

FREE DVD



# CentOS 7

Enterprise community distro sponsored by Red Hat



# Tails 2.2

the amnesic in cognito live system

# EDUCATIONAL LINUX

Distros for the classroom

LINUX PRO



# LINUX

## MAGAZINE

MAY 2016

# EDUCATIONAL LINUX

## Delve into distros for the classroom

### 5 Fine Backup Tools for the Linux desktop

### Prometheus

Does this next-generation monitoring tool steal the fire?

### Build Your Own Search Engine



### Stellarium

Star gazing in the great open source planetarium

### SQLite

This database tool is already running on your smartphone or Linux

### Wayland

Exploring the new graphics display protocol

### PhotoFilmStrip

Turn still shots into a photo documentary



Issue 186	US\$ 12.99
May 2016	CAN\$ 13.99
0 74470 58049 2	

**HETZNER**  
ONLINE

**Dedicated Servers**  
**for all users.**

★ ★ ★  
ATTRACTING  
LOYAL CUSTOMERS  
OF ALL TYPES

Since 1997

**GAMER  
ON  
BOARD**

I ♥  
SOFTWARE  
DEVELOPERS

I ♥  
PIXELS



**MANAGER**

**WORLD'S BEST  
ADMIN**

I ♥ My  
Ceo



## Dedicated Root Server EX51

Intel® Core™ i7-6700  
Quad-Core Skylake Processor  
64 GB DDR4 RAM  
2 x 4 TB SATA HDD Class Enterprise  
Guaranteed 1 Gbit/s bandwidth  
100 GB Backup Space  
30 TB traffic inclusive\*  
No minimum contract  
Setup Fee \$109

monthly \$ **54**

## Dedicated Root Server EX51-SSD

Intel® Core™ i7-6700  
Quad-Core Skylake Processor  
64 GB DDR4 RAM  
2 x 500 GB SATA SSD  
Guaranteed 1 Gbit/s bandwidth  
100 GB Backup Space  
30 TB traffic inclusive\*  
No minimum contract  
Setup Fee \$109

monthly \$ **54**

\* There are no charges for overage. We will permanently restrict the connection speed if more than 30 TB/month are used. Optionally, the limit can be permanently cancelled by committing to pay \$1.30 per additional TB used.

[www.hetzner.de/us](http://www.hetzner.de/us)

All prices exclude VAT and are subject to the terms and conditions of Hetzner Online GmbH. Prices are subject to change. All rights reserved by the respective manufacturers, Intel, Intel Logo, Intel Core and Core Inside are brands of the Intel Corporation in the USA or other countries.

# OLD AND NEW

## Dear Linux Pro Reader,

In December 2015, the Hactivist group Anonymous declared war on US Presidential candidate Donald Trump, citing outrage over Trump's disparaging comments about Muslims and his proposal to ban all Muslims from entering the United States. The group launched a denial-of-service attack on the Trump-Tower.com website, online presence for the bombastic billionaire's signature real estate development, and took it offline temporarily. They also broke into his cell phone and exposed some voicemail messages from journalists.

The attacks didn't seem to have much affect on the candidate's campaign. Trump continued to draw large crowds and win primary victories. Then, another YouTube video appeared on March 4, apparently from a rival faction within Anonymous that was dissatisfied with the skills and choices of the first group. The much-doctored voice of a speaker behind the familiar Guy Fawkes mask states, "You may remember me as the one who organized the complete shutdown of Anonymous Operations, a branch of Anonymous that successfully waged cyberwar against Turkey, Hungary, CloudFlare, and Donald Trump. You may think I silenced these hackers as a Donald Trump supporter, but that is far from the truth. In fact [the attack on] Trump is the only thing they accomplished that we admired." The new group said it was renewing the attack on Trump, and not just as a war but as a "total war."

I'm certainly no fan of Donald Trump, but I'm also not a fan of cyber warfare and the silencing of politicians, no matter how outrageous their ideas. If you asked me to place a bet, I would say I don't think these new attacks are going to have much effect. Why? For one thing, the web does not really appear to be so critical to the Trump campaign. Candidates use websites to raise money, and Trump isn't raising any money – he's spending his own money. Another use for a candidate's website is to provide detailed policy papers, and Trump doesn't seem to have many detailed policies. Other candidates have been brought down by personal secrets made public or inappropriate comments revealed to the world, but it is difficult to imagine what they might discover about Donald Trump that would actually change public perceptions in a meaningful way.

The Barack Obama campaigns of 2008 and 2012 were known for a kind of Internet fluency that, according the experts, heralded a new approach to presidential campaigns for a new century. Donald Trump actually reverses that presumed tide of history by putting the emphasis back on conventional

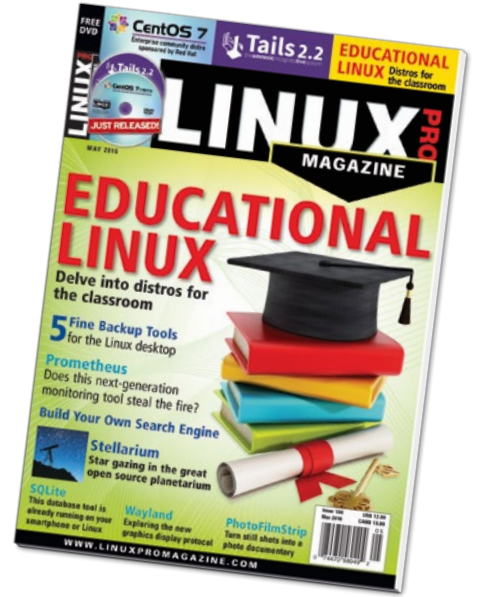
media forms like cable television. His deft use of Twitter notwithstanding, Trump is able to tell his story through conventional media because he actually *treats* conventional media as if it were social media. A stump speech, an interview, a sound bite from a candidate forum – Donald Trump makes the mainstream TV seem a little like his own personal flame blog, with a few recurring ideas and lots of space for rants about things (and people) he doesn't like. Who needs Facebook when you can make the news networks behave like your Facebook account?

We don't know how this current Anonymous faction will pursue its "total war," but one thing is certain: To stop Trump, they'll need a lot more than a phone hack and a denial-of-service attack on his website.

*Joe*

Joe Casad,  
Editor in Chief





# LINUX PRO MAGAZINE

## news

### 8 EMAIL FOUNDER DIES

- Ray Tomlinson, inventor of email, dies.
- Linux SQL Server announced.

### 9 MOUSE DONGLE ATTACK

- "Mousejack" attack allows access via mouse dongle.
- Open source in-vehicle infotainment system.
- Google ends support for Flash ads.

### 10 CYBERSECURITY INITIATIVE

- Executive orders to improve security on US computers.
- Increased funding for US HPC.
- BeeGFS cluster filesystem now open source.

### 11 LINUX BACKDOOR

- Fysbis malware does not need root access.
- Glibc bug could compromise security.

## 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.

# Educational Linux

Special Linux distros answer the call for putting computer literacy in the classroom.

**12 Educational Linux** The free Linux open source ecosystem is a strong contender for putting computer education in the classrooms. This article looks at six educational Linux distros.



## Community Notebook

### 90 Doghouse: FOSS Training

maddog examines what today's engineers and product managers need to know about Free and Open Software.



### 92 Kernel News

Some changes in software RAID maintainership and an overhaul of the kernel documentation system.



## HIGHLIGHTS

**44 PROMETHEUS**  
An alternative to classic monitoring solutions for complex environments.

**56 SQLITE TUTORIAL**  
Dive into the SQLite databases on all your devices.

**76 PHOTOFILMSTRIP**  
Build a documentary video from still photos with the Ken Burns effect.

**80 STELLARIUM EXTRAS**  
Delve deeper into the universe with this virtual planetarium.

# FEATURES

## 38 Do-It-Yourself Search Engine

Build your own search engine using Apache's Nutch web crawler and the Solr search platform.



## 50 Wayland

The X11 graphics protocol is showing some serious signs of age, but Wayland is poised to come to the rescue.

**44 Prometheus**  
Complex environments benefit from a different approach to network monitoring. Prometheus is an interesting alternative.

**56 SQLite Tutorial**  
Several databases likely reside on your desktop and smartphone. It is easy to manage this data or create a similar database yourself.

**62 Charly's Column – Doorbell Pi**  
A Raspberry Pi Zero, a noise detector, and a power outlet with a LAN connection alert you to visitors when you can't hear the doorbell.

**64 Perl – WiFi Connect Messages**  
Push notifications to your cell phone keep you up to date with the details of clients that log on or off of your home WiFi.

## LINUXUSER

- 70 Workspace – Fossil SCM**  
A software configuration management system for your next great software project.
- 76 PhotoFilmStrip**  
Video slideshows from still photos.
- 80 Stellarium Extras**  
Expand the Stellarium virtual planetarium.
- 86 Command Line – tre-agrep**  
All of grep's functionality with ambiguous or fuzzy searches and minimal regex.

## REVIEWS

- 20 Backup Solutions**  
Five programs for the desktop.
- 28 AfterShot Pro**  
Process and organize photos and RAW files.
- 34 Tool Tips**  
Tool tests on the fast track.



**CentOS 7**  
Enterprise community distro sponsored by Red Hat

- 64-bit install only
- Updated networking stack
- Live patching with kpatch
- Btrfs, OverlayFS, DNSSEC support



**Tails 2.2**  
the amnesic incognito live system

- Cryptographic tools
- Tor network
- Leave-no-trace desktop
- Anonymity and privacy

SEE PAGE 6 FOR DETAILS

# On the DVD



## CentOS 7 (1511), 64-bit INSTALL ONLY

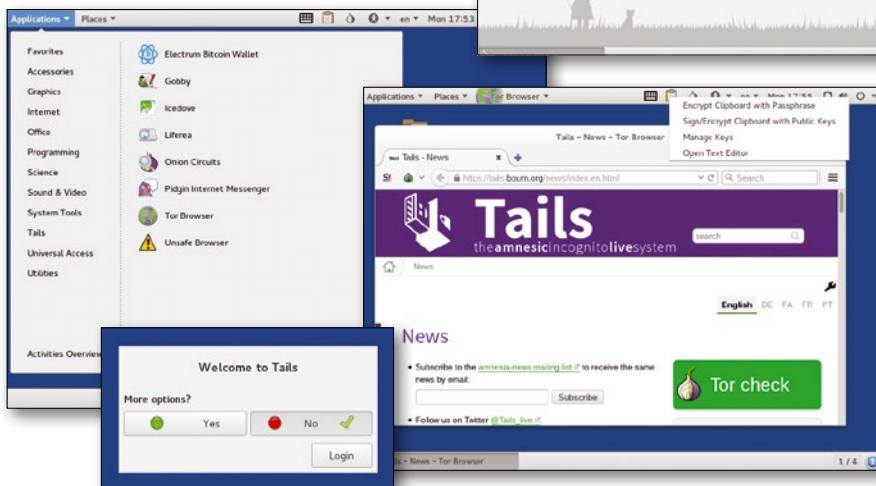
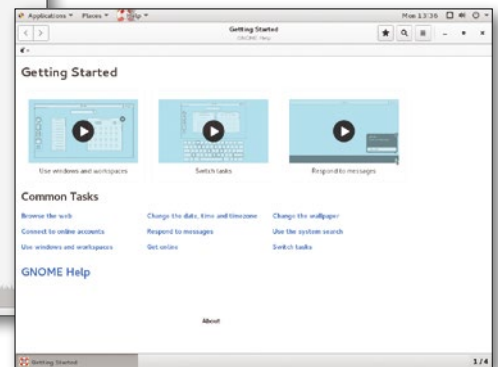
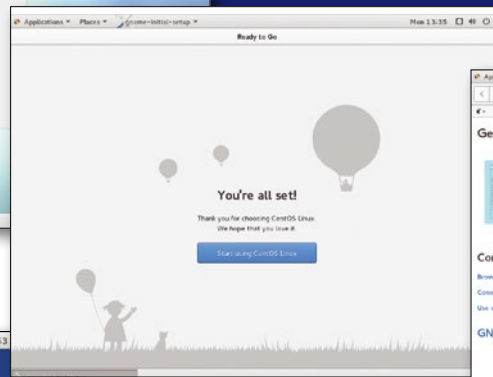
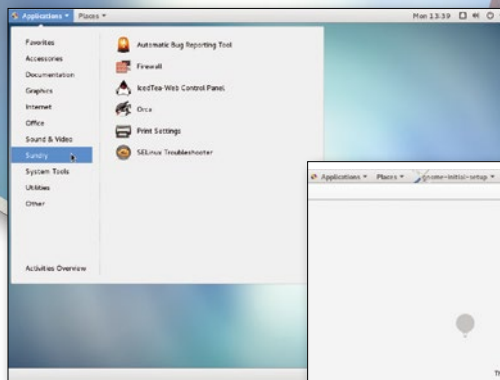
The version of CentOS 7 on this month's DVD is the third major release derived from Red Hat Enterprise Linux 7.2. This installation-only image offers a number of environments, including file, print, and web servers, as well as Gnome and KDE Plasma workspaces. CentOS 7 features:

- an updated networking stack
- TLS 1.1/1.2 support in some packages
- Btrfs, OverlayFS, and DNSSEC support
- live patching (kpatch)

## Tails 2.2, 32-bit Live

Tails is a Live operating system based on Debian that helps you preserve privacy and anonymity through the use of cryptographic tools, the Tor network for Internet activities, and leave-no-trace desktop policies. In addition to bug fixes, you'll find:

- Tor browser updates
- security fixes to glibc, cpio, LibreOffice, and libssh2
- support for viewing DRM-protected DVD videos
- KeePassX database saves after every change



## ADDITIONAL RESOURCES

- [1] CentOS Project: <https://www.centos.org>
- [2] CentOS 7 release notes: <https://wiki.centos.org/Manuals/ReleaseNotes/CentOS7>
- [3] CentOS forums: <https://www.centos.org/forums/viewforum.php?f=44>
- [4] Tails: <https://tails.boum.org>
- [5] Tails installation: <https://tails.boum.org/install/index.en.html>
- [6] Tor: <https://www.torproject.org>

Defective discs will be replaced. Please send an email to [cs@linuxpromagazine.com](mailto:cs@linuxpromagazine.com).



# APPS

## READY FOR THE CLOUD!



**1&1 CLOUD SERVER**  
**TRY IT**  
**TODAY!\***

With the **1&1 Cloud App Center** you can get your applications up and running in no time. Choose from more than **100 cutting-edge apps** and combine them with the superior speed and performance of the **best value cloud server around**.

- ✓ **Secure and powerful platform**
- ✓ **No prior server knowledge necessary**
- ✓ **Billing by-the-minute**



1 (877) 461-2631



**1and1.com**

\*1&1 Cloud Server is free for one month, after which the regular pricing starts from \$4.99/month. One-time setup fee of \$9.99 applies. 1&1 Cloud Server comes with a 30-day money back guarantee. If your package is not canceled within the first 30 days, you will be charged in accordance with your agreed billing cycle. Visit 1and1.com for full offer details, terms and conditions. Intel, the Intel Logo are trademarks of Intel Corporation in the U.S. and/or other countries. 1&1 and the 1&1 logo are trademarks of 1&1 Internet, all other trademarks are property of their respective owners. ©2016 1&1 Internet. All rights reserved.

# NEWS

Updates on technologies, trends, and tools

## THIS MONTH'S NEWS

### 08 Email Founder Dies

- Ray Tomlinson, inventor of email, dies
- Linux SQL Server announced

### 09 Mouse Dongle Attack

- "Mousejack" attack allows access via mouse dongle
- Open source in-vehicle infotainment system
- Google ends support for Flash ads
- More Online

### 10 Cybersecurity Initiative

- Executive orders to improve security on US computers
- Increased funding for US HPC
- BeeGFS cluster filesystem now open source

### 11 Linux Backdoor

- Fysbis malware does not need root access
- Glibc bug could compromise security
- More Online

## Inventor of Email Dies

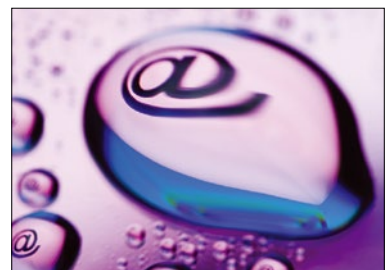
Ray Tomlinson, inventor of the global communication technique known as email, has died at the age of 74. The cause of death is not confirmed at this writing, but early reports suggest a heart attack. Tomlinson was born in Amsterdam, New York, in 1941 and attended Rensselaer Polytechnic University, graduating with a degree in electrical engineering in 1963. He earned a Masters degree from MIT in 1965.

***"I'm simply trying to preserve the world's supply of hyphens."***

Tomlinson invented email in 1971 while working at the research and development company Bolt, Beranek, and Newman (now BBN, a subsidiary of Raytheon) on supporting software for ARPANET, the predecessor of the Internet. He adapted a program called SNDMSG, which was intended to send messages to terminal users of a single time-sharing system, so it could send messages to other computers. Tomlinson chose the @ symbol as a handy indicator that would target the message to a single user of the destination machine.

When asked why he took the time in the middle of another project to develop the program for email, Tomlinson reported "because it seemed like a neat idea." After showing the new system to a colleague, he famously warned, "Don't tell anyone! This isn't what we're supposed to be working on."

According to the entry in Wikipedia, Tomlinson preferred "email" over "e-mail," stating in a 2010 interview, "I'm simply trying to preserve the world's supply of hyphens."



© Sebastian Duda, 123RF.com

## Microsoft Announces Linux Version of SQL Server

Microsoft has announced that it is creating a Linux version of its SQL Server database server tool. SQL Server has long been a core component of the Windows Server ecosystem, and the strange appearance of a Linux version underscores how far Microsoft has come in recent years to embrace the possibilities of Linux.

According to Microsoft VP AI Gillen, "This is an enormously important decision for Microsoft, allowing it to offer its well known and trusted database to an expanded set of customers. By taking this key product to Linux, Microsoft is proving its commitment to being a cross platform solution provider."

Of course, Microsoft doesn't do anything for free or for fun. The prevalence of Linux systems running in the cloud, including Microsoft's own Azure cloud, further confirms that Linux is here to stay. By porting SQL Server to the Linux platform, Microsoft can continue to exert an influence over SQL database systems, even as the underlying OS migrates to Linux.

The private preview version of SQL Server on Linux is available now, and Microsoft is aiming to have a release version ready by mid-2017.



## Strange New Attack Lets an Intruder Gain Access Through a Mouse Dongle

Researchers at Bastille Networks have discovered a flaw in some keyboard and mouse dongles that could allow an attacker located within 100 meters of the computer to execute commands on the system through the dongle. The computer receives the commands as if they were typed by the user sitting at the system.

The "Mousejack" attack capitalizes on a gap in the security of non-Bluetooth mice and keyboards from seven different vendors that use Logitech's Unifying wireless USB technology. According to Bastille, "Once paired, the Mousejack operator can insert keystrokes or malicious code with the full privileges of the PC owner and infiltrate networks to access sensitive data. The attack is at the keyboard level, therefore PCs, Macs, and Linux machines using wireless dongles can all be victims."

The Bastille Networks website has a dramatic video (<https://www.bastille.net/>) showing how the attack works. See the article at the Threatpost site for additional information: <https://threatpost.com/mousejack-attacks-abuse-vulnerable-wireless-keyboard-mouse-dongles/116402/>.



© Fernando Gregory, 123RF.com

## GDP-ivi: Linux In-Vehicle Infotainment Platform

The GENIVI Alliance has just released GDP-ivi9, an open source demo platform for in-vehicle infotainment developers and manufacturers. The GNU/Linux-based platform allows developers to showcase how the middleware of a free in-vehicle infotainment system will work.

This beta software comes in the shape of a downloadable virtual machine that simulates the inner workings of GENIVI's in-vehicle infotainment system. You can, for example, tweak the audio signals (music, phone, text to speech), try out the widgets used to create interfaces, or see a simulation of the 3D navigation system.

GDP-ivi is designed for the engineers who implement the system; that is, it is an interface to the middleware, not for the end users. The system showcases many GENIVI projects that are already available as free software: <https://at.projects.genivi.org/wiki/pages/viewpage.action?pageId=2785429>.

Members of the GENIVI Alliance include auto industry heavyweights such as BMW, Renault, Honda, Volvo, Jaguar, and more, alongside automotive component manufacturers and middleware, hardware, and service suppliers.

You can read more about GDP-ivi9 and download the virtual machine from GENIVI's open source site: <https://at.projects.genivi.org/wiki/pages/viewpage.action?pageId=11567210>.

## Google Announces End to Flash Ad Support

The embattled Flash framework just received another blow as ad giant Google announced an end to its support for Flash display ads. According to the announcement, Google's AdWords and DoubleClick digital marketing units will cease to accept ads built in the Flash platform starting June 30, 2016. Ads in the Flash format will no longer appear on the Google Display network after January 2, 2017.

Web surfers around the world are accustomed to the animated Flash ads that appear on websites. Google has revolutionized the ad business and now pushes ads out to millions of websites that didn't use to have them.

Flash has been in the news recently for numerous security problems, but even apart from the security issues, the Flash format is fast becoming obsolete. New features available in HTML5 provide a wide range of animation and graphic capabilities that used to require Flash.

Google's position in the market as the world's largest ad server means this change will likely result in other similar services pulling their support for Flash.

## MORE ONLINE

### Linux Pro Magazine

[www.linuxpromagazine.com](http://www.linuxpromagazine.com)

#### Off the Beat • Bruce Byfield

##### Wesnoth's Strange Legacy

I usually avoid massively multiplayer online role-playing games (MMORPGs). Instead, when I want more than solitaire or backgammon, I sometimes turn to The Battle for Wesnoth single-player campaigns, most of which can be finished in under an hour, giving me a convenient point at which to quit and do something else. However, that usually reliable strategy led me astray soon after Christmas 2015 when I discovered the Strange Legacy scenario.

#### Miguel de Icaza and His Ostracization from FOSS

Just before I settled down to write today, I read that Microsoft had acquired Xamarin, the company founded by Miguel de Icaza and Nat Friedman. To many, the news is the logical end to a story that has been unfolding for years now, and if the first cries of, "Traitors!" have not appeared on blogs and articles, then I expect they are only a matter of time.

#### Is free software too good?

I never thought I'd say this, but maybe free software and hardware would be more widely adopted if their standards were lower.

#### Productivity Sauce • Dmitri Popov

##### Instant Streaming from Linux to Chromecast with stream2chromecast

What's the easiest way to stream media files from a Linux machine to Chromecast? The correct answer is stream2chromecast. This simple tool makes the task of streaming media files from a Linux machine to a Chromecast device ridiculously easy.

#### Extension Watch: Beef up Privacy Protection with Decentraleyed for Firefox

uBlock Origin and Privacy Badger make a perfect combo for protecting your privacy and combating annoying ads. But these extensions can't protect you against more insidious ways of tracking your activities and collecting personal data through content delivery networks (CDN). Enter Decentraleyed.

#### twtxt: Microblogging for Hackers

Maybe you are not keen on using closed and tightly-controlled third-party microblogging services like Twitter. Or maybe you are just looking for a straightforward way to share your thoughts with the world directly from the command-line. In either case, twtxt might be something right up your alley.

## President Obama Launches New Cybersecurity Initiative

President Obama signed two executive orders to improve security on US government networks. The executive orders are reflected in the 2017 budget the president sent to Congress, which calls for \$19 billion in upgrades for government information technology.

One of the orders creates a Commission on Enhancing National Cybersecurity, which makes recommendations for building better security on both public and private networks. The second order launches the Federal Privacy Council, which will bring together privacy officials from 25 government agencies to help ensure more uniform protection of citizen information.

The initiatives are intended to promote better security and privacy practices throughout the government and to keep the emphasis on replacing out-of-date infrastructure. In his press conference launching the programs, Obama pointed out that the Social Security Administration system runs on a 1960s-era platform written in COBOL, which requires up to 400 people to maintain it.

The administration says the plan does not rely on increased funding but shifts existing priorities within the government. See the report at USA Today for more information on the announcement: <http://www.usatoday.com/story/news/politics/2016/02/09/obama-signs-two-executive-orders-cybersecurity/80037452/>.



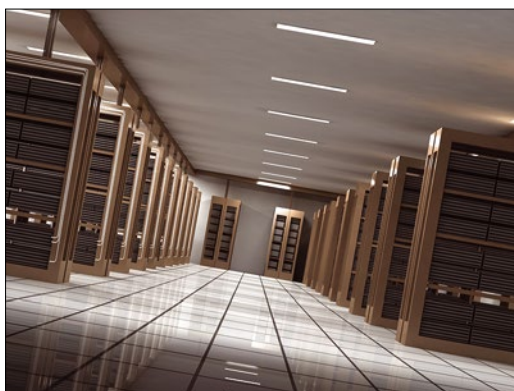
© Stanislaw Tokarski, 123RF.com

## President's 2017 Budget Increases HPC Spending by 50%

President Obama's budget for next year includes an increase in funding for HPC and future computing technologies of more than 50%. The funds for scientific computing are in the US Department of Commerce's National Institute of Standards and Technology (NIST) budget. Allotments for HPC and future computing technologies appear in several budget line items, but most notable is the budget for Measurement Science for Future Computing Technologies and Applications,

which sees an increase from \$13.6 million to \$25.6 million. The total discretionary spending for NIST is \$1 billion.

Of course, the president's proposed budget is a long way from being enacted, with the Republican-controlled congress vowing to challenge the president's funding priorities. Stay tuned for updates and see the article at the Scientific Computing website for additional information.



© Alejandro Mendoza, 123RF.com

## BeeGFS Parallel Filesystem Goes Open Source

The developers of the BeeGFS parallel filesystem have announced that BeeGFS is now available as open source. BeeGFS is a cluster filesystem for I/O-intensive workloads, with the focus on high performance and easy installation. The source code is available for download at the BeeGFS website.

## MORE ONLINE

### ADMIN HPC

<http://hpc.admin-magazine.com/>

#### Getting the Most from Your Cores

Jeff Layton

CPU utilization metrics tell you how well your applications are using your processing resources.

### ADMIN Online

<http://www.admin-magazine.com/>

#### Exploring dhcpy6d, a DHCP Server for IPv6 Networks • Konstantin Agouros

Even if you don't have to have DHCP on your IPv6 network, you might want it anyway. Dhcpy6d is a promising DHCP option for IPv6.

#### Ansible as an Alternative to the Puppet Configuration Tool • Martin Loschwitz

Automation is part of life in the data center, and Puppet is commonly regarded as the King of the Hill, but some users prefer the lean alternative Ansible.

#### Web-Based Reconnaissance • David J. Dodd

The recon-ng web reconnaissance framework is an important tool in penetration testing.

#### The SDFS Deduplicating Filesystem • Tim Schürmann

Deduplicating filesystems like SDFS store redundant data, such as that created by backups, only once, thereby saving valuable disk space. Additionally, the filesystem can distribute the data to be stored across multiple computer nodes.

The BeeGFS project, which is sponsored by the professional service company ThinkParQ, was already available free of charge and was typically distributed through turnkey solutions by ThinkParQ's network of international partners.

The project originally announced a desire to release the source code for BeeGFS in 2013 as part of the European Commission's DEEP-ER exascale project. According to ThinkParQ CEO Sven Breuner, "While some of our users are just happy with the fact that BeeGFS is so easy to install and doesn't require much attention, others really want to understand exactly what is happening under the hood to further optimize the runtime of their applications, improve their monitoring, or port it to other platforms like BSD. Also, being able to build BeeGFS for non-x86 architectures like ARM and Power is another important aspect that the community has been waiting for."

The GeeGFS developers are currently working with the European ExaNeST project to "develop and prototype evolved solutions for interconnection networks, persistent storage, and cooling, specifically on ARM systems."

See the press release at the BeeGFS site ([www.beegfs.com/](http://www.beegfs.com/)) for additional information.

## Linux Backdoor Doesn't Need Root Privileges

Researchers have discovered a new backdoor called Fysbis that is aimed at Linux machines. The new malware has extremely sophisticated properties, and experts suspect it might have come from the APT 28 cyber-espionage group, which reportedly has ties with Russia.

One insidious quality of Fysbis is that it doesn't require root access to start working for the attacker. The malware is apparently able to enter the system at a lesser security level and begin reconnaissance, performing tests and sending information back to a remote command and control center with clues for how to escalate privileges. According to a report at TechWorm, Fysbis can "... open a remote shell on the infected machine, run commands on the attacker's behalf, find, read, save, execute, or delete files, and log keyboard input."



© designpics, 123RF.com

## Major Bug in glibc Could Result in System Compromise

Google's Security team has reported a problem with the popular glibc library found in most Linux systems that could result in a serious security breach. The problem affects glibc version 2.9 and later. Specifically, the glibc client-side DNS resolver is vulnerable to a buffer overflow attack that could cause the system to access an attacker-controlled website or DNS server.

The post in the Google security blog reports that the glibc team was first alerted to the bug in July 2015 and that Red Hat has also been working on a fix for this problem. The best remedy is to update your systems and install the patch for CVE-2015-7547 as soon as possible.

If you are not immediately able to patch glibc, Google recommends you "... limit the response (i.e., via DNSMasq or similar programs) sizes accepted by the DNS resolver locally as well as ensure that DNS queries are sent only to DNS servers that limit the response size for UDP responses with the truncation bit set."



© Sergey Nivens, 123RF.com



### Linux systems for the education sector

# Back to School

The fact that Linux and its extensive open source ecosystem are free of charge is really a strong argument for financially strapped schools. This article looks at Linux in the classroom. *By Jörg Thoma*



**E**ducational computing is a big topic – and a big market around the world. Teachers and education experts are looking for new ways to deliver curriculum efficiently with maximum benefit for their students. Many of the available education platforms got their start before the widespread acceptance of Linux, which means Windows and Mac OS systems had a big head start on the market. But, open source innovation has opened the classroom doors to Linux. Several specialized Linux distros compete for the chance to serve curriculum to students.

One of the biggest reasons for considering Free software is that it offers lower cost – and sometimes even has *no* cost. But, FOSS provides other benefits for educational users. For instance, the adaptability and rich diversity of the open source ecosystem leads to an abundant variety of solutions tailored to specific interests and use cases. Also, an open environment reduces the possibility of vendor lock-in, which can lead to reduced choice and unnecessary expense with maintaining and updating the network.

A vast number of Free education-related applications are available for all versions of Linux. You'll find office applications, graphics, and presentation tools, learning tools, multimedia solutions, and a variety of educational games.

The modern school network, however, isn't just a room full of autonomous PCs. Leading classroom systems offer tools for collaboration, as well as teacher

#### WINDOWS PREFERRED

Learning programs created specifically for Windows, which some teachers unfortunately can't or don't want to do without, present an obstacle for Linux adoption in the education space. The Wine compatibility API [18] provides a solution for running Windows programs on Linux, or you can integrate Windows computers into an existing Linux-based infrastructure using the Samba protocol.

Nevertheless, many schools still favor Windows, even when the Windows version is hopelessly and dangerously outdated. Some schools in the world, for example, are still using Windows XP!



interaction and oversight. Some systems depend on thin client workstations and a terminal server model for supporting classroom activities.

One issue associated with rolling out open source software in schools is the general unfamiliarity with the Linux environment (see the box titled “Windows Preferred”). Each solution must provide the necessary documentation or background support to help teachers, students, and school IT staff use the systems productively.

This article tours some popular Linux distros designed for educational environments (Table 1). You’ll learn about Sugar [1], UberStudent [2], Edubuntu [3], DebianEdu [4], openSUSE Edu-Life [5], and UCS@school [6].

### Sugar

Fedora’s Sugar, which evolved from the One Laptop Per Child project [7] (see the box titled “OLPC”) is designed to make the

world of computers as simple and intuitive as possible for children. The user interface designed for pupils isn’t just aimed at first-graders but at older pupils, too. Sugar also demonstrates how difficult it is to devise a universal solution for pupils worldwide.

The user interface is kept pleasingly simple, and design is clearly aimed at mobile devices – in particular, the XO laptop, which was the target hardware for the OLPC developers. Sugar’s interface, however, is different from all other known GUIs (Figure 1). Anyone starting life in the IT world using Sugar will probably find it difficult to switch to other conventional user interfaces later.

The OLPC initiative wanted to familiarize as many children as possible with computers, especially those in regions of the world with a shortage of computer hardware. By Sugar’s own account, the system specifically targets immigrants who need to learn a new language and users with disabilities.

### OLPC

Sugar Linux isn’t just for the now-no-longer-produced XO-1 laptop from the OLPC initiative. Sugar on a Stick (SoaS, [8]) is a version of Sugar that is based on Fedora 23 and supported by the Fedora project. You’ll need to run SoaS on a laptop with a wireless chip that Sugar supports. There are also Sugar packages in the Ubuntu repositories, but they don’t offer all the applications provided by SoaS.

According to its own statements, the OLPC initiative has already achieved success in Ethiopia and Nepal and is used in Peru and Venezuela, among other places. The once-acclaimed project is no longer in the news, although the existing laptops are still being maintained. The brand name now belongs to the US-based Sakar, which manufactures a tablet with Android instead of laptops, and this tablet is more focused on leisure activities. Walmart sells the tablet with more than 100 games.

However, the creator of the OLPC project, Nicholas Negroponte, is continuing with his goal of global education in the non-profit, education-oriented organization XPRIZE.

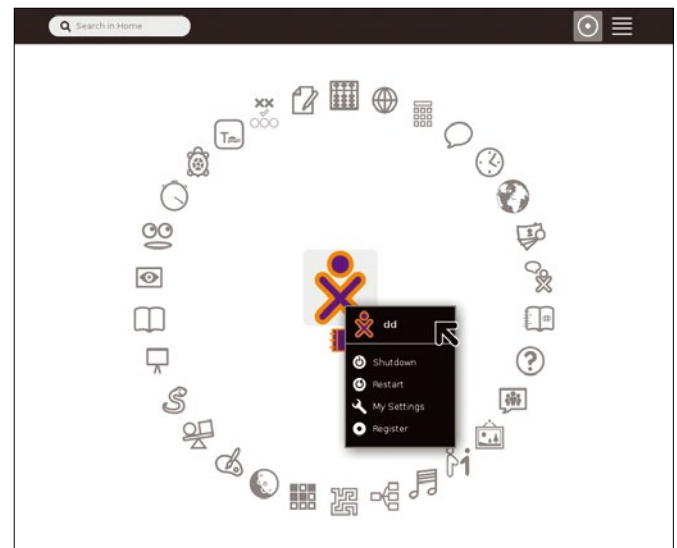
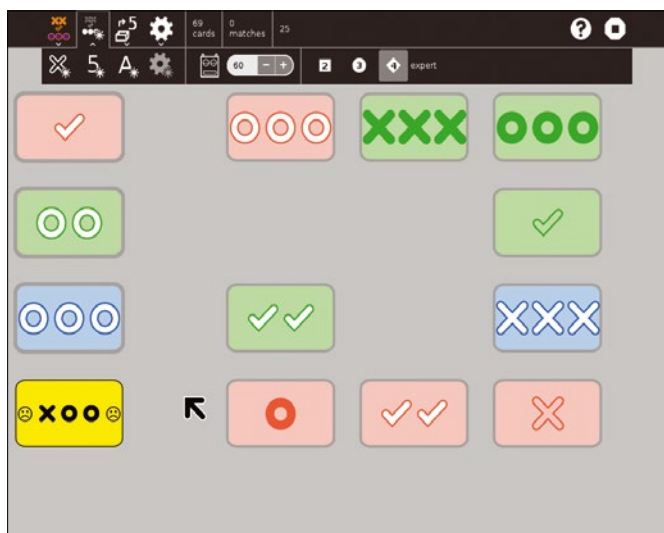


Figure 1: Fedora-based Sugar was originally developed for the OLPC project and has an idiosyncratic interface.

TABLE 1: Schoolroom Linux Features

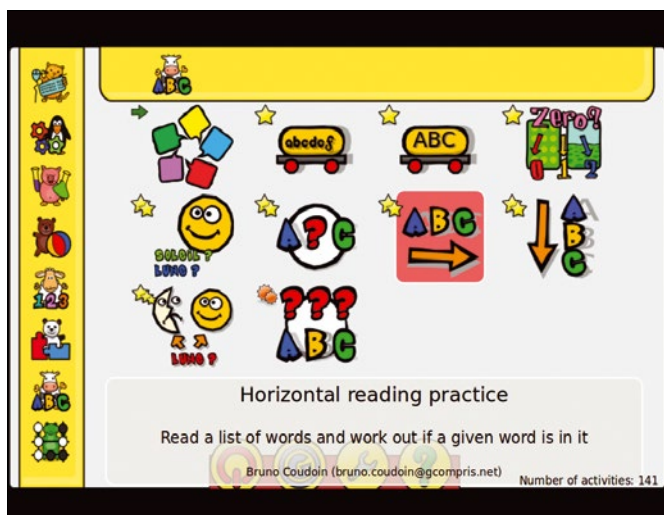
Feature	Sugar	Edubuntu	DebianEdu	openSUSE Edu-Life	UberStudent	UCS@school
Localization	Inadequate	Good	Good	Good	Defective	Good
Documentation	Good (English)	Good	Good	Good	Good (English)	Good
Target group	All ages	All ages	All ages	All ages	Secondary classes/students	Administrators
Application	Classes/schools	Classes	Classes	Classes	Individuals	Schools
Setup effort	Server: difficult; Clients: easy	Server: easy; Clients: complex user management	Server: easy; Clients: complex user management	Server: easy; Clients: complex user management	Server: —; Client: easy	Server: easy; Client: easy
LDAP support	Optional	Optional	Optional	Optional	No	Integrated
Mail server	Optional	Optional	Optional	Optional	No	Integrated
File server	Local or server	LTSP server	LTSP server	LTSP server	Local	Server
Windows/Mac clients	Optional with server	Optional	Optional	Optional	Optional	Yes
Management tools for clients	Moodle	Epoptes	GOsa	iTALC	No	Own solution
Desktops	Own solution	Unity	KDE	KDE	Xfce	KDE



**Figure 2:** The many non-text symbols in Sugar are used to help overcome language barriers, but they don't make it any easier to get started using the system.



**Figure 3:** Although UberStudent seems somewhat pretentious, it does provide a few useful applications that aren't available in other distributions.



**Figure 4:** The software collection GCompris for Kids provides plenty of games and fun.

The user interface provides only icons for the many applications (Figure 2). A number of learning, puzzle, and communication programs hide behind these icons. Sugar connects each user with a Jabber server by default. In any case, the system focuses heavily on the collaboration and communication between the students. Sugar keeps track of all activities locally in a journal, which contains elements such as images and texts.

If you don't have a central server, the Sugar can form a mesh network. Alternatively, you can set up a Sugar server based on CentOS. Teachers also can configure the network, including CentOS-based proxies like Squid (along with filters) or WWWoffle [9], which can deliver content despite shaky Internet connections. Sugar uses the free web-based Moodle learning platform, which groups the students and provides learning materials. Users can store their own activities on the server.

### UberStudent

Whereas Sugar is aimed at children, UberStudent tries to educate older students. Even the name seems a bit pretentious, and the description on the website even more so (Figure 3). The system, we learn, "pushes users to the head of the class." Anyone who scans through the website will receive requests for donations at almost obtrusively short intervals.

UberStudent is based on Ubuntu 14.04 but also provides the update manager from the Linux Mint project. The project initiator, Stephen Ewen, repeatedly leveled criticism at the desktops Unity and Gnome 3 in the past – his Linux distribution consequently uses Xfce with a specially designed interface.

Unlike systems intended for younger students, UberStudent plays down the simple math and reading and focuses on tools for research and writing, study aids, and "self-management" tools.

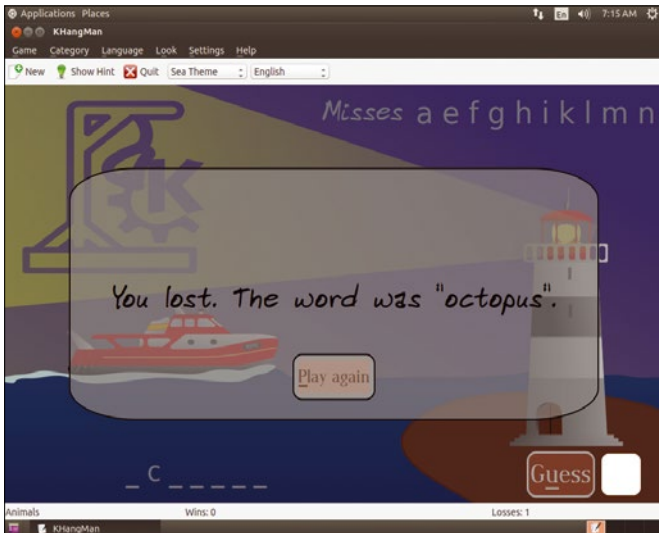
The default configuration comes with the Zotero research tool, the Cherry Tree note-taking application, several logic puzzles, and a number of tools with "encyclopedia" in the name, like "Encyclopedia of Philosophy" and "Encyclopedia of Mathematics." UberStudent offers a diverse collection of helpful applications for high school and college-age users; however, you won't find an accompanying set of management tools and server software. No wonder: UberStudent is not really intended for use in school classrooms but is more for learners' personal computers.

### Edubuntu

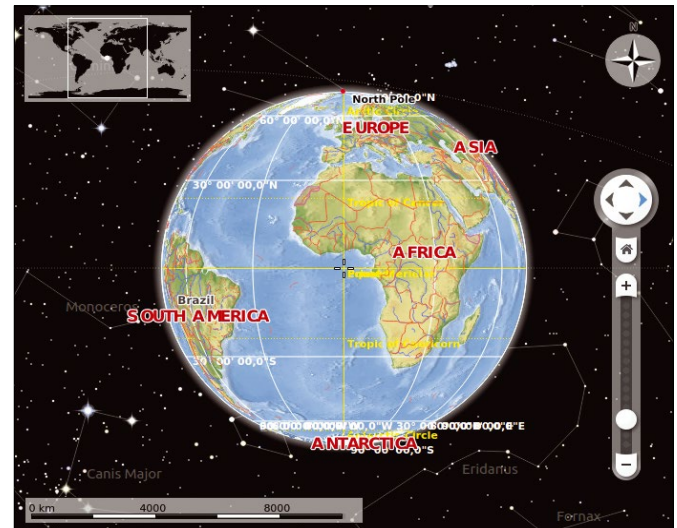
Edubuntu, which is classified as an official Ubuntu project, is significantly better networked and more extensive than UberStudent or Sugar.

Edubuntu includes puzzle and educational games and applications for children in preschool or primary grades in the preschool and primary packages. An essential difference from Sugar: The programs in Edubuntu use conventional menus which display both icons and text.

It is worth highlighting the very successful software collection GCompris [10], which includes a wealth of applications (Figure 4). The GCompris applications are separated into tools for learning the alphabet or for basic arithmetic operations. Games and puzzles like "Connect Four" or chess complement



**Figure 5:** KHangMan: Are the gallows really necessary? Things that are acceptable in one country might cause offense in another.



**Figure 6:** Edubuntu includes many applications for older students, such as a digital atlas.

the collection. You will also find software that describes the basics of power circuits or the water cycle.

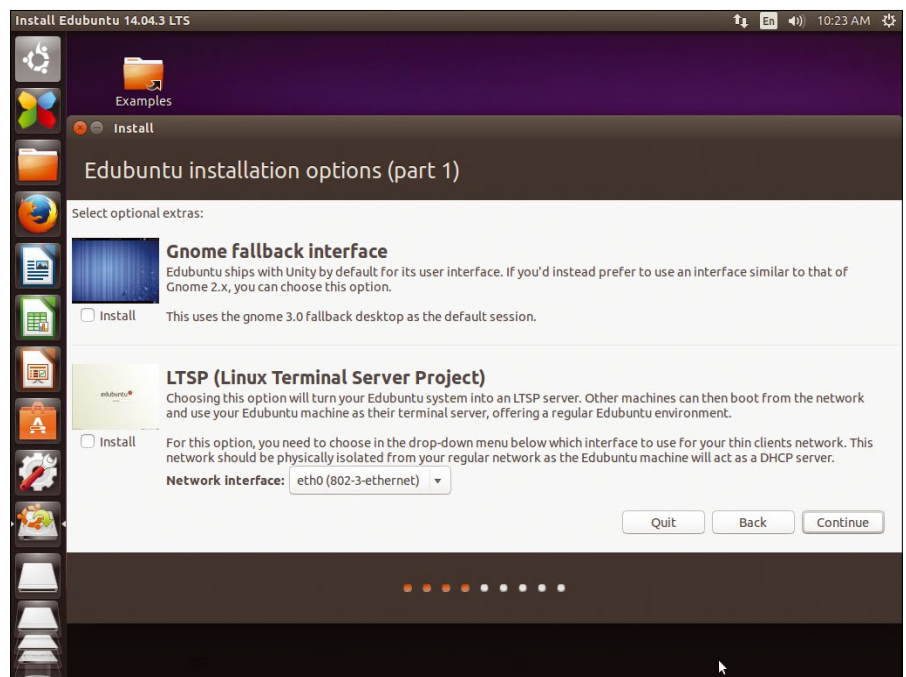
The Edubuntu developers appear to expect a lot from young children: When starting the application to experiment with the power circuit, a message appears, stating that the electric simulator Gnucep is missing, followed by a detailed description of how to install it. It's only in the final sentence that information is provided to say that Gnucep isn't absolutely necessary. The message alone is probably enough to confound primary students, and teachers need administrative rights to download the software.

One software package for pupils in first grade mainly consists of KDE applications, including the guessing game KHangMan (Figure 5). The application also serves as an example of the predominantly Anglo-Saxon orientation of many learning programs, such as Hangman. It is easy to see the problem with internationally standardizing the educational software – applications where children can learn to spell, do simple math, and use computers are really useful for grade-schoolers. However, the more advanced the students are, the more country-specific the requirements become.

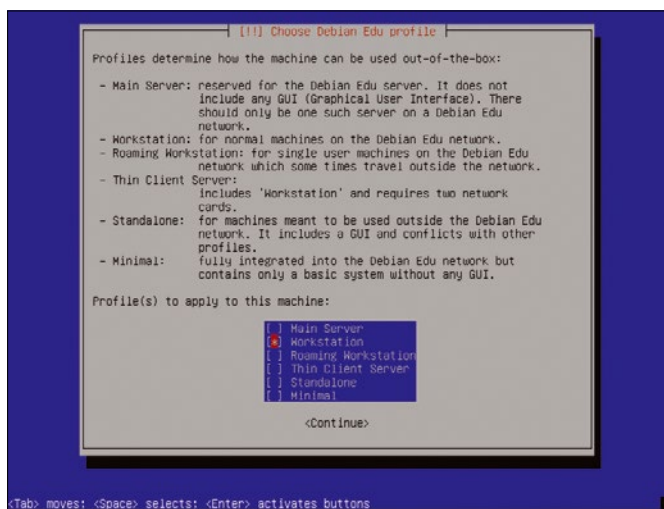
The supplementary packages secondary and tertiary make applications available for higher grades. Academic programs like Chemtool, which illustrates complex chemical compounds, is a good example of this category. Other tools act more as accessories and supplement the research, including the digital Atlas Marble (Figure 6) or Kstars, which invites you to explore the night sky. Linux also provides a massive selection of applications for budding programmers. Despite criticism, Edubuntu's software complements the pupils' textbooks well. Many programs also invite you to experiment.

A major difference that exists between Edubuntu and DebianEdu (which you will learn more about later in this article) is the underlying server infrastructure. Both systems might operate in a thin client environment, but Edubuntu targets an easy installation within a limited infrastructure using the Linux Terminal Server Project (LTSP, [11]), which loops through the graphical interface directly to the thin client (see Figure 7).

The server equipment limits the number of connected clients; an up-to-date computer with four CPU cores and 8GB of RAM can provide about 20 clients. It makes sense to use a server if you are operating Edubuntu in a classroom setting. The option to set up Edubuntu as an LTSP server is available upon installation and works afterwards without any need for further input.



**Figure 7:** Edubuntu tries to make it easy to get started with Linux – a terminal server sets up quickly and works without any further action.



**Figure 8:** DebianEdu used provides extensive learning software and various server solutions you can choose during the installation.

A decisive disadvantage in Edubuntu: Users work directly on the server. For example, during the tests for this article, an update message appeared, which the server obediently passes on to the thin clients. A student who gets hold of the admin password will have direct access to all their fellow pupils' data.

If system administrators want to create their own user accounts, they need to use self-constructed solutions. The default installation only provides the Ubuntu or Gnome user management. Epopotes [12] is available as a central remote maintenance and client-monitoring tool. However, configuration files need to be edited, and you need to start the Epopotes server via the command line. Basic knowledge of Linux is therefore a requirement.

It is also possible to connect diskless workstations to the Edubuntu LTSP server. The workstation takes over almost the entire role of a computer and operates as a Network Block Device (NDB). This setup relieves the server and reduces network traffic between workstations and servers because the workstation directly depends on the Internet. However, a separate DHCP server is required on the network. Finally, the workstation must also meet the system requirements for an Edubuntu installation.

### Ever Expandable

You can incorporate several Edubuntu LTSP servers in a cluster using LDAP and NAT. See the HowTo at the project website [13] – you'll need profound knowledge to make this work. The documentation is poor overall and is limited to a few scenarios. As with any Linux system, you can upgrade the software manually from Ubuntu's repositories and expand the system to taste.

### DebianEdu

DebianEdu (also known as Skolelinux) is a Debian-based educational distro

based founded in Norway. DebianEdu uses the KDE desktop and offers a wealth of learning and production software.

Edu (Figure 8) supports the use of terminal servers, which, in turn, operate their own thin clients. Edu also manages conventional workstations and laptops with Linux on the central server, which take care of both LDAP and DNS and DHCP. The server provides an internal email system that can be configured for external use.

The network also manages Windows and Mac computers. The graphical interface is GOsa (GONICUS system administration, [14]), which also offers user management via LDAP. The GOsa project was originally developed for the Limux project in the city of Munich. GOsa provides an extensive set of tools for managing users, groups, hosts, services, and workstations. You can set up a workstation from the central server via the network. The detailed manual explains how to perform the installation and also provides valuable tips for further configuration.

### openSUSE Edu-Life

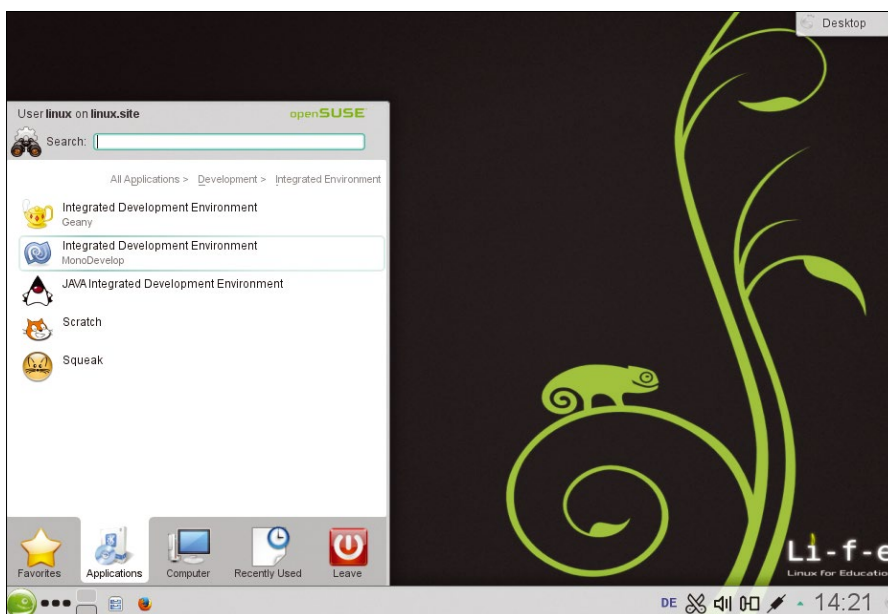
OpenSUSE also provides its own Linux for the education sector, in the form of Edu-Life (Linux for Education) (Figure 9). Like Edubuntu, the openSUSE derivative is available for schools as a live DVD with numerous applications for pupils of all ages.

Like Edubuntu, openSUSE's Edu-Life provides a terminal server based on LTSP. Instead of the remote control software Epopotes used in Edubuntu, the makers of openSUSE Edu-Life use iTALC [15]. In addition, the server also uses the Moodle educational platform [16] and the alternative ATutor educational management system [17].

OpenSUSE Edu-Life focuses on programming and includes numerous development environments by default, which help pupils with how to deal with Java, C, C++, Pascal, and Mono.

### UCS

Univention provides a platform that has been specifically compiled for schools through the server solution UCS@school (Fig-



**Figure 9:** Edu-Life is an openSUSE version specifically compiled for schools that provides extras such as onboard development tools.



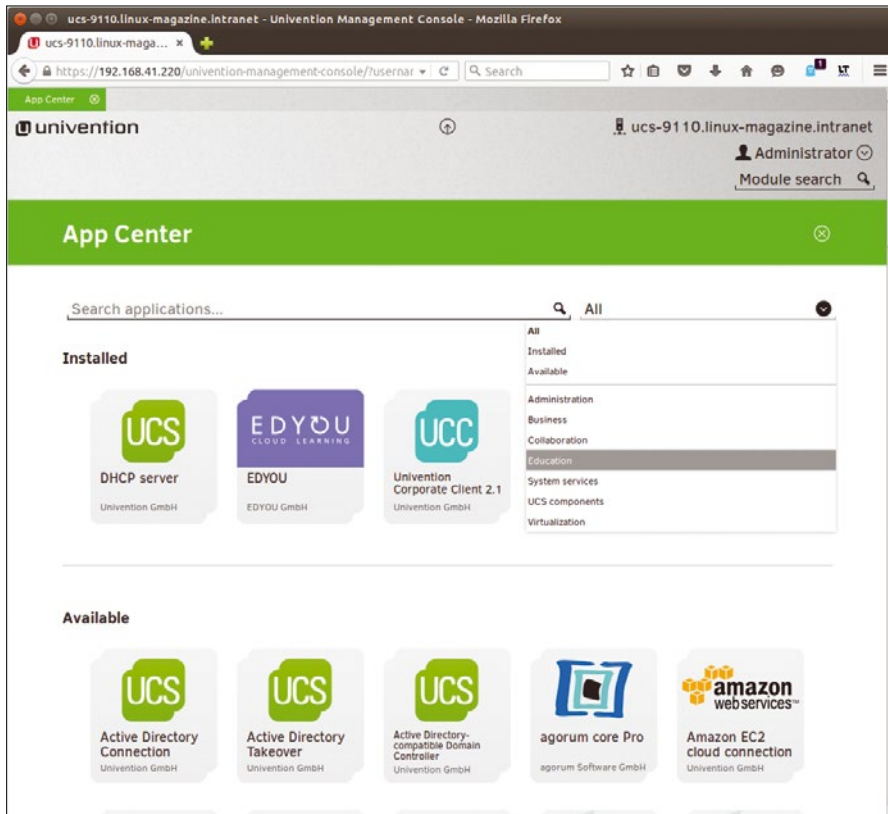


Figure 10: USC@School aims to make it as easy for admins to install and set up servers in schools; the company also provides support.

ure 10), which focuses on simple configuration. You can install and set up the software in just a few minutes. If desired, UCS@school operates as an Active Domain Controller and, in this role, provides roaming profiles for Windows computers. The system can also effortlessly integrate other client operating systems such as Linux or OS X.

The three most important services include a directory service, a central email server for teachers, and finally a data server including synchronization capabilities. At the same time, the IT infrastructure is increasingly outsourced to data centers. According to Univention, cloud services, in combination with their own IT infrastructure, now play a major role. Schools also consider external solutions such as Dropbox, ownCloud, Microsoft Office 365, or Google Apps for Business.

Univention provides a Debian-based Linux distribution called Univention Corporate Client [19] for clients, which administrators can distribute via the network. Important applications like LibreOffice are included by default, and you

# LOST YOUR BOOKSTORE?

## LET US BE YOUR BOOKSTORE

Browse our shop for single issues of *ADMIN*, *Linux Pro*, *Linux Magazine*, *Raspberry Pi Geek*, and *Ubuntu User* – delivered right to your door.

■ [shop.linuxnewmedia.com/single](http://shop.linuxnewmedia.com/single)

Better yet, subscribe, and you won't need a bookstore.

■ [shop.linuxnewmedia.com/subs](http://shop.linuxnewmedia.com/subs)



# shop.linuxnewmedia.com

DIGITAL AND PRINT EDITIONS AVAILABLE!





### SLOW MIGRATION INCREASES ACCEPTANCE

Switching to Linux isn't just a financial matter. IT teacher Marco Schneider recalls that, for his school, the desire to move away from proprietary solutions was based on the practical need to be free of a single vendor and avoid potential lock-in. The Microsoft solution used at his school in the past was too inflexible, and external support was too expensive. However, Schneider explains, migrating to Linux is a long-term project; a sudden switch is very difficult to implement – especially among teachers. Individuals also need to be taken along and shown Linux installations that are already working. It helps that many users are already familiar with products such as Firefox or LibreOffice.

Schneider says he didn't experience any problems with the hardware when switching. He and his colleagues had already been worked with Linux for a long time and managed to set up the IT infrastructure. The team made the decision to go with UCS@school because of time constraints, but also because the company provides support itself for its product. Schneider is gradually migrating all client machines to Linux. The client system also comes from Univention and provides a KDE interface. He recommends trying out all relevant software and experimenting with the operating system before beginning the transition.

can expand the range of software as required. In addition, UCS@school offers more quick-to-install apps.

You can also expand the server itself with monitoring software, an email infrastructure, and other network services. UCS@school requires the Univention Corporate Server as a base. In addition to the free, virtually unlimited, Core Edition, you find variants that cost 290-1690 euros depending on support options. You will also need to pay an additional one Euro per user per year for the UCS@school extension.

### Conclusions

Linux now offers several alternatives for the classroom, with a perfectly knitted solution for almost every scenario (see also "Slow Migration Increases Acceptance"). The educational software available for the Linux environment should occupy schoolchildren over several weeks without boring them with repetition. Even older pupils can draw from an extensive range of applications, which includes excellent academic programs as well as software for teaching programming.

Sugar has a key issue: The focus on unfamiliar symbols means the user interface is hardly intuitive. Teachers, in particular, must spend longer coming to grips with the desktop environment before they can offer assistance to their students, especially with complex applications like drawing. There is, however, detailed help. The localization in UberStudent is lagging

### BEYOND THE COST

Peripheral devices found in schools might cause some problems. Linux's CUPS printer support tool supports most printers, but scanners are sometimes more difficult to integrate. And a teacher can waste valuable teaching time rummaging through the system control because the user interface doesn't appear correctly on the projector.

The savings from using Free software might be well worth the expense of purchasing compatible hardware.

behind. Sugar and UberStudent may be more niche products, but they exemplify the complications associated with using free software.

Systems that lack support can cause a lot of frustration among both pupils and teachers (see the box titled "Beyond the Cost"). Teaching content is the top priority for teachers, who should have to spend as little time as possible explaining functions to pupils or rectifying deficiencies.

Edubuntu provides a good starting point – it might have only limited scope, but it can be used as a live system without excessive effort, which can give both teachers and pupils a first taste of Linux without the need to erect a complete infrastructure. DebianEdu is also a good option for starting with Linux, plus it can expand with a mature school server as necessary.

Anyone who prefers to use a finished and easy-to-maintain IT infrastructure can confidently use UCS@school, which – according to the manufacturers – is already in use in hundreds of schools. Univention doesn't just maintain the software, it also offers comprehensive support and has set up a community devoted to the server solution, with a website providing a forum as well as tips from colleagues. ■■■

### INFO

- [1] Sugar: [http://wiki.sugarlabs.org/go/Getting\\_Started](http://wiki.sugarlabs.org/go/Getting_Started)
- [2] UberStudent: <http://www.uberstudent.org>
- [3] Edubuntu: <https://www.edubuntu.org>
- [4] DebianEdu: <https://wiki.debian.org/DebianEdu/>
- [5] openSUSE Edu-Life: <https://en.opensuse.org/Education/Live>
- [6] UCS@school: <https://www.univention.com/products/univention-app-center/app-catalog/ucsschool/>
- [7] OLPC Project: <http://one.laptop.org>
- [8] Sugar on a Stick: [https://wiki.sugarlabs.org/go/Sugar\\_on\\_a\\_Stick](https://wiki.sugarlabs.org/go/Sugar_on_a_Stick)
- [9] WWWoffle: <http://www.gedanken.org.uk/software/wwwoffle/>
- [10] GCompris: <http://gcompris.net>
- [11] LTSP: <http://www.ltsp.org>
- [12] Epoptes: <http://www.epoptes.org>
- [13] LTSP-Cluster: [https://help.ubuntu.com/community/UbuntuLTSP/LTSP-Cluster\\_NAT\\_OpenLDAP](https://help.ubuntu.com/community/UbuntuLTSP/LTSP-Cluster_NAT_OpenLDAP)
- [14] GOsa: <https://oss.gonicus.de>
- [15] iTALC: <http://italc.sourceforge.net>
- [16] Moodle: <https://moodle.org/>
- [17] ATutor: <http://www.atutor.ca>
- [18] Univention: <https://www.univention.com/>
- [19] Wine: <https://www.winehq.org>
- [20] Univention Corporate Client: <https://www.univention.com/products/ucc/>

### AUTHOR

**Jörg Thoma** has been using Linux for 20 years and has been writing about it for almost as long. He loves everything about open source, still thinks the command line is useful, and relaxes while watching compiler messages.

# MOBILE USERS

search for us today at your digital newsstand!

Only a swipe away!

Download our convenient digital editions for your iPad, iPhone, or Android device.



## GOOGLE PLAY MAGAZINES

ADMIN Magazine

Linux Pro Magazine

Ubuntu User

Raspberry Pi Geek



## APPLE NEWSSTAND

ADMIN Magazine

Linux Pro Magazine

Ubuntu User

Raspberry Pi Geek



Visit our apps page for more information: [www.medialinx-shop.com/apps](http://www.medialinx-shop.com/apps)

## Desktop backup software

## Reinforcements

Backup strategies in IT are essential and expensive in terms of planning and administration, but individuals have simpler solutions. We look at five backup solutions for the desktop.

By Kristian Kießling

Experts regularly warn about the potential for data loss, but many users, including some IT professionals, have never had a serious hard drive crash, partly because they often discard their computers before their hard drives give up the ghost. Never having experienced a catastrophic loss of files, many users simply store data sporadically on an external disk; however, this isn't a real backup strategy.

More sophisticated users might employ Rsync [1] and Cron [2] scripts with their corresponding overhead, but graphical backup software, which often utilizes these command-line tools under the hood, is often a more satisfactory backup solution on the desktop. Linux users have a number of options suitable for automatic and continuous operation that allow them to save and restore files easily. These applications might even be worth using at work so that employees without root privileges can back up their data on an external hard disk or USB flash drive, allowing them to work at home or on the road.

In this review, I look at Déjà Dup [3], Back in Time [4], Sbackup [5], lucky-Backup [6], and Areca Backup [7].

### Considerations

The tools I discuss here aren't usually suitable for full system backups, including partitions. In such cases, you would turn to Clonezilla [8], ISO Master [9], or Partimage [10]. The desktop applications, on the other hand, store folders and files, either on the same computer (not recommended), on external drives, in the local network, on remote servers, or even with cloud providers, who also encrypt the data.

Another important thing to know when making backups is whether the software creates incremental, differential, or delta backups. Although incremental backups only temporarily store the modified files from one full backup to the next, saving disk space, the system has to put all the puzzle pieces back together when restoring them, which can take a while. Differential backups, on the other hand, collect the modified files, which takes up more space; how-

ever, the backups can be restored more quickly.

If you have to back up large files that grow or shrink, you will only want to transmit tiny changes and not the whole load each time (e.g., VM images). Delta or block-level backups, which split files into small pieces and secure them with rolling checksums, accommodate this situation. One drawback is that such backups can't be recovered manually in an emergency.

Those just wanting to back up two or three folders containing text, spreadsheets, and graphics don't need to worry about space problems. Consecutive full backups without compression will be fine in these cases. However, because even the home environment can involve significant volumes of data (e.g., video or RAW photography files), the software will have to deal with space issues. Does it even notice when a disk is full, and does it report that fact?

Also, how does it determine when to delete old backups? Your unique situation and requirements will determine your backup strategy, which will decide

what software you choose to implement this strategy.

In the tests for this review, I used an up-to-date Ubuntu Gnome 15.10. The packages available for this distro are typically available from the Debian repositories.

### Déjà Dup

Déjà Dup [3] is the obvious solution for simple, incremental backups. The desktop developers have integrated the software seamlessly in Gnome and Unity, and its simplicity is impressive. The Overview window after startup displays the *Restore* and *Back Up Now* options (Figure 1). Déjà Dup uses `deja-dup-monitor`, rather than Cron, to activate scheduled backups with the switch at the top right.

Being able to fine-tune the timing of backups has been sacrificed in favor of

usability. If you enable scheduled backups, you can back up either every day or every week – these are the only options. However, even that is more methodical than manually backing up data from time to time.

You can determine whether backups stay in the archive for *At least six months*, *At least a year*, or *Forever*. The last option is the default; however, Déjà Dup does delete old backups if the target location is running out of space.

Déjà Dup accepts various storage locations: whether SSH, Samba, NFS, FTP, or WebDAV. The data can even be moved to the Amazon (S3) and Rackspace cloud and can be encrypted (Figure 2) using GPG and symmetric ciphers.

Users can specify which files they want to back up in the *Folders to save* and *Folders to ignore* tabs. Inclusions and exclusions cannot be defined using regu-

lar expressions. The Help dialog, which can be accessed under *Backups*, indicates that Déjà Dup by default doesn't back up some folders, such as caches, thumbnails, and Flash content.

For Déjà Dup to work, you need to be logged in to your user session. Although it's possible to operate Déjà Dup from the command line, you would be better off turning to Duplicity [11], which Déjà Dup runs under the hood and which provides more options [12].

**Assessment.** Other than the well-thought-out design, Déjà Dup not only supports backups encrypted with GPG, it also pushes them to the cloud on request. Another important feature provided by the *Restore* function is that backed up files can be rolled back to an earlier version by right-clicking in the Nautilus file manager. In doing so, users receive a kind of version control, allowing them to recover files in a user-friendly manner.

One clear disadvantage is a lack of profiles: It isn't possible for several users to use one desktop or for one user to back up folders to different locations. The non-granular scheduling might not satisfy some users, either. Some users have complained about encrypted backups that can no longer be decrypted, but a manual emergency solution appears to be available from the command line for this situation [13].

### Back in Time

If you search for this name on the Internet, will find hits for the popular "Back to the Future" trilogy from the 1980s. Knowing how things went wrong in that movie is hopefully not an omen for the backup solution named Back in Time [4] (Figure 3), which you can find in the Ubuntu repositories in two versions: one for Gnome and one for KDE.

Unlike Déjà Dup, Back in Time uses Cron, runs jobs automatically with Anacron, and offers profiles. After starting, users automatically end up in the default *Main profile* in the Settings dialog, where they can access details about the backup process in the tabs. To add other profiles, click on the *Settings* icon in the toolbar.

Back in Time offers traditional fare in terms of backup targets: You can store files locally or via SSH, with an option to encrypt the payload in both cases. If you

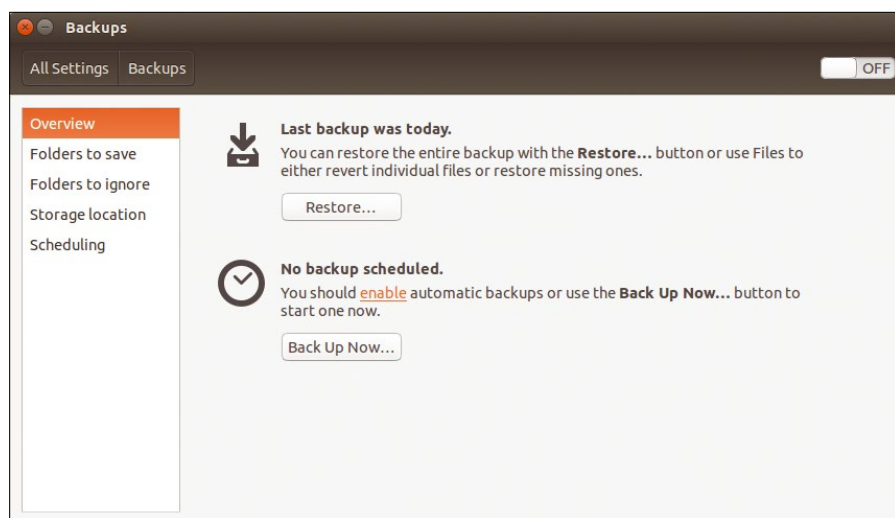


Figure 1: Impressively simple: The Déjà Dup startup window.

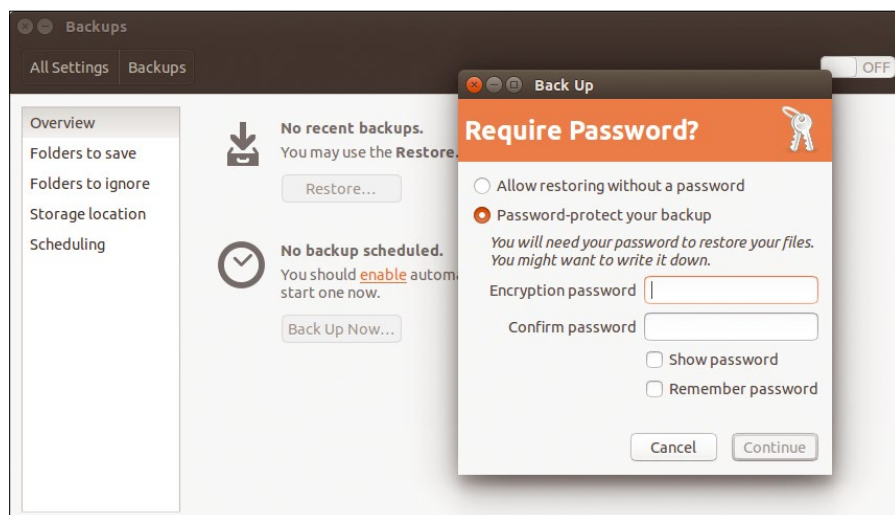


Figure 2: If needed, you can encrypt your backups. If you do, make sure to remember your password; otherwise, the recovery operation is very tricky.

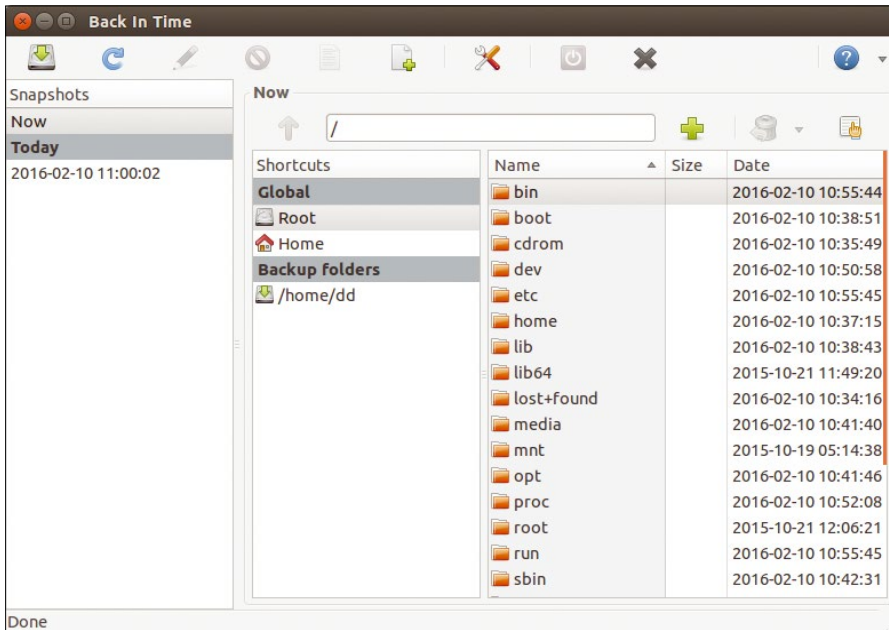


Figure 3: Back in Time shows the snapshots on the left and content on the right.

want to use Samba, you will have to use a workaround: Samba shares can be mounted on the computer with the backup software with `cifs-utils` and `/etc/fstab`.

The handling of automatic backup is significantly more flexible (Figure 4), ranging from *Every 5 minutes* to *Every Month*. Notebook owners are likely to be interested by the *Daily (anacron)* option, and the developers have even thought about external hard drive users with the *When the drive get connected (udev)* [sic] backup option. It is also possible to

back up *At every boot/reboot* or to stop the service with *Disabled*. The inclusion and exclusion of files works as before, except that you can use regular expressions for excluding folders and files.

Space management supports granular options. Back in Time deletes backups that are more than 10 years old by default, or if the remaining storage space drops below 1GB or the free inodes fall to less than two percent. Anyone who doesn't want to just get rid of old backups can click the *Smart remove* checkbox and choose how the software should

thin out its snapshots.

Even more choices are avail-

able in the *Options* and *Expert Options* tabs. For example, one option checks for changes and doesn't generate any new backups if none are found, and you can check for changes using checksums. Back in Time logs *All* by default, but the software still carries on with its work if it finds errors in the snapshots. Notifications are enabled and experts can also set the priority for a Cron job, throttle the Rsync bandwidth if desired, or preserve ACLs and extended attributes.

**Assessment.** Rsync, *cp*, and *diff* carry out the work in the background. Because Back in Time uses Rsync, local and remote filesystems need to support hard links for unchanged files. Initially, Back in Time excludes files with a prefixed period (dot files) from the backup. Compression is also missing, which is a real pain for large data sets.

On the other hand, Back in Time is much more flexible than Déjà Dup in some respects, especially concerning backup intervals and snapshot management. The restore function works well: You can just select a file, right-click, and then choose *Snapshots*. The previous revisions to the file now appear with the capture date and time. Right-clicking again on the selected candidates lets you *Diff* the contents of the file with another or restore it.

### Simple Backup

One of the weaker candidates in the test was Sbackup [5] (Figure 5), short for Simple Backup. The latest version, 0.11.6, was released some time ago and can be installed directly from Ubuntu re-

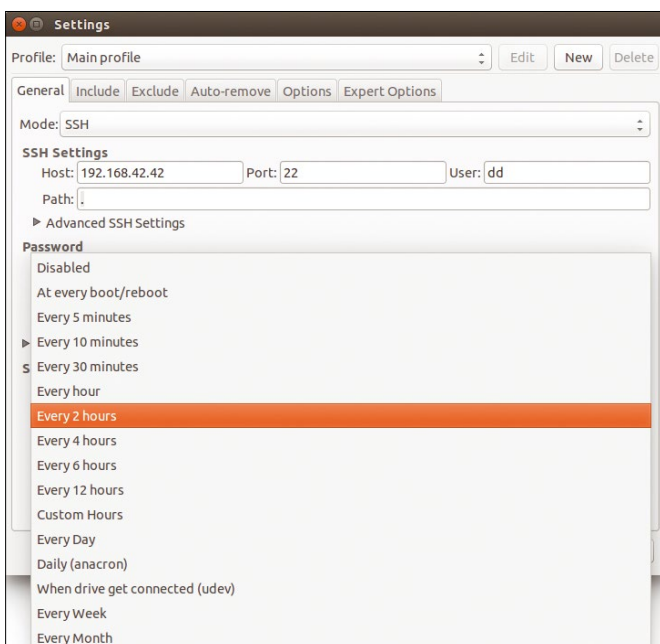


Figure 4: Back in Time has numerous backup intervals. It also responds if a user plugs in a USB flash drive.

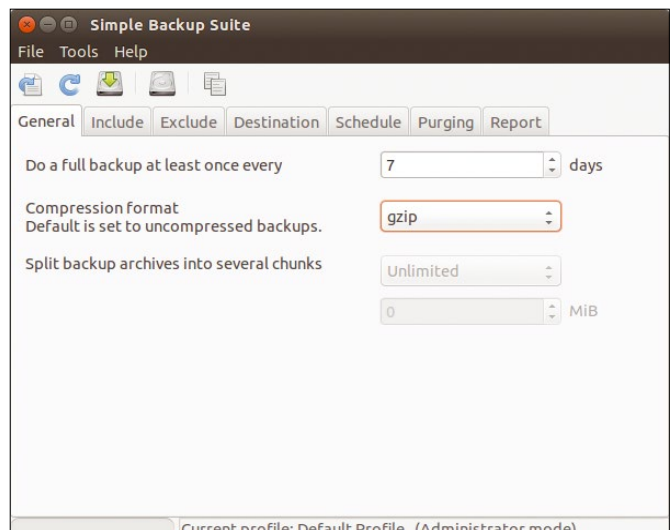
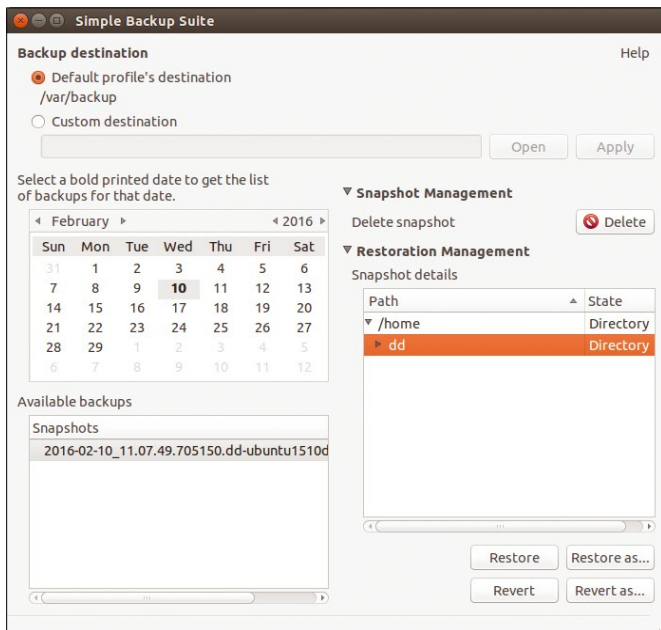


Figure 5: The options in Sbackup are limited.



**Figure 6:** This window could quickly become confusing with several backups.

positories. The Ubuntu wiki opens with a prominent caution about the software's reliability [14]. The lineup is also a bit confusing, with a backup daemon and two separate graphical interfaces: one for restoring and another for configuring. Furthermore, Sbackup distinguishes between normal and root users. The software performs automatic backups with Cron, so it needs to call up the `sbackup-config-gtk` configuration interface with root privileges.

Root and normal users generally have different default values – the differences are explained in the help section (F1). This includes the fact that compression is set to `gzip` and the *Include* tab integrates the `/var/`, `/home/`, `/usr/local/` and `/etc/` paths. Sbackup also excludes paths from the backup using various criteria, including regular expressions.

The documentation for Sbackup isn't particularly good. The Help dialog, for example, only shows a single image regarding the topic of restoration. Although Sbackup does not encrypt data, it does provide profiles.

After starting, you can determine how often you want Sbackup to perform a full backup, and in the meantime it executes incremental backups. Local directories (which include Samba and NFS mounts) and remote stores accessible via SSH are possible targets. To use SSH, you should at least register via SSH using the console on the remote host.

A status indicator in the notification area provides information about the progress of the backup. The recovery window is quite simple and can quickly become cluttered (Figure 6) if you have many backups. You can restore the entire backup or individual files, and this also works recursively. However, a diff function is missing.

### Assessment.

The warnings in various wikis are reason enough to stay away from Sbackup. The project doesn't seem to be very active anymore and has many weaknesses. Besides the confusing user guidance and multiple interfaces, complaints have been made by users about the software not reporting when a disk is full [15]. An option to encrypt data is also missing.

One positive Sbackup feature is support for logs and reports sent via email – Back in Time and Déjà Dup don't offer

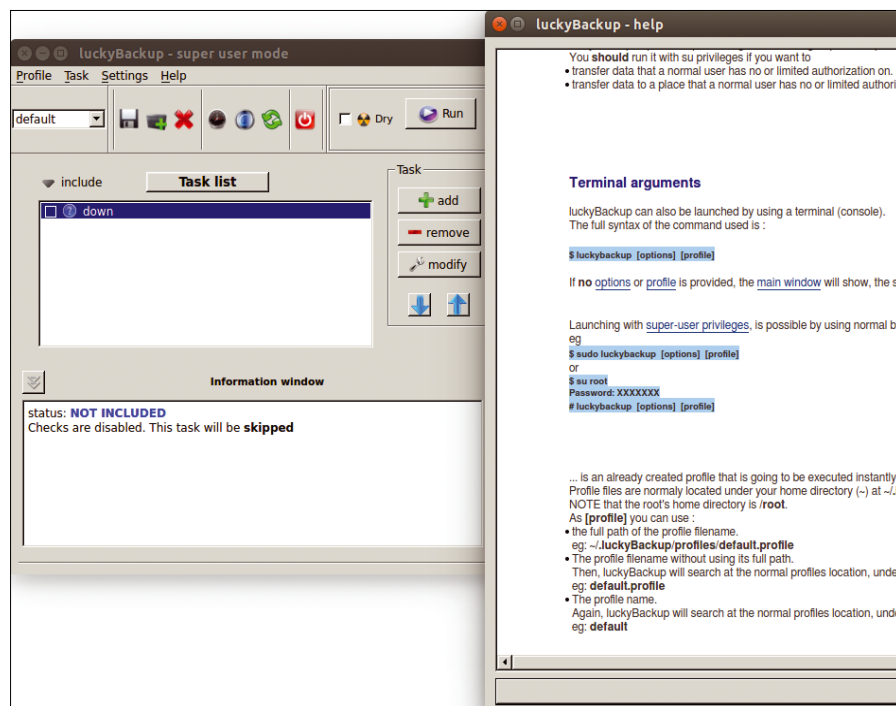
this. The option in the *General* tab to split an archive into several chunks is particularly helpful for those who store their data on hard drives with a FAT file-system that can only cope with files of a limited size. Nevertheless, considering the alternatives, you probably have little reason to use Sbackup.

## luckyBackup

The blaze of color that hits you when you open luckyBackup [6] (Figure 7) is quite something: Fonts in pink, green, and blue are a real feast for the eyes. The My First Website charm extends to the manual, which users can access by pressing F1. The developer obviously tried out all the HTML styles available – and didn't forget the smileys either.

But, enough about style, because presentation often says little about function. Indeed, luckyBackup provides some original ideas, along with some serious flaws, putting it in the same league as Sbackup rather than with the other two candidates, probably because no one really maintains the software anymore. The latest release, 0.4.8, is dated March 2014, and the developers are not providing any new updates or bug fixes.

At least the latest version is in Ubuntu's repositories. luckyBackup uses Rsync and a Qt interface with its own language settings. It doesn't just add profiles, it also imports and exports



**Figure 7:** The luckyBackup creators seem to love all things colorful.

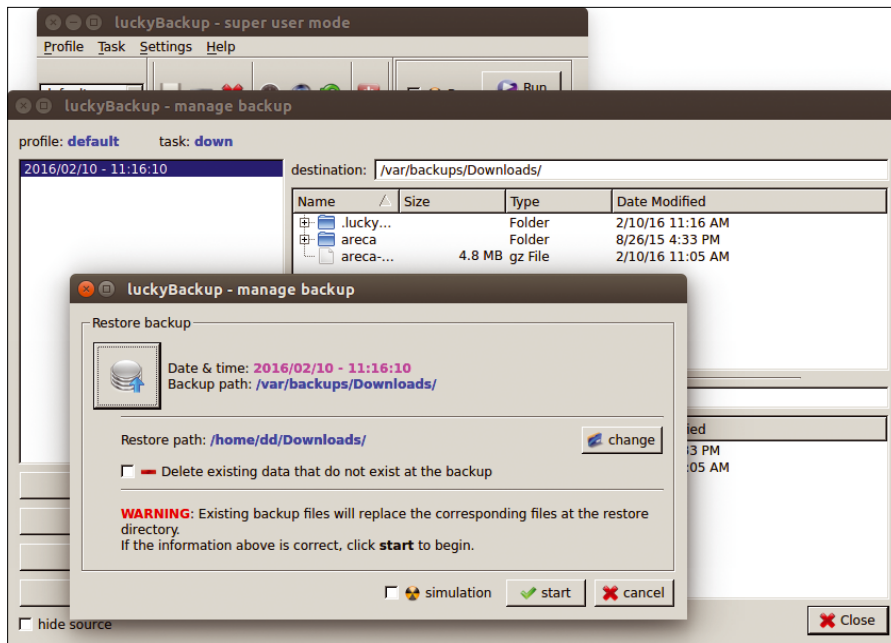


Figure 8: You can also choose to simulate a restore operation.

them. This means you can equip several computers quickly with a luckyBackup profile.

This program also has two modes: one for normal users and the second for superusers. The *Task List* in the middle of the GUI is where you line up different backup jobs.

If a job doesn't find its target directory, a corresponding message appears in the *Information window*. The software can also keep two computers in sync, as with Unison [16]. Although this function is similar to a backup, it mostly just caused confusion in this test.

A great option that isn't available in the programs I've already mentioned is the ability to give backups meaningful names. The pure date format on which the other programs rely doesn't always make searching easy when restoring, because all the file names are similar.

You can supplement a job with the *add* button in the Task pane to initiate a backup operation. In the Task properties dialog, *Backup Source inside Destination* is the backup function, whereas *Synchronize Source and Destination* keeps two directories in sync.

Clicking *Advanced* provides additional options that let you describe your backup, exempt files from the backup (with regex, too), supplement specific files, or make contact with a remote computer over SSH. You can sort out Rsync options manually in the *Com-*

*mand Options* tab, and scripts can be defined in the *Also Execute* tab, which can be run *Before* or *After* a backup. Finally, you can specify the maximum number of snapshots to be kept in the bottom right corner. If luckyBackup exceeds this number, it deletes older snapshots.

If you have defined one or more tasks, you can trigger the backup in the main window using *Run*. If you check *Dry*, luckyBackup runs a simulation and reports errors if found. Caution: The software executes the pre-scripts and post-scripts! Clicking the *Schedule* icon lets you create fixed backup intervals, which uses Cron. Once the software has performed its duties, it will send you an email with the report, if requested.

Finally, you can restore data by going to *Task | Manage Backup* (Figure 8). To do so, select a snapshot on the left and click *Restore*. You can create a diff by choosing *Calculate differences*.

**Assessment.** Most noticeable is that luckyBackup only provides very rudimentary ways to manage the space on the hard disk and delete old backups. You might need to define it as a task or write a task that runs before or after the backup, which isn't particularly convenient. Moreover, the software doesn't compress backups, doesn't encrypt, and doesn't really look after your storage space, either. On the other hand, it can be scripted to cut corners and feed parameters to Rsync. Although this is sure

to win some fans, the whole point of a graphical application becomes a bit lost.

## Areca Backup

Areca Backup is pretty much the opposite of luckyBackup in terms of presentation [7]. As a Java program, it needs `openjdk-7-jre`; works across platforms; provides a clearly structured, discreet, and almost boring interface; can be expanded with plugins; and otherwise leaves little to be desired. Despite being licensed under GPLv2, it isn't in Ubuntu's standard repositories. In Linux, you simply need to download and unpack the software and execute the associated script, `areca.sh`.

The list of features is impressive. First, Areca Backup has all forms of backups: full, incremental, and differential making it the tool of choice for large dynamic files. Choice of storage mode includes *Standard*, *Delta*, and *Image*, which produces a single image file and renews it as soon as the files of the source change. Successive archives can be combined to save space.

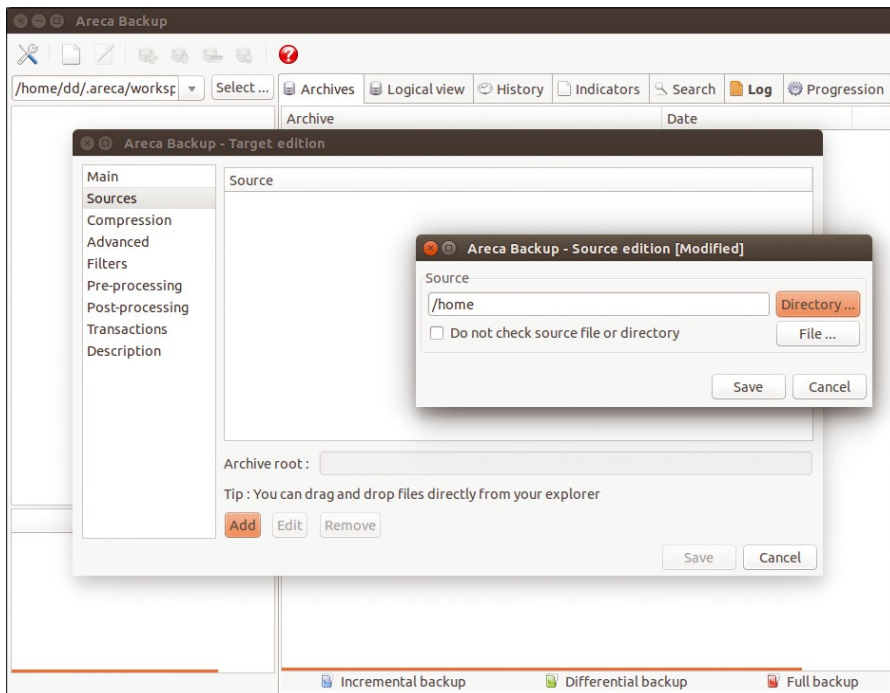
Areca Backup can also encrypt (AES128 or AES256) and compress data (Zip or Zip 64). Here, you also have the option to select character encoding and a compression factor. Network drives, USB keys, FTP(S) servers, and SFTP locations all serve as possible storage targets.

Users first need to create a new group (*Edit | New group*); then, right-click the group and choose *New Target*. A window will open (Figure 9) where you can enter the details for the backup operation, such as the target and the source folders. You can also choose to set up an accessible FTP or SFTP target in the *Main* tab, and *Sources* can be supplemented by dragging and dropping files from your file manager.

In the *Advanced* section, you can set up encryption and bandwidth throttling. The *Transactions* section is also interesting: If a backup aborts, Areca Backup continues at one of the transaction points (in KB). Transactions take place as individual operations to ensure the integrity of backups.

If the source and target are adequately defined, clicking *Save* stores them in the left window below the group name. If you right-click on the target, you'll have access to extra options, such as *Simulate backup* and *Wizards | Generate backup*





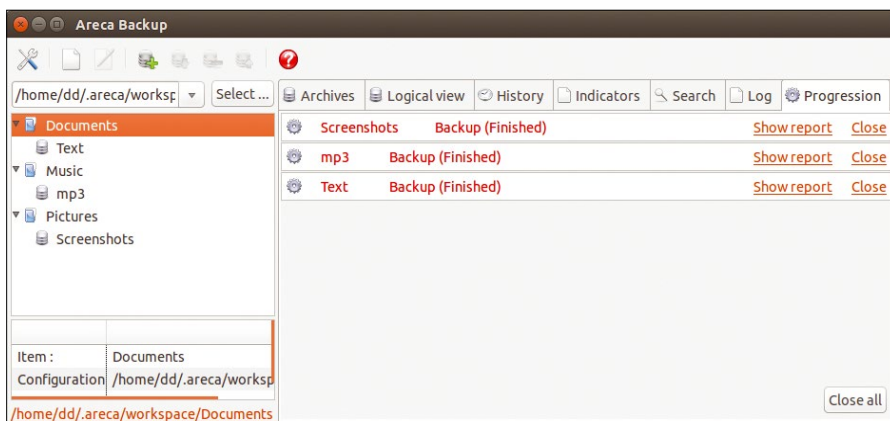
**Figure 9:** The details for the backup job can be defined via the Target window in the background.

shortcut, which means the backup process can be easily converted into a Cron job later. Clicking *Backup* then asks whether you want an incremental, differential, or full backup before Areca Backup gets on with its business. A progress bar provides feedback about the backup, and Areca lists error and success messages in the *Progression* tab (Figure 10).

The restore is handled in the *Logical view* tab. The backed up files appear here and can be both restored, viewed, or even played back (e.g. music files). The *Search* tab helps you find a particular file: Those who back up lots of files know what it's like to search for a needle

in a haystack during the restore process, so they will appreciate this feature. Last but not least, Areca generates reports from its backups and sends them to the user on request.

**Assessment.** Areca Backup leaves little to be desired. The rather unusual form of automation is probably because the program is written in Java, but the scripts that can be generated solve this problem in an elegant way. Very few plugins are available, although WebDAV or cloud support would be desirable features that could possibly be implemented. Areca Backup does not back up entire systems, but that isn't the aim of the software, either.



**Figure 10:** Profiles are known as groups in Areca Backup, and various backup jobs can be created in each group. The left side lists groups and targets, and you can follow the progress of the backup on the right.

## Conclusions

Table 1 shows the software versions used in this test, as well as the latest version numbers, along with the features provided by or missing from each backup solution.

If you're looking for a simple but efficient backup solution for desktops, you would be well served with Déjà Dup. It integrates well with Gnome and Unity, has a simple but functional interface, and encrypts the backup data. Of all the software tested, only Déjà Dup could push encrypted backups to the cloud. On the other hand, you can only roughly define automated backup intervals, and it was the only software reviewed that did not let you define profiles, which makes complex backup scenarios difficult.

Back in Time has a lot of settings for managing snapshots and defining intervals. For example, it reacts as soon as a user inserts a USB stick into the computer. Hard links make it possible to use lean incremental backups but also require specific filesystems on the target computer. A data compression option is missing, as well.

Sbackup has a rather awkward design and is therefore pretty inconvenient to operate, and it is one of the weakest programs tested. Development is apparently languishing. Data encryption is not possible, and although data restoration might work, it certainly isn't a spectacular solution. On the other hand, the program does provide email notification, but this is hardly likely to be a decisive criterion.

LuckyBackup is, at least in terms of presentation, the flamboyant personality among the backup programs. The developers couldn't decide whether they wanted to back up or synchronize files, so luckyBackup does both. You can fine-tune pretty much all the options, but this means the graphical program loses its simplicity; users can also switch to the command line.

The software is missing an option to delete existing backups intelligently, although the option to test the backup process in a dry run is useful. Snapshots can also be given titles or be tagged, which helps when restoring later.

Java is an island, as you can see from Areca, yet this program was the favorite test subject because it is designed clearly

and provides a lot of very useful features. Backup automation has been solved in a slightly unusual way but is practical.

The recovery function is also particularly successful: It isn't just possible to tag the snapshots, you can also search through them and search for files directly from the backup. Cloud support implemented through plugins would be appreciated.

Among the candidates, Déjà Dup and Back in Time solved the recovery process best, but they still have room for improvement when many backups are in the mix. Additionally, this review showed that users should test their backup processes in detail after setting them up because of all the variables involved (Samba shares, SSH passwords, disk space). ■■■

### INFO

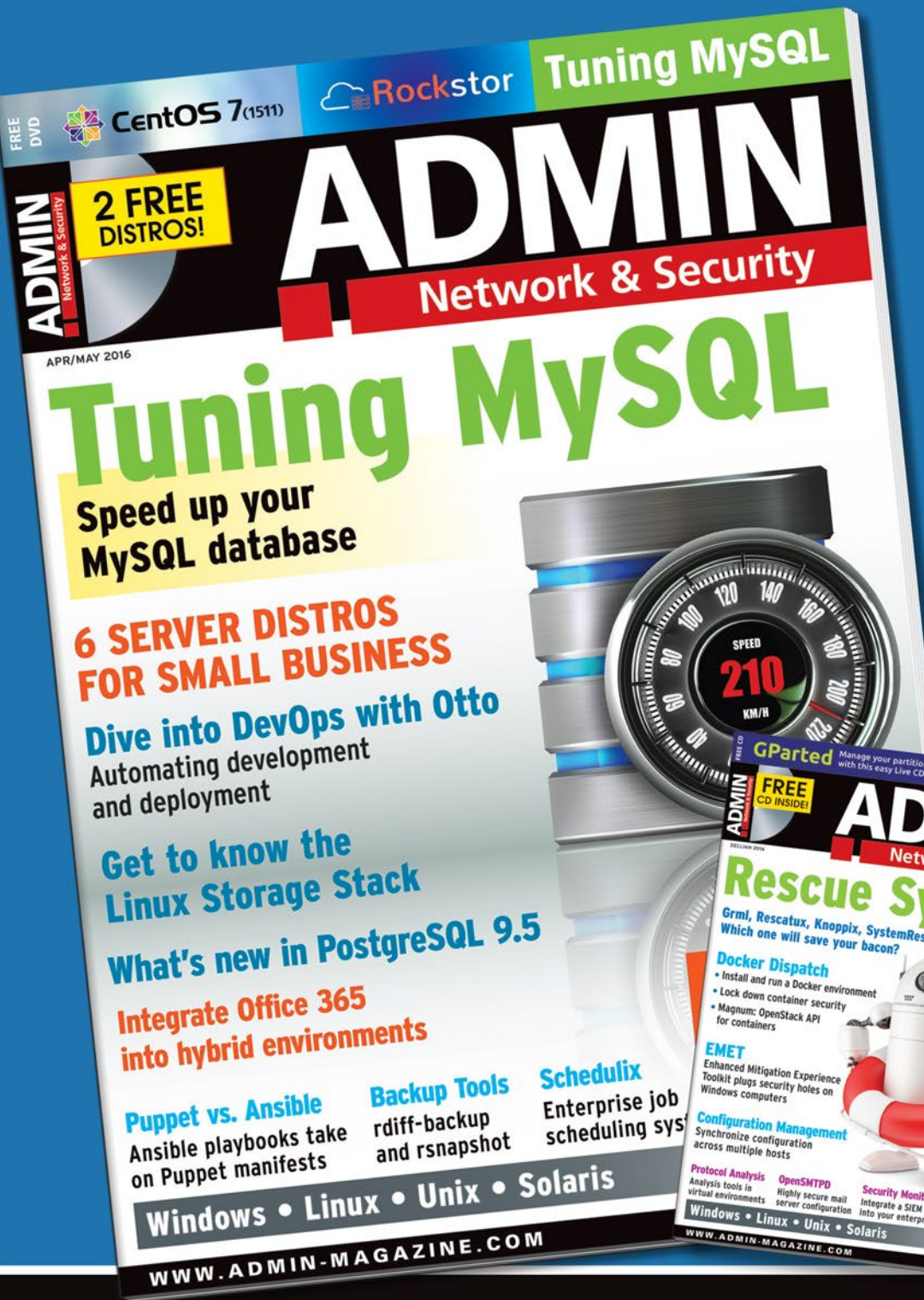
- [1] Rsync: <https://help.ubuntu.com/community/rsync>
- [2] Cron: <https://en.wikipedia.org/wiki/Cron>
- [3] Déjà Dup: <http://live.gnome.org/DejaDup>
- [4] Back in Time: <http://backintime.le-web.org>
- [5] Sbackup: <http://sf.net/projects/sbackup/>
- [6] luckyBackup: <http://luckybackup.sourceforge.net>
- [7] Areca Backup: <http://www.areca-backup.org>
- [8] Clonezilla: <http://clonezilla.org>
- [9] ISO Master: <http://www.littlesvr.ca/isomaster/>
- [10] Partimage: <https://www.partimage.org>
- [11] Duplicity: <http://duplicity.nongnu.org>
- [12] "Duplicity Cloud Backup" by Kristian Kißling, *Linux Pro Magazine*, issue 156, November 2013, pg. 20, <http://www.linuxpromagazine.com/Issues/2013/156/Duplicity-Cloud-Backup>
- [13] "Restore a broken Déjà Dup backup by hand" by David Huss (ato): <http://blog.atoav.com/2013/09/restore-broken-deja-dup-backup-hand/>
- [14] Ubuntu wiki for Sbackup: <https://help.ubuntu.com/community/BackupYourSystem/SimpleBackupSuite>
- [15] Space problems: <http://ubuntuforums.org/showthread.php?t=1122489>
- [16] Unison: <https://www.cis.upenn.edu/~bcperce/unison/>

**TABLE 1: Backup Solutions**

	Déjà Dup	Back in Time	Sbackup	luckyBackup	Areca Backup
Download	<a href="https://launchpad.net/deja-dup">https://launchpad.net/deja-dup</a>	<a href="https://github.com/bit-team/backintime">https://github.com/bit-team/backintime</a>	<a href="https://launchpad.net/sbackup">https://launchpad.net/sbackup</a>	<a href="http://luckybackup.sourceforge.net">http://luckybackup.sourceforge.net</a>	<a href="http://www.areca-backup.org">http://www.areca-backup.org</a>
License	GPL	GPL	GPL	GPLv3	GPL
Operating system	Linux	Linux	Sbackup	Linux, Windows Beta	Linux, Windows
Current version	34.1	1.1.2	0.11.6	0.4.8	7.5
Tested version	34.0	1.0.36	0.11.6	0.4.8	7.5
Automation	Own service	Cron	Cron	Cron	Cron
Interval	Daily/weekly/immediately	Fine-granular	Fine-granular	Fine-granular	Fine-granular
In the package sources	Preinstalled	Yes	Yes	Yes	No
Include/exclude option	Yes/yes	Yes/yes	Yes/yes	Yes/yes	Yes/yes
Profiles	No	Yes	Yes	Yes	Yes (groups)
System backup	No	No	No	No	No
Compression	Yes	No	Yes	No	Yes
Email notification	No	No	Yes	Yes	Yes
Encrypted backups	Yes	Yes	No	No	Yes
Tag backups	No	No	No	Yes	Yes
Restore with search	No	No	No	No	Yes
Deletion behavior	Few options	Very fine-granular	Simple and logarithmic	None	Not specified
Delta backups	No	No	No	No	Yes
Incremental backups	Yes	Yes	Yes	Yes	Yes
Differential backups	No	No	No	No	Yes
NFS	Yes	Yes, if locally mounted	Yes	Yes, if locally mounted	Yes
Samba	Yes	Yes, if locally mounted	Yes	Yes, if locally mounted	Yes
SSH/SFTP	Yes	Yes	Yes	Yes	Yes
WebDAV	Yes	No	No	No	No
FTP	Yes	No	Yes	No	Yes
Amazon S3	Yes	No	No	No	No
Rackspace cloud files	Yes	No	No	No	No

# ADMIN SOLUTIONS

DIGITAL EDITION NOW AVAILABLE AT  
[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)



- Quick & easy access
- Read PDFs offline
- Keyword search



**GO DIGITAL!**  
Subscribe NOW to our Digital Edition



Using Corel AfterShot Pro

# RAW MUSCLE

Process and organize photos and RAW files with Corel AfterShot Pro.

By Dmitri Popov

**W**hen it comes to picking a tool for processing and organizing photos, Linux users are spoiled for choice: digiKam, Darktable, RawTherapee, UFRaw – you have plenty of excellent RAW processing and photo management applications from which to choose. Thus, it might seem that opting for closed source commercial software like Corel AfterShot Pro doesn't make a lot of sense. And yet, this application offers plenty of powerful tools and advanced functionality that make it worth the asking price for some Linux photography enthusiasts.

Don't let the current version number of AfterShot Pro [1] fool you. This application is the continuation of the excellent Bibble Pro software, which was on version 5.2.3 when Corel took over its development and renamed it AfterShot Pro. In other words, AfterShot Pro is based on a mature and stable code base and features tried and tested graphical interfaces and tools (Figure 1). Maturity and stability are not AfterShot Pro's only attractions. The application is optimized for speed, and it runs blazingly fast, even on modest hardware.

Open source RAW decoders can give their proprietary counterparts a run for their money, but the RAW engine that powers AfterShot Pro can still yield better results with a minimum of manual tweaking. This can be particularly useful for photographers who want to get the most out of RAW files without too much work. AfterShot Pro will also appeal to users looking for advanced tools like adjustment layers, presets, versatile versioning functionality, tools for healing and cloning using regions, and high dynamic range (HDR) features. Add to this the ability to extend AfterShot Pro's default functionality via plugins, and you have a very compelling alternative indeed.

Although AfterShot Pro is a commercial product, it costs only a fraction of the price of its commercial competitors. The application is available in two editions: AfterShot Standard and AfterShot Pro. Among other things, Pro can handle more catalogs, layers, and custom correction tools. It also features the Perfectly Clear noise removal and powerful HDR functionality.

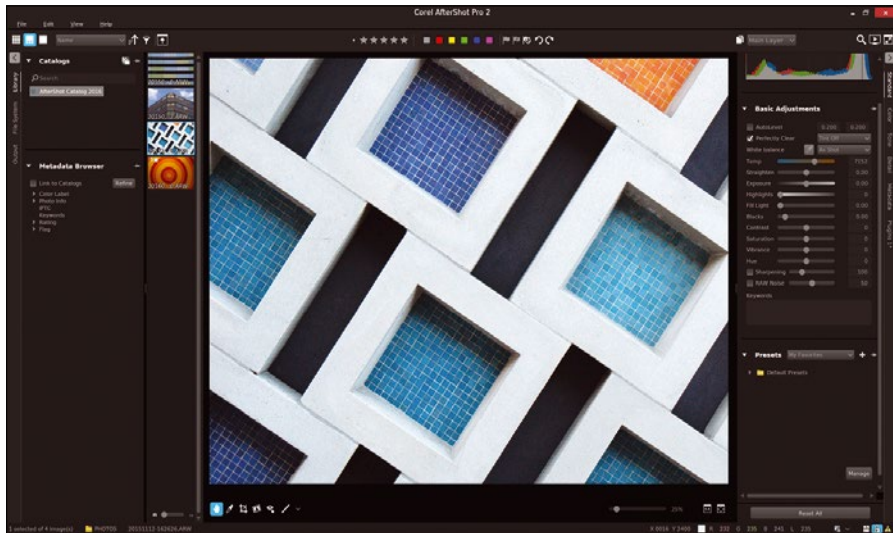
The AfterShot versus AfterShot Pro PDF document [2] provides a detailed de-

scription of the differences between the two editions. However, these differences are purely academic from a Linux user point of view, because only the Pro version of the application is available on Linux. Before you part with your money, though, you can download a 30-day trial version and see for yourself whether it fits your photographic needs. The trial version of AfterShot Pro (which can be unlocked later by purchasing a serial number) is available as DEB and RPM packages for both 32- and 64-bit architectures. Strangely, you can run the trial version in Standard mode, but you can't purchase a less expensive license key for it.

## Organizing and Managing Photos

AfterShot Pro has two modes for working with photos and RAW files. In the filesystem mode, you can access, manage, and edit photos without importing them first. In this mode, the application creates an accompanying XMP file (also called sidecar) for each photo. All modifications and data applied to the photo are stored in this file. To access photos and RAW files on the local disk, switch

Lead Image © Rene Walter, 123RF.com



**Figure 1:** Corel AfterShot Pro is a capable application for processing RAW files and organizing photos.

to the *File System* section in the left panel and navigate to the desired directory using the Directory View.

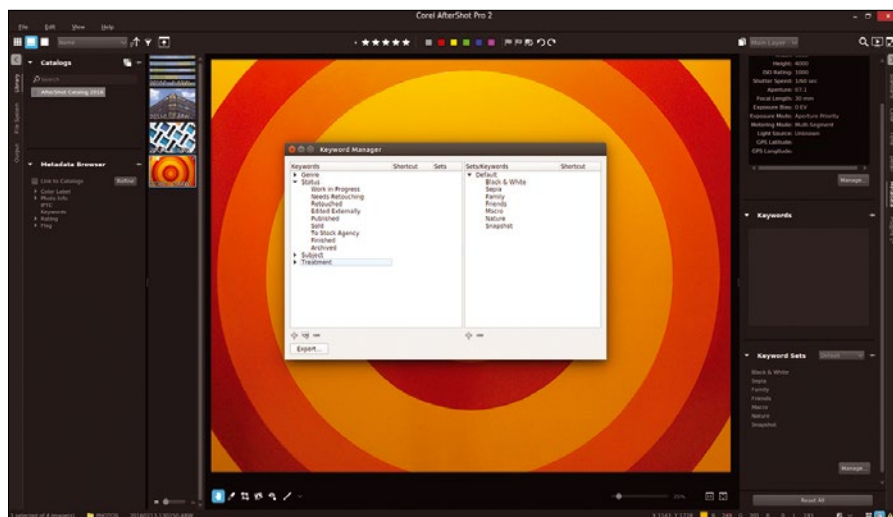
The filesystem mode can be useful for accessing and editing the occasional photo or two, but to get the most out of AfterShot Pro's capabilities, you should import photos into a catalog in the Library. When in filesystem mode, select the desired photos in the thumbnail bar, and choose *File | Import Selected Files*.

Alternatively, you can import an entire folder containing photos with *File | Import Photos from Folder*. During import, AfterShot Pro allows you to apply custom import settings, such as keywords and presets to all imported photos. Importing photos and RAW files into a catalog doesn't actually move the images: this simply creates a database and accompanying files for storing relevant data in a dedicated directory.

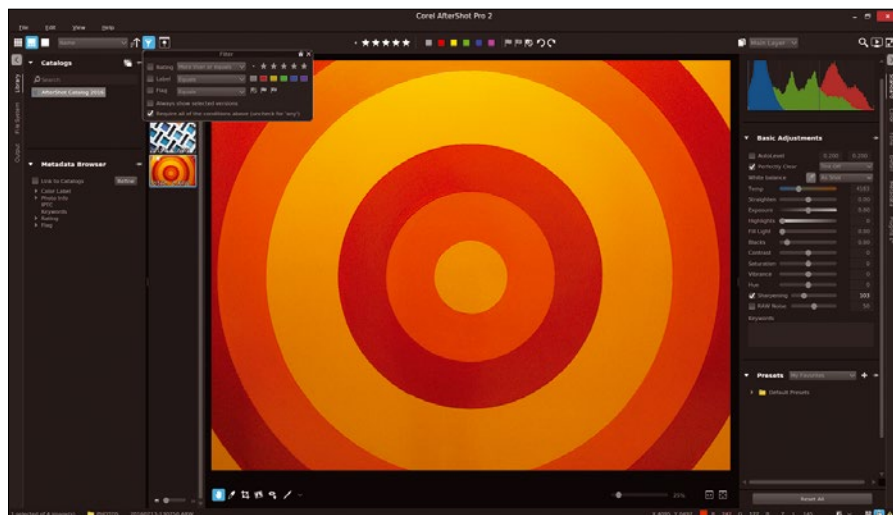
Importing photos into a catalog has several advantages. This allows the application to maintain a full editing history with every modification made to the photo over time. It also allows you to search and browse photos by their metadata, including EXIF and IPTC. The catalog also lets you maintain image stacks, where the master file and its versions are grouped together.

Like any decent photo management application, AfterShot Pro supports standard tools for classifying and describing photos, including ratings, color labels, and flags. You can also assign keywords to each photo by switching to the *Metadata* section in the right panel and entering the desired keywords into the Keywords field.

AfterShot Pro also lets you create sets containing lists of often-used keywords (Figure 2). For example, you can create the Travel set containing keywords like



**Figure 2:** AfterShot Pro allows you to create keyword sets.



**Figure 3:** Filter Tool lets you instantly filter images by specific criteria.

cityscapes, nightscapes, and streets by selecting *Manage* in the Keyword Sets section. In the Keyword Manager window, add the desired keywords and their children in the left pane. Next, create a new keyword set in the right pane. Then, add the keywords to the set by dragging them from the left pane.

The application's filtering functionality can be used to display photos that match a specified rating, color label, and flag. Click on the *Filter Tool* icon in the upper left corner in the main window and define the filtering criteria; the application then will automatically display matching images (Figure 3). The *Sorting* drop-down list next to the filter tool lets you sort photos by various criteria, including rating, focal length, ISO, and F-Number. The Metadata Browser section in the Library section of the left panel allows you to browse photos in the catalog by their metadata. Here, you can view photos by

camera model, aperture value, lens model, metering mode, and shutter speed.

### Editing Photos

AfterShot Pro supports non-destructive editing, meaning that none of the actions and modifications are applied directly to the original master file. The application saves all edits in the accompanying XMP sidecar file or in the catalog database. Additionally, AfterShot Pro makes it possible to create multiple versions of the original file. Creating a new version doesn't add a new image file: It simply stores a different set of edits for the master file. AfterShot Pro creates a new version of the master file as soon as you edit it. You can create additional versions at any time by selecting a thumbnail in the thumbnail bar and pressing the *Insert* key (or right-clicking on the thumbnail and choosing the appropriate command from the *Version* context menu).

All core adjustment tools in AfterShot Pro are tucked under the Basic Adjustments section that sits in the *Standard* tab in the right panel. Here, you'll find the usual assortment of tools for adjusting temperature, exposure, highlights, contrast, and saturation. In addition to this standard tool set, AfterShot Pro also features the proprietary Perfectly Clear tool. Billed as an intelligent image adjustment tool, Perfectly Clear makes it possible to improve photos with a single click. Although it's easy to dismiss this functionality as a gimmick, it is capable of producing decent results.

Preset support is de rigueur for any professional photo management application, and AfterShot Pro is no exception. The application comes with several ready-made presets in the Presets section under the *Standard* tab. Better still, AfterShot Pro lets you create custom presets based on the adjustments applied to the currently opened photo or RAW file (Figure 4). To create a new preset, press the *Add Preset* button in the Presets section, give the preset a descriptive name, select the adjustments you want to include, and hit *OK*.

To keep tabs on the presets, you can organize them into folders. To create a preset folder, press the *Add Folder* button at the bottom of the Presets section, give the folder a name, then drag and drop the desired preset items into the folder. All presets are saved as XMP files in the



Figure 4: AfterShot Pro supports custom presets.

~/AfterShot/Presets/ directory, so you can easily back them up and share with other AfterShot Pro users.

AfterShot Pro also allows you to copy and paste adjustments between images, which can be useful when you need to apply edits made to an image to another photo without creating a dedicated preset. To do this, select the source image in the thumbnail bar and choose *Edit | Copy Image Settings* (or *Copy Selective Image Settings* if you want to apply individual adjustments). Then, select the target image and choose *Edit | Paste Image Settings*.

The *Color* tab in the right panel contains color correction tools, including Curves, Color Balance, and Color Correction (Figure 5). The latter makes it possible to apply selective color adjustments to the image based on a selected color. You can select one of the default primary or

secondary colors in the color wells (color squares at the top of the Color Correction section) or pick any color from the image using one of the empty color wells. To do this, click on the empty color well. Click on the empty color square next to the *Hue* adjustment slider to choose a color from a color palette. Alternatively, enable the color picker by clicking on the *Selective Color Data Settings* icon next to the *Hue* slider, and click on the region of the image containing the desired color. You can then apply hue, saturation, luminance, and range adjustments to the selected color only.

Adjustment layers is where AfterShot Pro comes into its own, and the application makes use of layers for selective editing as well as healing and cloning adjustments (Figure 6). Selective editing can be useful when you need to make adjustments just to a specific area in the image.

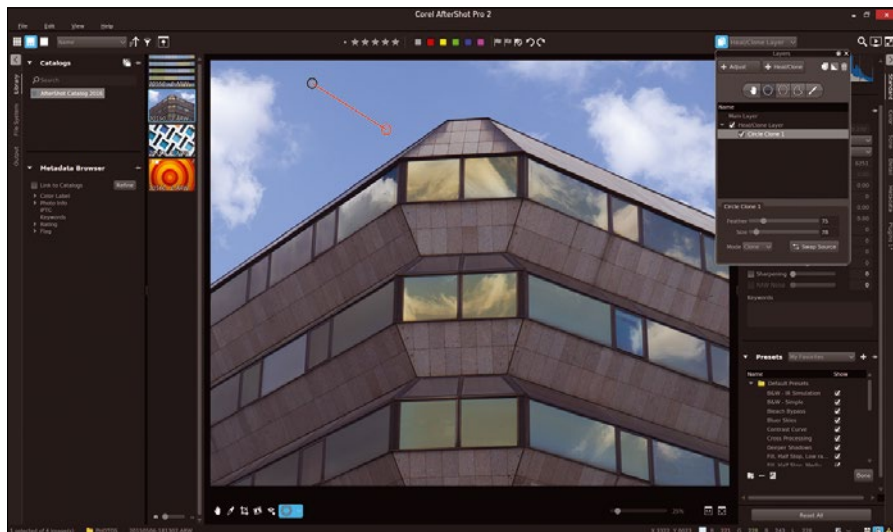


Figure 5: The application has all essential color adjustment tools, including Color Correction.

Suppose, for example, you need to lighten an underexposed area in a photo. Click the *Layer Manager* icon in the upper right corner of the top toolbar and press *Adjust* to add an adjustment layer. Use then a selection tool (you can choose between Circle, Polygon, Curve, and Brush) to define a selection area on the image. Once you've done that, all adjustments you make are applied to the selected area only.

The Heal and Clone adjustment layer works in a similar manner, but it's designed for removing minor spots and blemishes. To remove a spot, click the *Heal/Clone* button in the Layer Manager palette to create a new layer, choose the Circle selection tool, and click on the spot you want to fix. Choose the *Heal* or *Clone* option from the *Mode* drop-down list, and adjust the *Feather* and *Size* settings, if necessary.

AfterShot Pro's default functionality can be extended using plugins, and you can find a handful of free modules in the official plugin repository [3]. For example, AfterShot Pro lacks a perspective correction functionality; the zPerspector plugin fills the void (Figure 7). To install



**Figure 6:** AfterShot Pro uses adjustment layers for selective editing as well as healing and cloning actions.

this or any other plugin, download its .afzplug package, choose *File | Install Plugin* in AfterShot Pro, and select the downloaded file. Once installed, the plugin's interface can be accessed under the *Plugins* tab in the right panel.

Because AfterShot Pro stores all adjustments in XMP files or a database, you

need to run an export action to generate a processed photo. There are two ways to do that: Select the images you want to export in the thumbnail bar and choose *File | Export | Export Files* (or use *Ctrl + S*). In the Export Files dialog, specify the required settings (destination directory, image format, quality, size, etc.) and hit *OK*.

# Subscribe now!



Don't miss a single issue of the magazine that delivers the in-depth technical solutions you'll use everyday!



UK £ 49.90 | Europe € 79.90 | USA / Canada US\$ 99.95  
 Rest of World (by Airmail) US\$ 109.90  
 (DVD included with print edition)  
 Terms and conditions: <http://goo.gl/SSSQer>

[shop.linuxnewmedia.com/subs](http://shop.linuxnewmedia.com/subs)

# 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!



Specifying settings manually every time you export photos can quickly become a nuisance, so AfterShot Pro supports so-called batch presets, or ready-made export profiles. AfterShot Pro comes with several batch presets for common formats, such as 16-bit TIFF, 8-bit TIFF, and full-size JPEG. To export photos using a batch preset, choose *File | Export with Batch Preset* and select the desired preset item. Alternatively, you can use the keyboard shortcut assigned to the preset. The application lets you create custom batch presets, too (Figure 8). To create a batch preset, choose *File | Export with Batch Preset | New Batch Preset*, give the preset a name, and configure the available options.

## Final Word

In this article, I covered only some of the key features of AfterShot Pro, but it should give you a general idea of After-

Shot Pro's capabilities and help you to decide whether the application fits your photographic needs. Of course, adding AfterShot Pro to your photographic toolbox doesn't mean you have to give up other open source applications you already use. For example, you can still use digiKam's superior tools to import, rename, and organize photos, and then process individual RAW files in AfterShot Pro, when needed. ■■■

## INFO

- [1] Corel AfterShot Pro: [www.aftershotpro.com](http://www.aftershotpro.com)
- [2] AfterShot vs. AfterShot Pro: [img.aftershotpro.com/en/presskit/aftershot-version-matrix.pdf](http://img.aftershotpro.com/en/presskit/aftershot-version-matrix.pdf)
- [3] AfterShot Pro plugins: [www.aftershotpro.com/en/plugins](http://www.aftershotpro.com/en/plugins)
- [4] zPerspector plugin: [www.aftershotpro.com/en/plugins/zperspector](http://www.aftershotpro.com/en/plugins/zperspector)



Figure 7: zPerspector plugin in action.

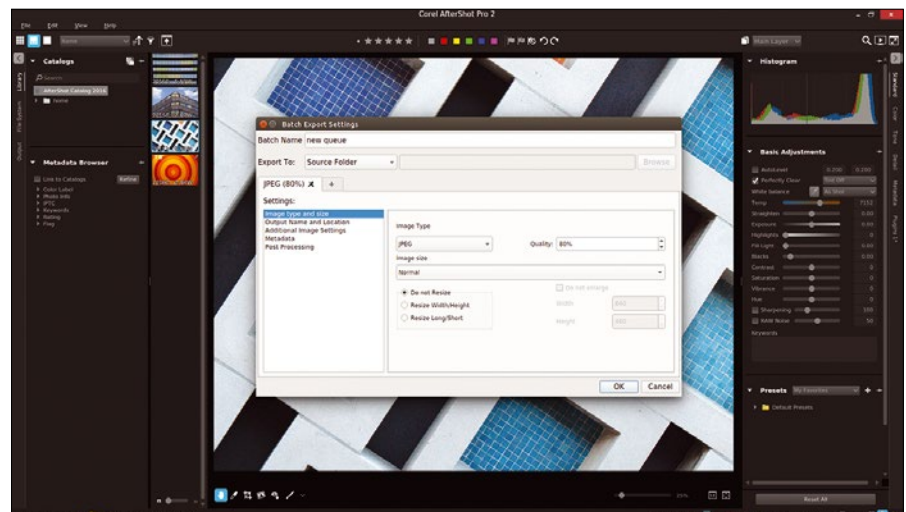


Figure 8: AfterShot Pro makes it possible to create custom batch output presets.



# Register Today!



## 2016 USENIX Annual Technical Conference

JUNE 22–24, 2016 • DENVER, CO

[www.usenix.org/atc16](http://www.usenix.org/atc16)

USENIX ATC '16 brings leading systems researchers together for cutting-edge systems research and unlimited opportunities to gain insight into a variety of must-know topics, including virtualization, system administration, cloud computing, security, and networking.



Co-located with USENIX ATC '16:

### SOUPS 2016

Twelfth Symposium on Usable Privacy and Security

JUNE 22–24, 2016

[www.usenix.org/soups2016](http://www.usenix.org/soups2016)

SOUPS 2016 will bring together an interdisciplinary group of researchers and practitioners in human computer interaction, security, and privacy. The program will feature technical papers, workshops and tutorials, a poster session, panels and invited talks, and lightning talks.

### HotCloud '16

8th USENIX Workshop on Hot Topics in Cloud Computing

JUNE 20–21, 2016

[www.usenix.org/hotcloud16](http://www.usenix.org/hotcloud16)

Researchers and practitioners at HotCloud '16 share their perspectives, report on recent developments, discuss research in progress, and identify new/emerging "hot" trends in cloud computing technologies.

### HotStorage '16

8th USENIX Workshop on Hot Topics in Storage and File Systems

JUNE 20–21, 2016

[www.usenix.org/hotstorage16](http://www.usenix.org/hotstorage16)

HotStorage '16 is an ideal forum for leading storage systems researchers to exchange ideas and discuss the design, implementation, management, and evaluation of these systems.



**Register by May 31 and save!**



Tool tests on the fast track *By Uwe Vollbracht*

# TOOL TIPS

## Term-Highlight 1.8.1

Function: Highlight patterns in terminal

Source: <http://sourceforge.net/projects/hlterm>

License: GPLv2

Alternatives: Grep, Supercat

```
Terminal - vollbracht@LMLab-1504.b:~
vollbracht@LMLab-1504.b:~$ hl -g hl -x2 -b 'positions' -x4 'string' - ~/extract/term
tips/Term-Highlight-1.8.1/Highlight.pm
#which will apply for the whole string
my ( $positions, $patterns, $string_ref ) = @_;
my $length = length $$string_ref;
push @$positions, ( [ 0, $length, 1, $i, $$patterns[ $i ] ],
while ( $$string_ref =~ /$$patterns[ $i ][ 0 ]/g )
push @$positions, ( [ $-[ 0 ], $+[ 0 ] - $-[ 0 ], 1, $i, $$pat
terns[ $i ] ],
my ( $positions, $patterns ) = @_;
for my $position( sort ByPositions @$positions )
my ( $positions, $string_ref, $tagtype ) = @_;
for my $position( sort ByPositions @$positions )
substr( $$string_ref, $position[ 0 ] + $offset, 0 ) = $colortag;
#string to process
#@positions contains the start and end positions of all found patterns in th
e current string
#do not change original string and return found positions if an array is exp
ected
=item Subj->Process( \$$string );
expects array of references to strings. Loads patterns to be processed.
expects reference to string. Makes substitution of color tags inside the
string. Returns count of found matches.
vollbracht@LMLab-1504.b:~$
```

Term Highlight – the name really says it all: The Perl script hl searches through text files for search patterns and color highlights the matches. If so desired, it will work like Grep and just output the lines with the matches. Optionally, Term Highlight can ignore the case and work recursively. The tool also understands regular expressions.

When defining the highlighting options and the search pattern, users can define the foreground and background color. The script reads from standard output or searches through files and directories explicitly specified by the user after the - sign.

User can store frequently needed highlighting in snippets in the ~/.hlrc configuration file. The unique designator is followed by the highlighting options and search patterns. After defining a snippet, you can enable it later on in the global call options following -s.

★★★★☆ Using the tool is easy but not always intuitive. If users forget the hyphen in front of the file list, hl just ignores the list. It is also somewhat unfortunate that the -b parameter has a different meaning in the global options and highlight options. ■■■

## BashBurn 3.1.0

Function: Burn CDs and DVDs at the command line

Source: <http://bashburn.dose.se>

License: GPLv2

Alternatives: Growisofs, Cdrecord

```
Terminal - vollbracht@LMLab-1504.b:~
+-----+
| BashBurn 3.1.0 |
+-----+
--(Audio Menu)
0) Burn Audio from Mp3s/Oggs/Wavs/FLACs
1) Burn Audio Directly
2) Copy Audio CD [CD to CD]
3) Copy Audio CD [CD to harddrive]
4) Burn M3U Playlist
5) Create MP3s from Wavs
6) Create Oggs from Wavs
7) Create FLACs from Wavs
8) Create MP3s from CD
9) Create Oggs from CD
10) Create FLACs from CD
11) Back

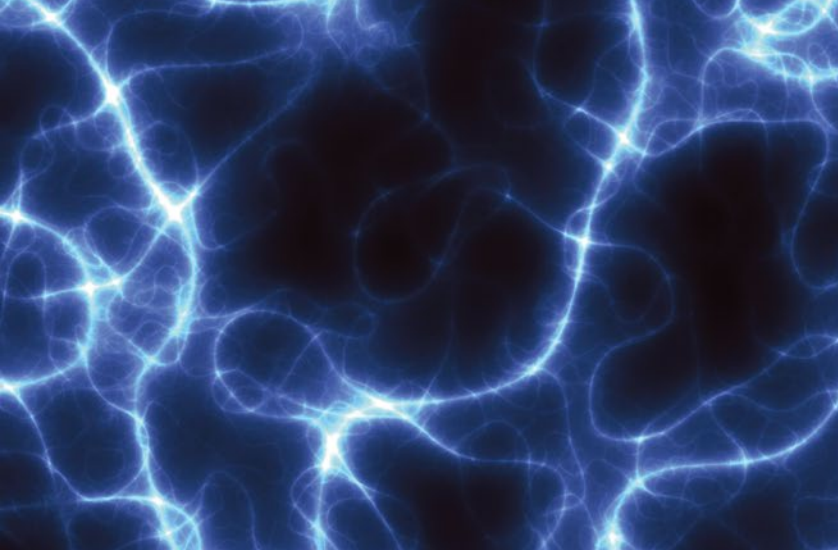
Your Choice? [0-11] |>
```

If you do not have a graphical burning tool, or you just prefer to use the shell, you will find a practical helper in the form of BashBurn. The shell script acts as a front end for tools like mkisofs, cdrecord, growisofs, cdparanoia, and others. It burns data CDs and DVDs, as well as music CDs and images; rips audio CDs; converts NRG files from the Nero application into standard ISO files; and much more.

When launched, the program comes up with a simple menu. You can use the program settings to define where the tool stores temporary files and set optional parameters for the underlying console tools.

The Data section contains entries for copying media or burning files and directories to disk, creating ISO images, burning existing images, and so on. The Audio section contains entries for copying CDs. You also have the option of converting WAV files to MP3, Ogg Vorbis, or FLAC; ripping a CD; or burning play lists.

★★★★★ In terms of the feature scope, BashBurn can easily keep up with graphical burning tools. The shell script also does a good job when you need to work on a remote system. ■■■



# Shop the Shop

shop.linuxnewmedia.com

Discover the past and invest in a new year of IT solutions at Linux New Media's online store.

Want to subscribe?

Need training?

Searching for that back issue you really wish you'd picked up at the newsstand?

## DIGITAL & PRINT SUBSCRIPTIONS



## SPECIAL EDITIONS



## TRAINING



- LPIC-1 LPI 101 - CompTIA Linux+ LX0-101
- LPIC-1 LPI 102 - CompTIA Linux+ LX0-102
- LPIC-1 - CompTIA Linux+ 101 + 102

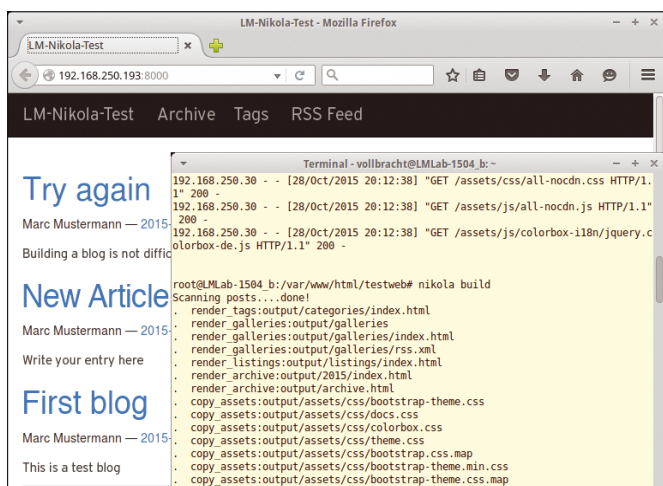
## Nikola 7.7.2

Function: Generate static websites

Source: <https://getnikola.com>

License: MIT

Alternatives: Jekyll, Hyde, Pelican



For security or performance reasons, many users rely on static content for their websites. Nikola provides support. The tool, written in Python, creates attractive pages, including links, site maps, RSS feeds, and more. It can also import content from platforms like WordPress, Tumblr, or Blogger. Nikola works with Python 2.7 and 3.3, but it needs many additional libraries despite the lean codebase. The developers recommend installing via the Python package manager pip.

To design a new website, you run `nikola init` and follow the wizard. You are prompted to provide details on the title, language, time zone, and the like. Then, you fill the framework with content. Options such as `new_page` or `new_post` help you do so. The `-e` switch launches the text editor defined in `$EDITOR`.

When done, type `nikola build` to compile, and `deploy` to publish the page. If you want to test your work locally in the browser before going live, you can launch the integrated web server by typing `serve`. Other useful options include `github_deploy` (uploads to GitHub), `status` (shows you when the content was last published), and `orphans` (finds orphaned pages).

★★★★★ Nikola is a very powerful tool. The project website offers a number of extensions and themes, and the comprehensive documentation leaves no questions unanswered. ■■■

### Steel 1.1

Function: Password manager for the shell

Source: <http://www.steelpasswordmanager.org>

License: GPLv3

Alternatives: KeePass

```
Terminal - vollbracht@LMLab-1504_b: ~/extract/tooltips
vollbracht@LMLab-1504_b:~/extract/tooltips$ steel -o /home/vollbracht/steel_db
Master passphrase:
vollbracht@LMLab-1504_b:~/extract/tooltips$ steel -l

Id          1
Title       server1
Username    admin
Passphrase  geheim
Address     server1.de
Notes      PDC
-----
Id          2
Title       server2
Username    webadmin
Passphrase  streng geheim
Address     www.mydom.de
Notes      content
-----
vollbracht@LMLab-1504_b:~/extract/tooltips$ steel -g 10
XocdoBtZE
vollbracht@LMLab-1504_b:~/extract/tooltips$ steel -a server2 paula www.mydom.de
e user
Enter new passphrase:
Retype new passphrase:
```

If you need to remember a large collection of password credentials, a password manager is the best place to keep them. In addition to various graphical variants, Linux also has something to offer command-line fans. Steel is a shell safe; the C program relies on SQLite and the Rijndael algorithm (256-bit).

You need to create a password database, and you can use the `-i` option to do this on first launching the program. The database contains a single table named *entries*. To add a new password, type `steel -a`. If you prefer to be helped rather than relying on your own creativity, you can type `-g` to generate a password automatically. Users state the password length as a number after the option.

The `-l` parameter lists the database content; `-f` lets you search for certain entries. To finish working with Steel, simply close the password database with `-c`. If you forget to do so, SQLite clients can read the content. You can set the master password each time you close.

Previous Steel versions showed passwords in the clear; the current version corrects this. Users now need to tell the tool to do this explicitly by specifying `-p`. If you want to migrate from a version before 1.0 to the new Steel, you first need to open the password database with the older version and then close with the newer version because the encryption algorithms are incompatible.

★★★★☆☆ Steel is fast and easy to use. The database is not permanently encrypted, which is a risk. ■■■

### Ukopp 5.8

Function: Backups on USB media

Source: <http://www.kornelix.com/ukopp.html>

License: GPLv3

Alternatives: Sysbak

Ukopp is a practical little GUI that creates backups on USB sticks and other pluggable media. A click takes you to the setup, where you can navigate the filesystem, select or exclude files, store the settings in recurring backup jobs, and more.

When launched, Ukopp comes up with a clear-cut window showing the most important functions at the top. In the main part of the window, you will see the status messages. Before creating your first backup, you need to define a job. To reduce the data volume, Ukopp only saves the changed data and directories in each run.

On the backup medium, Ukopp replaces the older versions with newer ones and keeps files that no longer exist on the source medium. In the job settings, you can define whether to keep older versions and, if so, how many on the backup medium; you can also define how long these versions are kept.

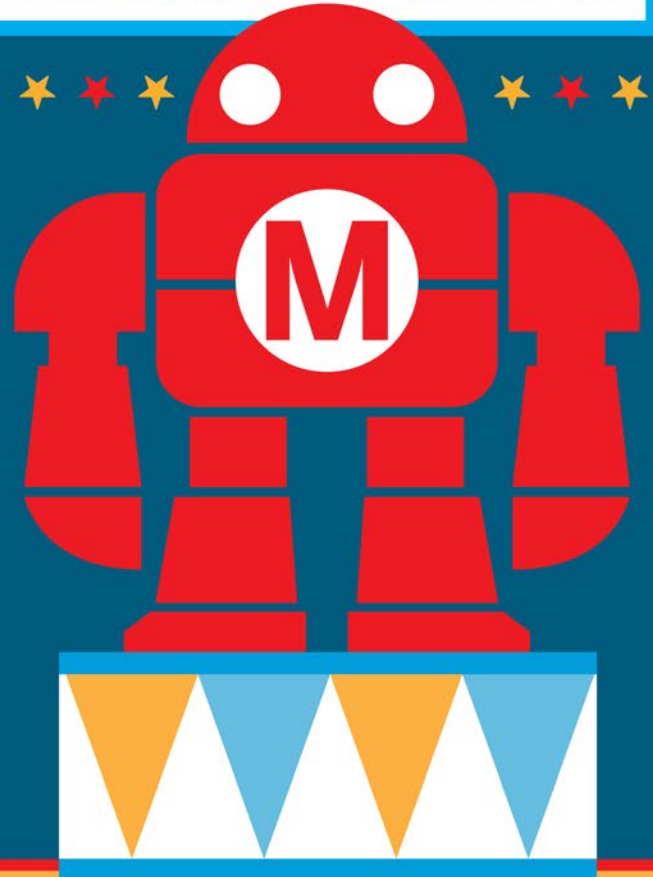
Ukopp uses the source directory structure and stores the execute and access permissions. It does not have functions for compression or encryption. If so desired, you can check the consistency of the backed up data after each run. Ukopp supports three modes here: full, increment, and compare.

★★★★☆☆ The backup program is clear-cut and handy. Because Ukopp does not compress the data, you need to have enough free space on the target drive. ■■■

PRESENTED BY 

# 11<sup>th</sup> ANNUAL **Maker Faire**<sup>®</sup> BAY AREA

GREATEST  
*Show*  
& **TELL**  
ON *Earth*



**MAY 20-22** SAN MATEO EVENT CENTER  
FRI 1-5PM, SAT 10AM-7PM, SUN 10AM-6PM

**OCT 1+2** 7TH ANNUAL WORLD MAKER FAIRE  
NEW YORK HALL OF SCIENCE

PRESENTING



GOLDSMITH SPONSORS



SILVERSMITH SPONSORS



GET YOUR  
**TICKETS**  
TODAY!

 [makerfaire.com](http://makerfaire.com) BROUGHT TO YOU BY  **Make:** MAGAZINE



An indexing search engine with Nutch and Solr

# Go Find It!

Build your own search engine using Apache's Nutch web crawler and Solr search platform.

By Markus Feilner and Sebastian Mogilowski

**E**MS, wikis, text files ... modern companies store important data in many different places, and that data must be accessible down to the tiniest detail through a single search. Commercial software vendors such as Google [1] offer tools that will index the data and store the index on an external server. But many organizations prefer to keep control of the search capabilities – for security and privacy reasons, but also to add flexibility and promote innovation and customization.

A handy constellation of open source tools from the Apache project will help you build your own search index for the assorted documents and data on your network: Nutch, Solr, Apache, and Lucene.

Nutch [2] is a powerful web crawler, and Apache Solr [3] is a search engine based on Apache Lucene [4]. You can

combine Nutch with Solr to create a complete search engine – a miniature Google, if you like.

The Nutch crawler uses HTTP and FTP to discover information. If you want Nutch to inspect your local files, you need to store the files on an HTTP or FTP server and point to the directories you want Nutch to crawl. Nutch fetches data that is then searched and indexed by Solr. Solr depends on the Apache Lucene search libraries and is written in Java, and it requires a Java Servlet container server. The Jetty Java Servlet container tool is installed by default, but many users prefer a more robust solution such as Apache Tomcat. (See “A Note of Caution” box for more info.)

This workshop shows how to build your own search engine using an Ubuntu 14.04.2 LTS system.

## Installing the Components

On Canonical's Enterprise Linux, Solr is available from the package sources; you only need to install Nutch manually (Listing 1, lines 1-4). Then back up Solr's default XML schema and replace it with the file supplied by Nutch (Lines 6 and 7).

By default, the server does not save the content of pages or documents it

finds. When it re-indexes, it transfers all the contents again. If you want to enable caching, you can do so in the `/etc/solr/conf/schema.xml` configuration file by changing the `stored="false"` entry in the following line:

```
<field name="content" type="text" ?
  stored="false" indexed="true"/>
```

line to "true"; then restart Tomcat by typing `service tomcat6 restart`.

## Configuring the Nutch Crawler

Although you can control the crawler's default behavior with the `/opt/nutch/conf/nutch-default.xml` file, it makes more sense to customize the `/opt/nutch/conf/nutch-site.xml` file with site-specific details.

The example in Listing 2 shows how you can configure the name of the HTTP agent. This name will appear in the web server's logfiles.

The `nutch-default.xml` file contains various settings that control the crawler's behavior. In `nutch-site.xml`, you need to do this:

```
<property>
  <name>file.content.ignored</name>
```

### A NOTE OF CAUTION

The crawler indexes data accessible to the daemons associated with the process. Depending on your security system, the search results could be more than you would want non-privileged users to see, so you might need to adjust your configuration to rule out access to highly secure files and directories.

**LISTING 1: Installing Solr and Nutch**

```
apt-get install solr-tomcat
wget http://www.eu.apache.org/dist/nutch/1.9/apache-nutch-1.9-bin.tar.gz
tar vfx apache-nutch-1.9-bin.tar.gz
mv apache-nutch-1.9 /opt/nutch

mv /etc/solr/conf/schema.xml /etc/solr/conf/schema.xml.orig
cp /opt/nutch/conf/schema.xml /etc/solr/conf/schema.xml
```

```
<value>>false</value>
</property>
```

Additionally, you need Nutch to remove any documents that the users have deleted in the meantime from the search engine's database:

```
<property>
  <name>db.update.purge.404</name>
  <value>true</value>
</property>
```

On a local network, with few servers and clients compared with the Internet, the five-second default setting between two requests to the same server leads to an unnecessarily large number of inactive threads, which slows down the search engine. The `fetcher.server.delay` is useful for ensuring that the search engine will not overload a server with requests:

```
<property>
  <name>fetcher.server.delay</name>
  <value>0.0</value>
</property>
```

It makes sense to disable this value and only re-enable it if problems occur.

**Large Documents**

On the Internet, it is sometimes useful to index large documents, but you need to be careful not to let the crawler get hung up on a gigantic tome with no useful information. Nutch lets you define the

`content.limit` class parameters that define the maximum size of the content that crawler processes (Listing 3). You can also define the length of the document title, say, to achieve a more informative view in the search results – the value is in characters not in bytes:

```
<property>
  <name>indexer.max.title.length</name>
  <value>150</value>
</property>
```

Another useful variable, `fetcher.threads.fetch`, defines the number of concurrent threads reading content. The `http.timeout` reduces the time the thread needs to wait for a request to time out.

**Indexed**

If you want to prevent a search engine server from accessing the Internet, you can define a very short HTTP timeout using a firewall. The search engine will find external URLs in the documents, but its attempts to resolve the URLs will fail if you set a sufficiently short timeout. The crawler will therefore find external URLs, but not reach them and thus not add them to the database.

For a cleaner approach, you can use the `regex-urfilter.txt` file in Nutch's `conf` directory. The `regex-urfilter.txt` file lets you define exceptions. (Nutch already has some default rules that prevent the crawler from reading unnecessary files such as CSS files or images.)

**LISTING 2: nutch-site.xml**

```
01 <?xml version="1.0"?>
02 <?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
03 <!-- Put site-specific property overrides in this file. -->
04 <configuration>
05   <property>
06     <name>http.agent.name</name>
07     <value>Company Search Agent</value>
08   </property>
09 </configuration>
```

The following command

```
-^(http|https)://www.wikipedia.com
```

stops Nutch from following links to <http://www.wikipedia.com>. It makes even more sense to define a whitelist and thus only permit individual servers:

```
+^(http|https)://intranet.company.local
```

Or you can specify multiple addresses using regular expressions:

```
+^http://([a-z0-9\-\A-Z]*\.)?
*.company.local/([a-z0-9\-\A-Z]*\/*)
```

The important thing is to correct the last line, which defines the general policy:

```
# accept anything else
+.
```

and replace it with:

```
# deny anything else
-.
```

You need to deny unspecified traffic in order for the list to serve as a whitelist.

**On Your Marks**

Everything is set up; the search can start – all you need to do is tell the crawler the starting point of its journey. You define the subdirectories and the file that contain the URLs in `/opt/nutch/`:

```
mkdir /opt/nutch/urls
echo "http://intranetserver.
company.local" >
/opt/nutch/urls/seed.txt
```

Any number of URLs are permissible.

**LISTING 3: File Lengths**

```
01 <property>
02   <name>file.content.limit</name>
03   <value>131072</value>
04 </property>
05 <property>
06   <name>http.content.limit</name>
07   <value>131072</value>
08 </property>
09 <property>
10   <name>ftp.content.limit</name>
11   <value>131072</value>
12 </property>
```

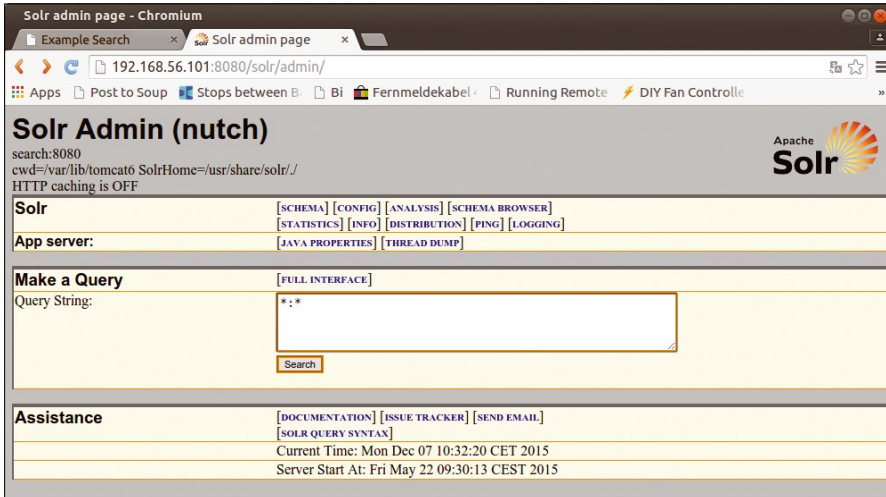


Figure 1: The Apache Solr admin interface is deliberately simple.

Then start the crawler:

```
export JAVA_HOME=  
/usr/lib/jvm/java-7-openjdk-amd64/jre/  
/opt/bin/crawl /opt/nutch/urls/   
/opt/nutch/IntranetCrawler/   
http://localhost:8080/solr/ 10
```

The first parameter in the crawl command specifies the directory containing the seed.txt file.

Nutch runs fetcher processes that load and parse the discovered content. The /opt/nutch/IntranetCrawler option specifies the directory in which you want Nutch to create this content. The address follows, including the Solr server port to which Nutch saves the results.

The number 10 at the end states the number of crawler runs. Depending on the pages it finds and the search depth, it can take some time for the command to complete. For initial tests, you might prefer a value of 1 or 2.

When the fetcher downloads and parses the results, it typically finds more links to more content. These links end up in its link database. On the next run, the crawler reads these URLs too and hands them over to the fetcher processes. This to-do list with links for the crawler grows very quickly at the start, because the crawler can only process a certain amount of content during each run.

Nutch breaks down the discovered links into segments, which it processes one by one. A segment only contains a certain number of links; for this reason, it can happen that the crawler creates new segments immediately after the

first one that it is processing. The new links found in this process end up in new segments.

The fact that the script has stopped running does not mean the crawler has

found all the content; some content may have been saved for fetcher processes later on.

After the first run, however, the content should all be available for displaying in the Solr web interface: `http://<Searchserver>:8080/solr/admin/` (see also Figure 1).

How often the crawler runs – every night, once a month, or on the weekend – is a question of the data volume and your need for up-to-date results. The important thing is that Nutch only finds data to which a link points – from a website or an indexed document. Non-linked documents virtually don't exist, except in the FTP or HTTP directory listings.

### Querying with jQuery

Admins typically integrate the Solr search directly into an existing intranet portal. Solr provides an API for this pur-

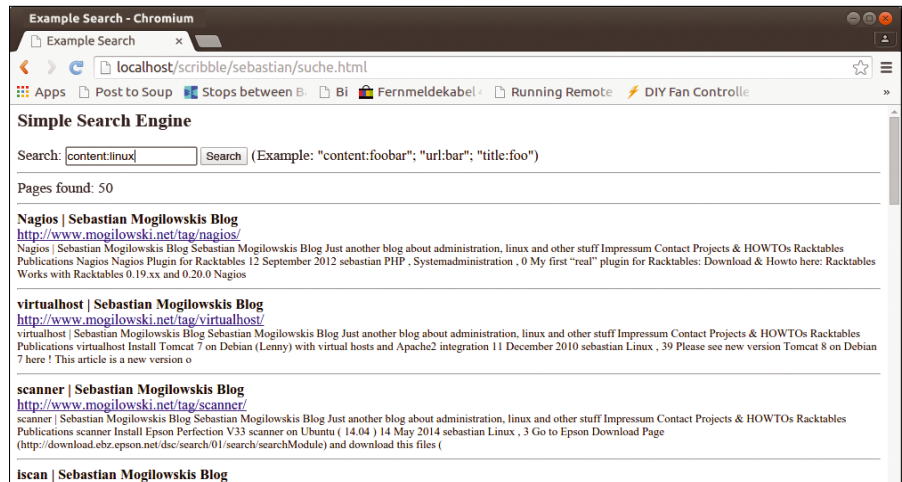


Figure 2: Thanks to jQuery, the search engine can be quickly and easily integrated into your own websites.

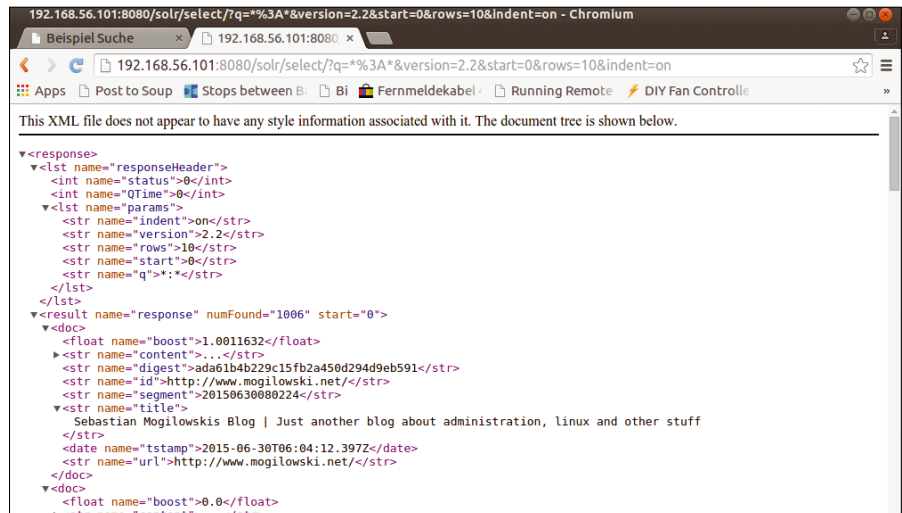


Figure 3: The jQuery script builds the HTML websites from this XML response.



## LISTING 4: jQuery Query

```

01 <html>
02 <head>
03 <title>Example Search</title>
04 </head>
05 <body>
06 <h3>Simple Search Engine</h3>
07 Search: <input id="query" />
08 <button id="search">Search</button> (Example:
09 "content:foobar"; "url:bar"; "title:foo")
10 <hr/>
11 <div id="results">
12 </div>
13 <script src="http://ajax.googleapis.com/ajax/libs/
14 jquery/1.4.2/jquery.min.js"></script>
15 <script>
16 function on_data(data) {
17     $('#results').empty();
18     var docs = data.response.docs;
19     $.each(docs, function(i, item) {
20         if (item.content.length > 400)
21             contentpart = item.content.substring(0,400);
22         else
23             contentpart = item.content
24         $('#results').prepend($(
25             '<strong>' + item.title + '</strong><br/>' +
26             '<a href="'+ item.url +' " target="_blank">'+
27             item.url + '</a>' +
28             '<br/><div style="font-size:80%;>' + contentpart +
29             '</div><hr/>'));
30     });
31
32     var total = 'Found page: ' + docs.length + '<hr/>';
33     $('#results').prepend('<div>' + total + '</div>');
34 }
35
36 function on_search() {
37     var query = $('#query').val();
38     if (query.length == 0) {
39         return;
40     }
41
42     var solrServer = 'http://SEARCHSERVER:8080/solr';
43     var url = solrServer + '/select?q=' +
44         encodeURIComponent(query) +
45         '&version=2.2&start=0&rows=50&indent=on&wt=
46         json&callback=?&json.wrf=on_data';
47     $.getJSON(url);
48
49     function on_ready() {
50         $('#search').click(on_search);
51         /* Hook enter to search */
52         $('body').keypress(function(e) {
53             if (e.keyCode == '13') {
54                 on_search();
55             }
56         });
57     }
58     $(document).ready(on_ready);
59 </script>
60 </html>

```

pose that manages database access and returns the search results. DIY queries with jQuery are a useful solution. Listing 4 shows the HTML code for a simple website with jQuery scripts (see Figure 2).

Figure 3 shows the server's response in the XML code transferred – a simple example with primitive search requests against the Solr back end.

In the search query, the user has many options. You can rank the page title higher than the content. The statement `content:(linux) title:(linux)^1.5` gives a match in the title one and a half times more weight than a match that is found in the document body. On the other hand, you can search for pages that contain the word “Linux” without the word “Debian.” In this case, you might still want to give the title preferential treatment:

```

content:(linux -debian) 2
title:(linux -debian)^1.5

```

Logical ANDs are easily achieved with a simple plus sign; OR-ing is the default; thus, `content:(+linux +debian)` searches for Linux and Debian.

Without the plus sign, the Solr-Nutch team would show you any documents containing *Linux* or *Debian*. Quotes let you search for complete terms (`content:("Linux Live USB Stick")`).

These simple forms of user interaction give you a good idea of the potential this option offers for jQuery and web programmers. What you will need for a professional-looking search engine front end is forms and point-and-click queries, as well as input validation and the ability to highlight the search key in the results. ■■■

## INFO

- 1] Google Search: <https://www.google.com/work/search/>
- 2] Nutch: <https://nutch.apache.org>
- 3] Solr: <https://lucene.apache.org/solr/>
- 4] Lucene: <https://lucene.apache.org>

## AUTHORS

**Markus Feilner** is a Linux specialist from Regensburg, Germany. He has worked with the free operating system since 1994 as an author, trainer, consultant, and journalist. The Conch Diplomat, Minister of the Universal Life Church, and Jedi Knight leads SUSE's document team in Nuremberg.



**Sebastian Mogilowski** is a computer scientist from Regensburg, Germany, who enjoys angling in his free time. He has been working part time since 2005 at a large data center, while also working as a freelance Linux administrator, consultant, and author. He focuses on open source system management, virtualization, and home automation.



# Subscribe today and join the revolution!

Each issue of Raspberry Pi Geek is an adventure, with ingenious applications and cool projects for Raspberry Pi, Arduino, and other maker-board systems!

Discover the secrets that will empower you to envision and build your own Raspberry Pi inventions.

## News

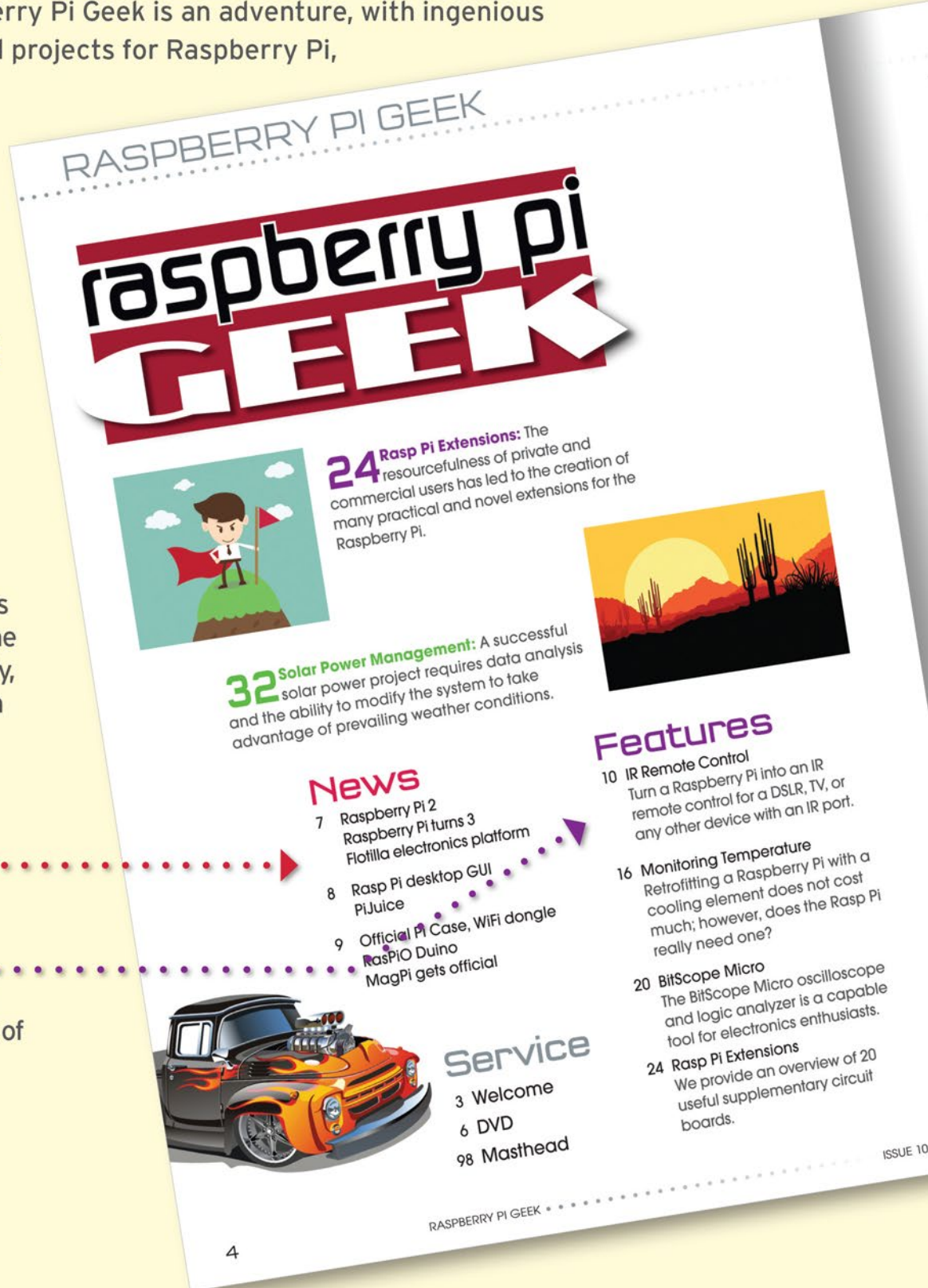
The latest developments from the front lines of the Raspberry Pi community, including information on new products and upcoming events.

## Features

Special articles on new technologies and topics of interest to the Rasp Pi community.

6 print issues with 6 DVDs or 6 digital issues for only

**\$59.95 £37.50 €44.90** [shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)



RASPBERRY PI GEEK

# raspberry pi GEEK



**24 Rasp Pi Extensions:** The resourcefulness of private and commercial users has led to the creation of many practical and novel extensions for the Raspberry Pi.

**32 Solar Power Management:** A successful solar power project requires data analysis and the ability to modify the system to take advantage of prevailing weather conditions.



## News

- 7 Raspberry Pi 2
- Raspberry Pi turns 3
- Flotilla electronics platform
- 8 Rasp Pi desktop GUI
- PiJuice
- 9 Official Pi Case, WiFi dongle
- RasPiO Duino
- MagPi gets official



## Service

- 3 Welcome
- 6 DVD
- 98 Masthead

## Features

- 10 IR Remote Control  
Turn a Raspberry Pi into an IR remote control for a DSLR, TV, or any other device with an IR port.
- 16 Monitoring Temperature  
Retrofitting a Raspberry Pi with a cooling element does not cost much; however, does the Rasp Pi really need one?
- 20 BitScope Micro  
The BitScope Micro oscilloscope and logic analyzer is a capable tool for electronics enthusiasts.
- 24 Rasp Pi Extensions  
We provide an overview of 20 useful supplementary circuit boards.

4

RASPBERRY PI GEEK

ISSUE 10





## Highlights

**10** **IR Remote Control:** Turn a Raspberry Pi into an infrared remote control.

**42** **Arduino Name Badge:** Build a programmable name tag with an Arduino brain.

**54** **Monitoring a Nest Box:** A nest box complete with camera keeps track of hatchlings.

**73** **Automobile Climate Control:** An old iPad and an Arduino Mega team up.

## Projects

- 32** **Solar Power Management**  
Tracking the sun and monitoring data to optimize a solar power system.
- 42** **Arduino Name Badge**  
A programmable name badge with a Steampunk theme.
- 46** **Modding the Robosapien, II**  
Hacking a Robosapien robot, Part 2: Control many motors by adding a port expander.

- 54** **Monitoring a Nest Box**  
A camera in a nest box lets you know how the newest hatchlings are faring.
- 62** **littleBits: callBit**  
A call button for a bed-ridden housemate.
- 68** **LED Shows**  
Two LED projects.
- 73** **Automobile Climate Control**  
Create an in-dash climate control app.

## Skills

- 80** **SwitchDoc Labs: Logic Analyzer**  
Use a logic analyzer to isolate problems in your hardware setup.

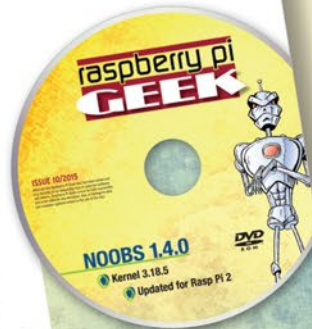
## Kid Stop

- 88** **Lego Projects**  
Lego Mindstorms NXT plus Rasp Pi equals cool projects.
- 92** **Scratch: PicoBoard Sensors**  
A PicoBoard connects Scratch projects to the physical world.



## Community

- 96** **Rasp Pi on the Space Station**  
We talk with Tim Peake, British ESA Astronaut, about the Raspberry Pis pegged for the International Space Station.



## NOOBS 1.4.0

- Kernel 3.18.5
- Updated for Rasp Pi 2

See p6 for details!

RASPBERRY PI GEEK

ISSUE 10

## Skills

Timely tutorials to help you build your skills with Linux and other underlying technologies.

## Kid Stop

Special projects for kids, including a new Scratch programming exercise in every issue.



## Projects

Do-it-yourself, real-world projects that let you learn by doing.

## DVD

Every print issue comes with a free DVD!

## Community

Interviews and reports on Raspberry Pi meet-ups around the world.

Get your Pi on!





### System monitoring for a new generation with Prometheus

# BIG WATCHER

Legacy monitoring solutions are fine for small-to-medium-sized networks, but complex environments benefit from a different approach. Prometheus is an interesting alternative to classic tools like Nagios. *By Martin Loschwitz*

**W**here monitoring is required, alerting and trending are never far away. Alerting plays a major role in practically any monitoring environment; the idea is to draw the administrator's attention to failures. And, trending is also important. Trending helps the admin detect potential bottlenecks at an early stage.

A quick look at the available monitoring solutions shows why Monitoring, Alerting, and Trending (MAT) are still an issue for many networks, particularly large and complex networks. Nagios, which has dominated the monitoring market for a long time, is a behemoth of complexity and comes with some inherent weaknesses.

Nagios alternatives such as Icinga have attempted to address some of the issues, but their scalability is limited.

The ballast of compatibility with Nagios and its plugins aggravates the situation. A state-of-art feature like trending was not exactly designed into the legacy Nagios. PNP4Nagios [1], a performance-tracking Nagios add-on, is one of the few options for useful trending with Nagios (Figure 1).

### SoundCloud as the Precursor

SoundCloud from UK was confronted with the challenge of implementing a monitoring solution. The company operates a streaming service along the lines of Spotify or Apple Music. The real challenge from the outset was to build a MAT system that would work reliably with thousands of nodes. Instead of combining existing components to create a better-than-nothing solution, Sound-

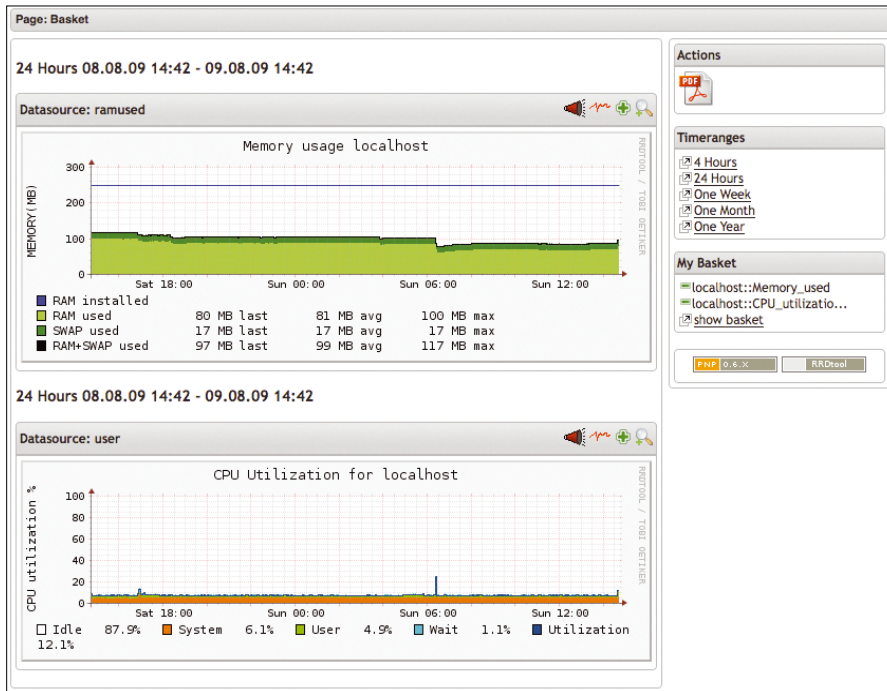
Cloud decided to explore unknown territory. The company chose to develop its own monitoring system and the result was Prometheus [2].

Compared with established solutions like Nagios, Prometheus has one very special feature: It comes with its own storage system to manage the data acquired from the network. Prometheus' internal database is based on the concept of the time series database. And, Prometheus tends to think more in the dimension of a complete metric rather than focusing on individual alerts. To understand what that means, I will take a short detour into the storage universe.

### How MAT Systems Manage Data

Classic monitoring systems, such as Nagios, do not have very sophisticated

Lead Image © Kurhan, 123RF.com



**Figure 1:** Legacy solutions such as PNP4Nagios generate heavy load when computing graphs and still take a huge amount of time – especially if you need to map longer periods of time.

data management, and they don't actually need it. The important thing with monitoring is whether a service is running properly *right now*. When you add the topic of trending, things start to become more difficult: Trending means you need long-term records relating to the availability of the service or the load on the existing infrastructure.

PNP4Nagios, for example, supports a database such as MySQL in the background in order to store the required values for a long period. MySQL is actually not designed for this kind of use, which can lead to problems. The volume of data you need to manage will grow extremely quickly in any large installation. The persistent storage on which all your trending data resides thus needs to scale just as easily as the entire platform. This is particularly true of the storage, but it also applies to the way in which the database handles a continuously increasing volume of data.

Also, preparing the data is a challenge: the data reaches the MAT system sorted in order of time, but at the other end, you'll need to output the data to reflect specific services. For example: the MAT system is regularly supplied with data points from its target systems for various services in consecutive order, such as "9AM: CPU load 1, RAM utilization 30 percent, and disc space usage 15 per-

cent." However, administrators will typically want to know what the CPU load looked like in a specific period, for example between 9AM today and the same time the previous morning.

Storing and manipulating large amounts of data in a database is an extremely resource-hungry process, and MySQL, in particular, loves taking its time with queries from tools like PNP4Nagios. A time-series database, such as the database used with Prometheus, offers an alternative approach.

Basically, a time series database is no more than a database that is designed for storing data in temporal relation. (See the box titled "Not the First, But the Best.") The data is converted by algorithms directly in the database. Prometheus is thus better equipped to take on a complex task such as trending thanks to its data model.

Typical monitoring and alerting is then no more than a side product: If no results are received for a specific metric over a period of time, the system assumes the service is not running correctly and sounds the alarm.

## Prometheus Modular Architecture

Under the hood, Prometheus relies on a modular architecture. The core of the application – that is, the time series data-

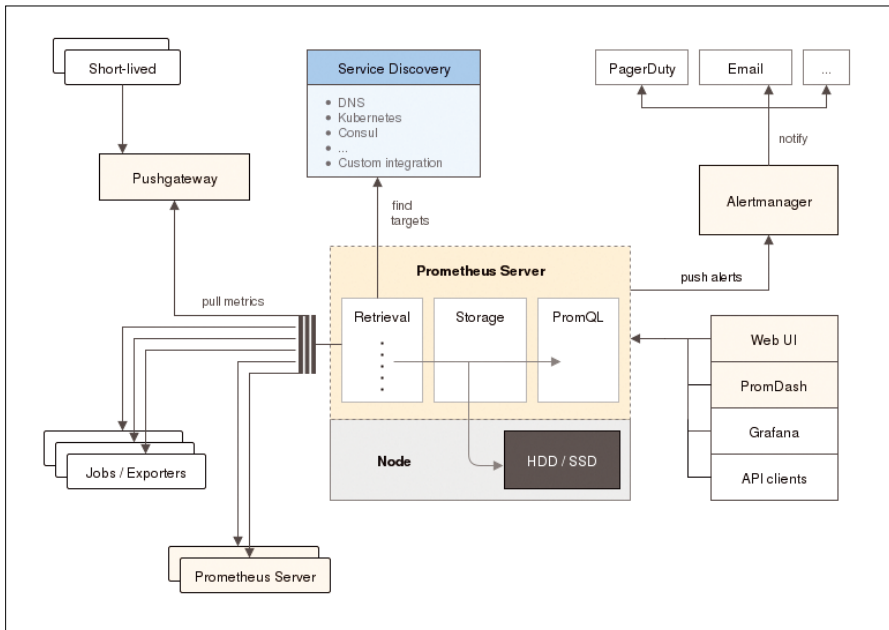
base – is programmed in Go, just like most of the applications in the Prometheus distribution. The database comes with its own web interface and a separate tool for alert management (the Alert Manager). Exporters for the target host are important – *exporter* is basically another word for agent: The node exporter, for example, logs various data for metrics such as CPU load or RAM usage on the host on which it is running, giving the Prometheus database the ability to pull this data when needed. If the service needs to push its data to the MAT system, you can deploy the push gateway, which fields the data from the services and stages it for the database.

At the heart of the system is the Prometheus server (Figure 2). The server handles many tasks, the most important of which is storing the measurement data acquired in the cloud. Although Prometheus comes from the cloud camp, the service is lagging behind in scalability. Although you can easily run any number of Prometheus instances within the same setup, in contrast to many other solutions, Prometheus does not rely on shared storage on the back end.

The Prometheus developers cite complexity as a reason for avoiding shared storage. They mention their competitor OpenTSDB as a negative example. Many admins would love to deploy OpenTSDB, but they are put off by the enormous overhead of running a complete Hadoop cluster.

### NOT THE FIRST, BUT THE BEST

Prometheus is not the first attempt to apply the time-series database model to network monitoring. Graphite [3] was around long before Prometheus, but its data model is not as mature. Influx DB [4], which is typically combined with a frontend such as Sensu, is even younger than Prometheus, but it addresses a different user group and, according to our tests, doesn't scale as well as Prometheus when faced with large volumes of data. And, then there is OpenTSDB [5], the Open Time Series Database, which fundamentally is very similar to Prometheus but requires external add-on components such as Hadoop. The fact that these external constraints do not apply to Prometheus is something that many admins really appreciate about the product.



**Figure 2:** The hub of the Prometheus system is the Prometheus server, which communicates with all the other components.

Instead, Prometheus relies on the sharding principle. You can configure multiple instances of the Prometheus server service to cover overlapping data areas. Before performing a search, the database determines the shard in which the data in question must reside and it only looks there.

At this level, you can replicate by letting logical pairs of servers collect the data from the same agent on the network. A record is thus available multiple times and still usable in scenarios in which one of the two nodes has failed.

The Prometheus developers are aware that there is a problem with this lack of a shared storage alternative. Right now, they are working on a solution that generates a superordinate instance for a cluster of Prometheus installations; the instance, in turn, picks up the data from the Prometheus shards.

This approach gives users centralized administration. And there are plans for the distant future: In the long term, the intent is for Prometheus to store data in OpenTSDB – and thus leverage its replication capabilities.

### Level DB and DIY

Prometheus handles data management with a combination of Level DB and a SoundCloud invention. The data model relies on Level DB for indexing, but SoundCloud designed its own approach to storing data structures. In the Sound-

Cloud developer blog, the developers explained that, in their opinion, there was no such thing as a storage engine capable of storing time series entries.

The Prometheus server is not just the storage back end that picks up data from exporters in the cluster. It is, at the same time, the port of call for users or other applications that want to read the data. To allow this to happen, the developers have devised their own query language: PromQL (Prometheus Query Language) works similarly to SQL for other databases, although it does have its own syntax (Figure 3).

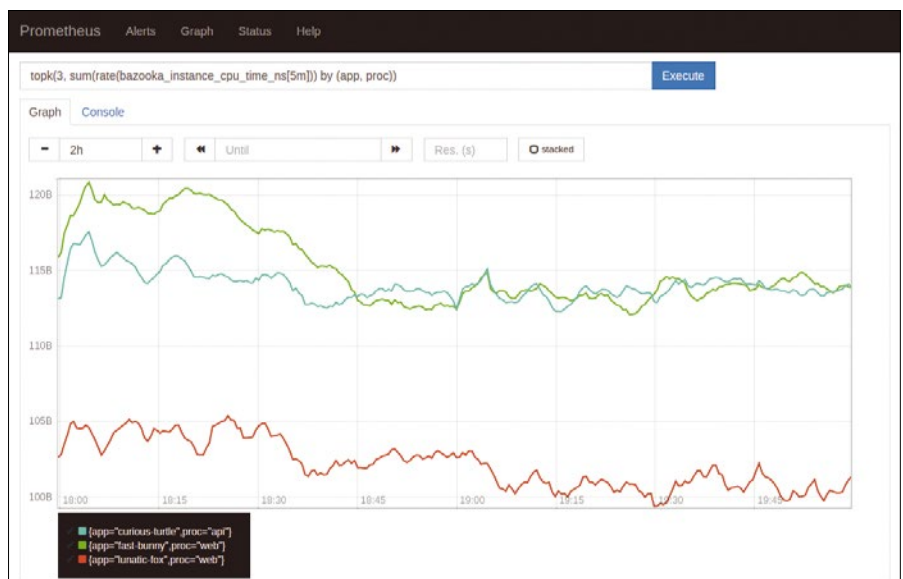
Each Prometheus instance offers a rudimentary web interface from which to submit queries. But, this local interface is not suitable for virtualization data. SoundCloud’s own dashboard is used for virtualization data. However, if you just want some initial information, Prometheus’ own interface will probably do the job.

### Exporters

The neatest monitoring solution is nothing if it is unable to make statements about the state of individual services in the cluster, and this data has to come from *somewhere*. Prometheus refers to the process of collecting data as scraping. Prometheus also has its own kind of service for scraping: the exporters.

Exporters do not simply check process states on the basis of the process ID, which is what simple Nagios plugins do; instead they take a more sophisticated approach. Within Prometheus, the exporters create a data structure for the service and then read the service’s metrics. In practical terms, Prometheus would retrieve the details of how many instances of haproxy were running over a certain period of time and alert if the value of this metric were too low.

Exporters are a curse and a blessing at the same time for administrators. On the one hand, developing an exporter for a specific server is far more complex than developing a minimal Nagios plugin. The exporter for HA Proxy offered by Prometheus itself makes this quite



**Figure 3:** PromQL is a query language for Prometheus, trimmed to the needs of a time series database.

clear: In addition to familiarity with Go, you also need in-depth knowledge of which metrics are actually usable for the target service. On the other hand, you need to be familiar with Prometheus' data structure because you cannot meaningfully create metrics without this knowledge.

On a more positive note, monitoring is far more precise than what simple Nagios plugins provide. A good example is the HA Proxy Exporter: in addition to a plain statement on the matter of traffic passing through the load balancer, the current version in Prometheus gives you detailed statements on availability and performance for individual pages of the front ends configured in HA Proxy. In other words, the effort put into developing exporters does pay dividends.

The good news is that pre-built exporters already exist for many classical services. You can either pick them up from Prometheus or check with the community, where other administrators are writing and publishing exporters. The Prometheus website for exporters [6] has an up-to-date list and links to exporters for MySQL, Bind, or Apache, just to name a few.

A few services, such as Etc, or various services from Google's Kubernetes, even offer native support for Prometheus. When Prometheus learns about the services, it automatically picks up the matching metrics.

## Service Discovery

Prometheus now offers two different approaches to automatic service discov-

ery. You can either employ DNS SRV records or use the cluster registry included with the Consul service management tool [7].

The idea behind this approach is simple: if you need to pick up the same metrics from 200 hosts, it doesn't make sense to force the admin to manually maintain a database of the existing machines when each new cluster node will need to be configured in exactly the same way for Prometheus. Using auto-discovery, Prometheus handles this task for the admin and saves a lot of work by doing so.

The node exporter has a special role – again, the exporter is offered directly by Prometheus. It provides information on each system, for example, the history of the CPU or RAM load.

## Push Gateway

All services cannot be meaningfully run with exporters. This is particularly problematic for services that generate data on the fly and need to offload the data immediately. One classic example is Cron, which handles a specific task and generates output while doing so, which Prometheus then needs to process. Prometheus has the push gateway to handle this kind of situation.

The gateway runs as a separate service with a RESTful interface and waits for the required values to arrive. If you want to deliver the Cron job output to the push gateway, all you need to do is append a pipe and a Curl call to the Cron job. The REST interface on the push gateway accesses the target host for Curl.

Of course, these gateways will scale horizontally; in other words you can run multiple gateways in parallel in the same installation. The Prometheus servers pick up the stored data from the gateways in exactly the same way they do with normal exporters. But, there is one slight drawback: The data output by the Cron job needs to exist in a format that the push gateway actually understands. In other words, you may need to tune the Cron output.

## Alert Manager

If something is going wrong in the cluster, administrators would naturally expect the monitoring system to tell them. Prometheus has two possible approaches. Prometheus itself offers a basic alerting function: if the value of a metric is outside of defined boundaries over a defined period of time, Prometheus sends a message to a previously defined address.

However, the developers expressly point out that this function is only basic alerting, and the number of tweaks required for modifying the alerting style to meet your local needs limits flexibility.

But Prometheus offers a second approach: instead of handing alerting itself, Prometheus can forward alerts directly to the Alert Manager. The Alert Manager is a separate component in the Prometheus universe (Figure 4), and it is probably one of the biggest construction sites in Prometheus right now.

When this issue went to press, the Alert Manager GitHub page [8] featured

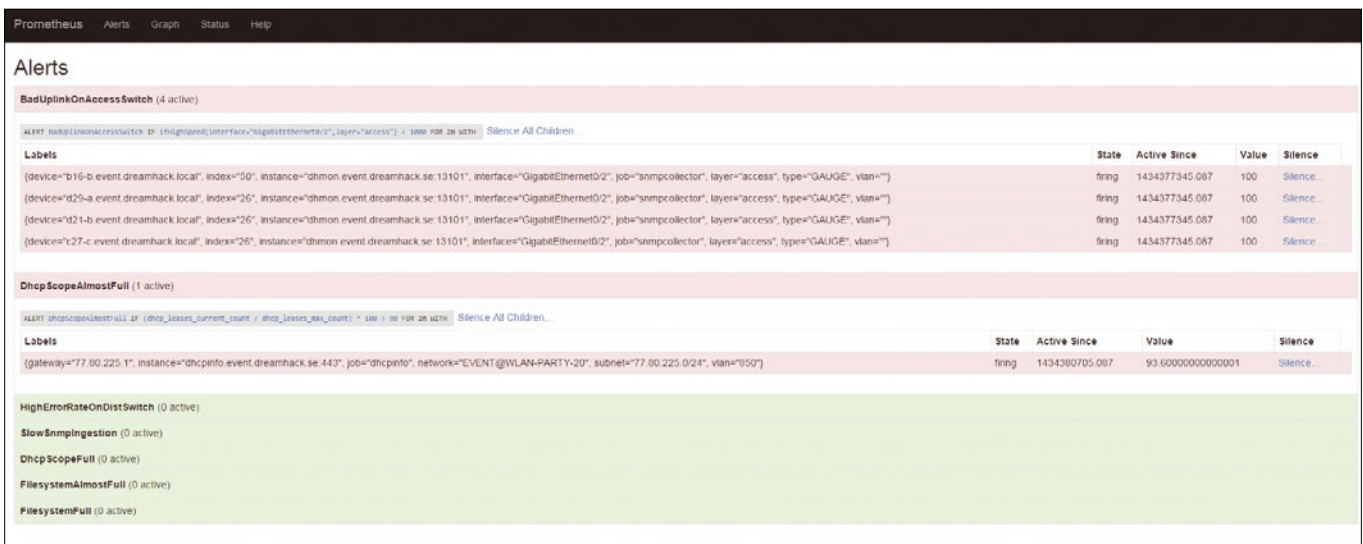


Figure 4: With the right choice of metrics and matching parameters, you can reach a “good” or “bad” decision for any situation.



Figure 5: Prom Dash is a helper that lets administrators create graphs based on Prometheus data.

a statement to the effect that the version in the master branch is a totally new development that is no longer compatible with the current stable version. If you want to take a look at, or test, Prometheus, you will still want to go for this version: the legacy version will disappear in the foreseeable future.

As an interface to Prometheus, the Alert Manager has its own API. When the Alert Manager receives an alert from Prometheus, it can process that alert in a variety of ways: the manager can mail a warning or forward the alert to the Pager Duty API.

The pager, in turn, supports various modules. Out of the box are Pager Duty modules for Slack, Hipchat, and various other chat protocols. But you could hitch up Pager Duty to a text service with a matching SMS module.

### PromDash

PromDash (Figure 5) is a powerful tool for rendering the metrics stored in Prometheus in a meaningful way. The tool's developers describe PromDash as a dashboard builder, thus setting it apart from standard dashboards.

PromDash does not try to give users prebuilt templates for charts but instead gives administrators a tool for creating these templates themselves. PromDash can thus rightly be described as a graphi-

cal helper that facilitates the use of the PromQL language.

After setting up the solution, you need to add the servers that will act as data sources in PromDash, then sit back and enjoy the ride. One of PromDash's features is the ability to annotate graphs or integrate them directly into other pages. For example, you could easily implement a status dashboard for the entire platform that would give you a centralized overview of all the systems you need to monitor.

The solution for configuring PromDash itself is also very smart: The PromDash module stores the configured dashboards in a MySQL database in the background. This design means the configuration and program are separate, so you can run any number of PromDash instances on the same host without you needing to manually synchronize the configurations.

### Conclusions

Prometheus and other time-series databases are ringing in a new era in monitoring. Although static constructs like Nagios or Zabbix are undeniably useful in small to medium-sized environments, they are not as functional as their new competitors when faced with massively scalable networks or cloud configurations.

Although Prometheus is still a very young product, its feature scope is already quite impressive. The data model devised by SoundCloud gives the application the ability to manage huge volumes of data in an efficient way, while at the same time supporting faster searching. The Alert Manager is also at an early stage of development, but it is already very flexible and promising.

The ability to use PromDash to convert data into at-

tractive graphics, which are then presented on custom dashboards, is very useful in the daily grind. Smart functions like automatic service detection on the network, or the ability to record the details of various metrics, round off the offering. Our lab tests indicate that you are likely to hear much more about Prometheus in the near future. ■■■

### INFO

- [1] PNP4Nagios: <http://sourceforge.net/projects/pnp4nagios>
- [2] Prometheus: <http://prometheus.io>
- [3] Graphite: <http://graphite.wikidot.com>
- [4] Influx DB: <https://influxdata.com>
- [5] OpenTSDB: <http://opentsdb.net>
- [6] Exporters: <http://prometheus.io/docs/instrumenting/exporters/>
- [7] Consul: <http://consul.io>
- [8] Alert Manager: <https://github.com/prometheus/alertmanager>

### AUTHOR

**Martin Gerhard Loschwitz** works as a cloud architect at SysEleven in Berlin. He is also an official member of the Debian project and has been a Debian developer for more than 12 years.

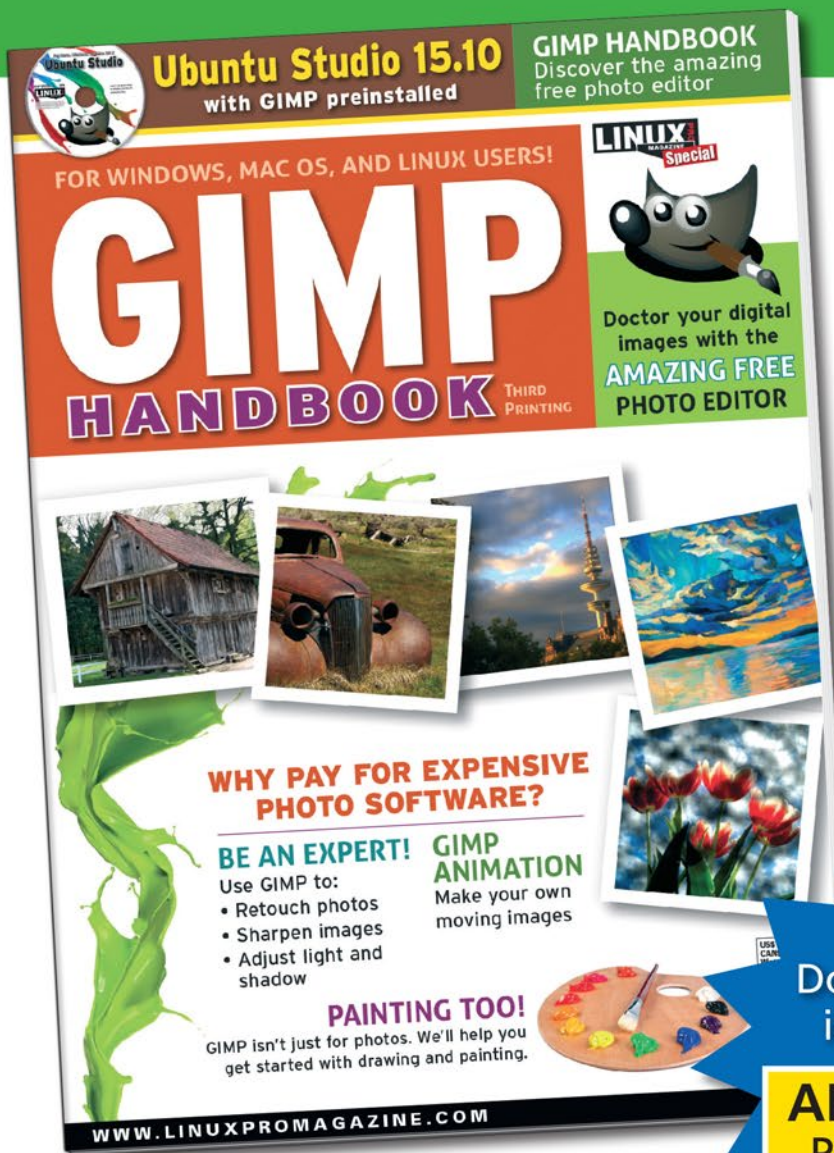
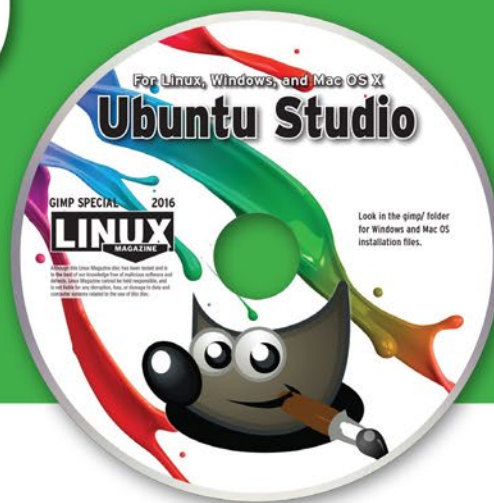




Shop the Shop

shop.linuxnewmedia.com

# GIMP HANDBOOK



**SURE YOU  
KNOW LINUX...**  
but do you know **GIMP?**

- Fix your digital photos
- Create animations
- Build posters, signs, and logos

**Order now and become an expert in one of the most important and practical open source tools!**

**GIMP**  
Doctor your digital images with the  
**AMAZING FREE PHOTO EDITOR!**

Order online:  
[shop.linuxnewmedia.com/specials](http://shop.linuxnewmedia.com/specials)



**FOR WINDOWS, MAC OS, AND LINUX USERS!**



Wayland display server protocol

# Changing of the Guard

The X11 graphics protocol is showing some serious signs of age, but Wayland is poised to come to the rescue. *By Ferdinand Thommes*

The X Window System has provided a framework for desktop graphics in Unix and Linux for more than 30 years. X has gone through many phases since it first ap-

peared in 1984, but it stabilized long ago. Version 11 of the X protocol (referred to as X11) has been around since 1987, and it is a fundamental part of the Linux landscape; however, computers

have changed a lot since 1987, and many experts believe the X Window System needs to be replaced. The Wayland display server protocol, developed since 2008 under the direction

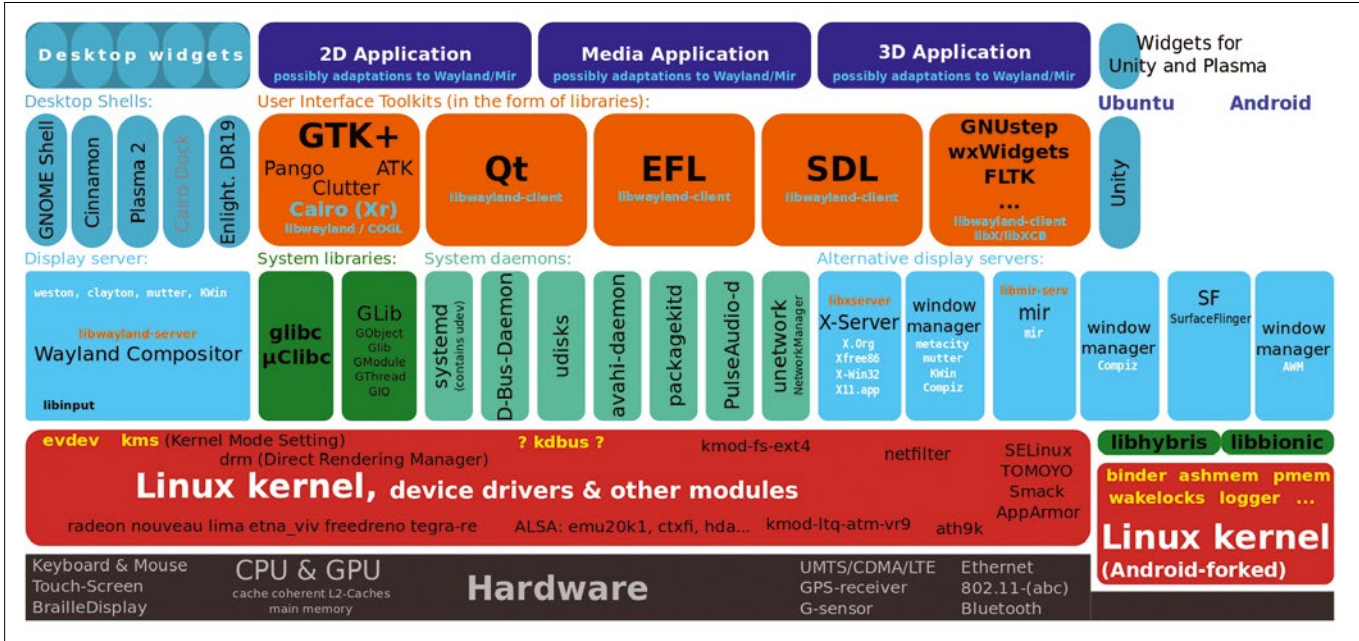


Figure 1: The display server between the kernel and hardware and the applications [4]. (CC BY-SA 3.0 [5])

Lead Image © Maddalena Delli, 123RF.com

of Intel employee Kristian Høgsberg, is primed to take over for X11. Wayland could eventually solve a load of problems that developers have to contend with when integrating Linux applications with a graphic desktop, but is Wayland equal to the task? This article looks at the state of Wayland integration.

### The Turning Point

X11 is a patchwork of code difficult to maintain and almost impossible to expand. Serious mistakes often emerge that have remained hidden in the code for years. For example, a security vulnerability in the font server dating from 1991 wasn't discovered until 2014 [1]. Rigorously practiced backward compatibility is also part of the reason X11 is not considered secure. Several of the core components carried over from the early days are no longer of any real use, but they have to be available by default.

The focus at the beginning of development was also completely different from what is expected of a modern graphics stack today. Modern graphics libraries now draw circles and rectangles or move windows, without having to resort to an X Server. The modern environment also supports shared libraries, which means graphics-capable applications don't have to carry around a ballast of graphics functions. Modern clients just expect the display server to allocate an area where they can write and display the content, which is what Wayland provides.

### Detours and Wrong Turns

The person perhaps most responsible for Wayland's position as an heir to X11 now no longer supports the Wayland protocol. In 2010, Mark Shuttleworth announced that Ubuntu would run with Wayland instead of X11 from version 12.04 onward [2]. This announcement gave the Wayland project, which originally was meant to prove only that X11 could be rebuilt without too much effort, strong standing as a possible heir apparent for X11. Shuttleworth later decided to develop his own, new Mir display server rather than support Wayland.

Most Linux distributions will eventually switch to Wayland, although X11 will be around for a few more years. The Wayland developers created XWayland, a slightly modified X server that serves as a compatibility layer, to eliminate

problems during the transition phase for applications that still require X11. The first major distribution that natively uses Wayland by default will probably be Fedora, which was considering making Wayland the default system in Fedora 24, although a recent blog post by developer Matthias Clasen [3] said, despite all the work put in toward that goal, Wayland was not quite ready for production distros.

### Wayland/X11 Differences

Wayland differs from X11 conceptually and functionally. In the Wayland environment, the compositor is the display server. In X11, on the other hand, the compositor is an external component that requires an additional processing step. Wayland is just the protocol, and Weston is the reference implementation of a compositor (Figure 1). The differences between Wayland, X11, and XWayland become clearer if you look at how the display servers pro-

cess an event such as a mouse click (Figures 2-4).

### Long and Short Routes

In X11, the kernel transmits a mouse click via the evdev driver [8]. The X server determines which window the event is intended for and sends it to the client responsible for the window. One problem with this approach is that the position of the window is controlled by the compositor and not by X itself. Because a mouse click can change elements in the window or open a new

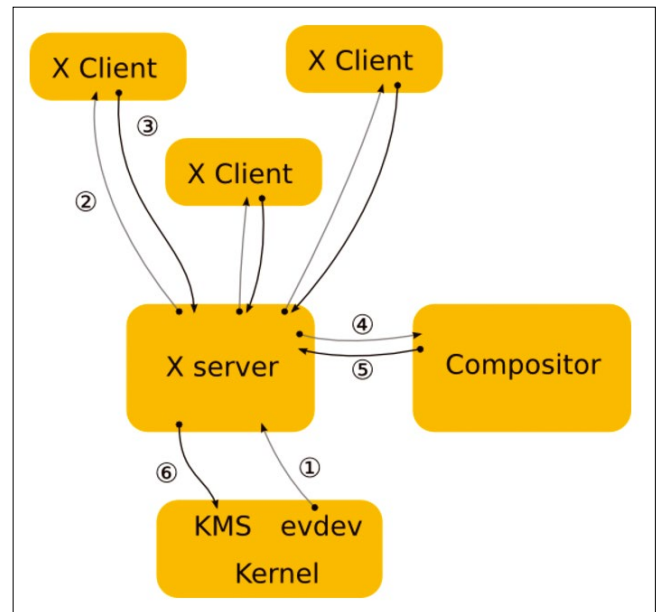


Figure 2: The X server processing an event from an input device [6].

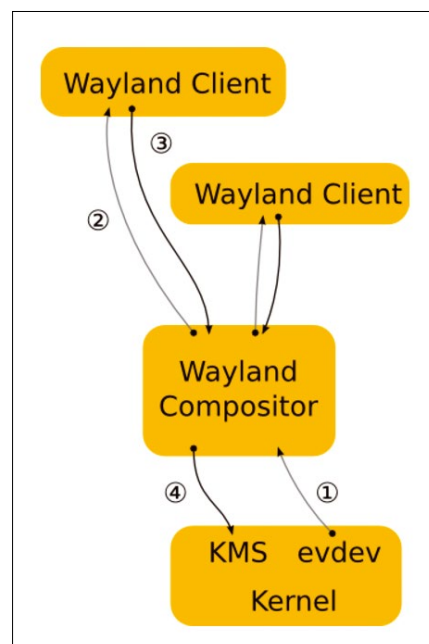


Figure 3: Wayland processing an event from an input device [6].

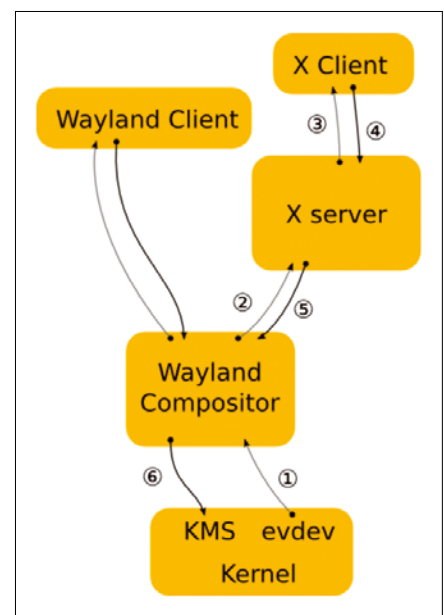


Figure 4: XWayland processing an event from an input device [7]. XWayland runs as a modified X Server in Wayland.

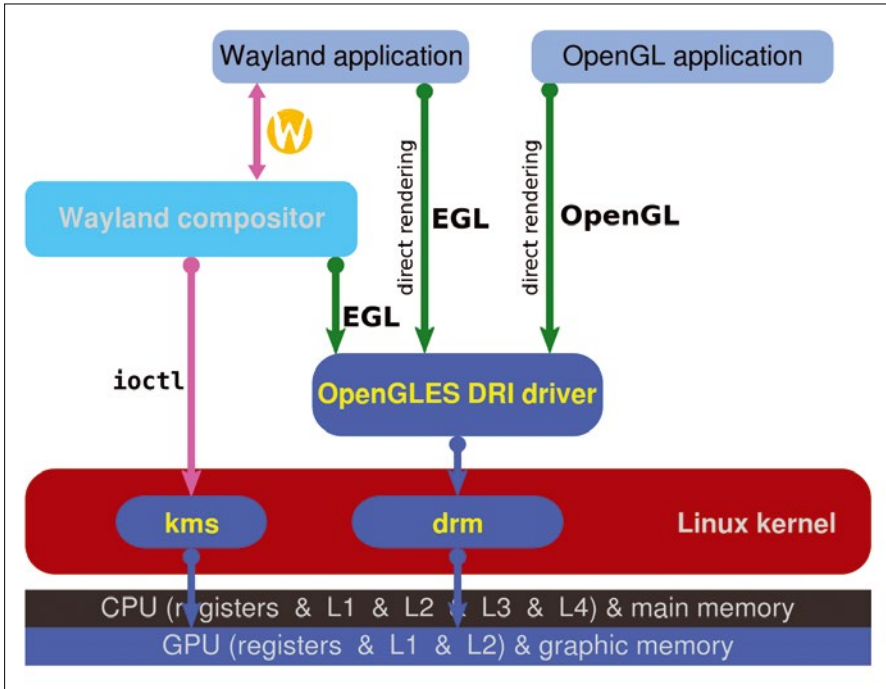


Figure 5: Short routes in Wayland [12]. (CC BY-SA 3.0 [5])

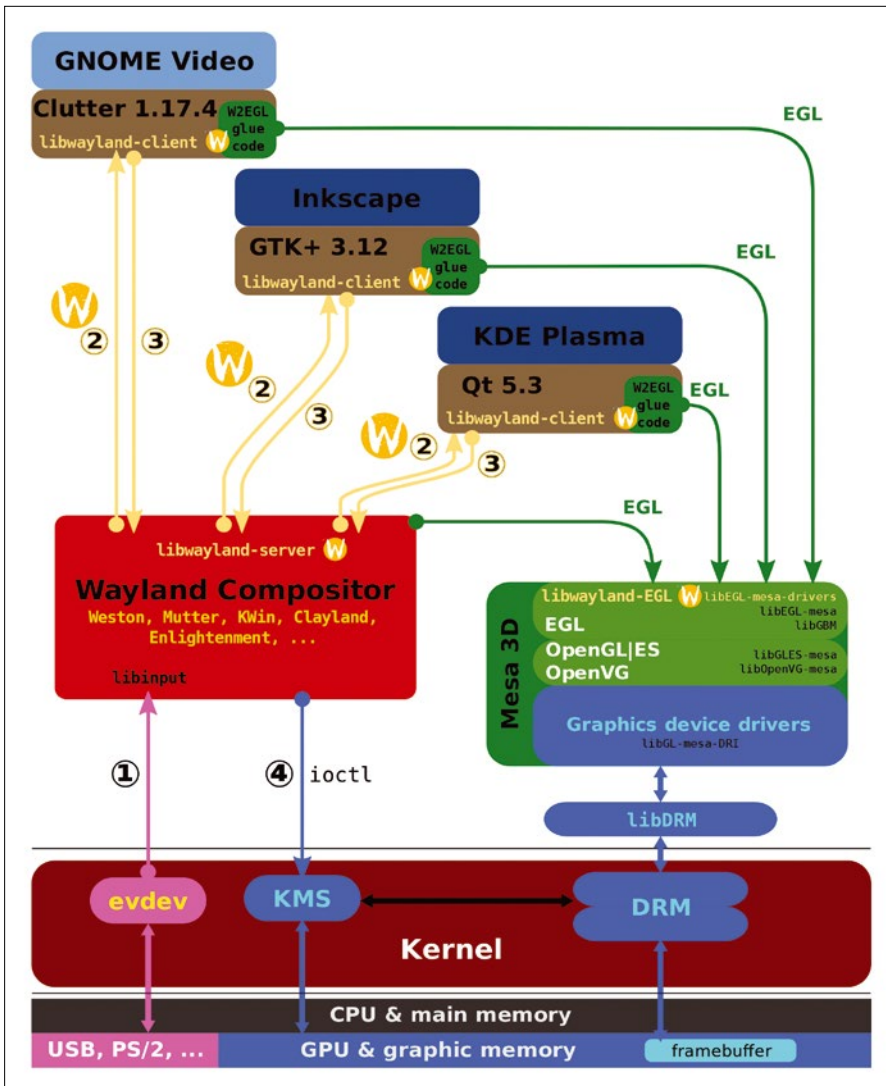


Figure 6: The Wayland compositor handles various clients [13]. (CC BY-SA 3.0 [5])

window, the client sends a request to render to the X server. The X server then forwards this request to the graphics driver, which delegates the rendering to the hardware. The X server then calculates the geometry and position of the area to be rendered and sends a report called a *damage event* to the compositor, which provides the information that something has changed in the window and that an area of the display must be redrawn. The commands for this redraw again have to go through the X server.

The X server carries around a lot of baggage, but it doesn't actually have much functionality that is still useful today. The X server is the middleman, and these days it just generates unnecessary steps between the application and the compositor and between the compositor and the graphics hardware.

In Wayland, the compositor is also the display server (Figure 5). The kernel hands over events directly to the compositor. Wayland allows the compositor to transmit events directly to the clients and then lets clients send back *damage events* directly to the compositor. For example, the compositor looks for which window the event is intended. Thanks to its scene graph [9], the compositor knows exactly which changes have already been made to the window (Figure 6).

The event is forwarded directly to the clients, and the clients calculate the rendering themselves and send feedback to the compositor, stating that the window has been updated. The compositor redraws the window and sends a system call (*ioctl*) to the Kernel Mode Setting (KMS) [10] to request a pageflip. The clients can't take over the rendering themselves because of the Direct Rendering Infrastructure (DRI) [11], which allows common video buffer storage for the client and server. This means that the client can link to a rendering library (e.g., OpenGL, Vulkan, or the rendering engines Qt or Gtk+), which then writes directly in the shared buffer.

The only supposed disadvantage of Wayland compared with X11 is that X11 is network capable, whereas Wayland is not. However, with the Remote Desktop Protocol (RDP) or Virtual Network Computing (VNC), users today have options for sending desktops or individual GUI application via the network. An RDP or

VNC server based directly on Wayland is also conceivable. For security reasons, however, much more focus is placed on one of X11's weaknesses: In the past, the X server corresponded directly with the hardware, so it traditionally ran with root privileges. Thanks to KMS, today the X server theoretically can run without root privileges, but this option is hardly ever used. Wayland, on the other hand, doesn't require root privileges because it communicates with the hardware via the kernel.

## Wayland's State of Development

Wayland and its reference compositor Weston are now up to version 1.9. The protocol itself is considered quite mature

and no longer receives new functions, so developers of compositors for KDE, Gnome, Enlightenment, Sailfish OS, Tizen, and others have time to finish and test their Wayland implementations. The Weston reference compositor [14], on the other hand, provides new experimental technologies and interfaces.

New features include atomic kernel mode switching [15], which promises a completely flicker-free view change. A library that can manage the input devices of various compositors and that aims to replace drivers such as *evdev*, *synaptics*, or *wacom* was already integrated with *libinput* in Weston 1.5 (Figure 7). Fedora has used *libinput* for managing input and output devices since version 22.

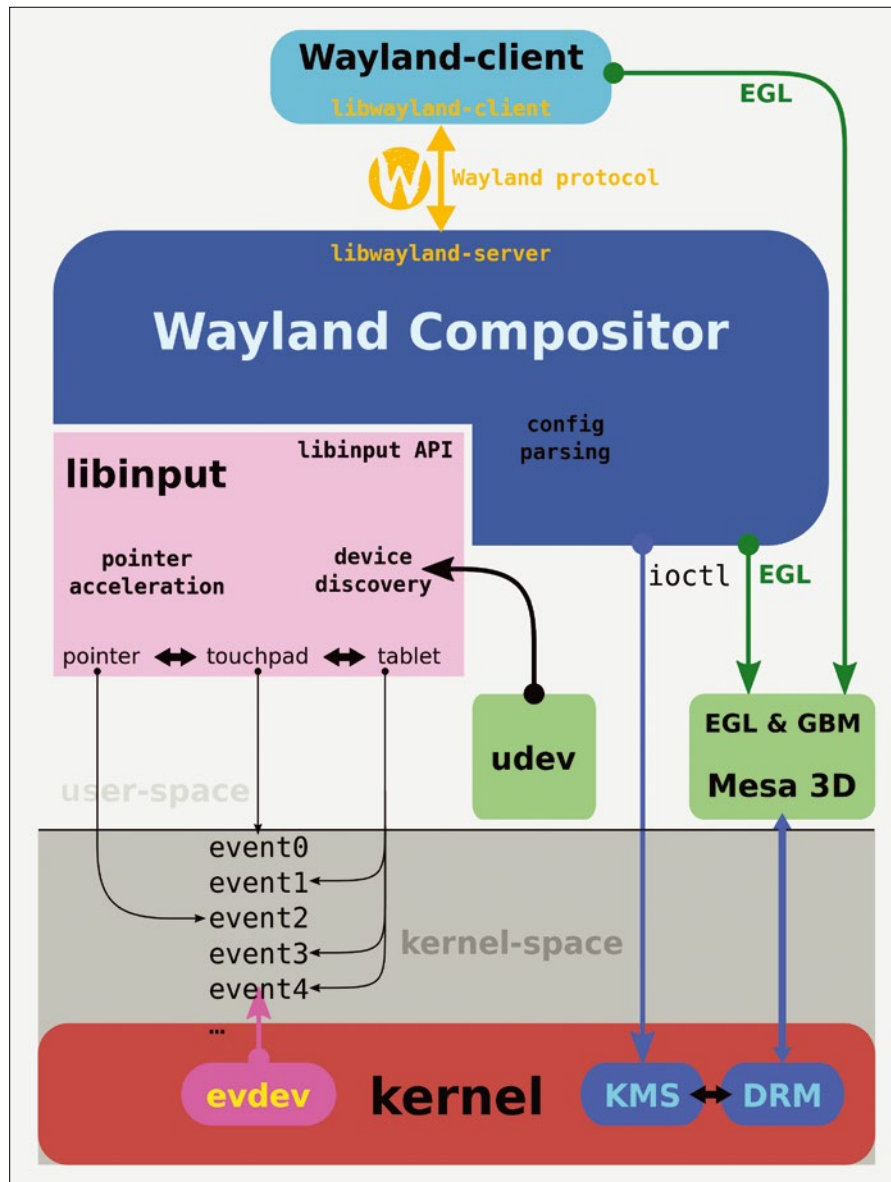


Figure 7: Libinput looks after input and output events and the devices behind them [16]. (CC-BY-SA-3.0 [5])

# GOT CLUSTER?

**ADMIN HPC** **HPC Up Close**

- Monitoring HPC Systems
- Watson in the Cloud
- SC14 Call for Workshop Proposals
- Data Visualization Tools

**Highlights**

**Monitoring HPC Systems: What Should You Monitor?**  
In rapidly growing HPC installations, you need to understand what is happening within the system to make improvements or simply to justify the...

**Watson in the Cloud**  
IBM launches three new services for its well-trained, Jeopardy-busting behemoth.

**SC14 Calls for Workshop Proposals**  
Entries should provide "interaction and in-depth discussions of stimulating topics of interest to the HPC community."

**Data Visualization Tools**  
Logfiles contain extremely useful traces of everyday events – if only they weren't so boring. Fortunately, special tools help you generate and...

**Most Read**

**Tuning I/O Patterns in Fortran 90**  
In the second article of this three-part series, we look at simple write examples in Fortran 90 and track the output with trace to see how...

**Build a Network Attached Storage System with FreeNAS**  
We provide an overview of FreeNAS and ZFS, its main filesystem, then show you how to set up and maintain a FreeNAS installation.

**XSEDE Releases Toolset**  
The new tools will improve compatibility of local campus clusters with the XSEDE supercomputer network.

**Further Reading**

- Fixing Disks with Parted Magic
- Exploring the Xeon Phi
- Tuning I/O Patterns in Python
- Blue Waters Open for Business

**Partner Link**  
Intelligent High Availability DAS for Windows

**Raspberry Pi Geek Offer**

Subscribe to Raspberry Pi Geek and receive either Issue 1 or Raspberry Pi Geek for FREE!  
Use code HOLSRRPG2013 at checkout.  
Click here for full details. (Offer expires 2014.)

Tune in to the HPC Update newsletter for news, views, and real-world technical articles on high-performance computing.

[hpc.admin-magazine.com/Newsletter](http://hpc.admin-magazine.com/Newsletter)

Fedora has offered experimental support for Wayland for more than a year. The Fedora Rawhide distribution takes that support one step further and, as of recently, supports Wayland by default. Users might need to edit the `/etc/gdm/custom.conf` file to receive a conventional X.Org session. Because not everyone wants to test Rawhide, the Fedora developers promise to implement Wayland bug fixes quickly so that users remain fairly close to the current state of development when updating their system.

### The Test Course

To test Wayland using a Live image of Fedora 23, the user needs to create a password after starting the Live medium to obtain the Wayland option. To do so, just click the down arrow in the top right corner of the desktop, select *Live System User | Account Settings*, and enter a password. Fedora places value on passwords that are as safe as possible and therefore makes you create secure passwords.

Once you've created a secure password, go to *Live System User | Log Out*, and you'll be taken to the login manager, where you can select the Wayland session by clicking the cogwheel to the left of *Sign In* (Figure 8). The next thing to do is import the latest updates. The quickest way to import updates is in a terminal using the

```
sudo dnf upgrade
```

command. If you want to check whether a Gnome Wayland session is running, you can use the

```
ls -l $XDG_RUNTIME_DIR/wayland-0
```



**Figure 8:** Wayland session in the Fedora 23 Gnome GDM login manager. In Fedora 24, the Wayland session will be *GNOME*.

command. A socket will be displayed in `/run/user/1000` if this command is successful.

### Limitations

Some limitations still exist in a Wayland test session. On the one hand, you currently still need to use free graphics drivers because the proprietary drivers adapted to Wayland from AMD and Nvidia still aren't ready.

If you want to test Wayland in a virtual environment, VirtualBox isn't currently an option because the virtual graphics driver isn't yet adapted to Wayland. You'd have to use KVM or VMware Player instead.

Anyone who also habitually copies and pastes using the mouse will be disappointed because this feature isn't yet supported. It can also be a bit annoying that the mouse pointer sometimes needs a few seconds to respond, and you could also experience problems when operating multiple monitors.

All Gnome core packages have already been migrated to Wayland. Other applications such as Firefox are still using the XWayland interlayer [17]. I wasn't able to find any visible differences between

native Wayland support and XWayland. The list of apps that still need XWayland (using the `x11clients` command) is not always complete.

### Conclusions

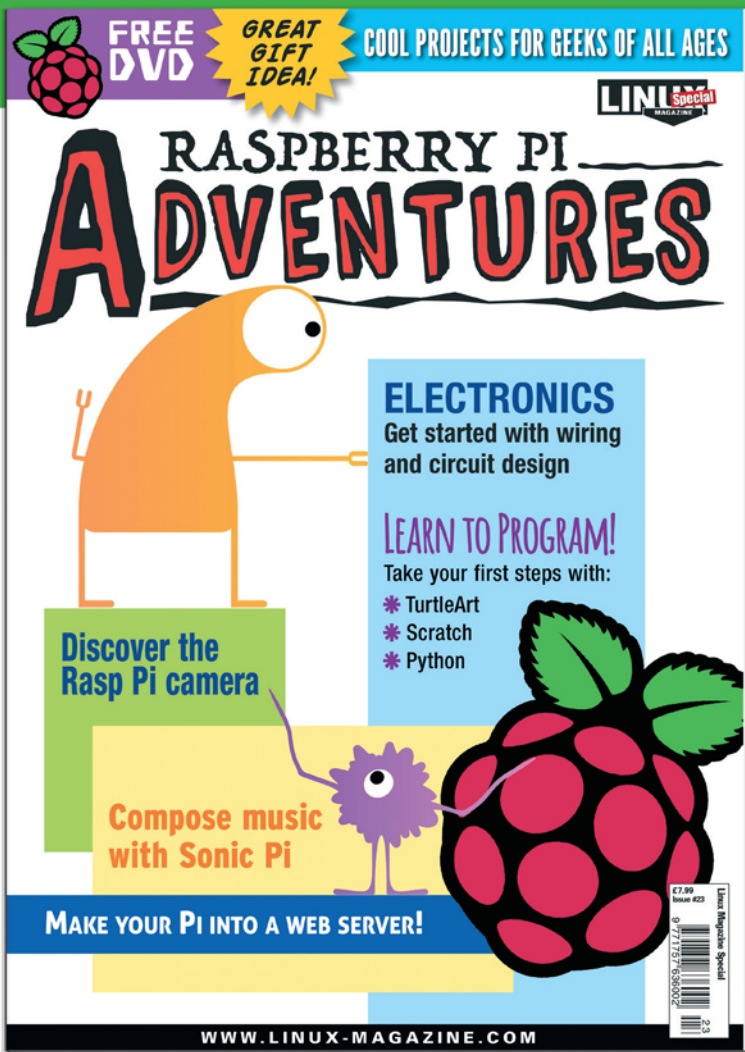
The appearance of windows and menus is very clean and flicker-free compared with tests conducted six months ago. A Wayland session runs stably in normal operation, but you can also still make it crash in rare cases. This is where X11 usually steps in, but sometimes you are sent back to the login manager. If you can cope with the restrictions described here, you really have no reason not to use Wayland. And, you can report any errors to the developers at Gnome Bugzilla [18].

After seven years of development, Wayland is now in the final sprint to the finish line. The intention of making Wayland a default in Fedora 24 was perhaps a bold plan, given the lack of proprietary drivers, gaps in functionality, and many errors still to chase down. More should be known later in 2016. In any case, X11 will still be around for years – even when Wayland becomes the default in future distributions. ■■■

### INFO

- [1] Errors in X11: <http://lists.x.org/archives/xorg-announce/2014-May/002431.html>
- [2] Shuttleworth on Wayland: <http://www.markshuttleworth.com/archives/551>
- [3] Matthias Clasen blog: <https://blogs.gnome.org/mclasen/2016/03/04/why-wayland-anyway/>
- [4] Display servers and UI toolkits, Shmuel Csaba Otto Traian: [https://en.wikipedia.org/wiki/Wikipedia:WikiProject\\_Desktop\\_Linux#/media/File:Free\\_and\\_open-source-software\\_display\\_servers\\_and\\_UI\\_toolkits.svg](https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Desktop_Linux#/media/File:Free_and_open-source-software_display_servers_and_UI_toolkits.svg)
- [5] Attribution-ShareAlike 3.0 Unported: <http://creativecommons.org/licenses/by-sa/3.0/>
- [6] X/Wayland architectures: <https://wayland.freedesktop.org/architecture.html>
- [7] XWayland: <https://wayland.freedesktop.org/xserver.html>
- [8] evdev <https://wiki.ubuntuusers.de/evdev>
- [9] Scene graph: [https://en.wikipedia.org/wiki/Scene\\_graph](https://en.wikipedia.org/wiki/Scene_graph)
- [10] KMS: [https://en.wikipedia.org/wiki/Mode\\_setting](https://en.wikipedia.org/wiki/Mode_setting)
- [11] DRI: [https://en.wikipedia.org/wiki/Direct\\_Rendering\\_Infrastructure](https://en.wikipedia.org/wiki/Direct_Rendering_Infrastructure)
- [12] Linux graphics drivers, Shmuel Csaba Otto Traian: [https://commons.wikimedia.org/wiki/File:Linux\\_graphics\\_drivers\\_DRI\\_Wayland.svg](https://commons.wikimedia.org/wiki/File:Linux_graphics_drivers_DRI_Wayland.svg)
- [13] Wayland display server protocol: [https://commons.wikimedia.org/wiki/File:Wayland\\_display\\_server\\_protocol.svg](https://commons.wikimedia.org/wiki/File:Wayland_display_server_protocol.svg)
- [14] Weston: [https://en.wikipedia.org/wiki/Wayland\\_\(display\\_server\\_protocol\)#Weston](https://en.wikipedia.org/wiki/Wayland_(display_server_protocol)#Weston)
- [15] Atomic KMS: [https://en.wikipedia.org/wiki/Mode\\_setting#Atomic\\_Mode\\_Setting](https://en.wikipedia.org/wiki/Mode_setting#Atomic_Mode_Setting)
- [16] Libinput for Wayland compositors, ScotXW: [https://commons.wikimedia.org/wiki/File:Libinput\\_for\\_Wayland\\_compositors.svg](https://commons.wikimedia.org/wiki/File:Libinput_for_Wayland_compositors.svg)
- [17] XWayland: [http://wayland.freedesktop.org/xserver.html#heading\\_toc\\_j\\_0](http://wayland.freedesktop.org/xserver.html#heading_toc_j_0)
- [18] Bug reports: <https://bugzilla.gnome.org/>

# RASPBERRY PI ADVENTURES



**COOL  
PROJECTS  
FOR GEEKS  
OF ALL  
AGES**



## RASPBERRY PI ADVENTURES

is a one-volume special edition magazine for curious Raspberry Pi beginners. This easy, hands-on guide starts with an introduction to computers and offers a series of special hands-on projects illustrating many of the most popular uses for the Raspberry Pi.



**ORDER YOUR VERY OWN ISSUE!**



ORDER ONLINE:

[shop.linuxnewmedia.com/se23](http://shop.linuxnewmedia.com/se23)



Discover how to use and probe a SQLite database

# Data Diver

Several databases likely reside on your desktop and smartphone, and it is easy to manage the data in these files or to create similar databases yourself. *By Marco Fioretti*

**S**QLite [1] (pronounced sequelite or S-Q-L-lite) is a public domain, embedded, relational database engine that runs on everything from smartphones to mainframes. If you use Linux or any other modern operating system, chances are good that you are already using at least one SQLite database. That alone is reason enough to learn the basics of SQLite, and it is in your interest to know not just how to back up that data, but how to generate, process, and analyze it in ways that would not be possible with other applications.

Moreover, you can install SQLite with almost zero configuration or manual work and then run it without root privileges. Finally, despite its simplicity, SQLite can handle even huge quantities of data, which means it may even help you on the job someday.

## What You Will Learn

In this tutorial, I explain what SQLite is, discuss how it works, and look at a few practical ways in which to use it. Although I show only a few quick examples of actual database queries, you can find plenty of those at the SQLite website or from online cheat sheets. Instead, I

focus on basic SQLite management, concepts, components, where you can find SQLite data on your desktop or smartphone, and why you might want to process it. Basic knowledge of the command line and shell scripts is all you need to take advantage of this tutorial.

Although I provide pointers to SQLite graphical interfaces, I mostly focus on what you can do at the SQLite command prompt, partly because the prompt is the only SQLite interface that is reliably available on every platform. The main reason, though, is that you can automate what you type a lot more easily than what you click.

## Relational Databases

Relational databases are highly

structured archives of data stored in one or more files in a low-level binary format. The data is stored in tables of rows and columns (Figure 1). Each column contains data fields of the same type (e.g., title, author, publication date, or ISBN number in a library database), and

	model	lens	aperture	focalLength	focalLength35	exp
	Filter	Filter	Filter	Filter	Filter	Filter
801	COOLPIX S...	NULL	3.1	6.3	35.0	0.03
802	COOLPIX S...	NULL	3.1	6.3	35.0	0.03
803	COOLPIX S...	NULL	3.1	6.3	35.0	0.03
804	COOLPIX S...	NULL	3.5	7.4	41.0	0.03
805	COOLPIX S...	NULL	3.5	7.4	41.0	0.03
806	DMC-FX35	NULL	2.8	4.4	25.0	0.00
807	DMC-FX35	NULL	2.8	4.4	25.0	0.00
808	DMC-FX35	NULL	2.8	4.4	25.0	0.00
809	DMC-FX35	NULL	5.6	17.6	100.0	0.00
810	DMC-FX35	NULL	2.8	4.4	25.0	1.0

**Figure 1:** Like all other relational databases, SQLite organizes its data in tables (shown in the SQLite DB Browser). Each column is devoted to one type of data, and each row holds all data related to a specific object (here, a digital photograph).

Lead Image © Leo Blanchette, 123RF.com



each row (or record) groups all the data referring to the same entity (e.g., all data about a specific book).

The low-level textual commands to create, alter, or filter tables and records are written in a structured query language (SQL), which varies slightly from database engine to database engine. Although graphical interfaces hide SQL commands from the user, they are still used to talk with the database engine.

### Features and Use Cases

SQLite is easy to manage and run because its database structure is very simple and it does not use the client-server model. Instead of requiring a server that always waits for requests from clients, SQLite is an ordinary program that launches only when needed. The core code fits everywhere because, even though it can manage terabytes of data, it has a very small memory footprint and no external dependencies.

Although executable SQLite programs are specific to an operating system, each SQLite database is a single, 100 percent cross-platform file that always has the same format. You can back it up as you do any other file or put it on a USB key for access from any computer, without the need for administrative privileges. If you want a read-only database, you just

remove write permission from the corresponding files.

All of these features taken together mean that the SQLite database format is explicitly promoted and ready for use, both as a searchable archival format (but more sophisticated than TAR or ZIP files) and as a general-purpose application file format (much like a DOC, ODT, or PDF file). Rather than developing a unique file format and all the code to handle it, developers can include the SQLite library in their source and use it to store everything their program needs in a SQLite database, from configuration to user data.

Now that I have explained what SQLite is, I will look at how it uses the SQL language. The main concepts and tools you need to know, which are much easier to understand than their names suggest, are pragmas, metacommands, data types, commands, operators, and functions.

### Pragmas and Quotation Marks

A PRAGMA statement is a SQL extension that changes the general high-level behavior of the SQLite library or checks the status and properties of a data structure (Figure 2). One example of pragma is

```
PRAGMA auto_vacuum = FULL;
```

which tells SQLite to enable automatic execution of the VACUUM command, which then rebuilds and compacts database files to save disk space.

SQLite uses single quotes to enclose literal strings and double quotes for keywords or column or table names. Assuming your database has a column named *platform*, you may write:

```
platform='linux'
platform="linux"
```

But the first statement means “[do something] if the field called *platform* has the value *linux*”; the second statement says “[do something] if the field called *platform* has the same value as the field called *linux*.”

### Metacommands

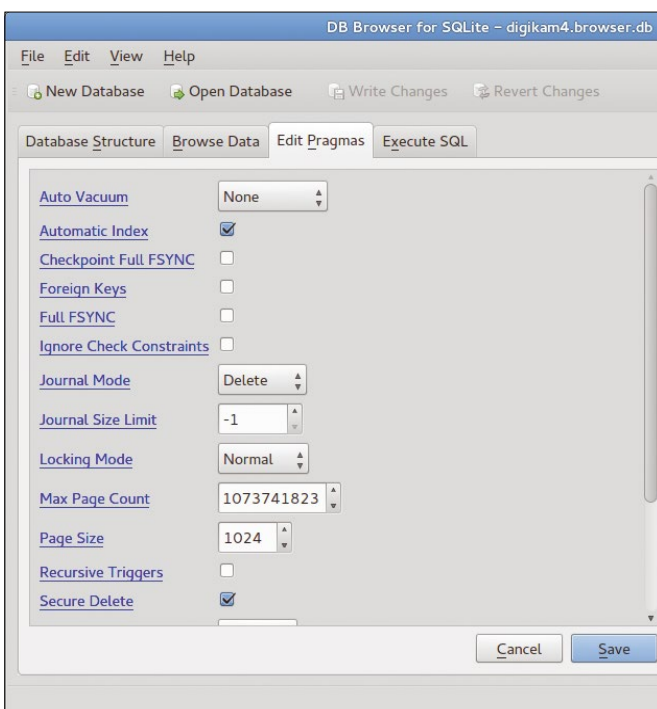
SQLite metacommands, called dot commands because they begin with a dot, don’t add to or fetch data from SQLite tables. They provide high-level information, shortcuts to frequently used metaqueries, or change the way query results are presented, as shown in Table 1.

The same table also shows how to add comments to your SQLite code: you can use two dashes (--), which means the parser ignores everything that follows to the end of the line. For multiline comments, you enclose text between the strings /\* and \*/, as you would in C programs.

### Data Types

In a SQLite table, each column can store data of different types, but each column has its own preferred storage class (Figure 3). This class is assigned to each column when you create a table, such as:

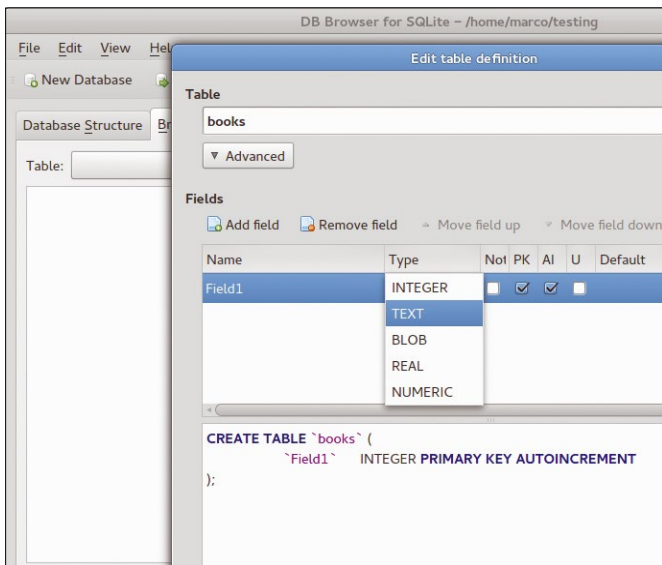
```
CREATE TABLE books
(bookid INTEGER PRIMARY KEY,
```



**Figure 2:** SQLite pragmas are high-level directives that change the behavior of SQLite. In DB Browser, you can easily check their values and change them as needed.

**TABLE 1:** Some SQLite Metacommands

Dot Command	Action
.databases	List all active databases
.help	List all metacommands
.import FILE TABLE	Import data from FILE into TABLE
.output FILE	Send output to FILE
.output stdout	Send output to the screen
.read FILE	Execute the SQL commands in FILE
.schema	List the complete structure of the current databases
.schema TABLE	Only show structure of the requested table
.separator STRING	Change column separator
.show	Show current settings
.tables	List all the database tables



**Figure 3:** Each column inside a SQLite table has a data type, which is specified when you create the table. You can do this directly by writing SQL code (bottom), or you can use DB Browser to generate the same code for you (top).

```
title TEXT NOT NULL,
author TEXT NOT NULL,
price REAL NOT NULL);
```

The *books* table has four columns: the first, which is also the index, is an integer; the last column is a floating-point number, and the others are text strings.

Dates deserve an extra word because they can be stored in three interchangeable formats:

- REAL – the number of days since noon, Greenwich mean time, on November 24, 4714BC
- INTEGER – the number of seconds since 1970-01-01 00:00:00 UTC
- TEXT – a string (e.g., "YYYY-MM-DD HH:MM:SS.SSS")

Speaking of text strings, beware! Although the SQLite core library can store strings with any encoding, by default it correctly compares and orders only ASCII characters in a case-insensitive way. This design choice makes the SQLite core code as small and fast as possible. Support for non-ASCII characters is delegated to external libraries, which in practice may be linked already into the SQLite software packaged for your Linux distribution. SQLite can also store BLOBs, which are raw sequences of bytes (e.g., images).

### Operators and Functions

In all variants of SQL, operators are tools that let you compare data among records

### USING SQLITE AT THE COMMAND LINE

To get a SQLite command prompt, enter `sqlite3 <dbn>`, where `<dbn>` is the name of the file with the database you want to access. If you don't specify a name, SQLite creates a temporary database; if the name you enter doesn't exist, SQLite creates it. Ctrl+D stops the program, and Ctrl+C stops a SQL statement.

or process them in several ways (e.g., to make calculations or reformat text strings). SQLite supports all the basic arithmetic (e.g., equality, the four operators (+ - \* /), inequalities, etc.) and logical operators (e.g., AND and OR), just to name the most common.

Some operators are combinations of simpler ones, like BETWEEN. For example, although the two statements

```
SELECT * FROM budget
WHERE (cost BETWEEN 10 AND 20);
SELECT * FROM budget
WHERE ((cost >= 10) AND (cost <= 20);
```

both mean the same thing – that is, “show me all data of all the expenses in the table called *budget* whose cost is between 10 and 20 dollars” – the first version is faster both to write and execute.

The last SQLite operator to introduce is LIKE, which is used to compare strings. In the statement

```
SELECT * FROM sales
WHERE customer LIKE '%William_';
```

the final underscore means that any *single* character in this position is valid, whereas the percent sign means that any *sequence* of characters in this position is valid. Therefore, the statement will display the sales made to customers with names like John Williams, B. Williamx, Mike T. Williamm, and so on.

Functions are operators that directly “extract” one number or string from (combinations of) other numbers or strings, for example:

- `abs(X)` returns the absolute value of the number X.
- `round(X, Y)` returns the number X rounded to Y decimal digits.
- `substr(X, Y, Z)` returns Z consecutive characters of string X, starting from position Y.
- `substr('Linux', 3, 2)` returns the string 'ux'.

SQLite has many more core and aggregate functions that are available by default, as well as functions you can use by loading extensions.

### Adding, Updating, and Selecting Records

The whole point of a database is to store data in an ordered way, keep it current, and above all search and filter the data with any combination of criteria. This work is done with the SQL commands INSERT, UPDATE, DELETE, and SELECT. In the next examples, assume you keep a catalog of all your books in the *books* table defined earlier. Listing 1 shows how you would add, update, and delete a book from the SQLite command prompt. (See the “Using SQLite at the Command Line” box.)

The first command adds a book to the database, and the second command corrects its price (assuming the automatically generated ID of that book record is 21). The third command removes all the records in the table with the same title and updated price.

In the next statement, the SELECT command, which has already been introduced, finds all the records that match certain conditions in the specified table

#### LISTING 1: Simple SQLite Statements

```
sqlite> INSERT INTO books(title, author, price)
VALUES('Lord of the rings', 'Tolkien' 25.50);
sqlite> UPDATE books SET price=20.50 WHERE bookid=21;
sqlite> DELETE FROM books WHERE ((title='Lord of the rings') AND (price=20.50));
```

and displays them grouped or sorted according to other criteria:

```
sqlite> SELECT title as TITLE,
sum(price) AS 'total_expense'
FROM books
WHERE author='Tolkien'
GROUP BY book_title
ORDER BY total_expense DESC;
```

In this case, it produces a list of how much you have spent on each Tolkien book in your library:

- in total (courtesy of the `sum()` function)
- even if you have more than one edition of the same book (thanks to the `GROUP BY` clause)
- from the most to least expensive (because of the `ORDER BY` clause in descending, `DESC`, order)

You can `SELECT` from multiple tables simultaneously by combining them with `JOIN` statements, as I will demonstrate in a moment.

## Practical Examples

The first cool thing you might want to do is get all or part of your data out of the database, formatted in ways that you can then easily process and reuse with other software.

For example, you can save to a text file all the SQL statements that would create a perfect copy of your database by using the `.output` and `.dump` meta-commands:

```
sqlite> .output mybackup.sql
sqlite> .dump
sqlite> .exit
```

Even better, you can save the same commands in a text file (or have other scripts generate them for you) and then run them with the `sqlite3` tool from the Linux command line:

```
# sqlite3 < commands.sqlite
```

To restore or clone a SQLite database from the dump, type (or, again, put it into a shell script) this command, instead:

```
# sqlite3 new.db < mybackup.sql
```

The file passed to `sqlite3` can contain any sequence of commands, including `SELECT` statements. Also, the data written to the file set by `.output` can be formatted in any of the ways supported by another very useful dot command called `.mode`: You can export your data as HTML tables, comma- or tab-separated files that are immediately usable in any spreadsheet, and five or six other formats.

## Managing Application Data

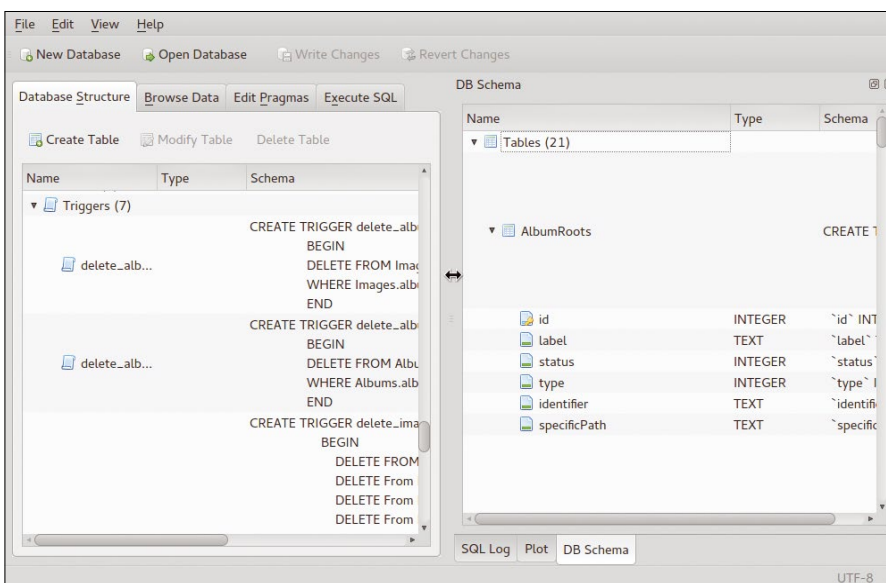
As I said at the beginning, several programs you already use on your desktop run SQLite under the hood. Knowing how to manage SQLite databases can help you get more from those programs than they offer you from their own interfaces. As just two of many examples,

I'll look at Firefox and digiKam. If you look at the hidden folder in which Firefox stores its data, you will find a number of SQLite databases, which all have the `.sqlite` file name extension. With the following simple shell script, you can periodically clean and compact all of these files to help Firefox run a bit faster:

```
cd $HOME/.mozilla/firefox/*.default
for i in *.sqlite
do
echo "VACUUM;" | sqlite3 $i
done
```

digiKam, on the other hand, stores location and all other metadata for each of the pictures it manages in one SQLite database called `digiKam4.db` (Figure 4).

Assume, for example, that you have hundreds of scanned photographs from the past 30 years scattered among many different digiKam albums. What do you do if your grandparents ask for a copy of all pictures that were taken at your parents' house? Should you find all the




**Figure 4:** DB Browser is quite useful whenever you need to study and understand the structure of an existing SQLite database that you did not create yourself. This screenshot shows some of the tables from the digiKam database.

## 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



[www.admin-magazine.com/newsletter](http://www.admin-magazine.com/newsletter)

corresponding albums manually? If they are properly geotagged (which is much easier than you might think, but that is a topic for another tutorial), you can tell SQLite to find those pictures for you.

Looking at the tables in the digiKam database with the metacommands `.tables` and `.schema` shows that the location of each album and the photograph file names are kept in the tables *Albums* and *Images*, respectively; another table, *ImagePositions*, stores the latitude and

longitude of each image. In these conditions, asking SQLite to list the locations of all the pictures whose latitude is the same as that of your parents' home (43,33 in this example) is relatively simple (Listing 2).

The `SELECT` command concatenates the three tables with the `JOIN` statement, producing the `photolist.txt` file with lines like those shown in Listing 3: a file, for example, that you can easily use in a `tar` command or in a script that copies the photographs onto a DVD. On the other hand, if you use `.mode html`, you can export your data as an HTML file (Figure 5).

### Graphical SQLite Interfaces

Command-line tools are the most efficient, but not the best for every situation. The most convenient graphical interfaces for managing SQLite databases on Linux may be LibreOffice Base and Kexi [2]. Several drivers can connect a SQLite database with the LibreOffice Suite; the easiest to set up at the time of writing was the ODBC Driver Extension [3]. Kexi, on the other hand, can deal with SQLite out of the box.

If you want a graphical interface, however, I recommend you at least test drive DB Browser for SQLite [4], as shown in the figures for this article. Like its command-line counterpart `sqlite3`, DB Browser is a multiplatform application and is available as a binary package for all the most common Linux distributions. For complex, raw queries, as well as heavy data processing tasks, I continue to prefer the command-line tool, but if you just need to tweak your SQLite configuration or study the structure of an existing database, DB Browser is a better choice.

### Conclusion

If you need a database that is simple to set up or want better access to the data that your software is already handling with SQLite, you now know how to proceed. The next step is to look at the official SQLite documentation, especially the "SQL Syntax" and "Core SQL Functions" sections [5], and bookmark some handy cheat sheets that will help you create queries, starting with the resources listed in the Info section [6]-[9]. Enjoy your portable databases! ■■■

### AUTHOR

**Marco Fioretti** is a freelance author, trainer, and researcher based in Rome, Italy. He has worked with Free/Libre Open Source Software (FLOSS) since 1995 and on open digital standards since 2005. Marco also is a Board Member of the Free Knowledge Institute (<http://freeknowledge.eu>) and a team member of the Digital DIY project (<http://www.didiy.eu>).



### INFO

- [1] SQLite: <http://sqlite.org/>
- [2] Kexi: [www.kexi-project.org](http://www.kexi-project.org)
- [3] LibreOffice SQLite ODBC driver: <http://extensions.libreoffice.org/extension-center/libreoffice-sqlite-odbc-driver>
- [4] DB Browser: <http://sqlitebrowser.org>
- [5] SQLite documentation: [www.sqlite.org/docs.html](http://www.sqlite.org/docs.html)
- [6] SQLite syntax reference: [http://www.sqlite.org/lang\\_expr.html](http://www.sqlite.org/lang_expr.html)
- [7] "My SQLite cheat sheet" by Niklas Ottosson: <http://blog.niklasottosson.com/?p=1342>
- [8] "15 SQLite3 SQL commands explained with examples" by Ramesh Natarajan: <http://www.thegeekstuff.com/2012/09/sqlite-command-examples>
- [9] "Searchable SQLite3 cheat sheet" by richardjh: <http://www.cheatography.com/richardjh/cheat-sheets/sqlite3/>



**Figure 5:** An HTML gallery of all (and only) the pictures from digiKam albums from a certain location.

#### LISTING 2: Selecting digiKam Images by Latitude

```
sqlite> .separator "/"
sqlite> .mode list
sqlite> .output photolist.txt
sqlite> SELECT relativePath, name from ImagePositions as P JOIN
(Albums as A JOIN Images AS I ON A.id = I.album) ON P.imageid =
I.id WHERE latitude = '43,33' ORDER BY relativePath, name;
```

#### LISTING 3: photolist.txt

```
[marco@localhost ~]$ head -2 photolist.txt
/marco/1978/family/birthdays/20080815-grandma-birthday/200808151951.jpg
/marco/2002/holidays/20021225-christmas-dinner/200212251505.jpg
```

 Find us on  
**Facebook**  
<http://www.facebook.com/linuxpromagazine>

# Your Roadmap to the Open Hardware Revolution ...

An exciting world of projects, tips, and skill-building tutorials awaits you in every issue of Raspberry Pi Geek.

Order your subscription today and tune in to the revolution!

[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

## Print Sub

Carry our easy-to-read print edition in your briefcase or backpack - or keep it around the lab as a permanent reference!

## Digital Sub

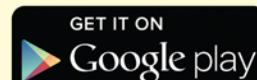
Our PDF edition is a convenient option for mobile readers.

6 print issues with 6 DVDs or  
6 digital issues for only

**\$59.95 £37.50 €44.90**



We're also in  
Google Play  
and the  
iTunes Store!



iTunes Store



Google US



Google UK

### The sys admin's daily grind: Charly's doorbell Pi

# Ring, Ring!

When Charly puts on his headphones at home, he often fails to hear the doorbell. So, he dreamed up a solution with a Raspberry Pi Zero, a noise detector, and a power outlet with a LAN connection. *By Charly Kühnast*

Sometimes I sit in my home office using a headset, listening to some good old electric guitar music or taking part in a teleconference – which is naturally slightly less invigorating. As a consequence of my acoustic escapism, I tend not to hear anyone who rings the doorbell.

To restore my social presence, I opted for the smallest version of the Raspberry Pi, the Zero. The Pi needs to detect the bell circuit closing and output a visual alert. In Germany, doorbells use 8 to 12V alternating current. You could convert this to DC and tune it down to a Pi-compatible voltage, but this would involve a mess of wire in the doorbell housing.

So, I went for a noise sensor. These things are very simply made, need a supply voltage of 5V, and send a signal via the output pin when they pick up a noise [1]. A blue rotary potentiometer lets me set the noise level. The sensor and the Raspberry Pi both fit into the doorbell housing (Figure 1).

Because of its proximity to the acoustic event, I can set the sensor's switching threshold to a fairly high level – it will not be tripped by the kids shouting or the dog going mad in the hallway.

### Lights and Tweets

The small Python program from Listing 1 evaluates the signal in an infinite loop. The GPIO Zero library [2], which I used here, might not support noise

sensors, but its `Button()` function is all I need to evaluate the sensor's short voltage pulse. Luckily, it includes a de-bounce feature: `bounce_time=2` summarizes all the signals the Raspberry detects within two seconds.

When the delivery man rings the bell now, the script `bell.sh` launches, which actuates a power outlet connected to my LAN, which switches on a lamp in my office. At the same time, the script sends

a tweet on Twitter. I know you can buy wireless doorbells down at the hardware store, but the doorbell Pi is more my style. ■■■

#### LISTING 1: Doorbell Script

```
01 #!/usr/bin/env python3
02 from gpiozero import Button
03 from signal import pause
04 import os
05
06 def bell_rang ():
07     os.system("/usr/local/shellscripts/bell.sh")
08
09 button = Button(21, bounce_time=2)
10 button.when_pressed = bell_rang
11 pause()
```

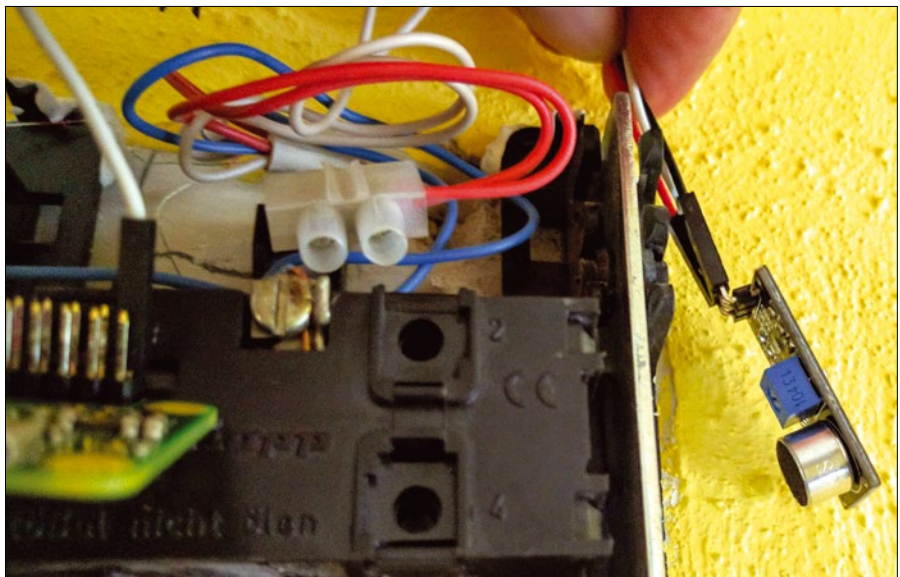


Figure 1: Noise sensor (right) and Raspberry Pi (left) have plenty of space in the doorbell housing.

#### INFO

- [1] Source for noise sensor: <http://www.amazon.de/gp/product/B00N1TSM7K>
- [2] GPIO Zero: <http://gpiozero.readthedocs.org>

#### CHARLY KÜHNAST

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.

Now Appearing on

# APPLE NEWSSTAND

New age convenience...

Our inspired IT insights  
are only a tap away.

Look for us on  
Apple Newsstand  
and the iTunes store.

Download  
a FREE issue of  
each publication  
now!



Monitor login attempts on your home WiFi

# Private Reception

Push notifications to your cell keep you up to date with the details of any clients that log in or out of your home WiFi. *By Mike Schilli*

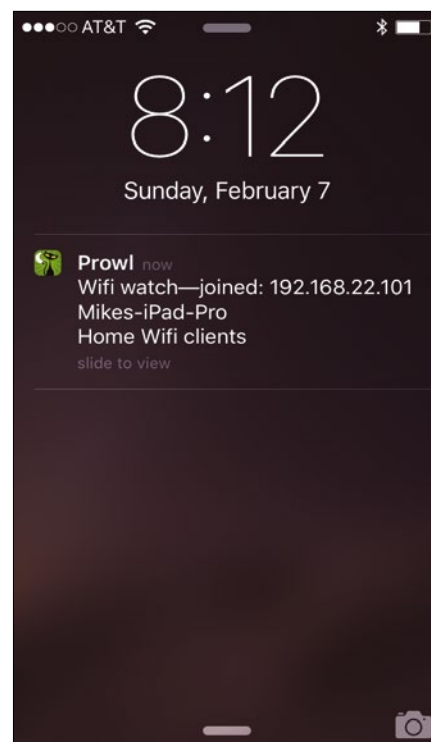


**W**henever the lights on my router start to flash like crazy when I'm not actually doing anything on the network, I always suspect that one of my neighbors might have cracked my WPA2 password and is making mischief on my Internet account. I live in a large city on the second floor of a house half way up a steep hill. My guess is that a couple of hundred people can receive my WiFi router's signal – and that potentially includes a couple of good for nothings who would love to mess around with it.

The scripts I look at this month regularly pick up a list of wireless devices logged in to my WiFi router and text my cell phone when a new client logs onto or off of my home WiFi network.

Instead of programming a new cell

app for this, Prowl for the iPhone or Notify my Android for Android devices has a web interface that lets scripts send messages. The Prowl servers then ensure

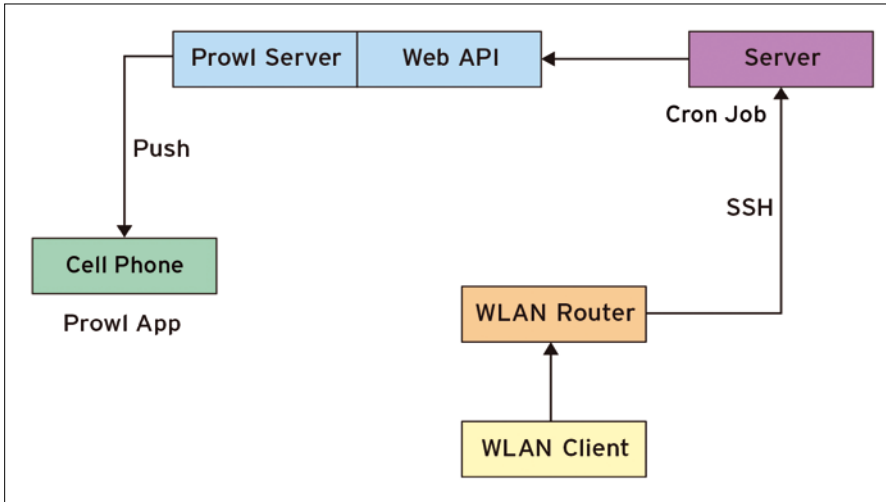


**Figure 1:** One glance at the cell phone reveals who has just logged on to the home WiFi network.

### MIKE SCHILLI

Mike Schilli works as a software engineer in the San Francisco Bay Area. He can be contacted at [mschilli@perlmeister.com](mailto:mschilli@perlmeister.com). Mike's homepage can be found at <http://perlmeister.com>.





**Figure 2:** The router pushes active DHCP data to a hosting service, where a cron job notifies the Prowl server of events via its web API. Prowl then sends a push notification to phones with the Prowl app installed.

that the events are forwarded to all the endpoints, with the Prowl app running under the same user account. This will even work with a locked phone; in this case, the messages just briefly flash up on the lock screen (Figures 1 and 2).

### Unusable Toolchain

How did I get the Asus RT-66U router, on which I immediately installed the free DD-WRT software after my purchase, to monitor active WiFi users and call the web service on the Prowl server in case

of changes? The process that handles the dynamic DHCP IP addresses on the router goes by the name of `dnsmasq`; it stores the DHCP leases currently in use in the `/tmp/dnsmasq.leases` file. The column format used here (Listing 1) lists the expiration time in seconds for each lease, the MAC address of the device, the IP address assigned to it, and a name the device uses to describe itself [1].

On a Linux distribution, it would be relatively simple to use an infinite loop to detect incoming or outgoing clients and send an HTTP request to the Prowl API in the case of changes. However, the DD-WRT distro on the router has a very limited selection of Unix tools, and solving the normally trivial problem of searching through the lease file with Perl and firing off HTTP requests, if needed, becomes a genuine brain teaser.

Although I could install Perl on the router, along with a couple of CPAN modules, the more the configuration were to deviate from a plain vanilla DD-WRT distribution, the more error prone the upgrade results would be. A simple shell script that uses the BusyBox tools on the router will hopefully still run in a couple of years without manual changes, and it's fairly easy to install via the DD-WRT GUI. The GUI stores the script in the router's NVRAM, pushes it into the

#### LISTING 1: `dnsmasq.leases`

```

1 86400 74:da:42:1b:44:a7 192.168.20.148 raspberrypi *
2 86400 e8:80:2e:e9:11:a9 192.168.20.130 MikesiPhone e8:80:2e:e9:11:a9
3 86400 10:68:3a:17:4a:ba 192.168.20.131 android-oba143110ae5ee34 10:68:3a:17:4a:ba
4 86400 00:51:b6:76:a1:b6 192.168.20.134 Mikes-Macbook 00:51:b6:76:a1:b6
  
```

#### LISTING 2: `lease-push.sh`

```

01 #!/bin/sh
02 file=/tmp/dnsmasq.leases
03 var=lease-file
04
05 HOME=/tmp/root
06
07 cd /tmp
08 uuencode <<'EOT'
09 begin 664 keyfile
10 M5V%S(&EC:"!N:6-H="!W96G#GRP@;6%C:'O@;6EC:"!N:6-H="!H96G#
GRX@
11 [...]
12 51\.V=&4N($]D97(@4V-H:6QL97(N
13 `
14 end
15 EOT
16
17 cd /tmp/root/.ssh
18 uuencode <<'EOT'
19 begin 664 known_hosts
20 M3&ER=6T@3&%R=6T@3,.V9F9E;' -T:65L+B!797(@;FEC:'!S(&AA="P@
9&5R
21 [...]
22 >(&AA="!N:6-H="!V:65L+B!$;VYA;&O@5')U;7`N
23 `
24 end
25 EOT
26
27 cd /tmp
28 while [ "forever" ]
29 do
30     sum=`/usr/bin/shasum $file | cut -d " " -f1`
31     stored="$(/usr/sbin/nvram get $var)"
32
33     if [ "$stored" != "$sum" ]
34     then
35         /usr/bin/scp -i keyfile $file
perlsnapshot@<somehoster.com>;
36         /usr/sbin/nvram set $var=$sum
37     fi
38
39     sleep 30
40 done
  
```

filesystem on reboot, and has it fire up its infinite loop.

### Tongues of Angels

The shell script in Listing 2 thus makes do with `ssh` to upload the file with the issued IP addresses to an account on my hosting provider's server, where a cron job regularly checks them for changes and contacts the Prowl server using Perl and an API key obtained from Prowl's website.

Even this task isn't totally simple; after all, the `scp` client on the BusyBox-based distribution is not fully functional. It takes the tongues of angels to get it to authenticate against the hosting provider's server using a private key and to send the file to the server.

After installing a private key generated using `ssh-genkey` for the SSH process on the DD-WRT distribution, the started process just disappears after reporting that some internal string somewhere in the depths of its implementation is too long. Thanks a lot, BusyBox! If you google this message, you will discover that the pared down SSH version uses a special format for private keys. Instead, a tool named `dropbearkey` on the DD-

WRT router generates a compatible key. For this, log into the router via `ssh` as root after enabling the SSH daemon in the GUI below *Services*. The call

```
dropbearkey -t rsa -f keyfile
```

generates the private key in the `keyfile` file.

The `dropbearkey` tool displays the matching public key on its standard output. If you copy this text string into the `.ssh/authorized_keys` file on the server, and open the connection on the router by typing

```
ssh -i keyfile <servername>
```

at the command line, you do not need to enter a password and can thus automate the process.

### Hard Wiring

Unfortunately, the filesystem on the router does not survive a reboot because it resides in volatile memory, and all the information is rebuilt from NVRAM after the reboot. This is why programmers embed all required information in the scripts that need to run on the router.

Listing 2 relies on the slightly ancient `uudecode` to handle the binary data from the special private key file; its unreadable mess of data, with a width of precisely 61 characters, fits perfectly into a source code listing. The following call created the data:

```
cat keyfile | uuencode -f keyfile
```

I manually copied the output into Listing 2. The same procedure as for the private key applies to the `known_hosts` file, which lists the SSH host keys for known hosts. The first call to `ssh` to open the connection to a new server interactively prompts you on the terminal to decide whether the server seems to be trustworthy. If you say `y`, SSH stores the host key in `~/.ssh/known_hosts` file. The `uuencode` data in lines 20-22 were created by running `uuencode` against the `known_hosts` file and cutting and pasting its output into Listing 2.

### Infinite Future

The infinite loop in lines 28-40 of Listing 2 checks the DHCP lease file on the router every 30 seconds to discover whether the list with the MAC addresses extracted using the `cut` command results in a different `sha1sum` footprint than the previous run. To maintain that value even through a reboot, line 36 stores the current value in the router's non-volatile RAM using `nvrnm set`. `nvrnm get` in line 31 picks the value up again to compare it with the current value from the lease file and transfers the current file to the server in case of changes.

To make sure the router runs the script after every reboot, I dropped it in the web UI in the *Startup* text field below *Administration | Commands* and pressed the *Save Startup* button at the bottom of the page (Figure 3). If something goes wrong, it isn't easy to discover why; the slimmed down commands on DD-WRT do not provide very much in the way of detailed error messages, but if you log in to the router using SSH and run the script manually at the command line, you can normally find the culprit.

### Messenger

The router thus ensures that the current version of the lease file is always available on a Linux server where a cron job,

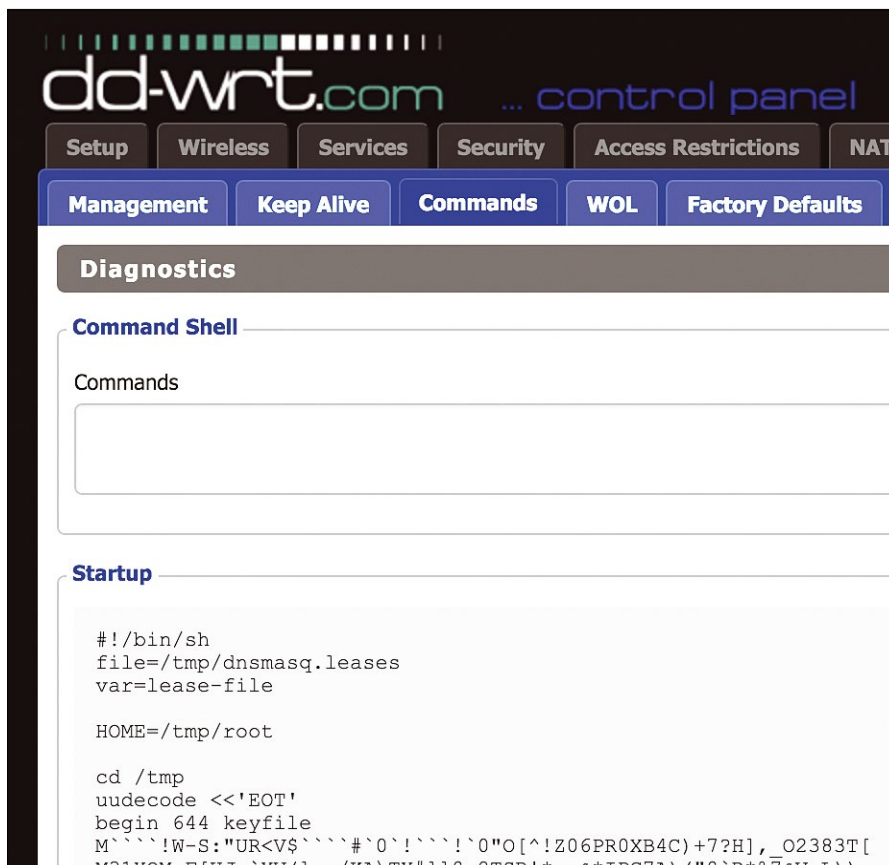


Figure 3: This script periodically sends changed lease entry files to the server.

### LISTING 3: prowl-test

```

01 #!/usr/local/bin/perl -w
02 use strict;
03
04 use WebService::Prowl;
05 my $ws = WebService::Prowl->new(
06     apikey => "xxxxxxxxxxxxxxxxxxxxxxxxx");
07
08 $ws->verify || die $ws->error();
09
10 $ws->add(
11     application => "Perl Snapshot",
12     event       => "Just a test.",
13     description => "Huzzah, it works!",
14     url         => "http://linux-magazin.de",
15 );
    
```

```
**** /home/perlsnapshot/lease-notify
```

running once a minute, detects the changes and notifies the Prowl web service about events, like clients joining or leaving the WiFi.

The Prowl service requires users to register on the Prowl website [2] with a username and password, where you do not need to provide an email address; however, it is useful later if you forget your password. An account entitles you to obtain API keys with which applications that want to send events authenticate against the web service. The Prowl server forwards the text messages to all instances of the app registered under the account. The app costs \$2.99 for an iPhone, but registering on the website to pick up an API key and forwarding the push notifications are both free for up to 1,000 events sent per day.

### Link If Needed

The simple test included in Listing 3 shows the procedure that uses the CPAN `WebService::Prowl` module, which nicely abstracts access to the Prowl web service. It requires an API key, which is available on the website under the API Keys tab as a text string for the logged-in user (Figure 4). The `verify()` method in line 8 checks whether communication to the Prowl server is working, and the `add()` method in line 10 accepts the four text strings for the message to be sent over the air.

The first three parameters provide the name of the sending application, the type of event, and a short description. The fourth parameter, `url`, carries a

URL that the cell phone is supposed to access when the user taps on the event and then consents to launching the browser. In Figure 5, you can see the push notification for the test event from Listing 3 just arriving on my locked iPhone 5. Figure 6 shows the dialog that appears when the user taps on the event.

On the server side, the script in Listing 4 runs as a cron job. It maintains a persistent data repository in the `leases.dat` file, to which it binds the `%leases` hash in line 14 using `tie`. After `lease-notify` is

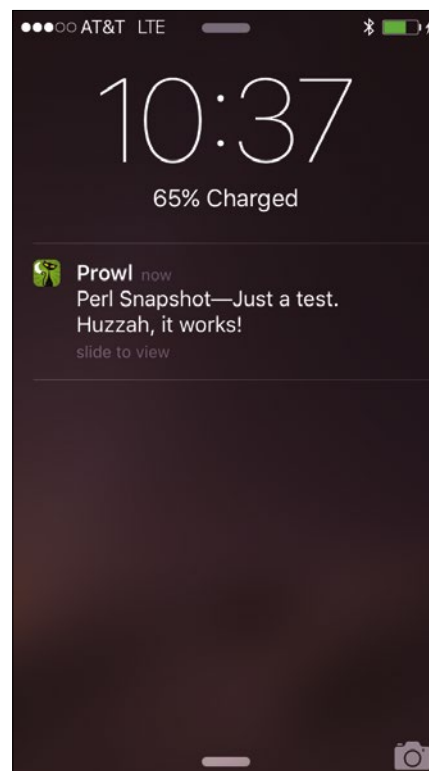


Figure 5: The iPhone showing the test message on the lock screen.

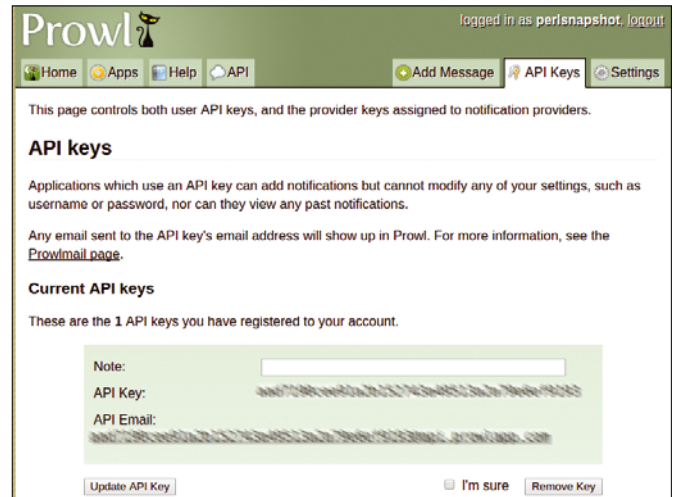


Figure 4: Free API keys from Prowl.

launched, the program looks for a file named `dnsmasq.leases`, loads its data using the CPAN `Path::Tiny` module, and then iterates through its text with `lines()` in line 20. The `split()` function in line 22 breaks down the space-separated fields into the remaining lease duration, MAC address, issued IP address, and device name.

### Coming and Going

Listing 4 then stores all the discovered MAC addresses in the persistent hash, `%leases` and maps them to the matching

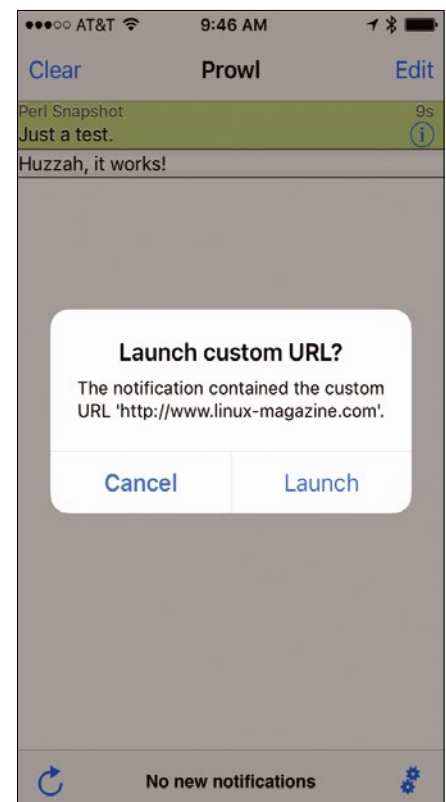


Figure 6: Opening the URL for the event.

## LISTING 4: lease-notify

```

01 #!/usr/bin/perl -w
02 use strict;
03 use lib '/home/perlsnapshot/perl5/lib/perl5';
04 use Path::Tiny;
05 use DB_File;
06 use WebService::Prowl;
07 use File::Basename;
08
09 my $in_file = "dnsmasq.leases";
10 my $db_file = "leases.dat";
11 my $API_KEY = "xxxxxxxxxxxxxxxxxxxxxxxx";
12 my $APP_NAME = "Wifi watch";
13
14 tie my %leases, 'DB_File', $db_file;
15
16 my $file = path( $in_file );
17
18 my %found = ();
19
20 for my $line ( $file->lines ) {
21     my( $secs, $mac, $ip, $name ) =
22         split " ", $line;
23
24     $found{ $mac } = 1;
25
26     if( !exists $leases{ $mac } ) {
27         $leases{ $mac } = "$ip $name";
28         notify( "joined: $leases{ $mac }" );
29     }
30 }
31
32 for my $mac ( keys %leases ) {
33     if( !exists $found{ $mac } ) {
34         notify( "left: $leases{ $mac }" );
35         delete $leases{ $mac };
36     }
37 }
38
39 untie %leases;
40
41 #####
42 sub notify {
43     #####
44     my( $event ) = @_;
45
46     my $ws = WebService::Prowl->new(
47         apikey => $API_KEY );
48
49     $ws->verify || die $ws->error();
50
51     $ws->add(
52         application => $APP_NAME,
53         event => $event,
54         description => "Home Wifi clients",
55         url => "",
56     );
57 }

```

IP addresses and device names. It thus knows which devices existed on the previous run, and which have just been added since then. Line 27 then adds the new entries to the hash's persistent memory, and line 28 fires off the events defined in lines 42-57 with `notify()`.

A similar script branch handles devices that existed in the previous run and which thus exist in `%leases` but are now missing in the current run. The volatile `%found` hash stores these. If a discrepancy is discovered, line 34 then sends a message stating that the device has disappeared.

The `notify()` function basically looks like the test script introduced in Listing 2. It uses the Prowl API key stored in `$API_KEY` at the start of the script and only adds the application name, the event type (`joined` or `left`), and the description, while it leaves the URL field empty.

Because my low-budget hosting service does not allow root access, I installed the CPAN modules required by

the script locally in the home directory below `perl5`; `cpansm` does this automatically if it notices that it cannot manipulate `/usr/lib` because it is lacking the necessary permissions. But for the script to find the modules installed there, line 3 in Listing 4 needs to add this path explicitly using `use lib`.

## No Invasion – Yet

It was very reassuring to see that the script only discovered known devices on my WiFi during the beta testing phase, but at least now I am perfectly prepared for a full-scale over-the-air attack. Another thing I noticed is that some devices suddenly connect to the WiFi network in the middle of the night, even though they are switched off and lying somewhere in the corner of the room, one example being my Kindle Paperwhite ebook reader, most likely checking for available software updates.

It would be fairly easy to improve the script to know which MAC address be-

longed to which device, which you could easily handle using a hash in `lease-notify`. The text messages would then use device names designed to reflect the situation on the home network, which would make them much easier to understand when received.

## Acknowledgment

I thank my co-worker Tristan Horn, whose idea it was to display devices joining and leaving the home network on a phone, and he also wrote an application for this purpose that integrates the whole enchilada in a far more professional UniFi system [3]. ■■■

## INFO

[1] Listings for this article:

<ftp://ftp.linux-magazine.com/pub/listings/magazine/186>

[2] Prowl: <http://www.prowlapp.com>

[3] Connecting push notifications with a UniFi system: <https://tris.net/software/unifi-logreader/>

# MORE UBUNTU!



**Can't get enough Ubuntu? We've got a whole lot more!**

**Ubuntu User** is your roadmap to the Ubuntu community. In the pages of **Ubuntu User**, you'll learn about the latest tools, best tricks, and newest developments in the Ubuntu story.

**Ubuntu User** helps you explore the treasures of open source software within Ubuntu's expansive repositories. We'll bring you exclusive interviews with Ubuntu leaders, keep you current on the exciting Ubuntu community, and answer your most perplexing Ubuntu questions. Learn how to choose a video editor, find the perfect tool to customize your desktop, and configure and manage Ubuntu systems using the best admin tools.

**DON'T MISS ANOTHER ISSUE!**



**HUGE SAVINGS OFF THE NEWSSTAND PRICE!**

**SUBSCRIBE NOW: SHOP.LINUXNEWMEDIA.COM**



Software configuration management with Fossil

# Fossilized Code

Get started with Fossil, a beginner-friendly software configuration management system that includes everything you need to work on your next great software project.

By Dmitri Popov

## 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, Spanish, and Russian magazines and websites. Dmitri is an amateur photographer, and he writes about open source photography tools on his Scribbles and Snaps blog at [scribblesand-snaps.wordpress.com](http://scribblesand-snaps.wordpress.com).

Git and GitHub are by far the most popular choices when it comes to writing and managing code. Because of their popularity, however, it's easy to overlook other software configuration management (SCM) systems that may prove to be a better fit for your next coding project. Case in point: Fossil [1], a lightweight SCM that is easy to master and offers a range of genuinely useful features and functionality.

Fossil's most important advantages compared with other SCMs are the integrated bug tracker, wiki, and so-called technotes. Better still, Fossil features a built-in web interface that gives you instant web access to the repository and all its tools. In other words, Fossil offers everything you need to document your code, track bugs, and collaborate right out of the box.

Fossil offers plenty of features and creature comforts that make it particularly suited for individual developers and casual coders who need a straightforward way to keep track of their small-scale coding projects. If you are one of them, then Fossil is right up your alley.

## First Steps with Fossil

The Fossil application is a self-contained, standalone executable, and it's available for all major platforms. If you run a mainstream Linux distribution, chances are Fossil is available in the official software repositories. Alternatively, you can compile the latest version of Fossil yourself using the following commands (replace `x.xx` with the latest version number)

```
wget https://www.fossil-scm.org/
download/fossil-src-x.xx.tar.gz
tar xzvf fossil-src-x.xx.tar.gz
cd fossil-src-x.xx
./configure --with-openssl=none
make
```

The `--with-openssl=none` parameter in the `./configure` command omits HTTPS functionality (which you probably don't need when using Fossil on your local Linux machine). Before compiling Fossil on Ubuntu, you need to install the `zlib` library manually:

```
sudo apt-get install zlib1g-dev
```

Once the `make` process is finished, move the resulting `fossil` binary to the directory in your `PATH` (e.g., `/usr/bin`).

The first step after you've installed Fossil is to initialize a repository in the directory that contains files for the coding project (i.e., the source tree) that you want Fossil to manage. Switch to the directory and run the `fossil init` command followed by the desired repository name, for example:

```
fossil init foo.fossil
```

The created repository database contains all check-ins (snapshots of the source tree at any point in time). In other words, unlike other SCM systems that store repository data across multiple files in a dedicated directory, Fossil keeps a source tree and version control data in a single database that acts like a regular file. You can rename, copy, and move it at will.

Next, open the repository to the working directory and add the project files to the repository:

```
fossil open foo.fossil
fossil add .
```

The last command adds all files in the current directory and subdirectories recursively. If you want to exclude certain files from being tracked, you can use the `fossil settings ignore-glob` command:

```
fossil settings ignore-glob "*.tasks,*.log,.directory"
```

This operation will ignore files matching the specified criteria for the current repository only, but you can add the `--global` parameter to apply the defined rule globally:

```
fossil settings ignore-glob "*.md,*.log,.directory" --global
```

If the repository contains images, fonts, and compiled executables, you can use the `fossil settings binary-glob` command to configure Fossil to treat them as binary files for committing and merging:

```
fossil settings binary-glob "*.jpg,*.png" --global
```

Finally, you can create the initial commit using `fossil commit`:

```
fossil commit -m "Initial commit"
```

As you would expect, Fossil also allows you to clone existing repositories using the dedicated `fossil clone` command. In addition to local repositories, Fossil supports cloning of remote repositories via HTTP and SSH. All you have to do is to point the `fossil clone` command to a repository file on a remote machine and specify the desired name for the cloned repository file:

```
fossil clone ssh://user@remotehost/path/to/foo.fossil foo.fossil
```

The clever part is that, when you commit modifications in a cloned Fossil repository, the system automatically pushes changes to the remote original repository. This feature is called Auto-Sync, and it provides an easy and hands-free way of keeping the original repository and its clones in sync. (See the “From Git to Fossil” box for more information.)

To pull recent changes from the remote repository, use the `fossil pull` command followed by the `fossil update` command.

Once you’ve created or cloned a repository, your typical Fossil workflow may look something like this:


- Add, edit, and delete some files in the source tree.
- Run the `fossil extra` command to list new files. This step is optional.
- Run `fossil status` to list modified files. This step is optional.
- Run `fossil addremove` to add the new files to the repository and remove the deleted files.
- Run the `fossil commit -m "Commit message"` command to commit the changes.

For a quick view of the differences between the current and previous versions of a file, run `fossil diff`:

```
fossil diff foo.php
```

## FROM GIT TO FOSSIL

Thanks to Fossil’s import functionality, migrating existing Git repositories to Fossil is as easy as it gets. Switch to a directory under Git control, and run the

```
git fast-export --all | 
fossil import --git foo.fossil
```

command (replace `foo.fossil` with the desired name for the resulting Fossil repository). That’s all there is to it.

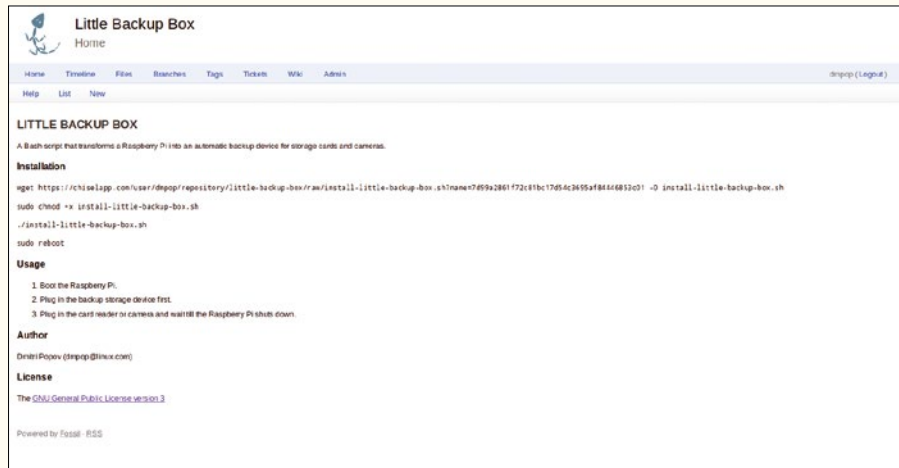


Figure 1: Fossil’s web interface.

When you need to cancel changes made to a file, use the `fossil revert` command

```
fossil revert foo.php
```

to go back to the previous version.

## Working with the Web Interface

The built-in interface provides an easy way to configure and work with the current repository from the convenience of your preferred browser (Figure 1). Fossil has two commands for launching the web interface. The `fossil ui` command launches the web interface with the default user logged in. This command also binds to the loopback IP address (127.0.0.1), so it can’t be used to serve content to a different machine.

The `fossil server` command doesn’t bind to the loopback IP address, and you need to log in with credentials automatically generated when you initialized the repository using the `fossil init` command. By default, Fossil’s web interface runs on port 8080, but you can change that using the `-port` parameter. Here are examples of using both commands:

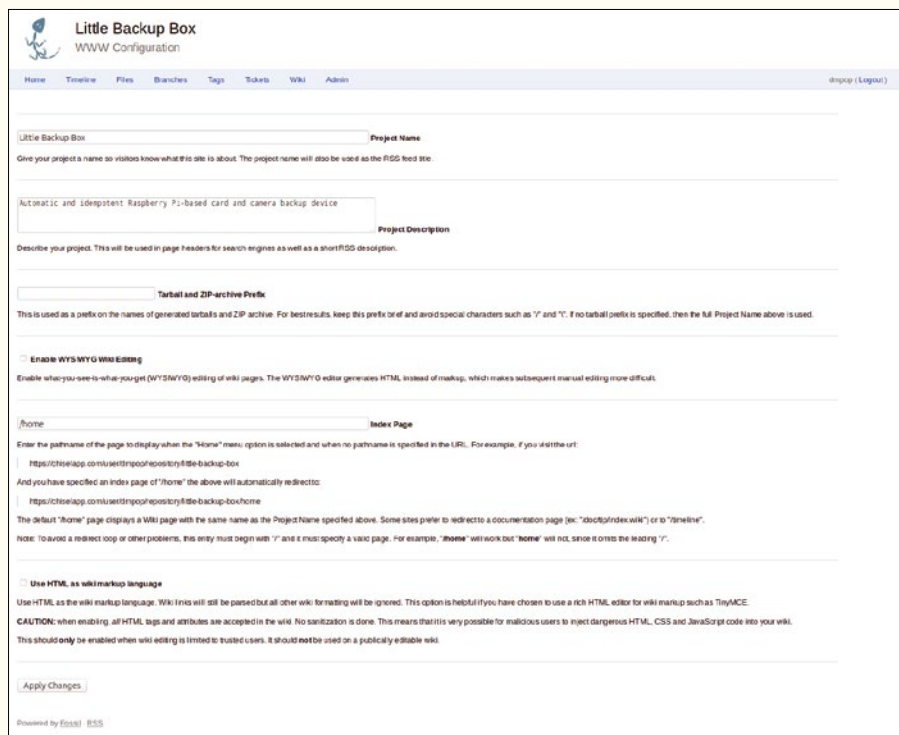


Figure 2: Configuring the repository’s basic settings.



```
fossil ui foo.fossil
fossil server foo.fossil
fossil ui foo.fossil -port 8381
```

When you run one of these commands, the web interface automatically opens in the default browser. Before you familiarize yourself with the interface and its functionality, it's a good idea to configure the repository's few basic settings (Figure 2). You can switch to the *Admin* area to access the list of configurable items. To begin, you might want to give the project a name and provide a short description, which can be done in the Configuration section. It also makes sense to change the automatically generated password for the default admin user. To do this, switch to the *Users* section, select the default user, specify the desired password, and press *Apply Changes*.

Speaking of users, Fossil comes with several default user accounts: anonymous, nobody, developer, and reader. All users, whether they are logged in or not, inherit the privileges of the nobody account. Anyone can log in to the repository as an anonymous user with the password shown on the homepage. Logged in users inherit all anonymous and nobody privileges. Knowing this can help you configure access rights for visitors, users, and contributors alike.

Fossil's web interface provides access to several modules, including *Timeline*, *Tickets*, and *Wiki*. Fossil's wiki engine powers practically all pages in the repository, and the Wiki section lists all existing pages and lets you easily add new ones. If you've ever tried to work with a wiki, you shouldn't have problems figuring out how to use Fossil's wiki functionality. Note that the wiki supports both native Fossil Wiki markup as well as Markdown.

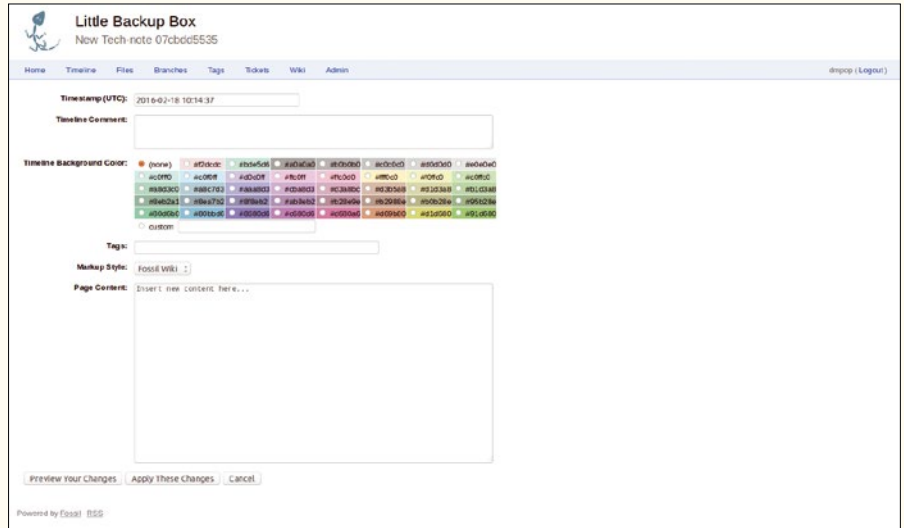


Figure 3: In addition to regular wiki pages, Fossil supports technotes.

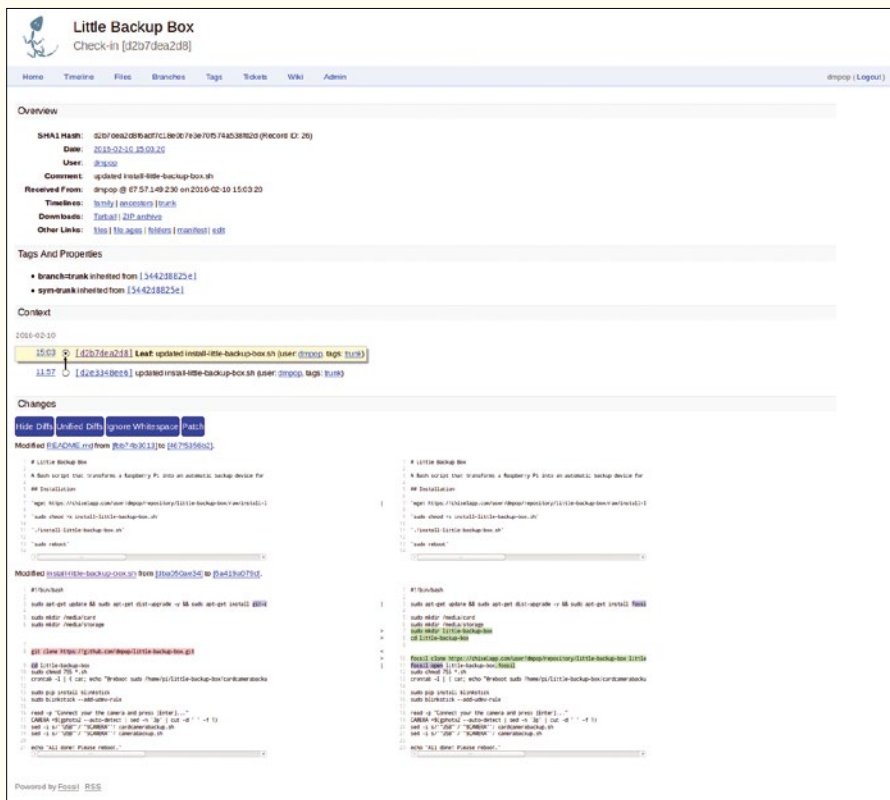


Figure 4: Check-in page offers detailed info and lets you compare files side by side.

In addition to regular wiki pages, Fossil lets you create so-called technotes (Figure 3). These are special wiki pages that – instead of page names – are associated with specific points in time. Each technote appears as an item on the Timeline (more about the Timeline later), and clicking on the timeline link will display the text of the technote. A technote features editable content, timeline entry text, and a timestamp. It's also possible to assign a color label and multiple tags to a technote.

As the name suggests, the Timeline section displays repository events, such as check-ins and wiki page edits, in chronological order. Each check-in entry (Figure 4) has a hyperlinked label consisting of the first 10 digits of the check-in's hash value. Click on the label to see detailed info about changes in this check-in. The two buttons on the page – *Show Unified Diffs* and *Show Side-by-Side Diffs* – can be useful for viewing differences between the current file and its previous version. The *Show Unified Diffs* button displays and highlights the modification inside the file, whereas the *Show Side-by-Side Diffs* button displays two versions of the file with highlighted changes side by side.

The Tickets section (Figure 5) provides tools for filing and tracking issues. Creating and submitting a ticket is rather easy, because you only have to provide a few simple bits of information in the default ticket form. Fossil also allows you to customize the default ticket form to your specific needs. For example, if you want to add a new item to the *Type* drop-down list, you can easily do this by navigating to *Admin | Tickets | Common* and inserting the desired item in the *set type\_choices* list.

In a similar manner, you can modify the default priority, severity, resolution, and status choices and tweak the default New Ticket and Edit Ticket pages using the appropriate links in the *Admin | Tickets* section. The *All Tickets* item runs an SQL query and presents the result as a nicely formatted table of all tickets. If you have a working knowledge of SQL, you can customize the report by clicking on the *edit* link and modifying the default SQL query. And, you can create additional reports using the *New report format* link.

## Fossil Hosting with Chisel

You don't need to set up a dedicated server for serving Fossil repositories via HTTP or SSH: Any machine or virtual private server accessible via the Internet will do just fine. But, if you are not keen on hosting Fossil repositories on your own server, Chisel [2] has you covered. This free service lets you to host an unlimited number of repositories and will back them up every week to keep them safe. The application that powers Chisel is open source and the service is available free of charge.

## Final Word

Although Fossil doesn't enjoy the same popularity as Git, this software configuration management system has a lot going for it. It's easy to master, it offers a range of useful tools and functionality out of the box, and its repositories are easy to manage and host. If you are curious how Fossil compares with Git, you might want to take a look at the Fossil vs. Git page [3] and the Fossil vs. Git article [4]. ■■■

### INFO

- [1] Fossil CMS: [fossil-scm.org](http://fossil-scm.org)
- [2] Chisel: [chiselapp.com](http://chiselapp.com)
- [3] Fossil vs. Git page:  
[www.fossil-scm.org/fossil/doc/trunk/www/fossil-v-git.wiki](http://www.fossil-scm.org/fossil/doc/trunk/www/fossil-v-git.wiki)
- [4] Fossil vs. Git article:  
[pietersz.co.uk/2015/03/fossil-vs-git](http://pietersz.co.uk/2015/03/fossil-vs-git)

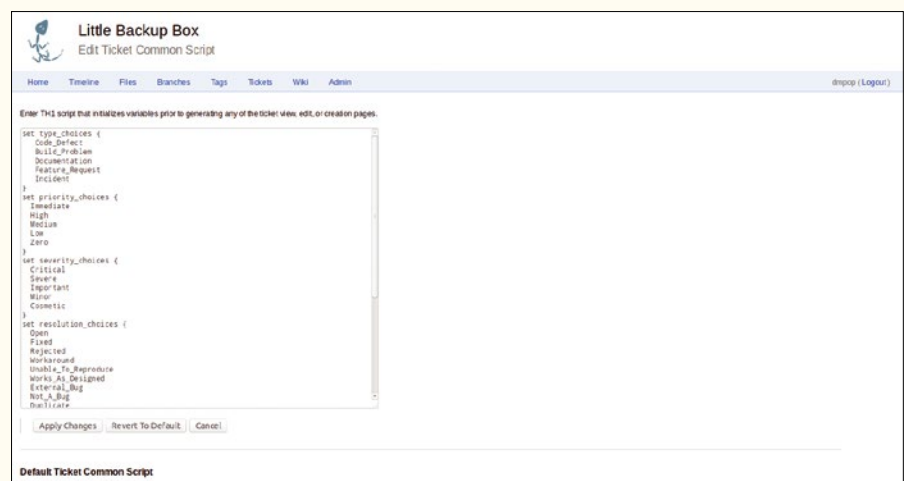


Figure 5: Customizing ticket options.

# IT Highlights at a Glance

The image shows a woman sitting on the floor with a laptop, looking at the screen. The screen displays a collage of IT news newsletters. The newsletters shown are:

- ADMIN HPC**: Features articles like "Spanning Tree Protocol" and "Further Reading" on HPC topics.
- ADMIN Update - Hottest Links**: Lists top admin tools, high availability solutions, and security updates.
- LINUX UPDATE**: Focuses on exploring the world of Linux, featuring articles on KDE's graphics editor and the benefits of free software.
- RASPBERRY PI GEEK**: Issue 07, featuring articles on the Raspberry Pi ecosystem and a "FREE FROM XP" special offer.

Too busy to wade through press releases and chatty tech news sites? Let us deliver the most relevant news, technical articles, and tool tips – straight to your in box. Subscribe today for our excellent newsletters:

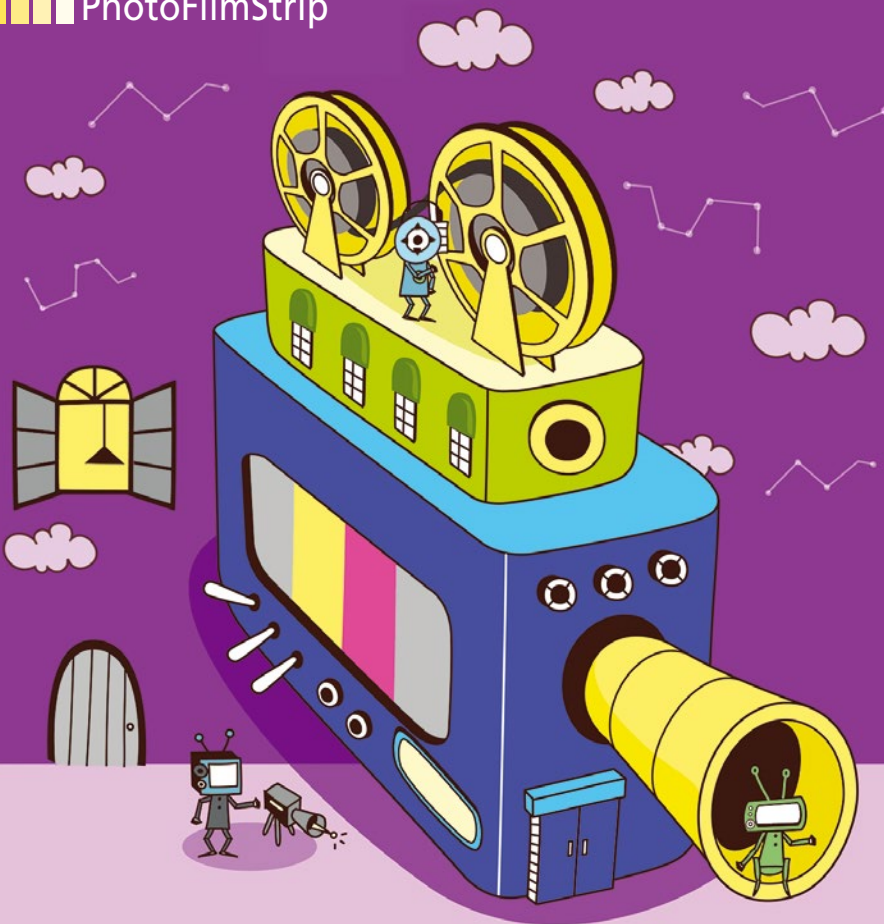
- ADMIN HPC
- Linux Update
- ADMIN Update
- Raspberry Pi

and keep your finger on the pulse of the IT industry.

Admin and HPC: [www.admin-magazine.com/newsletter](http://www.admin-magazine.com/newsletter)

Linux Update: [www.linuxpromagazine.com/mc/subscribe](http://www.linuxpromagazine.com/mc/subscribe)

Raspberry Pi: [www.raspberry-pi-geek.com/mc/subscribe](http://www.raspberry-pi-geek.com/mc/subscribe)



## Animated slideshows with PhotoFilmStrip

# Frame by Frame

Easy-to-use PhotoFilmStrip produces high-quality videos and offers plenty of useful features. *By Mario Blättermann*

### MARIO BLÄTTERMANN

Mario Blättermann is responsible for translations and integrating new templates and documentation with the gLabels project. He also works as a translator for the Gnome project and builds packages for Fedora.

Evening slideshows were often highly technical operations in the glory days of analog photography. Along with crossfades using two projectors or one projector with two lenses, it was also possible to provide both the soundtrack and the control signals for the picture change using a multitrack tape recorder. The results were pretty impressive and far superior to tediously skimming through an album.

But, no matter how you looked at it, the slideshow as such remained a static image. Now, however, you can not only compile slideshows and play movies using software, you can also create the illusion of a real video by panning and zooming in on images. This is much

more than just a gimmick; it can be used specifically to draw the viewer's attention to certain parts of an image that might otherwise go unnoticed. Whereas it only used to be possible to create such effects using professional rostrum cameras [1] – which are placed above the subject to facilitate zooming and panning – today, the software takes care of both processing the frames and rendering playable movies.

PhotoFilmStrip [2] is one such program that breathes new life into your photos with the help of the Ken Burns effect. The name comes from American filmmaker, Ken Burns [3], who, although he might not have invented the technique, he perfected it and earned numerous awards for his use of it.

After starting the program for the first time, you can create the basic framework for a new filmstrip via the *File | New Project* menu or directly by pressing the large plus icon on the home screen. The default settings of a 16:9 widescreen format and a total length of 30 seconds are just right for your first attempt (Figure 1). (See the “Installation” box for details.)

The new project is empty to start. The *Import Pictures* item is hidden in the *Tools* menu, but you have a corresponding icon in the toolbar as well. Simply select multiple images (or even just one) in the selection dialog, and PhotoFilmStrip will present an overview, as in Figure 2, with the first image to be processed in the top section.

You can now adjust the motion by panning and zooming the frame superimposed by two image views. No matter where you position the mouse pointer on the frame, the image's aspect ratio is maintained and prevents the video from showing black bars later. However, even if you don't do anything and just export the filmstrip to the target format, the result will still be impressive. Miraculously enough, PhotoFilmStrip seems to detect which parts of an image are acceptable and which parts shouldn't be used in the video.

An instant export also gives you a good impression of the preset fade speed. You can start rendering the video via *Tools | Render filmstrip* or by using the corresponding button in the toolbar. Various disk formats are available in the settings window (Figure 3) beside *Format*, ranging

from the ancient video CD up to the no longer state-of-the-art DVD.

There are also high-definition videos: You can set one of the MPEG4 variants as the format and then choose between *HD*, with 720 image lines, and *Full-HD*, with 1080 image lines, from the *Profile* field, which should no longer be grayed out. To make fine adjustments, click the button to the right of the *Format* field, where you can change the bit rate and other parameters.

## Off the Cuff

The shortcut buttons that lie between the two image views are very handy. The top button produces a *Random motion*; in principle, this is what PhotoFilmStrip does anyway when reloading images. The next two buttons down copy the left selection to the right and vice versa. This means that the optimal segment found is kept; the image remains static and isn't panned or zoomed. If the fourth button is pressed, PhotoFilmStrip exchanges the two selections.

## INSTALLATION

Sufficiently up-to-date packages for PhotoFilmStrip are only available for a few distributions. Debian users will find it in the Multimedia repository; openSUSE and Fedora have RPMs in the openSUSE Build Service. The Debian package can also be used in Ubuntu, but a real Ubuntu version is no longer maintained.

The dependency list for installation from source is pleasantly short. Because it's a Python program and nothing needs to be compiled, version 2.7 (not 3!) of Python will usually be enough together with the associated developer files, usually found under the name *python-dev* or *-devel*. You can use the `python setup.py build` command to create the file structure, which you then perpetuate in the system with

```
sudo python setup.py install
```

At run time, PhotoFilmStrip still depends on other software. The MEncoder from the MPlayer portfolio is required for rendering the video. On top of that, version 2.8.12 or newer of wxGTK is also required – the newer versions based on GTK3 also work.

One of the Python modules PIL (*python-imaging*) or Pillow (*python-pillow*) take on various image-processing functions, depending on the defaults in the respective distribution. Version 3.0.0 of Pillow is currently useless. It wasn't possible to save newly created projects on Fedora 23 after adding pictures. Debian and its derivatives are still using the "real," but no longer actively supported Python Imaging Library – that's why this problem didn't come up in the test with Ubuntu 15.10.

The penultimate button opens a settings window for pixel-perfect adjustment of the motion paths (Figure 4). The *Location* specification relates to the upper left corner of the selection frame, and *Size* relates to the horizontal width. If you can't select precisely enough using the mouse, you can winkle out the very last grain

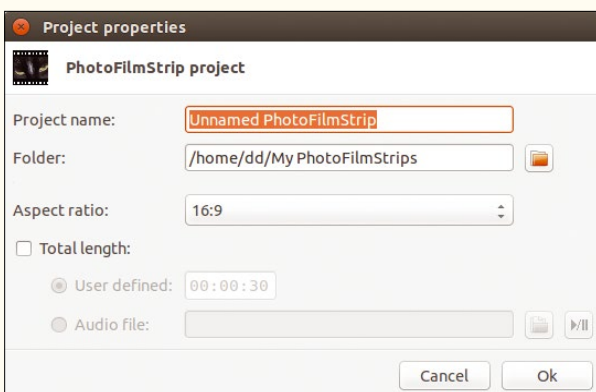


Figure 1: Lay the foundations for your new slideshow in the Project properties dialog.

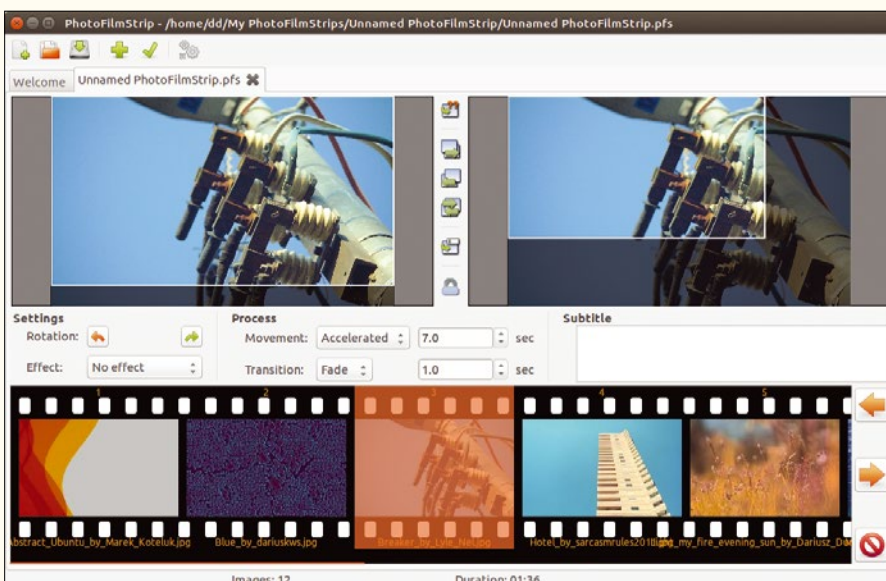


Figure 2: The main window appears tidy and clean after importing the image.

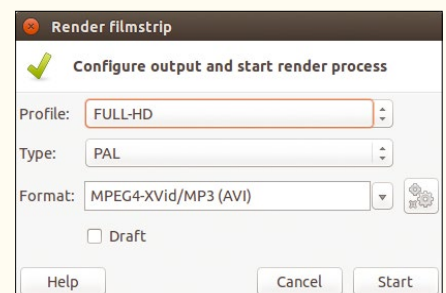
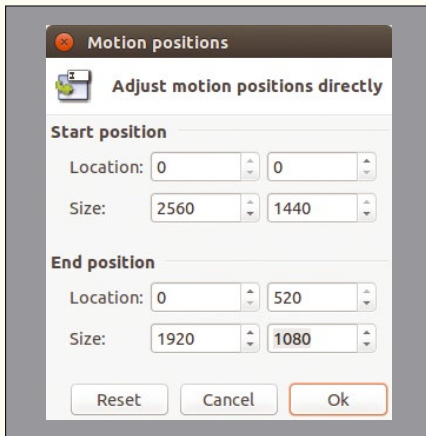


Figure 3: PhotoFilmStrip has no lack of formats.



**Figure 4:** Effects can be implemented even more precisely, thanks to fine adjustments.

of accuracy here. The lock icon at the bottom of the vertical button bar ultimately expands the operating range of the selection frame beyond the physical boundaries of the image. However, this creates black areas, which are usually more of a hindrance, but create additional leeway in certain cases.

Zooming and panning are far from the only possible options in PhotoFilmStrip. A few buttons under the output image offer ways to help get your image into shape. The arrow buttons should be pretty self-explanatory – they help put a misaligned image into the correct position. No other truly breathtaking image effects are hiding in the *Effect* drop-down, except options to convert to a black-and-white image or to a yellowed sepia image.

The *Movement* and *Transition* drop-downs to the right affect the transition between two images and the type of panning and zooming itself. The preset styles *Accelerated* and *Fade* already deliver attractive results. If necessary, you can change the values here or scale back the effect to the level of a conventional analog slideshow. Note, however, that the settings are only valid for the respective frame and can't be semi-automatically or fully automatically adopted for other images.

The filmstrip at the bottom of the window sorts your photos like a file manager when images are imported – usually alphabetically by file name. You can rearrange them if you want by simply dragging and dropping them using the mouse. It is easy to click the image to be edited, or you can toggle it in the filmstrip using the arrow keys on the right.

## Sounds

Presuming you're not wanting a high-resolution silent movie, you can add an audio file (*File | Properties*) when rendering. The length of the finished video is determined by the length of the audio file. By playing around with the length of the file and the number of images, you can influence the duration of the frame without having to intervene directly in the *Fade* parameters.

Although it isn't possible to record a comment directly into an audio file, you can use a little trick: First, render the video without sound. Then, when playing it, launch an audio recording program and record your comments to an audio file of any format. Next, render the video again, indicating the audio file. The comments and image will then always run in sync.

## Image Captions

An input field labeled *Subtitles* is available under the view of the target image on the right. Don't expect too much from it, however; the entries might be placed correctly

```
dd@dd-ubuntu14041d: /
dd@dd-ubuntu14041d:/$ photofilmstrip-cli --help
Usage: photofilmstrip-cli [options]

Options:
  --version          show program's version number and exit
  -h, --help        show this help message and exit
  -p PROJECT, --project=PROJECT
                   specifies the project file
  -o PATH, --outputpath=PATH
                   The path where to save the output files. Use - for
                   stdout.
  -t PROFILE, --profile=PROFILE
                   0=VCD, 1=SVCD, 2=DVD, 3=Medium, 4=HD, 5=FULL-HD
                   [default: 3]
  -n VIDEONORM, --videonorm=VIDEONORM
                   n=NTSC, p=PAL [default: p]
  -f FORMAT, --format=FORMAT
                   0=Single pictures, 1=MPEG4-XVid/MP3 (AVI),
                   2=MPEG4-XVid/AC3 (AVI), 3=MPEG(1/2)-Video (MPG), 4
                   =Flash-Video (FLV), 5=Motion-JPEG (AVI), 6=PyGame,
                   7=Cairo [default: 1]
  -d, --draft       enable draft mode - Activate this option to generate a
                   preview of your PhotoFilmStrip. The rendering process
                   will speed up dramatically, but results in lower
                   quality.

dd@dd-ubuntu14041d:/$
```

**Figure 5:** The command-line version of PhotoFilmStrip has just a few options.

in the appropriate image, but they are too small. It certainly isn't a replacement for proper subtitling software. There are no configuration options other than the text input itself. Color adjustments would be helpful but are not available. With the use of programs such as Gnome Subtitles or even a simple text editor to change the font color or formatting (bold, italic, underline), you can edit the subtitle file, which has a .srt suffix and is found in the same folder as the video after rendering.

## In the Terminal

A command-line option is normally a welcome addition for a graphical program. After all, batch processing large numbers of images in this way is often easier to manage.

With the command input `photofilmstrip-cli`, PhotoFilmStrip expects a project file and an output path (Figure 5). The terminal is thus only used for rendering a video, not for processing the imported images per se. It would have been useful to have special options here to be passed to MEncoder working in the background, but this isn't currently possible.

## Conclusions

PhotoFilmStrip is easy to use, works as expected, and produces high-quality videos, without having to intervene too much in the production process, and it has plenty of useful features if you want your video to be more elaborate. The subtitle function needs a bit of improving; otherwise, PhotoFilmStrip is fully operational. (See the "Alternatives" box for more information.)

Thanks to wxPython, the program is available in Windows as well as Linux and BSD, which is convenient for those who wanderer between worlds, allowing them to continue working on a project begun on one operating system, on a different operating system. ■■■

## ALTERNATIVES

PhotoFilmStrip isn't the only option. Imagination [4] is also capable of the Ken Burns effect. Although development seem to have stopped recently, the current version 3.0 also produces HD video.

DVD Slideshow [5] is another competitor that has unfortunately been in a deep sleep for four years. Added to this is the fact that it's a terminal program, which is not normally the best choice for video editing, and involves a steep learning curve. The package includes five Bash scripts that control external tools like FFmpeg, DVDAuthor, or ImageMagick via command-line options. The Ken Burns effect is realized in ImageMagick. By nature, ImageMagick isn't really suitable for

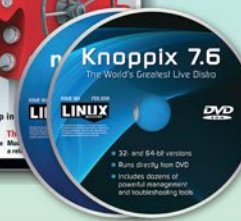
specifically editing individual images. Its advantages are rather in batch processing large numbers of images.

Also of mention is the repeatedly expressed desire in relevant forums and mailing lists for LibreOffice Impress to be equipped with such a feature. For now, however, you need to continue coping with the existing ways to spice up a presentation. Aside from the Ken Burns effect, a reasonable way for exporting as a video file is still missing. The only format currently available, Flash, is already on the decline, which means that the all-around office suite will hardly be competition for more specialized programs in the foreseeable future.

## INFO

- [1] Rostrum camera: [https://en.wikipedia.org/wiki/Rostrum\\_camera](https://en.wikipedia.org/wiki/Rostrum_camera)
- [2] PhotoFilmStrip: <http://www.photofilmstrip.org/>
- [3] Ken Burns: [https://en.wikipedia.org/wiki/Ken\\_Burns](https://en.wikipedia.org/wiki/Ken_Burns)
- [4] Imagination: <http://imagination.sourceforge.net/>
- [5] DVD Slideshow: [http://dvd-slideshow.sourceforge.net/wiki/Main\\_Page](http://dvd-slideshow.sourceforge.net/wiki/Main_Page)

# Risk-Free Trial!



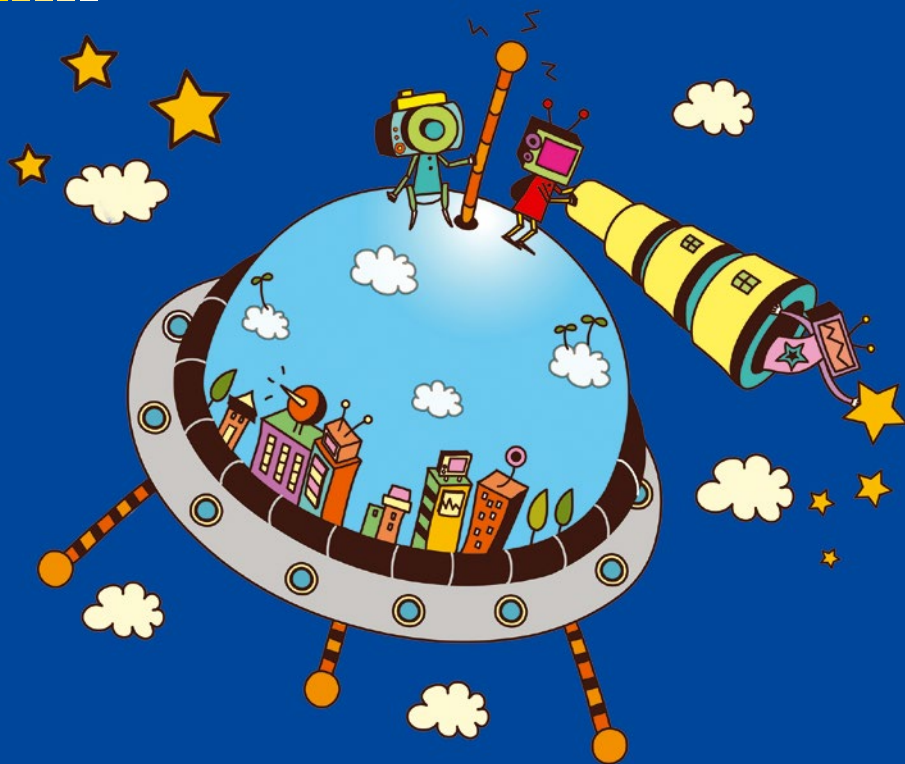
**GET IT NOW!**  
SAVE TIME ON DELIVERY WITH OUR PDF EDITIONS (DVD NOT INCLUDED)

**ORDER YOUR TRIAL NOW!**

UK £ 3, Europe € 3, USA / Canada US\$ 3, Rest of World (by Airmail) US\$ 9

[shop.linuxnewmedia.com/trial](http://shop.linuxnewmedia.com/trial)

Terms and conditions: <http://goo.gl/SSS0er>



## Extending the Stellarium virtual planetarium

# Star Power

Expand the Stellarium virtual planetarium with new objects and environments in just a few simple steps. *By Roland Pleger*

**T**he Stellarium planetarium program shows a realistic sky in 3D, calculates and displays star movements, and lets you prepare and evaluate observations. A previous article [1] introduced the software and looked at some of its capabilities. In this article, I focus on extending the software with plugins, additional materials, and little-known features.

### Installation

Most distributions offer Stellarium version 0.13.3 in their repositories, which is the version I discuss here. Version 0.14.2 was released January 2016 [2]. You can install the program on newer computers without problem; however, it isn't possible on older systems if they don't support OpenGL 2.1 or later. Your only recourse then is to install the legacy program, version 0.9.

### AUTHOR

Dr. Roland Pleger is a physicist and Wikipedia author. He believes that open source and open content prepares the way for a better world.

The directories for the program's files are based on the Linux standard: The data resides in `/usr/share/stellarium/`, and you should not change the data in this directory or its subdirectories. Only on rare occasions with a compelling need should you access this directory with root privileges. The program stores user-space data below `~/.stellarium/`. Naturally this directory is empty immediately after the installation.

To track notifications, start the program in the terminal by entering `stellarium`. The landscape often fills up the screen before you can finish reading all the output; if necessary, you can re-read the messages in the `~/.stellarium/log.txt` file.

Depending on the time of day, you'll see either a virtual, cloudless sky or the night sky (Figure 1). Vertical and horizontal bars with buttons appear when you move the mouse to the bottom left corner. If you click the small black triangles there, the bars stay on the screen permanently.

Clicking the rectangle with outward-pointing arrows in the center of the bottom bar (or pressing F11) toggles between full-screen and window mode; clicking the icon at the bottom right terminates the program.

You can determine the visible display detail using the top two icons in the vertical menu to select the location (top) and time of day (next from top). An observer in Casper, Wyoming, on March 4, 2016, would see the sky shown in Figure 1 looking southwest at 21:56.

### Constellations

The icons on the left in the bottom menu bar turn on the constellations. Asterism names differ depending on the culture chosen. With the *Western* default setting, names of stars and constellations in the northern hemisphere were derived from Greek and Arabic astronomy; in the southern hemisphere, names were assigned during medieval seafaring in the Age of Discovery.

To select the cultural basis of names, use the third icon from the top in the vertical menu column, click the *Starlore* tab, then check the *Use native names for planets from current culture* box at the bottom.

Constellations reside in the `/usr/share/stellarium/skycultures/western` subdirectory as 512x512-pixel black and white photos. In Figure 2, I replaced the



warrior Orion with a butterfly image from Wikipedia [3] (Figure 3). The stars that make up the belt now represent the body of the butterfly, and Rigel and Betelgeuse form the tips of the wings. Stellarium interprets black as transparent, and the PNG images do not need an alpha channel. The program also accepts images in JPG format, and it doesn't seem to require square images.

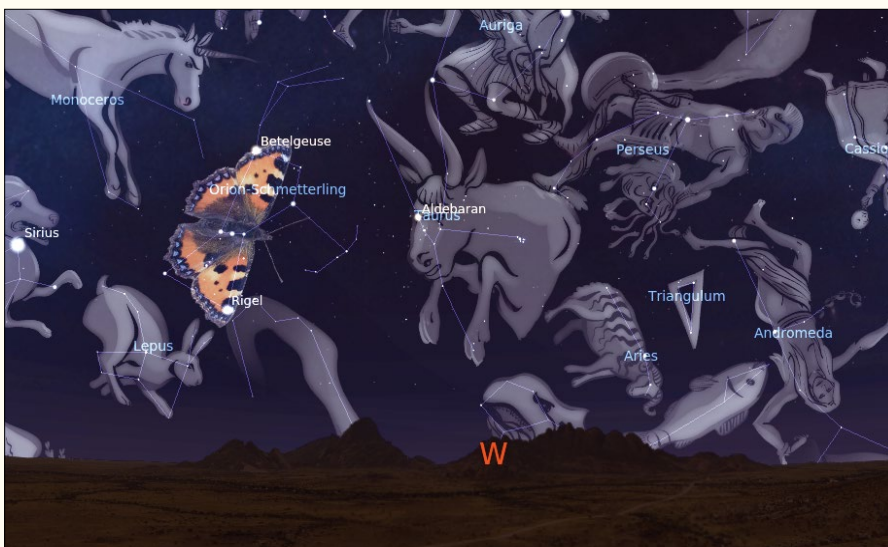
The easiest way to project the butterfly on the sky is to replace the image file `/usr/share/stellarium/skycultures/western/orion.png` with your own photo of the same name. A better approach, however, would be to expand the program without deleting anything. The "Own Constellations" section summarizes the steps for doing this.

## Your Own Landscape

Stellarium provides a series of landscape images over which the stars stretch. A sample image [6] is fine for experimenting with your own environment [7]. A Stellarium panorama image has edges that fit together seamlessly with a cylindrical equidistant projection (*Sky and viewing option window | Markings | Cylinder*) that covers 360 degrees in width and 180 degrees in height (i.e., a width-to-height ratio of 2:1). In this kind of projection, longitudes map to equally spaced vertical lines and latitudes map to horizontal lines.

To create your own panorama for Stellarium from an original image, you first need to change the width of the image to 4048 pixels without distortion. Then, embed the photo in a 4048x2024-pixel transparent background so that the horizon is in the middle (Figure 4). Any transparent areas on the bottom of the image must be colored; otherwise, stars will shine through. The sky, then, is made transparent. Finally, save the image in PNG format.

To integrate the new panorama in Stellarium, you must first create the `stellarium/landscapes/mountain` user directory. The name `mountain` refers to the new environment. Copy the edited `Panorama - spitzkoppe_4048.png` in this example - to `mountain/spitzkoppe_4048.png`.



**Figure 2:** Stellarium displays constellations on request, which you can replace with your own pictures as required. (Landscape GFDLv1.2 [4])



**Figure 1:** Stellarium usually starts in full-screen mode. The controls only appear if you move the cursor into the bottom left corner.

## OWN CONSTELLATIONS

To begin, copy the `/usr/share/stellarium/skycultures/western/` directory to `~/stellarium/skycultures/western`. To do this, create the `skycultures` subdirectory.

Now, rename the `western` subdirectory (e.g., to `westernm`), edit the `stellarium/skycultures/westernm/info.ini` file, and replace the name `western` there with the name of the new directory, `westernm`. Next, replace the `orion.png` image file with your own image.

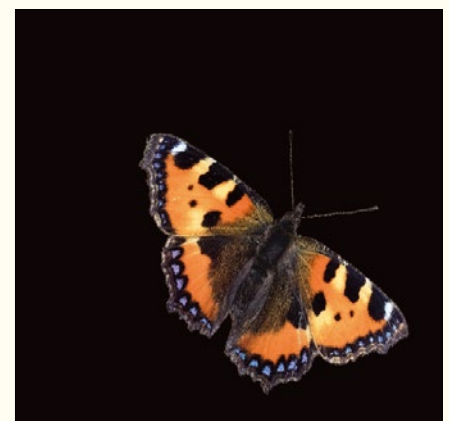
In `constellation_names.eng.fab`, change the line

```
Ori "Orion" _("Orion")
```

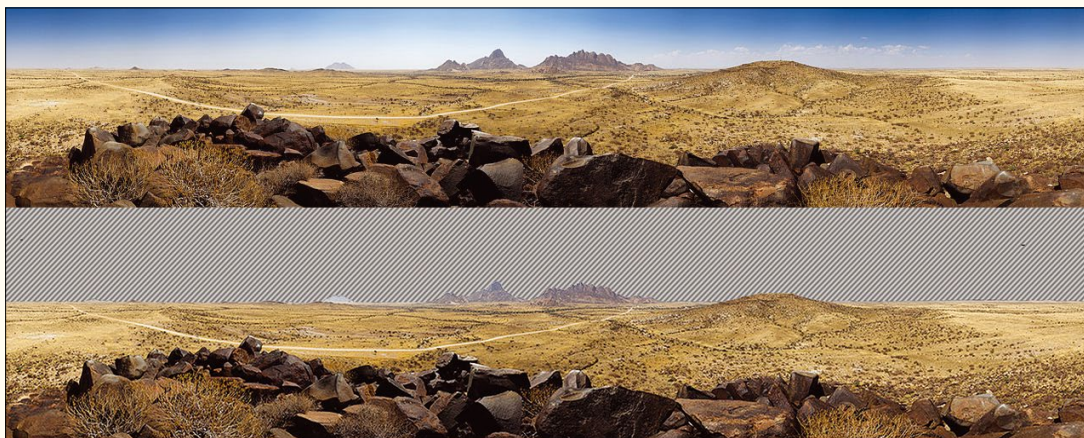
to:

```
Ori "Orion" _("Orion Butterfly")
```

The butterfly thus has its own name. After restarting, a new entry called *Westernm* can be found in the *Starlore* tab. If you select it, the butterfly appears in the sky.



**Figure 3:** This butterfly against a black background is suitable as a new image for the Orion constellation. (CC-BY-SA 4.0 [5])



**Figure 4:** A panorama that you can install in a few steps in Stellarium. (Top) original panorama. (Bottom) scaled image with transparent sky.

More data relating to the environment is found in the `mountain/landscape.ini` file (Listing 1). Line 5 specifies the projection, and line 6 the name of the image file. Lines 8 and 9 refer to images that the software places over the panorama, as required. They are commented out in this example. To integrate new images correctly, just create them in the

### LISTING 1: landscape.ini

```
01 [landscape]
02 name = mountain
03 author = Ikiwaner / Wikipedia
04 description = Spitzkoppe; source: https://commons.
    wikimedia.org/wiki/File:Spitzkoppe_360_Panorama.jpg
05 type = spherical
06 maptex = Spitzkoppe_4048.png
07 ;maptex_illum = Spitzkoppe_illum.png
08 ;maptex_fog = Spitzkoppe_fog.png
09
10 [location]
11 planet = Earth
12 latitude = -21d52'09"
13 longitude = +15d11'59"
14 altitude = 1000
```

same way as the panorama. It is important that the length-to-width ratio matches. Lines 11 to 14 define the location; but Stellarium doesn't evaluate this information. After restarting Stellarium, a new entry called *Mountain* appears in the *Sky and viewing options window* icon (third from the top) below the *Landscapes* tab.

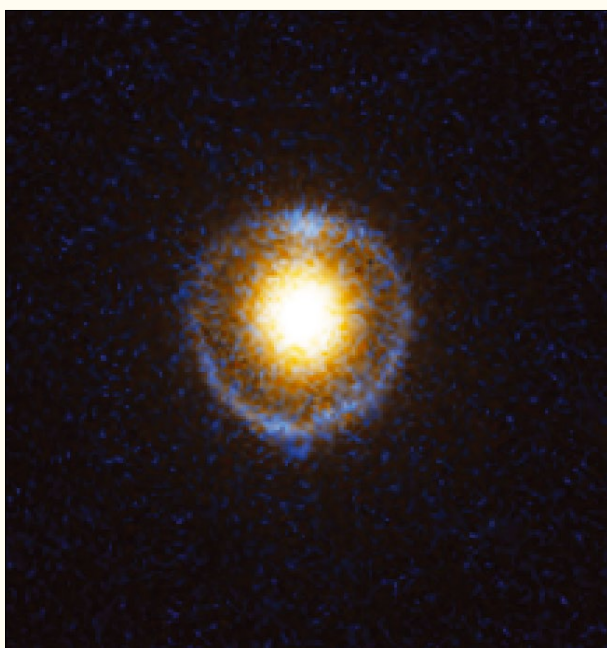
### Own Galaxies

Not only can you gaze at stars using Stellarium, but whole galaxies, too. To do so, you need to integrate high-resolution telescope images, which you can find in the `/usr/share/stellarium/nebulae/default` directory. However, version 0.13 is only familiar with the default subdirectory, so you must make a copy of the high-res images in the new directory, `~/stellarium/nebulae/default`.

Online, you can find an updated compilation [8] to save in the new directory. The error messages in the logfile or in the terminal when launching the software point out that the software isn't able to assign some names correctly, with mismatches in some cases and missing files or file name extensions in others. At the time of publication, I still hadn't received an answer to my query as to whether some of the images might be protected by copyright.

Stellarium is already familiar with all large galaxies; however, it is missing those that only appear intermittently because of their distance from Earth. One example is the object `SDSS J162746.44-005357.5`. This object shows two galaxies that lie exactly on an axis with the earth but are several billion light years away from each other. This arrangement causes the front galaxy to bend the light of the back galaxy, in line with the general theory of relativity, forming an Einstein ring [9] (Figure 5). The front galaxy appears as a point, the rear one as a ring – here with a diameter of two arcseconds. For comparison, terrestrial telescopes cannot capture less than one arcsecond because of fluctuations in the atmosphere; however, the Hubble Space Telescope manages 0.05". Stellarium is already familiar with the object of interest. If you enter `SDSS J162746.44-005357.5` into the search window, the program will guide you to the right place, although nothing is there yet to be seen. Stellarium doesn't do the search itself; instead, it asks the SIMBAD database [10].

Therefore, you have to place the image of the Einstein ring into the correct position and at the correct size in the Stellarium night sky. The basic procedure is the same as for the constellations; the "Fitting the Einstein Ring" box lists the specific steps. The main work this time involves determining the coordinates. A rough estimate is enough to demonstrate the principle: The diameter of the



**Figure 5:** Hubble image of Einstein ring `SDSS J162746.44-005357.5`. Image: NASA, ESA, A. Bolton (Harvard-Smithsonian CfA) and the SLACS Team.

## FITTING THE EINSTEIN RING

To begin, create the `~/stellarium/nebulae/default` subdirectory and copy the image of the Einstein ring there. Next, adjust the `~/stellarium/nebulae/default/textures.json` file by expanding the JSON list to include:

```
{ "imageCredits": { "short": "Einstein ring SDSS J162746.44-005357.5; Source NASA/ESA http://hubblesite.org/gallery/album/exotic/pr2005032g/" }, "imageUrl": "EinsteinSDSSJ162746.png", "worldCoords": [[ [-113.0556, -0.9001], [-113.0572, -0.9001], [-113.0572, -0.8985], [-113.0556, -0.8985]] ], "textureCoords": [[ [0,0], [1,0], [1,1], [0,1]] ], "minResolution": 0.28, "maxBrightness": 16.0 }
```

After restarting, you'll find the image in the sky.

Einstein ring is two arcseconds. Figure 5 therefore approximately covers an extension of six arcseconds. According to SIMBAD, the coordinates of the Einstein ring are 16hr 27min 46.447sec right ascension, -00°53'57.56" declination.

After converting the coordinate values to decimal degrees [11], the vertices of the image are the coordinates in the `worldCoords` line (in the boxout). The `imageUrl` line lists the name of the image file and `maxBrightness` the brightness. The value `mag 17` would be correct. I chose `mag 16`. Stellarium suppresses objects with a magnitude smaller than 16 (higher numbers represent smaller magnitudes).

The data in the `textures.json` text file are in JSON format. Syntax errors such as comma or bracket mistakes will prevent all the images from being integrated, resulting in an error message in the terminal:

```
WARNING: Can't parse JSON description: "/home/rp/.stellarium/nebulae/default/textures.json"
```

You can check position and size of the added object by loading the *Angle Measure* plugin at startup (*Configuration window | Plugins*). After a restart, the tool will appear as an angle symbol in the bottom menubar. The diameter of the ring should be two arcseconds (Figure 6), if you have estimated the edge coordinates correctly. Figure 7 shows guide stars that can help find the tiny object.



Figure 6: If everything worked, you'll be able to find the Einstein ring as an image in the sky with the correct diameter.

## Conclusions

Normal galaxies stretch out over many arcminutes. In the Einstein ring example, Stellarium illustrates the persistence required to find tiny structures in the sky like Einstein rings with a diameter of only a few arcseconds. Presumably, only very few users manage to come across the object by chance.

All told, it is pleasingly easy to familiarize yourself with the stars using Stellarium. By making a few small modifications, you can quickly adapt the software to include new environments and objects. ■■■

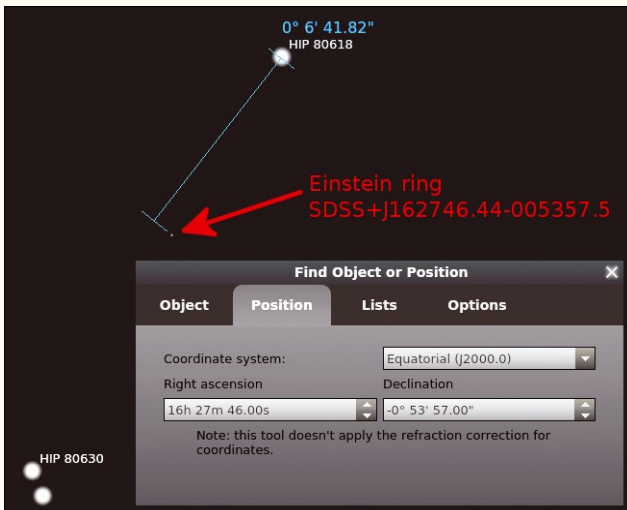


Figure 7: Use the guide stars to find the Einstein ring.

## INFO

- [1] "Explore the Night Sky with Stellarium" by Karl Sarnow, *Ubuntu User*, issue 16, 2013, <http://www.ubuntu-user.com/Magazine/Archive/2013/16/Explore-the-night-sky-with-Stellarium>
- [2] Stellarium at Launchpad: <https://launchpad.net/~stellarium/+archive/ubuntu/stellarium-releases>
- [3] Small tortoiseshell (butterfly) by Zeynel Cebeci: [https://commons.wikimedia.org/wiki/File:Aglais\\_urticae\\_-\\_Small\\_tortoiseshell\\_02.jpg](https://commons.wikimedia.org/wiki/File:Aglais_urticae_-_Small_tortoiseshell_02.jpg)
- [4] GNU Free Documentation License, version 1.2: [https://commons.wikimedia.org/wiki/Commons:GNU\\_Free\\_Documentation\\_License,\\_version\\_1.2](https://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License,_version_1.2)
- [5] Attribution-ShareAlike 4.0 International: <https://creativecommons.org/licenses/by-sa/4.0/deed.en>
- [6] 360° Panorama over the small Spitzkoppe, the Spitzkoppe, and the Erongogebirge, Namibia, by Ikiwaner: [https://commons.wikimedia.org/wiki/File:Spitzkoppe\\_360\\_Panorama.jpg](https://commons.wikimedia.org/wiki/File:Spitzkoppe_360_Panorama.jpg)
- [7] Code and images: <ftp://ftp.linux-magazine.com/pub/listings/magazine/186>
- [8] File with updated galaxies: <http://barry.sarcasmogerd.com/stellarium/uploads/stellariumnebula-1.ZIP>
- [9] Einstein ring: <http://hubblesite.org/newscenter/archive/releases/2005/32/image/g/>
- [10] SIMBAD Astronomical Database: <http://simbad.u-strasbg.fr>
- [11] Conversion of right ascension and declination in degrees: <https://www.swift.psu.edu/secure/toop/convert.htm>

# SUBSCRIBE NOW



## The New IT

New tools, new threats, new technologies...Looking for a guide to the changing world of system administration?

[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

# AND SAVE 30%

## Explore the new world of system administration

It isn't all Windows anymore - and it isn't all Linux. A router is more than a router.

A storage device is more than a disk. And the potential intruder who is looking for a way around your security system might have some tricks that even you don't know. Keep your network tuned and ready for the challenges with the one magazine that is all for admins.

Each issue delivers technical solutions to the real-world problems you face every day. Learn the latest techniques for better:

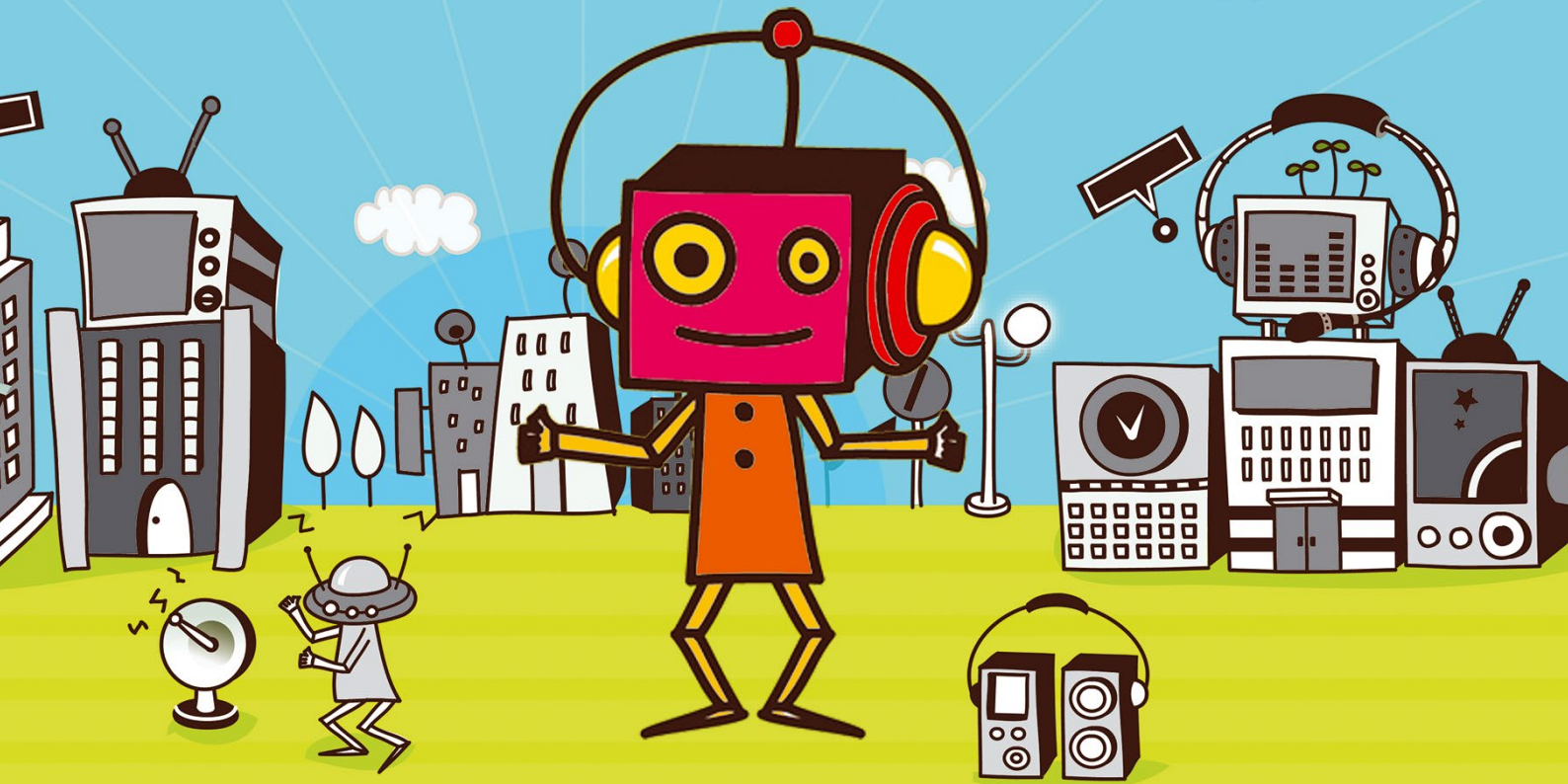
- network security
- system management
- troubleshooting
- performance tuning
- virtualization
- cloud computing

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

**REAL-WORLD  
PROBLEMS  
SOLVED!**

**ADMIN**  
Network & Security

[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)



## Using fuzzy searches with tre-agrep

# A Grep Replacement?

Tre-agrep has all of grep's functionality but can also do ambiguous or fuzzy searches without deep knowledge of regular expressions. *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>

**G**rep [1] is a standard command-line tool. It searches files for regular expressions, then displays any lines that include a match. In expert hands, grep can be a flexible tool, but gaining expertise can take years of practice. Nor do related commands like egrep [2] or fgrep [3] make grep any easier to use. For these reasons, those lacking expertise might want to check out TRE [4], which includes a reimplement of agrep (approximate grep) [5] as a command-line utility. tre-agrep is a grep-like tool that has all of grep's functionality but can also do ambiguous or fuzzy searches that are much easier to learn.

Grep and tre-agrep share similar options, such as --ignore-case and --count. However, the logic of their searches can be different. (I say "can be" because often both commands have multiple ways of getting the same result.) To give a simple example, imagine that you are searching for files that contain both "Linux," and "Linus." Using grep, you would probably use regular expressions one way or the other. Probably the simplest would be:

```
grep 'Linu.' *.txt
```

Here, the period in `practi.e` indicates that any character can be substituted for it, giving results with both "practice" and "practise." This use of a regular expression is relatively simple, but it must be entered and positioned accurately. If it were more complicated, newer users might be put off by a series of familiar and unfamiliar characters used with non-standard syntax.

```
bruce@nanday:~/Notes$ tre-agrep -l 'Linux' *.txt
licensing.txt:It is released under the second version of the GNU General Public L
icense (GPL). When the third version of the GPL was written, Linus Torvalds, as l
eader of the kernel project, refused to changed the license..bruce@nanday:~/Notes
$
```

**Figure 1:** Set to search for “Linux” or a word with a one-character difference, tre-agrep locates “Linus.”

By contrast, with tre-agrep, the command is more likely to use an option for ambiguity:

```
tre-agrep -l 'Linux' *.txt
```

The option here means that the results should include those with one character different from the string “Linux” – a command that requires both less precision and less user knowledge, but perhaps at the price of more irrelevant results (Figure 1). Moreover, the entered command would find typos anywhere in the string, not just in the second-to-last letter. Notice, too, that, both commands begin displaying the results with the name of the file and end with the current account and the file path.

Usually, tre-agrep displays the first result that matches the search. If you want more than the first search result, you can specify the number of errors. For example, if you set the command to look for results with four errors, results with three errors will not show, so you might want to make several searches with minor differences.

The original version of agrep was developed in 1988-2001 by Udi Manber and Sun Wu. Originally written for Unix, this version was widely ported to other operating systems, but it’s rare in Linux distributions, because for years, it was released under a non-free license. Since 2014, it has been released under the ISC Open Source License [6], but either the new license is not recognized as free, or the change has gone unnoticed, because Debian still includes it in the non-free section of its repositories.

Today, the most common version is tre-agrep, written in 2002-2004 by Ville Laurikari. Tre-agrep uses a different library from the Manber and Wu version and is released under a BSD license. Most distributions include it in their repositories, although not as part of the default installation.

When used without any options, tre-agrep’s output is identical to grep’s. However, it is the options that make tre-agrep’s results different. All tre-agrep’s options come under one of three categories: options for approximations, regular expressions, and output filtering and formatting.

### Options for Setting Approximations

Approximations or fuzzy logic are at the heart of tre-agrep. The man page describes the number of differences as the cost (based on the Levenshtein distance [7]), which is a count of the number of characters that a command using approximation options can depart from the precise string entered in the command. By default, a missing, an

**TABLE 1:** Common Regular Expressions

Character Keys	Meaning
.	Any single character
*	Any any number of characters, or none
^	The following regular expression at the start of a line
\$	The following regular expression at the end of a line
[ ]	Any of the characters in the brackets
\	Turn off the next character’s meaning as a regular expression
\<	Characters at the start of a word
\>	Characters at the end of the word
?	One or zero instances of the preceding regular expression

extra, or a substituted character all have a cost of 1, although you change these costs with `--delete-cost=NUM (-D NUMBER)`, `--insert-cost=NUMBER (-I NUMBER)`, or `--substitute-cost=NUMBER (-S NUMBER)` to reflect your needs.

The concept of cost is used without explanation in the command's help, but its usefulness of the concept soon becomes clear enough. Cost is a way to judge output records and sort through them. Most of the time – although not always – the lower the cost, the closer the result is likely to be to your intention. Conversely, the higher the cost, the greater the chance that an output record is relevant. However, if you know, for example, that relevant results are most likely to be a substitution, you can set the cost of substitutions to 0, lowering their cost and making them easy to find with an output option such as `--best-match` or `--show-costs` (see below).

If you are not interested in changing the cost of approximations, the concept of fuzzy results is straightforward. The most useful option for approximations is `-#`, which should be replaced in a command by a digit between 0 (an exact match) and 9 errors – with “error” being the name for any deviation from the string entered as part of the command. You can also further filter output records via `--max-errors=NUMBER (-E NUMBER)`. These are simple but powerful options, and they are easily remembered.

### Options for Regular Expressions

Regular expressions are search patterns, in which characters stand for other groups of characters in files, the contents of files, or locations in a file [8]. Both `grep` and `tre-agrep` can use the same standard set of regular expressions (Table 1).

Regular expressions can be entered directly into the string part of the command. However, ambiguity sometimes can be reduced by using the option `--regexp=PATTERN (-e PATTERN)`. In particular, this option can be useful if a search includes a hyphen (-), which might be misinterpreted as introducing an option, or a forward slash (/), which might be read as introducing a directory.

As in `grep`, a search for regular expressions can be refined in several ways. With `--ignore-case (-i)`, a regular expression treats lower and upper case letters the same, both in a search pattern and in the names of input files. With `--literal (-k)`, the search pattern is read as though it has no special characters in it. You can also use `--word-regexp (-w)` to match only whole words, or `--invert-match (-v)` to select records that do not match the regular expression you entered. These refinements can help filter results, but they can add another level of complexity; therefore, unless you have a special need, you might first prefer to focus only on using regular expressions until you are comfortable with basic patterns.

### Output Options

Some of `tre-agrep`'s output options are less well known than those for approximations, but some can be almost as useful. Some are identical to `grep`'s, such as `--quiet (-q)`, which suppresses output, letting you know only that a match has been found, or `--files-with-matches (-l)`, which lists only the names of files with matching re-

```
bruce@nanday:~/Notes$ tre-agrep -2 --count 'practice' *.txt
Licensing.txt:0
new-notes.txt:1
notes2.txt:0
notes.txt:1
```

**Figure 2:** The `--count` option lists the number of matches in each file in the directory.

```
new-notes.txt:3-11:In pratcice, grep requires the memorization of regular epressi
ons to be used efficiently. tnotes.txt:3-11:In pratcice, grep requires the memori
zation of regular epressions to be used efficiently. tbruce@nanday:~/Notes$ █
```

**Figure 3:** Here, the `--show-position` option shows two matches in nearly identical files, each of which starts three characters from the start of a file and ends 11 characters from the start.



sults. Still another option shared with grep is --count (-c), which only tells you the number of matches in each file, but does not display them (Figure 2).

However, by far the most useful option for filtering results is --best-match (-B), whose option displays only the records with the lowest cost – that is, those closest to the string you entered in the command. By using this option, especially with approximations, you can reduce the results through which to scroll, although possibly at the cost of missing serendipitous results.

Another way to judge results is to add --show-cost (-s), which displays the cost directly after the file name at the start of the result. By seeing how far a result differs from the string you enter, you might be able to judge each result's reliability and usefulness.

Other output options format rather than filter results. For example, --color (--colour) is almost always useful, because it highlights results in the output strings, using the GREP\_COLOR environment variable. Similarly, you can use --show-position (Figure 3) to prefix each output record with the start and end of the record (the first character of the record and the first character after the match). You might also help organize results by prefacing each output reference with the name of the file in which it is located, using --with-filename (-H). As you continue with your work, you might also find it useful to number each output record by adding the option --record-number (-n).

## Choosing tre-agrep

Note that tre-agrep is not a complete replacement for grep. Grep has several options that tre-agrep lacks, including options for defining how directories and binary files are handled, as well as the ability to stop after a set number of results. Such options would be equally useful in tre-agrep.

Still, in typical searches, tre-agrep can be a drop-in replacement for grep, and because its ability to handle regular expressions is roughly equal to grep's, that should not be surprising.

However, what truly sets tre-agrep apart is not only its ability to handle approximations, but the simplicity with which approximations are implemented. Aside from a few basic examples, to get the most from regular expressions requires neither extensive memorization nor frequent consultations with man or web pages. Neither alternative holds much appeal for average users.

By contrast, approximations in tre-agrep usually get results without the need to be familiar with regular expressions. For that reason alone, you might consider tre-agrep as grep's replacement. ■■■

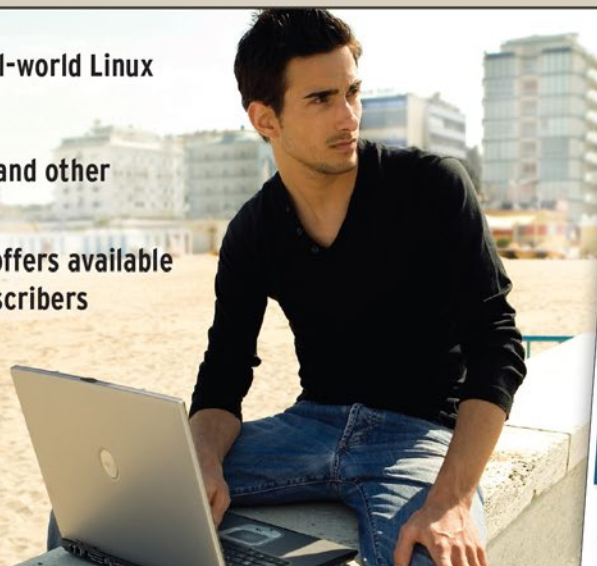
## INFO

- [1] grep: <https://en.wikipedia.org/wiki/Grep>
- [2] egrep: [http://linux.about.com/library/cmd/blcmdl1\\_egrep.htm](http://linux.about.com/library/cmd/blcmdl1_egrep.htm)
- [3] fgrep: [http://linux.about.com/library/cmd/blcmdl1\\_fgrep.htm](http://linux.about.com/library/cmd/blcmdl1_fgrep.htm)
- [4] TRE: <https://github.com/laurikari/tre/>
- [5] agrep: <https://github.com/Wikinaut/agrep>
- [6] agrep copyright: <https://raw.githubusercontent.com/Wikinaut/agrep/master/COPYRIGHT>
- [7] Levenshtein distance: [https://en.wikipedia.org/wiki/Levenshtein\\_distance](https://en.wikipedia.org/wiki/Levenshtein_distance)
- [8] Regular expressions: <http://www.cyberciti.biz/faq/grep-regular-expressions/>

# LINUX UPDATE

Need more Linux? Our free Linux Update newsletter delivers insightful articles and tech tips to your mailbox twice a month.

- Original articles on real-world Linux
- Linux news
- Tips on Bash scripting and other advanced techniques
- Discounts and special offers available only to newsletter subscribers



Ft Photography, Fotolia

[www.linuxpromagazine.com/mc/subscribe](http://www.linuxpromagazine.com/mc/subscribe)

## Developing certification and training for FOSS managers

# What Should You Know?

maddog examines what today's engineers and product managers need to know about Free and Open Software. *By Jon "maddog" Hall*

**A**t the Linaro Conference in Bangkok, one of the keynotes was dedicated to "License Compliance." The speakers were two people I have known for a long time: Shane Coughlan, from the Free Software Foundation, Europe, and Harald Welte, who has been associated with Free Software since his teenage years and who (as Harald himself pointed out) "is no longer a teenager ...."

The topic of license compliance has been discussed many times. After Harald explained why he personally started spending his own time and money pursuing companies who did not follow the GPL licensing, he was asked what could be done to make this less of an issue, but he had no real answer.

I stood up and volunteered an answer: "It will stop when the business people of tomorrow are taught the issues of Free and Open Source Software [FOSS] in the universities of today."

I recently started working with universities to train computer engineering students in the use and creation of FOSS. Some universities are incorporating this training into their mandatory engineering curriculum. However, some universities are treating the topics as electives (allowing their students to opt out of the training) or not offering the training at all, which often means that organizations like Linaro have to develop their own training to take good "closed source" computer engineering students and turn them into good Free and Open Source engineers.

Modern-day software engineers should know something as basic as using a distributed source code control system, or making sure that the licenses of the code being integrated into a product are compatible with the other code being using, or even understanding the business model of the final product.

However, ignorance of these issues does not end with engineers and often is magnified by managers who were trained at university 30 years ago, before the concepts of Free and Open Source Software were widely known.

Such topics also include managing a product through its life cycle when some or most of the code comes from a community of people not necessarily tuned to the needs of a corporation but instead have metrics that determine only when code is worthy of being released. Most likely FOSS

business models were not taught to these managers when they were students – only the business models of proprietary software.

The reality, however, is even worse because upper-level managers, such as CTOs or board members, do not even know what questions to ask when hiring the managers who in turn will be hiring (or trying to hire) FOSS software engineers. Because it is unlikely that these upper-level managers will be going back to university to learn these skills, it is probably worthwhile to develop training specifically aimed at providing the information to this level of management.

Of course, many books, magazine articles, and web pages discuss the subject of how to be a good "FOSS manager." (I have written many of these articles myself.) However, the material tends to be scattered, and much of the problem is that these managers are not even aware of what they do not know.

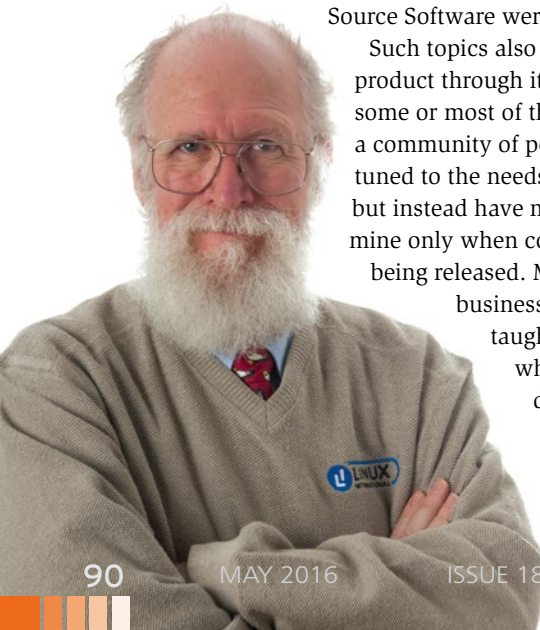
This situation mirrors the state of affairs in 1999 when we formed the Linux Professional Institute (LPI). We then had upper-level managers who needed Linux professionals but did not even know the questions to ask that would help them determine the competency of the people applying for jobs. LPI studied the issues, created the objectives of each level of professionalism, and then created the tests that certified that knowledge. Perhaps that type of certification could help in the hiring of FOSS managers today.

This brings me to the question: What knowledge should a manager of FOSS people and products have? I have mentioned software licenses, but what about issues of "upstreaming code"? Surely a manager of FOSS should know what that means and understand not only the legal issues but also the value of upstreaming code to the product itself. Should the manager be familiar with examples of FOSS business plans? Should the FOSS manager know that the schedules around FOSS projects might be more flexible than they are used to having with paid, dedicated engineers? Do they know that collaboration with their competitors is expected, and not just a nice saying?

I would like LPI to develop a certification for Managers of Free and Open Source Software and to have the FOSS community help determine what should be tested in that certification. I think it might make everyone's life easier in the long run, and from the number of engineers taking pictures of Harald's slides, I think many of the engineers might agree. ■■■

### THE AUTHOR

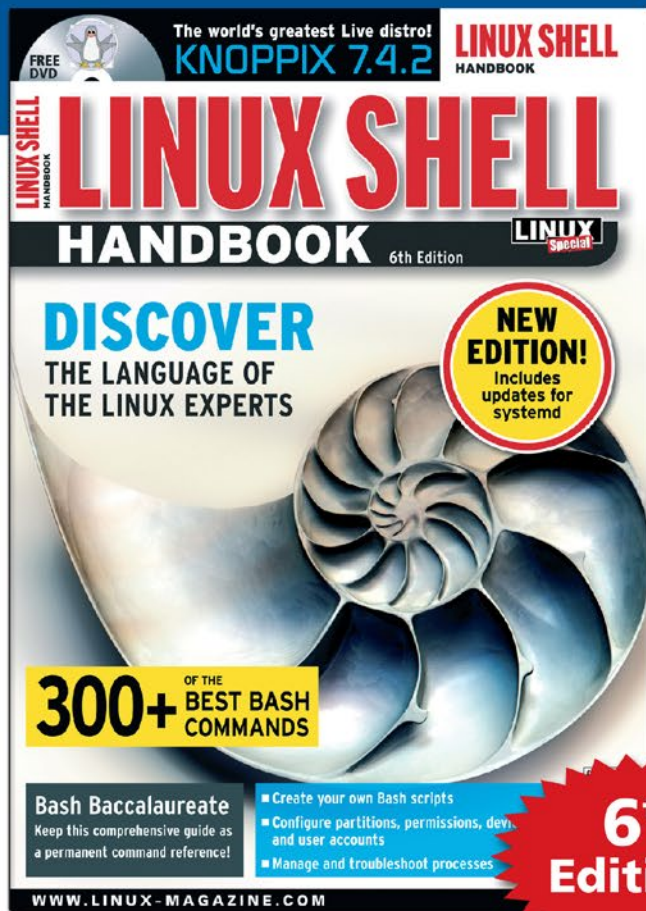
Jon "maddog" Hall is an author, educator, computer scientist, and free software pioneer who has been a passionate advocate for Linux since 1994 when he first met Linus Torvalds and facilitated the port of Linux to a 64-bit system. He serves as president of Linux International®.



Shop the Shop

[shop.linuxnewmedia.com](http://shop.linuxnewmedia.com)

# EXPERT TOUCH



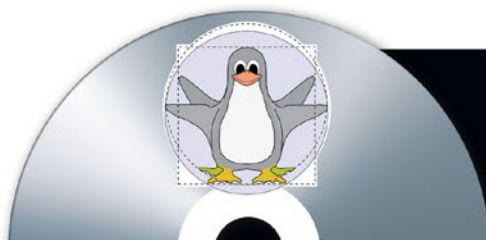
Linux professionals stay productive at the Bash command line – and you can too!

The Linux Shell special edition provides hands-on, how-to discussions of more than 300 command-line utilities for networking, troubleshooting, configuring, and managing Linux systems. Let this comprehensive reference be your guide for building a deeper understanding of the Linux shell environment.

You'll learn how to:

- Filter and isolate text
- Install software from the command line
- Monitor and manage processes
- Configure devices, disks, filesystems, and user accounts
- Troubleshoot network connections
- Schedule recurring tasks
- Create simple Bash scripts to save time and extend your environment

The best way to stay in touch with your system is through the fast, versatile, and powerful Bash shell. Keep this handy command reference close to your desk, and learn to work like the experts.



**FREE DVD INSIDE!**

The world's greatest Live distro!

**KNOPPIX 7.4.2**

ORDER ONLINE:

[shop.linuxnewmedia.com/specials](http://shop.linuxnewmedia.com/specials)

# Zack's Kernel News

Chronicler Zack Brown reports on the latest news, views, dilemmas, and developments within the Linux kernel community.

By Zack Brown

## Software RAID Maintainership

Neil Brown has stepped down as the software RAID maintainer after 15 years at the post. He explained that he'd been losing interest in the project lately and had not been able to attract any potential replacements. As a result, he decided not to push too hard for a clean handoff, but instead to simply step down and allow the vacuum to be filled by whoever ends up having the greatest need to fill it.

In terms of how he envisioned software RAID maintainership in the future, Neil said, "I think it would be great if a (small) team formed to continue to oversee md [Multiple Device driver] rather than just a single individual (or maybe the dm [Device Mapper] team could extend to include md??). If I had managed to be part of a team rather than "going it alone" for so long, I might feel less tired of the whole thing now." And he added:

*A few people have been mentioned to me in earlier less-public conversations. Any of them may well be suitable, but I would rather they named themselves if interested.*

*So, I'm hoping to get one or more volunteers to be maintainer:*

- to gather and manage patches and outstanding issues,
- to review patches or get them reviewed
- to follow up bug reports and get them resolved
- to feed patches upstream, maybe directly to Linus, maybe through some other maintainer, depending on what relationships already exist or can be formed,
- to guide the longer term direction (saying "no" is important sometimes),
- to care, but also to be aware that maintainership takes real effort and time, as does anything that is really worthwhile.

*This all applies to mdadm as well as md (except you would ultimately \*be\* upstream for mdadm, not needing to send it anywhere). Even if a clear team doesn't form, it would be great if different people maintained mdadm and md.*

*One part of the job that I have put a lot of time in to is following the linux-raid@vger.kernel.org list and providing support. This makes people feel good about md and so more adventurous in using it. Consequently, I tend to hear*

*about bugs and usability issues nice and early (well before paying customers hit them in most cases), and that is a big win.*

*In recent times, I've been doing less of this and have been absolutely thrilled that the gap has been more than filled by other very competent community members. Not developers in particular, but a number of md users have been providing excellent support. I'd particularly like to highlight Phil Turmel who is very forthcoming with excellent advice, but he is certainly not the only one who deserves a lot of thanks. So "Thank you" to everyone who answers questions on linux-raid.*

*This would be a good place for any future maintainer to hang out to receive wisdom as well as to provide support.*

*I will still be around. I can certainly help out in some sort of mentor role, and can probably be convinced to review patches and comment on designs. But I really want to head towards spending less time on md (there are so many other interesting things to learn about).*

*So: if anyone is interested – please announce yourself, ask questions and start doing things. I have no clear idea about how a transition will happen. That is really up to you (plural). Take the bull by the horns and start \*being\* a maintainer(team). I won't get in your way, and I'll help where I can.*

There were a lot of replies from people who were either interested in joining a maintainer team or were interested in bowing out gracefully. Phil Turmel hinted that he might join in, though he said his kernel skills might need polishing. Shaohua Li said that he'd be interested in joining the team and had support from his employer, Facebook. Jonathan Brassow suggested that a bunch of Red Hat employees might be interested in joining a maintainer team. Jes Sorensen also offered to be a part of the team and also endorsed Shaohua as an important potential teammate. Artur Paszkiewicz also volunteered to join the team.

Neil was very happy to see this level of interest and offered some parting advice: "The first question is where do you send your patches to get the appropriate review and upstream acceptance. Alasdair or Mike (DM), Jens (Block), Andrew Morton (anything), and Linus (everything) are all defensible choices

## 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**.

for upstreaming (I've submitted through Andrew in the past, but through Linus exclusively once I figured out git). That is really something you and they would need to negotiate though."

And, he added, "If you have specific questions about anything (md internals, git workflows, whatever) do feel free to ask. I am trying to write some comprehensive documentation on the internals of MD – though it is slow going. Once that is mostly done, I'll make it available and it should at least serve as a good structure in which to place answers to further questions." Finally, Neil said, "I plan to submit a pull request to Linus for the 4.5 merge window and then stop queuing patches." And, thereafter, he planned to "only look at other patches if they already have an independent 'Reviewed-by'."

In a concluding post, Neil announced, "Jes Serensen will be coordinating mdadm (once I make a release ... tomorrow?) and Shaohua Li will be looking after the kernel/md side. ... Only one person (at a time) can queue patches, but several can collaborate at development and support and bug fixing and testing." Jes then announced the creation of a new mdadm Git repository [1].

## Kernel Documentation System Overhaul

Jonathan Corbet posted a set of patches – intended only as proof-of-concept so far – to convert the Linux kernel to use AsciiDoc for its documentation, instead of DocBook. This would make it easier for contributors, who could use a plain-text markup language instead of XML. It would also allow the makefile to produce HTML, man pages, PDFs, and a lot of other documentation formats. Jonathan posted a link to the Git repository [2] for anyone wanting to tinker around with it.

Jani Nikula announced that he'd been doing some work that overlapped with Jonathan's (Jonathan's tree actually included a patch from Jani, which was the impetus for the whole project). Jani said, "I first took roughly the same approach as you did. I was really impressed with the speed and the beauty of the produced HTML. The trouble is, neither asciidoc nor asciidoctor can produce chunked (split to several pages) HTML directly. This is a showstopper for the gpu document which turns into 1.3MB of HTML, which looks pretty but is a pain to navigate. To do chunked output, you have to output DocBook and handle that like we do now. So while I would like to have asciidoc generate HTML directly for speed and beauty, I ended up going the asciidoc to DocBook path. The upside is all the output formats are supported."

Jani added, "One significant difference between our approaches is that I ditched docproc out of the equation. Instead of having the docproc ! directives in the asciidoc, I opted for using asciidoc's native include macro, with specially crafted filename suffixes to specify what parts of the source to include. Those files are then generated as intermediate asciidoc files using kernel-doc, with dependencies set right. There's a bunch of scripting involved, but it's pretty straightforward."

Daniel Vetter voiced some concern about the need to parallelize doc creation. The sheer size of Linux kernel documentation made this an important issue and related to Jani's consideration of "chunked" output.

Jani replied: "asciidoc -> HTML is about as fast as asciidoc -> XML, and with parallel kernel-doc, it really is fast. Sadly, the XML-HTML part still takes forever.

"I just want to emphasize that we can get parallel kernel-doc with either pipeline. It is also possible to enable both pipelines, i.e.



have a fast path HTML generation with few external dependencies and the Swiss army knife slow path with XML.”

Jonathan was excited to collaborate with Jani, saying, “Overlapping work is just how this kernel thing works :)” He proceeded to look over Jani’s work so far, and said:

*I would \*really\* like to get the XML step out of the processing path if possible. It adds complexity, makes it harder for others to build the docs, makes things more fragile, and slows it all down. It seems to me that it should be possible to do that.*

*The issues, it seems, are splitting the output files and format support. The latter isn’t really an issue, I don’t think; there are tools to do all kinds of format conversions. The only one that’s even slightly weird is man, and kernel-doc already has some (unused, I think) provisions for doing that. We could generate man pages directly without much pain.*

*For HTML-page splitting, we can see if the tools can help us, consider splitting the template files, or do the splitting in a post-processing step. Docproc (or whatever replaces it) could also maybe do that work. It doesn’t seem to me something that should force the inclusion of an entire XML-based processing step.*

Jonathan added:

*I have some thoughts on the whole thing ...*

- *I would like to format directly to HTML if at all possible. It seems it should be possible to get a table of contents into the files, and the feedback I got was that a TOC would be enough for navigation – it would not be necessary to split the files at that point. We might still want to try to figure that out too, though. In any case, this isn’t a show stopper, in that we can change it anytime if a better way shows up. But I’d like to have it in mind.*
- *Asciidoc templates and processing should happen in a new directory (perhaps imaginatively called “asciidoc”); having them in a directory called “DocBook” seems a little weird. More importantly, though, I’d like to separate them out as a fresh start, and not mess with the existing DocBook templates until we decide we don’t need them anymore. If we could end up with a cleaner, simpler makefile in the process, that would be a bonus.*

- *I’m not sold on the new inclusion mechanism. Creating thousands of little files and tracking them for dependencies and such doesn’t seem like a simplification or a path toward better performance. I would like to at least consider keeping the direct-from-source inclusion.*
- *Insisting on EXPORT\_SYMBOL being in the same file doesn’t seem like it’s going to work for now; that could maybe change after Al’s work goes in, which could be fairly soon.*

Jonathan asked Jani, “Do you think you can find some time over the next month for this? I’ll try to shake loose some time too, but, well, \$EXCUSES ...”

Daniel advocated a slow approach, converting the front-line documentation to asciidoc but having the makefile produce DocBook XML, which could then use the existing toolchain to produce the final output. He thought that once this was in the source tree, they could gradually remove the DocBook dependencies in later kernel releases.

Jani wanted to get that issue nailed down right away, before deciding on any further course of action. He asked:

*Is the aim to have asciidoc -> HTML only or dual asciidoc -> HTML and asciidoc -> XML -> whatever? Or independent asciidoc -> HTML first, with the existing DocBook on the side until everything’s converted? Something else?*

*Direct asciidoc -> HTML has the problem I mentioned that there is no chunked output. If the source is big (as-is or via asciidoc includes) the output is big. The current gpu.tmpl turned way too big. We could alleviate that by splitting the source documents into smaller pieces (in gpu.tmpl case, it’s desirable no matter what), and tying them together via cross-references and TOC rather than asciidoc includes.*

*The problem with this, in turn, is that I don’t really know how automatic cross-referencing between kernel-doc comments would turn out then (e.g., i915 kernel-doc references a symbol in drm core kernel-doc after gpu.tmpl split) as asciidoc would process the files independently. A kernel-doc comment writer shouldn’t have to know which document the referenced symbol is in .... We could do post-processing I guess, but I’d really like to get rid of the homebrew aspects here.*

*Is it acceptable to have dead links when referencing symbols outside of the*

document in question, for the time being, until someone figures out a nice way to do this?

Jani added, “in my dream world, you could have asciidoc files anywhere in the Documentation tree, with a Makefile per directory identifying which ones should be processed as asciidoc. I might even name them all .txt, and you wouldn’t have to rename existing ‘almost markup’ plain text files to have them processed, just fix the markup and update the Makefile. (FWIW asciidoc suggests .txt extension, though asciidoctor suggests .adoc or .asciidoc.) I think this would better promote a gradual transition to lightweight markup, with easier to re-view patches.”

Jonathan said that asciidoc-to-HTML alone wasn’t viable, since people wanted other formats. He said that he personally would love to see epub, and he knew of others who preferred man pages. He suspected someone would like to see PDF output, though he doubted anyone had successfully generated kernel docs in PDF format for quite some time.

In terms of splitting the files up, Jonathan suggested “generate a TOC and use CSS to place it correctly,” which was an idea that had come up in private conversation and might alleviate the need for chunked output.

Keith Packard and Jonathan both felt that Daniel’s slower, short-term approach of having asciidoc translate to DocBook was not their preference. Keith felt it was important to reduce the number of links on the toolchain.

Keith also took a deeper look at asciidoc and reported:

*It is a stream processor with a stack. The output is generated with some simple templates, one for each back end. Here’s the xhtml1.1 template for section level sections: (sect1 in the .conf file):*

```
[sect1]
<div class="sect1"{
  style? {style}{role? {role}}">
  <h2{id? id="{id}"}>
    {numbered?{sectnum} }{title}</h2>
  <div class="sectionbody">
    |
  </div>
</div>
```

*The contents of the section get inserted at the |; it’s nesting, so [sect2] bits*

would get expanded while being processed.

*Each asciidoc back end has dramatically different functionality. It’s pretty clear to me that the ‘docbook’ back end has the best support for larger documents as that provides ‘book-scale’ processing bits. I’ve recently written a book in asciidoc using the docbook back end, and the html and pdf results are quite comparable. Using the html back end from asciidoc yields a significantly different result.*

*I think it should be pretty easy to hack asciidoc to add diversions to hold TOC contents while generating the rest of the doc and then replay the diversion into the final document.*

Keith went on to say, “At the end of the document, we’d have some way of wrapping the diversion in suitable additional bits to complete the TOC, which would then be formatted by CSS.

“This same technique could be used to create lists of figures and tables.

“The goal would be to create an html document which could be used without javascript, and that would work without css as well.”

Jani was reluctant to take Keith’s ideas too far, however. He said, “One of the chief complaints with the current pipeline (and some of the proposals) has been the need to install lots of tools with lots of dependencies. I would like to avoid the need to install bleeding-edge tools and stick to what’s already widely available in distros. Thus, I would like to avoid hacking asciidoc for our needs.”

Keith agreed with that, but he said “that means using docbook for now; the native html output from asciidoc is simply not usable for anything more complicated than a short web page. However, getting ready to collapse the pipeline by eliminating docbook seems like a good medium-term goal.” Meanwhile, Keith did hack up the asciidoc sources somewhat and sent his patches to the upstream maintainers.

The discussion ended there, but the kernel folks are in the process of descending on a set of tools and wrangling them into the form they want. This would resemble the relationship between the kernel and GCC, GDB, and libc, with cross-over contributions, and the occasional controversy as the kernel’s needs fail to overlap with the goals of the tools’ native developers. ■■■

## INFO

[1] mdadm Git repository: [git://git.kernel.org/pub/scm/utills/mdadm/mdadm.git](https://git.kernel.org/pub/scm/utills/mdadm/mdadm.git)

[2] AsciiDoc for Linux kernel: [git://git.lwn.net/linux.git](https://git.lwn.net/linux.git)

# 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](mailto:events@linux-magazine.com).



### Embedded Linux Conference

**Date:** April 4–6, 2016

**Location:** San Diego, California

**Website:** <http://events.linuxfoundation.org/events/embedded-linux-conference/>

ELC, the technical conference for companies and developers using Linux in embedded products, presents the 12th year of sessions dedicated exclusively to embedded Linux and embedded Linux developers.

### Apache: Big Data

**Date:** May 9–11, 2016

**Location:** Vancouver, British Columbia

**Website:** <http://events.linuxfoundation.org/events/apache-big-data-north-america>

Apache projects are the foundation of many Big Data platforms. Join other professionals working in Big Data, ubiquitous computing, and data engineering and science to accelerate the state of the art.

### ApacheCon North America

**Date:** May 11–13, 2016

**Location:** Vancouver, British Columbia

**Website:** <http://events.linuxfoundation.org/events/apachecon-core-north-america>

Join the open source community to learn about and collaborate on the technologies and projects driving the future of open source, web technologies, and cloud computing.

## EVENTS

Embedded Linux Conference	April 4–6	San Diego, California	<a href="http://events.linuxfoundation.org/events/embedded-linux-conference">http://events.linuxfoundation.org/events/embedded-linux-conference</a>
Cloud Expo Europe	April 12-13	London, England	<a href="http://www.cloudexpo-europe.com/">http://www.cloudexpo-europe.com/</a>
Smart IOT London	April 12 - 13	London, England	<a href="http://www.smartiotlondon.com/">http://www.smartiotlondon.com/</a>
Cloud Security Expo	April 12 - 13	London, England	<a href="http://www.cloudsecurityexpo.com/">http://www.cloudsecurityexpo.com/</a>
Linux Storage Filesystem and MM Summit	April 18–19	Raleigh, North Carolina	<a href="http://events.linuxfoundation.org/events/linux-storage-filesystem-and-mm-summit">http://events.linuxfoundation.org/events/linux-storage-filesystem-and-mm-summit</a>
Vault Linux Storage and Filesystems Conference	April 20–21	Raleigh, North Carolina	<a href="http://events.linuxfoundation.org/events/vault">http://events.linuxfoundation.org/events/vault</a>
Open Source Data Center Conference	April 26–28	Berlin, Germany	<a href="https://www.netways.de/en/events_trainings/osdc/overview/">https://www.netways.de/en/events_trainings/osdc/overview/</a>
Grazer Linuxtage 2016	April 29-30	Graz, Austria	<a href="https://www.linuxtage.at/">https://www.linuxtage.at/</a>
Re:publica	May 2-4	Berlin, Germany	<a href="https://re-publica.de/">https://re-publica.de/</a>
Open Tech Summit	May 5	Berlin, Germany	<a href="http://opentechsummit.net/">http://opentechsummit.net/</a>
Apache Big Data North America	May 9–11	Vancouver, BC, Canada	<a href="http://events.linuxfoundation.org/events/apache-big-data-north-america">http://events.linuxfoundation.org/events/apache-big-data-north-america</a>
DrupalCon North America	May 9–13	New Orleans, Louisiana	<a href="https://events.drupal.org/neworleans2016">https://events.drupal.org/neworleans2016</a>
ApacheCon Core North America	May 11–13	Vancouver, BC, Canada	<a href="http://events.linuxfoundation.org/events/apachecon-core-north-america">http://events.linuxfoundation.org/events/apachecon-core-north-america</a>
Computex Taipei	May 31–June 4	Taipei, Taiwan	<a href="https://www.computextaipei.com.tw/">https://www.computextaipei.com.tw/</a>
MesosCon North America 2016	June 1 - 2	Denver, Colorado	<a href="http://www.cloudsecurityexpo.com/">http://www.cloudsecurityexpo.com/</a>
Eclipsecon	June 7 - 9	Toulouse, France	<a href="https://www.eclipsecon.org/france2016/">https://www.eclipsecon.org/france2016/</a>
Pi and More 9	June 11	Trier, Germany	<a href="http://piandmore.de/en">http://piandmore.de/en</a>
SLAC	June 15 - 17	Berlin, Germany	<a href="https://www.heinlein-support.de/secure-linux-administration-conference">https://www.heinlein-support.de/secure-linux-administration-conference</a>
ISC High Performance	June 19–23	Fankfurt, Germany	<a href="http://www.isc-hpc.com/">http://www.isc-hpc.com/</a>



# 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](mailto: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](http://www.linux-magazine.com/contact/write_for_us).

## AUTHORS

Mario Blättermann	76
Zack Brown	92
Bruce Byfield	86
Joe Casad	3, 8
Markus Feilner	38
Marco Fioretti	56
Jon "maddog" Hall	90
Kristian Kißling	20
Charly Kühnast	62
Martin Loschwitz	44
Sebastian Mogilowski	38
Dr. Roland Pleger	80
Dmitri Popov	28, 70
Mike Schilli	64
Jörg Thoma	12
Ferdinand Thommes	50
Uwe Vollbracht	34

## CONTACT INFO

### Editor in Chief

Joe Casad, [jcasad@linux-magazine.com](mailto:jcasad@linux-magazine.com)

### Managing Editor

Rita L Sooby, [rsooby@linux-magazine.com](mailto:rsooby@linux-magazine.com)

### Localization & Translation

Ian Travis

### News Editor

Joe Casad

### Copy Editor

Amber Ankerholz

### Layout

Dena Friesen, Lori White

### Cover Design

Dena Friesen, Lori White

### Cover Image

© destinacigdem, 123RF.com and Noel Powell, Fotolia.com

### Advertising – North America

Ann Jesse, [ajesse@linuxnewmedia.com](mailto:ajesse@linuxnewmedia.com)  
phone +1 785 841 8834

### Advertising – Europe

Penny Wilby, [pwilby@sparkhausmedia.com](mailto:pwilby@sparkhausmedia.com)  
phone +44 1807 211100

### Publisher

Brian Osborn, [bosborn@linuxnewmedia.com](mailto:bosborn@linuxnewmedia.com)

### Marketing Communications

Darrah Buren, [dburen@linuxnewmedia.com](mailto:dburen@linuxnewmedia.com)  
Linux New Media USA, LLC  
616 Kentucky St.  
Lawrence, KS 66044 USA

### Customer Service / Subscription

For USA and Canada:  
Email: [cs@linuxpromagazine.com](mailto: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](mailto:subs@linux-magazine.com)

Phone: +49 89 99 34 1167

Fax: +49 89 99 34 1198

[www.linuxpromagazine.com](http://www.linuxpromagazine.com) – North America

[www.linux-magazine.com](http://www.linux-magazine.com) – Worldwide

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 © 2016 Linux New Media USA, LLC.

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 USA, LLC, 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

Published in Europe by: Sparkhaus Media GmbH, Putzbrunner Str. 71, 81749 Munich, Germany.

Issue 187 / June 2016

# SSDs

**Approximate**

UK / Europe

May 07

USA / Canada

Jun 03

Australia

Jul 04

**On Sale Date**

Solid State Drives (SSDs) promise faster storage with fewer complications, but you'll need some special tools for managing your SSD storage systems. Next month, we look at tools and tuning techniques for SSDs on Linux networks.

## 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](http://www.linux-magazine.com/newsletter)

# 8th Annual Southeast Linuxfest



# 2016

SHERATON AIRPORT, CHARLOTTE, NC

JUNE 10  
JUNE 12

**FREE**



@SELINUXFEST



YOUTUBE.COM/  
USER/  
SOUTHEASTLINUXFEST



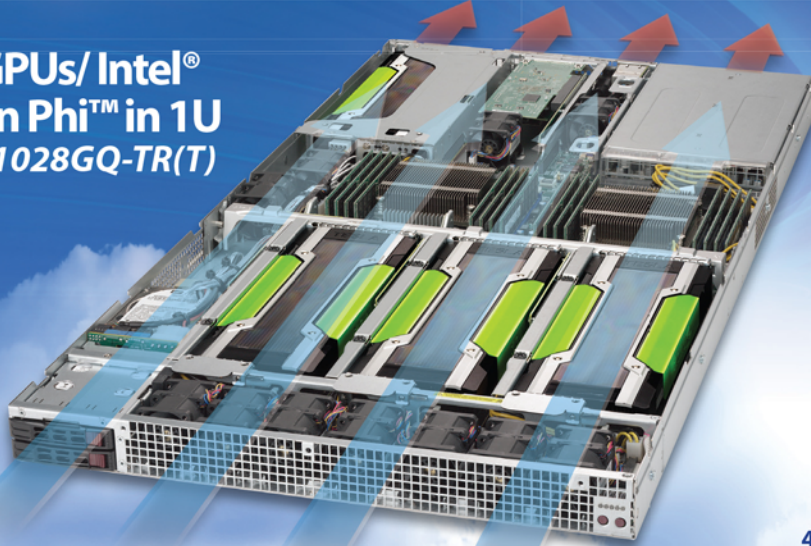
FB.COM/  
SELINUXFEST



# GPU Solutions

Maximum Acceleration · Highest Efficiency · Application-Optimized

4 GPUs/ Intel®  
Xeon Phi™ in 1U  
SYS-1028GQ-TR(T)



2/3 GPUs/ Intel® Xeon Phi™ in 1U  
SYS-1028GR-TR(T) (shown)  
SYS-1018GR-T/SYS-5018GR-T



4/6 GPUs/ Intel® Xeon Phi™ in 2U  
SYS-2028GR-TR(T)/TRH(T)



8 GPUs/ Intel® Xeon Phi™ in 4U  
SYS-4028GR-TR(T)



20 GPUs/ Intel® Xeon Phi™ in 7U  
SBI-7128RG-X/F/F2



4 GPUs/ Intel® Xeon Phi™ in 4U/Tower  
SYS-7048GR-TR



## Maximum Acceleration for Highest-Performance Workloads, with models that support:

- Up to 36 Cores per Node dual Intel® Xeon® Processor E5-2600 v3 product family
- No GPU/ Intel® Xeon Phi™ preheating
- Direct connect to GPU/ Intel® Xeon Phi™ w/o redriver or switch or PCI-E extension cable
- 1.5TB DDR4-2133MHz memory, 24 DIMM slots per Node
- 4/6/8/20 GPUs/ Intel® Xeon Phi™ in 1U/2U/4U/7U
- 160W TDP CPUs and 300W TDP GPUs/ Intel® Xeon Phi™
- Redundant Titanium Level (96%+) power supplies & intelligent cooling controls
- IPMI 2.0 Server Management and GPU status monitoring



Learn more at [www.supermicro.com](http://www.supermicro.com)

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

