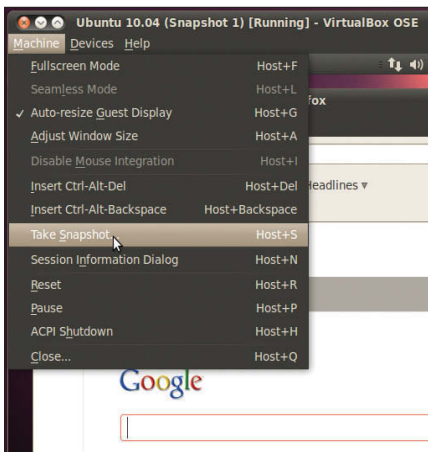


Figure 4: The user creating a snapshot of the virtual machine state.

store it in a couple of seconds. This approach means that web surfers can use a freshly installed Ubuntu with an open browser for embarrassing web searches (Figure 3).

The torrid details of what Google knows about its users are always revealed when you enter a search key and the autocomplete function does the rest. After all, nobody wants to be reminded that they searched for “athlete’s foot” just a couple of days ago, not to mention the influence that the search term has on long-term targeting by Google’s personalized advertising machine.

Paranoid penguin friends will not want to keep the cookies they received for longer than necessary or to leave any tracks in the browser history. Because there’s no way of knowing exactly what the browser does behind your back in terms of caching and other data storage,

it only makes sense to fire up a virtual machine quickly for sensitive searches and then to restore the virtual machine’s original status when you’re done.

This approach isn’t entirely watertight; if you’re combating a mastermind, they might find data snippets on the virtual hard disk, and the IP address of your router will appear in Google’s

logs, but it does raise the bar considerably. If you need more privacy, you might like to use Tor [3] and also shred against the virtual machine disk, which is a single file.

Shortly after installing Linux from the DVD, users who are interested in protecting their privacy will launch the Firefox browser in the brand-new guest system and create a snapshot of the virtual machine (Figure 4).

After completing his secret mission, the user then powers off the virtual machine using the *Stop* command and then switches back to the original state by selecting the *Restore* function in the VirtualBox snapshot menu. The next time you launch the virtual machine you will have an open browser that doesn’t remember a thing – just as if you’d jumped back in time.

Automated Command

Instead of clicking around in the menus to launch the VirtualBox GUI every time you want to research something, you might prefer to use a Perl script that automates the process of selecting and booting the virtual machine. When you’re done, you press the Enter key in the script to shut down the virtual machine and restore its original state. Luckily, VirtualBox offers the *VBoxManage* tool, which gives you complete control at the command line. Listing 1 shows the Perl script that selects the virtual machine named “Ubuntu 10.04” and its “Browse” snapshot, which I prepared previously (see the list of snapshots in Figure 1).

The script uses the *tap* command from the CPAN *Sysadm::Install* module to issue shell commands. Of course, it might be easier to implement the program as a shell script, but experienced Perl programmers know it’s only a matter of time until the functional scope of a shell script has grown to the extent that you will want to reimplement it in Perl for fear of not being able to maintain the monster.

The script uses the subcommands *startvm* and *ctrlvm poweroff* provided by *VBoxManage* to start and shut down the virtual machine. It restores a snapshot both at the start and at the end of the script to make quite sure that the virtual machine boots into the snapshotted “Browse” state, even if somebody has messed around in the Snapshot menu of the VirtualBox GUI in the meantime.

Limited Networking

VirtualBox virtual machines can also be managed in headless mode without any screen output. An invisible SSHD dae-

LISTING 1: browse

```
01 #!/usr/local/bin/perl -w
02 #####
03 # browse - VM for browsing
04 # Mike Schilli, 2011
05 # (m@perlmeister.com)
06 #####
07 use strict;
08 use Sysadm::Install qw(:all);
09 use Log::Log4perl qw(:easy);
10
11 Log::Log4perl->easy_init(
12     $DEBUG);
13
14 my $vbm = "VBoxManage";
15 my $vm = "Ubuntu 10.04";
16
17 tap $vbm, "snapshot", $vm,
18     "restore", "Browse";
19 tap $vbm, "startvm", $vm;
20
21 print
22     "Press Enter for shutdown";
23 <STDIN>;
24
25 tap $vbm, "ctrlvm", $vm,
26     "poweroff";
27 tap $vbm, "snapshot", $vm,
28     "restore", "Browse";
```

mon running on the guest system lets you execute commands from the host system or exchange files between guest and host.

That said, VirtualBox launches guest systems in NAT (Network Address Translation) mode by default, assigning them an address on the virtual 10.x.x.x network, and using address translation to communicate with the host's local network in a similar approach that most routers in private homes take to let the local devices with their 192.168.x.x addresses communicate with the Internet. This works fine from the virtual machine to the local network; however, the host and the devices on the local network can't open a connection to the virtual machine. The Network Adapters dialog hidden behind the icon with the two terminals bottom right on an active virtual machine (Figures 5 and 6) allows you to change this behavior.

Bridge Opens Firewall

If you change the setting from NAT to Bridged Mode, the virtual machine re-

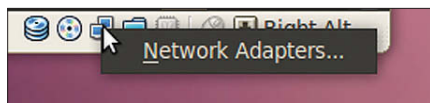


Figure 5: Right-clicking the icon with the two terminals leads to a dialog box for setting up the network adapter.

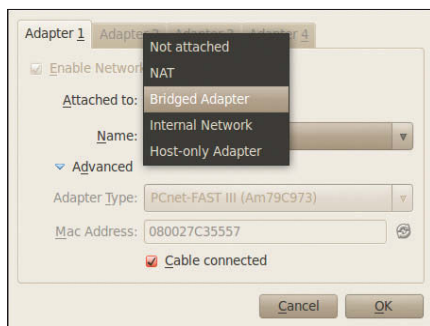


Figure 6: The Bridged Adapter in the network dialog supports bidirectional communications between the virtual machine and the rest of the world.

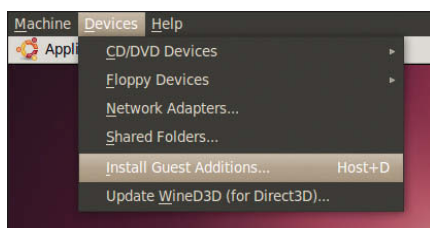


Figure 7: Installing Guest Additions gives you useful utilities for VirtualBox VMs.

trieves an IP address from the DHCP server on the local network (e.g., 192.168.0.135), becoming a peer communication partner. If you launch an SSHD daemon with

```
sudo apt-get  
install  
openssh-server
```

on the virtual machine, clients on the host or on the local network that log in to the virtual machine (ssh 192.168.0.135) with their user IDs are granted access. If the user also copies their SSH public key (e.g., ~/.ssh/id_rsa.pub) to the ~/.ssh/authorized_keys file on the virtual machine, they don't even need a password – and this is important for automated scripts, assuming the private key was created without a passphrase.

However, it's not always easy to find out which IP address a specific virtual machine has picked up when it launches. VirtualBox offers a method via guest properties; an extension that you first need to install on the active virtual machine after selecting *Devices | Install Guest Additions* (Figure 7).

After doing so, the virtual machine downloads an ISO file off the Internet, mounts the file like a CD drive on the virtual machine, and executes a shell script stored on the CD that triggers an orgy of kernel module builds.

Manual Attention Needed

Unfortunately, this approach didn't work in the VirtualBox version that I was using, 3.1.6; I had no alternative but to download VBoxGuestAdditions_3.1.6.iso manually to the virtual machine, mount the file using a `mount -o loop`, install the dkms Ubuntu package, and then execute the `VBoxLinuxAdditions-x86.run` shell script from the command line (Figure 8). Note that this will not work with Ubuntu 11 as the guest system if the host system is still running Ubuntu 10.04, which seems to require platform parity.

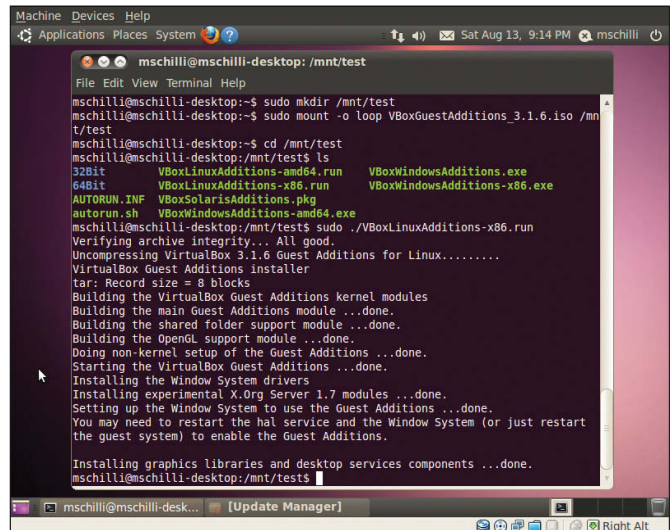


Figure 8: Installing the Guest Additions on 10.04 takes a fair bit of manual TLC.



Figure 9: The guestproperty command shows you which IP address the specific virtual machine has picked up from the DHCP server.

After restarting the virtual machine, the host can issue the `VBoxManage guestproperty enumerate "Ubuntu 10.04"` command and, hidden in a mass of other information, receive the line "Name: /VirtualBox/GuestInfo/Net/0/V4/IP, value: 192.168.0.135", which gives you the IP used by the virtual machine. If you run `get` for the `guestinfo` path shown above instead of `enumerate`, you are sent the IP address of the active virtual machine directly (Figure 9).

Smoking Chimneys

Another use case for virtual machines is testing CPAN modules in pristine environments. Assuming CPAN modules do a good job of specifying their dependencies on other modules, a CPAN shell will handle the installation without breaking a sweat. Unfortunately, some authors forget to specify modules they have on their development machines that a vanilla Perl installation will not include. This causes much frustration among users – except for cases in which one of today's ubiquitous and automatically activated CPAN smoke tests identify a problem and notify the author via email.

A Perl installation stored as a virtual machine snapshot that the module programmer can boot with the `cpan-smoke` script (Listing 2) is a big help here. You just pass in your newly created CPAN distribution tarball and let a CPAN shell test whether the new version is installable on the vanilla system with a connection to CPAN. If the module survives the accompanying test suite without error, it passes the smoke test and you have a good chance that the module will install on similar systems.

Branches on the Snapshot Tree

For this to happen, developers need to configure a CPAN shell on a virtual machine with retrospectively installed Guest Additions, create a snapshot, and call the snapshot “CPAN Smoke.” Figure 10 shows two branches on the snapshot tree of the Ubuntu 10.04 virtual machine.

The originally created snapshot, `Browse`, inherits from a fresh installation and is independent of any snapshots you

might create on the basis of the second branch, `CPAN Smoke`.

To make sure the installation also works with the mini CPAN shell `cpanm` in the local directory, the CPAN Smoke snapshot also installs the CPAN `local::lib` module. The best way of doing this on a fresh Ubuntu system in a VM is to issue the command

```
$ sudo apt-get install liblocal-lib-perl
```

which relies on the package manager to install `local::lib` on the virtual machine in the main perl module tree only accessible by root. The command

```
$ eval $(/usr/bin/perl -Mlocal::lib)
```

issued in the shell sets the shell variables required to install additional CPAN modules in the user’s home directory, accessible by the user’s regular ID. You can install the mini-shell locally via the “big” CPAN shell:

```
$ cpan App::cpanminus
```

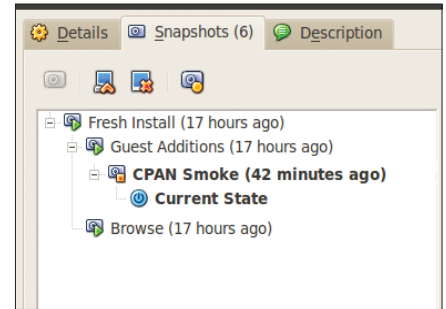


Figure 10: A new CPAN Smoke snapshot, derived from a vanilla Ubuntu install with the Guest Additions, plus a configured and ready-to-run CPAN shell.

After doing so, the `cpanm` command for installing other modules will be available in your `$PATH` inside the VM:

```
$ cpanm --version
cpanm (App::cpanminus) version 1.4008
```

Now you create the snapshot and store it as “CPAN Smoke.” If you call `cpan-smoke` on the host and pass in a CPAN tarball created with `make tardist` in the module directory, the script will restore the snap-



Online Training with Linux Magazine Academy Preparing for the LPIC exam - the easy way



GET YOUR LINUX KNOW-HOW CERTIFIED WITH LPIC:
Professional training for the exams LPI 101 and 102

- hardware settings
- package management
- partitioning and filesystems
- shell environment
- automate system administration
- network configuration
- security administration tasks
- troubleshooting

**Equivalent to CompTIA Linux+
powered by LPI**

For more information and to order: academy.linux-magazine.com/lpic

LISTING 2: cpan-smoke

```

01 #!/usr/local/bin/perl -w
02 #####
03 # cpan-smoke - VM for CPAN
04 # module smoke testing
05 # Mike Schilli, 2011
06 # (m@perlmeister.com)
07 #####
08 use strict;
09 use Sysadm::Install qw(:all);
10 use Proc::Simple;
11 use Net::Ping;
12 use Log::Log4perl qw(:easy);
13
14 my ($tarball) = @ARGV;
15
16 die "usage: $0 tarball"
17   if !defined $tarball;
18
19 Log::Log4perl->easy_init(
20   $ERROR);
21
22 my $vbm = "VBoxManage";
23 my $vbh = "VBoxHeadless";
24 my $vm = "Ubuntu 10.04";
25 my $ippath = "/VirtualBox" .
26   "/GuestInfo/Net/0/V4/IP";
27 my $snap = "CPAN Smoke";
28
29 # in case it's up in
30 # foreground mode
31 tap $vbm, "controlvm", $vm,
32   "poweroff";
33 tap $vbm, "snapshot",
34   $vm, "restore", $snap;
35
36 my $proc =
37   Proc::Simple->new();
38 $proc->start(
39   "$vbh --startvm '$vm'");
40
41 END {
42   $proc->kill();
43 }
44
45 INFO
46   "Waiting for VM to come up";
47
48 while (!$proc->poll()) {
49   DEBUG
50     "Waiting for VM process";
51   sleep 1;
52 }
53
54 my $ip;
55
56 while (!defined($ip = ip()))
57 {
58   DEBUG "Waiting for IP";
59   sleep 1;
60 }
61
62 my $ping = Net::Ping->new();
63 while (!$ping->ping($ip)) {
64   DEBUG "Waiting for Ping";
65   sleep 1;
66 }
67
68 INFO "VM is up: $ip";
69
70 tap "scp", $tarball,
71   "$ip:/tmp/$tarball";
72 sysrun "ssh", $ip,
73 qq{eval `$(/usr/bin/perl
74   -Mlocal::lib);
75   cpanm /tmp/$tarball};
76
77 tap $vbm, "controlvm", $vm,
78   "poweroff";
79 tap $vbm, "snapshot",
80   $vm, "restore", $snap;
81 #####
82 sub ip {
83   #####
84   my ($stdout) = tap $vbm,
85     "guestproperty",
86     "get", $vm, $ippath;
87
88   if (
89     $stdout =~ /Value: (.*)/)
90   {
91     return $1;
92   }
93
94   return undef;
95 }

```

shot, start the virtual machine, and wait until its network configuration is complete and responds to ping.

The `ip()` function (Listing 2, lines 82-95) looks for the IP address of the virtual machine you are using by sending a `guestproperty` request to VirtualBox. A regular expression grabs the relevant IP address from the verbose output.

The `VBoxHeadless --startvm` command of line 39 launches the virtual machine without any GUI output and waits in the foreground until the user kills the program by pressing `Ctrl + C`. The script uses the `CPAN Proc::Simple` module to send the `startvm` command into the background and remembers its PID so the `END` snippet triggered at the end of the script sends its `kill()` command to the right program in line 42, thus shutting down the headless virtual machine.

Smoke Signals?

Line 70 copies the tarball into the virtual machine's `/tmp` directory, and line 74

launches the `cpanm [tarball]` command, which picks up the tarball, unpacks it, and issues `make` commands to test and install. If `cpanm` discovers CPAN modules that the module specifies as dependencies, it downloads them automatically from CPAN and proceeds to install them.

To make sure the shell in the virtual machine finds the `cpanm` command, searches for installed CPAN modules in the local `~/perl5` path, and installs new modules in this path, the `ssh` command in line 72 issues the `eval` command shown earlier to set the required shell variables before running the `cpanm` command itself.

The output from the `cpanm` installation script is directed to the standard output of the calling `cpan-smoke` script, thanks to the `Sysadm::Install` module's `sysrun` function, pointing diligent developers to first-hand information on how their latest creation behaves in an pristine environment. If you do see a smoke signal, you can assume something is wrong;

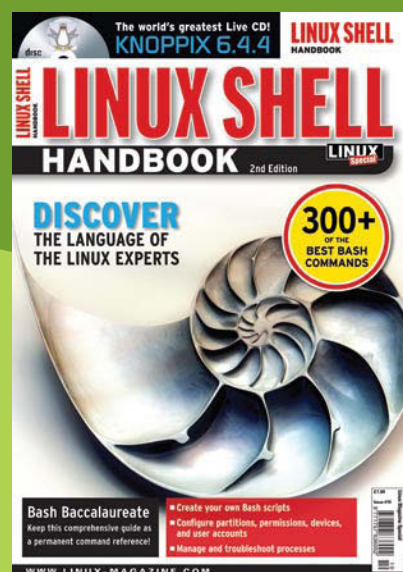
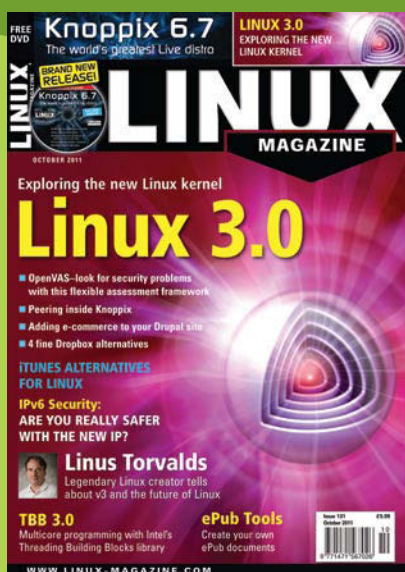
most likely, the dependencies on other CPAN modules need careful review.

In today's age of the cloud, where virtual machines such as Amazon's EC2 are available on the web, even for those on a budget, a local virtualization solution like VirtualBox is still an attractive option. A detailed instruction manual for newcomers – compared with the typically less structured offerings on the Internet – is available [4], although it describes the more or less obsolete VirtualBox version 3.1. ■■■

INFO

- [1] Listings for this article: <http://www.linuxpromagazine.com/Resources/Article-Code>
- [2] VirtualBox Project Page: <http://www.virtualbox.org>
- [3] Tor: <http://www.torproject.org/>
- [4] Romero, Alfonso V. *VirtualBox 3.1: Beginner's Guide*. Packt Publishing, 2010 (Kindle edition)

OUR BEST DEAL EVER!



THREE FANTASTIC PRODUCTS FOR ONE LOW PRICE OF US\$ 115.85*

*Quoted price for US readers – look online for similar discounts in your country!

- ▶ US\$ 40 off the subscription price
- ▶ US\$ 120 of the newsstand price

LIMITED TIME

- ▶ 1 Year subscription to Linux Magazine or Linux Pro (12 issues)
- ▶ 1 Year subscription to ADMIN Magazine (4 issues)
- ▶ Linux Shell Handbook

shop.linuxnewmedia.com/hattrick



Authoring documents with Mallard

On Topic

Simple syntax and a modular structure make Mallard the ideal starting point for creating topic-based documentation. *By Mario Blättermann*

If you develop GUI-based software, it only makes sense to rescue your faithful users if they get stuck. Writing help documents can be pesky. The typical DocBook format is generally regarded – and rightly so – as complicated, and it requires a huge learning curve.

Many developers and documentation specialists are turning to Mallard as an alternative for creating user help documentation. Mallard is similar to DocBook if you look at the logical structure, but it uses far fewer tags and considerably simplifies the syntax.

Mallard [1] is specifically designed for creating help docs. In other words, Mallard is ideal for situations in which the information is organized into small, task- or topic-oriented elements. The design forces the author to work in a topic-based way and not just string information together in a block of static text.

According to the Mallard project, good online help has three important characteristics:

- Users can navigate to information, following a path that fits their mental model.
- Users can find information they need in a topic without reading any more background information than necessary.
- Users can learn more if they choose by following links to more in-depth information.

Mallard helps you create documents that embody these characteristics. The emphasis is on small, bite-sized topics interconnected through links. The XML-based format that Mallard uses makes it easy to extend help docs by inserting new topics without modifying the existing structure.

The Mallard project defines a format that can then be viewed by any tool ca-

pable of viewing Mallard files. The project recommends the Gnome Yelp help viewer as an example of a Mallard-ready viewer.

Getting Started

Listing 1 shows a minimal Mallard file. The index is the root on which all the other page files are based. The headers indicate that this is a guide from the content viewpoint, a task stylistically, and an index from an organizational point of view. The latter is important to ensure that other pages assume their rightful places in the documentation. The remaining lines tell you something about the author and the license. After a short lead-in, you come to the texts that you will see later on. The `<section>` block creates a section, which can be empty apart from the title.

Adding `id="introduction"` to the opening XML tag ensures that the parser cre-

ates a page file with this ID exactly here and nowhere else in the main document. Refer to Listing 2 for the structure of the page file.

The third line in Listing 2 defines the ID, which is confirmed by supplying a link in line 6. You can add more topics and sub-topics in the same way.

These listings demonstrate the simple Mallard syntax. If you are not worried about giving your text elements a sophisticated look – at least for the time being – you will have no trouble getting started. Later on, you can add style information inside the opening XML tags, which will be processed by other tools.

Mallard makes life easy for the user here, too: Instead of using different tags to express the various GUI elements (e.g., `<gui label>`, `<gui menuitem>`, and `<gui button>` in DocBook), you just use the succinct `<gui>`. If needed, you can then add, for example, `style="button"` in the opening tag to define that you want a button in the GUI.

You also have the option of adding a `` tag for information that will not be displayed later on, but which is very useful for various processing tools. Additionally, you can extend Mallard by adding elements from external namespaces, virtually without limits.

Because you need to create a page file for every topic, and typically every sub-topic, you might confuse hardened DocBook users at first, because they're used

to a single large file that only keeps the license declaration somewhere externally. Separating items has its benefits, though: If you assign intuitive names to the individual files, it will be easier for you to find parts that you want to modify, delete, or supplement later on.

At the same time, links facilitate the task of integrating external files, even if they are from third parties. For example, if the developers write a separate help text for the plugins available for a program and reference this cleanly as part of the main documentation, the external topic will be shown as an integral part without you having to change a single character in the manual. If the plugin is not installed, the corresponding help page doesn't exist, and the manual reader will not waste time searching for functions that don't exist.

As an author, you will find that the format doesn't pose any more problems than other markup languages. Although you will search in vain for a WYSIWYG editor tailored for the syntax, the Gnome

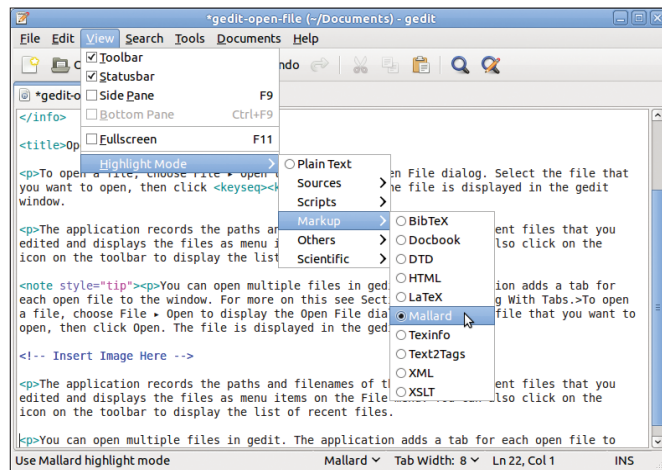


Figure 1: Gedit has supported Mallard for some time now.

standard editor, Gedit (Figure 1) understands the format and its characteristics. A text snippet plugin for Gedit gives you default tags and offers to close any tags you have opened.

The Emacs editor also identifies the syntax and displays it correctly. If your favorite editor doesn't speak Mallard, you can enable syntax highlighting for generic XML. This approach is typically fine to help you find your way around the tag jungle.

Assembly Line

The source text alone doesn't make a manual: You first have to process it. This step is where Mallard loses some points compared with its secret role model, DocBook. Apart from using the Gnome

LISTING 1: Minimal Mallard File

```
01 <page xmlns="http://projectmallard.org/1.0/"
02     type="guide" style="task"
03     id="index">
04
05 <info>
06 <title type="text">Sample documentation</title>
07 <credit type="author">
08 <name>John Doe</name><email>max@online.de</email>
09 </credit>
10 <license>
11 <p>Creative Commons Share Alike 3.0</p>
12 </license>
13 </info>
14
15 <title>Sample application</title>
16
17 <section id="introduction">
18 <title>Introduction</title>
19 </section>
20 </page>
```

LISTING 2: Page File Structure

```
01 <page xmlns="http://projectmallard.org/1.0/"
02     type="guide"
03     id="introduction">
04
05 <info>
06 <link type="guide" xref="index#introduction"/>
07 <credit type="author">
08 <name>John Doe</name><email>max@online.de</email>
09 </credit>
10 <license>
11 <p>Creative Commons Share Alike 3.0</p>
12 </license>
13 </info>
14
15 <title>What is <app>sample application</app>?</title>
16 <p>
17 <app>Sample application</app> is a program
18 with many interesting functions.
19 </p>
20 </page>
```

help browser Yelp and exporting to HTML, where the former internally relies on the latter, virtually no display programs understand Mallard.

If you have ever converted a DocBook document into a PDF document with a professional layout, you might be in for disappointment with Mallard. Although a conversion is being developed [2], work on it isn't making much progress. Right now, the design makes it difficult to create attractive hard copy.

The ability to translate texts with the use of gettext PO files is understood in Mallard. The popular xml2po does not do a good job of this, but a replacement is in sight. The ITS tool [3] from the Gnome treasure trove now lets you automate the process of adding comments for the translator and hiding elements not designed for translation in the PO files.

The predecessor wasn't all that choosy and put everything in the PO files – whether it was program code, commands, or content for translation. The new method reduces the error count in foreign language versions, thus helping to improve the not always impeccable reputation of manuals.

Gnome Rebuilt

Not so long ago, DocBook was the tool of choice for authoring manuals in Gnome. Since the introduction of Mal-

lard, this has changed fundamentally, and it was not just the simplified syntax that prompted the move. The intent was to avoid having documents look like theses and to give them a more wiki-like layout to facilitate searching for information. One of the early adopters was the Empathy program, which offered a Mallard-formatted manual as early as Gnome 2.28 in fall 2009 (Figure 2).

Behind the facade were also a couple of innocuous-looking changes. Newly written documentation was published under a Creative Commons license, rather than under the GFDL, as was the case previously. This approach facilitates the propagation of texts because the documentation teams of, for example, Fedora and Ubuntu, also chose this type of license. This change means nothing can prevent authors from exchanging and mutually integrating documents.

Changing a license also means seeking the agreement of all the previous authors, if the project wants to continue using existing content. In many cases, the effort involved in doing so would be too high a price to pay. But, it does give maintainers an option for preventing obsolete content from being published. At the same time, they can impose a topic-based structure.

Right now, the official Gnome servers offer 42 manuals in the new format, and the figures are rising. External projects

such as Déjà Dup [4] and Simple Scan [5] have also adopted Mallard. As of Gnome 3.2, when integrating external help pages, it doesn't matter which installation prefix the main program and plugin are installed under. The browser will render a page installed in `~/local` just as easily as one located in `/usr`. Thus, you can install plugins in your home directory without having root privileges – including correctly integrated help pages.

In general, you could say that Gnome is the motor that drives Mallard, although Mallard

is not a Gnome-only project but has a universal claim.

Future

Because of the restrictions mentioned, Mallard will probably not be able to oust DocBook in the long term – simply complement it. The developers respond quickly to feature requests, but the field of application is restricted to displaying text onscreen. It's anybody's guess whether there will ever be a usable LaTeX/PDF export feature. If you need to write technical documents for printing, you have no alternative to DocBook – unless you happen to use LaTeX already. But, if you need to write topic-based user manuals that offer your readers the convenience of a familiar wiki-style layout, Mallard is your tool of choice.

Development is still very active. As soon as they finish one new feature, they plow on with the next. Work is in progress on glossaries [6], which allows you to highlight terms in the text and link them automatically with explanations. This feature was prompted by a commercial user, which clearly demonstrates the wide acceptance the format now finds.

One hopes Mallard will find its way into more projects in the future and that it will be developed into more than a feasibility study. The Xfce desktop project is on the right track [7]. Currently only the Xfce bar and terminal programs have Mallard manuals, with files provided directly in HTML by a help browser, but this does nothing to detract from the basic idea of topic-based help. ■■■

INFO

- [1] Mallard project: <http://projectmallard.org>
- [2] Converting to LaTeX: <http://gitorious.org/+projectmallard/projectmallard/mal2latex>
- [3] ITS Tool: <http://itstool.org>
- [4] Déjà Dup: <http://launchpad.net/deja-dup/>
- [5] Simple Scan: <http://launchpad.net/simple-scan>
- [6] Shaun McCance on glossaries: <http://blogs.gnome.org/shaunm/2011/07/07/mallard-glossaries/>
- [7] Mallard for Xfce: <http://wiki.xfce.org/documentation>
- [8] Author's documentation: <http://mariobl.fedorapeople.org/Mallard-en/>

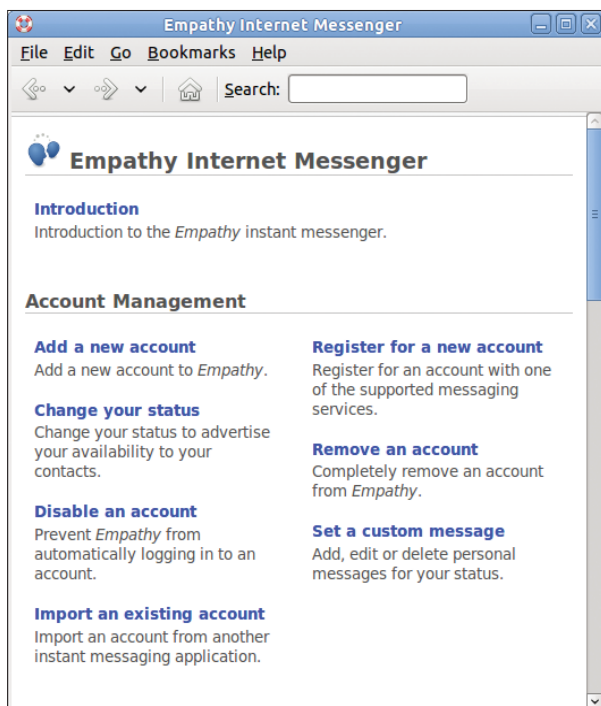


Figure 2: A Mallard-formatted manual offered with Gnome 2.28.

The sys admin's daily grind: OpenNetAdmin

Hooked on Tables

Last month, the Havege daemon helped organize the chaos in this column. Today, Charly attempts to organize the network – a tale of suffering in three chapters. *By Charly Kühnast*

People who want to keep track of the continually growing number of network objects have a whole world of software products at their disposal. These products are generically referred to as IPAM (IP Address Management) tools, and none of them is really popular. Admins initially resist using them before finally giving in. This process typically takes several years and is divided into three levels.

Level one: The network is still pristine. Administrators have set it up themselves and are familiar with every single server, router, and switch – in fact they are on a first-name basis with everything on the net: They even know the IP addresses by heart.

Level two: The network has grown considerably, and DMZ number seven is planned. Virtual servers are springing up like mushrooms. The administrators – typically several of them at this stage – still painstakingly keep a list in a text file or a spreadsheet to stay on top of everything. Level 2 is normally very tenacious until something important blows up because of an error in the manually main-

tained spreadsheet – say, because you forgot a server and assigned its IP to something else.

Level three: The admin crew finally admits the mission is likely to fail if they stick to the spreadsheet method. Their search for software support might just take them to the OpenNetAdmin (ONA) website [1], which prompts the seeker to “Stop using that spreadsheet!” The program’s authors obviously understand what they are doing.

Web GUI and CLI

ONA has become a very powerful tool in the course of its development; it can manage a more-or-less unlimited number of (sub)networks, including the matching DNS and VLAN information. Even more impressive is the way it helps deal with the daily niggles. The Ajax-based web interface is neat, fast, and nice to use, but nobody forces you to use it. In fact, ONA is quite happy if you feed it at the command line.

That said, the web interface does add some value. When you check out a subnet, it displays a bar that indicates how

many addresses on the subnet are already in use. You can assign roles to network components. This is important for infrastructure components for which many switches handle routing. ONA will also let you assign an IP address to multiple devices at the same time. Many IPAM tools prematurely suspect an absent-minded administrator at this stage, but if you manage a cluster with “migrant” IP addresses, you need this function. You can also assign URLs to any object you create, which is useful if you want to monitor the status in Nagios. ONA can also generate a unique URL for every object if needed. For example, if I create a DNS server with the name *Dns1.example.com* in ONA, I can view its overview page by accessing the following URL (Figure 1):

```
http://127.0.0.1/ona/?search=Dns1.example.com
```

The ability to enter the eight-digit MAC address in any form is also useful; that is, the separator can be a colon, hyphen, or nothing. ONA’s parser reliably converts your input into a standard format. ONA might not have the most features, but it doesn’t bug you or go on strike while you’re using it, and these are the best prerequisites for its victory in the battle of the spreadsheet. ■■■

INFO

[1] OpenNetAdmin:
<http://opennetadmin.com>

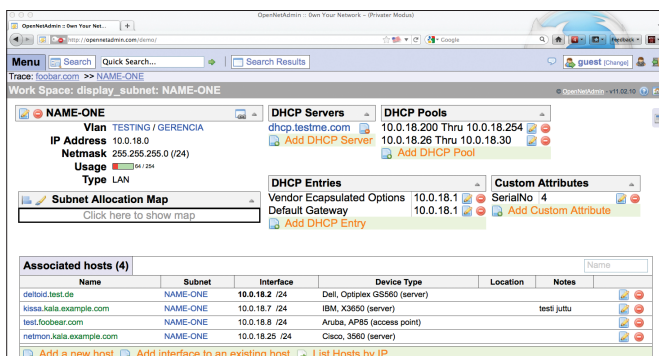


Figure 1: The OpenNetAdmin web GUI offers an attractive host and IP address inventory.

AUTHOR

Charly Kühnast is a Unix operating system administrator at the Data Center in Moers, Germany. His tasks include firewall and DMZ security and availability. He divides his leisure time into hot, wet, and eastern sectors, where he enjoys cooking, freshwater aquariums, and learning Japanese, respectively.



System start with systemd

Full Steam Ahead

systemd is a group of developer solutions for revolutionizing the boot process. But will this turbocharger really add speed? *By Tim Schürmann*

For many years, a tool by the name of SysV Init controlled the system start. It had the task of enabling all the services and applications required for operations. Unfortunately, the tool just doggedly launched one process after another, and the process took some time on modern, function-packed Linux systems. Over time, many alternatives have been released – all of which aimed to do a better and faster job.

Cure-All

One particularly promising candidate goes by the name of systemd [1], which was written by Red Hat employee Lennart Poettering, who also authored NetworkManager. Although systemd just celebrated its first birthday in April, the software has already made the grade for Fedora 15. Other distributions have announced that they will be adopting systemd, too; wherever you look, reviews are very enthusiastic.

systemd uses a mix of new and legacy designs to accelerate the boot process. And, according to its inventor, systemd not only accelerates the start process, it also monitors the services and programs it starts, takes care of mounting filesystems, replaces the time-controlled cron service, and acts as a session manager. At the same time, the tool retains compatibility with existing SysV Init scripts, which have probably grown massively over the years and which admins can keep for now.

To launch all the required components as quickly as possible, systemd recycles some ideas from Mac OS X (or, its Mac OS X counterpart, launchd, to be more precise) and uses a couple of special Linux kernel functions. Poettering's program will thus not port easily to other Unix systems, such as FreeBSD.

Initially, systemd only enables the services you actually need. For example, the CUPS printing system and its dae-

mon only start working if you have a printer attached or if an application wants to print something. Programs that systemd doesn't launch can't block the system start process.

systemd simply starts all the remaining services at the same time. Unfortunately, many services have dependencies. For example, the Avahi network configurator needs D-Bus, which in turn requires syslog.

If worst comes to worst, all of the services involved will start sequentially. systemd uses a simple trick to work around this problem. To understand how it works, you first need to take a quick look under the hood.

Connectors

Linux programs use sockets to communicate with services. You could compare them with the gates at an airport: An airplane docks at the gate and unloads its payload at the terminal. In a similar

fashion, each service provides a socket that gives other programs the option to dock with and offload requests for the service.

The trick is to provide these sockets before the services are fully active. In keeping with the metaphor, this is like having a couple of gangways on the airfield, although the terminal still hasn't been completed. If a plane arrives, the passengers can disembark onto the gangway and wait for the building work to be completed. In the meantime, the plane can take off again.

systemd does exactly this with the `syslog` service, which writes any messages it receives over to a logfile. systemd creates a socket for `syslog` at boot time, just in case. If a program wants to offload an error message, it just throws the data at the socket. If `syslog` isn't running, the messages are temporarily cached.

As long as enough space is left in the buffer, other programs don't need to wait for `syslog` to start; they can just get on with their own work. After `syslog` launches, it picks up the messages that accumulated in the buffer and processes them.

It's a good thing that the Linux kernel manages these queues. systemd only has to set up all the required sockets and then launch all the corresponding services in parallel. This approach doesn't just save administrative overhead, and thus time, it also has a couple of pleasant side effects: If a service dies and

quits, the socket is still there. Requests from programs are thus not lost, but cached.

This means you can restart or replace the service on the fly – just as you can with an update – without the program even noticing. You can even open the sockets and not start the service until messages arrive (*Start on Demand* or *On Demand Loading*). systemd thus assumes the role of programs such as `inetd`.

Many services, and in particular GUI-based programs, use D-Bus as their preferred communication method instead of sockets. Fortunately, the method I just described also works for D-Bus services (bus activation): systemd simply registers a couple of service names with D-Bus before going on to launch the corresponding programs.

Big Brother

If a critical service crashes, systemd needs to restart it as quickly as possible. Because a service can clone itself and launch other programs as needed, it has traditionally been difficult to ascertain the demise of a service.

systemd provides an elegant solution to this problem with the relatively new Linux kernel control groups, or `cgroups` for short. The operating system uses `cgroups` to group programs, or more precisely processes. systemd now locks every service it starts in a `cgroup`. If the service starts other programs or forks itself, as is the wont of web servers or SSH

daemons, the child processes end up in the same `cgroup`. If no active processes exist in a group, the service has either terminated or crashed. systemd then restarts it.

systemd doesn't just take care of services; it can also mount partitions and check them for errors. To make sure these time-consuming processes run in the background and in parallel with all the other services, systemd relies on `Autofs` [2] here. If a program attempts to save to a device that is not yet available, the kernel buffers the request. Once the filesystem becomes available, the kernel hands over the data.

This principle lets you share, for example, your home partition in a network using Samba, although `fsck` is still checking for errors. Additionally, systemd monitors individual directories and mounts the matching filesystem when a program tries to write to them.

Shell Ersatz

Most distributions still use shell scripts to launch individual services. Starting external programs and subshells takes a huge amount of computational time. On top of this, many scripts use recursive constructs and redundant commands; they are error prone and can become difficult to maintain. This explains why `Poettering` views scripts as a major nuisance and consistently avoids them in systemd.

Maintainers should replace the scripted function with real functions, preferably written in C, and thus with programs that execute quickly. Or, they can build the scripts into the daemons themselves. systemd handles important and common functions itself.

systemd can currently set the hostname, mount filesystems, and set the language (`system locale`). It reads the settings for this from the well-known configuration files – or almost.

In some cases, distributions use different files. The system name on Fedora is stored in `/etc/sysconfig/network`, for example. On the other hand, `openSUSE` uses `/etc/HOSTNAME`, and `Debian` uses `/etc/hostname`. In this case, the systemd makers decided on one file: systemd expects the hostname in `/etc/hostname`. This can be seen as an attempt by the systemd developers to standardize distributions through the backdoor.

TABLE 1: Important Directories and Configuration Files

File	Content
<code>/etc/hostname</code>	Hostname of the system
<code>/etc/vconsole.conf</code>	Keyboard layout and console font
<code>/etc/locale.conf</code>	Language settings (locale)
<code>/etc/modules-load.d/*.conf</code>	Kernel modules the system loads at boot time
<code>/etc/sysctl.d/*.conf</code>	Configuration for <code>Sysctl</code> parameters
<code>/etc/tmpfiles.d/*.conf</code>	Configuration for all files the system creates, removes, or cleans up on booting
<code>/etc/binfmt.d/*.conf</code>	Configuration for binary formats to launch Java, Mono, and Wine programs directly
<code>/etc/os-release</code>	Name and other information about the distribution (replaces <code>/etc/fedora-release</code> and similar files)
<code>/etc/machine-id</code>	Computer ID
<code>/etc/machine-info</code>	(Meta)information about the computer
<code>/run</code>	Where programs and services should store temporary data that doesn't belong in <code>/tmp</code> , including socket information or lock files. <code>/run</code> is thus a replacement for <code>/var/run</code> , but a temporary directory (<code>tmpfs</code>).

TABLE 2: Unit Types

Type	Description
service	Service (typically in the form of a daemon).
socket	Encapsulates a socket. Each socket has a matching <code>service</code> unit that launches automatically when a program connects with the socket.
device	Device.
mount	Mountpoint.
automount	Automount point on the filesystem. Each <code>automount</code> unit has a matching <code>mount</code> unit, which <code>systemd</code> mounts when a program accesses the directory.
target	Groups other units, which then act as a single unit.
snapshot	Works like <code>target</code> and stores the service state. This makes it possible to temporarily shift the system state to <code>Panic</code> and then return to the normal work environment.
swap	Manages swap files and partitions.
timer	Enables services at certain times or at certain intervals; the specification uses the cron syntax.
path	Activates units depending on whether specific files exist or a spool directory has reached a certain fill level

TABLE 3: Conditions for a Service

Instruction	Meaning
After	The service would like to launch after the specified unit.
Require	The service relies on the specified unit.
Wants	The service would like the unit to be running.
Conflicts	The service does not cooperate with this unit.

Table 1 provides a brief overview of the most important configuration files and directories, all of which were created in coordination with the distributions, by the way. For more information on this, check out Lennart Poettering's blog [3].

Unitarian

`systemd` refers to any task that it manages as a *unit*. A unit could comprise, for example, the CUPS printing service, or it could relate to mounting the home directory. Both units obviously require differ-

ent actions, which explains why each unit has a specific type.

CUPS is a service and thus takes the `service` type, whereas mounting the home directory is a `mount` type. To make sure `systemd` actually learns about the unit, it needs a matching configuration file. The file has the same name as the unit, and it comprises a freely selectable name and the type.

Thus, it would make sense to call the configuration file for the printing service `cups.service`. Besides `service` and `mount`, `systemd` also uses other categories, as detailed in Table 2.

Home Worker

To enable a certain service at boot time, you need to create a matching configuration file. Listing 1 shows the `noisemeasuring.service` file as a small example of software for a noise-measuring station.

LISTING 1: noisemeasuring.service

```
[Unit]
Description=This service records flight noise.
After=syslog.target

[Service]
ExecStart=/usr/bin/noisemeasurement
Restart=on-abort

[Install]
WantedBy=multi-user.target
```

LISTING 2: Creating a Fork

```
[Service]
ExecStart=/usr/bin/noisemeasurement -d
Type=forking
Restart=on-abort
```

LISTING 3: Using D-Bus

```
[Service]
Type=dbus
BusName=de.dfld.noisemeasurement
ExecStart=/usr/bin/noisemeasurement
```

The configuration files use a structure similar to the popular `.desktop` files. The `[Unit]` section contains some generic information about the service, which includes a short human-readable `Description`.

The `systemd` software uses the `syslog` service to report issues and sends the data it records to files on disk. With `systemd`, services can always expect the filesystem to exist, so there is no need to make sure. This just leaves the `syslog` dependency, which is handled by the `After=` parameter in the third line. For this, you can just enter the unit on which the program depends.

If multiple dependencies exist, simply list the unit names separated by blanks. This list can be fairly long, and to avoid too much typing, target units let you group multiple units under a single banner. `systemd` also comes with a couple of special target units in place [4]. They include `syslog.target`, as used in Listing 1. This unit simply starts a `syslog` implementation.

Because `systemd` activates as many services as possible in parallel, the instruction here only serves to propose a (start) order and not to enforce it. Besides `After`, `systemd` also supports the instructions in Table 3.

The next section `[Service]` provides some information about the service itself. It thus only exists in `.service` type configuration files. `ExecStart=` lets you specify the name of the program file, which is `/usr/bin/noisemeasurement` in this example. `systemd` calls this when it launches the service.

As generally required by `systemd`, the software runs in the foreground. If the service daemon has to run in the background or create a fork, it uses the `Type=forking` parameter to tell `systemd` (Listing 2). And if the service uses D-Bus, `systemd` has a variant to cover that, too (Listing 3).

The `BusName=` parameter states the D-Bus name of the software. `Restart=on-abort` makes sure that `systemd` restarts the service if it terminates for some reason.

The last section `[Install]` in Listing 1 tells `systemd` when, and under what conditions, you want to launch the service. In the example, the software runs (`WantedBy`) after `systemd` enables the `multi-user.target` unit. This encapsu-

Attend the second Android Developer Conference



AnDevCon II

The Android Developer Conference

November 6-9, 2011 • San Francisco



30+ Expert Speakers
**70+ Technical Classes
and Workshops**

Google KEYNOTE!

Chet Haase and
Romain Guy present:

"Android Awesomeness"



"A lot of useful, cutting-edge information."

—Alfred Mirzagitov, Sr. Software Engineer, Webroot

"AnDevCon had lots of great information,
excellent speakers and a coherent program."

—Paul Verger, Software Developer, Pico Software

"There were great presentations with very
professional lecturers. Go for it!!!"

—Eyal Zmora, Software Engineer, NDS Technologies

Register Early and SAVE!

Download the complete course listing at
www.AnDevCon.com

A BZ Media Event



TABLE 4: Systemctl Commands

Command	Meaning
daemon-reload	Reparse the configuration
start Unit	Start Unit
stop Unit	Stop Unit gracefully
kill Unit	Kill the Unit (loss of data possible)
status Unit	Query the status of Unit
disable Unit	Disable Unit and then boot at system start time on request

```
$ sudo systemctl isolate multi-user.target
```

This example enables all the units required for multiuser mode without a GUI. For everyone still thinking in terms of run levels and SysV Init, systemd provides matching targets that simulate the legacy behavior. For example,

```
$ sudo systemctl isolate runlevel5.target
```

changes to a system state with a GUI. At boot time, systemd also enables the default.target, which is a symbolic link that points to another configuration file. On Fedora 15, this is still the runlevel5.target, although the developers will be moving to the systemd counterpart, graphical.target, in the near future.

systemd also adds a new command for gracefully shutting down and powering off the system:

```
$ sudo systemctl --force poweroff
```

The legacy shutdown and reboot commands still work; systemd translates them (see the “Legacy” box). If you pre-

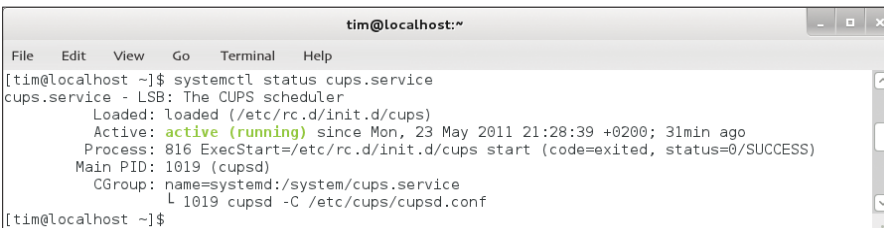


Figure 1: The systemctl status cups.service command gives you the details of the printing service.

lates all the services required to operate a multiuser system.

Fire Away

The complete configuration file is now saved as noisemeasurement.service in the /etc/systemd/system/ directory. This is where you should put all of your own configuration files; the system’s own configuration files are stored in /lib/systemd/system/. The following command,

```
$ sudo systemctl daemon-reload
```

tells systemd to parse the modified configuration. You can then finally start the new service as follows:

```
$ sudo systemctl start noisemeasurement.service
```

The systemctl shows that the command is running, but this time you need the status keyword instead of start (Figure 1). For an overview of the software commands, refer to Table 4.

Whodunit?

The Systemctl tool is useful in many situa-

tions. Typing systemctl without parameters returns a list of all active services (Figure 2). Pay special attention to the ACTIVE column, which shows whether a service is currently active or inactive, or whether an issue occurred trying to start it (maintenance). systemctl status Unit gives you more information on the status.

The isolate parameter tells systemd to create the units for a specific target:

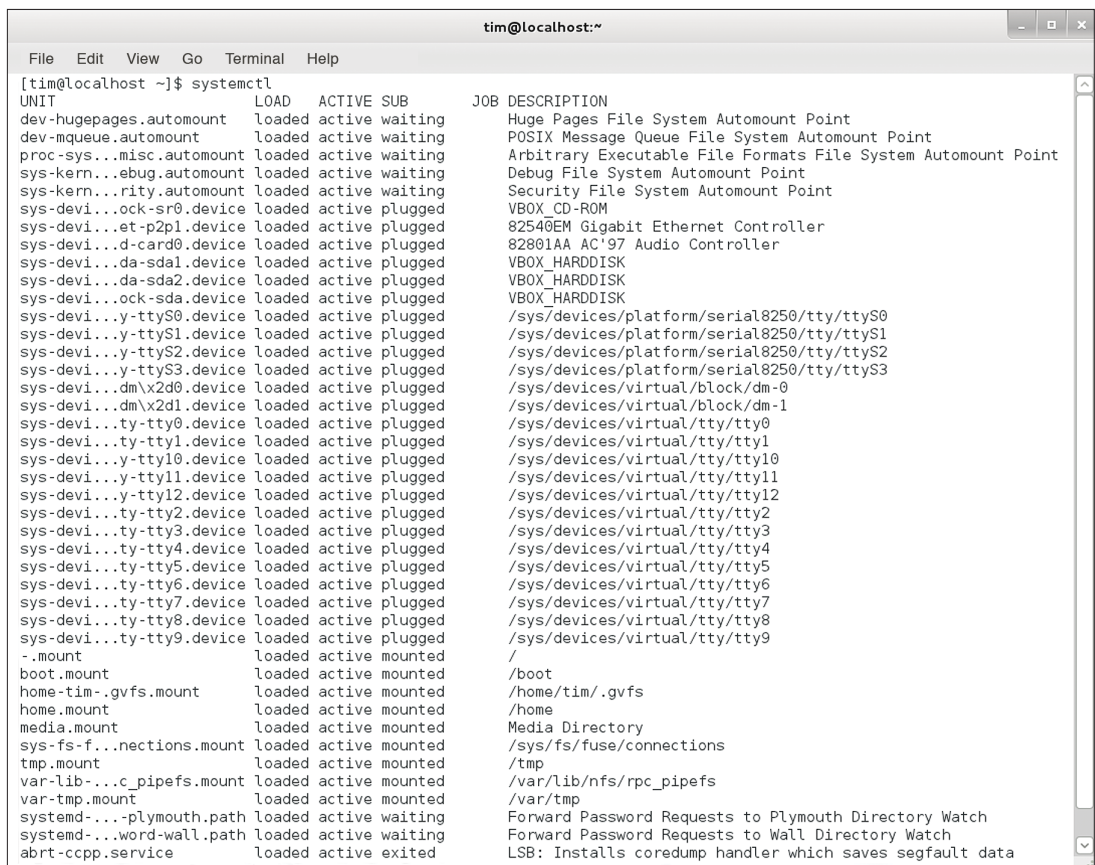


Figure 2: Systemctl displaying a list of existing services, their states, and descriptions.

fer a graphical interface instead of Systemctl, you can use Systemadm from the `systemd-gtk` package (Figure 3). `systemd` logs its own system start. `systemd-analyze blame` tells you which service took

how long to start (Figure 4), and the command

```
$ sudo systemd-analyze plot > results.svg
```

gives you a neat chart to illustrate. The resulting output is an SVG image in the `results.svg` file, which you can view in Inkscape or any capable browser (see Figure 5).

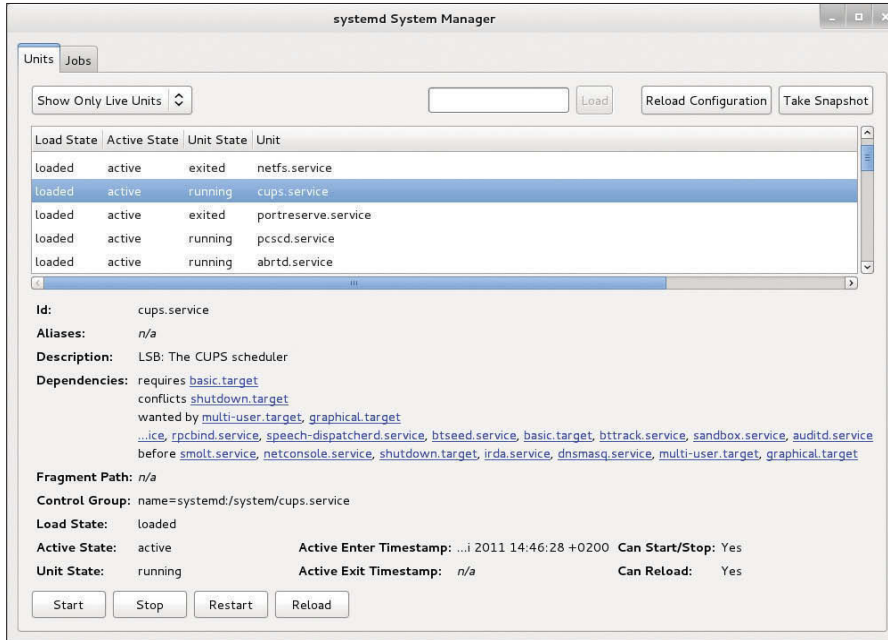


Figure 3: Systemadm lets you point and click to enable and disable units.

The Future

Although the configuration file in Listing 1 looks lean and mean at first glance, `systemd` is more complex than its competitor Upstart. If you want to understand the tool, you have no choice but to battle through many pages of text, which Lennart Poettering published bit by bit in his blog [5] and in the current 45 man pages [6].

`systemd` requires service maintainers to modify their software, thus causing additional programming – although the overhead is not too large. Administrators also need to say goodbye to the popular shell scripts. All of this promises a fast system boot. In a small test, the last Fedora RC booted about five seconds slower than Ubuntu 11.04, which uses the competitor Upstart, under comparable conditions.



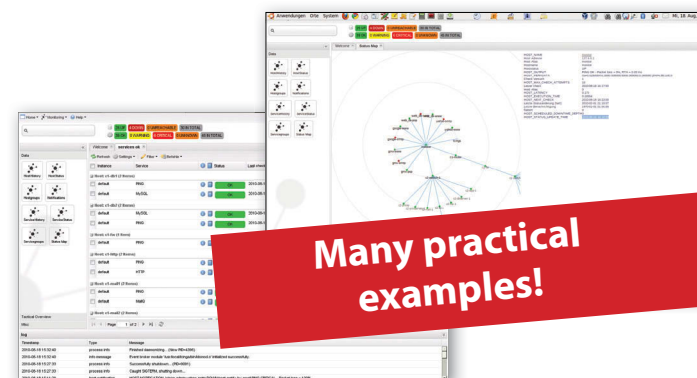
Online Training at Linux Magazine Academy

Monitoring with Nagios

Getting your IT under control the easy way:

- Web front-end
- Monitoring Windows/Linux/Unix
- Structuring the configuration
- Monitoring SNMP-components
- Nagvis, Grapher V2, and ND02DB add-ons

20%
off for current subscribers!



Many practical examples!

For more information and to sign up: academy.linux-magazine.com/nagios

```

tim@localhost:~$ systemd-analyze blame
3979ms NetworkManager.service
3090ms sendmail.service
2978ms udev-settle.service
2594ms cups.service
2302ms abrt-d.service
2263ms acpid.service
2230ms irqbalance.service
2220ms mcelog.service
2208ms lvm2-monitor.service
2201ms portreserve.service
2199ms iptables.service
2196ms iptables.service
2146ms sshd.service
1579ms cpuspeed.service
1500ms abrt-ccpp.service
1497ms mdmonitor.service
1479ms rpcbind.service
1459ms vboxadd-x11.service
1302ms rpcidmapd.service
1071ms nfslock.service
1059ms netfs.service
992ms udev.service
921ms systemd-vconsole-setup.service
701ms fedora-readonly.service
579ms pcsd.service
477ms hwclock-load.service
468ms remount-rootfs.service
456ms rpcgssd.service
451ms fedora-sysinit-hack.service
440ms bluetooth.service
416ms media.mount
415ms auditd.service
415ms console-kit-log-system-start.service
411ms fedora-sysinit-unhack.service
405ms system-remount-api-vfs.service
391ms system-user-sessions.service
349ms systemd-tmpfiles-setup.service
327ms fedora-storage-init.service
318ms dbus.service
318ms sandbox.service
293ms systemd-sysctl.service
    
```

Figure 4: systemd remembers which services took how long to start.

However, these results don't really allow any conclusions to be drawn about systemd: Upstart has been around for a

LEGACY

To retain compatibility with the SysV Init system, systemd parses the legacy init scripts, then it simply views them as additional sources of configuration files and converts the scripts it parses to units internally.

In a similar fashion, systemd parses and interprets other well-known configuration files, including `/etc/fstab`, using its entries as `mount` or `automount` units.

If you have written a start script for the legacy SysV Init and want to convert this into a service file, you might be interested in an option for running scripts to prepare this step before the program start. But, you can give up on this with systemd, or almost: The `ExecStartPre=` variable in the `[Service]` section gives you a workaround:

```

ExecStartPre=/bin/rm -f /var/log/measurements
    
```

systemd runs the command that follows the equals sign, which can be a shell script, before enabling the service that follows `ExecStart=`. Similarly, `ExecStartPost=` tells systemd to run the specified com-

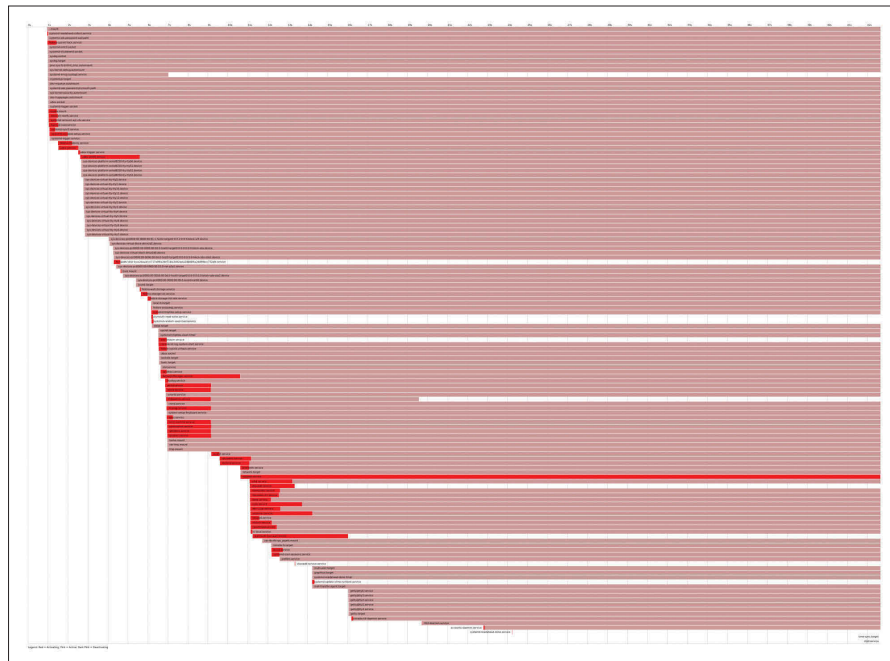


Figure 5: If needed, you can create a chart that displays the system start and the times required for it.

couple of years and has been optimized accordingly; systemd has just celebrated its first birthday and is still subject to constant change.

systemd is becoming more popular with distributors. Fedora 14 included systemd as an optional package; as of Fedora 15, it is used by default. Open-

SUSE, Debian, and other distributions are also considering their systemd options.

If you want to try systemd for yourself, links to packages for various distributions are on the systemd website – including some for Ubuntu, although Ubuntu's distributor, Canonical, still backs its own horse, Upstart.

Because of widespread support, systemd is likely to become the standard soon. Whether doing without flexible scripts and prescribing functions that are exclusive to Linux really are a step forward, remains to be seen. ■■■

INFO

- [1] systemd: <http://freedesktop.org/wiki/Software/systemd>
- [2] Autofs in the Debian Wiki: <http://wiki.debian.org/AutoFs>
- [3] New configuration files: <http://0pointer.de/blog/projects/the-new-configuration-files>
- [4] Overview of target units: <http://0pointer.de/public/systemd-man/systemd.special.html>
- [5] Lennart Poettering's blog: <http://0pointer.de/blog/>
- [6] systemd man pages: <http://0pointer.de/public/systemd-man/>
- [7] systemd for Developers: <http://0pointer.de/blog/projects/socket-activation.html>

REAL SOLUTIONS FOR REAL NETWORKS

FREE DVD
fedora 15 f
MULTIBOOT
4 LIVE VERSIONS OF FEDORA 15!
ADMIN Network & Security
We compare 7 IPsec options
NEW!
ADMIN
Network & Security
Secure Networking
MEET NAGIOS CREATOR
ETHAN GALSTAD
Will your IPsec work with other implementations?
Nagios IPMI
Monitor your hardware
Going virtual:
■ Put your apps in the cloud with Google's App Engine
■ Virtual updates and Citrix Provisioning Services
PAC automator
Manage local and remote connections
ADMIN Issue 04
US\$ 15.99
CANS 17.99
0 74470 86640 4

**FREE CD
Inside:
4 Live Versions
of Fedora 15!**

Each issue delivers technical solutions to the real-world problems you face every day.

Learn the latest techniques for better:

- network security
- system management
- troubleshooting
- performance tuning
- virtualization
- cloud computing

on Windows, Linux, Solaris, and popular varieties of Unix.

ORDER ONLINE AT shop.linuxnewmedia.com

Klaus Knopper answers your Linux questions

Ask Klaus!

By Klaus Knopper

Dynamic Doldrums

? Hi. I have decided to part with Microsoft, and I want to help the Linux community without going through a hundred screens of information and signing in blood (i.e., Ubuntu's "community" site). I couldn't get any information on how to deal with my problem scanning the net with Google on any of the Linux sites. I would think someone would have run across this issue and there would be a fix page. So, I thought maybe you knew a "better" way.

I have been loading Linux of one version or another on a partition on my Microsoft machines, without losing my OEM-issue operating system. I now have an HP G-62-227CL notebook straight from Hewlett-Packard loaded with all its "stuff." I got my tools together to squeeze in another partition and load Linux – no joy, it wouldn't work. So, I tried several tricks to no avail. After a little investigation, I discovered that I have a "dynamic" disk, and it won't allow me to install another OS (Ubuntu in this case). I'd have to wipe the disk and re-install everything.

Right, not in this lifetime, particularly with Windows 7 and all the other junk, without knowing if it will even work on a static disk. Well, I finally discovered a workaround. I installed an external hard drive with a USB cable and loaded Ubuntu on the external drive. It works, but to boot, I have to escape during the loading sequence into the BIOS and select the boot menu and select the external drive to boot to Linux. A real pain, but it works for now.

Best, Mike

💡 I may have two helpful add-ons to your solution.

You might, first of all, be able to set the BIOS to boot from the USB disk first, when it is present. For many computers, this setting is NOT in the *Boot priority* menu, but in the *hard disks* settings. You

can change the setting of the internal disk to be seen as the SECOND drive and set the USB disk (which probably has to be plugged in at that time for the BIOS to recognize it as "second hard drive") as the FIRST drive. After saving this hard disk setting, the boot priority menu will automatically display the USB drive instead of the internal hard disk in the chosen sequence. Then, after leaving the BIOS with *Save and Exit*, your computer should boot from the USB drive if it is plugged in. I can't promise that Windows won't be confused by the new setting, though.

Second, for Windows "dynamic disks," forums recommend some proprietary toolkits running from CD for conversion of dynamic disks to "normal" disks or partitions, but because I have doubts about their legality, I cannot recommend any.

Backing up all data and reinstalling Windows without the "dynamic disk" feature, then installing Linux on a second partition, still seems to be the most reliable way if you want to have a dual-booting system.

KVM Virtualization

? When starting the KVM virtualization engine, I get a message saying that there is no file `/dev/kvm` present. What's missing?

💡 KVM is a hardware-accelerated open source virtualization engine, which allows you to run Linux, Windows, and other operating systems inside a Window (or switchable full screen) under Linux. I use it for testing Knoppix and other distributions because it is very fast and has great scripting capabilities, supports virtual networks and network bridging, and is fully supported by the vanilla Linux kernel.

To use the hardware-accelerated feature (meaning that your CPU can run programs from the guest system natively with almost no overhead or performance loss), your CPU needs the virtualization extension. You can find out if it's present with:

```
egrep '^flags.*(vmx|svm)' /proc/cpuinfo
```

If this command produces no results, you can still run software-only virtual-



KLAUS KNOPPER

Klaus Knopper is the creator of Knoppix and co-founder of LinuxTag expo. He currently works as a teacher, programmer, and consultant. If you have a configuration problem, or if you just want to learn more about how Linux works, send your questions to: klaus@linux-magazine.com

ization with Qemu, but it will be very slow. To activate hardware acceleration for KVM, if supported by your hardware, you need to load either `kvm_intel` or `kvm_amd`, depending on your CPU type. Loading the wrong one will just fail with no bad effects:

```
sudo modprobe kvm_intel
sudo modprobe kvm_amd
```

After the module has been loaded, `udev` will create the `/dev/kvm` character device, and KVM should now start in full speed.

VirtualBox and Knoppix

? In Knoppix 6.7.0, I get an error message as soon as I start a virtual machine with VirtualBox. Starting the kernel modules with

```
sudo /etc/init.d/virtualbox start
```

also fails with an error message about an “invalid module format.” What’s wrong?

💡 It’s all my fault. ;-)

I forgot to recompile the `vboxdrv` kernel module, which is supposed to be loaded by the VirtualBox init script you mentioned above. The error message is somewhat misleading, but it indicates that the compile environment of the module is incompatible with the one the kernel was compiled with.

For Knoppix 6.7.0, you can recompile the VirtualBox kernel module on the fly with

```
sudo dpkg-reconfigure virtualbox-dkms
```

and load the module with:

```
sudo /etc/init.d/virtualbox start
```

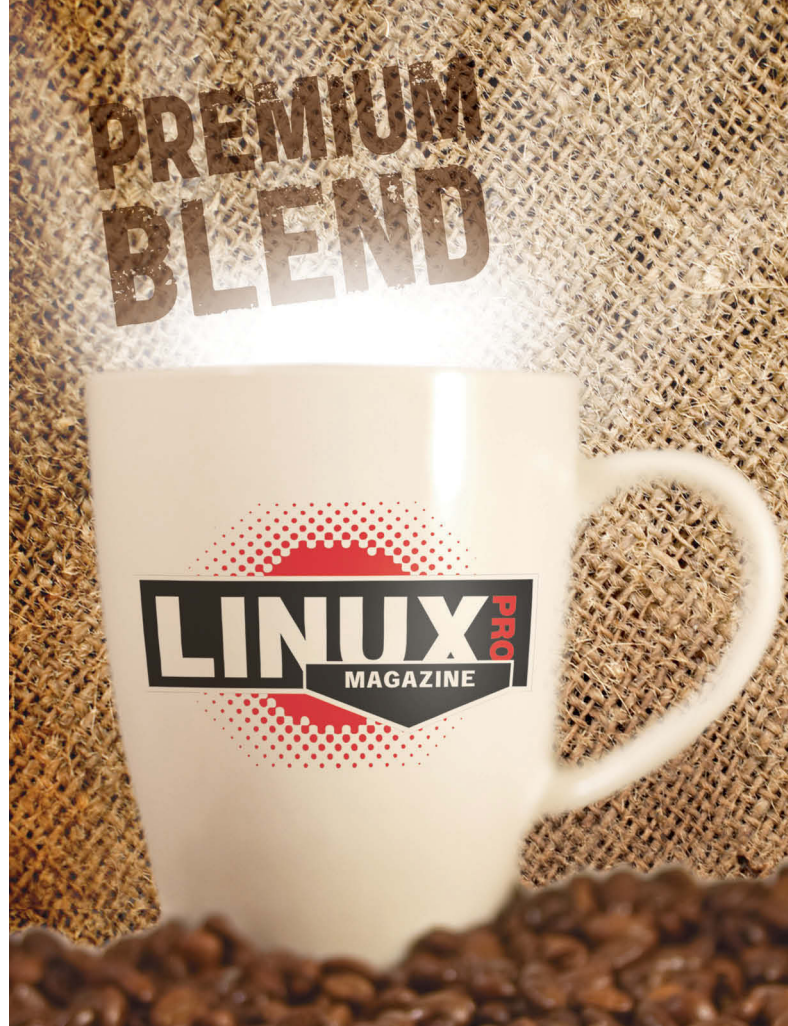
If you are using the “persistent image” feature of Knoppix after installation to a flash drive of at least 8GB, the repaired module is fixed permanently.

? When starting the Knoppix DVD in VirtualBox, whether from a real DVD drive or using an ISO image, Knoppix hangs at the “starting udev...” message. After waiting for about 10 minutes, the boot process continues. What’s wrong?

💡 I have been debugging this weird effect for a while now. It only seems to happen in VirtualBox – no other virtualization is affected.

It seems that VirtualBox does something with the virtual text console when any kind of text output is in progress (like the “Progress bar” in Knoppix, possibly console output in other distros, too) that keeps the `udev` daemon from starting; that is, the `udev --daemon` just sits there and does nothing until some kind of timeout occurs.

This looks like a strange error in the virtualization engine to me, and I have not seen this behavior on real hardware



LINUX PROS READ LINUX PRO

Enjoy a rich blend of tutorials, reviews, international news, and practical solutions for the technical reader.

Subscribe now
3 Issues

\$3

GET IT
NOW!
SAVE ON
DELIVERY TIME
WITH OUR
PDF EDITION

Print or PDF Edition

shop.linuxpromagazine.com/trial

yet. Luckily, there is a workaround: Booting with the kernel command line

```
knoppix splash
```

disables the progress bar (it is supposed to load a splash picture instead, if one is present in the initial ramdisk), and this makes udevd start immediately under VirtualBox.

I will change the progress bar in Knoppix 6.7.1 to be single-threaded, which will also make the VirtualBox problem go away without the “splash” workaround.

Mageia Play

? Klaus: I was pleased to receive the Mageia 1 DVD included with the September 2011 issue, but a little disappointed to find that my attempt to install it on my (admittedly ancient) IBM ThinkPad T22 (900MHz PIII with 512MB memory) failed.

This is because the mkinitrd step times out before it is able to complete. Apparently, this is because the kernel used while running the installer is compiled with CONFIG_HZ_1000, which slows things down too much on older machines.

I didn't find a step-by-step solution listed on the Internet, so here, for anyone experiencing the same problem, are the steps I followed after the initial installation failed:

1. Boot from the DVD and select *Rescue system* from the menu.
2. Select the *Mount your partitions under /mnt* from the rescue menu.
3. Select *Go to console* from the rescue menu
4. At the shell prompt, type:

```
mount --bind /proc /mnt/proc
mount --bind /sys /mnt/sys
cp -ar /dev /mnt/dev
chroot /mnt
cd /boot
/sbin/mkinitrd -f
```

```
initrd-2.6.38.7-desktop-1.mga.img
2.6.38.7-desktop-1.mga
exit
rescue-gui
```

5. Select *Reboot* from the rescue menu and reboot from the DVD.
6. Select *Install Mageia 1* from the initial menu.
7. Complete the language and license steps.
8. Select *Upgrade Mageia (Official) 1*.
9. Complete the keyboard, additional media, summary, and update steps, finally clicking the *Reboot* button.

Essentially, this process builds the initrd image, and the “upgrade” from the DVD completes the bootloader installation.

Note that attempting

```
mount --bind /dev /mnt/dev
```

fails – hence the use of cp instead.

Following successful installation it appears that more suitable kernel packages are available for download. Regards, David

? Thanks for your mini-remastering how-to for Mageia. For those who don't know: Mageia is a fork of Mandriva Linux run by a non-profit organization [1].

I believe that mount --bind might have failed just because the mount command from the small initial rescue system does not support all standard options. Typing cp -a /dev into the chroot environment will create the necessary device files there, which should not require too much space.

PS2 Port Failure

? From a reader of the Spanish Linux Magazine:

Hi, I have a problem with my Ubuntu 10.10. First, I installed Windows 7 Professional and later the 64-bit version of Ubuntu 10.10 on a different partition with its corresponding swap. The Ubuntu installation worked flawlessly: It

detected all the devices, installed all the updates, etc. Unfortunately, after some time, the port PS2 failed. I know it is not the physical device because they work on Win7.

Under Ubuntu, they freeze and don't allow you to use them in any way. I have tried removing and reinstalling them, but there is no change. The system works normally if I connect only USB devices.

My motherboard is an Asus P7P55D PRO. Thanks.

? This doesn't seem to be a commonly known problem with this board. However, PS2 claims an interrupt that is difficult to share with other devices, which may cause side effects with other installed devices. Can you have a look into the output of dmesg (using a USB keyboard, in case your PS2 stops working when the problem occurs)?

When PS2 suddenly drops dead, you should see a message like “lost interrupt” (usually something in the range of interrupt 12). This can be related to BIOS settings as well as different interrupt-handling schemes in Linux, which could explain why you have not seen this in Windows yet. Or maybe you were just lucky. :-)

Table 1 shows the boot options I would recommend trying for working around interrupt problems. If one of these stabilizes PS2, you can add the option to your bootloader's kernel APPEND line to make it permanent. ■■■

INFO

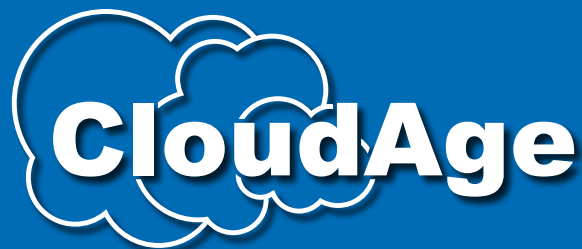
- [1] Mageia: <http://www.mageia.org>
- [2] Advanced Configuration and Power Interface: http://en.wikipedia.org/wiki/Advanced_Configuration_and_Power_Interface
- [3] Advanced Programmable Interrupt Controller: http://en.wikipedia.org/wiki/Advanced_Programmable_Interrupt_Controller

TABLE 1: Boot Options for Interrupt Problems

Boot Command	Description
<code>linux acpi=noirq</code>	Use ACPI [2] for everything but interrupts.
<code>linux noapic</code>	Don't use the APIC [3].
<code>linux nolapic</code>	Don't use the CPU-internal local APIC.
<code>linux pci=irqmask=0x0e98</code>	Mask out some potentially problematic interrupts.

GOING TO THE CLOUD?

How will you get there?



The cloud portal for IT
specialists — up close,
all technical, all cloud.

cloudage.admin-magazine.com

Powered by:





Tag, retrieve, share: phTagr

Tag Line

Use phTagr to publish and share images on the web with consummate ease. *By Dmitri Popov*

Piwigo and Gallery are probably the most popular open source web-based applications for hosting image galleries on your own server, but they are not the only fish in the sea. Take phTagr [1], for example. This application has all the essential features you would expect from a decent

image and video sharing application wrapped into a lightweight and user-friendly interface. The application's real strength, though, lies in its tagging and searching capabilities, which make it a solid solution for maintaining a web-based image and video archive.

Getting Started

Like Piwigo and Gallery, phTagr runs on the Apache/MySQL/PHP stack. So, to deploy the application, you need a server with the required components. Or, you can use the XAMPP software [2], which offers all the necessary components in one neat package. Before you download and install phTagr, you might want to tweak the maximum upload limit specified in the `php.ini` file. The default 2MB limit is too low if you plan to upload high-resolution images. To increase the limit, open `php.ini` in a text editor as root (in XAMPP, the path to the file is `/opt/lampp/etc/php.ini`), locate the `upload_max_filesize` parameter, and adjust its value – for example:

```
upload_max_filesize = 25M
```

While you're at it, locate the `post_max_size` parameter and adjust its value, too:

```
post_max_size = 25M
```

To enable the changes, restart the Apache server.

phTagr relies on ExifTool and ImageMagick tools for processing images and their metadata, so you should install these tools on your server. On Debian and Ubuntu, this can be done by executing

```
apt-get install \
    libimage-exiftool-perl \
    imagemagick
```

DMITRI POPOV

Dmitri Popov has been writing exclusively about Linux and open source software for many years, and his articles have appeared in Danish, British, US, German, and Russian magazines and websites. Dmitri is an avid amateur photographer, and he writes about open source photography tools on his Scribbles and Snaps blog at scribblesandsnaps.wordpress.com.

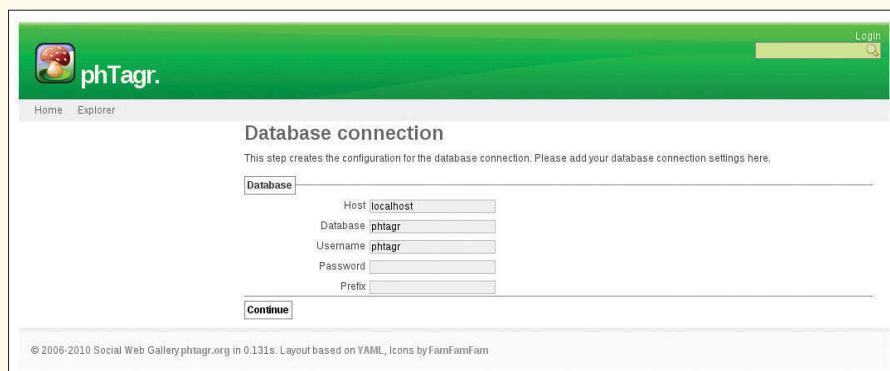


Figure 1: phTagr installation is a simple three-step procedure.

as root. You also must create a separate MySQL database; in XAMPP, you can use the phpMyAdmin tool to do so.

With all the pieces in place, you can install phTagr. From the website, grab the latest release, unpack the downloaded archive, and move it to the document root of your server (in XAMPP, it's /opt/lampp/htdocs/). Now point your browser to <http://127.0.0.1/phtagr> (replace 127.0.0.1 with the actual address of your server) to set up phTagr. The setup procedure consists of three simple tasks: specifying a database connection, creating an administrator account, and configuring external programs (Figure 1). The last step is optional; you need to configure additional programs only if you want to use phTagr to manage videos.

Once phTagr is up and running, you should tweak its default system settings. To do this, click on the *System* link in the upper-right corner of phTagr's page. In the *General* section, provide a name and description for your phTagr installation. If you want to use Google Maps with phTagr, you have to enter your Google Maps API key in the *External Programs* section (Figure 2). If you don't have a key, you can obtain one by signing up for the Google Maps API [3].

Setting Up Users and Groups

Like any photo sharing application worth its salt, phTagr can handle multiple users and organize them in groups with different access privileges, so your next task is to populate it with users and groups. To add a user, click on the *Preferences* link in the upper-right corner, switch to the *User Accounts | Add user* section, specify the required user info, and press the *Create* button. You can then specify the user's role (you can choose between user, system operator, and administrator), as well as an optional upload quota and expiration date (Figure 3). When editing a new or existing user, you can also specify an external directory in the *External Paths* sub-section. This directory acts as a place the user can drop off image files and then import them into phTagr. This feature is indeed nifty, and you can use it to collect and import photos from multiple users. When you add images from an external directory to phTagr, the application doesn't actually import the files; rather, it links to them. Thus, you can keep your images separate from the application and save a lot of time when uploading large photos. This approach also allows you to override the upload limit.

phTagr provides basic support for groups, which are useful for organizing and managing users. To set up a group, switch to the *Groups | Create group*, then give the group a name and description. Using the available options, you can make the group hidden or shared (Figure 4). Additionally, you can enable moderation for new group members. To add existing users to the created group, click on the group's name in the *Groups* list, enter the desired username in the *Username* field in the *Add user* section, and press *Add*. Finally, the *Access Rights* section lets you configure the default access rights. Here, you can define who can view and download images (e.g., everyone, registered users, or group members), as well as who can add tags and edit metadata. To assign a specific group to all incoming images, select the group you want from the *Default image group* drop-down list.

Using phTagr

Once you have configured phTagr's settings to your liking, you are ready to populate the application with images. phTagr of-

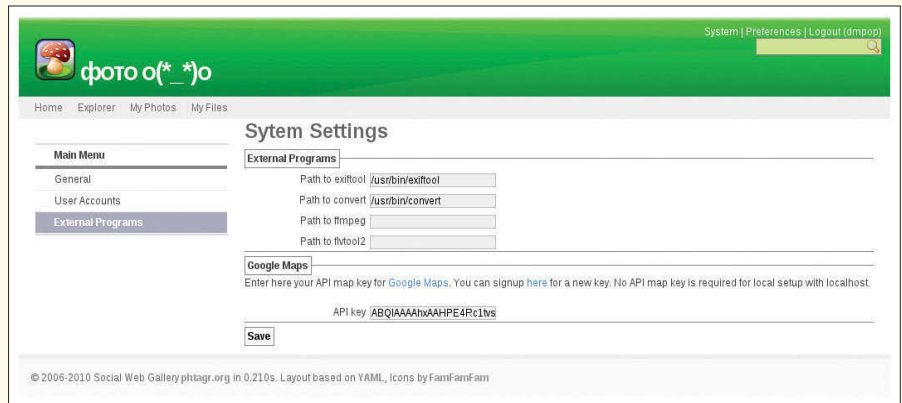


Figure 2: Configuring phTagr's system settings.

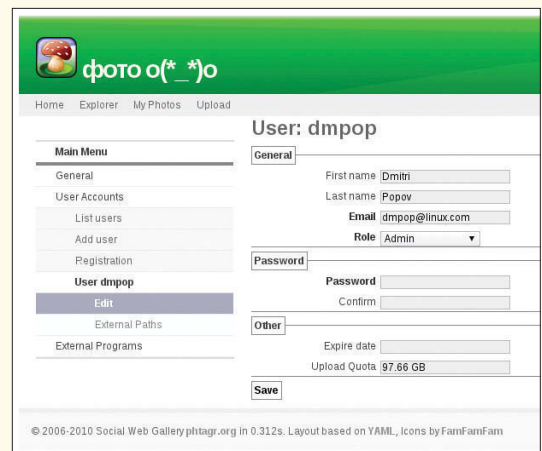


Figure 3: Adding a user.

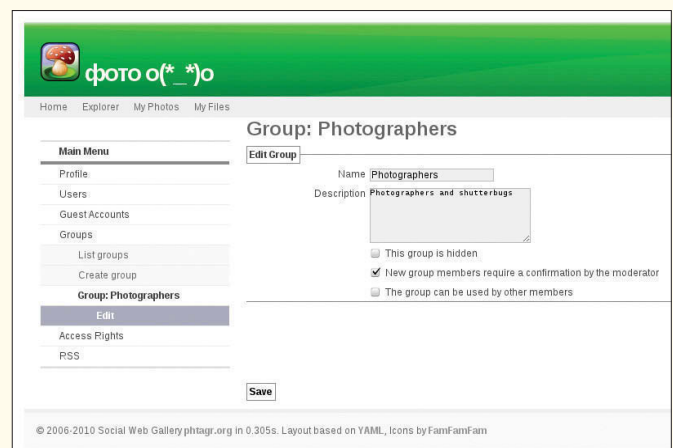


Figure 4: Creating a group.

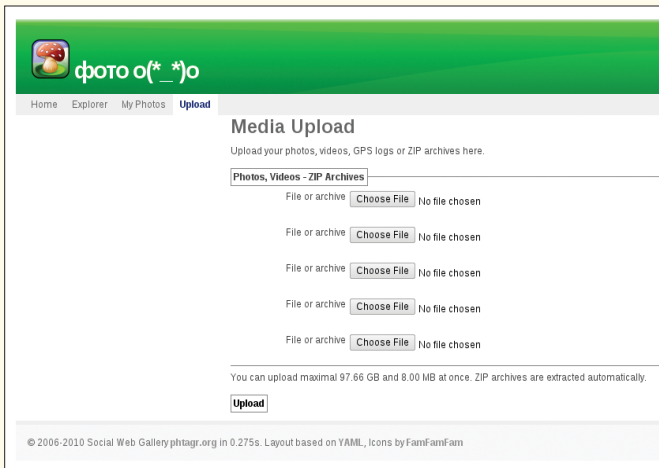


Figure 5: Using the Upload form, you can download images, videos, and ZIP archives containing multiple files.

fers two ways to do that. The most straightforward approach is to use phTagr's upload form, which you can access by clicking on the *Upload* link in the navigation bar (Figure 5). Here, you can upload multiple images individually. Alternatively, you can pack all the desired files as a ZIP archive and feed it to phTagr. This way, you can upload multiple images in one fell swoop. Keep in mind, though, that the size of the ZIP archives must not exceed the maximum download size value specified in the `php.ini` configuration file.

Additionally, if you configured external directories, you can add photos using the built-in file browser that is accessible via the `http://127.0.0.1/phtagr/browser`. To make the browser easier to access, just go to the *Preferences* page, switch to the *Profile* section, and enable the *Show advanced file browser* option. This action replaces the *Upload* entry in the navigation bar with *My Files*.

To evoke the file browser, click on the new menu entry (Figure 6), then navigate to the directory you want, select the desired photos, and press *Import* to add them to phTagr. You can also use the file browser to download selected files to your machine or to delete them.

One of phTagr's best features is the ability to handle metadata stored in images. So,

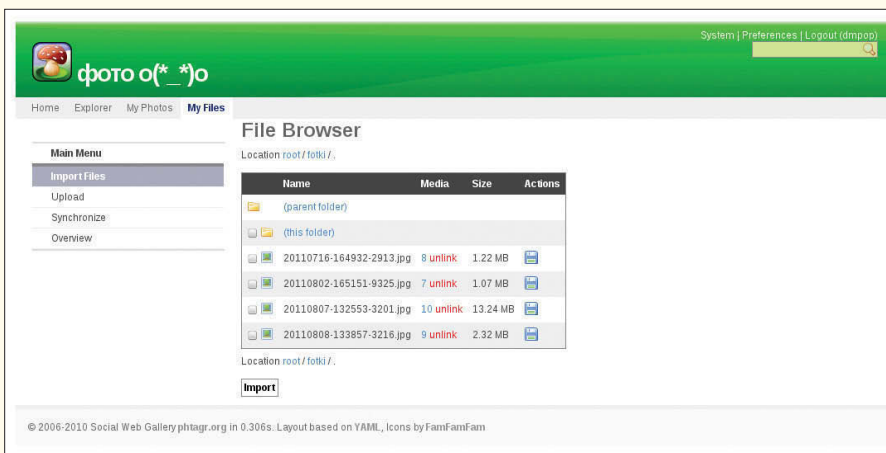


Figure 6: With the file browser, you can add images from external directories.

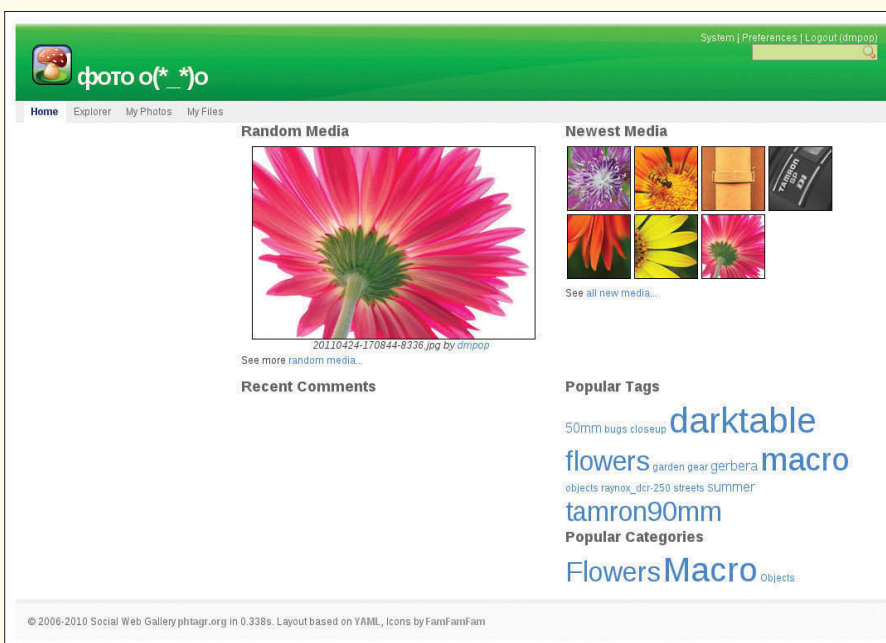


Figure 7: phpTagr's Home page.

if you tagged your photos and added geographical coordinates to them in a desktop photo management application like digiKam, all of the metadata will be imported into phTagr along with the images.

phTagr features two key pages: Home, Explorer, and My Photos. The Home page displays a large preview of a random image, as well as thumbnails of the recently added image files (Figure 7). The Popular Tags section displays a tag cloud culled from tags in the imported photos, whereas the Recent Comments section shows the latest comments. The Explorer page allows you to browse your images, and it offers several handy search and editing tools in the *Main Menu* sidebar (Figure 8).

The *Tags* list, for example, lets you quickly filter photos by tag. The number next to each tag in the list indicates how many images have the specific tag. Click on the tag you want, and phTagr displays matching images. Instead of tags, you can use the categories listed in the *Categories* section.

The *Advanced Search* feature allows you to construct more advanced search queries. Here, you can include and exclude the tags you want, and you can search for images with a specific date range or location. The *Advanced* subsection lets you search images by name, group, and visibility. The My Photos page is practically identical to the Explorer page, but it only shows images belonging to you.

To view an image, you can simply click on its thumbnail. Besides the image itself, the image page provides image info divided into two sections: *General* and *Media Details* (Figure 9). The latter section contains basic image and Exif data, such as file size and creation date, view count, ISO, aperture, and shutter settings.

The *General* section shows the date, tags, categories, and location of the image. Several so-called action buttons let you edit the tags and access rights of the currently viewed image or download the image and view it on the map (provided the image contains geographical coordinates).

When you edit the info of an image that comes from an external directory, the changes are stored in the database back end. To save them to the image, you need to click on the *Synchronize* action button.

Instead of editing tags, access rights, and other information for each individual image, you can use phTagr to process multiple images in one go. To do this, select the desired images on the Explorer or My Photos page and use the *Metadata* and *Access Rights* sections at the bottom of the page to specify the desired changes. When you are done, press *Apply*. To synchronize the changes made to images from external directories, switch to the file browser and run a sync operation in the *Synchronization* section.

Wrap-up

phTagr might not be as feature-rich as Piwigo or Gallery, but it does have several handy tools that make it an excellent solution for sharing and managing images and videos. The ability to manage images stored in external directories and easy-to-use yet powerful filtering and searching capabilities are just a few of phTagr's highlights. So, if you are in the market for an application that can help you put your images and videos on the web, you ought to give phTagr a try.

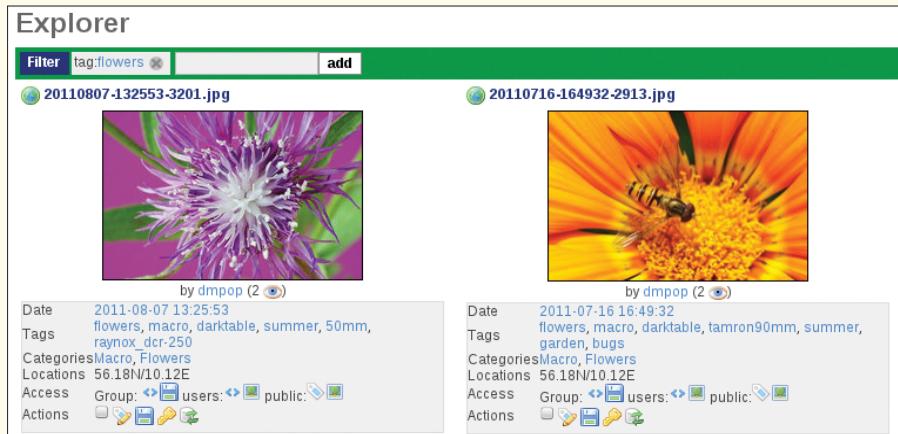


Figure 8: The Explorer page sports search and filtering tools.

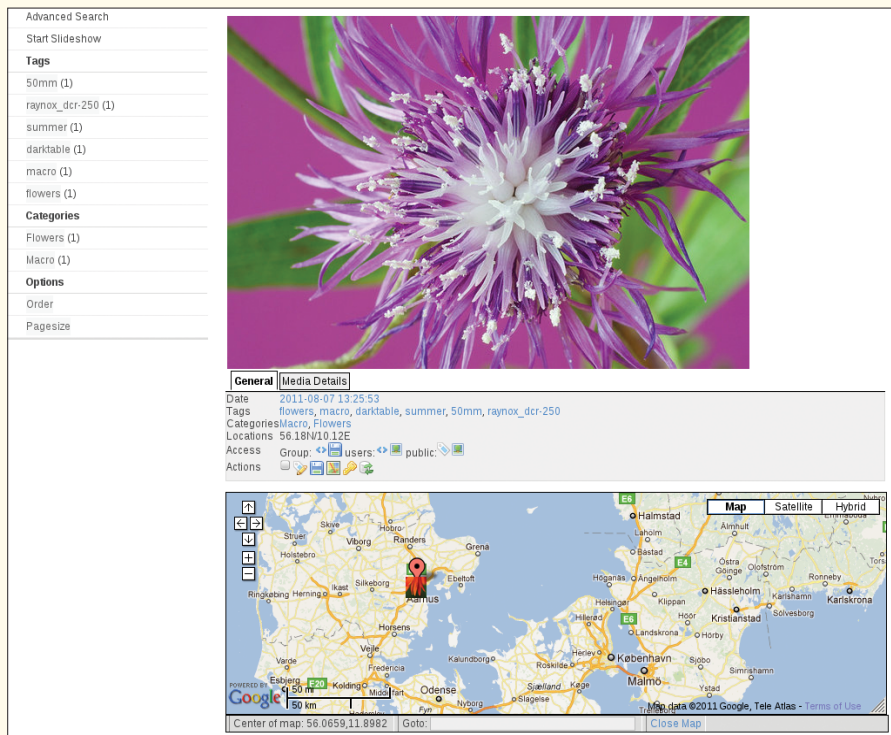
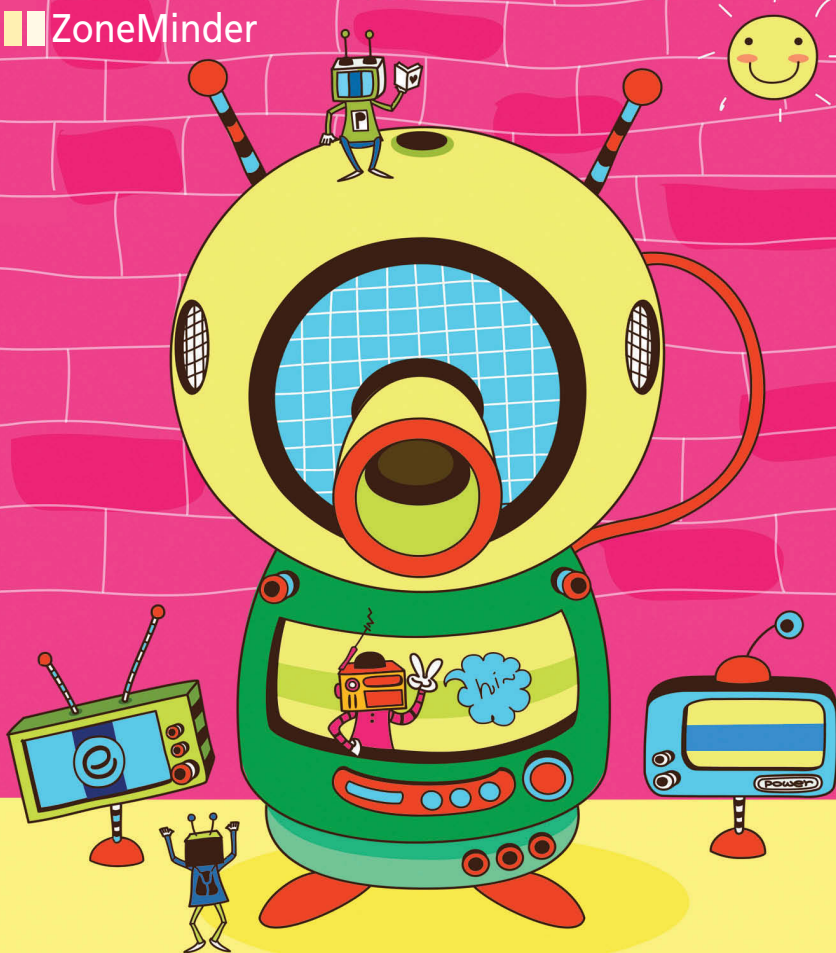


Figure 9: Viewing an image and editing its info.

INFO

- [1] phTagr: <http://www.phtagr.org/>
- [2] XAMPP: <http://www.apachefriends.org/en/xampp.html>
- [3] Google Maps API: <http://code.google.com/apis/maps/signup.html>



Video monitoring with ZoneMinder

Magic Eye

Vandalism, property damage, and theft are serious worries for many people, but Linux and free video monitoring software give you a fair chance of catching the culprits.

By Erik Bärwaldt

Stolen bikes, scratched cars, graffiti on the walls – all directly in front of your own home. When police investigations into these types of crimes fail because of a lack of conclusive evidence, a video monitoring system can work miracles. Effective monitoring of strategic points doesn't need to cost an arm and a leg. With just a Linux PC, the free ZoneMinder [1] control software, and some standard webcams, you can take control for just over a hundred dollars.

Preparation

The utility value of a video monitoring system hinges on the quality of the cam-

eras you use. To have any chance of video sequences being accepted as evidence in court, they need to have high resolution for a crisp image when you zoom in to the important details. As you can imagine, not every camera is suitable for video monitoring. Also, you need to take location into consideration. If you want to monitor outdoor areas, the camera must be designed for outdoor use; you don't gain anything if a rain shower kills the device. If you occasionally need to change the area of observation, you need a pan/tilt camera with motors that move the lens in two axes to monitor different areas. Pan/Tilt/Zoom (PTZ) cameras support even more flexible use. If needed, they let you zoom in to the area you are monitoring.

If the areas to monitor are in dark areas, or if you need to monitor at night, you also need an infrared light source. Expensive, professional surveillance cameras often include several infrared light-emitting diodes that give you a range of up to 10 meters in poor weather conditions. If you are using the far cheaper and typically smaller webcams to monitor an interior location, or your surveillance camera doesn't have infrared diodes, special infrared spotlights are available as external components.

Linux Support

Although dedicated pan/tilt surveillance cameras all look similar from the outside, they have major technical differences. The lenses are practically identical in all of the models available today, except for the pixel count, but many cameras only support other operating systems and do not work on Linux. Cameras that rely on a TV interface card in the computer (TV tuner) and analog S Video cameras can be used in combination with ZoneMinder in many cases, assuming you have Linux drivers for the TV or capture card. However, you should first check out the program's compatibility list [2] – especially if you intend to deploy several different cameras in parallel both indoors and out.

Missing drivers for Linux can (still) cause issues with popular USB webcams. In particular, you will want to steer clear of older cameras with lower resolutions. Cameras of this kind are not typically capable of producing useful evidence. Thanks to various projects, web-

cam support under Linux has improved significantly in the past few years. If you use a USB camera, your first port of call should be the GSPCA [3] and UVC Video [4] projects, where you can discover whether your camera candidates actually work with Linux. Because the kernel has included both drivers since version 2.6.27, the cameras rated as compatible by the projects will work with Linux without the need for manual intervention.

Finally, you should also make sure your monitoring PC has sufficient hard disk capacity. High-resolution video recordings with a high frame rate and large color space consume a huge amount of space.

Software Installation

ZoneMinder is a highly modular application with a web-based control interface. Correspondingly, you need to fulfill a number of conditions before you can use the program. Besides Video4Linux to support cameras and TV cards and the Apache web server with MySQL, PHP, and Perl, you also need to install the FFmpeg and libjpeg packages for recording still and moving images.

Fortunately, most popular Linux distributions and their derivatives now include packages for ZoneMinder that let you install the program directly from repositories using Synaptic, MCC, or YaST, which typically resolve any dependencies. I tested the system on Ubuntu 10.10. A special Live CD variant of ZoneMinder is out there, but I didn't test it exhaustively because it is based on Xubuntu 8.10, which is no longer maintained and which doesn't have drivers for many of the current crop of webcams and video capture cards because of its age.

To simplify installation and configuration of the video surveillance system, which is somewhat complex because of the many packages and codecs it requires, the ZoneMinder forum offers a script [5] that both reliably automates this task and modifies the memory configuration. To begin, copy the script to the clipboard and then save it with an editor. To make it executable, run `chmod` then become root and run the script, which makes all the necessary changes in a single pass.

Whenever you boot your system from now on, the Apache web server will automatically launch; after opening a web browser, you can view the ZoneMinder interface by typing `http://127.0.0.1/zm` in your browser's address bar.

Getting Started

When first launched, the software comes up with a fairly spartan interface that you still need to configure (Figure 1). Pay attention to the value shown in the *Disk*: field. This gives you the currently occupied disc space as a percentage. If you intend to operate multiple cameras, this value needs to be as low as possible to prevent the high-resolution recordings from consuming all of your disk space, even at acceptable frame rates.

If you have attached some cameras to the system, you now need to integrate them with ZoneMinder. I tested the system with three different USB webcams by Logitech

SURVEILLANCE CAMERAS: Tips and Tricks

To use ZoneMinder effectively, you need to choose both your cameras and their locations carefully. Because limited storage capacity does not allow you to store all surveillance data, most surveillance cameras don't start to record until they detect some movement. For this reason, you will want to set up your cameras to monitor a quiet location.

In outdoor use, you need to avoid trees or bushes covering a major part of the surveillance area. Both will move in the wind and trigger events all the time. If keeping unimportant moving objects out of the camera's surveillance range is not an op-

tion for you, then you will need to define your own zones in ZoneMinder (see the "Alarm Zones" section).

Surveillance cameras should never monitor an area at a very acute angle from above; people standing in this area will be very hard to identify in a bird's eye view. Instead, the camera should monitor from an elevated position to the side.

If you use hard-wired cameras, make sure the wires are not easily noticeable and reachable by unauthorized persons. Cutting the wire will disable the surveillance camera, leaving ZoneMinder in the dark.

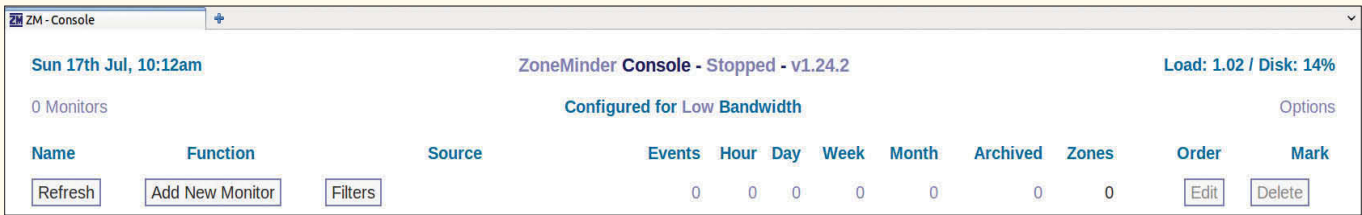


Figure 1: The spartan web interface after launching ZoneMinder.

and Chicony. To introduce the cameras to ZoneMinder, click the *Add New Monitor* button in the main window. The program now opens another window with four horizontal tabs at the top. The *Source* tab (Figure 2) has the hardware parameters you need to modify.

ZoneMinder makes the setup process easy thanks to a utility that reads the camera's technical data. To do this, enter the following command in a terminal window (Figure 3):

```
$ zmu -d Device-Name -V 1 -q -v
```

The device name is typically `/dev/videoX`, where *X* is a number starting from zero. The program now shows you a list of the device's specifications. To discover the parameters you need, first look for the *Window Attributes* section and note the values in the *Width* and *Height* lines for resolution in pixels. In the *Picture Attributes* section,

look at the *Palette* line and take note of the color palette. Finally, at the end of the program output, you will find the *Format* entry that lists the TV standard (NTSC, PAL, SECAM).

Armed with this information, you can make the required entries in the ZoneMinder *Source* configuration window. Make sure the entry in the *Capture Method* field reads *Video For Linux version 2*; this supports more or less any recent camera. After completing the entries, press the *Save* button,

Operating Modes

Next, go to the *General* tab (Figure 4), where you can type an intuitive name for the camera location in the *Name* text box before defining the device's operating mode. Typically, you do not need to change the *Source Type* entry of *Local* for webcams. For IP cameras that you connect to your intranet, you will need to select *Remote* here instead. The *Function* option *Monitor* shows the video signal but doesn't record anything. Choosing *Modect* enables motion detection and throws an event when ZoneMinder registers a movement. Recording starts automatically in this case and does not stop until ZoneMinder decides that the motion has finished. The next movement in front of the camera generates another event and starts recording again.

The third option, *Record*, starts a permanent recording, independent of mo-

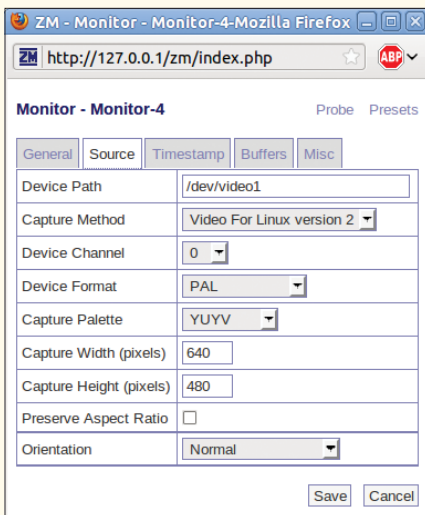


Figure 2: The ZoneMinder interface lets you quickly complete all the necessary settings.

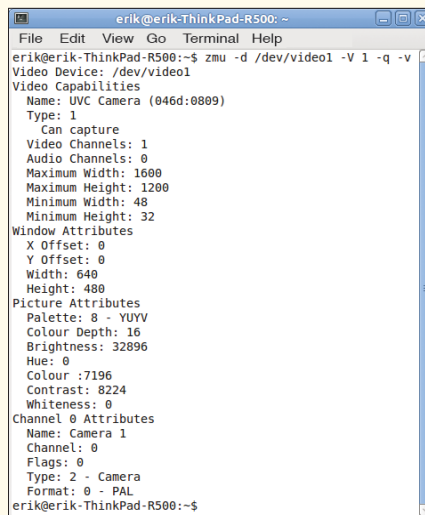


Figure 3: Discovering the correct settings for a camera.

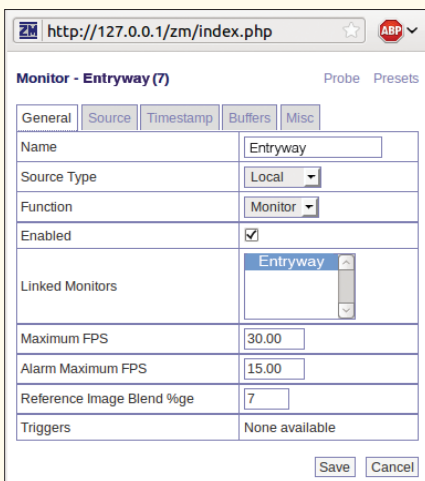


Figure 4: The setup dialog.

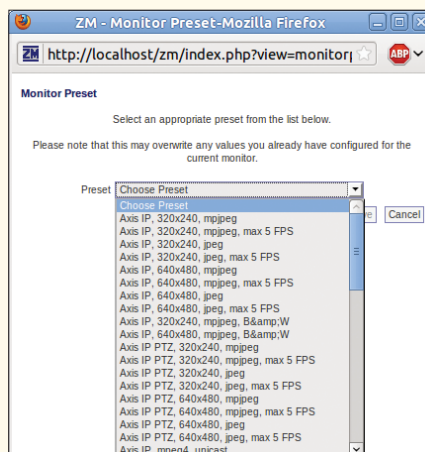


Figure 5: Presets facilitate the PTZ and IP camera setup.

IF YOU USE LINUX, YOU NEED LINUX MAGAZINE!

SAVE UP TO 85%

Linux Magazine keeps the emphasis on real-life, practical techniques and delivers a premium blend of technical articles.

Learn about the latest products, tools and strategies for sys admins, programming in the Linux environment, advanced desktop techniques, and much, much more!

GET IT NOW!
SAVE TIME ON DELIVERY WITH OUR PDF EDITIONS (DVD NOT INCLUDED)



FREE DVD

Knoppix 6.7
The world's greatest Live distro

BRAND NEW RELEASE!

LINUX

MAGAZINE

OCTOBER 2011

Exploring the new Linux kernel

Linux 3.0

- OpenVAS—look for security problems with this flexible assessment framework
- Peering inside Knoppix
- Adding e-commerce to your Drupal site
- 4 fine Dropbox alternatives

STUNNING ALTERNATIVES

FREE DVD
With each print issue



Get 3 issues only £3 / €3 / US\$ 3!*

*Depending on the region you live in.

shop.linuxnewmedia.com/trial

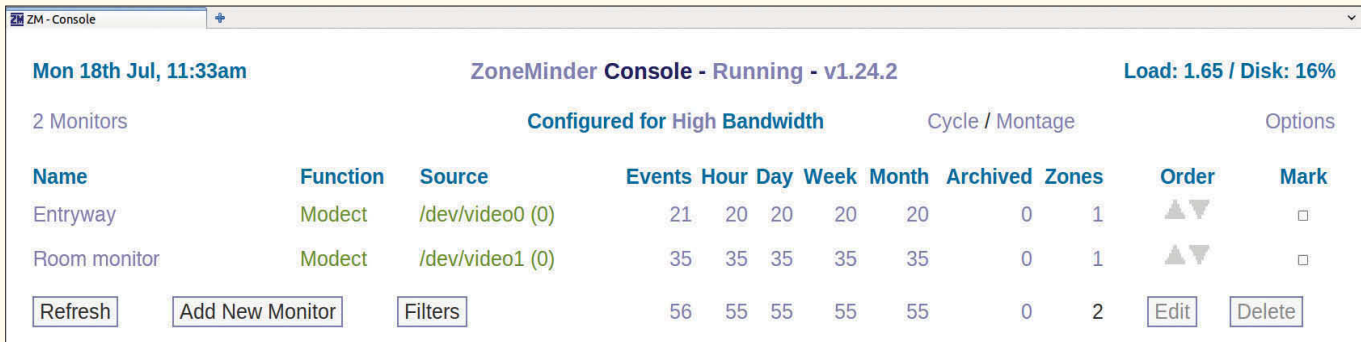


Figure 6: An at-a-glance view of the active cameras.

tion. If you don't manually restrict the recording duration, the software will record all the video signals from this camera, thus creating a huge amount of data. The *Mocord* option combines *Record* and *Modect*. ZoneMinder records permanently and generates an event whenever it starts recording, but if it detects motion in front of the camera, it outputs this event in uppercase letters in the event list. The final option, *Nodect* supports the use of external motion detectors, which many cameras include out of the box. In this case, the camera, rather than ZoneMinder, triggers the events.

After selecting the operating mode, check the *Enabled* box to switch on the camera. To define the number of images per second you want ZoneMinder to create, you then need to add entries in the *Maximum FPS* and *Alarm Maximum FPS* fields. The first entry relates to normal operation, the second to an alert case. The maximum frames per second rate for an alert scenario should be set higher than for normal operations, so you can follow the action more closely in the case of an event and also record more detail. If you have a powerful computer and enough disk space, values of up to 30 fps make sense here. After completing these details, press *Save*. The camera is now ready for action.

ZoneMinder has a large number of presets with meaningful defaults for the professional IP or PTZ cameras and video cards that the software supports. Press *Presets* in the top right of the configuration window to find them. In the *Preset* drop-down box (Figure 5), select your camera model and the basic settings and press *Save*. ZoneMinder will use these settings, although you do need to specify the IP address for the camera and the operating mode in the *Function* field. Typically, no other changes need to be made.

Sources and Targets

In your browser window, ZoneMinder now lists the image sources in a list. Note the entries in the *Source* column (Figure 6): Devices in green type are working correctly. If you see a camera with red type, you have a problem – typically a configuration error. In this case, click the *Name* for this entry, and the camera image should appear in the monitor window. If the window is black, or if you just see an empty frame, you need to reconfigure the camera. For lower powered models, it often helps to experiment with the resolution and the frame rate.

The monitor window (Figure 7) shows both the current camera image and a list of events with duration and number of frames. When you click on an event in the *Name* column, it plays back. The top section of the window (Figure 8) gives you generic data for the specified image sequence and a menu in which to process the images, if needed.

Because ZoneMinder initially stores events as a sequence of images in a temporary directory, you need to perform a second step if you want to generate a video from them. To do so, click the *Video* entry in the menu. In the window that appears (Fig-



Figure 7: ZoneMinder gives you an event list of the scenes captured by the camera.

Shop the Shop

shop.linuxnewmedia.com

Want to subscribe?

Need training?

Searching for that back issue you really wish you'd picked up at the newsstand?

Discover the past and invest in a new year of IT solutions at Linux New Media's online store.

shop.linuxnewmedia.com

DIGITAL & PRINT SUBSCRIPTIONS



SPECIAL EDITIONS



TRAINING



Nagios

- LPIC-1 LPI 101 - CompTIA Linux+ LX0-101
- LPIC-1 LPI 102 - CompTIA Linux+ LX0-102
- LPIC-1 - CompTIA Linux+ 101 + 102
- Bundle Monitoring with Nagios



Figure 8: The playback window menu.

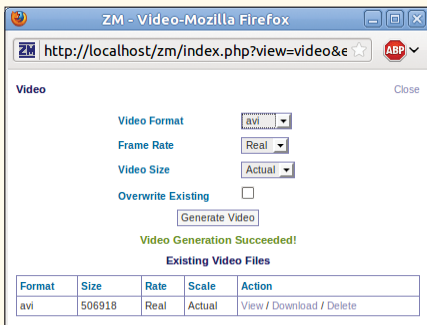


Figure 9: The dialog for creating a video from a sequence of images.

ure 9), choose the settings you need. ZoneMinder supports several popular video formats and lets you choose the size of the video (in the *Video Size* field) and the frame rate. For documentation purposes, you will want to keep the *Frame Rate* entry of *Real*, which corresponds to a real-time recording.

After making any necessary changes, press the *Generate Video* button. After creating the movie, you will see a table with the sequence data. In the right-hand column select *Action* to define what happens to the video sequence. Because you have created, but not permanently saved, the movie, you need to select the *Download* command in the *Action* field. ZoneMinder now assigns a name to the video and stores it in the current user's download directory on disk.

Alarm Zones

To avoid triggering events around the border of a surveillance area – for example a neighboring street with heavy traffic in a large-scale surveillance area – you can define alarm zones based on the camera image.

To do so, click the *Zones* column for the camera in the ZoneMinder main window. This displays a window with a table showing four coordinates on the *X* and *Y* axes below the current camera image. ZoneMinder superimposes a green grid on the camera image defined by the four coordinates in the table. To modify the area in which alarms are triggered, just drag points, and to modify the shape of the image, add more points (Figure 10).

The table on the left of the camera image is for various other settings related to triggering events or alarms. Adjusting the sensitivity of the software keeps every insignificant movement from throwing an event. ZoneMinder displays alerts in sequence in the event list, so you can identify critical events at a glance. Again, you need to press the *Save* button to keep your changes.

Conclusions

ZoneMinder lets you set up a professional video surveillance system that is hard to beat in terms of wide-ranging hardware support and flexible configuration options. The software offers useful performance even in larger projects with many cameras. It will run on older computer hardware thanks to its frugal use of resources, despite the typically exacting requirements of video processing.

The only downside to the ZoneMinder software is the lack of localization for users who need a foreign language setup and the somewhat convoluted configuration for some USB webcams. Because ZoneMinder is a free program and still under constant development, you can expect the maintainers to clean up these minor weaknesses in a future version. ■■■

INFO

- [1] ZoneMinder: <http://www.zoneminder.com/>
- [2] ZoneMinder compatibility list: http://www.zoneminder.com/wiki/index.php/Hardware_Compatibility_List
- [3] GSPCA compatibility list: <http://lwn.net/Articles/291036/>
- [4] Linux UVC video compatibility list: <http://www.ideasonboard.org/uvc/#devices>
- [5] Configuration script: <http://www.zoneminder.com/forums/viewtopic.php?t=16628>

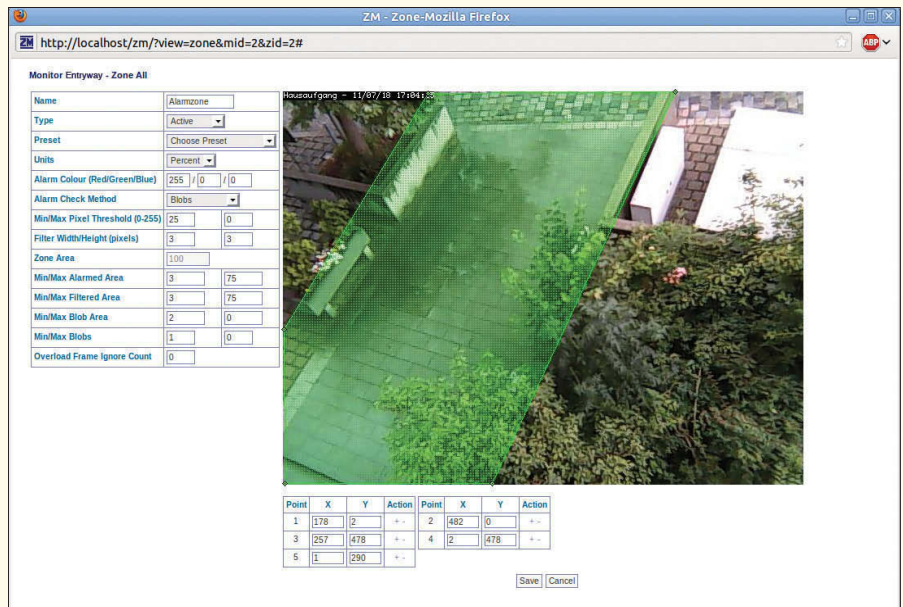


Figure 10: Alarm zones let you define the areas to monitor.



Google

FREE CD!

 **peppermint**
Web-Centric Operating System **two**

GOOGLE SPECIAL



Get Started with Google+

Fall 2011

Google

Discover Google+

Log in to the exciting service that is redefining the way you connect.



AVAILABLE AT YOUR NEWSSTAND!
OR ORDER ONLINE AT SHOP.LINUXNEWMEDIA.COM
(SELECT SPECIAL EDITION)

Mobile

10 cool smartphone apps

Free Tools!

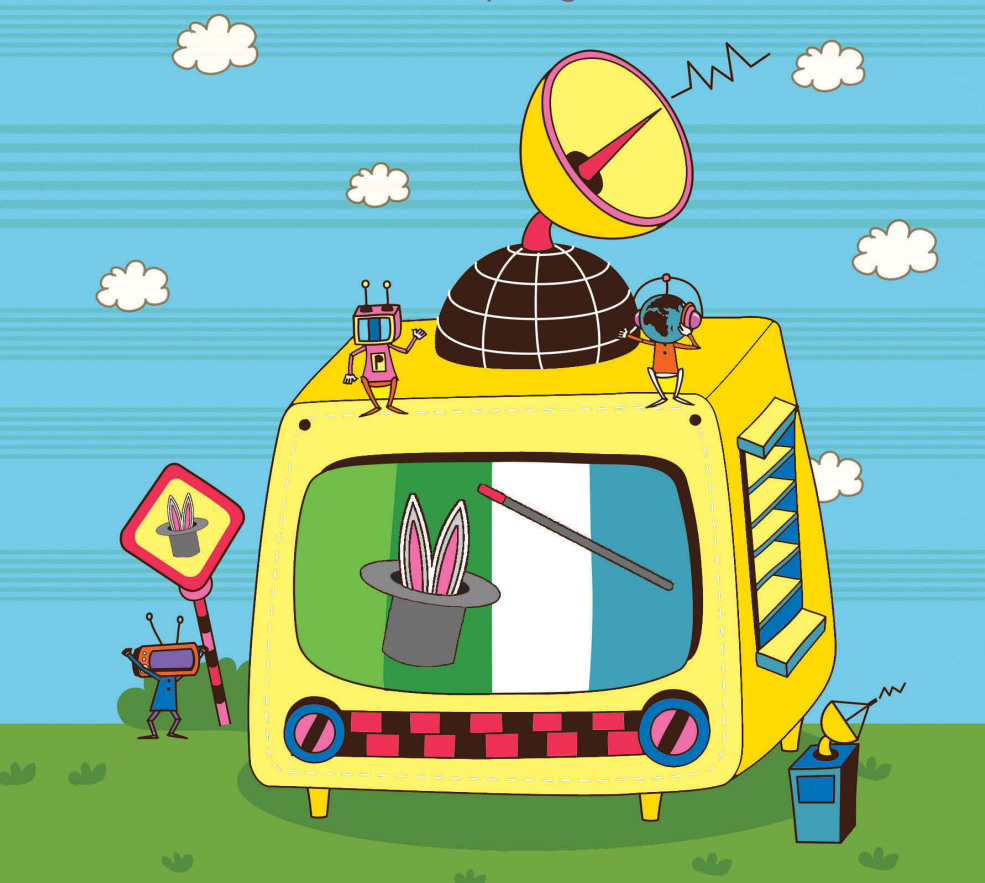
- Manage online shopping from a single web interface
- Build your own blog
- Create custom illustrations
- Improve your site's search rank

Manage Your Photos

Keep your pix in the cloud!
You can even touch up your digital images with Google's Picnik photo editor



US\$ 15.99
CAN\$ 17.99



The magic mantra: Compiling software from source

Going to the Source

If your distro's repository doesn't have the app you need, just repeat: `configure`, `make`, `make install`. *By Bruce Byfield*

BRUCE BYFIELD

Bruce Byfield is a computer journalist and a freelance writer and editor specializing in free and open source software. In addition to his writing projects, he also teaches live and e-learning courses. In his spare time, Bruce writes about Northwest coast art. You can read more of his work at <http://brucebyfield.wordpress.com>

One of the greatest strengths of GNU/Linux is its package management systems. Thanks to `apt-get` and `Yum`, you can download most applications in under 10 minutes, as well as all the other software you need to run it. However, some applications might not be in your distribution's repositories. That's when you need to install from the source code using the commands that my late colleague Joe Barr called the magic mantra: `configure`, `make`, and `make install`.

You might need these commands for several reasons. An application could be in such rapid development that its programmers have no time to make packages. At other times, your distro's repositories might not carry an application because it uses a non-free license or because the application is new and no one

has packaged it yet. At other times, a project's website might carry a newer version of an application than that included in your distribution, with features the distro's version doesn't have.

In any of these circumstances, you might want to install from source. However, be warned that source installations can be riskier than remaining within your distro's repositories. Source code can be buggier than software from a repository, and it might not be from a trusted source.

However, the greatest risk when installing from source is the potential to damage your system. When you install a package, compatibility is usually taken care of for you, but when you install from source, you are bypassing the package management system, and nothing prevents you from crippling your system.

For this reason, you should be careful what you install from source. As a general rule, be cautious about installing or updating major libraries, headers, and system utilities. You might want to be careful, too, about anything that affects your desktop.

By contrast, games and productivity software should create few, if any problems beyond overwriting an installation made with the package manager. Be aware, too, that installing from source isn't nearly as efficient and simple as using a package manager.

Preparing to Install

Most packages you install from source will be written in C or C++. To compile, you might need some packages that weren't included in your system installation, unless the installer had an option for configuring for development or advanced users. On Debian-based systems like Ubuntu, these packages are `gcc`, `cpp`, `binutils`, `libc-dev`, `g++`, and `libstdc++` and their dependencies. On Red Hat-based systems like Fedora, the list is the same, except that `glibc-devel` replaces `libc-dev` and `gcc-c++` replaces `g++`.

Before installing from source, take the time to read the project website. Many projects, like Gnumeric [1], list the dependencies needed for an application to run. If they are listed, you can generally minimize installation problems by making sure you have the version listed for each dependency.

Source code is generally available as a compressed TAR file from its project's website. Most of these TAR files will create a subdirectory into which the files are opened, but not all do, so you should create a separate directory as the target for your download. To change to the download directory and uncompress the TAR file, enter

```
tar -xf TARFILE.tar.gz
```

or

```
bzip2 -df TARFILE.tar.bz2
```

depending on whether the file is compressed with gzip or bzip2.

Next, go to the directory to which the files were uncompressed and check for a text file for instructions. Usually, such a file is named `CONFIGURE`, `README`, or a similar name. Source code from the GNU Project in particular can contain a generic set of instructions in a file called `INSTALL`. Whatever the name, it will usually tell you whether you should use the basic mantra, or some variation of it. Once you have read the file, you are ready to begin compiling the source code.

Step One: ./configure

The configure script prepares for compiling in two ways. First, it checks that your system has all the dependencies needed to compile (Figure 1). For example, one of the first checks configure usually does is check whether the system has the GCC compiler installed and running correctly. Because all this checking is just preparation, you do not need to be root to run configure.

Second, it creates the files with which to compile the program. The most important of these files is `Makefile`, which contains the basic instructions for compiling. Additionally, it might create a number of header files [2]. Other files that configure creates include `config.status`, which you can run to recreate the current configuration, and `config.log`, which you can use to help debug if you are an expert.

Each configure script is specific to the program it is written for, although if you examine the scripts, you will often see similarities, even in unrelated programs.

Typically, the configure script is in the top level of the directories you have uncompressed. In many cases, all you need to do is change into the directory that contains configure and enter `./configure`. You do not even have to run the script as root because you are not making any actual changes to your system.

However, you might want to run `./configure --help` first (Figure 2), especially if you know that your system has any peculiarities, such as placing files in an unusual location. You might also want to run an option such as `--no-create`, to see what happens without creating any output files.

```
checking for curl-config... /usr/bin/curl-config
checking curl/curl.h usability... yes
checking curl/curl.h presence... yes
checking for curl/curl.h... yes
checking for libcurl header... no
checking whether pthreads work with -pthread... yes
checking searching for pthread library... using -lpthread
checking for joinable pthread attribute... PTHREAD_CREATE_JOINABLE
checking if more special flags are required for pthreads... no
checking for the Boost Version...
checking for boost header... checking for Boost libraries... -lboost_thread
-mt -lboost_program_options-mt
checking if -fvisibility-inlines-hidden is broken... no
checking whether g++ supports -fvisibility=hidden... yes
```

Figure 1: One of the main purposes of the configure script is checking your system's resources.

```

or .deb users: apt-get install dejagnu
or .rpm users: yum install dejagnu

Gnash should still compile even with these warnings.
If it doesn't, report the warnings as a bug.

[root@fedora gnash-0.8.9]# ./configure --help
'configure' configures gnash 0.8.9 to adapt to many kinds of systems.

Usage: ./configure [OPTION]... [VAR=VALUE]...

To assign environment variables (e.g., CC, CFLAGS...), specify them as
VAR=VALUE. See below for descriptions of some of the useful variables.

Defaults for the options are specified in brackets.

Configuration:
-h, --help                display this help and exit
--help=short              display options specific to this package
--help=recursive          display the short help of all the included package
s
-V, --version              display version information and exit
-q, --quiet, --silent     do not print 'checking ...' messages
--cache-file=FILE        cache test results in FILE [disabled]

```

Figure 2: The configure script's help file lists options to fine-tune installation.

After you run `./configure`, look at the last few lines of the output to see the results (Figure 3). If you see an error, scroll back in the output to see exactly what the errors might be. Some error messages will be very specific, telling you exactly what packages you need to configure successfully. Others can be vague, so you have to second-guess them. For example, a line saying that the GCC compiler doesn't work might indicate that gcc isn't installed or it might mean it's corrupted.

In many cases, you can solve the problem by using your package manager to install or upgrade whatever system resources are missing. For example, if a header file is missing for a specific program, you can usually add it by installing a package that has the program's name followed by `-dev` or `--devel`. When you think you have solved the problem, run `./configure` again and repeat the trouble-shooting process until you have a message that configure was successful.

Step Two: make

The second step is to compile an executable that you can install. As with the configure script, you do not need to log in as root user – after all, you are still in the preparation stages.

Best practice suggests you should start with the command `make check` in the same directory in which you ran `configure`. This command runs any tests that come with the package, and is one more assurance that all will go smoothly.

However, if you have successfully configured, you can usually just enter `make`. Like the configure script, this command produces considerable output as it first searches for resources, in various subdirectories then uses them to build the executable (Figure 4). Unlike `configure`, `make` is considerably slower to run. Whereas `configure` typically

takes only a few minutes, even a small executable can take five to 10 minutes to compile, and large ones like Gnome or LibreOffice can take hours on the fastest machine.

As with `configure`, what you are mostly interested in is the results in the last lines of the output. Unfortunately, though, `make` errors are even harder to work with than `configure` errors because they give very little information to help you solve problems besides just a warning that something couldn't be done.

This lack of information makes troubleshooting difficult, but you can try several generic solutions. If you are using the latest nightly build for a program, the most likely problem is a bug that has crept into the code, and you can try another build. Another possibility is that one or more libraries on your system needs updating or downgrading to be compatible with the program – although whether you care to hunt down the problem depends on how important the program is to you.

You might also search the web to see if others have had the same problem. If you are lucky, you might find a patch for the source code that corrects the problem.

If you think you have solved the problem, go to the directory to which you uncompressed the source code, and enter the command `make clean` to remove the

```
WARNING: You need to have 'swfc' from SWFTools installed
to run some of the tests in Gnash testsuite.
You can install it from http://www.swftools.org/
or .deb users: apt-get install swftools
WARNING: you need as3ccompile from SWFTools
to run some of the tests in Gnash testsuite.
You can install it from http://www.swftools.org/
WARNING: You need the gettext package installed to use translations.
Required for building a package or 'make distcheck'
Install it from http://www.gnu.org/software/gettext/
or .deb users: apt-get install gettext
or .rpm users: yum install gettext
WARNING: You need the dejagnu package installed to get a summary
report after running 'make check'
Install it from http://www.gnu.org/software/dejagnu/
or .deb users: apt-get install dejagnu
or .rpm users: yum install dejagnu

Gnash should still compile even with these warnings.
If it doesn't, report the warnings as a bug.
```

Figure 3: Look at the end of the configure script's output to see how successful the script has been. This run has some problems that should probably be taken care of, but not enough to prevent basic success.

```
[root@fedora gnash-0.8.9]# make
Generating revno.h (0.8.9 20110824 none)
make all-recursive
make[1]: Entering directory `/home/bruce/Downloads/gnash-0.8.9'
Making all in desktop
make[2]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
make[2]: Nothing to be done for `all'.
make[2]: Leaving directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
Making all in libbase
make[2]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/libbase'
sed -e 's|@DEFAULT_FLASH_PLATFORM_ID[@]|LNX|g' -e 's|@DEFAULT_FLASH_MAJOR_VE
RSION[@]|10|g' -e 's|@DEFAULT_FLASH_MINOR_VERSION[@]|1|g' -e 's|@DEFAULT_FLA
SH_REV_NUMBER[@]|999|g' -e 's|@DEFAULT_STREAMS_TIMEOUT[@]|60|g' -e 's|@DEFAU
LT_SOL_SAFEDIR[@]|~/gnash/SharedObjects|g' './gnashrc.in' >gnashrc
make[2]: Leaving directory `/home/bruce/Downloads/gnash-0.8.9/libbase'
Making all in libmedia
```

Figure 4: The `make` command compiles the source code, searching recursively for resources to include in the executable.

remnants of your last effort to compile. You might also remove the Makefile, just to be safe. If you want to start the process from the beginning, use `make distclean` to remove any files created by the configure script as well.

Step Three: make install

Now, you are ready to install the program on your system. Because most applications are stored in directories that ordinary users cannot write to, log in as root user, change to the directory you have been working in, and enter the command `make install` (Figure 5).

When you are done, the only sign is a return to the command prompt. If you change to your ordinary account, you should be able to run the program you have just installed.

Because most of the difficulties occur in the first two steps, `make install` should be quick and trouble-free. Try to run `make install` as an ordinary user, and you will quickly see a *Permission Denied* message when you try to write to a restricted directory such as `/usr/share/bin`.

The most common problem you will encounter is likely to be that the program files were installed in an unusual place that is not in your path, such as `/opt`. The output from `make install` will tell you where to look. If you can run the program using the full path, then either add the path to your environment with

```
export PATH="$PATH:[DIRECTORY]"
```

or else start over again using

```
./configure --prefix=DIRECTORY
```

to specify a folder in your path in which to install.

Uninstalling and Alternative Systems

How long you persist in trying to install from source depends on luck, patience, and how important the program is to you. Should you decide to give up or decide that the program isn't for you, just return to the directory from which you installed as root user, and run `make uninstall` (Figure 6). Like `make install`, this command gives no indication of its success or failure, but you can check to see whether the uninstall is complete by looking at the paths in the Makefile and, if necessary, deleting any files manually.

These processes should apply to perhaps two-thirds of the source code available for download. But, increasingly, alternatives are being used. Today, some source code can be installed by shell scripts, or even generic installers similar to those common in Windows. Still others, like KDE, use CMake [3], which has its own set of procedures, or, like Pysol, are written in a language like Python in which you run a script rather than pre-compile.

However, if nothing else, trying the magic mantra a few times will give you more of an appreciation of what goes into preparing a program to run. Should you hit any difficulties, experimenting with `./configure`, `make`, and `make install` can also give you a renewed thankfulness for package managers. ■■■

```
[root@fedora gnash-0.8.9]# make install
Generating revno.h (0.8.9 20110824 none)
make install-recursive
make[1]: Entering directory `/home/bruce/Downloads/gnash-0.8.9'
Making install in desktop
make[2]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
make[3]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
test -z "/usr/local/bin" || /bin/mkdir -p "/usr/local/bin"
/usr/bin/install -c gnash-gtk-launcher '/usr/local/bin'
test -z "/usr/local/share/applications" || /bin/mkdir -p "/usr/local/share/applications"
/usr/bin/install -c -m 644 gnash.desktop '/usr/local/share/applications'
test -z "/usr/local/share/icons/hicolor/32x32/apps" || /bin/mkdir -p "/usr/local/share/icons/hicolor/32x32/apps"
/usr/bin/install -c -m 644 ../packaging/gnash.xpm '/usr/local/share/icons/hicolor/32x32/apps'
```

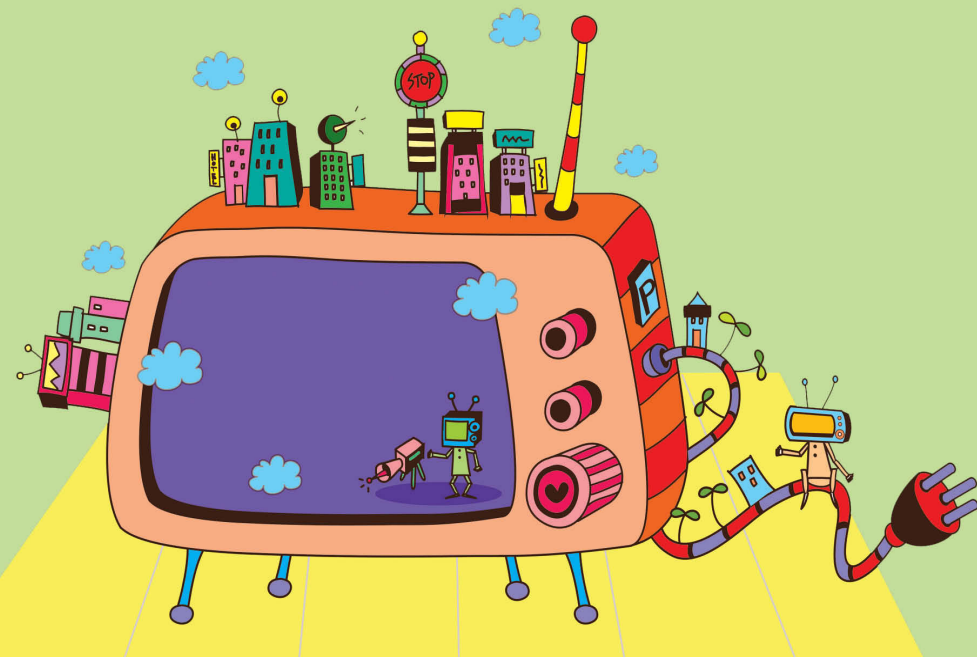
Figure 5: Run `make install` as root to add the compiled code to your system and notice what directories it is using in case you ever need to uninstall manually.

```
[root@fedora gnash-0.8.9]# make uninstall
Making uninstall in desktop
make[1]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
( cd '/usr/local/bin' && rm -f gnash-gtk-launcher )
( cd '/usr/local/share/applications' && rm -f gnash.desktop )
( cd '/usr/local/share/icons/hicolor/32x32/apps' && rm -f gnash.xpm )
( cd '/usr/local/share/gnash' && rm -f gnash-splash.swf )
make[1]: Leaving directory `/home/bruce/Downloads/gnash-0.8.9/desktop'
Making uninstall in libbase
make[1]: Entering directory `/home/bruce/Downloads/gnash-0.8.9/libbase'
( cd '/usr/local/include/gnash' && rm -f accumulator.h SimpleBuffer.h extension.h GnashNumeric.h jmtree.h GnashSleep.h gmemory.h SharedMem.h sharedlib.h tree.hh tu_file.h IOChannel.h Socket.h GnashSystemFDHeaders.h GnashSystemNetHeaders.h GnashSystemIOHeaders.h GnashFileUtilities.h ClockTime.h WallClo
```

Figure 6: The `make uninstall` command removes the software you have installed from source.

INFO

- [1] Gnumeric: <http://projects.gnome.org/gnumeric/downloads.shtml>
- [2] Header files: http://en.wikipedia.org/wiki/Header_file
- [3] CMake: <http://www.cmake.org/>



Cut the cruft from presentations with deck.js

On Deck

Fussing with presentation software is no fun, and all the fuss doesn't really improve your presentation. If you really want to simplify, look at deck.js. *By Joe "Zonker" Brockmeier*

JOE "ZONKER" BROCKMEIER

Joe "Zonker" Brockmeier is a freelance technology journalist, editor, community builder, and Software Freedom advocate. Brockmeier was the openSUSE Community Manager from 2008 to early 2010. Brockmeier now writes for *Linux Pro Magazine* and a number of other technology publications. His personal site can be found at <http://www.dissociatedpress.net/>, and you can follow him on Twitter and Identi.ca as @jzb.

I've given more than my share of presentations over the years. On Linux, the default option for presentations is usually LibreOffice or OpenOffice.org. I have to admit, though, I've never much liked the OOo/Libre presentation tool, although it *will* work in a pinch, and it's *great* for Microsoft Office refugees.

The OOo/LibreOffice developers did an excellent job creating a drop-in replacement for Microsoft PowerPoint, if that's what you want. Unfortunately, I never liked the original PowerPoint, so a clone doesn't really make me happy either.

As a result, I'm always on the lookout for a new presentation program. I've tried a few alternatives, but I wind up going back to LibreOffice because the others are just too clunky or too much work. I have even tried the the closed source Keynote program, part of Apple's

iWorks suite, and it's not much better than the others.

When I read about deck.js (Yes, sadly, no initial cap), I was intrigued. Despite the annoying web developer habit of ignoring the naming convention for proper names, it grabbed my attention.

What is deck.js?

As you've probably surmised from the name, deck.js <http://imakewebthings.github.com/deck.js/> is a JavaScript-based presentation program. (Hence the .js in its name.) In a nutshell, deck.js allows you to write presentations in plain-old HTML and theme them to get the look and feel that you want.

The slides are nothing more than HTML, so to create slides, you just need a text editor. That's it. To display slides, you need a supported web browser – preferably with a full-screen mode, although a full-screen display isn't absolutely necessary.

This sparse description might sound like a deck.js presentation would be, well, pretty homely. In fact, the default set of themes and transitions is *fantastic* for a standard presentation. I wish I'd had deck.js handy when I was working with Novell and doing lots of openSUSE presentations.

Advantages to deck.js

Many Linux users are at least passingly familiar with HTML. If you know some HTML, it will take you almost no time to become familiar with deck.js formatting. If you know some JavaScript and CSS, you'll quickly become an expert.

Deck.js also allows extensions, so you can add features if you happen to know JavaScript – or you can find extensions created by other users.

Because it's all HTML, deck.js makes it is easy for folks to share presentations and make changes. Developers can even check their deck into version control, meaning you could easily share a preso via GitHub or another service. This feature is great for projects that may have two, ten, or one hundred ambassadors visiting shows or Linux User Groups to talk about their project.

Another thing I like about deck.js is that it is easy to include images, sound, and video. If you're online, you can easily include a YouTube video in a presentation, which is a royal pain if you're

talking with LibreOffice Impress. It probably goes without saying, but the other nice thing about deck.js is that it is open source. It's dual-licensed under the GPL and MIT licenses – which should cover almost any open source project that might want to think about using deck.js.

Another advantage? You can use, or re-use, deck.js for websites and online presentations. It is, after all, just HTML, CSS, and JavaScript.

Get Started

Now that you're chomping at the bit to get started, I'll show you how.

You can find the source code on the GitHub website [1], along with the readme for deck.js. The only dependencies that you will need to obtain are the source code, a JavaScript library called Modernizr [2] and jQuery [3]. Modernizr is actually included with deck.js.

Download the tarball from GitHub and expand it in a directory where you can start tweaking the slides. Alternatively, you could fork the project on GitHub and work with it there if you're familiar with Git.

A sample deck is included in the introduction folder. You can view it in your browser just by opening the index.html file.

Editing deck.js

To get started with a basic deck.js presentation, copy the index.html file from the introduction. Don't try to memorize all the meta and link tags, but be sure you edit the title tag each time!

Specify the slide themes with the link tag:

```
<link rel="stylesheet"
id="style-theme-link"
href="../../themes/style/web-2.0.css">

<link rel="stylesheet"
id="transition-theme-link"
href="../../themes/transition/horizontal-slide.css">
```

Note that I just grabbed the preceding example right out of the introduction code. The code for a standard slide looks like this:

```
<div class="slide" id="title-slide">
<h1>Let's Start with deck.js</h1>
</div>
```

That's it. A div tag using the slide class, along with the header, is all you need to produce a slide that looks like Figure 1.

Look good? It also comes in neon (Figure 2).

These slides might not be quite as spiffy as some of the themes you see for Keynote or PowerPoint, but I like them. (Frankly, the default themes shipped with Impress, don't impress.) The deck.js look is clean and easy to read, and it puts the focus on the presentation.

Of course, you can include bullets in your presentation. In fact, if you use the slide class, you'll be able to unveil one bullet at a time, just like with the fancy presentation software:

```
<ul>
<li class="slide">bullet 1</li>
<li class="slide">bullet 2</li>
</ul>
```

That's way simpler than doing the same thing in Impress or PowerPoint. You can also use the blockquote element to get a very nice block quote, like the following:

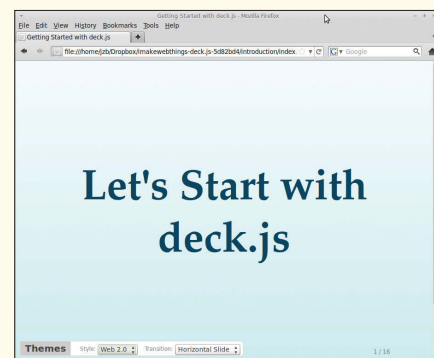


Figure 1: Building a deck.js title slide.

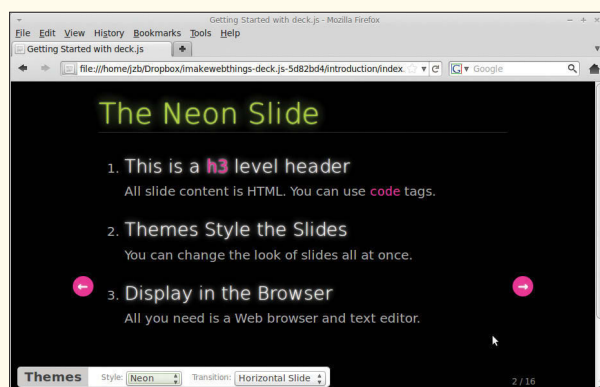


Figure 2: Add some flash with the deck.js Neon theme.

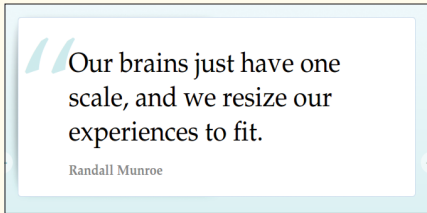


Figure 3: The deck.js blockquote in action.

```
<blockquote>
<p>Quote goes here</p>
<p><cite>Author</cite></p>
</blockquote>
```

That's all there is to it. (Figure 3). Finally, you can include images with the `img` tag and video with the `iframe` element.

It's really simplicity itself. If you can stumble through basic HTML, you can whip out a presentation in deck.js in about an hour – assuming, of course, that you already know what you want to say.

Extensions

The deck.js tarball includes several extensions. You have to specify these extensions in the headers as well, as shown in Listing 1.

The `goto` extension lets you hit `g` and then skip to a slide number. The `hash` extension allows you to link to individual slides with a (you guessed it) hash.

The `menu` extension gives you a menu overview of all the slides. Just hit `m` and you can see a grid that displays all slides. Just click the slide you want to visit or hit `m` again (Figure 4).

Finally, the `navigation` extension adds the spiffy wee arrows to the presentation if you don't want to use the keyboard to navigate – or if you want to cue people visually when you're posting the preso online.

Really, there's no good reason *not* to include these extensions 99.9% of the time, so just include them by default.

Disadvantages to deck.js

I should note that deck.js has a couple of disadvantages. The primary disadvantage to deck.js is that you can't easily create a PDF out of a presentation and send it to people; nor is it easy to print if you need to hand a printed version to somebody. Well, that could be a disadvantage to some, and a selling point for others who'd like to see their office use less paper. It's also a problem that you can't upload the presentation to SlideShare and similar services.

You probably won't win friends and influence people if you're working in a corporate environment and produce a deck using deck.js – at least not without talking it over with colleagues first.

To put it another way, deck.js is non-standard and you might not want to be non-standard element in certain work environments.

Put the Focus Where It Belongs

I want to address one non-technical point in closing and emphasize why I like deck.js.

The software makes you focus on your own presentation, not your slides. With a few exceptions, most presentations don't need to be wordy and full of bullet points and graphs, etc. If you're doing the Q4 sales presentation, of course, you're going to have an ugly set of slides, and the focus will be on the information.

But when you're doing any kind of presentation that requires you to sell an idea to the audience – whether it is to sell a product or convince a room full of engineers to try a different development model – the bulk of your time should be spent *away from the computer* developing your story and your delivery.

Your presentation should have a narrative flow. Start with the intro, tell the audience what they're going to learn, and then tell your story. Don't just belt out facts or recite from your slides. The slides are a visual aid, and the focus should be on you – not the slides.

So, the next time you have a presentation, use deck.js and keep it simple. ■■■

LISTING 1: Extensions

```
<link rel="stylesheet"
href="..../core/deck.core.css">
<link rel="stylesheet"
href="..../extensions/goto/deck.goto.css">
<link rel="stylesheet"
href="..../extensions/menu/deck.menu.css">
<link rel="stylesheet"
href="..../extensions/navigation/deck.
navigation.css">
<link rel="stylesheet"
href="..../extensions/status/deck.status.
css">
<link rel="stylesheet"
href="..../extensions/hash/deck.hash.css">
```

INFO

- [1] GitHub: <https://github.com/imakewebthings/deck.js>
- [2] Modernizr: <http://www.modernizr.com/>
- [3] jQuery: <http://jquery.com/>

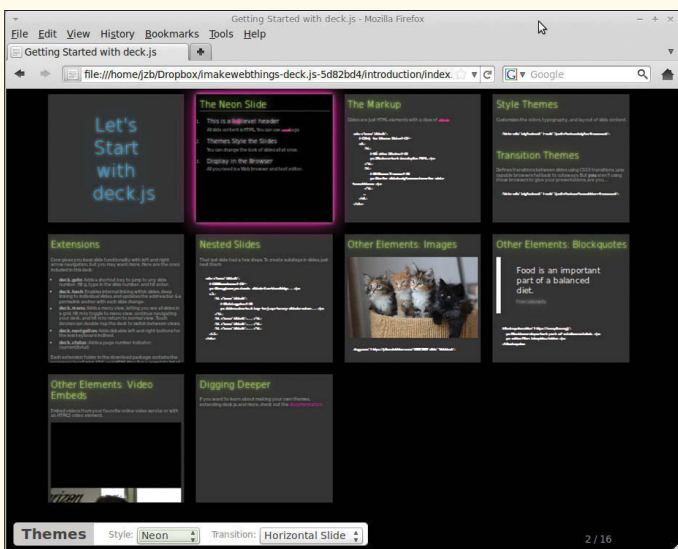


Figure 4: See all the slides together in menu mode.

LinuxCon North America

By Jon “maddog” Hall

At the recent LinuxCon North America [1] in Vancouver, Canada, we celebrated the 20th anniversary of the start of the Linux Kernel project. Many of the best developers in the community were there, and we had “gala” events celebrating the past 20 years. Although the Linux kernel tends to dominate aspects of computing such as high-performance computing, embedded systems design, and servers, several people acknowledged that we still have not enjoyed “the year of the Linux desktop.”

This particular conference stirred up a lot of retrospection. Jon Corbett, who normally does a bang-up job telling us what is new in the Linux kernel, went back a bit in time and showed some of the early timeline issues that almost derailed kernel development. Fortunately, the community was ready to tackle those issues, which tended to be more procedural and structural, concerning how to scale development, than technical, concerning how to scale the kernel itself.

Dr. Irving Wladawsky-Berger of IBM gave a fascinating view into some of the early meetings that IBM had on the Linux kernel and the decisions that led them to invest that first billion dollars. What was really news to me was the effort IBM made to investigate and view Linux from every angle before making that decision, and one of the main considerations was that “Linux would run on everything.” IBM realized that the community would make Linux run on everything from the very smallest computers to the very largest, and to them, this created real value for the OS that could not be matched by any other operating system.

Richard Fontana, a lawyer for Red Hat, gave a good presentation of legal issues that have plagued the Unix, and then the Linux, community since the earliest days. I was a little disappointed that, although he had a very good description of Unix legal issues and why distribution of Unix source code was impractical because of AT&T source code licensing, he did not mention the commercial versions of Unix and their place in the ecosystem.

Jim Zemlin, Executive Director of the Linux Foundation, moderated a panel that included me, Dan Frye of IBM, and Eben Moglen of the Software Freedom Law Center. We reminisced about some of the early days of Linux, but with a view

toward the next 20 years. Eben waxed long on the issue of software patents, and some people thought his message was a little “dark.” I tend to think that Eben’s message was not dark enough, and people have to double or quadruple their efforts to get true patent reform before it is too late.

All of the keynote speeches were streamed live on video, and if you go to the LinuxCon site and sign up for video streaming, you will be able to watch them.

Two of the main sessions I found very interesting had a common theme addressing performance.

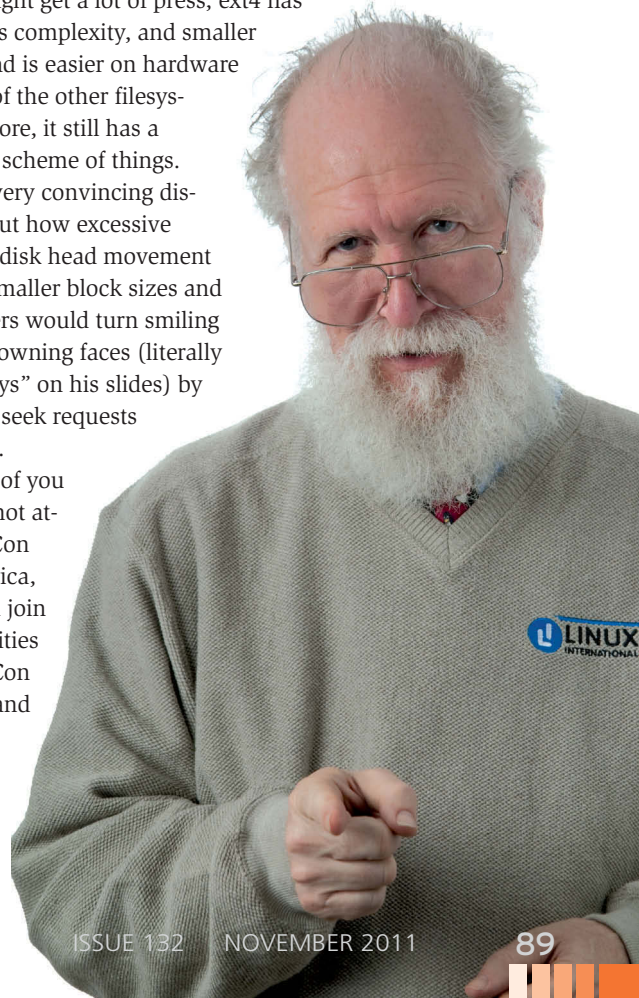
The first session was by Khoa Huynh of IBM. He was benchmarking storage in KVM-based clouds because a customer complained of poor performance. By modeling the customer’s systems, he determined that writing small chunks of data was causing the underlying disks to write back huge blocks for every small write that was issued. When the customer wrote larger chunks of data at one time, they did not notice the problem. This was a fine example of good engineering.

I also appreciated a talk by Ted Ts’o of Google about the continuing value of the ext4 filesystem. Ted presented the argument that although other filesystems, such as Btrfs, SquashFS, OCFS, and XFS, might get a lot of press, ext4 has stability, less complexity, and smaller code size and is easier on hardware than some of the other filesystems; therefore, it still has a place in the scheme of things. Ted gave a very convincing discussion about how excessive amounts of disk head movement caused by smaller block sizes and more pointers would turn smiling faces into frowning faces (literally with “Smileys” on his slides) by stacking up seek requests very rapidly.

For those of you who could not attend LinuxCon North America, you can still join in the festivities with LinuxCon Europe [2] and LinuxCon Brazil [3] later this year. ■■■■

INFO

- [1] LinuxCon North America:
<http://events.linuxfoundation.org/events/linuxcon>
- [2] LinuxCon Europe:
<http://events.linuxfoundation.org/events/linuxcon-europe>
- [3] LinuxCon Brazil:
<http://events.linuxfoundation.org/events/linuxcon-brazil>



Zack's Kernel News

Chronicler Zack Brown reports on the latest news, views, dilemmas, and developments within the Linux kernel community.

By Zack Brown

Rebooting Kerneloops.org

The kerneloops.org site has been down for a long time, but it may be getting some love in the near future. Jiri Slaby first reported that the site was broken back in June and was followed by a small chorus of interested folks who were also interested. For several months, there was no news, and Alan Cox speculated that the site might simply be gone for good.

Finally in August, Arjan van de Ven invited people to help fix the code. Andi Kleen remarked that it had been a very useful service, and that it would be good to get it working again. A couple of people expressed interest, including Bastien Roucaries, Arnaud Lacombe, and Josh Hunt, but, as of this writing, the site still gives the same error Jiri initially reported – a kernel version mismatch when trying to view recent kernels.

The idea behind kerneloops.org has been to take advantage of the oops reports that often get sent to the linux-kernel mailing list and other lists by folks who experience crashes and want help debugging them. Each oops report is a highly structured wad of data, generated by the kernel, that tries to give as much info as it can about why the system crashed. Although oops reports are somewhat human-readable, they are designed also to be parsed by scripts.

In addition to whatever help people get from the folks when they post an oops report to a mailing list, each of these dying kernel breaths are collected in the little glass kerneloops.org bottle and analyzed for statistics, which are then presented on the site. The stats let people know which kernels are particularly prone to crashes, and which parts of any given kernel are causing the most trouble. Essentially, it helps the kernel folks figure out where to direct their energies.

It'd be nice to have kerneloops.org up and running again. Maybe someone, or a few someones, will take it over and make it well again.

Long-Term Plan for Long-Term Kernels

Greg Kroah-Hartman announced his intention to revamp his approach to long-term kernel releases. These releases are like the stable kernel series, but they continue to be actively maintained for much longer than other stable ker-

nels. Typically, the long-term kernel releases are highly valued by Linux distributions, embedded systems developers, and other corporations who base their business on Linux. Greg posted a request for suggestions as to how he should organize the long-term maintenance process. A big discussion ensued.

His plan essentially involved picking a favored kernel release every year, to be maintained as a long-term stable series, and then dropping support for that kernel after two years.

Greg also gave a nice history of long-term kernel releases:

2.6.16 became a "long-term" kernel because my day job (at SUSE) picked the 2.6.16 kernel for its "enterprise" release, and it made things a lot easier for me to keep working at applying bugfixes and other stable patches to it to make my job simpler (applying a known-good bunch of patches in one stable update was easier than a set of smaller patches that were only tested by a smaller group of people).

Seeing that this worked well, a cabal of developers got together at a few different Linux conferences and determined that based on their future distro release cycles, we could all aim for standardizing on the 2.6.32 kernel, saving us all time and energy in the long run. We turned around and planted the proper seeds within the different organizations and, lo and behold, project managers figured that this was their idea and sold it to the rest of the groups and made it happen. Right now, all of the major "enterprise" and "stable" distro releases are based on the 2.6.32 kernel, making this trial a huge success.

A big response followed his announcement. Everyone has a stake in stable kernels. David Lang came up with some suggestions. He pointed out that just picking a kernel every year on a time clock might run into some problems. He said he'd seen a lot of kernel releases that turned out just to be not quite right – just a bit off in a general sense, in a way that backporting fixes was not really going to help.

He suggested selecting long-term kernels from the pool of stable kernels that had already been out for a few months and that showed signs of non-wonkiness. Greg liked this idea.

ZACK BROWN

The Linux kernel mailing list comprises the core of Linux development activities. Traffic volumes are immense, often reaching 10,000 messages in a week, and keeping up to date with the entire scope of development is a virtually impossible task for one person. One of the few brave souls to take on this task is Zack Brown.

Ben Hutchings also pointed out that Greg's idea of "dropping" support after two years was not ideal. He suggested handing off the long-term development to a volunteer after the two-year mark and only "dropping" support if no volunteer came forward. But, it seemed that this approach was actually Greg's default, and he confirmed he'd be handing the 2.6.32 kernel off to Willy Tarreau around February 2012, instead of just dropping support.

Greg had initially justified his two-year estimate by saying that consumer electronics were often replaced after two years, but some folks, including Ben, felt that two years was a lot less than what a lot of vendors needed. And, Jeremiah C. Foster said that the automotive industry needed at least 10 years of support for the entertainment software typically embedded in modern cars.

Greg's response was, "Isn't that the job of the distros and commercial OSVs today? Are they somehow not doing this job well? Do they need help from the community instead to help define, implement, and maintain this for them?" He invited Jeremiah or anyone from the automotive industry to contact him to figure the issue out, and Jeremiah started setting up a meeting. But, Jeremiah did also add, "The 10-year model may be coming to an end. Over the air updates, once a year dealer updates, and mileage-based service updates are all now opportunities to ship bug fixes and potentially new features. So, it looks like a 5-year model, or even shorter, might be usable." He added, "Currently, there are agreements between car makers and their software partners and, as these groups both get more familiar with open source, hopefully they'll be better prepared to work with the mainline kernel maintainers and to be open about their requirements. Right now, unfortunately, they are not able to do that."

Overall, Greg's announcement was met with a lot of discussion and encouragement, and a lot of folks were enthusiastic to continue talking about things during the Kernel Summit and various other gatherings.

Promoting Wireless Networks

Luis R. Rodriguez liked the Electronic Frontier Foundation's effort to promote wireless networks and invited folks to contribute to a wiki he created [1].

There were a number of comments; most seemed to be saying that an open wireless network would have too many technical and legal problems to ever be feasible. Still, Luis posted a link to a new mailing list set up by the EFF [2].

The idea of an open wireless network, supported by individuals rather than big companies, seems like something to work toward if possible. The reliance on wired technology, and on the good will of corporate service providers, seems like something to move away from. ■■■

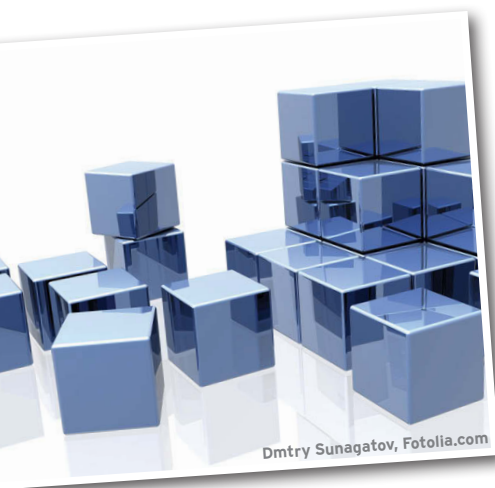
INFO

[1] Open Wireless Movement: <http://wireless.kernel.org/en/developers/OpenWirelessMovement>

[2] Tech: <https://lists.openwireless.org/mailman/listinfo/tech>



Projects on the Move



Drupal and WordPress often hog the spotlight, but other content management systems offer impressive performances. This month, we look at the latest Joomla release and check out **concrete5**. By *Rikki Kite*

Although I wouldn't call myself a web designer, over the years I've managed websites running on eZ Publish and Drupal, and I've created several sites running WordPress, which I think is the most user-friendly site to hand off for non-technical clients to maintain. Recently, I started using Joomla, which powers the advancedclustering.com website. At first, I found Joomla to be less user-friendly than WordPress, more inviting than eZ Publish, with a learning curve comparable to Drupal (although newer versions of eZ Publish are pretty Drupal-like). In short, Joomla is *different* from other content management systems, so it takes some getting used to; however, it is feature-rich, has an active community, and comes with excellent documentation.

"Each CMS has its own unique strengths and weaknesses, but for me Joomla's unique strength lies in the breadth of the community and the range of available extensions – over 8,000 – that give site builders unrivaled versatility with the type of websites that they can build," says Brian Teeman, co-founder of Joomla.

Joomla 1.7 rolled out in July and was the first release on the new six-month release cycle [1]. In addition to security updates, the new release offers the option to specify the site name before or after the page title, a "remove installation folder" button as part of the installation process, and the ability to add a background image to the custom HTML module.

To test Joomla without installing it, users can create a demo account, which includes 30 days of free hosting, video training, and support [2]. The Joomla site also

RIKKI KITE

Rikki Kite is Director of Marketing at Advanced Clustering Technologies in Kansas City, Kansas. In addition to *Linux Magazine* and *Linux Pro Magazine*, Kite has been published in *Sys Admin Magazine* and *UnixReview.com*. She can be contacted at rkite@advancedclustering.com.

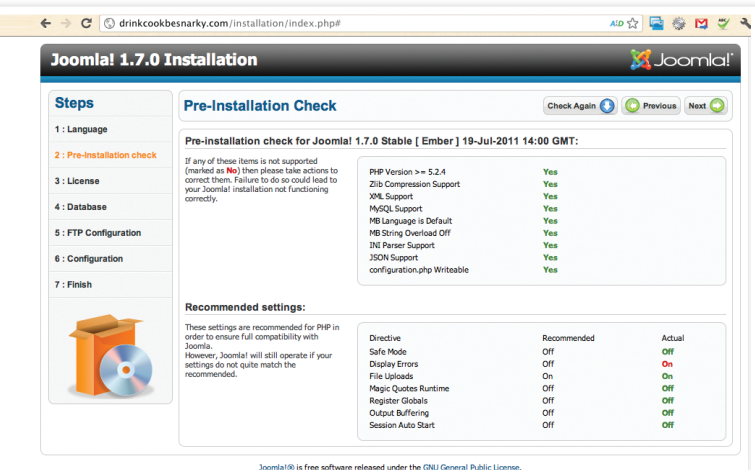


Figure 1: A systems check before takeoff.

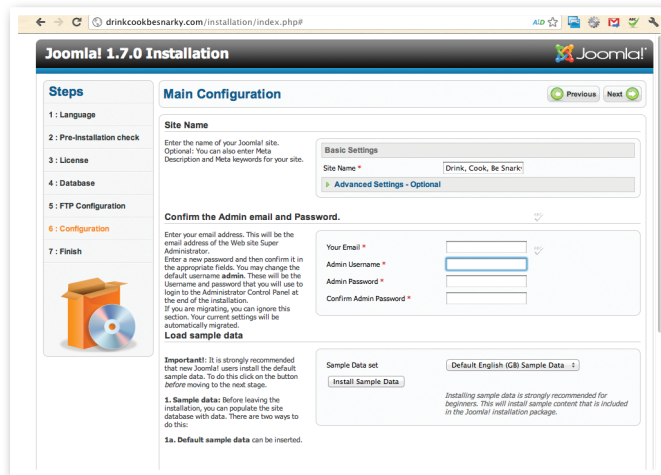


Figure 2: Enter your site name and admin details.

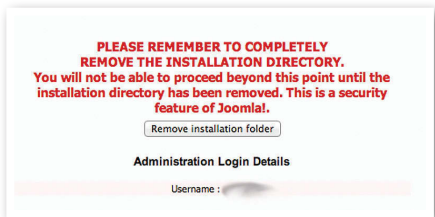


Figure 3: Install complete; the installation directory is ready for removal.

provides detailed install manuals, videos, documentation, and a Joomla online magazine [3]. During the Joomla install, the first screen offers the option to choose a language, and the next screen provides a pre-installation check and recommended settings (Figure 1).

In the next three screens, you will add your database configuration, FTP configuration (optional), and your main configuration, which allows novices to install a sample data set to get started (Figure 2). After congratulations for a successful install, the next screen offers the new button to remove the installation directory (Figure 3).

Clicking the *Site* button takes you to your new site with the sample data set (if you chose that option). The sample data set offers guides for beginners, upgraders, and professional users (Figure 4). Clicking the beginners tab opens a page with links to information about extensions, components, modules, templates, plugins, and more. The bottom, left-hand corner of the sample page offers the login and admin options (Figure 5). After you log in, the fun begins because you have finally found the admin panel.

The sample data already populated on the site allows you to play with all the admin options and get a feel for Joomla before you tweak the theme or add extensions (which is beyond the scope of this article). Clicking the *Menus* option in the top tool bar opens up the sample menus, which I promptly deleted (with the exception of the *Main Menu*). To edit the sample articles, click on *Content* in the top menu (Figure 6). *Banners* under the *Components* tab show up in the bottom of the sample site as the three *Support Joomla!* boxes (Figure 4).

Under *Extensions*, you can manage extensions, add modules, and manage plugins, templates, and languages. The final option, *Help*, in the administration tool-

bar, provides links back to the Joomla documentation and community.

Whether you are a novice or an experienced web designer, Joomla is packed with features, well documented, and powered by an international community of contributors. The learning curve can be a bit intimidating, but the extensive documentation, active community, and a little hands-on experience will help you get your personal or business site up and running quickly.

concrete5

Released in 2008, the open source concrete5 content management system is the offspring of Concrete CMS, created by Portland developers Franz Maruna and Andrew Embler [4]. The project site says that concrete5 is a CMS made for marketers but built for geeks.

How does concrete5 compare to other well-known content management systems, such as Drupal, WordPress, eZ Publish, and Joomla? “As a website editing tool, I think concrete5 compares exceedingly well to everything you mentioned,” Embler says. “They all have their strengths; ours is in-context editing: you navigate your website as normal, and change content right from the page on which it appears.”

Embler says that concrete5 is also great as a framework for building interactive websites with custom components and evolving requirements. “If you’re certain your site will never be more than a blog, a store, or a community news archive and your design requirements aren’t that exact, other software will probably meet those needs with less tinkering,” he says, “But if you wish to add some or all of these capabilities in an elegant way to a website that looks good, concrete5 is a great choice.”

Although concrete5’s e-commerce add-on is full-featured and capable, Embler says that he’d probably choose a

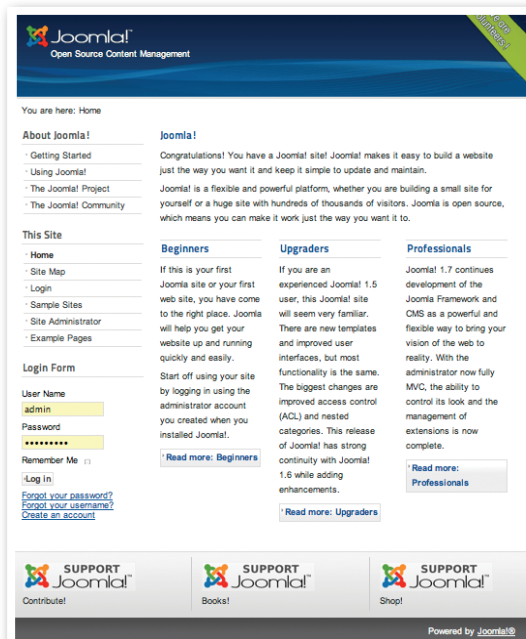


Figure 4: Sample data offers beginners a preview of Joomla in action.

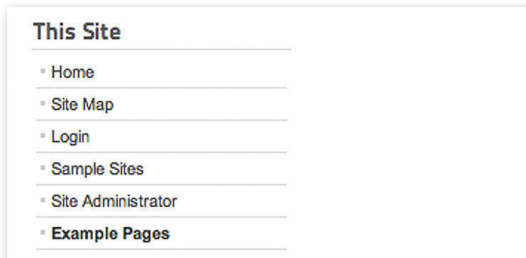


Figure 5: Click *Login* to dig into the configuration fun.

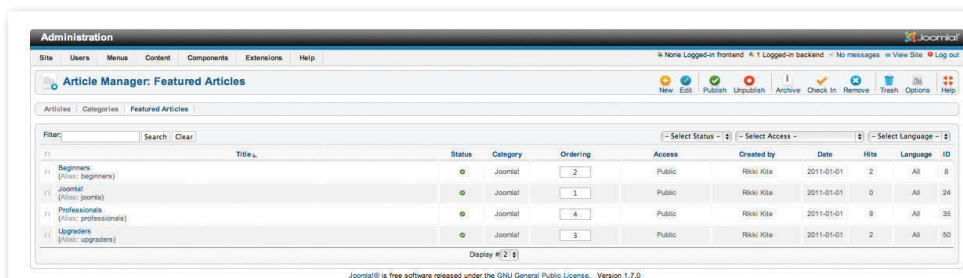


Figure 6: Sample feature articles also offer handy, practical tips for using Joomla.

Community Notebook

Free Software Projects

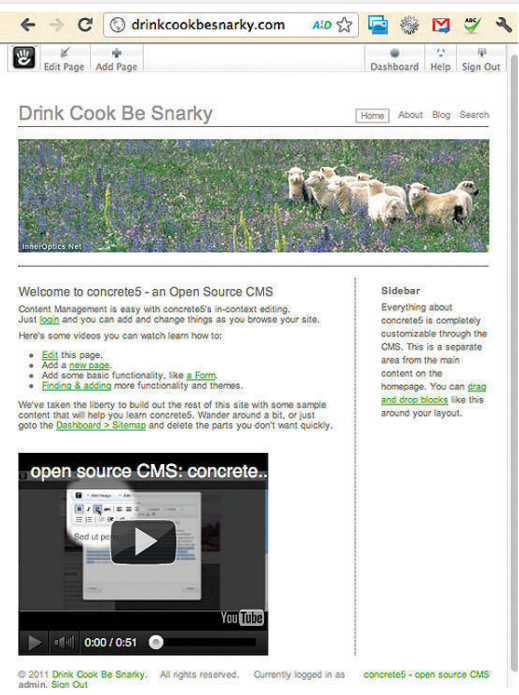


Figure 7: concrete5 installed and ready to roll.

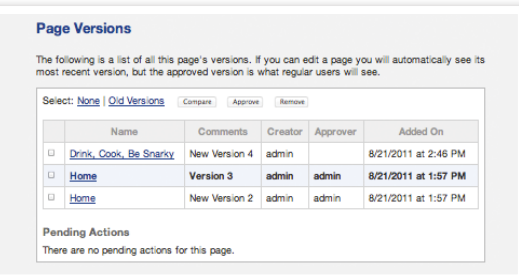


Figure 8: See previous versions of your website pages.

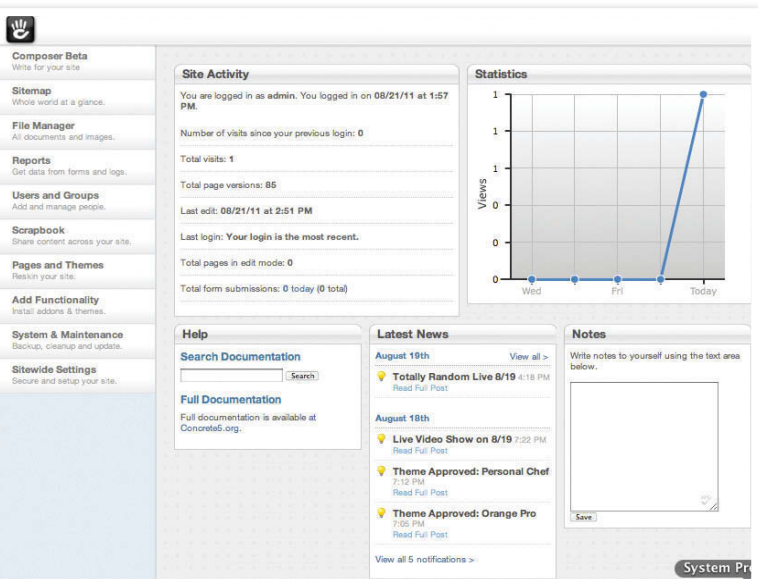


Figure 9: This well-organized CMS won't require a lot of head scratching.

different application if he were building a large e-commerce website with tens of thousands of products. "If I were building a custom web application without a lot of content management, I'd choose a dedicated framework like the Zend Framework or Symfony," he says, "We're also not a good fit if you're building a site where every single page has a significantly different design footprint, or if your site is built for Flash."

To try concrete5 without installing it, set up a demo install on the project site [5]. If you would prefer to install it, you can download concrete5 from the website, where you will also find documentation, forums, add-ons, themes, and installation help.

My concrete5 install was pretty straightforward and any questions I had were answered by searching the project website. I chose to install with sample content, which brings up a welcome page with helpful links, video, and a streamlined toolbar (Figure 7).

Clicking *Edit Page* in the top-left hand menu opens an expanded toolbar with *Properties*, *Design*, *Permissions*, *Versions*, and *Move/Delete* buttons. *Properties* opens a window with fields for the site name and description. *Design* opens a window with options for page layout and theme selection, including a link to get more themes; however, you'll need to connect your site to the concrete5 community site before you can access add-ons and themes.

You can select the *Permissions* button to set user permission levels and use the

Versions button to see page version archives (Figure 8).

Selecting *Dashboard* in the top-right side of the screen opens a window with a file manager, reports, user and group management, a scrapbook, pages and themes, an option to add functionality, backup and update options, a sitewide settings option, and an overview of your site activity (Figure 9). Clicking on elements within your web pages opens up windows for editing (Figure 10).

concrete5 version 5.5 should be released before the end of 2011. "It should include a number of performance improvements, improved role-based permissions, a significantly redesigned dashboard and installation experience, and a few things we'll probably keep secret for a little while yet," Emblar says.

I enjoyed the clean, minimalist design of concrete5 and the intuitive interface. If Joomla, Drupal, WordPress, and eZ Publish are overkill for your needs and you want something fresh and functional, try concrete5. ■■■

INFO

- [1] Joomla: <http://www.joomla.org/>
- [2] Joomla demo: <http://demo.joomla.org/>
- [3] Getting started: <http://www.joomla.org/about-joomla/getting-started.html>
- [4] concrete5: <http://www.concrete5.org/>
- [5] concrete5 demo: <http://www.concrete5.org/about/trial/>

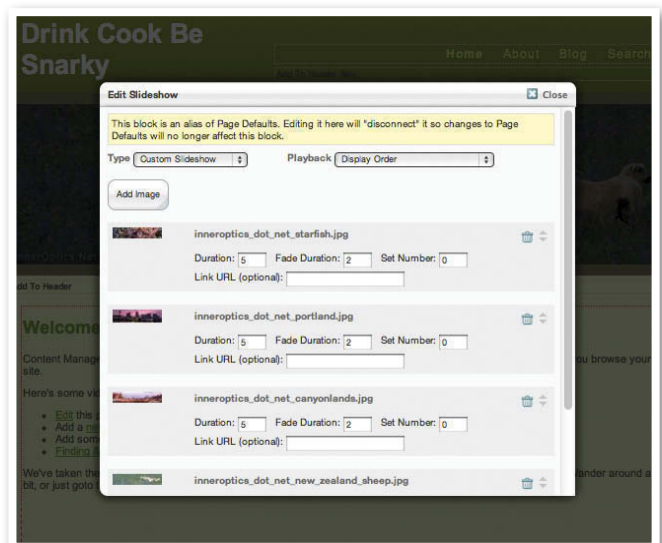


Figure 10: Editing content is easy when you can just click on it and open a new window.



Seattle, WA
Connecting communities
through HPC

November 12-18, 2011

**Seattle, WA
Washington State
Convention Center**

SC11, the premier international conference on high performance computing, networking, storage and analysis, will convene **November 12-18, 2011 in Seattle, Washington.** This year's conference will place an emphasis on bringing together communities to facilitate information exchange, discussions and new collaborations for research and education related to innovating high performance computing applications and advancing scientific discovery and scholarship.

Conference firsts:

- a full day of technical sessions on Friday, November 18, the last day
- the most Technical Program submissions ever
- a State of the Practice track
- a Scientific Visualization Showcase
- the largest number of exhibitors ever—nearly 300

SC11 Thrust:

Data Intensive Science—the conference thrust—focusing on the challenges and opportunities for addressing the exponential growth and demands in the generation and analysis of data.

Technical Program Focus:

Sustained Performance—achieving real, measurable productivity within an elapsed time using cutting-edge advances in high performance computing, networking,

storage and analysis across a diverse range of science and engineering disciplines.

SC11 will continue to recognize exceptional individual accomplishments through the awards and fellowship programs. The Communities Program offers opportunities for people new to SC to engage through the Education Program, Broader Engagement, International Ambassadors, Student Volunteers, Student Cluster Challenge, and the annual Student Job Fair.

Join the SC11 conversation on Facebook, Twitter and LinkedIn and look for podcasts, videos and blogs.

We look forward to your participation!

sc11.supercomputing.org

Sponsors:



SC11 returns to Seattle, a city noted as a conference favorite by attendees. Why?

- Hotels are within walking distance of the convention center
 - The downtown area is vibrant—even after dark
- Experiences abound—Pike Place Market, Pioneer Square, Puget Sound, Space Needle, and much more.

FEATURED EVENTS

Users, developers, and vendors meet at Linux events around the world. We at *Linux Magazine* are proud to sponsor the Featured Events shown here. For other events near you, check our extensive events calendar online at <http://linux-magazine.com/events>.

If you know of another Linux event you would like us to add to our calendar, please send a message with all the details to events@linux-magazine.com.



Hacker Halted USA 2011

Date: Oct 21-27, 2011

Location: Miami, FL, USA

Website: www.hackerhalted.com/2011

The Hacker Halted security conference offers 3 days of intensive conference sessions, plus live training with security experts and other extras for IT professionals.

FSCONS 2011

Date: Nov 11-13, 2011

Location: Gothenburg, Sweden

Website: fscons.org

The yearly FSCONS is the Nordic countries' largest gathering for free software. FScons brings together roughly 250-300 participants, primarily from northern Europe.

linux.conf.au 2012

Date: Jan 16-20, 2012

Location: Ballarat, Australia

Website: linux.conf.au

linux.conf.au is a technical meeting dedicated to Linux and FLOSS. The conference features 5 days of community-driven miniconfs, lectures, tutorials, and social events.

EVENTS

Droidcon London 2011	October 06-07	London, UK	uk.droidcon.com/
LibreOffice Conference 2011	October 12-15	Paris, France	conference.libreoffice.org/
Blackberry DevCon 2011	October 18-20	San Francisco, CA, USA	www.blackberrydevcon.com/
Contact Summit 2011	October 20	New York, NY, USA	contactcon.com/
6th International Workshop on Plan 9	October 20-21	Madrid, Spain	iwp9.org/
Hacker Halted USA 2011	October 21-27	Miami, FL, USA	www.hackerhalted.com/2011
Linux Kernel Summit 2011	October 23-25	Prague, Czech Republic	events.linuxfoundation.org/events/linux-kernel-summit
LinuxCon Europe	October 26-28	Prague, Czech Republic	events.linuxfoundation.org/events/
Apps World North America 2011	November 01-02	New York City, NY, USA	www.apps-world.net/northamerica/
ApacheCon NA 2011	November 07-11	Vancouver, Canada	na11.apachecon.com/
Libre Software World Conference 2011	November 09-10	Zaragoza, Spain	www.asolif.es/?page=fechas_lswc_2011
Grace Hopper Celebration of Women in Computing	November 09-12	Portland, OR, USA	gracehopper.org/2011/
FSCONS 2011	November 11-13	Gothenburg, Sweden	fscons.org/
SC11: SuperComputing 2011	November 12-18	Seattle, WA, USA	sc11.supercomputing.org/
Open Source Developers Conf. 2011	November 16-18	Canberra, Australia	osdc.com.au
Linuxday Vorarlberg 2011	November 26	Dornbirn, Austria	www.linuxday.at
TakeDownCon - Mobile & Wireless Security	December 02-07	Las Vegas, NV, USA	www.takedowncon.com/?page_id=165
LISA '11	December 04-09	Boston, MA, USA	www.usenix.org/events/
LinuxCon Brazil 2011	December 07-08	São Paulo, Brazil	events.linuxfoundation.org/events/linuxcon-brazil
linux.conf.au 2012	January 16-20	Ballarat, Australia	linux.conf.au/
Cloud Expo Europe 2012	January 25-26	London, UK	www.cloudexpo-europe.com
The Mobile Show	April 17-18	Dubai, UAE	www.terrapinn.com/2012/the-mobile-show/

CALL FOR PAPERS

We are always looking for good articles on Linux and the tools of the Linux environment. Although we will consider any topic, the following themes are of special interest:

- System administration
- Useful tips and tools
- Security, both news and techniques
- Product reviews, especially from real-world experience
- Community news and projects

If you have an idea, send a proposal with an outline, an estimate of the length, a description of your background, and contact information to edit@linux-magazine.com.

The technical level of the article should be consistent with what you normally read in *Linux Magazine*. Remember that

Linux Magazine is read in many countries, and your article may be translated into one of our sister publications. Therefore, it is best to avoid using slang and idioms that might not be understood by all readers.

Be careful when referring to dates or events in the future. Many weeks could pass between your manuscript submission and the final copy reaching the reader's hands. When submitting proposals or manuscripts, please use a subject line in your email message that helps us identify your message as an article proposal. Screenshots and other supporting materials are always welcome.

Additional information is available at:
http://www.linux-magazine.com/contact/write_for_us.



AUTHORS

Erik Bärwaldt	74
Mario Blättermann	54
Joe "Zonker" Brockmeier	86
Zack Brown	90
Gavin W. Burris	14
Bruce Byfield	82
Mela Eckenfels	28
Jon "maddog" Hall	89
Olaf Hübner	43
Rikki Kite	92
Klaus Knopper	66
Charly Kühnast	57
Hans-Peter Merkel	38
Dmitri Popov	70
Mike Schilli	48
Tim Schürmann	32, 58
Kurt Seifried	46

CONTACT INFO

Editor in Chief

Joe Casad, jcasad@linux-magazine.com

Managing Editor

Rita L Sooby, rsooby@linux-magazine.com

Production Editor

Lori White, lwhite@linux-magazine.com

Contributing Editors / Test Lab Staff

Uli Bantle, Andreas Bohle, Jens-Christoph Brendel, Hans-Georg Eßer, Markus Feilner, Oliver Frommel, Marcel Hilzinger, Mathias Huber, Kristian Kißling, Jan Kleinert, Thomas Leichtenstern, Jörg Luther

Localization & Translation

Ian Travis

Proofreading & Polishing

Amber Ankerholz

Layout

Judith Erb, Dena Friesen, Lori White

Cover Design

pinball digitalagentur,
lohfert@pinball-digitalagentur.de

Advertising

www.linux-magazine.com/Advertise

United Kingdom and Ireland

Penny Wilby, pwilby@linux-magazine.com
phone +44 1787 211 100

North America

Ann Jesse, ajesse@linuxpromagazine.com
phone +1 785 841 8834
Eric Henry, ehenry@linuxpromagazine.com
phone +1 785 917 0990

All other countries

Petra Jaser, pjaser@linuxnewmedia.com
phone +49 89 9934 1124

Corporate Management (Vorstand)

Hermann Plank, hplank@linuxnewmedia.de
Brian Osborn, bosborn@linuxnewmedia.com

Publisher

Brian Osborn, bosborn@linuxnewmedia.com

Marketing Communications

Julie Miller, jmiller@linuxnewmedia.com

Product Management

Hans-Jörg Ehren, hjehren@linux-magazine.com

Customer Service / Subscription

For USA and Canada:
Email: cs@linuxpromagazine.com
Phone: 1-866-247-2802
(Toll Free from the US and Canada)
Fax: 1-785-856-3084

For all other countries:
Email: subs@linux-magazine.com
Phone: +49 89 9934 1167
Fax: +49 89 9934 1199

Linux Magazine – Putzbrunner Str 71
81739 Munich – Germany

www.linux-magazine.com – Worldwide

www.linuxpromagazine.com – North America

www.linuxmagazine.com.au – Australia

www.linux-magazine.co.uk – United Kingdom

While every care has been taken in the content of the magazine, the publishers cannot be held responsible for the accuracy of the information contained within it or any consequences arising from the use of it. The use of the disc provided with the magazine or any material provided on it is at your own risk.

Copyright and Trademarks © 2011 Linux New Media Ltd.

No material may be reproduced in any form whatsoever in whole or in part without the written permission of the publishers. It is assumed that all correspondence sent, for example, letters, email, faxes, photographs, articles, drawings, are supplied for publication or license to third parties on a non-exclusive worldwide basis by Linux New Media unless otherwise stated in writing.

Linux is a trademark of Linus Torvalds.

Microsoft, Windows, and the Windows logo are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

Printed in Germany

Distributed by COMAG Specialist, Tavistock Road, West Drayton, Middlesex, UB7 7QE, United Kingdom

Approximate

UK / Europe

USA / Canada

Australia

Nov 07

Dec 02

Jan 02

On Sale Date

Issue 133 / December 2011

Web Tools

New tools, new techniques, new smart devices, and lots of new businesses going on line every day – sounds like a good time for a visit to the ever-changing world of web admin and development. Next month, we focus on some important tools for building and managing a new generation of websites.

Preview Newsletter

The Linux Magazine Preview is a monthly email newsletter that gives you a sneak peek at the next issue, including links to articles posted online.

Sign up at: www.linux-magazine.com/newsletter

Denis Tevelov, 123RF.com



**New
Version
BackupEDGE 3.0.3
with SharpDrive™
Support**

FULL LINUX SERVER BACKUPS TO ALMOST ANY MEDIUM WITH BARE METAL RECOVERY

BackupEDGE™ creates easy, verified data backups with available high strength encryption and key management. Choose the storage medium best suited to protect your critical data and systems, including...

- * **SharpDrive™** removable disk / cartridge disk / flash drive storage.
- * **Blu-ray Disc™** writers.
- * Amazon S3 Internet cloud storage.
- * Network Attached Storage (intranet or Internet) using FTP / FTPS Backups.
- * Tape drives (including changers/libraries).
- * Iomega® REV® drives (including loaders).
- * Standard CD and DVD writers.
- * Any combination.

RecoverEDGE™ bare metal disaster recovery is fully supported from any medium type. It is included with every copy of *BackupEDGE* for Linux.



What's a SharpDrive?

SharpDrive is our name for **any** USB hard drive, flash drive, or cartridge drive when used as a backup / bare metal recovery device for Linux. Capabilities include...

- * Wizard-based media formatting / initialization.
- * Instant File Restore™
- * Hot plug multiple devices at once, with automatic selection based on most available space.
- * Variable compression, optional encryption
- * Multiple archives per medium with lazy reclamation.
- * Direct booting into bare metal recovery.
- * MySQL™ hot backups.
- * Mix and match media types as needed.
- * Unlimited scheduling flexibility.

Browse to: <http://www.microlite.com> right now for additional information about **BackupEDGE 3.0.3!**

Only *BackupEDGE* has an advanced, **standards-based storage format** allowing capture and protection of ALL of your data. Features include:

- * Pathnames - up to 5,000 characters
- * Integrity checksums (file level)
- * 9 selectable compression levels (file level)
- * Archive / subset-of-archive encryption (file level)
- * Access control list (ACL) support

Transparent Media™ technology allows virtually instantaneous restore of individual files and directories from most archive resources, and the fastest possible tape positioning and restore.

Microlite Corporation has been in business since 1983. We've been protecting SMB and edge-of-enterprise UNIX and Linux servers since 1987. Visit our website today to download a fully functional 60-day evaluation copy of *BackupEDGE*.

<http://www.microlite.com>

Resellers Wanted



Cut Execution Time by >50% with WhisperStation-GPU

Delivered ready to run new GPU-enabled applications:



WhisperStation with 4 Tesla Fermi GPUs

Design	Simulation		BioTech
3ds Max	ANSYS Mechanical	MATLAB	AMBER
Bunxspeed	Autodesk Moldflow	ACUSIM AcuSolve	GROMACS
Shot	Mathematica	Tech-X GPUlib	NAMD, VMD
Adobe CS5			TeraChem

Integrating the latest CPUs with NVIDIA Tesla Fermi GPUs, Microway's WhisperStation-GPU delivers 2x-100x the performance of standard workstations. Providing explosive performance, yet quiet, it's custom designed for the power hungry applications you use. Take advantage of existing GPU applications or enable high performance with CUDA C/C++, PGI CUDA FORTRAN, or OpenCL compute kernels.

- ▶ Up to Four Tesla Fermi GPUs, each with: 448 cores, 6 GB GDDR5, 1 TFLOP single and 515 GFLOP double precision performance
- ▶ Up to 24 cores with the newest Intel and AMD Processors, 128 GB memory, 80 PLUS® certified power supply, and eight hard drives
- ▶ Nvidia Quadro for state of the art professional graphics and visualization
- ▶ Ultra-quiet fans, strategically placed baffles, and internal sound-proofing
- ▶ New: Microway CL-IDE™ for OpenCL programming on CPUs and GPUs

Microway's Latest Servers for Dense Clustering

- ▶ 4P, 1U nodes with 48 CPU cores, 512 GB and QDR InfiniBand
- ▶ 2P, 1U nodes with 24 CPU cores, 2 Tesla GPUs and QDR InfiniBand
- ▶ 2U Twin² with 4 Hot-Swap MBs, each with 2 Processors + 256 GB
- ▶ 1U S2050 servers with 4 Tesla Fermi GPUs

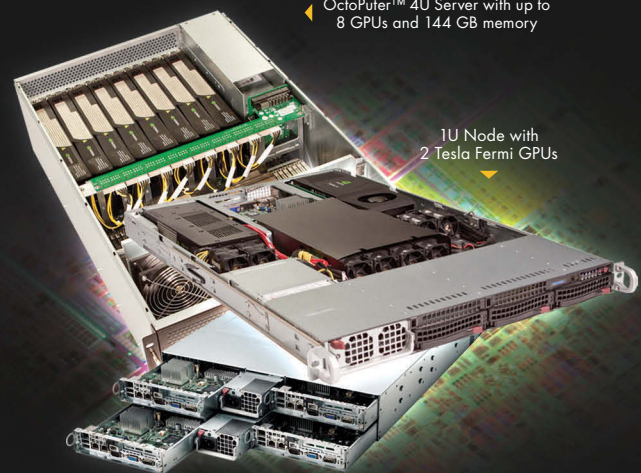
Microway Puts YOU on the Cutting Edge

Design your next custom configuration with techs who speak HPC. Rely on our integration expertise for complete and thorough testing of your workstations, turnkey clusters and servers. Whether you need Linux or Windows, CUDA or OpenCL, we've been resolving the complicated issues – so you don't have to – since 1982.

Configure your next WhisperStation or Cluster today!

microway.com/quickquote or call 508-746-7341

Sign up for technical newsletters and special GPU promotions at microway.com/newsletter



OctoPuter™ 4U Server with up to 8 GPUs and 144 GB memory

1U Node with 2 Tesla Fermi GPUs

2U Twin² Node with 4 Hot-Swap Motherboards
Each with 2 CPUs and 256 GB



GSA Schedule
Contract Number:
GS-35F-0431N



Visit Microway at SC11 in Booth 2606