RSpec | Virtualization | iSCSI SAN | AMQP | IPv4 Anycast | HIPL

LINUX JOURNAL

Since 1994: The Original Magazine of the Linux Community
NOVEMBER 2009 | ISSUE 187 | www.linuxjournal.com

Virtualization Shootout

Solve Mobility Problems with HIPL

How Big Is Your Infrastructure?

Improve Availability with Anycast

Use AMQP for Enterprise Messaging

Storage Speed and Redundancy with iSCSI SAN

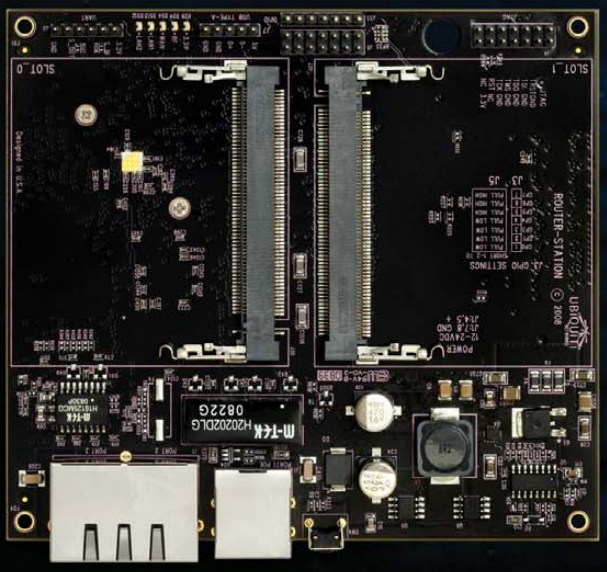
PLUS:

POINT/COUNTERPOINT:
Ext3 vs. XFS

$5.99US $5.99CAN
11>
0 09281 03102 4

UBIQUITI NETWORKS

Embedded Wireless Dream Machines.

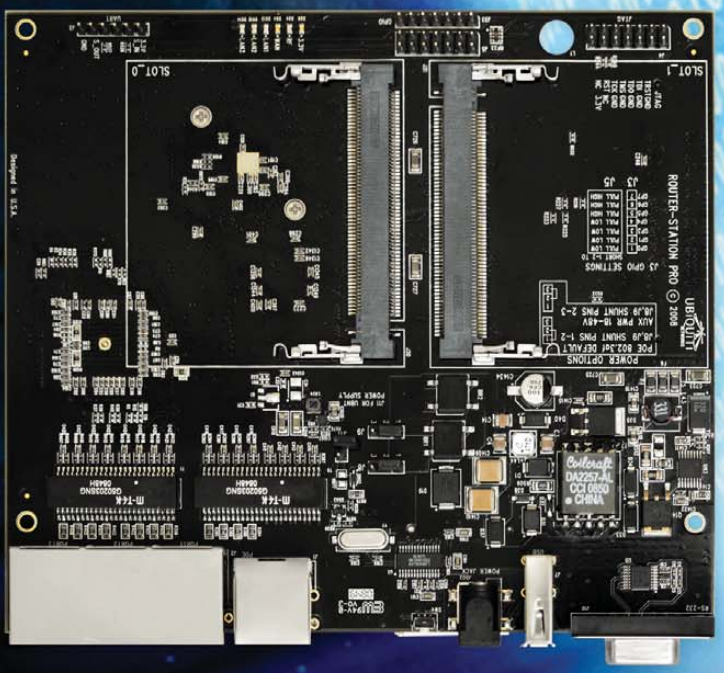


RouterStation

Featuring a fast 680MHz MIPS 24K CPU, 64MB RAM, and 16MB Flash; RouterStation provides a excellent horsepower for a variety of processor intensive multi-radio system applications.

Up to 3 mini-PCI radios, 3 10/100 ethernet interfaces, a 5A power supply for multiple hi-power card support, USB 2.0, and enhanced temperature operating performance and ethernet ESD protection for carrier applications.

MSRP $59



RouterStation Pro

In response to the outstanding demand for our initial RouterStation OEM platform, Ubiquiti Networks announces the RouterStation Pro. Breakthrough Price/Performance with a $79 USD MSRP.

Pro Version Enhancements:

- 48V 802.3af Power Over Ethernet
- 4-Port Gigabit Ethernet Switch
- 256MB RAM
- On Board SDIO Support
- On Board, USB 2.0, RS232/dB9, and DC power jacks

MSRP $79

Prices in USD. Ubiquiti Networks, Inc. Copyright © 2009 All Rights Reserved.

www.ubnt.com

Hurry, offer ends October 31st!

PERSONAL HOSTING

Everything you need for a perfect site.

1&1® Home Package:
- 2 FREE Domains
- 150 GB Web Space
- UNLIMITED Traffic

50% off! First 6 months

~~$6.99 per month~~

$3.49 per month*

Hurry, offer ends October 31st!

BUSINESS HOSTING

Powerful solution for professional websites.

1&1® Business Package:
- 3 FREE Domains
- 250 GB Web Space
- UNLIMITED Traffic

50% off! First 6 months

~~$9.99 per month~~

$4.99 per month*

Special Offer: .net and .info domains just $4.49 for the first year!*
More special offers are available online. For details, visit www.1and1.com

*Offers valid through October 31, 2009. Setup fee, minimum contract term, and other terms and conditions may apply. Visit www.1and1.com for full promotional offer details. Program and pricing specifications, availability and prices subject to change without notice. 1&1 and the 1&1 logo are trademarks of 1&1 Internet AG, all other trademarks are the property of their respective owners. © 2009 1&1 Internet, Inc. All rights reserved.

MEMBER OF united internet

Call **1-877-GO-1AND1**

Visit us now **www.1and1.com**

1&1®

CONTENTS

NOVEMBER 2009
Issue 187

FEATURES

INFRASTRUCTURE

44
IBM INFOSPHERE STREAMS AND THE UPPSALA UNIVERSITY SPACE WEATHER PROJECT
Dealing with massive amounts of real-time data.
Shawn Powers

50
USE LINUX AS A SAN PROVIDER
Linux: the Swiss-Army knife of technologies.
Michael Nugent

54
ADVANCED MESSAGE QUEUING PROTOCOL (AMQP)
Enterprise messaging with Python and AMQP.
Joshua Kramer

62
IPV4 ANYCAST WITH LINUX AND QUAGGA
Improve availability with anycast!
Philip Martin

ON THE COVER

- **Virtualization Shootout, p. 40**
- **Solve Mobility Problems with HIPL, p. 70**
- **How Big Is Your Infrastructure?, p. 44**
- **Improve Availability with Anycast, p. 62**
- **Use AMQP for Enterprise Messaging, p. 54**
- **Storage Speed and Redundancy with iSCSI SAN, p. 50**
- **Point/Counterpoint: Ext3 vs. XFS, p. 76**

Cover image from **antwrp.gsfc.nasa.gov/apod/ap080601.html**

Orion iX-N4224

The **Orion** features unrivaled cooling efficiency, power efficiency and excellent storage density!

Superior Savings On Energy Costs

Designed for storage-intensive applications and virtualization, the iX-N4224 4U storage server series delivers incredible performance, storage capacity, and energy savings with adaptability and superior hardware. Each iX-N4224 comes with a **Gold Level power supply**, boasting a high 93% energy efficiency. Powerful Intel® Xeon® 5500 series quad core processors **intelligently save power during low-use periods** and increase performance when systems require it. Intel® Xeon® 5500 series processors include virtualization technologies to lead the way in performance, scalability, and simplified server management and migration.

The iX-N4224 supports up to 144GB of DDR3 1333 energy efficient RAM and utilizes three 5000 RPM cooling PWM fans and two 5000 RPM rear exhaust PWM fans. iX-N4224 servers offer up to 48 terabytes of storage with 24 hot-swappable SAS/SATA drive bays in a 4U configuration. Storage sizes for the iX-N4224 are customizable, with 250MB, 500MB, 750MB, 1TB, and 2TB hard drives available.

The iX-N4224 provides the ideal solution for applications requiring maximum storage capacity and power savings. For particularly storage-hungry applications, Western Digital® offers 2TB WD™ RE4-GP hard drives, which offer lower power use during idle times, a 64 megabyte cache, up to 25% increased performance, and a savings of up to $10 per drive on yearly power costs. Each hard drive is equipped with improvements to rotary vibration tolerance and calculates optimum seek speeds to lower power consumption, noise, and vibration. These drives also require less power and time to start up, allowing more drives to start spinning simultaneously due to the decrease in the current each drive requires. Equipping the Orion iX-N4224 4U storage server with the WD™ RE4-GP drives provides unparalleled storage capacity and power efficiency.

For more information about the Orion Series visit *http://www.iXsystems.com/Orion*.

Features

- Dual 64-Bit Socket 1366 Quad-Core or Dual-Core Intel® Xeon® Processor 5500 Series
- 24 x 3.5" SAS/SATA Hot-swappable Drive Bays
- 1200W high-efficiency (1+1) redundant power supply (Gold Level 93%)
- 100% Cooling Redundancy
- Dual Intel® 5520 chipsets with QuickPath Interconnect (QPI)
- Up to 144GB DDR3 1333/1066/800 SDRAM ECC Registered Memory (18 DIMM Slots)
- 2 PCI-E 2.0 x16, 4 PCI-E x8 (1 in x16 slot), and 1 PCI-E x4 Expansion Slots
- Intel® 82576 Dual Port Gigabit Ethernet Controller
- Optional 2x Internal Fixed 3.5" HDD or 2x fixed 2.5" HDD + DVD
- Matrox G200eW Graphics
- Remote Management - IPMI 2.0 + IP-KVM with dedicated LAN

Call 1-800-820-BSDI Today!

iXsystems™

intel Xeon® inside™

Powerful. Intelligent.

800-820-BSDI
http://www.iXsystems.com
Enterprise Servers for Open Source

Intel, the Intel logo, and Xeon Inside are trademarks or registered trademarks of Intel Corporation in the U.S. and other countries.

CONTENTS

NOVEMBER 2009
Issue 187

COLUMNS

20 REUVEN M. LERNER'S AT THE FORGE
RSpec for Controllers

24 DAVE TAYLOR'S WORK THE SHELL
Exploring Lat/Lon with Shell Scripts

28 MICK BAUER'S PARANOID PENGUIN
DEFCON: One Penguin's Annual Odyssey

32 KYLE RANKIN'S HACK AND /
Dr hjkl and Mr Hack

76 KYLE RANKIN AND BILL CHILDERS' POINT/COUNTERPOINT
Ext3 vs. XFS

80 DOC SEARLS' EOF
The Hacking of Infrastructure. And Vice Versa.

REVIEW

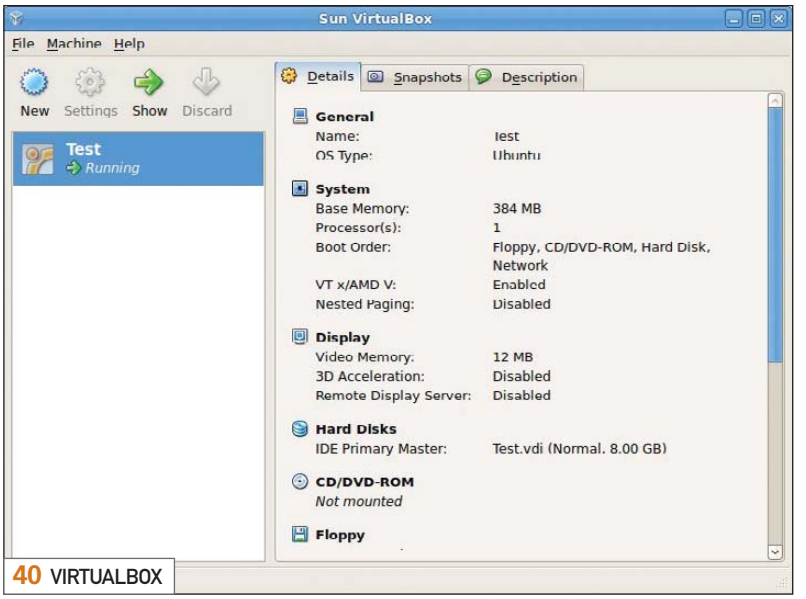
40 VIRTUALIZATION SHOOTOUT: VMWARE SERVER VS. VIRUTALBOX VS. KVM
Bill Childers

INDEPTH

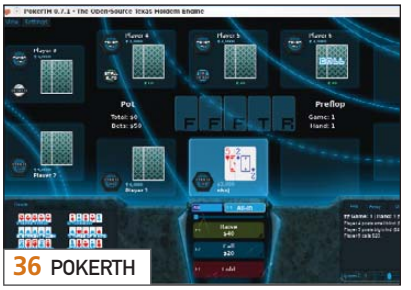
70 HOST IDENTITY PROTOCOL FOR LINUX
HIPL gives your Linux box a name.
Abhinav Pathak, Andrei Gurtov and Miika Komu

IN EVERY ISSUE

8 CURRENT_ISSUE.TAR.GZ
10 LETTERS
14 UPFRONT
34 NEW PRODUCTS
36 NEW PROJECTS
65 ADVERTISERS INDEX
78 MARKETPLACE



40 VIRTUALBOX


36 POKERTH

37 X-MOTO

Next Month: EMBEDDED

What better Christmas present could there be than an issue of *Linux Journal* focused on embedded systems? According to the S. Claus research center, there is nothing better.

Grab an eggnog and read about using Player with a CoroWare robot. Find out how to get your embedded system to boot faster with tips from MontaVista. And, if you're new to embedded systems, read our intro article on what goes into a typical embedded system.

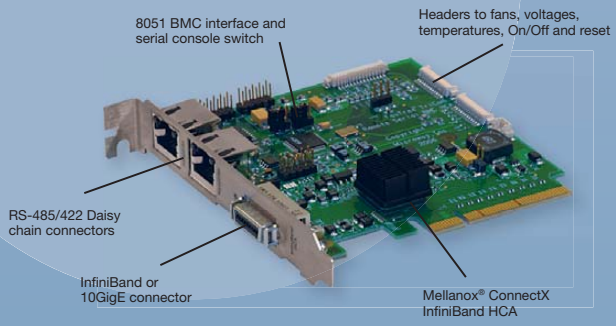
Up here, we think snow for Christmas, but if you live down under, you're probably thinking "it's not the heat, it's the humidity". If that's you, don't miss our article on controlling humidity with an embedded Linux system.

USPS *LINUX JOURNAL* (ISSN 1075-3583) (USPS 12854) is published monthly by Belltown Media, Inc., 2211 Norfolk, Ste 514, Houston, TX 77098 USA. Periodicals postage paid at Houston, Texas and at additional mailing offices. Cover price is $5.99 US. Subscription rate is $29.50/year in the United States, $39.50 in Canada and Mexico, $69.50 elsewhere. POSTMASTER: Please send address changes to *Linux Journal*, PO Box 16476, North Hollywood, CA 91615. Subscriptions start with the next issue. Canada Post: Publications Mail Agreement #41549519. Canada Returns to be sent to Bleuchip International, P.O. Box 25542, London, ON N6C 6B2

Your Applications Will Run Faster
With Next Generation Microway Solutions!

TriCom™X

- QDR/DDR InfiniBand HCA
- ConnectX™ Technology
- 1µsec Latency
- Switchless Serial Console
- NodeWatch™ Remote Management



Teraflop GPU Computing

For Workstations and HPC Clusters

- NVIDIA® Tesla™ GPU with 240 Cores on One Chip
  - CUDA™ SDK
- NVIDIA® Quadro® Professional Graphics
- AMD® FireStream™ GPU
  - Stream SDK with Brook+

NumberSmasher®

Large Memory Scalable SMP Server

- Scales to 1 TB of Virtual Shared Memory
- Up to 128 CPU Cores
- 8U System Includes 32 Quad Core CPUs
- QDR 1 µsec Backplane

FasTree™X

- Mellanox® InfiniScale™ IV Technology
- QDR/DDR InfiniBand Switches
- Modular Design
- 4 GB/sec Bandwidth per Port
- QSFP Interconnects
- InfiniScope™ Real Time Diagnostics

Call the HPC Experts at Microway to Design Your Next High-Reliability Linux Cluster or InfiniBand Fabric.

508-746-7341

Sign up for Microway's Newsletter at
www.microway.com

Microway®
Technology you can count on℠

LINUX JOURNAL™
Since 1994: The Original Magazine of the Linux Community

Digital Edition Now Available!

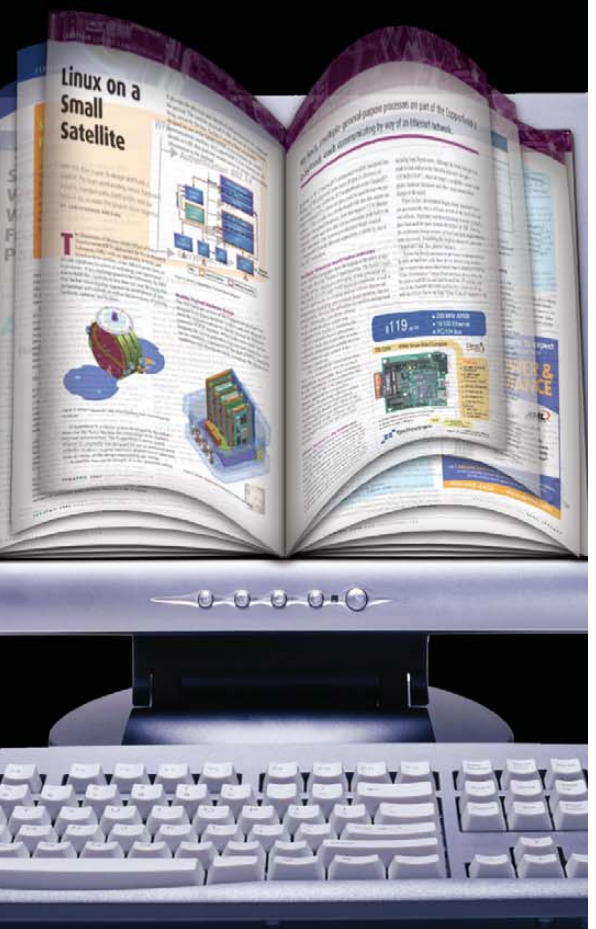
Read it first
Get the latest issue before it hits the newsstand

Keyword searchable
Find a topic or name in seconds

Paperless archives
Download to your computer for convenient offline reading

Same great magazine
Read each issue in high-quality PDF

Try a Sample Issue!
www.linuxjournal.com/digital



LINUX JOURNAL

Executive Editor	Jill Franklin jill@linuxjournal.com
Senior Editor	Doc Searls doc@linuxjournal.com
Associate Editor	Shawn Powers shawn@linuxjournal.com
Associate Editor	Mitch Frazier mitch@linuxjournal.com
Art Director	Garrick Antikajian garrick@linuxjournal.com
Products Editor	James Gray newproducts@linuxjournal.com
Editor Emeritus	Don Marti dmarti@linuxjournal.com
Technical Editor	Michael Baxter mab@cruzio.com
Senior Columnist	Reuven Lerner reuven@lerner.co.il
Chef Français	Marcel Gagné mggagne@salmar.com
Security Editor	Mick Bauer mick@visi.com
Hack Editor	Kyle Rankin lj@greenfly.net
Virtual Editor	Bill Childers bill.childers@linuxjournal.com

Contributing Editors
David A. Bandel • Ibrahim Haddad • Robert Love • Zack Brown • Dave Phillips • Marco Fioretti
Ludovic Marcotte • Paul Barry • Paul McKenney • Dave Taylor • Dirk Elmendorf

Proofreader Geri Gale

Publisher	Carlie Fairchild publisher@linuxjournal.com
General Manager	Rebecca Cassity rebecca@linuxjournal.com
Sales Manager	Joseph Krack joseph@linuxjournal.com
Associate Publisher	Mark Irgang mark@linuxjournal.com
Webmistress	Katherine Druckman webmistress@linuxjournal.com
Accountant	Candy Beauchamp acct@linuxjournal.com

Linux Journal **is published by, and is a registered trade name of, Belltown Media, Inc.**
PO Box 980985, Houston, TX 77098 USA

Reader Advisory Panel
Brad Abram Baillio • Nick Baronian • Hari Boukis • Caleb S. Cullen • Steve Case
Kalyana Krishna Chadalavada • Keir Davis • Adam M. Dutko • Michael Eager • Nick Faltys • Ken Firestone
Dennis Franklin Frey • Victor Gregorio • Kristian Erik • Hermansen • Philip Jacob • Jay Kruizenga
David A. Lane • Steve Marquez • Dave McAllister • Craig Oda • Rob Orsini • Jeffrey D. Parent
Wayne D. Powel • Shawn Powers • Mike Roberts • Draciron Smith • Chris D. Stark • Patrick Swartz

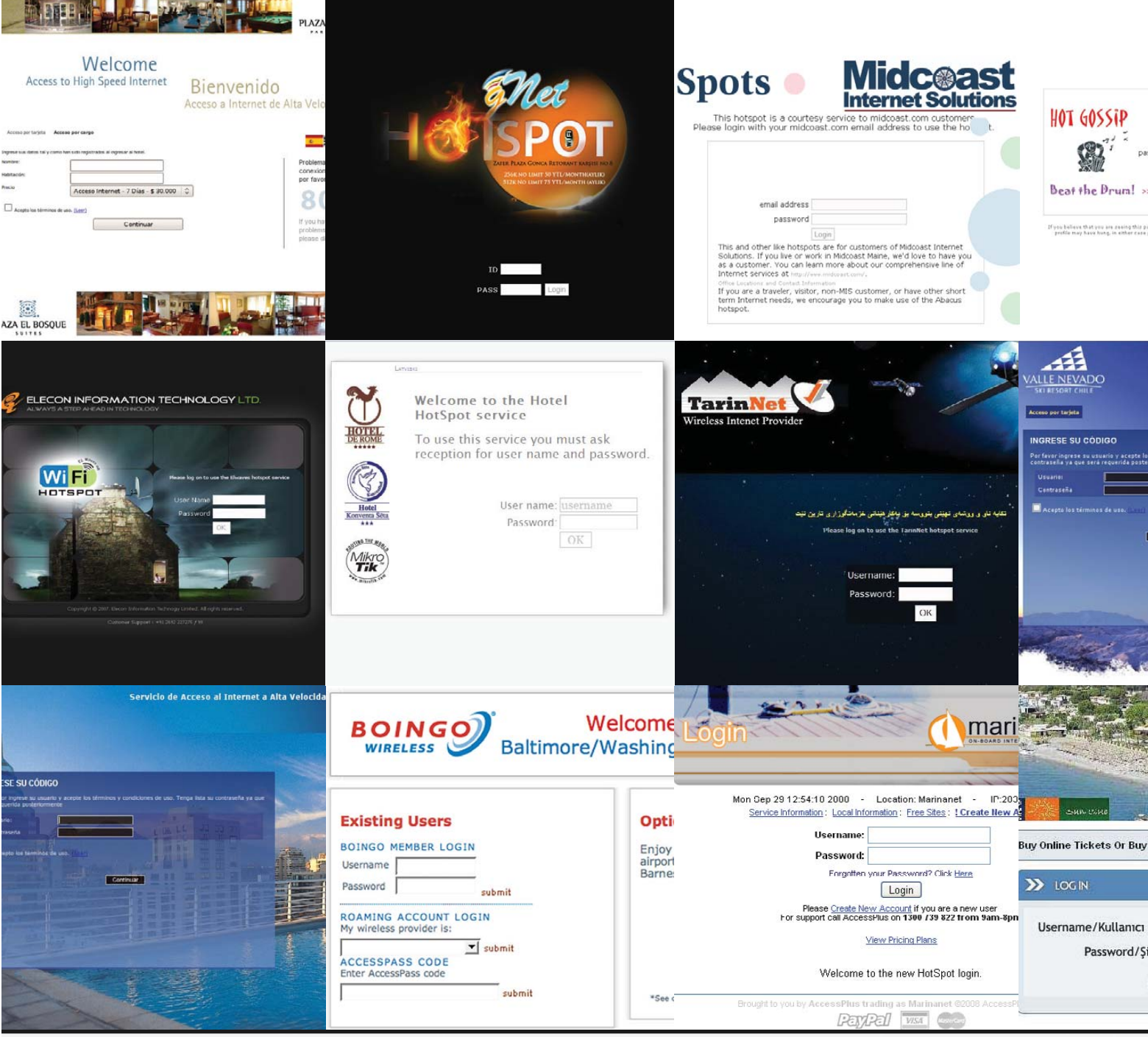
Editorial Advisory Board
Daniel Frye, Director, IBM Linux Technology Center
Jon "maddog" Hall, President, Linux International
Lawrence Lessig, Professor of Law, Stanford University
Ransom Love, Director of Strategic Relationships, Family and Church History Department,
Church of Jesus Christ of Latter-day Saints
Sam Ockman
Bruce Perens
Bdale Garbee, Linux CTO, HP
Danese Cooper, Open Source Diva, Intel Corporation

Advertising
E-MAIL: ads@linuxjournal.com
URL: www.linuxjournal.com/advertising
PHONE: +1 713-344-1956 ext. 2

Subscriptions
E-MAIL: subs@linuxjournal.com
URL: www.linuxjournal.com/subscribe
PHONE: +1 818-487-2089
FAX: +1 818-487-4550
TOLL-FREE: 1-888-66-LINUX
MAIL: PO Box 16476, North Hollywood, CA 91615-9911 USA
Please allow 4–6 weeks for processing address changes and orders
PRINTED IN USA

LINUX is a registered trademark of Linus Torvalds.

PRINTED WITH SOY INK®



MikroTik powered Hotspots around the world

MikroTik RouterOS powered hotspots are everywhere. From Internet cafes in desert towns of Africa, to futuristic airport lounges in the US and five star hotels in the Mediterranean. Mikrotik can power your systems too. Free evaluation installations are available in our download section.

www.mikrotik.com

SHAWN POWERS

All Your Blades Are Belong To Us

I often say geeks rule the world, and we could take over if we wanted to—we just have better things to do with our time. This issue, we explore that notion and show Linus we're well on our way to world domination. I could go on and on, but I think regular folk might get scared if they realized the degree of access we have to information—*mwahahahaa*. Putting aside my aspirations to be a modern-day Lex Luthor, this month, we talk about infrastructure. Let's face it, Linux rules the roost when it comes to infrastructure. Heck, even a large percentage of Windows servers are really just virtual machines running on top of a Linux hypervisor.

Bill Childers gets us going with that very topic. If you're planning to virtualize much of your existing server room, picking a hypervisor can be the hardest step. Bill compares and contrasts VMware Server, VirtualBox and KVM. In my own server room, I have only one Windows server, and the fact that it runs on top of a Linux hypervisor makes me smile. Bill doesn't stop there, however; he also argues again this month with Kyle Rankin. Kyle seems to think XFS is the best filesystem to use, while Bill is convinced ext3 is still king. I try to stay out of their little spats, but their discussion is enlightening to read.

Every server room needs storage. For many of us, that's just a few hard drives in a RAID array. As needs grow, however, single-server storage solutions don't scale that well—enter SAN. Usually, that means lots of money to an already expensive infrastructure, but Michael Nugent shows us how to create a Linux-based SAN for a fraction of the cost. Along with the need for large storage solutions, comes the need for redundancy. We also have an article on IPv4 Anycast, where Philip Martin explains how to add availability for mission-critical services. (Anyone that has experienced the "network hang" of a downed DNS server will appreciate the notion of high availability!)

Infrastructure extends outside our precious server closets though, and sits on our desks, in our backpacks and even our pockets. When traveling from location to location, changing networks can be frustrating. Abhinav Pathak, Andrei Gurtov and Miika Komu show us a bit about Host Identity Protocol for Linux and how we can keep our identity no matter where we go. In a similar vein, Joshua Kramer demonstrates the Advanced Message Queueing Protocol, which allows applications to communicate with each other regardless of location. Even if you are telecommuting from the "clouds", it's important to be connected. A good infrastructure knows no geographical limits—which brings us to an interview I conducted this month....

Linus may be happy with Linux dominating the world, but quite frankly, some people have bigger goals in mind. The IBM InfoSphere Streams Project aims a bit higher, and using Linux as its underlying base, it gathers information about space weather. The amount of data is so great, it has to be analyzed in real time. I like the sound of "Interplanetary Domination" quite a bit, so Mitch Frazier and I took the bull by the horns and interviewed the folks at IBM. I enjoyed the interview; hopefully, you will too.

What about our regular cast of columnists? They're all here this month too. Reuven M. Lerner continues telling us about RSpec, Dave Taylor shows us how to manage latitude and longitude from inside a shell script, and Mick Bauer describes the ultimate conference for hackers, DEFCON. Speaking of hackers, Kyle Rankin tries to explain why arrow keys have no place in our lives as Linux users and strives to turn us all into die-hard vim users. I'm already mostly with him, but I'll admit I use arrow keys. I guess that makes me a n00b.

So although your coffeepot might not be running a Linux kernel and your dishwasher doesn't instant message you when the cleaning cycle is complete, that time is coming sooner than you think. What will our intergalactic infrastructure be based on? My guess is Linux. This month, you can get a jump start on that transition and perhaps have a say on whether your refrigerator will have an ext3 or XFS filesystem—at least, that's what Bill and Kyle are hoping for.■

Shawn Powers is the Associate Editor for *Linux Journal*. He's also the Gadget Guy for LinuxJournal.com, and he has an interesting collection of vintage Garfield coffee mugs. Don't let his silly hairdo fool you, he's a pretty ordinary guy and can be reached via e-mail at shawn@linuxjournal.com. Or, swing by the #linuxjournal IRC channel on Freenode.net.

servers DIRECT®

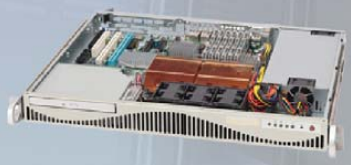
MORE PRODUCTS, BETTER SERVICE, GUARANTEED.

GO STRAIGHT TO THE SOURCE!

1.877.727.7887 | www.ServersDirect.com

YOUR HIGH PERFORMANCE COMPUTING HAS ARRIVED.

The ServersDirect® Systems with the Intel® Xeon® Processor helps you simplify computing operations, accelerate performance and accomplish more in less time



STARTING AT **$899**

ENTRY LEVEL INTELLIGENT SERVER

SDR-S1341-T00 is among our most cost-effective 1U Xeon Servers, and it is ideal for large high-performance computing deployments

STARTING AT **$959**

APPLICATION SERVER

Refresh your servers with new **SDR-S1337-T02** powered by Intel® Xeon® processor 5500 series, based on intelligent performance, automated energy efficiency and flexible virtualization.

SDR-S1343-T04

STARTING AT **$1,099**

1U INTEL® XEON® PROCESSORS 5500 SERIES SERVER W/ 4X 3.5" HOT-SWAP SATA DRIVE BAYS

- Supermicro 1U Rackmount Server with 560W Power Supply
- Supermicro Server Board w/Intel® 5520 Chipset
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 96GB DDR3 1333/ 1066/ 800MHz ECC Reg.DIMM
- 4x 3.5" Hot-swap SATA Drive Bays
- Intel® 82576 Dual-Port Gigabit Ethernet Controller

SDR-S2311-T08

STARTING AT **$1,159**

2U INTEL® XEON® PROCESSORS 5500 SERIES SERVER W/ 8X 3.5" HOT-SWAP SAS/SATA BAYS

- Supermicro 2U Rackmount Server with 560W Power Supply
- Supermicro Server Board w/Intel® 5500 Chipset
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 24GB DDR3 1333/ 1066/ 800MHz ECC Reg.DIMM
- 8x 3.5" Hot-swap SATA Drive Bays
- Dual Intel® 82574L Gigabit Ethernet Controller

SDP-IP308-T10

STARTING AT **$1,599**

PEDESTAL INTEL® XEON® PROCESSORS 5500 SERIES SERVER W/ 10X HOT-SWAP (OPT.) SATA BAYS

- Intel Pedestal Chassis w/ 750W (1+1) Power Supply
- Supermicro Server Board w/Intel® 5520 Chipsets
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 96GB DDR3 1333/ 1066/ 800MHz ECC Reg./unbuffered DIMM
- Option 10x 3.5" Hot-swap SATA Bays
- Intel® 8257EB Dual-port Gigabit Ethernet Controller

SDR-S4313-T24

STARTING AT **$1,899**

4U INTEL® XEON® PROCESSORS 5500 SERIES SERVER W/ 24X 3.5" HOT-SWAP SAS/SATA BAYS

- Supermicro 4U Rackmount 900W (1+1) Red. Power Supply
- Supermicro Server Board w/ Dual Intel® 5520 Chipsets
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 144GB DDR3 1333/ 1066/ 800MHz ECC Reg. DIMM
- 24x 3.5" Hot-swap SATA Drive Bay
- Intel® 82576 Dual-port Gigabit Ethernet Controller

SDR-S3305-T16

STARTING AT **$1,979**

3U INTEL® XEON® PROCESSORS 5500 SERIES SERVER W/ 16X 3.5" HOT-SWAP SAS/SATA BAYS

- 3U Rackmount Server with 1+1 900W Red. Power Supply
- Supermicro Server Board w/ Dual Intel® 5520 Chipsets
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 96GB DDR3 1333/ 1066/ 800MHz ECC Reg.DIMM
- 16x Hot-swap SAS/SATA Drive Bays
- Intel® Dual 82576 Dual-Port Gigabit Ethernet (4 ports)

SDR-C9303-T50

STARTING AT **$4,339**

9U INTEL® XEON® PROCESSORS 5500 NEHALEM SERIES SERVER W/ 50X HOT-SWAP SATA II / SAS BAYS

- 9U Chassis with 1620W Redundant Power Supply
- Supermicro Server Board w/ Dual Intel® 5520 Chipsets
- Support up to Dual Intel® 5500 series Xeon® Quad/Dual-Core, with QPI up to 6.4 GT/s
- Support up to 144GB DDR3 1333/ 1066/ 800MHz ECC Reg. DIMM
- 50 x 3.5"Internal SATA Drives Trays
- Intel® 82576 Dual-port Gigabit Ethernet Controller

10th year

SERVERS DIRECT CAN HELP YOU CONFIGURE YOUR NEXT HIGH PERFORMANCE SERVER SYSTEM - CALL US TODAY!

Our flexible on-line products configurator allows you to source a custom solution, or call and our product experts are standing by to help you to assemble systems that require a little extra. Servers Direct - your direct source for scalable, cost effective solutions.

1.877.727.7886 / www.ServersDirect.com

Intel, Intel logo, Intel Inside, Intel Inisde logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, Pentium, and Pentium III Xeon are trademarks of Intel Corporation or it's subsidiaries in the United States and other countries.

intel Xeon inside™

Powerful. Intelligent.

letters

Pidgin Is Not a GNOME App

In reading the Cooking with Linux column in the September 2009 issue, I ran across mention of Pidgin, a relatively popular instant-messaging client. I'd like to point out that referring to Pidgin as a GNOME application is wrong. Pidgin is not a GNOME application and not a part of the GNOME Project, nor does it have any GNOME dependencies. Using GTK+ does not make something a GNOME application. Empathy is the blessed GNOME IM application.

I find it disappointing that people insist on referring to Pidgin as a GNOME application, when we have no involvement with GNOME. [Note, the author of this letter is a Pidgin developer.]

--

John Bailey

LJ Videos Rock

I have really been enjoying the Tech Tips and other videos on the *Linux Journal* Web site. I discovered these gems after adding the *LJ* feed to my home page. To me, these videos are the most exciting and useful addition *LJ* has made in years.

My question is this: what video capture and editing tools/devices do the *LJ* staff use to create these videos? I see that Shawn and Mitch contribute videos quite frequently, so I'm curious what recommendations the folks at *LJ* have for us readers to create video tutorials of our own. I know this would be a great way to document things at work or share some useful tips with the community.

--

Tom H

If you go back in the archives, both Mitch and I show our screencasting methods. I must admit, however, I've adjusted the way I do videos quite a bit and still vary from day to day. If I'm showing something that isn't graphic-intensive, I'll use a VM and capture with either xvidcap on my Linux machine or with SnapzProX on my Macintosh. Then, I'll do final edits with Final Cut or Kino. I also usually voiceover after recording the videos to avoid "Ums".

One of the things I want to check out soon is the idea of Web-based screen-capture tools. As I'm usually on different computers and different platforms all the time, it would be nice to have a consistent interface. Although I haven't tried them yet, sites like **www.screentoaster.com** *look promising.—Ed.*

Autoconf, Automake, Libtool

Almost every Linux build from source uses the familiar:

```
tar -xvf ...
./configure
make
```

However, getting started with the magic tools (autoconf, automake and libtool) is similar to finding yourself lost in the "maze of twisty little passages, all alike".

How about a tutorial series on getting started with the tools? I realize there is a complexity that cannot be satisfied with a brief tutorial, but some hints, tips and examples (samples of good practice) would be very helpful in shining light on the right path.

If the mention of the "twisty little maze" didn't give it away, let's just say I'm not exactly new to programming and system administration. I *can* use the source, Luke.

--

BRWms

That's not a bad idea. Perhaps we can get someone to contribute a few tech tips for the Web on the process. Thanks for the suggestion!—Ed.

KDE 4 Does Not Disappoint

In regard to the September 2009 letter to the editor titled "Disappointed with KDE 4" from Christian H., I must clarify some points and make some corrections to Christian's initial view of KDE 4.

I too was "raised" on KDE 3.x. I installed it in Debian—3.5.5 I think it was. So I saw some of his same points when I first decided to switch to KDE 4. Fortunately, most of them are simply false or non-issues.

First, Christian writes that KDE 4 has lost the ability to put icons on the desktop; this is simply false. There are, in essence, three ways to do so: 1) a folder view widget, 2) by dragging the application icons to your desktop and 3) by right-clicking the desktop, going to appearance settings, and changing from a widget desktop to folder view. True, this is "confusing" for new users, but the KDE 4 desktop is as robust and feature-filled as previous releases. If you read the documentation, this does indeed exist though. Pressing Alt-F1 shows the KDE handbook at any time.

Although the argument of a widget-based desktop will continue to be fought, widget-based desktops are an exciting and new take on the desktop, and KDE 4 has managed to push the development of the desktop in directions that *no one* has gone before.

You can autohide the kicker, which is now called simply a panel, in KDE 4. Click on the configuration cashew on the panel (you may have to unlock your widgets first) and click more options. There is your autohiding panel. In KDE 4.3 there is also "Windows can go

above" and "Windows can go below", the latter of which I have never seen any other desktop do. Saying KDE 4 is less configurable than KDE 3 is simply not true. There are many more options and ways to configure the KDE 4 panel than there ever was for the KDE 3 kicker.

You say that Konqueror is no longer included in the desktop. This also is untrue. Konqueror is, indeed, no longer the *default* file manager, but that was so its role could be more fine-tuned as simply a Web browser. KDE 4.2 and 4.3's (which was released last week) Konqueror is simply ages ahead of KDE 3's Konqueror, and it is faster and renders most of the sites that Firefox does. The only time I ever have to use Firefox is when working with the Fedora Koji Build System (which 99.9% of users won't have to do). And, if you really want Konqueror back as your file manager, you can change it in the System Settings dialog Default Applications. This is where you would have changed it in KDE 3, so why not look there in KDE 4?

You say you are waiting for KDE 5, and that could still be years away. The head KDE developers say that KDE 4 is the track we are on for a while, and in my opinion, it is a very exciting track. Take a look at KDE 4.3; it should be in most distributions right now. The best way to make KDE 4 the best desktop for users is by participating. You can do so *very* easily—by contributing art, contributing to user base and, most important, filing bugs on **bugs.kde.org** when you find them, so they can be fixed.

Also, KRunner is absolutely an amazing tool that is, in my opinion, like GNOME Do on steroids, and with a better interface too. [Note: the author of this letter is a KDE contributor and member of the Fedora KDE Special Interests Group.]

--

Ryan Rix

LJ Tech Tips

Hey all, your site is great. I use Miro to catch the tech tips—wonderful. I was hoping that maybe sometime you might do a bit on putting your /home/username folder under some type of revision control. I am using Dropbox to copy my dot files over to a backup manually, but it is not very handy. Anyway, love the show and this site. Great work making Linux more accessible.

--

Shawn Bright

Thanks for the compliments! As to /home directory revision controls, I basically handle that with backuppc. Although a bit longer than a video tech tip could handle, setting up backuppc isn't too painful, and it keeps snapshots for as long as you have space. My favorite feature is how fast you can restore older versions of a file. It has a great Web interface and "one-click restore".—Ed.

Re: Linux on the Desktop, Part II

In the September 2009 issue, Cary's letter to the editor was very good, and I realized that when confronted with the positions that Windows works better out of the box for the computer-user masses than does Linux, I need to counter that a Linux-based machine purchased from a commercial source, such as Linux Certified, EmperorLinux or even Dell (and so on), is going to provide just as painless of an initial experience as would Windows from a commercial source such as Dell, HP or Acer (and so on). One caveat in the experience difference is in arcane peripherals, such as limited-production film-strip scanners, thermal printers and other specialty products. The manufacturer's lack of incentive to produce Linux drivers creates an experience void that scares away many a member of the computer masses. Nonetheless, Cary's letter was enlightening.

--

Edward Comer

I tend to ramble about things like this in length, so I'll try to restrain myself. I still think Linux supports so much *hardware out of the box, that it makes Windows look silly in comparison. That said, there are some areas where the Linux end user suffers, like you mention, due to manufacturers' lack of support. As geeks, we see the problem. End users just see it as a limitation (which it is).*

Another big problem is familiarity. Computers aren't "new and cool" anymore; everyone knows how to use them. Most people are familiar with Windows, and other stuff is scary because it's different. Apple has the same problem. Although I think OS X provides a much better user experience than Windows, Apple still suffers with low percentages in spite of its enormous marketing budget. I think people like me need to keep getting Linux into schools, where next-generation purchasers will gain familiarity with it.—Ed.

Cross-Platforming Teachers!

Shawn Powers' column on the pairing of open-source software and Windows [see Shawn's Current_Issue.tar.gz in the September 2009 issue] could not have been more perfectly timed for me. I am a high-school teacher who is a longtime Linux user. For years, I have advocated open-source programs and the Linux OS as viable alternatives to the Windows software my school district spends so much money to license each year. Just two days before I read Shawn's column, I taught a workshop for other teachers in my district of Florida. The subject of this workshop? Open-source software. We covered OpenOffice.org 3, Firefox and a plethora of other open-source programs that are available in Windows versions. The response from my teacher-students was overwhelmingly positive. I actually saw jaws drop open as these teachers realized that programs like OpenOffice.org give them functionality as good and often better than the commercial software they were accustomed to using. They also were impressed with Linux itself, because I chose to run the programs on a bootable version of Ubuntu I'd customized for the class with the programs I wanted to showcase.

For those of us who have been advocates for Linux for a long time now, perhaps the current recession has a bright side. My school district, which always has been disinclined to consider the merits of open-source software, soon will be switching from Windows server software to Apache servers and from Microsoft Office 2007 to OpenOffice.org 3. Why? Like many school systems in America, we are operating with less money than ever, so open-source software is suddenly appealing to the higher-ups in our district who are trying to save every cent. I'm hoping that once our teachers move to OpenOffice.org and see the quality and usability of open-source software, some of them also will be receptive to the idea of an open-source operating system. Since I taught my workshop two days ago, I have had several e-mail messages from teachers who tell me they've downloaded Ubuntu

9.04 and are trying it out. As Shawn Powers' column suggests, the road to converting others to Linux might very well begin by showing them the merits of open-source software in Windows or on the Mac. Who knows where they might go from there?

--
Mike Creamer

That's great news! I'm still trying similar things here at our school, but I made some horrible mistakes that burned a lot of bridges in the past. Unfortunately, there are many teachers who think "Linux" is that stupid thing Mr Powers likes so much. Stories like yours are very encouraging. I hope everyone reads it and is motivated to try something similar. This year, I'm rolling out LTSP5 on Ubuntu 9.04. I'm hoping the shiny factor helps with more people falling in love.—Ed.

Really?

Hello from sunny Sweden. In your They Said It column on page 17 of the August 2009 issue, you claim that IBM's Chairman Thomas Watson said, "I think there is a world market for maybe five computers." But, did he?

Wikipedia calls it: "Although Watson is well known for his alleged 1943 statement: 'I think there is a world market for maybe five computers', there is scant evidence he made it." The author Kevin Maney tried to find the origin of the quote, but has been unable to locate any speeches or documents of Watson's that contain this, nor are the words present in any contemporary articles about IBM. And, there's more here: **en.wikipedia.org/wiki/Thomas_J._Watson#Famous_misquote**. It's a fun quote to be sure, but I don't think it has the ring of truth. Anyway, love the magazine!

--
Daniel Lundh

Transcripts Please?

Please provide transcripts of the videos; this'll save on loading time (besides, Flash is pretty buggy on many 64-bit platforms), and make search easier and command copy & paste a cinch.

--
Jaco

This isn't the first time someone has asked for transcripts. The problem is that often the transcripts would be, "See this does that, and then see what happens here", which isn't terribly useful without the video.

We have tried to address the problem a bit by having more tech tips on the site in text format as opposed to all video. Hopefully, between the two, everyone will get a bit of something.—Ed.

PHOTO OF THE MONTH

Have a photo you'd like to share with *LJ* readers? Send your submission to publisher@linuxjournal.com. If we run yours in the magazine, we'll send you a free T-shirt.

Linux Journal **Stickers: $5.00; Vinyl Magnet Sheet: $6.00; Tux Guarding My Beer: Priceless. Submitted by Fred Richards**

LINUX JOURNAL

At Your Service

MAGAZINE

PRINT SUBSCRIPTIONS: Renewing your subscription, changing your address, paying your invoice, viewing your account details or other subscription inquiries can instantly be done on-line, **www.linuxjournal.com/subs**. Alternatively, within the U.S. and Canada, you may call us toll-free 1-888-66-LINUX (54689), or internationally +1-818-487-2089. E-mail us at subs@linuxjournal.com or reach us via postal mail, Linux Journal, PO Box 16476, North Hollywood, CA 91615-9911 USA. Please remember to include your complete name and address when contacting us.

DIGITAL SUBSCRIPTIONS: Digital subscriptions of *Linux Journal* are now available and delivered as PDFs anywhere in the world for one low cost. Visit **www.linuxjournal.com/digital** for more information or use the contact information above for any digital magazine customer service inquiries.

LETTERS TO THE EDITOR: We welcome your letters and encourage you to submit them at **www.linuxjournal.com/contact** or mail them to Linux Journal, 1752 NW Market Street, #200, Seattle, WA 98107 USA. Letters may be edited for space and clarity.

WRITING FOR US: We always are looking for contributed articles, tutorials and real-world stories for the magazine. An author's guide, a list of topics and due dates can be found on-line, **www.linuxjournal.com/author**.

ADVERTISING: *Linux Journal* is a great resource for readers and advertisers alike. Request a media kit, view our current editorial calendar and advertising due dates, or learn more about other advertising and marketing opportunities by visiting us on-line, **www.linuxjournal.com/advertising**. Contact us directly for further information, ads@linuxjournal.com or +1 713-344-1956 ext. 2.

ON-LINE

WEB SITE: Read exclusive on-line-only content on *Linux Journal*'s Web site, **www.linuxjournal.com**. Also, select articles from the print magazine are available on-line. Magazine subscribers, digital or print, receive full access to issue archives; please contact Customer Service for further information, subs@linuxjournal.com.

FREE e-NEWSLETTERS: Each week, *Linux Journal* editors will tell you what's hot in the world of Linux. Receive late-breaking news, technical tips and tricks, and links to in-depth stories featured on **www.linuxjournal.com**. Subscribe for free today, **www.linuxjournal.com/enewsletters**.

Polywell Linux Solutions

More Choices, Excellent Service, Great Value!

Serving the Industry for More Than 20 Years

Quiet Performance NAS Storage

8TB $1,999

12TB $2,999

24TB $5,999

- Dual Gigabit LAN
- RAID-5, 6, 0, 1, 10
- Hot Swap, Hot Spare
- Linux, Windows, Mac
- E-mail Notification
- Tower or Rackmount

NVIDIA GeForce

Silent Eco Green PC

NVIDIA® GeForce Graphics
Energy efficient, Quiet and Low Voltage Platform. starts at $199

The Best Terminal PC

LD-001

Fanless Silent Mini-ITX PC

with Solid State Drive DC Power In

ITX-10A Thin-PC $199

ITX-30A with PCI Riser

Nvidia® ION™ PC

ITX-40A with Blu-Ray or DVD

ITX-1000 with WiFi Option

Excellent for Linux Appliance

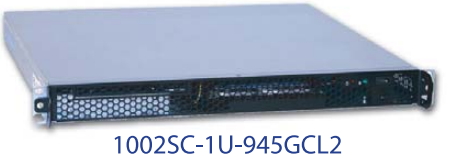
4U-24Bay 48TB Storage Server

Hardware RAID-6, NAS/iSCSI/SAN Storage
Mix SAS and SATA, 4 x GigaLAN or 10Gbit LAN

4024SS

Mini-1U Server for Data Center ISP

Intel Dual-Core or Quad-Core Processor, Dual GigaLAN
4GB to 8GB RAM, 2 x 500GB RAID HD
Linux Server Starts at **$499**



1002SC-1U-945GCL2

Polywell OEM Services, Your Virtual Manufacturer
Prototype Development with Linux/FreeBSD Support
Small Scale to Mass Production Manufacturing
Fulfillment, Shipping and RMA Repairs

- 20 Years of Customer Satisfaction
- 5-Year Warranty, Industry's Longest
- First Class Customer Service

888.765.9686
linuxsales@polywell.com
www.polywell.com/us/Lx

POLY POLYWELL

Polywell Computers, Inc 1461 San Mateo Ave. South San Francisco, CA 94080 650.583.7222 Fax: 650.583.1974

NVIDIA, ION, GeForce and combinations thereof are trademarks of NVIDIA Corporation. Other names are for informational purposes only and may be trademarks of their respective owners.

diff -u

WHAT'S NEW IN KERNEL DEVELOPMENT

Transcendent memory, called **tmem**, is a virtual form of RAM that can be given to user programs in copious amounts, provided those programs are okay with the fact that the tmem may vanish without warning. The **Xen** folks have implemented tmem for Xen, and now they want to provide a generic API for the kernel to make tmem available to any program that wants it. **Dan Magenheimer** and other Xen folks have been working on some patches, and it looks as though the kernel people are open to the tmem concept, so long as certain security issues are addressed. Security concerns actually drown out most other discussions, so it remains to be seen what technical problems remain before tmem could be included in the kernel tree.

Andrew Morton has taken over temporary maintainership of the **MMC** code. **Pierre Ossman** has stepped down as maintainer, and no one stepped up, so Andrew said he'd do it for now. **Ian Molton**, **Matt Fleming**, **Roberto A. Foglietta** and **Philip Langdale** all stopped just short of actually volunteering to be the new maintainer, though they all said they'd like to be CCed on all MMC patches. One benefit of the maintainership change over was that a bunch of MMC patches bubbled up that had been lying dormant for too long. **Paul Mundt**, **Ohad Ben-Cohen** and **Adrian Bunk** all submitted or pointed to MMC patches to be considered.

kernel.org may be getting some new **mailing-list software**, written by one of the kernel.org admins, **Matti Aarnio**. Aside from the fact that this is clearly a very fun project for him, the reasons behind it are not so clear. His code improves on majordomo security, and there are various other enhancements, but he also could have fed those features as patches to majordomo or one of the other popular list-handling tools around. One thing is clear. If kernel.org adopts a brand-new list-handling tool, a lot of other places will use it too.

PramFS, the nonvolatile RAM-based filesystem, keeps state across reboots, just like a normal filesystem. **MontaVista** tried to get it in the kernel back in 2004, but it was rejected because MontaVista was trying to get a patent on the algorithms. Now **Marco Stornelli** and **Daniel Walker** have said that MontaVista has abandoned its patent effort, and Marco wants to submit the code for inclusion again. But, it turns out that this is not a full-featured filesystem. There's no support for symbolic links, and there are other technical questions as well. One obvious question that was asked during the discussion was why PramFS was necessary at all. Why not just extend an existing filesystem to support nonvolatile RAM? **Pavel Machek** led the charge against PramFS and argued vehemently against accepting the PramFS code as is. He saw no justification for the project and said that before it even could be considered, it would have to implement modern features, such as journaling and other features that come standard with many newer filesystems today.

Microsoft has GPLed its **Hyper-V** drivers, and it will allow the in-kernel versions of that code to be the canonical versions. Future Microsoft contributions will be made as patches to those kernel drivers, rather than as full releases of their own. **Greg Kroah-Hartman** announced the occasion, praising Microsoft's **Hank Janssen**, **Haiyang Zhang** and **Sam Ramji**, as well as numerous non-Microsoft people, for helping get this done. Some of the Microsoft people, including Hank, said they intend to continue their work on these drivers as community contributors.

—ZACK BROWN

They Said It

Don't worry about what anybody else is going to do. The best way to predict the future is to invent it.

—Alan Kay

Premature optimization is the root of all evil (or at least most of it) in programming.

—Donald Knuth

We're even wrong about which mistakes we're making.

—Carl Winfeld

`filter(P, S)` is almost always written clearer as `[x for x in S if P(x)]`.

—Guido van Rossum on Python

Lisp has jokingly been called "the most intelligent way to misuse a computer". I think that description is a great compliment because it transmits the full flavor of liberation: it has assisted a number of our most gifted fellow humans in thinking previously impossible thoughts.

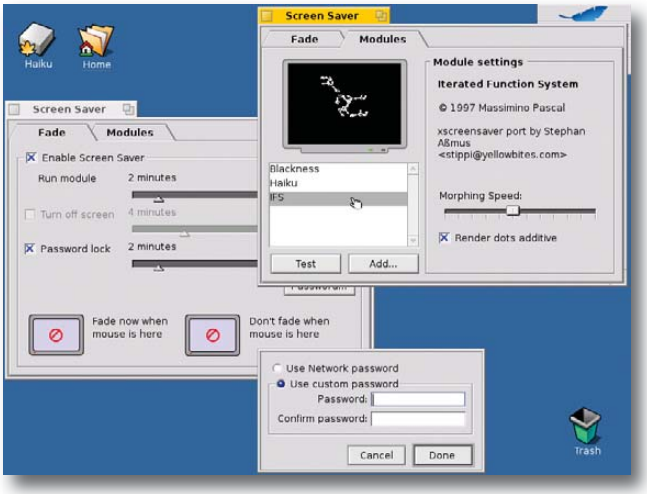
—Edsger Dijkstra, CACM, 15:10

A government big enough to give you everything you want, is big enough to take away everything you have.

—Thomas Jefferson

NON-LINUX FOSS

Haiku is a free and open-source operating system designed to be compatible with BeOS. BeOS was the operating system that ran on computers built and sold by Be, Inc., in the 1990s and also on Apple's PowerPC reference platform. BeOS was designed for working with digital media and took advantage of modern hardware. It worked on multiprocessor systems and extensively used multitasking and multithreading. BeOS was not built to look like another *nix system and neither is Haiku. It is not based on Linux nor does it use the X Window System or GNOME or KDE.



The Haiku Screen Saver Preferences Applet (from www.haiku-os.org)

Haiku is written in C++, as was BeOS before it, and the operating system API is object-oriented. As of 2008, Haiku can be compiled from within Haiku itself. As of 2009, there is a native GCC4 port that now allows numerous applications to be ported to Haiku. A Java port for Haiku also is in progress.

Haiku began in 2001 and was named OpenBeOS until 2004, when the name was changed to avoid problems with the original trademarks (and also because the original name required too many Shift-key presses). Haiku is released under the MIT license. Haiku currently is bootable and usable, but it has not reached version 1.0 yet (R1 in Haiku speak).

—MITCH FRAZIER

LinuxJournal.com

This month's *Linux Journal* is all about infrastructure. Want a broader view? Visit us at LinuxJournal.com for more of our editors' insights on infrastructure as it applies to Linux, open source and Web technology.

These articles should get you started:

- "Comparing Hard and Soft Infrastructure": **www.linuxjournal.com/content/comparing-hard-and-soft-infrastructure**
- "Understanding Infrastructure": **www.linuxjournal.com/content/understanding-infrastructure**
- "Building a Multisourced Infrastructure Using OpenVPN": **www.linuxjournal.com/article/9915**
- "Why Internet & Infrastructure Need to Be Fields of Study": **www.linuxjournal.com/content/why-internet-infrastructure-need-be-fields-study**

—KATHERINE DRUCKMAN

LJ Index November 2009

1. Number of open-source C files available on the Internet (duplicates removed): **11,500,000**
2. Number of open-source Java files: **10,600,000**
3. Number of open-source C++ files: **8,640,000**
4. Number of open-source PHP files: **3,960,000**
5. Number of open-source Perl files: **1,820,000**
6. Number of open-source Python files: **1,570,000**
7. Number of open-source Ruby files: **952,000**
8. Number of open-source FORTRAN files: **374,000**
9. Number of open-source COBOL files: **9,000**
10. Number of open-source "Hello World" programs: **198,000**
11. Number of open-source versions of stdio.h: **4,000**
12. Number of open-source files containing "TODO:" comments: **1,640,000**
13. Number of open-source files containing "FIXME:" comments: **1,230,000**
14. Number of open-source files containing the word "hack": **901,000**
15. Number of open-source files containing the "F word": **88,800**
16. Number of Linux distros listed on linux.org: **220**
17. Number of Linux distros listed on distrowatch.com: **309**
18. Result count difference between Yahoo and Google searching for "Linux": **1,023,000,000**
19. US National Debt as of 08/03/09, 11:18:07am MST: **$11,595,953,181,678.30**
20. Change in the debt since last month's column: **$94,411,207,892.70**

Sources: *1–15: Google Code Search (www.google.com/codesearch) / 16: www.linux.org/dist 17: distrowatch.com / 18. Yahoo and Google (Yahoo returns the higher count) 19. www.brillig.com/debt_clock / 20. Math*

People, Research, Excellence

THE PAST FEW MONTHS in this space, I've covered specific utilities and how they can be used, sometimes in quite interesting ways. This month, I instead look at a task and see what utilities are available to accomplish it. People who do scientific computational work tend to use several pieces of software in series. This software could span the entire computer age in terms of how old it might be. The usual work flow involves taking some initial data and feeding it as input to a program, in order to do the first computational step. The output then is fed as input to another program, in order to complete the second computational step. This process continues until the final results are reached. The problem with this method is that the programs used at each computational step probably were written by completely different groups, possibly decades apart. This means the researcher may need to do some kind of transformation to get the output from one computational step into the proper format to be used as input for the next computational step.

One simple yet common problem is the use of different field separators in a data file. In some cases, fields may be separated by commas. In other cases, they may be separated by tab characters. If you have to change from one to the other, you can use the tr utility:

```
tr "," "\t" <data_file_1 >data_file_2
```

The above replaces every comma in data_file_1 with a tab and writes the results into data_file_2. This works well for replacing single characters or even classes of characters. Say you had a really old piece of FORTRAN code that expected all letters to be uppercase. You could accomplish that with the following:

```
tr "[:lower:]" "[:upper:]" <data_file_1 >data_file_2
```

But, what if you have some more-complicated translation to make? A more general-purpose utility to use for this is sed, the Stream EDitor. With sed, you can make substitutions with the s command. For example, you can achieve the same result as above, converting commas to tabs, by running:

```
sed -e "s/,/    /g" data_file_1 >data_file_2
```

(The blank space after the second forward slash is a tab character.) Remember: to type a tab character in the bash shell, you need to type C-v TAB. Using this command, you can translate any kind of separator into any other kind of separator. And don't think it can't happen to you. I personally have seen separators like |*| or %*% in the wild. You never know what some previous person is going to think is a good idea.

So, now you have your data fields separated with the correct separator, but what if you need only some of this data? The output file you are massaging may have more data than you need for the next computational step. What can you do? The cut and paste utilities can be used for this purpose. You can cut selected columns out of the data file with:

```
cut -f1,3 data_file_1 >data_file_2
```

This cuts columns 1 and 3 and dumps them into data_file_2. It assumes that the field separator is a tab character. If you've used a different separator character, you can use the -d option. For example, the following cuts the file up using comma separators:

```
cut -f1,3 -d "," data_file_1 >data_file_2
```

If you have the opposite problem, you can use the paste utility to glue together data from multiple files. Say you have two data files containing the parts required for the next computational step. You can glue them together with:

```
paste data_file_1 data_file_2 >data_file_3
```

This assumes that you want to use a tab character as the field separator. If you want to use another character, such as a comma, you can use the -d option, like this:

```
paste -d "," data_file_1 data_file_2 >data_file_3
```

Another very useful utility can be used to do this type of job, awk. With awk, you can pull out only the data you need. For example, say your output file has three columns of data, but the next computational step requires only columns 1 and 3. With awk, this becomes a very simple task by executing the following:

```
awk '{print $1,$3}' data_file_1 >data_file_2
```

This example assumes that the initial field separator in the data_file_1 is a tab character. You get columns 1 and 3, with a comma as the field separator, dumped into the data_file_2. If you want to keep the tab character, use the following instead:

```
awk '{print $1"\t"$2}' data_file_1 >data_file_2
```

If your initial data file, data_file_1, uses a comma as a field separator, you can tell awk this with the -F option:

```
awk -F "," '{print $1,$2}' data_file_1 >data_file_2
```

With these options, you can do the field separator translation and the cut function both in one step.

With awk, you can do even more impressive data massaging. Say you need to use the average of the three columns as input to the next computational step. Do the following:

```
awk '{print ($1+$2+$3)/3}' data_file_1 >data_file_2
```

awk makes an entire programming language available, and it can be used for very complex data massaging. Hopefully, this short introduction shows some of the possibilities available for your data management tasks and helps smooth the work flow between computational steps.

—JOEY BERNARD

Netbooks Only: Operating Systems for the Little Guys

I have a love/hate relationship with Netbooks. On the upside, they are inexpensive, portable and beefy enough to run most applications. On the downside, they have small screens and small keyboards. Although I'm not normally a fan of non-standard desktops, a few Netbook-specific Linux distributions make a valid case for their existence—especially on small screens.

Ubuntu Netbook Remix Although not the first, UNR is one of the most "open" of the Netbook-specific interfaces. Underneath it is running (of course) Ubuntu, but the interface is designed for tiny screens (Figure 1). Its large icons and lack of tiny menus make it easy to navigate, even from low resolutions. Some of the applications still are awkward at small screen sizes, but UNR does a nice job of making the most of screen real estate. Check it out at www.ubuntu.com/GetUbuntu/download-netbook.

Figure 1. Ubuntu Netbook Remix runs Ubuntu underneath this Netbook-friendly interface.

Moblin Although still in beta, Moblin is a Linux distribution built from the ground up designed for mobile devices (Figure 2). More than merely a wrapper around an underlying operating system, Moblin offers an entirely unique user experience. I personally had a difficult time figuring out how to use Moblin, but it's early in development and available to try out. To download a current live image, visit moblin.org/downloads.

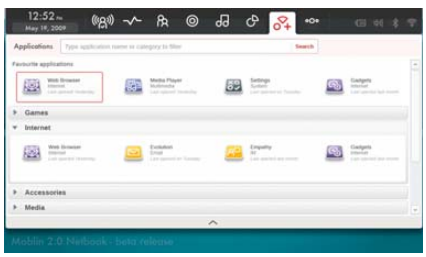


Figure 2. Moblin is designed to work on a variety of handheld devices, but Netbooks are one of the current foci.

Jolicloud Jolicloud is not only early in development, but is actually in a closed beta program. I heard about the operating system on Twitter, and when I checked it out, it was rather impressive (Figure 3). Although the interface does appear to be very user-friendly, and it's designed for a cramped screen size, the real difference with Jolicloud is its goal to move all your information to the "cloud". Operating systems like gOS have attempted to do the same, without much success. Hopefully, Jolicloud's sync/cache method to switch between Netbooks and desktops will be able to move us effectively to Web-based applications without sacrificing the need for off-line usability. You can sign up for the closed testing phase at my.jolicloud.com/account/invitation.

Figure 3. Jolicloud looks impressive. Only time will tell whether it will provide a dependable Web-immersed environment.

Chrome OS: the Future, or Too Little Too Late? A while back, Google announced its upcoming operating system Chrome. It remains to be seen whether the Chrome OS will be a dominant force in the Netbook market or whether it will face the same limited fanfare that the Chrome browser receives. It also begs the question of whether Google waited too long for such an endeavor. With options like UNR, Moblin and soon Jolicloud, Netbooks finally might become more than a novelty. I just hope hardware manufacturers can make some decent keyboards for them!

—SHAWN POWERS

YOUR TECHNOLOGY SHOULD WORK AS FAST AND SMART AS YOU DO.

And choosing Linux should never limit your technology options.

We have more than 500 Service Providers serving more than 12 million end-users in 125 countries with our Linux-based solution.

Talk to the people who know Linux.

Talk to Parallels.

Parallels™
Optimized Computing

offer@parallels.com
425.282.6448
www.parallels.com

SuperGamer, 8GB of Linux-Only Gameplay

I admit, I'm one of those people who dual-boots so I can play video games. I've tried running programs like CrossOver Games in order to feed my need for fragging, but in the end, it seems I always have to install Windows to enjoy some real gaming fun. Thankfully, I'm not the guy in charge of things worldwide, because the folks over at **www.supergamer.org** have created a bootable, dual-layer DVD full of native-running Linux games. Yes, I said native. Check out the impressive list of preinstalled games you'll get when you download the ISO:

This is SuperGamer's official screenshot. Notice all the game icons on the bottom of the screen.

- *Quake Wars*
- *Doom 3*
- *Prey*
- *Unreal Tournament*
- *Quake 4*
- *Savage 2*
- *Postal 2*
- *Enemy Territory*
- *Penumbra Black Plague*
- *Sauerbraten*
- *Urban Terror*
- *Soldier of Fortune*
- *Torcs*
- *Tremulous*
- *AlienArena*
- *True Combat*
- *America's Army*
- *Nexus*
- *OpenArena*
- *PlaneShift*
- *Drop Team*
- *Frets On Fire*
- *Chromium B.S.U.*
- *Mad Bomber*
- *X-Moto*
- *BZ Flag*
- *Mega Mario*
- *Glaxium*
- *GL-117*
- *NeverBall*
- *NeverPutt*
- *Super Tux*
- *PPRacer*

So much for gaming being a Windows-only adventure! Based on Vector Linux, SuperGamer is ready to perform on all modern video cards without additional downloads. Just pop it in, boot it up, and frag.

—SHAWN POWERS

A Bit of Welcomed Scumm on Your Linux Machine

This might make me sound like an old fogey, but I really do miss the old games like *Space Quest*, *The Curse of Monkey Island* and *Return to Zork*. The problem isn't that I don't have the games anymore, but rather that they were designed for my 386 computer running DOS. Thankfully, I'm not alone in my fits of nostalgia. The developers over at **www.scummvm.org** have reproduced the "Script Creation Utility for Maniac Mansion" developed by Lucas Arts and packaged it into a virtual machine (thus, ScummVM). That virtual machine is open source and available for just about any platform you can imagine.

It's important to note that ScummVM doesn't come with any actual games. You either need to purchase the old games it supports from eBay or look in your closet for those stacks of old game disks you used to play as a kid (or as an adult, for some of us). What ScummVM does provide is a platform for playing those old games and even introducing your kids to games they'll probably never have a chance to play otherwise. At OSCON in July 2009, ScummVM was announced as SourceForge's Community Award winner in the category of "Best Project for Gamers". If you've never checked it out, now is a good time.

As for me? I think it's about time I introduced *Putt Putt & Fatty Bear* to my kids. Hopefully, they enjoy the games as much as my brother-in-law did when he was growing up. *Monkey Island*? Yeah, I think I'll play that one myself.

—SHAWN POWERS

Figure 1. If you wonder what the mustached car could possibly be saying, you'll have to play *Putt Putt Joins the Parade*. Thankfully, it's installable on ScummVM.

ABERDEEN
SERVERS AND STORAGE

HOW MUCH VMWARE DO YOU NEED?

VMware Certified NAS and Servers

The scalable AberNAS is the only Network Attached Storage appliance in its class to carry the VMware® Ready Certification. Aberdeen enables VMware-centric organizations of any size to affordably take advantage of AberNAS's robust features such as; superior RAID controller speed, increased scalability and optional NAS-to-NAS mirroring.

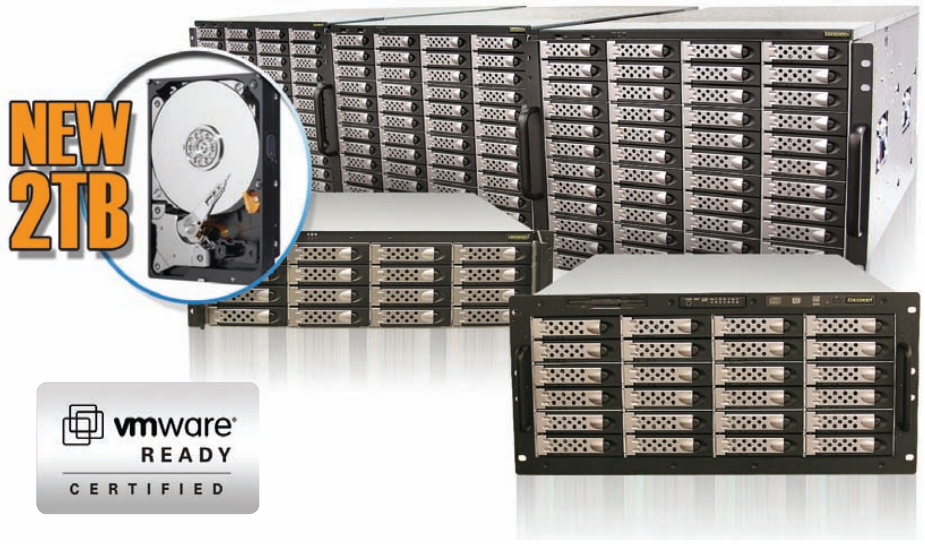
VMware Certified Aberdeen Stirling servers offer a lower cost of ownership, while eliminating concerns of compatibility complications. Stirling servers possess exceptional benefits designed to deliver improved operational efficiency within a dynamic VMware environment.

ABERNAS

VMware Certified NAS Up To 100TB

- Quad-Core Intel® Xeon® Processor 5500 Series
- 3GB/6GB 1333MHz DDR3 memory
- Supports both SAS & SATA Storage Drives
- RAID 0, 1, 5, 6, 10, 50, 60 Capable
- SAS Expansion Ports & iSCSI Target Capable
- Dual/Quad Gigabit Ethernet Ports
- **5-Year Warranty**

1U - 8TB NAS Starting at $4,995
2U - 24TB NAS Starting at..... $10,495
3U - 32TB NAS Starting at..... $12,995
5U - 48TB NAS Starting at..... $17,995
6U - 64TB NAS Starting at..... $23,495
8U - 100TB NAS Starting at..... $33,995



intel Xeon inside™
Powerful. Intelligent.

STIRLING 123

1U Dual Xeon Server

- Up to 2x Quad-Core Intel® Xeon® Processor 5400 Series
- 32GB FBDIMM max memory
- Up to 4 SATA or SAS drives
- 650W redundant power
- Pre-installed VMware® ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $1,845

STIRLING 132T

1U Twin Node Server

- Up to 2x Quad-Core Intel Xeon Processor 5400 Series
- 64GB FBDIMM max memory
- Up to 2 SATA drives
- 980W high efficiency power
- Pre-installed VMware ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $2,675

STIRLING 209

2U Xeon Server

- Single Quad-Core Intel Xeon Processor 5500 Series
- 24GB DDR3 max memory
- Up to 6 SATA drives
- 400W high efficiency power
- Pre-installed VMware ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $1,495

STIRLING 244

2U Quad Xeon Server

- Up to 4x Six-Core Intel Xeon Processor 7300 Series
- 192GB FBDIMM max memory
- Up to 6 SATA or SAS drives
- 1200W redundant power
- Pre-installed VMware ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $5,275

STIRLING 269

2U Dual Xeon Server

- Up to 2x Quad-Core Intel Xeon Processor 5500 Series
- 144GB DDR3 max memory
- Up to 8 SATA or SAS drives
- 720W redundant power
- Pre-installed VMware ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $2,095

STIRLING 444

4U Quad Xeon Server

- Up to 4x Six-Core Intel Xeon Processor 7300 Series
- 192GB FBDIMM max memory
- Up to 5 SATA or SAS drives
- 1200W redundant power
- Pre-installed VMware ESXi on Disk-on-Module
- **5-Year Warranty**

Starting at $5,345

Intel, Intel Logo, Intel Inside, Intel Inside Logo, Pentium, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. For terms and conditions, please see www.aberdeeninc.com/abpoly/abterms.htm. lj031

888-297-7409
www.aberdeeninc.com/lj031

RSpec for Controllers

More on RSpec's "outside-in" approach to testing.

REUVEN M. LERNER

RSpec is a popular testing framework for Ruby programmers that works on the principle of behavior-driven development (BDD). BDD distinguishes itself from test-driven development (TDD) in that it looks at programs from the outside, rather than from the inside, considering code as a user or observer, as opposed to an implementer. In the BDD world, you don't implement tests, but rather specifications; if the specification passes, the code is doing what it is supposed to do.

As with many things in the Ruby arena, RSpec has become particularly popular among users of the Rails framework for Web development. Last month, I discussed RSpec in the context of testing Rails models (that is, classes that connect to the relational database). This month, I look at the slightly more complicated case of controller testing. Controller testing is more complicated because it requires that you consider a few more cases, or at least different cases. Now you have to consider inputs from the outside world, in the form of HTTP requests. It also introduces the need for mocks and stubs, objects you can use to test your controllers without having to create real objects (and the database that sits behind them).

Controller testing is more complicated because it requires that you consider a few more cases, or at least different cases.

This month, I examine some of the ways the RSpec testing framework allows you to test controllers in your Ruby on Rails applications. Along the way, I consider what it means to test controllers and how much you might want to test them. Finally, I take a quick look at the world of mocks and stubs, and show how they can help improve your testing.

A Simple Application

Last month, I started building a simple appointment calendar as an example. As it happens, I implemented only a small part of that appointment calendar, creating a single person model, which you can use to represent the people with whom you will meet. Now, let's create appointments as well:

```
./script/generate rspec_model appointment starting_at:timestamp \
    ending_at:timestamp person_id:integer location:text notes:text
```

As you might expect, you will enhance your model files by linking them together, indicating that each person has_many appointments, but that each appointment belongs_to one person. That'll allow you to use Ruby's object-oriented syntax to retrieve person.appointments, or appointment.person.

Now that you have two models in place, you should do something with them. One obvious thing to do is list today's appointments. In BDD fashion, let's write a spec that describes what the system should do; you actually will implement the code afterward.

The spec will describe how you want to be able to see a list of appointments. Let's assume that the specs for the models (people and appointments) are in place, and that you now can concentrate on your controllers. Basically, you want an appointment controller whose index action shows all current appointments. You can do that by generating such a controller:

```
./script/generate rspec_controller appointment index new create show
```

Create a controller named appointment, along with a few actions named similarly to a purely RESTful controller (which this is not). Now, open up spec/controllers/appointment_controller_spec.rb, which is the location of the spec file for this controller, and you will see a number of simple specs, one for each of the methods you've defined. As I explained last month, RSpec's power is its readability, with "describe" blocks that indicate an overall context, "it" blocks that describe specifications, and then individual assertions, which are written as "something SHOULD be-something". The initial, automatically generated spec for the index action, thus, looks like this:

```
describe "GET 'index'" do
  it "should be successful" do
    get 'index'
    response.should be_success
  end
end
```

Mocking

The response object is given automatically in controller specs, and it allows you to do such things as check for success. The thing is, you also want this index action to retrieve (and display) all the current

appointments in your database. How can you test for that?

One way is to load your database with a bunch of fake data, or "fixtures", and actually retrieve the data from the database. But hey, you're trying to test the controller here, not the database—so going to the database is going to be massive overkill.

What you can do instead is tell Ruby you expect the controller to request a bunch of appointment objects. Indeed, it should request all the appointments in the database, as per your specification. So long as it does that, you can rest assured that the action's specification has been met.

You can do this by switching your normal Appointment object with a mock, sometimes called a test double object. This mock object allows you to check that the right things are happening, while staying within your program. For example, if you want to make sure that Appointment.find(:all) is being invoked, modify your spec to read as follows:

```
describe "GET 'index'" do
  it "should be successful" do
    appointments = [mock(:appointment), mock(:appointment)]
    Appointment.should_receive(:find).and_return(appointments)
    get 'index'
    response.should be_success
  end
end
```

Here, two lines are added before the invocation of "get 'index'". In the first line, you create an array of two mock objects, each of which will claim (if asked) that it is an instance of Appointment. It isn't a real appointment object, of course, but rather a thin layer meant just for testing. You will create two such objects, so you can pretend that there are multiple appointments in your database.

The next line is even more interesting. It says that Appointment (the class) should expect to receive the find method at some point. Notice that the placement here is important; if you were to put this mocking line after the invocation of GET, it would be too late. Instead, set up the mock such that the GET method can do things appropriately. If the mock doesn't receive an invocation of "index",

Expert included.

AMD Opteron™ 64

As a Sales Engineer for Silicon Mechanics, Scott is here to consult with you, understand your processing and storage needs, and then configure the systems that will work for you. Energy efficiency, compute density, reliability, and serviceability are just some of the things he makes a priority. That's why Scott is excited about the Bladeform 8840 Blade for the Bladeform 8100 Series Blade Server Platform. Each Bladeform 8840 Blade supports four of the latest Six-Core AMD Opteron™ 8000 Series processors. With 24 cores per blade and 10 blades per enclosure, you can pack the power of 240 cores in just 7U of space.

At the same time, Scott is impressed that the Bladeform 8100 provides 90%+ high-efficiency redundant power supplies for operating cost reduction and earth-friendly computing. The Bladeform 8100 Platform with the Bladeform 8840 Blade is a perfect choice for mission-critical enterprise applications as well as scale-out and high performance computing environments.

When you partner with Silicon Mechanics, you get more than a high-efficiency AMD solution — you get an expert like Scott.

SILICON MECHANICS

visit us at www.siliconmechanics.com
or call us toll free at 866-352-1173

See the Silicon Mechanics Bladeform 8100 and the 8840 Blade at www.siliconmechanics.com/8840

Silicon Mechanics and the Silicon Mechanics logo are registered trademarks of Silicon Mechanics, Inc. AMD, the AMD Arrow logo, AMD Opteron, and combinations thereof, are trademarks of Advanced Micro Devices, Inc.

RSpec exits with a fatal error. Indeed, using BDD methods, that's exactly what I can expect to see after I run RSpec:

```
Spec::Mocks::MockExpectationError in 'AppointmentController
 ➥GET 'index' should be successful'
<Appointment(id: integer, starting_at: datetime, ending_at:
 ➥datetime, person_id: integer, location: text, notes:
 ➥text, created_at: datetime, updated_at: datetime)
 ➥(class)> expected :find with (any args) once,
 ➥but received it 0 times
```

In other words, the example above says you want Appointment to have its "find" method called, but that never happened. Thus, add that invocation of find to the index action:

```
def index
  Appointment.find(:all)
end
```

Now the spec passes (thanks to the mock object), and you have functionality. What could be better? Well, perhaps you want to test the output you see in the view that displays that object. I'm not going to go into it here, but RSpec allows you to test views as well, using a similar mechanism that looks at the resulting HTML output.

Indeed, I have begun to scratch only the surface of what is possible with RSpec's mocking mechanism. You can stub out specific object methods, allowing you to use models without their overhead or dependencies. For example, you could replace calls to "find" with a mock object that you return, and ignore any calls to "save"—thus, allowing you to work with real models, but faster and more reliably.

You also can imagine how you could test your ability to retrieve models that are associated with one another using mocks. For example, the "index" method probably would be useless if it displayed only appointments. You probably would want to show the person with whom the appointment was scheduled. That requires traversing a foreign key association, which you easily can take care of with stub objects that you then reference from within your mock.

Now, you might be wondering if all this would be possible with either fixtures or factories. The answer is yes, and different developers have used fixtures and factories successfully over the years. I generally find fixtures to be the most natural of the bunch to understand and to use, but the fact that they go through the database and require that I set up and coordinate each of the individual objects begins to take its toll as a project gets larger. I also enjoy using factories and have been experimenting (as I mentioned a few months back) with different factory classes.

But, the more I'm exposed to mocking, the more I wonder if the entire factory class is necessary, or if I simply can use mocks and stubs to pinpoint and use the functionality that interests me. I'm sure other developers are thinking about these considerations as well, and I hope the plethora of options available to Ruby developers will improve and encourage the culture of testing that is already so strong in the Ruby community.

Conclusion

RSpec's "outside-in" approach to testing takes a bit of getting used to, but I increasingly have found it to be a method that forces me to think harder about my code, as well as about my testing strategy. That said, I'm not sure if I really have a strong preference for RSpec over similar BDD-style tools, such as Shoulda, which works with Ruby's traditional Test::Unit system. The bottom line is that you should try to include as much automated testing as possible in any software you design—not only because it will benefit your users, but also because it will benefit you as a developer.■

Reuven M. Lerner, a longtime Web/database developer and consultant, is a PhD candidate in learning sciences at Northwestern University, studying on-line learning communities. He recently returned (with his wife and three children) to their home in Modi'in, Israel, after four years in the Chicago area.

Resources

The home page for RSpec is **rspec.info**, and it contains installation and configuration documentation, as well as pointers to other documents.

The Pragmatic Programmers recently released a book called *The RSpec Book*, written by RSpec maintainer David Chelimsky and many others actively involved in the RSpec community. If you are interested in using RSpec (or its cousin, the BDD tool Cucumber), this book is an excellent starting point.

An RSpec mailing list, which is helpful and friendly, but fairly high volume, is at **groups.google.com/group/rspec**.

Finally, a good introduction to RSpec and mocking is in *The Rails Way*, one of my favorite books about Rails, written by Obie Fernandez. This book describes mocking both within the context of RSpec and as a general testing tool when developing Rails applications.

The Apache Software Foundation and Stone Circle Productions are pleased to present

ApacheCon US 2009

The official conference of the ASF bringing together Open Source developers, power users, programmers, and technologists from around the world.

Sign Up Today!

Receive 10% OFF Full Conference and Single Day Passes. To receive the 10% discount, use discount code "Community"

http://www.us.apachecon.com

Hackathon | Expo | MeetUps | BarCamp | Trainings | Conference | BoFs

November 2 - 6

Marriott Oakland City Center
Oakland, California

Sponsors: thawte, The Apache Software Foundation, Aster, cloudera, facebook, FUSE, HALO, Hotwax Media, hp invent, HIPPO, LinkedIn, lucid IMAGINATION, MARK LOGIC, spring source, WSO2

2009 marks the 10th anniversary of the ASF. ApacheCon will be a unique and exciting celebration of ten years of great technology, the evolution of an amazing community, and a look at the future of Open Source and The Apache Software Foundation.

The conference features a full schedule, including many FREE events, along with inside looks at some of the ASF's most popular established projects, as well as cutting-edge technologies and industry topics that include:

- Web servers and servlet containers: Apache HTTP Server and Tomcat
- Cloud computing frameworks/tools: Hadoop, Lucene, and Solr
- OSGi-based development: Felix and ServiceMix
- Enterprise Web Services and SOA frameworks: CXF, Tuscany, ActiveMQ, OFBiz and Camel
- Databases, content management and support technologies: CouchDB and WebDAV
- Web application frameworks and build automation models: Sling, Wicket, and Maven
- Business and community issues such as Open Source licensing, business models, documentation, and collaborative development
- A big blow-out Anniversary Party and Reception … and much more!

Join us at ApacheCon to usher in the next decade of "The Apache Way"

DAVE TAYLOR

Exploring Lat/Lon with Shell Scripts

Never get lost at the command line again.

With the rise of geolocation systems on mobile devices (think "around me" on the Apple iPhone), a consistent method of measuring points on Earth has become quite important. The standard that's used is latitude and longitude, which measure the distance north or south of the equator and the distance east or west of the prime meridian (which goes through Greenwich, England). Your GPS devices all understand this notation, as does Google Maps, Yahoo Maps, MapQuest and so on.

From a shell scripting perspective, we're interested in both being able to identify lat/lon for a point on the Earth and then, armed with that information, to see if we can calculate the distance between two points on the planet.

Figure 1. The White House

The first seems almost insurmountably hard until you learn that Yahoo Maps has a very simple API that lets you specify a URL that includes a street address and returns an XML object that includes its lat/lon values.

Let's start by creating a simple script where you can specify a street address and it will output lat/lon values.

Where Is This Place?

For example, you might be familiar with 1600 Pennsylvania Avenue, Washington, DC. I know you've seen pictures of the place. What's its lat/lon?

```
$ u='http://api.maps.yahoo.com/ajax/geocode'
$ a='?appid=onestep&qt=1&id=m&qs=1600+pennsylvania+ave+washington+dc'
$ curl "$u$a"
```

```
YGeoCode.getMap({"GeoID"      : "m",
                 "GeoAddress" : "1600 pennsylvania ave washington dc",
                 "GeoPoint"   : {"Lat" : 38.89859,
                                 "Lon" : -77.035971},
                 "GeoMID"     : false,
                 "success"    : 1} ,1);
<!-- xm6.maps.re3.yahoo.com uncompressed/chunked
     Tue Aug  4 12:16:51 PDT 2009 -->
```

Note that the output actually comes back as two lines; the data above, and in the other examples, has been reformatted to make it more readable.

Skim that return object, and you'll see Latitude = 38.89859 and Longitude = -77.035971. Feed those two into Google Maps as "38.89859,-77.035971" as a check, and you'll find the image shown in Figure 1.

You guessed it, it's the street address of the White House.

Let's start by creating a simple script where you can specify a street address and it will output lat/lon values.

Scripting Our Solution

The first part is easy: take whatever was specified

on the command line, and "recode" it to be URL-friendly. Then, append that to the Yahoo API URL, and output the results of a curl call:

```
#!/bin/sh

url='http://api.maps.yahoo.com/ajax/geocode'
args='?appid=onestep&qt=1&id=m&qs='
converter="$url$args"

addr="$(echo $* | sed 's/ /+/g')"
curl -s "$converter$addr"
exit 0
```

Let's test it with a different address this time:

```
$ sh whereis.sh 2001 Blake Street, Denver, CO
YGeoCode.getMap({"GeoID"       : "m",
                 "GeoAddress" : "2001 Blake Street, Denver, CO",
                 "GeoPoint"   : {"Lat" : 39.754386,
                                 "Lon" : -104.994261},
                 "GeoMID"     : false,
                                 "success"    : 1}, 1);
<!-- x1.maps.sp1.yahoo.com uncompressed/chunked
     Tue Aug  4 12:37:44 PDT 2009 -->
```

You can figure out what's at this address if you like. More important, you can see that this simple four-line script does the job—sort of.

Cleaning Up the Output

What we really want, however, is to extract just the lat and lon values and toss everything else out. This can be done with a bunch of different tools, of course, including Perl and awk, but I'm a rebel, so I use cut instead.

To do this, we need to count the double quotes (") in the output block. The 12th double quote is immediately before the latitude value, and the 15th is immediately after the longitude value. If we just worked with that, we would get:

```
$ sh whereis.sh 2001 Blake Street, Denver, CO | cut -d\" -f13-15
:39.754386,"Lon":-104.994261},
```


SILICON MECHANICS
visit us at www.siliconmechanics.com
or call us toll free at 866-352-1173

intel Xeon inside
Powerful. Intelligent.

Carlos knows a good thing when he sees it, but as one of the Silicon Mechanics Quality Assurance Experts, that's his job. When it comes to the new Storform iServ R515, Carlos sees more than just the care and attention to detail that go into the production of your custom-configured, versatile storage server.

The new R515 from Silicon Mechanics is a 4U server with 24 hot-swap SAS/SATA drives, optional integrated SAS controller, 6 PCI expansion slots, and redundant power. It features the new Intel® Xeon® processor 5500 series with Intel Turbo Boost Technology that delivers additional performance automatically when you need it. It also offers smart features like energy-saving DDR3 memory. With 2 internal SATA drive bays, you may want to think about incorporating Intel X25-E Extreme Solid State Drives for fast, reliable, energy-efficient OS drives. And Carlos will make sure that the servers we build and ship to you are exactly what you ordered.

When you partner with Silicon Mechanics, you get more than cost-effective, flexible, high-capacity storage—you get an expert like Carlos.

For more information about the Storform iServ R515 visit www.siliconmechanics.com/R515

Expert included.

Silicon Mechanics and the Silicon Mechanics logo are registered trademarks of Silicon Mechanics, Inc. Intel, the Intel logo, Xeon, and Xeon Inside, are trademarks or registered trademarks of Intel Corporation in the US and other countries.

Okay, so that's most of the work. Better, though, is to specify two different specific fields (13,15 rather than 13-15):

```
$ sh whereis.sh 2001 Blake Street, Denver, CO | cut -d\" -f13,15
:39.754386,":-104.994261},
```

That's 99% of what we want. Now we just need to clean up the noise. To do that, I'll jump back into the script itself, rather than experimenting on the command line:

```
curl -s "$converter$addr" | \
     cut -d\" -f13,15 | \
     sed 's/[^0-9\.\,\-]//g'
```

And testing:

```
$ sh whereis.sh 2001 Blake Street, Denver, CO
39.754386,104.994261,
```

Almost. Really, really close. But, that last comma is not wanted. Hmmm....

Because Earth is an oblate spheroid, not a perfect sphere, I expect this will have some small level of error, but let's proceed and see where we get.

Okay! To delete the last comma, we simply need to add a second substitution to the sed statement, so that the full sed expression is now:

```
sed 's/[^0-9\.\,\-]//g;s/,$//'
```

(The invocation is substitute/old-pattern/new-pattern/.)

Now we've got what we set out to create initially. Let's try it with yet another address:

```
$ sh whereis.sh 1313 S. Disneyland Drive, Anaheim CA
33.814413,-117.924424
```

Yep, that's the parking structure for Disneyland in California.

Distance between Two Points

Now comes the hard part of this, actually. We can get the lat/lon of any address we desire, but calculating the distance between two points is a bit more tricky, as the mathematics involved is rather hairy, because what we're basically going to do is measure relative to the circumference of Earth.

I found a formula in JavaScript on-line as a starting point:

```
var R    = 6371;         // kilometers
var dLat = (lat2-lat1);
var dLon = (lon2-lon1);
var a    = Math.sin(dLat/2) * Math.sin(dLat/2) +
           Math.cos(lat1.toRad()) * Math.cos(lat2.toRad()) *
           Math.sin(dLon/2) * Math.sin(dLon/2);
var c    = 2 * Math.atan2(Math.sqrt(a), Math.sqrt(1-a));
var d    = R * c;
```

In this case, the circumference is R, and it's 6,371km. Because Earth is an oblate spheroid, not a perfect sphere, I expect this will have some small level of error, but let's proceed and see where we get.

To accomplish any sophisticated mathematics in a Linux shell, we're pretty much stuck with bc, but it's plenty powerful enough for this task, even if it's a bit clunky.

As an example, here's how you'd set the value of pi within a bc script:

```
pi=$(echo "scale=10; 4*a(1)" | bc -l)
```

The first stumble we have is that bc wants to work with radians, not degrees, but the lat/lon values we're getting are in degrees, so we need to convert them.

But before we do that, here's the intermediate output we seek, as we now need to work with two addresses, not just one:

```
$ sh farapart.sh \
  "1600 pennsylvania ave, washington dc" \
  "1313 s. disneyland drive, anaheim, ca"

Lat/long for 1600 pennsylvania ave, washington dc

= 38.89859, -77.035971

Lat/long for 1313 s. disneyland drive, anaheim, ca

= 33.814413, -117.924424
```

Next month, we'll crack open the script to see how I am working with two addresses at the same time and splitting it into the four variables we'll later need. Then, we'll look at how to use bc to do the math.■

Dave Taylor has been involved with UNIX since he first logged in to the on-line network in 1980. That means that, yes, he's coming up to the 30-year mark now. You can find him just about everywhere on-line, but start here: www.DaveTaylorOnline.com. In addition to all his other projects, Dave is now a film critic. You can read his reviews at www.DaveOnFilm.com.

Rob Purdie
Project Manager
economist.com
amnesty.org

Lullabot-Powered

The most super powered sites in the world are created in Drupal, by you and Lullabot.

Lullabot™ New Lullabot Learning Series training DVDs at Lullabot.com

MICK BAUER

DEFCON: One Penguin's Annual Odyssey

Thousands of hackers in the same Las Vegas hotel? Sounds like a party to Mick!

Last month, I wrote a case study on Linux desktop system hardening, in the form of a step-by-step walk-through of how I prepared my Ubuntu laptop for DEFCON 17, the annual hacker's convention in Las Vegas that features one of the world's most hostile public wireless LANs. Well, you'll be happy and perhaps surprised to learn that my laptop came through unscathed.

But, you may wonder, was Mick exposed to cutting-edge developments in information security? Did he get invited to any elite skybox parties? And, doesn't this sort of reporting normally belong on a blog instead of languishing for a few months through the lengthy print process to which magazines are subject?

DEFCON has represented, for nearly two decades, an attempt to build some sort of understanding between the hacker community (in the broadest sense), law enforcement and the IT professions (certainly IT security).

I'll answer the last question first. In the past, I've covered DEFCON on LinuxJournal.com under my hacker pseudonym Darth Elmo. (No, I'm no more scary as a hacker than my handle implies, although I'm working on it.) But this time, I thought it might be interesting to cover DEFCON, which really is one of the most important annual events in my field, in a little more depth. I wanted not merely to *report* on DEFCON, but also to touch just a bit on some ongoing paradoxes and conflicts in information security that always seem to leap out at me at DEFCON.

In short, I wanted to write a DEFCON article that people still would find relevant and interesting a few months after the actual event. You be the judge!

Background

DEFCON, in case you aren't familiar with it, is an annual conference for the "security underground" held by Jeff Moss, aka The Dark Tangent (aided by scores of volunteers) in Las Vegas, Nevada, every summer since 1993 in late July or early August. It's run for and by self-identified "hackers", which is to say, technology's more creatively minded researchers, problem solvers and boundary pushers.

The term hacker, of course, has a lot of baggage. In mainstream English usage, it typically means "computer criminal". However, in the original meaning of the term, hackers are simply people who explore the limits of what is possible in computer systems, networks and other complex systems. Hackers are technologists who are driven to understand the full truth of what a given network, software application, device or operating system is really capable of doing (or being made to do), regardless of what its manuals, specifications or even its creators say.

Penetration testing, the art of breaking into systems or networks in order to document and demonstrate their various vulnerabilities, is one of the most visible and interesting applications of that kind of exploration, although it represents only a subset of what hacking is about. But penetration testing, and the skills involved in its practice, is somewhat problematic. Some hackers can and do cave in to the temptation to use their skills illegally or unethically, and even those who don't tend to be treated with suspicion by more conventionally minded IT professionals (not to mention law-enforcement representatives).

DEFCON has represented, for nearly two decades, an attempt to build some sort of understanding between the hacker community (in the broadest sense), law enforcement and the IT professions (certainly IT security). It isn't the oldest hacker conference, but according to longtime DEFCON insider Dead Addict, it probably was the first hacker convention to invite law-enforcement representatives and journalists to attend deliberately, and to encourage them to give presentations too.

In this column, I discuss my own perspective on DEFCON. DEFCON has changed a lot even just in the eight years I've been going (and even more over

the past 16), but in my opinion, it remains the single-most important event in my profession, imperfect though it unquestionably is.

Presentation Highlights

To start off, a bit of reporting is in order. At DEFCON, you really can't discuss culture separately from technology, since the whole point of the exercise is to celebrate their convergence. Furthermore, as always, I saw some very cool and interesting things.

In "Is Your iPhone Pwned?", Kevin Mahaffey, John Hering and Anthony Lineberry (whom I interviewed in the August 2009 issue) described a WAP push attack that, although easily detected and traced by carriers, can be used to open arbitrary links and windows on mobile browsers. They gave an excellent overview of mobile device security, highlighting difficulties caused by incompatibilities between different providers' implementations of mobile platforms and devices.

Moxie Marlinspike, in his talk "More Tricks for Defeating SSL", described a new "null prefix" attack that can be used to create fraudulent certificate signing requests (CSRs) that could result in attackers obtaining legitimately signed certificates for domains they don't own. Moxie's talk created a lot of buzz, and at least two other presentations referred to his work, including Dan Kaminsky's and Sam Bowne's.

Moxie is also author of the SSLstrip tool, which is sort of an HTTPS-to-HTTP proxy that can be used to capture SSL-encrypted data via man-in-the-middle attacks. He had presented on SSLstrip just a few days earlier at Black Hat Briefings 2009, a large commercial security conference that always precedes DEFCON. Sam Bowne gave a chilling but engaging demonstration of SSLstrip in his presentation "Hijacking Web 2.0 Sites with SSLstrip", also demonstrating Rsnake's "Slowloris" tool for denial-of-service-attacking Apache Web servers.

While we're on the topic of SSL attacks, Mike Zusman gave a talk called "Criminal Charges Are Not Pursued: Hacking PKI", in which he demonstrated a way to use ordinary Domain Validation (DV) SSL certificates in man-in-the-middle (MitM) attacks against sites that use Extended Validation (EV) certificates. It was easy to see how Zusman's attack could be combined with SSLstrip and the null prefix attack.

As you can see, man-in-the-middle attacks against SSL were a very hot topic at DEFCON 17. At this point you may be wondering, "oh great screaming goats, can I ever use eBay safely again?" The good news is, yes, probably.

MitM attacks work only when attackers can insert themselves logically upstream of the victim and downstream of the Web site the victim is trying to reach. In some contexts, this is relatively easy—on a public Ethernet, like at a hotel or on some kinds of Wi-Fi hotspots (never mind exactly how for right now, although I may write a future column on ARP spoofing). But the chances of someone doing this on your home DSL network or at your workplace are probably fairly slim.

Still, I hope this cluster of presentation topics serves as a wake-up call to Web developers who mix clear text (HTTP) and encrypted (HTTPS) content, which makes this sort of attack much harder for end users to detect, and to Certificate Authorities who need to figure out better ways of screening certificate signing requests.

It may, of course, simply be that somebody needs to figure out a better way of securing Web traffic than SSL (or TLS) as we know it. Even without attempting MitM attacks, phishers frequently are successful in luring users who don't even notice that their fake e-commerce and on-line banking look-alike sites lack any SSL at all. SSL and TLS represent an important enabling technology for making the WWW useful for shopping, banking and other sensitive transactions. We wouldn't be using the Web for those things today had it not been for SSL/TLS. But, it isn't at all certain whether SSL can evolve to address emerging threats satisfactorily.

As is so frequently the case with DEFCON, some of the best talks I attended weren't explicitly technical. In "The Year in Computer Crime Cases", Jennifer Stisa Granick of the Electronic Frontier Foundation used two recent court cases to illustrate a rash of recent attempts to widen inappropriately the definition of "unauthorized access" in the US Computer Fraud and Abuse Act. Jason Scott, in his talk "That Awesome Time I Was Sued For Two Billion Dollars", gave a breath-takingly profane and funny account of a spurious lawsuit filed against him over an electronic book archived on his site **www.textfiles.com**.

And, in a conference characterized by very large venues filled to capacity, Adam Savage of the TV show *MythBusters* really packed the house, giving an entertaining and inspiring account of the role of failure in his career. Savage, an expert in special effects and industrial design, may not be as obvious a candidate for speaking at a hacker conference as Ms Granick, a longtime legal advocate in criminal cases involving hackers, or Mr Scott, a noted hacker historian and archivist. But with his highly creative approach to problem solving and his eloquence and empathy in describing the challenges faced by everyone who works with complicated systems, Savage connected convincingly and resoundingly to the DEFCON crowd and received a very warm welcome (and a standing ovation).

I also saw good presentations on security challenges in cloud computing, techniques and patterns of stock-scam spammers, quirks of the credit reporting system and on Metasploit's new WMAP module for attacking Web applications. And, I was very pleased to attend a talk by my old friend and former employer Richard Thieme, hackerdom's most prominent cultural attaché.

Some of the presentations I attended weren't very good—sad to say, I even walked out on a couple. DEFCON always has been somewhat hit and miss with regard to consistency of presentation quality. But the good ones were *very* good, and they easily outnumbered the less-good ones. In all my years attending DEFCON, I've *never* felt it was a wasted trip. Besides, prematurely exiting one or two presentations is usually the only way I can find time to check out the DEFCON vendor area, which provides one-stop shopping for all your hacker-fashion, lockpicking and wireless hardware needs.

A Couple Dissonances

Maybe because DEFCON invites such high expectations, a few things bothered me. Some are peculiar to DEFCON; others probably are characteristic of hacker culture as a whole. Either way, these observations are offered in a wholly constructive spirit. Nothing worthwhile is worth being complacent about.

The thing that bothered me most consistently about DEFCON this year was the behavior and attitude of many (emphatically not all) of the "red shirt goons". In case you're unfamiliar with them, all members of DEFCON's volunteer staff are called goons, whether they're serving as actual physical-security goons like the red shirts, manning the information desk or running the massive DEFCON LAN infrastructure. All goons have T-shirts proclaiming their DEFCON goon status, but only the physical security crew's shirts are red.

I'm privileged to call many of these goons friends. In fact, it was the "original goon", Conal Garrity, who first urged me to give DEFCON a try many years ago. I've seen my goon friends work incredibly long hours with little sleep, irregular meals and little else in the way of extrinsic rewards for their efforts. They're an amazing group of people.

So maybe I was disproportionately bothered by seeing a small number of the red shirts being disrespectful to the point of being counterproductive, in their efforts to manage the large crowds that attended DEFCON 17. At various times I saw some of these guys yelling at attendees, calling them names, insulting their intelligence and making vague threats (though their preferred punishment seemed to be "more yelling").

One prominent goon even interrupted a presentation I was enjoying to harangue the crowd because there had been an incident concerning one person trying to bungee jump off the hotel's roof and another involving someone with a concealed handgun on the casino floor. The only problem was I'm pretty sure none of the hundreds of people who had up until this point been respectfully listening to Sam Bowne's talk had even *heard* of these incidents, let alone contributed to them in any way. I understand the goon was frustrated and stressed, but he took it out on the wrong people.

The crowds I saw at DEFCON this year were certainly large, but not unruly nor even particularly uncooperative. Certain goon antics seemed disproportionate. When I described some of them to a nonhacker friend later, his reaction was "sounds like Barney Fife syndrome". I had to reluctantly agree that yes, it did seem as though authority had gotten to some of these guys' heads just a tiny bit.

Another thing that occasionally struck me was the paradox of DEFCON elitism. On the one hand, in many ways DEFCON represents one of the most inclusive, accepting and open atmospheres I experience in any context. Everybody is welcome: hackers, cops, feds, nerds, script kiddies, lawyers, teachers, students, reporters—even vendors. Boundaries of race, nationality, socioeconomics, creed or sartorial style generally do not apply at DEFCON.

And yet, there's definitely an in-crowd. DEFCON parties abound, which are, as with parties the world over, frequently about who is *not* invited as much as who *is*. This shows up in all sorts of contexts, including the speaking schedule itself, but it's subtle, and over the years I've had trouble putting my finger on the real shape, extent and nature of DEFCON elitism. To talk of elitism at such an essentially inclusive event as DEFCON really is a bit of a paradox.

Obviously nepotism figures into practically any human endeavor, so maybe it's no big mystery. But I've observed that many if not most of those who seem to be in the DEFCON in-crowd are more oriented toward attacking things than defending them. I suppose this isn't very surprising, given the way DEFCON markets itself—one of the official DEFCON T-shirts this year featured the slogan "hack everything!"

Why wouldn't a hacker conference concern itself primarily with new attack techniques? After all, as I've just described, much of the content that made the biggest impression on me this year involved attacks. Exposure to new attacks and vulnerabilities provides valuable insights to those of us who defend networks and systems for a living.

So, I don't mean to suggest DEFCON should set

some sort of quota on attack-oriented material. However, I do think it's a shame that there's *less* of a focus on defense at DEFCON nowadays than there used to be. For example, both times I presented at DEFCON (in 2002 and 2003), my talk was included in the "Defense" track—a track that was phased out years ago. Maybe it's time to bring it back. Maybe more people need to submit DEFCON proposals involving compelling, cutting-edge defensive techniques.

And maybe, if we hackers want the world to give us more credit for the constructive things we do, and if we want people ever to accept the broader definition of hacker as creative problem solver, we need to do a little more to avoid giving the impression that we're almost exclusively creative problem *makers*.

So perhaps I'm less worried about nepotism per se—which in one form or another is inevitable in *anything* that relies so heavily on volunteers—than I am about its particular effects and ramifications. DEFCON simply needs more defense-oriented people it its in crowd. And I'm prepared to serve in that capacity myself, even if that means having to present at DEFCON year after year in multiple tracks, schmooze at all hours with prominent feds and attractive celebrity lawyers and accept one free beer after another at crowded, hot parties. You know where to find me, guys!

Conclusion

In all seriousness, DEFCON already is remarkably good, even incomparable. I can't over emphasize that for my friends and I who attended it, volunteered at it and presented at it, DEFCON 17 was a tremendous success—educational, thought-provoking, relevant, unpredictable, exhilarating at least as often as it was frustrating and, above all, fun.

In the words of Richard Thieme, who at the time wasn't sure whether he was quoting Simple Nomad or Bruce Potter, "For the system to work, it must never grow up and it must make us smile." Here's to the scene's never growing up. I hope to see you at DEFCON 18!■

Mick Bauer (darth.elmo@wiremonkeys.org) is Network Security Architect for one of the US's largest banks. He is the author of the O'Reilly book *Linux Server Security*, 2nd edition (formerly called *Building Secure Servers With Linux*), an occasional presenter at information security conferences and composer of the "Network Engineering Polka".

Resources

The DEFCON Web Site (including links to presentation materials for DEFCON 17 and also for DEFCONs past): **www.defcon.org**

Moxie Marlinspike's Web Site (where you can get SSLstrip and Moxie's paper on Null Prefix Attacks): **www.thoughtcrime.org/software.html**

Jason Scott's Archive of Hacker Lore Dating from the Era of BBSes: **www.textfiles.com**

ASA COMPUTERS

Want your business to be more productive?
The ASA Servers powered by the Intel Xeon Processor provide the quality and dependability to keep up with your growing business.

Hardware Systems for the Open Source Community - Since 1989.
(Linux, FreeBSD, NetBSD, OpenBSD, Solaris, MS, etc.

1U Server - ASA1401i
- 1TB Storage Installed. Max – 3TB.
- Intel Dual core 5030 CPU (Qty-1), Max-2 CPUs
- 1GB 667MGZ FBDIMMs Installed.
- Supports 16GB FBDIMM.
- 4X250GB htswap SATA-II Drives Installed.
- 4 port SATA-II RAID controller.
- 2X10/100/1000 LAN onboard.

2U Server - ASA2121i
- 4TB Storage Installed. Max – 12TB.
- Intel Dual core 5050 CPU.
- 1GB 667MGZ FBDIMMs Installed.
- Supports 16GB FBDIMM.
- 16 port SATA-II RAID controller.
- 16X250GB htswap SATA-II Drives Installed.
- 2X10/100/1000 LAN onboard.
- 800w Red PS.

3U Server - ASA3161i

- 4TB Storage Installed. Max – 12TB.
- Intel Dual core 5050 CPU.
- 1GB 667MGZ FBDIMMs Installed.
- Supports 16GB FBDIMM.
- 16 port SATA-II RAID controller.
- 16X250GB htswap SATA-II Drives Installed.
- 2X10/100/1000 LAN onboard.
- 800w Red PS.

5U Server - ASA5241i
- 6TB Storage Installed. Max – 18TB.
- Intel Dual core 5050 CPU,
- 4GB 667MGZ FBDIMMs Installed.
- Supports 16GB FBDIMM.
- 24X250GB htswap SATA-II Drives Installed.
- 24 port SATA-II RAID. CARD/BBU.
- 2X10/100/1000 LAN onboard.
- 930w Red PS.

8U Server - ASA8421i

- 10TB Storage Installed. Max – 30TB.
- Intel Dual core 5050 CPU.
- Quantity 42 Installed.
- 1GB 667MGZ FBDIMMs.
- Supports 32GB FBDIMM.
- 40X250GB htswap SATA-II Drives Installed.
- 2X12 Port SATA-II Multilane RAID controller.
- 1X16 Port SATA-II Multilane RAID controller.
- 2X10/100/1000 LAN onboard.
- 1300 W Red Ps.

All systems installed and tested with user's choice of Linux distribution (free). ASA Collocation—$75 per month

2354 Calle Del Mundo,
Santa Clara, CA 95054
www.asacomputers.com
Email: sales@asacomputers.com
P: 1-800-REAL-PCS | FAX: 408-654-2910

Intel®, Intel® Xeon™, Intel Inside®, Intel® Itanium® and the Intel Inside® logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Prices and availability subject to change without notice. Not responsible for typographic errors.

intel Xeon® inside™
Powerful. Efficient.

Dr hjkl and Mr Hack

Arrow keys, schmarrow keys—some of the best programs out there let you move around from the home row just like vim intended it.

KYLE RANKIN

Without diving headfirst into an ancient Linux holy war, let me set the record straight. I am a card-carrying, home-row hugging, Esc-key hammering, vim user. If you love Emacs, JOE, ed, Kate, gedit or the magnifying-glass-and-magnet approach to text-file editing, that's fine, and I'm not here to judge. It's just that for me, once I got over the initial vi(m) learning curve, I started looking for other tools that take the same approach to key bindings. Specifically, I am talking about the h, j, k and l keys and how you can use them to move left, down, up and right, respectively, in a document. What I found was that most of my favorite tools either already had vi-style key bindings or there was a simple way to enable them. Some programs even offered advanced bindings that closely mimic vim in a number of ways. In this column, I highlight some programs that either have vi key bindings or can be made to have them with a few simple steps.

I am a card-carrying, home-row hugging, Esc-key hammering, vim user.

Before I start talking about specific programs, I probably should explain why navigation with hjkl is better than with the arrow keys. It's a dirty secret among vim users that many people just use the arrow keys and backspace to edit their documents. The main reason hjkl navigation is great is that all of those keys are on the home row. In case you never took formal typing, the home row is the asdfghkl row of keys on a qwerty keyboard. If you learn to touch type, you are taught to rest your fingers on this row by default. This means the hjkl keys are within easy reach, but every time you reach for the arrow keys you have to move your right hand off the home row. Now, if you aren't a touch typist, that isn't a big deal. But if you are, it is almost as disruptive as reaching for the mouse. Granted, I know it is awkward at first, but if you are a vim user and touch type at all, it's worth it to force yourself to use hjkl for navigation until it becomes second nature.

In case you are new to vi key bindings, here are some of the main keys that you'll find work similarly in other programs:

- h — move left
- j — move down
- k — move up
- l — move right
- ^ — move to the beginning of a line
- $ — move to the end of a line
- G — move to the bottom of the document
- g — move to the top of the document (gg in vim)
- w — move the cursor ahead one word
- b — move the cursor behind one word
- / — enter search mode
- n — go to the next search result
- N — go to the previous search result

Paging Programs

A number of standard command-line programs use vi-style navigation out of the box, and the first I want to mention is less. The less program allows you to page through a text file, and if you needed yet another reason to use less instead of more, use it because j and k will move down and up a document. In addition, you can type G to scroll to the very bottom of a document and g (gg in vim) to move to the very top. As with vim, you also can press / to type a search term, and press Enter, then press n and N to find the next and previous matches, respectively. Like with less, by default, you can scroll through man page output with the same keys.

Screen also can use vi key bindings to navigate through its copy mode. Screen is an amazing shell program that allows you to open multiple shell sessions and detach and re-attach to them. If you've started using screen after being used to a regular terminal session, you likely ran into the strange behavior screen exhibits when you press Shift and PgUp and PgDn (or use the scroll bar) to scroll up and down

through the output. In screen, if you want to view output that has scrolled past the top of the terminal, simply press Ctrl-A Esc to enter copy mode. Within copy mode, now you can use the arrow keys (shame on you) or hjkl to scroll around the output. As with less, you also can use g and G to scroll to the top and bottom of the output. When you are done scrolling, simply press the q key to exit copy mode.

Even bash itself can be set so that you can navigate the command line in true vi style. In your bash shell, just type `set -o vi`. Now, keep in mind that once you enable this option, you will have to enter insert mode (press the i key) to insert text just like in vi. And, if you want to use h or l to move the cursor left or right, or w or b to move forward or back a word, you will have to press Esc to leave insert mode. For those of you who tried this and want to undo it, simply press i to enter insert mode, and then type `set -o emacs`.

Firefox

Unfortunately, Firefox doesn't use vi key bindings by default (although Google Reader does), but it's not surprising that this can be fixed with a Firefox plugin. The Vimperator plugin (**vimperator.org/trac/wiki/Vimperator**) is extensive enough to deserve a column of its own (in fact, send me an e-mail at lj@greenfly.net, if you'd be interested in that). Essentially, once the plugin is installed, your entire Firefox session turns into a modal vi-style session. Not only can you use hjkl, g, G and so forth to navigate pages, but also when you are in a text field, Vimperator actually moves into insert mode! You even can record and play back macros just like in vim. Vimperator adds a bunch of other features to make keyboard-only Web browsing not only possible, but also preferable to the mouse. If you are a vim lover and haven't installed Vimperator yet, I highly recommend it.

Essentially, once the plugin is installed, your entire Firefox session turns into a modal vi-style session.

Mutt

Here's yet another opportunity for me to add one more reason I love mutt as an e-mail program—it's practically vim's key-binding cousin. In fact, when you first start using mutt, you'll notice that when in doubt, you often can just press the same keys you'd use in vi to do something similar in mutt. The only place you might become confused initially is once you open an e-mail message and read it. By default, the j and k keys switch to the next and previous e-mail message in your folder, even when an e-mail is open, so you do have to teach yourself to use Enter and backspace to scroll through the body of an e-mail message.

Netris

Netris is a great command-line *Tetris* clone available on most major Linux distributions. One thing that always bugs me about *Netris* is that although it uses much of the home row to rotate and move shapes in the game, the keys are just slightly off from what you'd expect them to be in vi. Luckily, you can change the key bindings when you start *Netris*, so for true vi keys execute:

```
netris -k "hkl j"
```

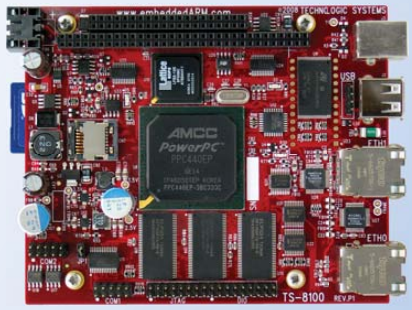
Doing the above causes h to move pieces left, l to move them right, k to rotate them, j to make a piece drop faster and the spacebar to drop a piece to the bottom immediately. My *Netris* score was much improved once I could play it like vi.

As you dig around both command-line and GUI programs, you'll find that a surprising number of them at least support hjkl, if not more-extensive vi key bindings. I've listed only some of my favorites here, but the next time you open a program, press j a few times—you just might be surprised when the program scrolls down.■

Kyle Rankin is a Senior Systems Administrator in the San Francisco Bay Area and the author of a number of books, including *Knoppix Hacks* and *Ubuntu Hacks* for O'Reilly Media. He is currently the president of the North Bay Linux Users' Group.

TS-8100
Ultra Reliable
w/ 128MB ECC RAM

Qty 1 $269 Qty 100 $229



shown w/ optional SD Card

400 MHz PowerPC
w/ Floating Point Unit

- POE ready
- Dual-execution unit
- Double-precision FPU
- Multifunctional PC/104 connector
- 12K LUT customizable FPGA
- 512MB NAND Flash
- 1 USB Host, 1 USB Device (12 Mb/s)
- Boots Linux 2.6 in < 2 seconds
- Fanless < 4W, sleep mode < 1mW
- Regulated 5-28V power input
- 5 10 bit ADC
- 4 COM ports
- SPI & DIO
- RS485/RS422
- 2 SDHC sockets
- 2 10/100 ethernet
- RTC & WatchDog
- 2 DMX Channels

TS-8150 available w/ extra data acquisition features

- Over 20 years in business
- Never discontinued a product
- Engineers on Tech Support
- Open Source Vision
- Custom configurations and designs w/ excellent pricing and turn-around time
- Most products ship next day

Technologic SYSTEMS

We use our stuff.
visit our TS-7800 powered website at
www.embeddedARM.com
(480) 837-5200

TRENDnet's TEW-654TR Wireless N Travel Router Kit

Your shoulder will thank you for traveling with TRENDnet's TEW-654TR Wireless N Travel Router Kit, a device that its maker calls the world's smallest 300Mbps wireless 802.11n router. This little guy measures a mere 6.4 x 8.2 x 1.9cm and comes with a carrying case, a thin 1-meter Ethernet cable, an Energy Star Certified external power adapter and an alternate USB cable to power the router from a computer. The router also features Access Point and Access Point Client modes and offers the latest in wireless encryption to protect valuable data. An advanced Multiple Input Multiple Output antenna technology delivers high-speed wireless connectivity and broad coverage that minimizes dead spots.

www.trendnet.com

Bluelounge's Refresh

Forget your old device charger in your hotel room with élan after picking up Bluelounge's new Refresh charging station, which can simultaneously re-juice up to four devices of nearly any kind. Refresh has six universal connectors in one compact location—namely two iPod/iPhone connectors, a Micro USB, a Mini USB and two USB sockets. Users can extend their device options by plugging in their own connectors, and short USB connector cables are available from Bluelounge.

www.bluelounge.com

Colfax International's HPC Cluster Computing Bundles

COLFAX DIRECT

In an effort to expand the accessibility of HPC, Colfax International announced availability of two new low-cost HPC cluster computing bundles, which include InfiniBand switches and adapters provided by Mellanox Technologies and Platform Computing's Platform Cluster Manager. The new bundles improve application performance and productivity in enterprise and data centers by adding 20Gb/s (Bundle 1) or 40Gb/s (Bundle 2) InfiniBand connectivity and simplify cluster operation through a fully integrated software stack. They further enable more companies to take advantage of the performance, low-latency and efficiency benefits of InfiniBand and the ease of use provided by Platform Cluster Manager, the latter of which "allows a user to build a cluster in hours versus weeks", says Colfax. A 10Gb/s bundle also is available.

www.colfax-intl.com

iX Systems' iX-Green Neutron Server Line

Go green and save green with iX Systems' new iX-Green Neutron, a server line that its maker says "is optimized for high-performance applications and provides the lowest power consumption on the market". The iX-Green Neutron models iX-GN1204, iX-GN1208 and iX-GN 2216 utilize power-saving DDR3 memory, 2.5" SAS and/or SATA drives and are equipped with high-efficiency (86%–93%) power supplies, all designed to reduce data-center costs without sacrificing performance. The series also leverages Intel's Xeon Processor 5500 series to boost performance, speed and energy efficiency over previous generation processors (12% at peak performance and 47% when idle), in part due to the way it interacts with power-saving DDR3 memory. The 5520 chipset introduces Intel QuickPath technology, which allows high-speed point-to-point links to navigate shared memory swiftly, distributed amongst the processors, greatly increasing efficiency and thereby cutting back on memory power utilization as well. The systems run FreeBSD.

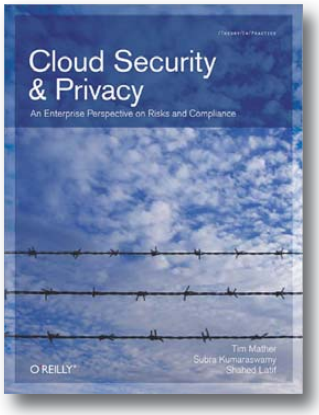
www.ixsystems.com

IPBrick for Oracle

Get Oracle running out of the box with IPBrick for Oracle, an appliance loaded and configured with Oracle Enterprise Linux, Oracle Database and Application Server. IPBrick asserts that its product offers greater simplicity than Microsoft Windows servers with automatic installation taking around 20 minutes, functional configuration via a Web interface that does not require Linux knowledge and simple recovery taking around 30 minutes. The company says firms can save money by not needing Linux experts to install and manage the system. The server also integrates with Microsoft Active Directory.

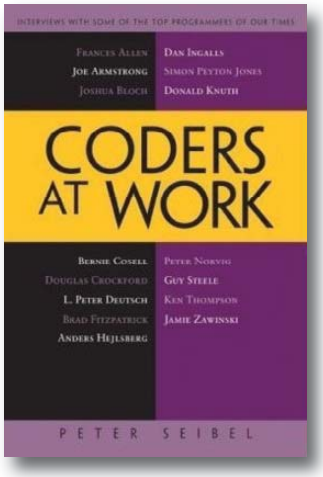
www.ipbrick.com

Tim Mather, Subra Kumaraswamy and Shahed Latif's *Cloud Security & Privacy* (O'Reilly)



If you are planning on putting cloud computing to work in your organization, you'll want to consider picking up the new O'Reilly book *Cloud Security & Privacy: An Enterprise Perspective on Risks and Compliance*. The title is penned by Tim Mather, Subra Kumaraswamy and Shahed Latif. Written for readers as diverse as business managers, IT personnel, service providers and investors, the book walks through the steps needed to ensure that Web applications are secure and data is safe, as well as addresses regulatory issues, such as audit and compliance.

www.oreilly.com



Peter Seibel's *Coders at Work* (Apress)

Learn how some of the world's most interesting computer programmers "tick" with Peter Seibel's new book *Coders at Work* from Apress. Editor Seibel whittled an original list of 284 names down to 15 that made it into the book. The interviews focus on how these programmers tackle the day-to-day work of programming while revealing how they became great programmers, how they recognize programming talent in others and what kinds of problems they find most interesting. Some of the interviewees include Frances Allen, the first female winner of the Turing Award and IBM fellow; L. Peter Deutsch, author of Ghostscript; Brendan Eich, inventor of JavaScript; Simon Peyton Jones, co-inventor of Haskell; Donald Knuth, creator of TeX; and Ken Thompson, inventor of UNIX. The book is for programmers interested in new approaches and points of view that can be gleaned from leaders in the field.

www.apress.com

Fixstars' Y-HPC

The new, updated v2.1 of Fixstars' Y-HPC for Sony PlayStation 3, dubbed by the company as the world's only commercial, cross-architecture cluster construction suite, is now available. This release's key improvement is the addition of the next generation of ps3vram for fast, temporary file storage or swap using PS3 video RAM. This version of ps3vram, says Fixstars, is up to 50% faster than prior versions and is automatically enabled as swap. Also included are the new features found in Yellow Dog Enterprise Linux v6.1, such as updated kernel v2.6.28, IBM Cell SDK v3.1.0.1, improved ps3vram support and Libfreevec. Fixstars says that the monumental improvements in compute performance from Y-HPC v2.1 will allow existing and new PlayStation 3 clusters to tackle problems never before believed to be practical.

us.fixstars.com

Please send information about releases of Linux-related products to newproducts@linuxjournal.com or New Products c/o *Linux Journal*, PO Box 980985, Houston, TX 77098. Submissions are edited for length and content.

Fresh from the Labs

PokerTH—Quality *Texas Hold'em*

www.pokerth.net

During this decade, *Texas Hold'em* rapidly has become one of the most popular variants of poker across the globe. Taking center stage in such films as *Casino Royale* and being the main feature in the World Series of Poker, *Texas Hold'em* is now the coolest game around. However, unless you want to use some kind of tacky on-line game, it's hard to find a sim that feels any good. Well, *PokerTH* steps up to the plate quite nicely indeed. According to its Web site: "*PokerTH* is a poker game written in C++/Qt4. You can play the popular *Texas Hold'em* poker variant against up to nine computer opponents or play network games with people all over the world. This poker engine is available for Linux, Windows and Mac OS."

Installation The *PokerTH* downloads page has a number of binary packages, and *PokerTH* also is included in a number of repositories. A distro-neutral binary package also is available as a tarball, along with a binary installer and source.

For those going with source, grab the latest tarball, extract it and open a terminal in the new folder. Chances are, you won't have all the needed libraries, so install the following as recommended by the *PokerTH* Web site:

- Qt version >= 4.4.3, 4.5.1 recommended.
- zlib version 1.2.3.
- libcurl version >= 7.16.
- gnutls (version 2.2.2).
- libboost_thread, libboost_filesystem, libboost_datetime, libboost_program_options, libboost_iostreams, libboost_asio and libboost_regex (version >= 1.36, 1.38.0 recommended).
- libSDL_mixer and libSDL.

I also had to install libqt4-dev. Once you have all the needed dependencies, enter the following commands:

```
$ qmake-qt4 pokerth.pro
$ make
$ sudo make install
```

When the installation is over, run *PokerTH* with this command:

```
$ pokerth
```

If you're lucky, it will be in your system's menu, and there also may be a new desktop icon.

Note the helpful guide of possible hands for new players on the left and a chance meter on the right.

A helpful feature for new players is that it actually displays what hands you should play for on the left, along with the terminology.

Usage The starting screen will have the options to start a local game, an Internet game, create a network game or join a network game. Obviously, you'll want to learn the game's interface before playing anyone else, so choose Start Local Game. A screen appears with a default number of players, starting cash, blind settings and game speed. Unless you really know what you're doing, stick with the given defaults.

I'm assuming you know the basics of *Texas Hold'em* here, but even if you don't, the game goes to some lengths to make the learning process fairly intuitive. As soon as you're in the actual game screen, you'll be right in the action with two cards dealt. Here you can bet for more cash on the table, check/call or fold. As soon as you make your choice, things move on to the next player and the round continues.

As the game moves on, more bets can be placed before everyone has called/checked, more cards are revealed, and the round finishes with someone winning the pot. Above the Raise, Call and Fold buttons is also the option to go all-in. The field with the numbers next to the All-In button with the slider below lets you adjust how much you want to bet/raise, rather than being stuck with the game's defaults.

A helpful feature for new players is that it actually displays what hands you should play for on the left, along with the terminology. On the right are some tabs with brilliant features.

The first tab contains a log of all that's happened so far. The second tab has what actions to choose when you're away from the computer. And best of all, the last tab has a dynamically updated chance section, telling the mathematical chance you have of getting each kind of hand.

This last feature is particularly of use for new players, because it tells what chance you have of getting the hand you're after, so you don't need to be a mathematical savant. This is great for getting a feel for the game and avoiding stupid errors. Once you've played for a while, the dynamics and mathematics of *Texas Hold'em* should start to come more intuitively.

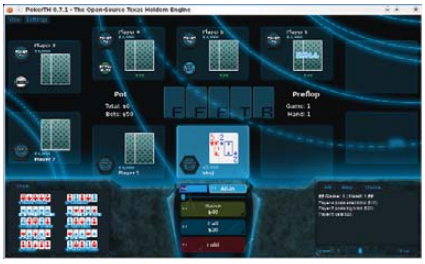
Playing the computer becomes tedious after a while (the computer is all math and no instinct), and you'll be wanting to play some humans soon. Close the game, go back to the *PokerTH* main screen, and choose Internet Game. You'll be taken to a screen with lots of

games from which to choose. Pick an open game that has a decent number of players, and when the host is ready to start, the game will proceed.

Playing on-line, you get a time limit and a chat window. Note the four color suits—great for avoiding mixups in quick situations like this.

The first major difference you'll notice between on-line play and local play is that a timer bar is applied in on-line mode, which takes care of tardy players. If you are going to be away from the computer for any amount of time, it's worth changing your settings in the Away tab on your right. An on-line chat tab also is available—great for a social game like poker. When you're ready to leave, press the Lobby button on the bottom-right corner. I'll let you work out the rest of the game from here.

It's definitely worth having a look at the game's Web site, where users have made a number of themes and additions. I find the default theme a little bland, but some other themes are quite snazzy. Some card themes have four different colors as well, which really helps you differentiate between suits quickly when it's midnight and you have a head full of whiskey!



Some of the themes, such as Stardust (shown here), look snazzy and make *PokerTH* a real class act.

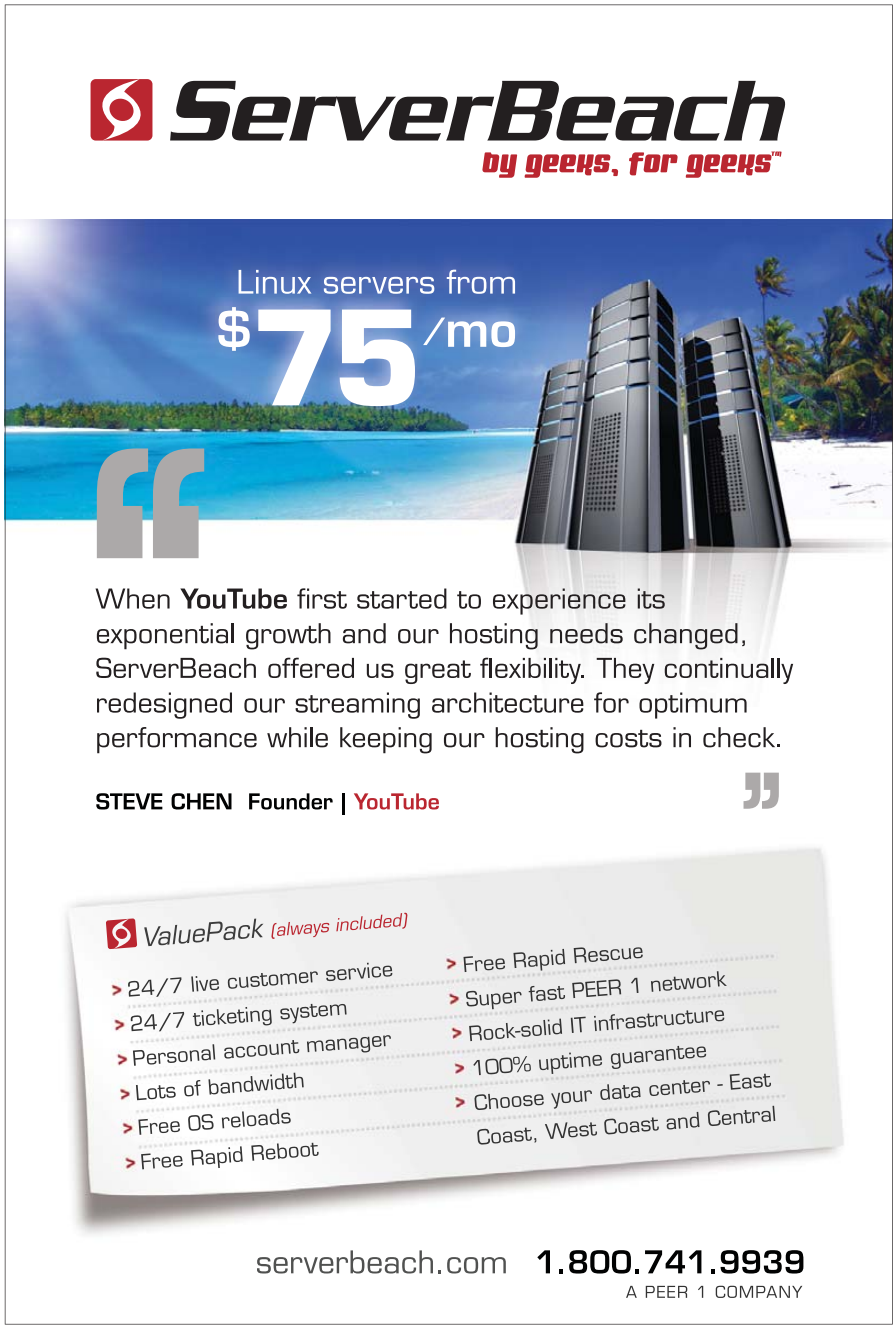
I'm sure *Texas Hold'em* fans will love *PokerTH*. Its use of open protocols, such as IRC, should help its longevity, and its large fan base is testament to this (I've never once had trouble finding a game on-line). It's a great poker sim for newbies and veterans alike, and I highly recommend it.

X-Moto—Motorbike 2-D Platformer

xmoto.sourceforge.net

Something that's been making a buzz at SourceForge, *X-Moto* appears to be a remake of an old DOS game I used to play, *Action Supercross*, along with its earlier Windows remake, *Elastomania*. Some tweaks have been made to the the physics engine, new community levels have been added, and it's now about ten times harder! This game is a physics engine gone mad, allowing for some truly hilarious moves and addictive gameplay. The actual premise of the game and its controls are simple, yet the dynamics and gameplay are complex. The result is that despite its simplistic look, this actually is one of the hardest games I've ever played.

Installation If you look on the Web



site, lots of installation methods are available, and installing *X-Moto* is easy. Packages are provided for various distros in deb, rpm and Slackware form, and *X-Moto* is provided in a number of repositories, so it's well worth having a look in your package manager first. And, for those who like to do things the hard(er) way, a source tarball also is provided.

In terms of requirements, you'll need: aclocal, SDL_mixer, liblua, libode0, build-essential, sqlite3, zlib, libjpeg, libpng, libbz2, glu, SDL_ttf, liblualib50 and libcurl. For the source, download the latest tarball, extract it, open a terminal in the new folder and enter:

```
$ ./configure
$ make
$ sudo make install
```

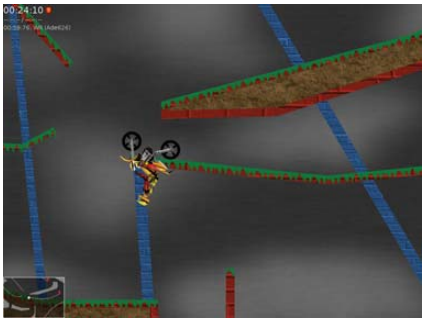
To run the game, look in your system's game menu, or enter:

```
$ xmoto
```

Usage When the game first starts, it prompts you to connect to the Net. I recommend saying yes, because there's a whole swag of things you can do when connected, such as rate levels, get new ones and so on. If you don't like being connected, you can turn it off with F8 anyway. Let's get playing. Click Levels, and in the menu below in the Level Packs tab, look under All Levels and choose a level. I recommend working your way through the "aeRo's Training" levels first and going onto harder levels from there.

Once you're in the game itself, the bike is controlled purely with the arrow keys and the spacebar—that's it. The up key controls the bike's throttle, and the down key controls the brakes. The right key rotates/pitches the bike clockwise, and the left key, anti-clockwise. The spacebar flips the direction of the bike between left and right, and this can be performed anywhere, at any time. When you need to restart the level, press Enter (you'll be doing this a lot). You'll probably notice that each level has a timer, and each level has a high score Internet-wide. If you're an *X-Moto* fanatic, I'm sure you'll wanna take someone down!

Although this game may look simple at first glance, beginners will have trouble just keeping the bike upright. Full throttle starts generally result in a wheelie, and if you don't "feel" the physics engine and its inertia, you'll quickly flip the bike over and land on your head. The main things to grasp are what the active objects are in this game, how everything is controlled and what this will allow you to do.



The wacky physics engine allows you to do some truly mind-boggling stunts!

First, there's the rider's head. Don't hit it on anything, or the level ends, and you have to restart. However, things also can go through the rider's body. Now this might sound strange, but it allows for some truly hilarious possibilities, such as hanging upside down on a rail with the wheels on top. The suspension reacts in real time and is a big part of the physics engine. The tires will react in time with the suspension also, so pay constant attention to your terrain. Remember that like most motorbikes, this is rear-wheel drive only, so if you try to feed on power when you're only on the front wheel, it won't do any good. Braking, however, works on both wheels.

Braking is very important in some often unexpected ways, because many puzzles require you to lock your wheels and actually flip the bike over. Don't forget that the body of the rider itself also will flip forward under braking inertia, not just the bike, so if you pull up too hard and too late before a wall, the rider might hit his head, even though the bike is still okay.

Learning to feel the actual game is really important—get to grips with the physics engine, especially on tricky hills where the bike can tip over and you lose all of your momentum—you want your reactions to come naturally. Learn to use opposing power, flipping the bike around and applying power in the reverse direction, as braking isn't always the answer. Don't glue down the accelerator. Some puzzles require quick dabs on the keyboard, and just about every level requires a lot of delicacy. Finally, remember that when the bike is riding upside down hanging on a ledge (yes, it's crazy, but it's what the game is all about), the wheel will be turning in the opposite direction to when the bike is upright. This is counter-intuitive at first, but you'll get used to it.

I've covered only basic single-player stuff here, but this game has many more features and some very clever adaptations, especially in the scripted levels, which really show off what this crazy engine can do. If you check the Web site, you'll notice it has a very extensive community and, most important, a level editor. Try making your own levels, and explore the dynamics of this game intimately.■

In-game scripting allows people to make their own modifications freely, sometimes with elaborate results as shown here.

Look on-line and you'll find some truly crazy levels!

John Knight is a 25-year-old, drumming- and climbing-obsessed maniac from the world's most isolated city—Perth, Western Australia. He can usually be found either buried in an Audacity screen or thrashing a kick-drum beyond recognition.

Brewing something fresh, innovative or mind-bending? Send e-mail to newprojects@linuxjournal.com.

USENIX Upcoming Conferences

23rd Large Installation System Administration Conference (LISA '09)

Sponsored by USENIX and SAGE in cooperation with LOPSA and SNIA

NOVEMBER 1–6, 2009, BALTIMORE, MD, USA
http://www.usenix.org/lisa09

Symposium on Computer-Human Interaction for Management of Information Technology (CHIMIT '09)

Sponsored by ACM in association with USENIX

NOVEMBER 7–8, 2009, BALTIMORE, MD, USA
http://www.chimit09.org/

ACM/IFIP/USENIX 10th International Middleware Conference

NOV. 30–DEC. 4, 2009, URBANA CHAMPAIGN, IL, USA
http://middleware2009.cs.uiuc.edu/

First Workshop on Sustainable Information Technology (SustainIT '10)

Co-located with FAST '10

FEBRUARY 22, 2010, SAN JOSE, CA , USA
http://www.usenix.org/sustainit10
Submissions due: November 9, 2009

8th USENIX Conference on File and Storage Technologies (FAST '10)

Sponsored by USENIX in cooperation with ACM SIGOPS

FEBRUARY 23–26, 2010, SAN JOSE, CA , USA
http://www.usenix.org/fast10

3rd USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET '10)

Co-located with NSDI '10

APRIL 27, 2010, SAN JOSE, CA, USA
http://www.usenix.org/leet10
Submissions due: February 25, 2010

2010 Internet Network Management Workshop (INM '10)

Co-located with NSDI '10

APRIL 27, 2010, SAN JOSE, CA, USA
http://www.usenix.org/inm10
Paper registration due: November 30, 2009

7th USENIX Symposium on Networked Systems Design and Implementation (NSDI '10)

Sponsored by USENIX in cooperation with ACM SIGCOMM and ACM SIGOPS

APRIL 28–30, 2010, SAN JOSE, CA, USA
http://www.usenix.org/nsdi10

2nd USENIX Workshop on Hot Topics in Parallelism (HotPar '10)

JUNE 14–15, 2010, BERKELEY, CA, USA
http://www.usenix.org/hotpar10
Submissions due: January 24, 2010

2010 USENIX Annual Technical Conference (USENIX '10)

JUNE 23–25, 2010, BOSTON, MA, USA
http://www.usenix.org/usenix10
Submissions due: January 11, 2010

USENIX Conference on Web Application Development (WebApps '10)

JUNE 23–25, 2010, BOSTON, MA, USA
http://www.usenix.org/webapps10
Submissions due: January 11, 2010

19th USENIX Security Symposium (USENIX Security '10)

AUGUST 11–13, 2010, WASHINGTON, D.C., USA

9th USENIX Symposium on Operating Systems Design and Implementation (OSDI '10)

OCTOBER 4–6, 2010, VANCOUVER, B.C., CANADA

SOFTWARE

Virtualization Shootout: VMware Server vs. VirutalBox vs. KVM

A comparison of three virtualization solutions: VMware Server, VirtualBox and KVM—each has its strengths and weaknesses. BILL CHILDERS

Virtualization is a buzzword that's been making its way around the corporate IT circles for a few years. On paper, virtualization sounds great—you can make full use of those unused CPU cycles, leverage a particular machine to its fullest potential, and save power and space at the same time. Many people think virtualization is good only in the corporate data center; however, several software packages run just fine on desktop- and laptop-class Linux machines, as well as servers. In this article, I put three of them through their paces: VMware Server, VirtualBox and KVM.

"But wait!" you may exclaim, "Why aren't you evaluating Xen too?" The answer is simple. Xen, although extremely powerful, is more of an enterprise-class virtualization solution and may be overkill for the average Linux user. If you're going to be building a data center or a service that will be exposed to customers on the Internet, that's when you should consider Xen. This is one of the reasons Ubuntu officially supports KVM, rather than Xen, as its open-source virtualization solution, and I follow that reasoning here.

First, I should define a couple terms for the purposes of this article. A host is a physical machine running one of the virtualization solutions. A guest, virtual machine or VM is the virtual machine running inside the virtualized container provided by the host.

Because this is a shootout, I assign point values to categories, and the product with the most points wins the shootout. The values range from 1 to 3, with 1 being poor, 2 being average and 3 being excellent. All of the virtualization packages are installed on an Ubuntu 9.04 host. The categories are as follows:

- Ease of installation.
- Administrative tools.
- Capabilities.
- License.

VMware Server

VMware has been providing virtualization solutions for ten years, and as such, is the virtual 800-pound gorilla in the marketplace. With at least six virtualization products that span both the desktop and server markets, VMware has a package that will fit your needs. The product I review here is VMware Server 2.0. It's free (as in beer) and is very feature-rich.

Ease of Installation VMware Server ships as a 507MB Windows executable, a 465MB RPM or a 466MB tarball. Because I'm installing on an Ubuntu machine, I use the tarball. Kicking off the installation is fairly straightforward on Ubuntu. Simply ensure that you've got the build-essential package installed, along with the headers for whatever kernel you're running. Then, untar the tarball and run `./vmware-install` as root, and follow along with the prompts. The installer will prompt you for the paths to where you want to install various things. It's acceptable to choose the defaults, as the installer chooses fairly sane locations.

One thing to note is that due to VMware's "free as in beer" license, you *must* get a serial number from the VMware site before you can run it. Make sure you have registered on the VMware site and have your serial number handy, as the installer will ask you for it near the end of the installation process.

Ease of installation score: 2. This is mostly due to VMware requiring some packages and asking many questions in the installer. It works well once you get it installed, and you can take the defaults on just about every question, but it is a little tedious.

Administrative Tools If you've used VMware Server 1.0 and haven't looked at 2.0 yet, you're in for a surprise. The 2.0 version of the product uses a Web-based administrative panel, compared with the "fat client" approach that the 1.0 product used (Figure 1).

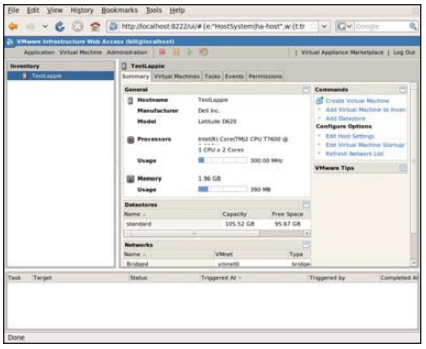


Figure 1. VMware Server Administrative Console

Everything in the admin console is easy to use. Creating a virtual machine is a simple matter, thanks to VMware's excellent form-based wizards. Simply fill in the blanks, and VMware will create an appropriate VM and get it ready for its first boot. VMware Server provides a

virtual console via its Web interface to the virtual machine as well (Figure 2). It requires installing a Firefox plugin, but the console works well and doesn't require a fat client. Unfortunately, the plugin doesn't work on the Mac version of Firefox.

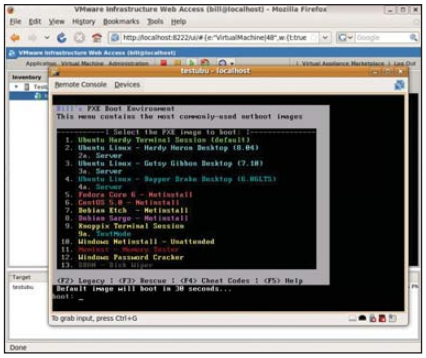


Figure 2. PXE Menu via the VMware Virtual Console

VMware also allows you to console to the machine remotely via VNC. This requires adding the lines `RemoteDisplay.vnc.enabled = "TRUE"` and `RemoteDisplay.vnc.port = 5900` to the virtual machine's configuration file (named <hostname>.vmx in the virtual machine storage directory).

In short, the VMware administrative console is excellent. The Web-based GUI is easy to navigate, and the tools work well on Linux or Windows. The ability to enable VNC access to a virtual machine's console without using the Web GUI could prove invaluable in certain administrative cases.

Administrative tools score: 3. VMware's experience in the field shows here, and VMware Server's historical connections to the GSX commercial product mean that the tools are best of breed.

Capabilities VMware Server is an extremely capable virtualization platform. Its ancestor is VMware's first-generation commercial server product, VMware GSX, so it has a great pedigree. VMware Server's key features include:

- The ability to run on standard x86 hardware, with or without hardware virtualization extensions.
- Two-processor Virtual SMP, allowing a single virtual machine to span two processors.
- A snapshot feature, allowing you to capture the state of a VM and then roll it back to that state.
- 64-bit support, on both the host and guest operating systems.
- Support for bridged, NAT and host-only network interfaces.
- Support for USB devices and controllers.

All these features mean that VMware Server is a great platform for personal experimentation or light business use. I've personally had a VMware Server host with a couple guest machines running continuously since 2007.

Capabilities score: 3. VMware has been building its feature set for years, and it shows here.

Licensing VMware Server has a proprietary license with appropriate EULA for this software. Although it's technically free, it's "free as in beer", meaning that though it costs nothing, you can't actually modify it. VMware does make some source code available, but it's not the entire source tree, only the parts that are GPL that VMware modifies. In order to use the software, you need to register on the VMware Web site and get a serial number in your name. Although this is available at no cost, it isn't "free" in the open-source sense.

Licensing score: 1. VMware's proprietary license and EULA mean you can't lift the hood and tweak it as you see fit, nor can you analyze the code for vulnerabilities. You're at the mercy of VMware. If Free Software is important to you, this license will give you fits.

VMware Server total score: 9.

VirtualBox

VirtualBox is a relative newcomer to the virtualization market, with its initial release in early 2007. VirtualBox originally was created by Innotek, but it has since been acquired by Sun Microsystems. Version 3.0 of the software was released recently and includes many new features.

Ease of Installation VirtualBox ships for Linux hosts as a native package for most distributions. There are packages for Ubuntu, Debian, OpenSUSE, Fedora, Mandriva, Red Hat, Turbolinux and PCLinuxOS 2007. Installing the software is as simple as downloading the package for your OS, then using your native package manager to install the package. On Ubuntu 9.04, the binary package is 43MB, and installation required the additional packages of libcurl3, libqt4-network, libqtcore4, libqtgui4 and python2.5, all of which are easily fetched via apt-get. Double-clicking on the package in Nautilus launches the Ubuntu Package Installer, which pulls in the dependencies automagically. In all, installation is straightforward, quick and easy. VirtualBox also maintains a repository for Debian-based distributions that you can add to your apt sources. Then you simply can `apt-get` the package (virtualbox-3.0) and its dependencies.

Ease of installation score: 3. The only way VirtualBox could be easier to install is if it were included in the Ubuntu apt sources out of the box.

Administrative Tools VirtualBox includes a native "fat client" for your host OS that allows you to manage your virtual machines (Figure 3). The client is easy to use, and it's wizard-based—much like the VMware admin console. Creating virtual machines is a snap, and VirtualBox gets kudos for making it as easy as VMware to spin up new virtual machines.

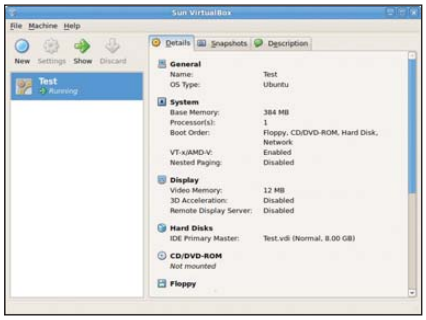


Figure 3. VirtualBox Admin Console

If you want to run your guest machines in headless mode, VirtualBox has that covered too. There is a VBoxHeadless management binary that will bypass the admin GUI and start an RDP server running for that particular VM. Once your VM is running in headless mode, you can point an RDP client to your physical host's port 3389 (by default, the port is also configurable), and you'll see the virtual machine's

console. This is very handy if you're not at the physical machine or can't tunnel X easily. Figure 4 shows a VM running with VirtualBox.

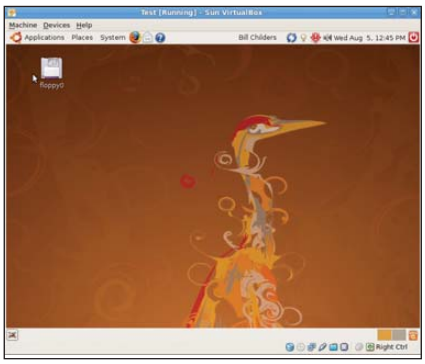


Figure 4. Booting an Ubuntu VM under VirtualBox

Administrative tools score: 3. VirtualBox includes excellent tools for creating and managing virtual machines. The fact that it's a native "fat client" rather than a Web GUI is slightly less convenient for multiplatform access, as compared to VMware, but every bit of functionality is there and easy to use.

Capabilities VirtualBox may be a young project, but it certainly doesn't lack features. It compares with VMware handily in many areas, such as the following:

- Support for bridged, NAT and host-only networking.
- Two-processor virtualized SMP.
- 64-bit support for both hosts and guests.
- Snapshot capability for easy capture and rollback.

Unlike VMware, VirtualBox is available in both a proprietary and open-source edition. The open-source edition is released under GPL, but it doesn't include the following features that are available only in the proprietary version:

- The headless RDP server is not available in the open-source edition.
- There is no virtualized USB support in the open-source edition.
- Because USB and RDP support aren't included, the proprietary version's USB-over-RDP feature isn't in the open-source edition.
- The virtualized serial ATA disk controller isn't in the open-source edition. Disks appear as either SCSI or IDE devices.

Capabilities score: 3. VirtualBox nearly matches VMware Server feature for feature.

Licensing As mentioned above, VirtualBox ships two different versions of its product: a proprietary version and an open-source edition. The proprietary version is licensed under the VirtualBox Personal Use and Evaluation License (PUEL), and although you are asked to register the software when it's first launched, it's not required. The open-source edition is covered under the GPL, and it's truly open source, though it does omit the four features I mentioned previously. If you do decide to run the open-source edition, be advised that it doesn't come as a binary package, only source code, so you will have to build it yourself. Building it yourself isn't terribly painful, as the folks at VirtualBox have supplied fairly good instructions.

Licensing score: 2. VirtualBox's PUEL license on the more feature-rich version isn't open source, but VirtualBox does make most of the source code available and provides instructions on how to build the code if you don't want to succumb to the evils of proprietary licensing.

VirtualBox total score: 11.

KVM

KVM is the Kernel-based Virtual Machine, and it is a virtualization technology that's fully open source and integrated into Linux. Ubuntu ships its distribution to be KVM-ready out of the box, and several other distros do as well. KVM isn't quite as simple as the other two products...yet, but it is very capable.

Ease of Installation KVM isn't as easy as VirtualBox or VMware to install. First, you must ensure that your hardware is compatible with KVM. Although VirtualBox and VMware will install on most machines with x86 processors, KVM requires that the processor support Intel-VT or AMD-VT extensions, and that those extensions are enabled in the BIOS. Once that's confirmed, you need to install some packages. Because my host machine is Ubuntu 9.04, I just run apt-get:

```
$ sudo apt-get install kvm                \
                       libvirt-bin        \
                       ubuntu-vm-builder  \
                       qemu               \
                       bridge-utils       \
                       virt-manager
```

Next, you need to add your user to the libvirtd group, and log out and back in for your group membership to take effect:

```
$ sudo adduser bill libvirtd
```

To confirm that your system is ready, run virsh, a shell interface to manage virtual machines. If you get a connection error, your system isn't ready to run KVM yet:

```
$ virsh -c qemu:///system list
Connecting to uri: qemu:///system
 Id Name                 State
----------------------------------
```

The default network configuration in KVM is NAT. If you want to use a bridged interface, you need to perform the additional step of manually setting up a br0 device on the host machine. (See Resources for a link to how to do this on an Ubuntu host.) You may need to do several more steps, depending on what you're trying to achieve.

Ease of installation score: 1. Compared to VMware and VirtualBox, KVM requires way too much work. Setting up bridged networking should be a drop-down in a dialog box and not require part of its own wiki page.

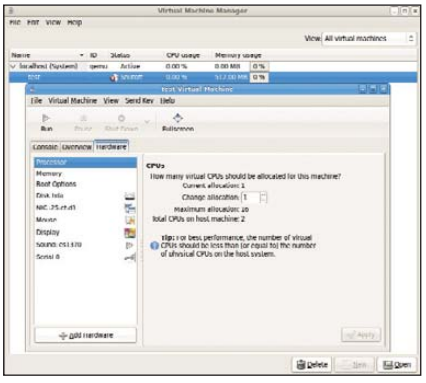


Figure 5. virt-manager in Action

Administrative Tools KVM's administration tool on Ubuntu is called virt-manager (Figure 5). In order for virt-manager to address things like bridged interfaces correctly, it should be run as root. virt-manager is fairly nice and easy to use, and it presents you with a wizard-based interface for virtual machine creation. Unfortunately, only the basics are supported for virtual machine creation and configuration. KVM also allows you to get a console on the virtual machine via the virt-manager tool, but it doesn't provide you with headless RDP or VNC abilities like the others. To enable some of the more-advanced features on your guest machines, you need to edit the XML definitions for those VMs.

Administrative tools score: 1. If it were possible to give a 1.75, I would. The tools are adequate for the task but still need a bit of work before I'd call them average. However, KVM is a rapidly developing target, so things most likely will improve with time.

Capabilities KVM's capabilities aren't yet on a level with the other two packages in this shootout. The framework for the functionality may be there, in some cases, but it may be hard to configure and use. KVM doesn't implement virtual USB ports or some of the other hardware that VMware and VirtualBox do. The lack of a headless capability also limits its usefulness in certain situations, such as a collocated environment.

Capabilities score: 2. KVM is adequate for most virtualization tasks, but it doesn't particularly shine at any of them due to the current limitations on what it can virtualize. The ability to have virtualized USB ports and headless connection options would beneficial.

Licensing KVM's shining point is its licensing model. It's completely open source—most parts are GPL or LGPL licenses. This means it's truly free (as in speech), and your favorite Linux distributions are free to package it and ship it as a ready-to-run feature.

Licensing score: 3. It's hard to beat open source.

KVM total score: 7.

Conclusion

And the winner is...VirtualBox! The combination of ease of installation, its excellent feature set, top-notch admin tools and flexible licensing nudged this contender ahead of the rest. Of course, any of these three tools probably will meet your virtualization needs, but if you're starting off fresh, give VirtualBox a try. You'll be pleasantly surprised, and who knows...you may just start virtualizing everything!■

Bill Childers is an IT Manager in Silicon Valley, where he lives with his wife and two children. He enjoys Linux far too much, and he probably should get more sun from time to time. In his spare time, he does work with the Gilroy Garlic Festival, but he does not smell like garlic.

Resources

VMware Server Home Page: **www.vmware.com/products/server**

VMware Server Source (Modified): **www.vmware.com/download/server/open_source.html**

VirtualBox Home Page: **www.virtualbox.org**

VirtualBox Source Code: **www.virtualbox.org/wiki/Downloads**

VirtualBox Editions: **www.virtualbox.org/wiki/Editions**

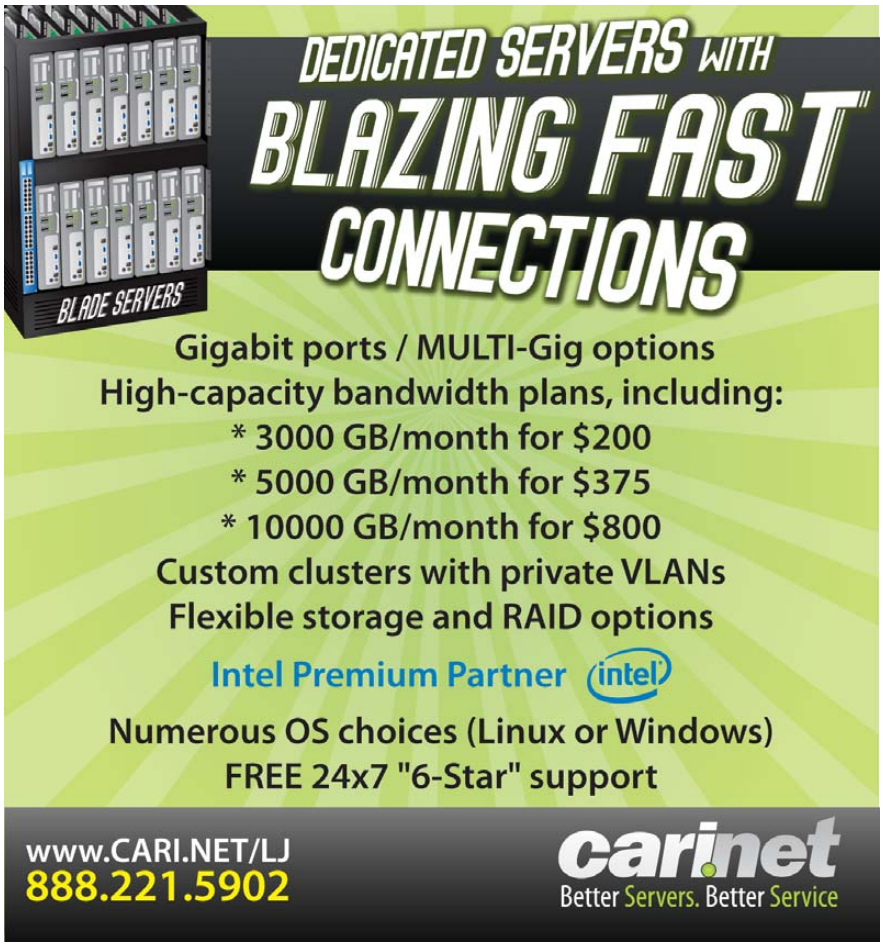
VirtualBox PUEL License: **www.virtualbox.org/wiki/VirtualBox_PUEL**

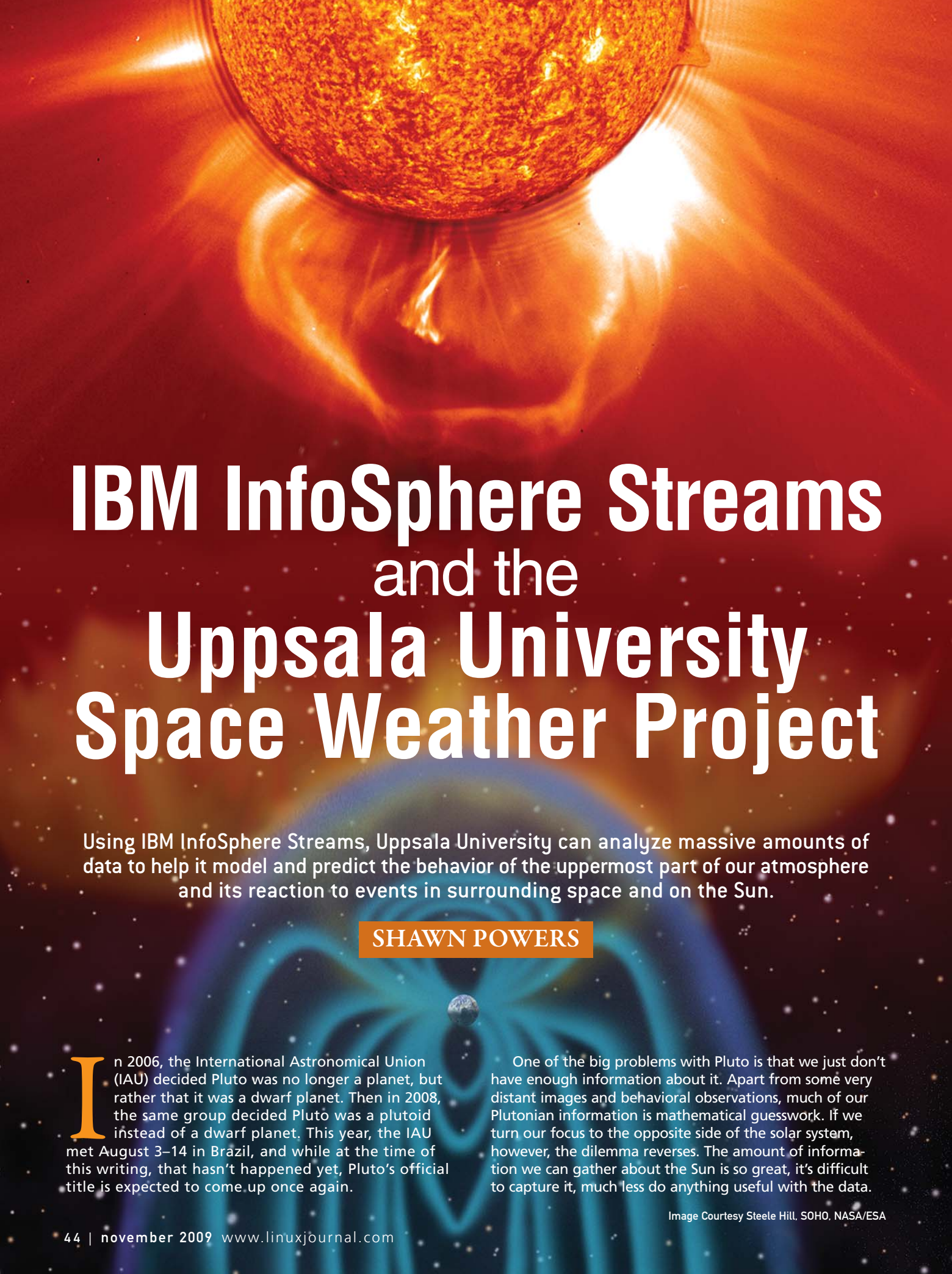
KVM Home Page: **www.linux-kvm.org**

Running KVM on Ubuntu: **https://help.ubuntu.com/community/KVM**

Network Config for KVM on Ubuntu: **https://help.ubuntu.com/community/KVM/Networking**

Comparison Matrix of Virtual Machines: **en.wikipedia.org/wiki/Comparison_of_platform_virtual_machines**





IBM InfoSphere Streams and the Uppsala University Space Weather Project

Using IBM InfoSphere Streams, Uppsala University can analyze massive amounts of data to help it model and predict the behavior of the uppermost part of our atmosphere and its reaction to events in surrounding space and on the Sun.

SHAWN POWERS

In 2006, the International Astronomical Union (IAU) decided Pluto was no longer a planet, but rather that it was a dwarf planet. Then in 2008, the same group decided Pluto was a plutoid instead of a dwarf planet. This year, the IAU met August 3–14 in Brazil, and while at the time of this writing, that hasn't happened yet, Pluto's official title is expected to come up once again.

One of the big problems with Pluto is that we just don't have enough information about it. Apart from some very distant images and behavioral observations, much of our Plutonian information is mathematical guesswork. If we turn our focus to the opposite side of the solar system, however, the dilemma reverses. The amount of information we can gather about the Sun is so great, it's difficult to capture it, much less do anything useful with the data.

Image Courtesy Steele Hill, SOHO, NASA/ESA

On January 8, 2008, Solar Cycle 24 started. Although that might seem insignificant to most people, in about three years, it will be reaching its peak (Figure 1). Solar storms, or space weather, can have a very significant effect on modern society. These invisible outbursts can take out satellites, disrupt electrical grids and shut down radio communications. There is nothing we can do to avoid solar storms; however, early detection would make it possible to minimize the effects. And, that's what researchers at Uppsala University in Sweden are trying to do.

The problem is the amount of data being collected by the digital radio receivers—to be precise, about 6GB of raw data *per second*. There is no way to store all the data to analyze later, so Uppsala teamed up with IBM and its InfoSphere Streams software to analyze the data in real time.

LJ Associate Editor Mitch Frazier and I had an opportunity to speak with both IBM and Uppsala, and we asked them for more information on how such a feat is accomplished. We weren't surprised to hear, "using Linux". Here's our Q&A session, with some of my commentary sprinkled in.

Shawn & Mitch: What hardware does it run on?

IBM & Uppsala: InfoSphere Streams is designed to work on a variety of platforms, including IBM hardware. It runs clusters of up to 125 multicore x86 servers with Red Hat Enterprise Linux (RHEL). The ongoing IBM research project, called System S, is the basis for InfoSphere Streams and has run on many platforms, including Blue Gene supercomputers and System P.

S&M: Will it run on commodity hardware?

I&U: Yes, x86 blades.

S&M: What operating system(s) does it run on?

I&U: InfoSphere Streams runs on RHEL 4.4 for 32-bit x86 hardware and RHEL 5.2 for 64-bit x86 hardware.

S&M: Are these operating systems standard versions or custom?

I&U: They are standard operating systems.

S&M: What language(s) is it written in?

I&U: InfoSphere Streams is written in C and C++.

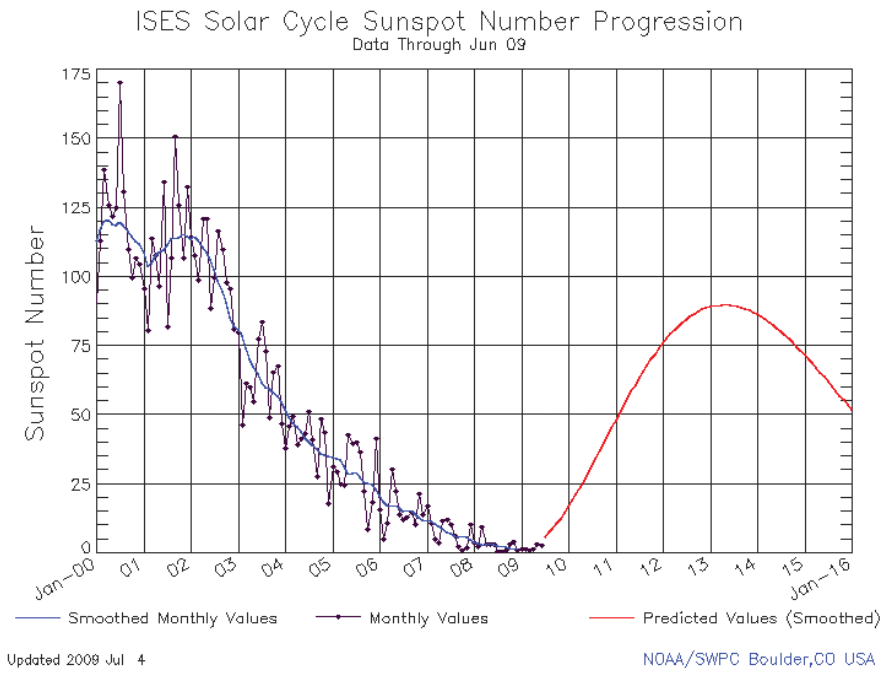


Figure 1. We are just beginning this solar cycle, which makes early detection particularly important. (Graphic Credit: National Oceanic and Atmospheric Administration, www.noaa.org)

S&M: How does a programmer interact with it? Via a normal programming language or some custom language?

I&U: Applications for InfoSphere Streams are written in a language called SPADE (Stream Processing Application Declarative Engine). Developed by IBM Research, SPADE is a programming language and a compilation infrastructure, specifically built for streaming systems. It is designed to facilitate the programming of large streaming applications, as well as their efficient and effective mapping to a wide variety of target architectures, including clusters, multicore architectures and special processors, such as the Cell processor. The SPADE programming language allows stream processing applications to be written with the finest granularity of operators that is meaningful to the application, and the SPADE compiler appropriately fuses operators and generates a stream processing graph to be run on the Streams Runtime.

[See Listing 1 for a sample of SPADE. Listing 1 is an excerpt from the "IBM Research Report—SPADE Language Specification" by Martin Hirzel, Henrique Andrade, Bugra Gedik, Vibhore Kumar, Giuliano Losa, Robert Soulé and Kun-Lung Wu, at the IBM Research Division, Thomas J. Watson Research Center.]

S&M: Is there a nontechnical user interface to it, or is all interfacing done by a programmer?

I&U: Currently, InfoSphere Streams does not have a nontechnical user interface for developers.

There is an IBM Research project that is working on providing a nontechnical user interface to allow business analysts to have programs generated and run based on information they are looking for. The project is called Mashup Automation with Run-time Invocation and Orchestration (MARIO, **domino.research.ibm.com/comm/research_projects.nsf/pages/semanticweb.Semantic Web Projects.html**).

MARIO allows business users to automate composition by letting them specify information goals, which are expressed as high-level semantic descriptions of desired flow output.

Listing 1. Example VWAP application in SPADE. VWAP, or "volume-weighted average price", is a common calculation in financial trading.

```
composite VWAP {
param
  expression<set<string>> $monitoredTickers :
                              { "IBM", "GOOG", "MSFT" };

type
  TradeInfoT   = decimal64 price, decimal64 volume;
  QuoteInfoT   = decimal64 bidprice,
                 decimal64 askprice, decimal64 asksize;
  TradeQuoteT  = TradeInfoT, QuoteInfoT,
  tuple<string ticker, string dayAndTime, string ttype>;
  TradeFilterT = TradeInfoT, tuple<timestamp ts, string ticker>;
  QuoteFilterT = QuoteInfoT, tuple<timestamp ts, string ticker>;
  VwapT        = string ticker, decimal64 minprice,
                 decimal64 maxprice, decimal64 avgprice,
                 decimal64 vwap;

graph
  stream<TradeQuoteT> TradeQuote = FileSource()  {
    param fileName  : "TradesAndQuotes.csv.gz";
          format    : csv, compressed, nodelays;
          columns   : irange(1,3), 5, irange(7,9), [11, 15, 16];
  }

  stream<TradeFilterT> TradeFilter = Functor(TradeQuote) {
    param  filter      :    ttype == "Trade"
                        && (ticker in $monitoredTickers);
    output TradeFilter  : ts = timeStringToTimestamp(dayAndTime);
  }

  stream<QuoteFilterT> QuoteFilter = Functor(TradeQuote) {
    param filter :   ttype == "Quote"
                  && (ticker in $monitoredTickers);
  }

  stream<VwapT, tuple<decimal64 sumvolume>>
    PreVwap = Aggregate(TradeFilter)
  {
    window TradeFilter  : sliding, count(4), count(1);
    param  groupBy      : ticker;
           perGroup     : true;
    output PreVwap      : ticker = Any(ticker),
                          vwap = Sum(price*volume),
                          minprice = Min(price),
                          maxprice = Max(price),
                          avgprice = Avg(price),
                          sumvolume = Sum(volume);
  }

  stream<VwapT> Vwap = Functor(PreVwap) {
    output Vwap : vwap = vwap / sumvolume;
  }

  stream<timestamp ts, decimal64 index>
    BargainIndex = Join(Vwap as V; QuoteFilter as Q)
  {
    window V            : sliding, count(1);
           Q            : sliding, count(0);
    param  equalityLHS  : V.ticker; // can also be written
                                    // as nested loop join:
           equalityRHS  : Q.ticker; // "condition: V.ticker==Q.ticker"
           perGroupLHS  : true;
    output BargainIndex :
                index = vwap > askprice*100.0
                        ? asksize*exp(vwap-askprice*100.0)
                        : 0.0;
  }

() = PerfSink(BargainIndex) { }

pragma
  debugLevel: trace;
}
```

MARIO uses existing information definitions and available information sources to generate possible applications that generate desired information goals. The optimal application is selected, deployed to the runtime and then the requested information results are displayed to the user.

S&M: Is the data captured and analyzed or just analyzed?

I&U: The benefits of stream computing overcome the problems associated with traditional analytics, which is slow, inflexible (in terms of the kinds of data it can analyze) and not well suited for capturing insights from time-sensitive events, such as tracking an epidemic or financial trading. With InfoSphere Streams, data can be captured and analyzed or just analyzed. Information can be analyzed and the data stored in files or in databases, or sent to other systems for storage. Summarized data and models also can be saved and stored. For example, an application analyzing hydrophone data to study marine mammal populations doesn't capture and store the endless hours of audio, only a model of the results. The model includes number, frequency and duration of visits by the marine mammals.

S&M: Is it open source?

I&U: No, InfoSphere Streams is not open-source code.

[Shawn notes: I'll admit, this was almost a deal-breaker for me. I was dragging out my soapbox as I contacted my IBM representative. It turns out, although InfoSphere Streams isn't open source, IBM does in fact contribute greatly to the Open Source community. I was gently reminded that IBM is a major kernel contributor, invests about $100 million annually in open-source development and "gives back" to the community it benefits from so greatly. I still would prefer InfoSphere Streams to be an open-source project; however, I suppose as long as IBM honors the GPL and is a good member of the Open Source

community, I'll put my soapbox away.]

S&M: What value does InfoSphere Streams bring to an organization (in other words, why would someone buy this)?

I&U: As the world becomes increasingly interconnected and instrumented, the amount of data is skyrocketing—and it's not just structured data found in databases, but unstructured, incompatible data captured from electronic sensors, Web pages, e-mail, audio and video. InfoSphere Streams enables massive amounts of data to be analyzed in real time, delivering extremely fast, accurate insights. These insights enable smarter business decision making and, ultimately, can help businesses differentiate themselves and gain competitive advantage.

[Shawn notes: Okay, I get it. I'm convinced InfoSphere Streams is more than a handful of Perl scripts. At this point, we were curious to hear more about the space project itself.]

S&M: What is the project name?

I&U: Swedish Institute of Space Physics REAL TIME High Frequency RADIO WEATHER STATISTICS AND FORECASTING. To put this project in context, it is part of the Scandinavian LOIS Project (**www.lois-space.net**), which in turn is an offspring of the major European Project, LOFAR (**www.lofar.org**).

S&M: What are the project's goals?

I&U: Using InfoSphere Streams, Uppsala University is analyzing massive volumes of real-time data to better understand space weather.

Scientists use high-frequency radio transmissions to study space weather or the effect of plasma in the ionosphere that can affect energy transmission over power lines, communications via radio and TV signals, airline and space travel, and satellites. However, the recent advent of new sensor technology and antennae arrays means that the amount of information collected by scientists surpassed the ability to analyze it intelligently.

The ultimate goal of the InfoSphere Streams Project is to model and predict the behavior of the uppermost part of our atmosphere and its reaction to events in surrounding space and on the Sun. This work could have lasting impact for future science experiments in space and on Earth. With a unique ability to predict how plasma clouds travel in space, new efforts can be made to minimize damage caused by energy bursts or make changes to sensitive satellites, power grids or communications systems.

S&M: Is it currently up and running?

I&U: A new generation of high-speed software defined triaxial digital radio sensors has been manufactured and is being tested to be deployed as part of this project. The InfoSphere Streams software is currently being

Powerful: Rhino

Rhino M6400/E6500
- Dell Precision M6400/ Latitude E6500
- 2.2-3.0 GHz Core 2 Duo or 2.5 GHz Core 2 Quad
- Up to 17" WUXGA LCD w/ X@1920x1200
- NVidia Quadro FX 3700M
- 80-500 GB hard drive
- Up to 16 GB RAM
- DVD±RW or Blu-ray
- 802.11a/g/n
- Starts at $1330

- High performance NVidia 3-D on a WUXGA widescreen
- High performance Core 2 Quad, 16 GB RAM
- Ultimate configurability — choose your laptop's features
- One year Linux tech support — phone and email
- Three year manufacturer's on-site warranty
- Choice of pre-installed Linux distribution:

Tablet: Raven

Raven X200 Tablet
- ThinkPad X200 tablet by Lenovo
- 12.1" WXGA w/ X@1280x800
- 1.2-1.86 GHz Core 2 Duo
- Up to 8 GB RAM
- 80-320 GB hard drive / 128 GB SSD
- Pen/stylus input to screen
- Dynamic screen rotation
- Starts at $2200

Rugged: Tarantula

Tarantula CF-30
- Panasonic Toughbook CF-30
- Fully rugged MIL-SPEC-810F tested: drops, dust, moisture & more
- 13.3" XGA TouchScreen
- 1.6 GHz Core 2 Duo
- Up to 8 GB RAM
- 80-320 GB hard drive
- Call for quote

EmperorLinux

...where Linux & laptops converge

www.EmperorLinux.com

1-888-651-6686

Model specifications and availability may vary.

Figure 2. One of the test stations that gathers data for the Uppsala University Project. This station is located in Växjö, Sweden. It includes tripole antennas and three-channel digital sensors along with a GPS antenna and receiver.

updated for this new hardware and is expected to be deployed with the new sensors in September/October 2009. Prior to the purchase of the new sensors, the project was up and running.

S&M: What type of data is being analyzed?

I&U: Massive amounts of structured and unstructured data from network sensors and antennas are being analyzed as part of this project. By using IBM InfoSphere Streams to analyze data from sensors that track high-frequency radio waves, endless amounts of data can be captured and analyzed on the fly. Over the next year, this project is expected to perform analytics on at least 6GB per second or 21.6TB per hour. The technology addresses this problem by analyzing and filtering the data the moment it streams in, helping researchers identify the critical fraction of a percent that is meaningful, while the rest is filtered out as noise. Using a visualization package, scientists can perform queries on the data stream to look closely at interesting events, allowing them not only to forecast, but also to "nowcast" events just a few hours away. This will help predict, for example, if a magnetic storm on the Sun will reach the Earth in 18–24 hours.

S&M: What type of hardware is being used to capture the data?

I&U: The project uses tri-axial electric dipole antennas (commonly known as tripole antennas) and triaxial magnetic-loop antennas (Figure 2).

S&M: What type of hardware is used in the system?

I&U: In addition to the equipment used to gather weather data, networking equipment to route the data to the Streams runtime, the InfoSphere Streams software has been developed on a 4-core x86 (Xeon) system, but can also be ported to IBM JS20 Blade Center (Power PC) and to a Lenovo ThinkPad X200s laptop.

S&M: What software, beyond InfoSphere Streams, is used?

I&U: RHEL and custom analytics written in C and C++ with some legacy FORTRAN code.

S&M: What is the user interface to the system?

I&U: InfoSphere Streams has a browser-based management console to manage the runtime. It allows people to deploy jobs, see how jobs are distributed across machines in the runtime cluster, see performance details and many other functions to manage the runtime environment. Output from the system can be streamed to various display and dashboarding applications to visualize the results of the real-time analytic processing (Figures 3 and 4).

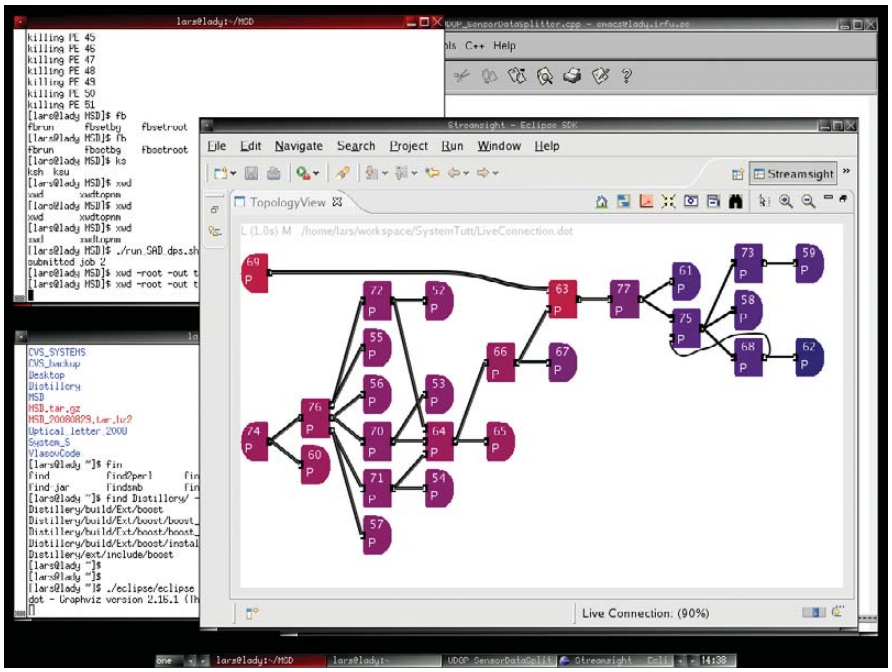


Figure 3. Streamsight is the administrative view of the running InfoSphere Stream. It allows people to visualize on which machines in the Linux cluster the various tasks are running, performance levels and other information. Each box represents a different type of analytic being run, and the lines represent data streaming between each task.

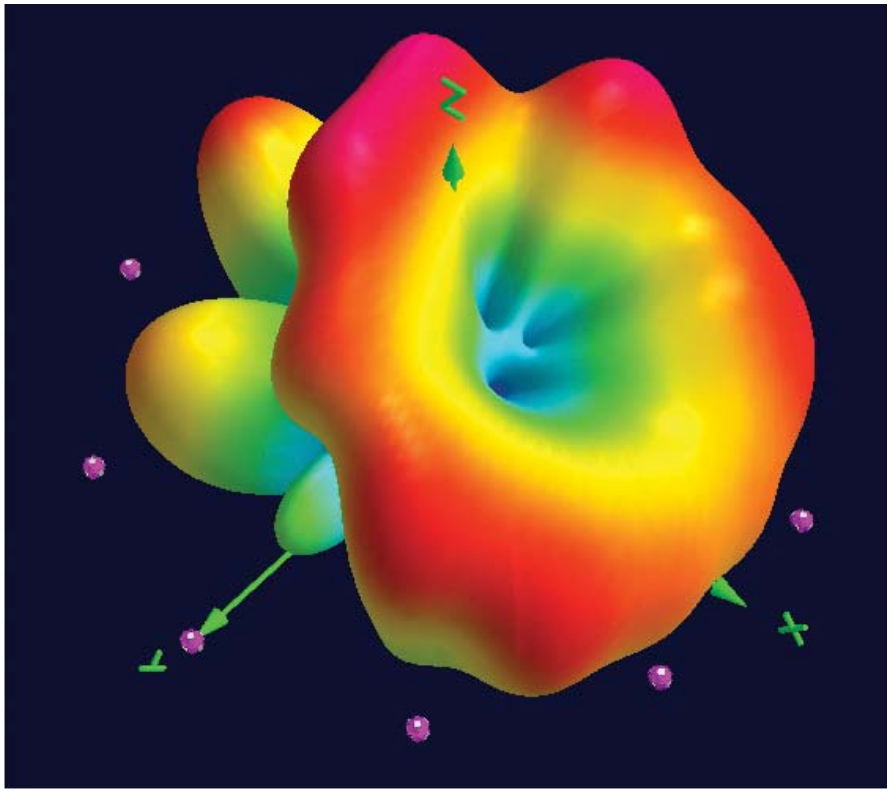


Figure 4. Visualizations of the findings the Uppsala uncovered using InfoSphere Streams.

S&M: How long did it take to develop the system?

I&U: The Space Weather application was developed over several months by a PhD candidate at Uppsala University. The larger LOIS Project has been ongoing for eight years.

S&M: How many developers/researchers were involved?

I&U: The PhD student was supported by four scientists and one research engineer from the LOIS team and several IBM researchers.

S&M: Is the space weather software open source?

I&U: No

[Shawn notes: Yes, we really did ask twice about how open these products are. We couldn't help ourselves.]

Conclusion

The Uppsala Space Weather Project is a prime example of how Linux is used as the underlying engine that makes the world go. In fact, Linux has become so mainstream in such projects, we specifically had to ask about what infrastructure the project used. Even then, the answer wasn't "Linux", but rather what version of Linux. Apparently, it was supposed to be obvious that the project would run in a Linux environment—that's the kind of presumptive attitude I like to see in the world!

Whether the information collected and analyzed by Uppsala will make a difference in how we weather Solar Cycle 24 remains to be seen. At the very least, we'll have more data about space weather than ever before in history. As to our little planet/dwarf planet/plutoid Pluto, sadly we'll have to wait until July 14, 2015 for more detailed information. The New Horizons satellite is racing there now to get more information on the little frozen body. It's hard to say how Pluto will be classified by the time it gets there, but nonetheless, we will be anxiously awaiting the data. When it finally arrives, it's pretty likely the data will be analyzed by Linux.■

Shawn Powers is the Associate Editor for *Linux Journal*. He's also the Gadget Guy for LinuxJournal.com, and he has an interesting collection of vintage Garfield coffee mugs. Don't let his silly hairdo fool you, he's a pretty ordinary guy and can be reached via e-mail at shawn@linuxjournal.com. Or, swing by the #linuxjournal IRC channel on Freenode.net.

LINUX™ JOURNAL

Linux News and Headlines Delivered To You

Linux Journal topical RSS feeds NOW AVAILABLE

linuxjournal.com/rss_feeds

Use Linux as a SAN Provider

At one-tenth the cost of the typical commercial appliance, Linux can deliver storage with speed and redundancy. Make the move toward a full-featured iSCSI SAN solution with what you already have in your server room.

Michael Nugent

Storage Area Networks (SANs) are becoming commonplace in the industry. Once restricted to large data centers and Fortune 100 companies, this technology has dropped in price to the point that small startups are using them for centralized storage. The strict definition of a SAN is a set of storage devices that are accessible over the network at a block level. This differs from a Network Attached Storage (NAS) device in that a NAS runs its own filesystem and presents that volume to the network; it does not need to be formatted by the client machine. Whereas a NAS usually is presented with the NFS or CIFS protocol, a SAN running on the same Ethernet often is presented as iSCSI, although other technologies exist.

iSCSI is the same SCSI protocol used for local disks, but encapsulated inside IP to allow it to run over the network in the same way any other IP protocol does. Because of this, and because it is seen as a block device, it often is almost indistinguishable from a local disk from the point of view of the client's operating system and is completely transparent to applications.

The iSCSI protocol is defined in RFC 3720 and runs over TCP ports 860 and 3260. In addition to the iSCSI protocol, many SANs implement Fibre Channel as a mechanism. This is an improvement over Gigabit Ethernet, mainly because it is 4 or 8Gb/s as opposed to 1Gb/s. In the same vein, 10 Gigabit Ethernet would have an advantage over Fibre Channel.

The downside to Fibre Channel is the expense. A Fibre Channel switch often runs many times the cost of a typical Ethernet switch and comes with far fewer ports. There are other advantages to Fibre Channel, such as the ability to run over very long distances, but these aren't usually the decision-making factors when purchasing a SAN.

In addition to Fibre Channel and iSCSI, ATA over Ethernet (AoE) also is starting to make some headway. In the same way that iSCSI provides SCSI commands over an IP network, AoE provides ATA commands over an Ethernet network. AoE actually is running directly on Ethernet, not on top of IP the way iSCSI does. Because of this, it has less overheard and often is faster than iSCSI in the same environment. The downside is that it cannot be routed. AoE also is far less mature than iSCSI, and fewer large networking companies are looking to support AoE. Another disadvantage of AoE is that it has no built-in security outside of MAC filtering. As it is relatively easy to spoof a MAC address, this means anyone on the local network can access any AoE volumes.

Should You Use a SAN?

The first step in moving down the road to a SAN is the choice of whether to use it. Although a SAN often is faster than a NAS, it also is less flexible. For example, the size of or the filesystem of a NAS usually can be changed on the host system without the client system having to make any changes. With a SAN, because it is seen as a block device like a local disk, it is subject to a lot of the same rules as a local disk. So, if a client is running its /usr filesystem on an iSCSI device, it would have to be taken off-line and modified not just on the server side, but also on the client side. The client would have to grow the filesystem on top of the device.

There are some significant differences between a SAN volume and a local disk. A SAN volume can be shared between computers. Often, this presents all kinds of locking problems, but with an application aware that its volume is shared out to multiple systems, this can be a powerful tool for failover, load balancing or communication. Many filesystems exist that are designed to be shared. GFS from Red Hat and OCFS from Oracle (both GPL) are great examples of the kinds of these filesystems.

The network is another consideration in choosing a SAN. Gigabit Ethernet is the practical minimum for running modern network storage. Although a 100- or even a 10-megabit network theoretically would work, the practical results would be extremely slow. If you are running many volumes or requiring lots of reads and writes to the SAN, consider running a dedicated gigabit network. This will prevent the SAN data from conflicting with your regular IP data and, as an added bonus, increase security on your storage.

Security also is a concern. Because none of the major SAN protocols are encrypted, a network sniffer could expose your data. In theory, iSCSI could be run over IPsec or a similar protocol, but without hardware acceleration, doing so would mean a large drop in performance. In lieu of this, at the very least, keeping the SAN data on its own VLAN is required.

Because it is the most popular of the various SAN protocols available for Linux, I use iSCSI in the examples in this article. But, the concepts should transfer easily to AoE if you've selected that for your systems. If you've selected Fibre Channel, things still are similar, but not as similar. You will need to rely more on your switch for most of your authentication and routing. On the positive side, most modern Fibre Channel switches provide excellent setup tools for doing this.

To this point, I have been using the terms client and server, but that is not completely accurate for iSCSI technology. In the iSCSI world, people refer to clients as initiators and servers or other iSCSI storage devices as targets. Here, I use the Open-iSCSI Project to provide the initiator and the iSCSI Enterprise Target (IET) Project to provide the target. These pieces of software are available in the default repositories of most major Linux distributions. Consult your distribution's documentation for the package names to install or download the source from **www.open-iscsi.org** and **iscsitarget.sourceforge.net**. Additionally, you'll need iSCSI over TCP/IP in your kernel, selectable in the low-level SCSI drivers section.

Setting Up the Initiator and Target

In preparation for setting up the target, you need to provide it with a disk. This can be a physical disk or you can create a disk image. In order to set up a disk image, run the dd command:

```
dd if=/dev/zero of=/srv/iscsi.image.0 bs=1 seek=10M count=1
```

This command creates a file about 10MB called /srv/iscsi.image.0 filled with zeros. This is going to represent the first iscsi disk. To create another, do this:

```
dd if=/dev/zero of=/srv/iscsi.image.1 bs=1 seek=10M count=1
```

Configuration for the IET software is located in /etc/ietd.conf. Though a lot of tweaks are available in the file, the important lines really are just the target name and LUN. For each target, exported disks must have a unique LUN. Target names are formatted specially. The official term for this name is the iSCSI Qualified Name (IQN).

The format is:

```
iqn.yyyy-mm.(reversed domain name):label
```

where iqn is required, yyyy signifies a four-digit year, followed by mm (a two-digit month) and a reversed domain name, such as org.michaelnugent. The label is a user-defined string in order to better identify the target.

Here is an example ietd.conf file using the images created above and a physical disk, sdd:

```
Target iqn.2009-05.org.michaelnugent:iscsi-target
       IncomingUser michael secretpasswd
       OutgoingUser michael secretpasswd
       Lun 0 Path=/srv/iscsi.images.0,Type=fileio
       Lun 1 Path=/srv/iscsi.images.1,Type=fileio
       Lun 2 Path=/dev/sdd,Type=blockio
```

The IncomingUser is used during discovery to authenticate iSCSI initiators. If it is not specified, any initiator will be allowed to connect to open a session. The OutgoingUser is used during discovery to authenticate the target to the initiator. For simplicity, I made them the same in this example, but they don't need to be. Note that both of these are required by the RFC to be 12 characters long. The Microsoft initiator enforces this strictly, though the Linux one does not.

Start the server using `/etc/init.d/iscsitarget start` (this may change depending on your distribution). Running `ps ax | grep ietd` will show you that the server is running.

Now you can move on to setting up the initiator to receive data from the target. To set up an initiator, place its name (in IQN format) in the /etc/iscsi/initiatorname.iscsi file (or possibly /etc/initiatorname.iscsi). An example of a well-formatted file would be the following:

```
InitiatorName=iqn.2009-05.org.michaelnugent:iscsi-01
```

In addition, you also need to modify the /etc/iscsi/iscsid.conf file to match the user names and passwords set in the ietd.conf file above:

```
node.session.auth.authmethod = CHAP
node.session.auth.username = michael
node.session.auth.password = secretpasswd
node.session.auth.username_in = michael
node.session.auth.password_in = secretpasswd
```

```
discovery.sendtargets.auth.authmethod = CHAP
discovery.sendtargets.auth.username = michael
discovery.sendtargets.auth.password = secretpasswd
discovery.sendtargets.auth.username_in = michael
discovery.sendtargets.auth.password_in = secretpasswd
```

Once this is done, run the iscsiadm command to discover the target.

```
iscsiadm -m discovery -t sendtargets -p 192.168.0.1 -P 1
```

This should output the following:

```
Target: iqn.2009-05.org.michaelnugent:iscsi-target
     Portal: 192.168.0.1:32360,1
          IFace Name: default
```

Now, at any time, you can run:

```
iscsiadm -m node -P1
```

which will redisplay the target information.

Now, run `/etc/init.d/iscsi restart`. Doing so will connect to the new block devices. Run `dmesg` and `fdisk -l` to view them. Because these are raw block devices, they look like physical disks to Linux. They'll show up as the next SCSI device, such as /dev/sdb. They still need to be partitioned and formatted to be usable. After this is done, mount them normally and they'll be ready to use.

This sets up the average iSCSI volume. Often though, you may want machines to run entirely diskless. For that, you need to run root on iSCSI as well. This is a bit more involved. The easiest, but more expensive way is to employ a network card with iSCSI built in. That allows the card to mount the volume and present it without having to do any additional work. On the downside, these cards are significantly more expensive than the average network card.

To create a diskless system without an iSCSI-capable network card, you need to employ PXE boot. This requires that a DHCP server be available in order for the initiator to receive an address. That DHCP server will have to refer to a TFTP server in order for the machine to download its kernel and initial ramdisk. That kernel and ramdisk will have iSCSI and discovery information in it. This enables the average PXE-enabled card to act as a more expensive iSCSI-enabled network card.

Multipathing

Another feature often run with iSCSI is multipathing. This allows Linux to use multiple networks at once to access the iSCSI target. It usually is run on separate physical networks, so in the event that one fails, the other still will be up and the initiator will not experience loss of a volume or a system crash. Multipathing can be set up in two ways, either active/passive or active/active. Active/active generally is the preferred way, as it can be set up not only for redundancy, but also for load balancing. Like Fibre Channel, multipath assigns World Wide Identifiers (WWIDs) to devices. These are guaranteed to be unique and unchanging. When one of the paths is removed, the other one continues to function. The initiator may experience slower response time, but it will continue to function. Re-integrating the second path allows the system to return to its normal state.

RAID

When working with local disks, people often turn to Linux's software RAID or LVM systems to provide redundancy, growth and snapshotting. Because SAN volumes show up as block devices, it is possible to use these tools on them as well. Use them with care though. Setting up RAID 5 across three iSCSI volumes causes a great deal of network traffic and almost never gives you the results you're expecting. Although, if you have enough bandwidth available and you aren't doing many writes, a RAID 1 setup across multiple iSCSI volumes may not be completely out of the question. If one of these volumes drops, rebuilding may be an expensive process. Be careful about how much bandwidth you allocate to rebuilding the array if you're in a production environment. Note that this could be used at the same time as multipathing in order to increase your bandwidth.

To set up RAID 1 over iSCSI, first load the RAID 1 module:

```
modprobe raid1
```

After partitioning your first disk, /dev/sdb, copy the partition table to your second disk, /dev/sdc. Remember to set the partition type to Linux RAID autodetect:

```
sfdisk -d /dev/sdb | sfdisk /dev/sdc
```

Assuming you set up only one partition, use the mdadm command to create the RAID group:

```
mdadm --create /dev/md0 --level=1 --raid-disks=2 /dev/sdb1 /dev/sdc1
```

After that, `cat` the /etc/mdstat file to watch the state of the synchronization of the iSCSI volumes. This also is a good time to measure your network throughput to see if it will stand up under production conditions.

Conclusion

Running a SAN on Linux is an excellent way to bring up a shared environment in a reasonable amount of time using commodity parts. Spending a few thousand dollars to create a multiterabyte array is a small budget when many commercial arrays easily can extend into the tens to hundreds of thousands of dollars. In addition, you gain flexibility. Linux allows you to manipulate the underlying technologies in ways most of the commercial arrays do not. If you're looking for a more-polished solution, the Openfiler Project provides a nice layout and GUI to navigate. It's worth noting that many commercial solutions run a Linux kernel under their shell, so unless you specifically need features or support that isn't available with standard Linux tools, there's little reason to look to commercial vendors for a SAN solution.■

Michael Nugent has spent a good deal of his time designing large-scale solutions to fit into tiny budgets, leveraging Linux to fulfill roles that typically would be filled by large commercial appliances. Recently, Michael has been working to design large, private clouds for SaaS environments in the financial industry. When not building systems, he likes sailing, scuba diving and hanging out with his cat, MIDI. Michael can be reached at michael@michaelnugent.org.

October 24, 2009
8 AM - 5 PM
Radssion Resort Orlando-Celebration
2900 Parkway Blvd
Kissimmee, Florida 34747

http://www.floridalinuxshow.com/

Join us for the hottest Linux show in Florida this coming October!

Donald L. Corbet "TheLinuxGuy®"
2009 Show Chairman

Keynote Speaker:
Mr. Robin "Roblimo" Miller

Bryan Smith
2009 Technical Director

Chris Crisafulli
2009 Community Relations Director

Dan Trevino
2009 Web Operations Director

Rod Sharp
2009 Show Director

Charitable Outreach

Children's Miracle Network®

We believe it's important to support worthwhile charitable causes. This year, Florida Linux Show will donate $1.00 of every $15.00 show admission to the Children's Miracle Network. Donation jars will also be available during the show for anyone who wishes to make an additional contribution.

Linux Presentation Areas
- Plug Computing
- Security
- Ubuntu, Ubuntu, Ubuntu
- Linux Thin Clients
- Virtualization
- Cloud Computing
- Interoperability
- Handheld Devices

Open Source Presentation Areas
- Open Source Software
- Open Source Solutions
- Solutions for Businesses
- Open Source Resources
- Open Source PBXs

Official Sponsors

GOLD

GSA PROPOSAL.COM

SILVER

O'REILLY®

BRONZE

TSHIRTBORDELLO.COM
Your place for dirty laundry

MEDIA

LINUX JOURNAL®

LINUX PRO MAGAZINE

LinuxQuestions.org

Supporting Florida LUGs: GatorLUG / Linux enthusiasts & professionals / Pensacola Linux User's Group / Florida Linux Users Group / Linux Users Group of Tallahassee/ Thomasville / Greater Orlando LUG

Florida Ubuntu Loco Team

OSS Atlanta.org

Open Source Software

ADVANCED MESSAGE QUEUING PROTOCOL (AMQP)

AMQP is an open standard for enterprise messaging, designed to support messaging for almost any distributed or business application.

JOSHUA KRAMER

What if, using a single service call, you easily could ask a computing cloud to give you the readings from thermometers in 100 different locations? Or, perhaps you'd like to know the status of the 89 servers under your control. In the past, you might have accomplished those things by writing a server dæmon. Your dæmon might have managed each of hundreds of connections, conducting specific operations on each connection. However, with the advent of AMQP and the Apache Qpid Project, it's possible to concentrate on the data processing and let another program handle the messaging.

AMQP is an innovative open messaging protocol. Created by John O'Hara and others at JPMorgan to replace proprietary products, the AMQ protocol defines both the wire-level formats and the behavior of messaging server and client software. Using the above example, you could send a single message to the AMQP server with a topic such as server_stats or thermometer_readings. The AMQP server listens for messages with those topics and routes the messages to the applications connected to the AMQP server.

A Bit of History

AMQP began in 2003 with John O'Hara at JPMorgan-Chase. O'Hara was looking for a messaging solution that provided high durability, extremely high volume and a high degree of interoperability. In the types of environments addressed with AMQP, there is an economic impact if a message is lost, arrives late or is processed improperly. With volumes greater than 500,000 messages per second, the requirements were high. The commercial products that were available at the time could not deliver the level of service required, and banks were known to develop their own enterprise middleware to fill in the gaps. However, developing enterprise middleware is complex and difficult, and bank middleware would come and go.

As he reflected on other highly successful protocols, such as Ethernet, TCP/IP and HTTP, O'Hara noted several similarities. Namely, each protocol was royalty-free and not encumbered by patents. Furthermore, the protocols had a strong specification created by an independent body. Freely available implementations of the protocol specifications allowed developers to pick them up and find interesting uses for them quickly. Strong governance and user-driven design made these protocols a technical and economic success.

With AMQP, O'Hara wanted to have a freely available implementation of the AMQ protocol in use in a mission-critical place at JPMorgan. With this goal in mind, he contracted with the iMatix Corporation to create the first implementation, OpenAMQ. This implementation then was put into production in a trading application with more than 2,000 users.

Today, many companies collaborate on AMQ. Several brokers are available, including RabbitMQ, OpenAMQ and Apache Qpid (also known as Red Hat MRG Messaging). In this article, I describe the Apache Qpid server. Up for discussion is the Qpid M4 release, and you can download it via the link in the Resources for this article. I also demonstrate how to compile and install the C++ version of the server and write example applications in Python.

Anatomy of a Server

Figure 1 depicts the anatomy of a Qpid server. It is important to know about three components of an AMQP server: local queues, server queues and exchanges.

The exchange determines message delivery based on the message header. Exchanges can provide different delivery schemes, such as direct (deliver this message to queue XYZ), publish-subscribe (deliver this message to all queues subscribed to topic spring.flowers) and XML (all messages that match XPath query Z go to Queue Y). A server queue is a queue that resides on the server and receives messages from the exchange. A local queue is a queue associated with an instance of an application. Local queues are bound to server queues, so any message delivered to the server queue appears on the local queue. More than one local queue can be bound to a server queue. This is handy when you have a farm of machines (or processes) responding to requests. In this case, messages will be delivered from the server queue to the local queue on a round-robin basis.

In addition to the server terms, note that programs reading from queues are called consumers, and those writing to exchanges are producers. This can become confusing when you have applications that act as both consumers and producers.

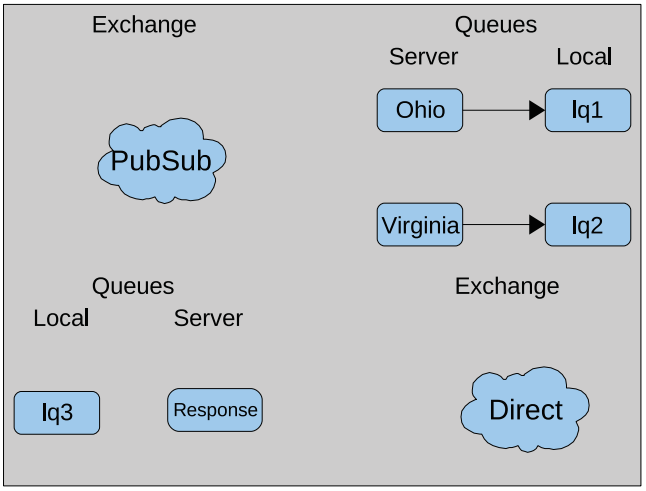


Figure 1. Anatomy of a Qpid Server Work Flow

Therefore, it makes sense to use standard terminology for clients and servers, where a client sends a request and expects a response on a reply queue, and a server listens for messages and responds as requested.

The Qpid broker discussed here comes with XML files describing the AMQP specification. These files define the formats used by the server and clients. The server and libraries use these specifications to formalize parameters, such as wire format, server commands and error messages. Managing these specifications outside the server allows you to maintain compatibility across different server vendors and different server versions. In theory, you should be able to replace a Qpid server with OpenAMQ, RabbitMQ or any other AMQP-compliant server and have it work out of the box. In practice, different servers support different versions of the specification or require different options. For example, the Qpid Java Client supports three versions of the protocol: 0-8, 0-9 and 0-10. However, the C++ client supports only 0-10 in its latest release. RabbitMQ, a competing AMQP broker, supports only 0-8 and 0-9 of the specification. Because of this, the best results are when using clients and brokers from the same product line.

Managing these specifications outside the server allows you to maintain compatibility across different server vendors and different server versions.

Apache Qpid and its commercial counterpart, Red Hat MRG Messaging, are versatile products. They offer many features not covered here. For example, you can use SSL and InfiniBand fabric as interconnects, and you can control how clients connect to your server via ACLs and authentication. I highly recommend the Red Hat MRG documentation for further reference on these features.

Installing the Server

First, a word about prerequisites. These examples were created on CentOS 5.2 with the standard development packages as well as Ruby. Also, note that certain versions of PyXML present conflicts that will break the tests run after installation.

To install the server, simply download the full M4 release from the URL noted in the Resources section of this article to your preferred development directory and un-archive the package. Once you have a directory structure, go to the server's directory by typing:

```
cd qpid-M4/cpp
```

Initially, there is no configure script; create it by running the `bootstrap` command. Once bootstrap completes, do the standard `configure`, `make` and `make install`.

One step the installation process does not perform is installing the AMQP specification files. These specification files are contained in the specs directory under qpid-M4. Copy the files found there to /usr/local/share/qpid/specs.

After installation, it's a good idea to run tests to ensure that all prerequisites have been satisfied. Start a new shell, change directories to /usr/local and `su` to root. Then, run the Qpid dæmon with the command:

```
sbin/qpidd -t --auth no
```

Once the broker is running, return to the original shell. Move from the cpp directory to the python directory contained within qpid-M4. Run the Python tests using:

```
run-tests -s 0-10-errata -I cpp_failing_0-10.txt
```

If the tests run and return no errors, proceed to install the Python modules by running this command as root:

```
python setup.py install
```

Writing Applications—A Simple Model

This example demonstrates a simple application used to query server status. The server script runs rpm to query the packages stored on the system and returns the list, with its PID, to the client. The program generating the requests is the client, and the server is a dæmon running on a "remote server". It has an event loop that waits for requests.

In this example, the scripts use a combination of two message-routing methods: publish-subscribe (pubsub) to deliver the requests to all listening servers and direct to route the replies directly to the calling client.

Listing 1 describes the client, which is fairly straightforward. First, the client reads the spec file and then creates the Qpid connection. The connection is made by creating a standard Python socket object and passing that object to the connection's constructor. The connection, in turn, provides a session object when the session() method is called.

Next, the client creates the reply-to server queue. Note that the reply-to server queue name contains the session ID. This gives each client a unique server queue. The queue then is bound to the amq.direct exchange, which uses queue names as its routing keys. Using the queue name for the server queue

Listing 1. Client-Side Python Program

```
#!/usr/bin/python

from qpid.util import connect, ssl
from qpid.connection import Connection, sslwrap
from qpid.datatypes import Message, RangedSet, uuid4
from qpid.queue import Empty
from qpid.spec import load

# First, load the correct specification file.
amqSpec = load('/usr/local/share/qpid/specs/amqp.0-10.xml')

# Now, connect to the server.
socket = connect("localhost", 5672)
connection = Connection (sock=socket,
                         spec=amqSpec,
                         username = "guest",
                         password = "guest")
connection.start()
session = connection.session(str(uuid4()))

# Declare the reply queue:
replyQueueName = "producerReply_" + session.name
replyQueue = session.queue_declare(queue=replyQueueName,
                                   exclusive=True,
                                   auto_delete=True)
session.exchange_bind(exchange="amq.direct",
                      queue=replyQueueName,
                      binding_key=replyQueueName)

# Declare a local queue to which we subscribe the reply-to queue
localQueueName = "producerLocalQueue_" + session.name
localQueue = session.incoming(localQueueName)
session.message_subscribe(queue=replyQueueName,
                          destination=localQueueName)
localQueue.start()

# Now, create a message with a request.
message_properties = session.message_properties()
message_properties.reply_to = session.reply_to("amq.direct",
                                               replyQueueName)
delivery_properties =
    session.delivery_properties(routing_key="SERVER_STATUS")
requestMsgText = "RPM_STATUS"

# Send the message and wait for a response.
session.message_transfer(destination="amq.topic",
                         message=Message(message_properties,
                                         delivery_properties,
                                         requestMsgText))

while True:
    try:
        message = localQueue.get(timeout=60)
        content = message.body
        session.message_accept(RangedSet(message.id))
        print content
    except Empty:
        print "No more messages!"
        break
```

Listing 2. Server-Side Python Program

```
#!/usr/bin/python

import subprocess
import os
from qpid.util import connect, ssl
from qpid.connection import Connection, sslwrap
from qpid.datatypes import Message, RangedSet, uuid4
from qpid.queue import Empty
from qpid.spec import load
from qpid.queue import Empty
from qpid.session import SessionException

# processRequest: this is what actually does the work.
def processRequest(requestMessage):

    print "Servicing Request"
    proc = subprocess.Popen('rpm -qa',
                        shell=True,
                        stdout=subprocess.PIPE,
                        )
    stdout_value = proc.communicate()[0]
    myPid = os.getpid()
    ret_value = "From Server PID " \
                + str(myPid) + ":\n" + stdout_value \
                + "---------------------------------\n"
    return ret_value

# First, load the correct specification file.
locSpec = load('/usr/local/share/qpid/specs/amqp.0-10.xml')

# Now, connect to the server.
socket = connect("localhost", 5672)
connection = Connection (sock=socket,
                        spec=locSpec,
                        username = "guest",
                        password = "guest")
connection.start()
session = connection.session(str(uuid4()))

# Declare the listening server queue and connect to server queue.
# Create server queue if it does not exist.

myPid = os.getpid()
serverQueueName = "serverListenQueue" + str(myPid)

localQueueName = "serverListenLocal_" + session.name
session.queue_declare(queue=serverQueueName,
                      exclusive=True)
session.exchange_bind(exchange="amq.topic",
                      queue=serverQueueName,
                      binding_key="SERVER_STATUS")
session.message_subscribe(queue=serverQueueName,
                          destination=localQueueName)
localQueue = session.incoming(localQueueName)
localQueue.start()

# Now, start an event loop.
while True:
    try:
        requestObj = localQueue.get(timeout=60)
        session.message_accept(RangedSet(requestObj.id))
        requestStr = requestObj.body
        requestProperties = requestObj.get("message_properties")
        replyTo = requestProperties.reply_to
        if replyTo == None:
            raise Exception("This message is missing " \
                            + "the 'reply_to'"  property, " \
                            + "which is required")
        responseMessage = processRequest(requestStr)
        props = session.delivery_properties(
                                routing_key=replyTo["routing_key"])
        session.message_transfer(destination=replyTo["exchange"],
                                 message=Message(props,
                                                 responseMessage))
    except Empty:
        continue
```

and delivering replies to the amq-direct exchange ensures that multiple copies of the server receive only their own replies.

After the server queue is declared, the program creates a local queue and subscribes it to the server queue. Once the local queue is subscribed, the program is ready to transmit a message.

The client then creates the request message. Because the program is using publish-subscribe, the routing key is set to the topic. In this case, the topic is SERVER_STATUS. Any server that is subscribed to the topic SERVER_STATUS will receive this particular message. The client also supplies the exchange type and the routing key for the reply-to fields. For this message, it is the amq-direct exchange and the name of the server queue that was created previously.

Finally, the client creates the message itself (the text "RPM_STATUS") and delivers it to the exchange. After the message is delivered, the client waits for a reply and prints the contents of the reply to the screen.

Listing 2 defines the server. This application will listen for messages with the topic SERVER_STATUS, run rpm to query the package contents of the system and send a reply. The first steps are similar to Listing 1 in that the server starts a connection and uses the connection to get a session and create a server queue. The server then subscribes the local queue, starts the queue, and the program is ready to respond to requests.

Using the queue name for the server queue and delivering replies to the amq-direct exchange ensures that multiple copies of the server receive only their own replies.

Listing 3. Multiserver Weather Client

```
#!/usr/bin/python

from qpid.util import connect, ssl
from qpid.connection import Connection, sslwrap
from qpid.datatypes import Message
from qpid.datatypes import RangedSet
from qpid.datatypes import uuid4
from qpid.queue import Empty
from qpid.spec import load

# First, load the correct specification file.
amqSpec = load('/usr/local/share/qpid/specs/amqp.0-10.xml')

# Now, connect to the server.
socket = connect("localhost", 5672)
connection = Connection (sock=socket,
                         spec=amqSpec,
                         username = "guest",
                         password = "guest")
connection.start()
session = connection.session(str(uuid4()))

# Declare the reply queue:
replyQueueName = "weatherReply_" + session.name
replyQueue = session.queue_declare(queue=replyQueueName,
                                   exclusive=True,
                                   auto_delete=True)
session.exchange_bind(exchange="amq.direct",
                      queue=replyQueueName,
                      binding_key=replyQueueName)

# Declare a local queue to which we subscribe the reply-to queue
localQueueName = "weatherLocalQueue_" + session.name
localQueue = session.incoming(localQueueName)
session.message_subscribe(queue=replyQueueName,
                          destination=localQueueName)
localQueue.start()

# Now, create messages with requests.

weatherStates = ['ohio', 'virginia']

for state in weatherStates:
    for i in range(1, 11):
        message_properties = session.message_properties()
        message_properties.reply_to = session.reply_to("amq.direct",
                                                       replyQueueName)
        routingKey = "weather." + state
        delivery_properties = session.delivery_properties(
                                    routing_key=routingKey)
        requestMsgText = "weather_report"
        session.message_transfer(destination="amq.topic",
                                 message=Message(message_properties,
                                                 delivery_properties,
                                                 requestMsgText))
        print "Sent message " + str(i) + " with key " + routingKey

while True:
    try:
        message = localQueue.get(timeout=60)
        content = message.body
        session.message_accept(RangedSet(message.id))
        print content
    except Empty:
        print "No more messages!"
        break
```

In the event loop, the server first receives a request from the local queue. If there is no request within the timeout value (60 seconds), the get() method will raise an Empty exception. Because the server needs to serve requests continually, the program catches the Empty exception and simply continues. When a message arrives, the server runs the processRequest method and constructs data with the method's return values. The reply message takes exchange and routing key information from the original message's reply-to field and then is delivered to the exchange.

A Slightly More Complex Model

With AMQP, it is possible to construct a queuing system that allows a server farm to respond to multiple different kinds of requests. This example considers weather prediction models. Here, there are different server clusters, with one cluster serving each state. In such a case, it would be extremely handy to be able to send requests to each farm from an arbitrary location.

This example requires three processes. The first process (the client) delivers requests, and it is fundamentally the same as the client in the previous example. It is different only in that it loops over a list to deliver ten weather requests for Ohio and ten requests for Virginia. On the receiving end, there are two servers: one for Ohio and one for Virginia. Each server subscribes to the amq.topic exchange with the routing key #.ohio or #.virginia. Furthermore, each server has the ability to subscribe to existing server queues or create those that do not exist.

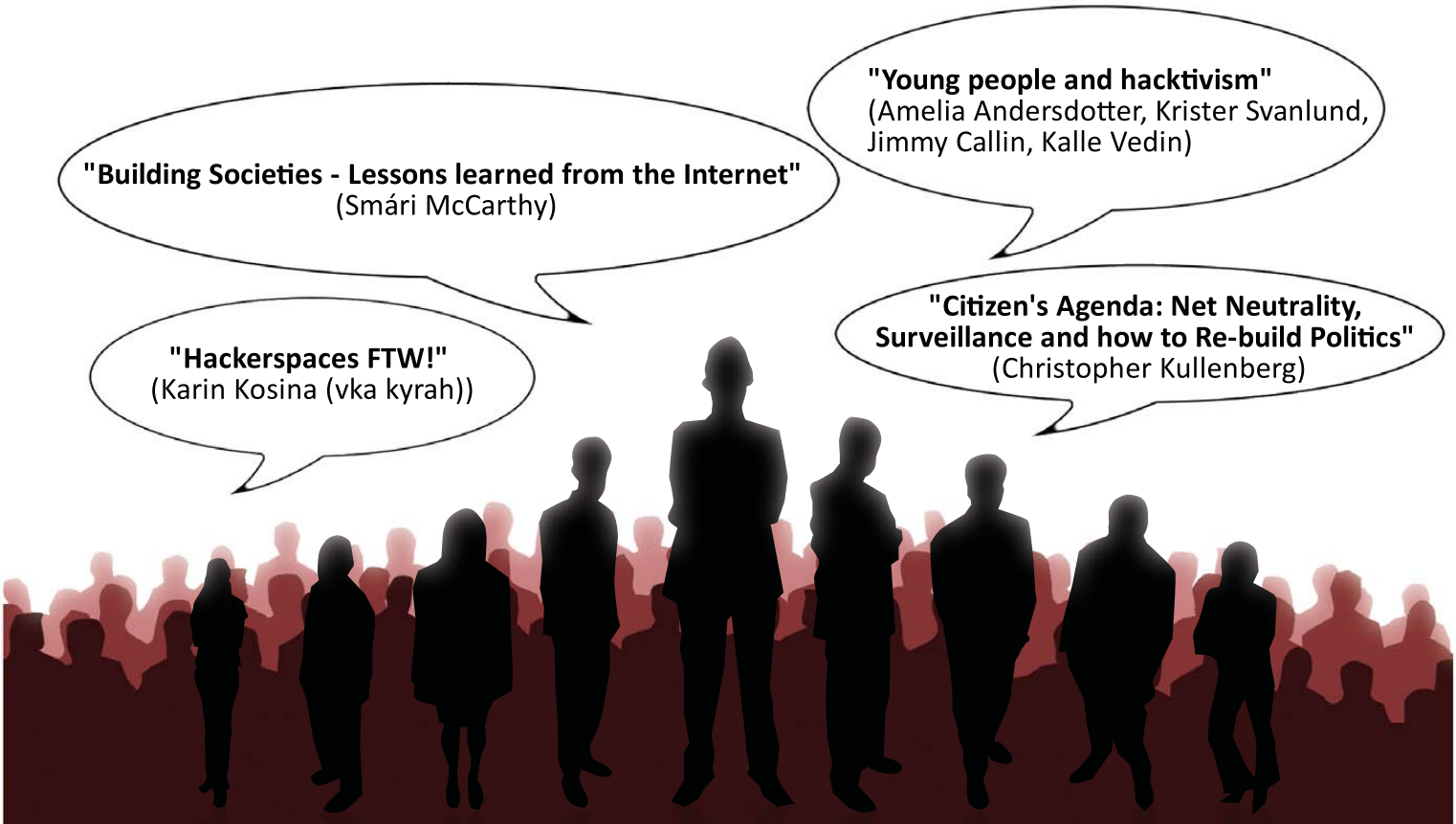
These routing keys contain wild cards. When the routing key contains a hash mark in place of text, the exchange will match any text where the hash mark resides. In this way, the weather predicting dæmons using #.ohio also would respond to requests for topic news.ohio and sports.ohio. Likewise, if a sports reporting dæmon had invaded the cluster and was listening for sports.#, the subscriptions for both the sports dæmon and the weather reporting dæmon for Ohio would match sports.ohio.

Listing 3 contains the client, and Listing 4 contains the server for Ohio. Create the server for Virginia by duplicating the server for Ohio and replacing all occurrences of Ohio with Virginia. (When you do so, make sure all routing keys have all lowercase characters.)

When you run this demonstration, run several copies each of the Ohio and Virginia servers. The messages for each state will be picked up in a round-robin manner by the respective

FSCONS 2009

Free Society Conference and Nordic Summit



fscons.org

Burgården Conference Center

Gothenburg, Sweden

13-15 November 2009

Listing 4. Multiserver Server Side (Ohio)

```
#!/usr/bin/python

import subprocess
import os
from qpid.util import connect, ssl
from qpid.connection import Connection, sslwrap
from qpid.datatypes import Message, RangedSet, uuid4
from qpid.queue import Empty
from qpid.spec import load
from qpid.queue import Empty
from qpid.session import SessionException

# ProcessRequest: this is what actually does the work.
def processRequest(requestMessage):
    print "Predicting the weather for Ohio"
    myPid = os.getpid()
    ret_value = "From Server PID " \
                + str(myPid) + ": Ohio is sunny and 70!"
    return ret_value

# First, load the correct specification file.
locSpec = load('/usr/local/share/qpid/specs/amqp.0-10.xml')

# Now, connect to the server.
socket = connect("localhost", 5672)
connection = Connection(sock=socket, spec=locSpec,
                        username="guest", password="guest")
connection.start()
session = connection.session(str(uuid4()))

# Declare the listening server queue and connect to server queue.
# Create server queue if it does not exist.

myPid = os.getpid()
listenTopic = "#.ohio"
serverQueueName = "serverListenQueueOhio"
localQueueName = "localQueue_" + str(myPid)
try:
    session.message_subscribe(queue=serverQueueName,
                              destination=localQueueName)
    localQueue = session.incoming(localQueueName)
    localQueue.start()
    print "Successfully attached to existing server queue."
except SessionException, e:
    print "Could not find server queue, so I am creating it."
    session = connection.session(name=str(uuid4()), timeout=0)
    session.queue_declare(queue=serverQueueName, exclusive=False)
    session.exchange_bind(exchange="amq.topic",
                          queue=serverQueueName,
                          binding_key=listenTopic)
    session.message_subscribe(queue=serverQueueName,
                              destination=localQueueName)
    localQueue = session.incoming(localQueueName)
    localQueue.start()
except Exception, e:
    print "Something broke unexpectedly."
    os.exit()

# Now, start a message loop.
while True:
    try:
        requestObj = localQueue.get(timeout=60)
        session.message_accept(RangedSet(requestObj.id))
        requestStr = requestObj.body
        print "Received message."
        requestProperties = requestObj.get("message_properties")
        replyTo = requestProperties.reply_to
        if replyTo == None:
            raise Exception("This message is missing the "
                            + "'reply_to' property, "
                            + "which is required")
        responseMessage = processRequest(requestStr)
        props = session.delivery_properties(
                              routing_key=replyTo["routing_key"])
        print "Responding to request."
        session.message_transfer(destination=replyTo["exchange"],
                              message=Message(props, responseMessage))
    except Empty:
        continue
```

instances of the server script. In turn, the client will print a listing of the weather forecasts with the server PIDs.

Conclusion

The AMQ protocol and its open-source implementations provide a solution for anyone requiring high-performance, versatile message communications. As I demonstrate here, using the Apache Qpid message broker is an easy way to achieve these goals. See my blog at **www.globalherald.net/jb01** for further discussion regarding this article.■

By day, Joshua Kramer is an integration specialist with Belron US, the autoglass company. By night, he creates unique social-networking presences using technologies such as Linux, Django and AMQP. Josh has a Bachelor's degree in Philosophy from Capital University and lives in rural Ohio.

Resources

"Is AMQP on the way to providing real business interoperability?" by Steven Robbins: **www.infoq.com/news/2008/08/amqp-progress**

"Toward a Commodity Enterprise Middleware: Can AMQP Enable a New Era in Messaging Middleware? A Look Inside Standards-Based Messaging with AMQP" by John O'Hara: **queue.acm.org/detail.cfm?id=1255424**

Source: **qpid.apache.org/download.html**

Lane Becker of Get Satisfaction speaks at Do It With Drupal 2008.

DO IT WITH DRUPAL

A 3 DAY SEMINAR
NEW ORLEANS, LA
DECEMBER 9 - 11, 2009

STILL GETTING YOUR HEAD AROUND DRUPAL?

Learn to harness the power of this flexible open-source social content management system and web application framework.

Do It With Drupal is in its second year and the 2009 event promises to be even better than last year.

- Learn from the world's top Drupal experts
- Examine and dissect successful Drupal sites
- Discover new site-building strategies
- Hear from community-building experts
- Connect with other Drupal professionals

"Your favourite community personalities, the developers of the most power-packed modules and experienced Drupal solution providers join together to teach you the tools to build amazing Drupal powered websites. I loved it! I can't wait to put what I learned at DIWD into action."
Matthew Nuzum
Ubuntu.com Webmaster

"Broad enough for beginners, focused enough for experts. A required event for the Drupal community."
Daniel Chvatik
CTO, Adulmec

REGISTER NOW
DoItWithDrupal.com

Lullabot powered

Registration price includes:
- Admission to all sessions and keynote presentations
- Social and networking events with Drupal's leaders and top talent
- Breakfast and lunch each day of the event
- Access to entire 2009 Do It With Drupal video archive

Early registration and hotel room discounts available.

IPv4 Anycast

with Linux and Quagga

Ease configuration headaches and improve availability with anycast.

Philip Martin

"DNS is down and nothing is working!" is not something anyone ever wants to hear at 3am. Virtually every service on a modern network depends on DNS to function. When DNS goes down, you can't send mail, you can't get to the Web, you can't do much—hopefully, your coffeemaker is not Web-enabled! Administrators do a lot of things to mitigate this risk. The traditional safeguard is to establish multiple DNS servers for a given site. Each DNS client on the network is configured with each of those servers' IP addresses. The chances of all of those servers failing in a catastrophic way are fairly small, so you have a margin of safety.

On the other hand, many stub resolvers will take only two DNS servers, making it nearly impossible to have any meaningful geographical dispersion in your DNS topology. DNS stub resolvers generally use the first of two configured DNS servers exclusively. Consequently, you end up with one server taking the entire query load and one idling, waiting for a failure. Not optimal, but hey, that's the price of redundancy...right? It doesn't have to be.

DNS redundancy and failover is a classic use case for anycast. Anycast is the concept of taking one IP address and sharing it between multiple servers, each unaware of the others. The DNS root nameservers make extensive use of anycast. There are currently 16 root nameserver IP addresses, only eight of which make use of anycast. There are 167 servers that respond to those 16 IP addresses.

Of course, anycast is not limited to DNS. It can be used to provide redundancy and failover for any number of stateless protocols and applications. Anycast might sound a little like multicast, but aside from the one-to-many, IP-to-endpoint relationship, they have very little in common. Multicast takes packets from one sender and delivers them to multiple endpoints, all of which subscribe to a single multicast address using a number of multicast-specific routing technologies. Anycast takes packets from one sender and delivers those packets to the "closest" of a number of possible endpoints using nothing more than standard unicast routing.

How Does It All Work?

Let's start with some terminology:

- An endpoint (also known as a node) is a server that responds to an anycast address and, by extension, provides services on that address.
- An anycast address is an IP address that has multiple endpoints associated with it. Anycast addresses can be from any part of the normal IPv4 address space.
- A service address is a unique IP address on a physical device on the system. Service addresses are used for administrative or monitoring access to anycast endpoints.
- IGP anycast refers to an anycast scheme confined to a single network (typically a larger network with multiple physical sites). I cover IGP anycast in this article.
- BGP anycast refers to an anycast scheme that spans multiple networks and can span the entire Internet. The DNS root servers use BGP anycast.

Anycast endpoints participate in whatever internal routing protocol is being run on your network. All endpoints for a given anycast IP advertise a host route (also known as a /32) for the anycast IP to the router. In other words, each endpoint announces that the anycast IP can be reached through it. Your routers will see the advertisements coming from the various servers and determine the best path to that IP address. Therein lies the magic. Because the IP address is

Of course, anycast is not limited to DNS. It can be used to provide redundancy and failover for any number of stateless protocols and applications.

advertised from multiple locations, your router ends up choosing the best path to that IP address, according to the metric in use by that routing protocol—meaning either the path with the fewest hops (RIP), the highest bandwidth path (OSPF) or some other measurement of network goodness. When you send a request to an anycast IP address, it will be routed to the single server with the best metric according to the routers between you and the server.

What if that server fails? If the host fails, it will stop sending out routing advertisements. The routing protocol will notice and remove that route. Traffic then will flow along the next best path. Now, the fact that the host is up does not necessarily mean that the service is up. For that, you need some sort of service monitoring in place and the capability to remove a host from the anycast scheme on the fly.

GENSTOR

Linux - FreeBSD - x86 Solaris - MS etc.

Proven technology. Proven reliability.

When you can't afford to take chances with your business data or productivity, rely on a GS-1245 Server powered by the Intel® Xeon® Processors.

Quad Core Woodcrest

2 Nodes & Up to 16 Cores - in 1U

Ideal for high density clustering in standard 1U form factor. Upto 16 Cores for high CPU needs. Easy to configure failover nodes.

Features:

- 1U rack-optimized chassis (1.75in.)
- Up to 2 Quad Core Intel® Xeon® Woodcrest per Node with 1600 MHz system bus
- Up to 16 Woodcrest Cores Per 1U rackspace
- Up to 64GB DDR2.667 & 533 SDRAM Fully Buffered DIMM (FB-DIMM) Per Node
- Dual-port Gigabit Ethernet Per Node
- 2 SATA Removable HDD Per Node
- 1 (x8) PCI_Express Per Node

Servers : : Storage : : Appliances

Genstor Systems, Inc.

780 Montague Express. # 604
San Jose, CA 95131
Www.genstor.com
Email: sales@genstor.com
Phone: 1-877-25 SERVER or 1-408-383-0120

intel Xeon inside™

Intel®, Intel® Xeon®, Intel® Inside® are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Naturally, myriad other details need to be worked out when designing an anycast scheme. The general concept is pretty simple, and small implementations are easy to set up. However, no matter what size implementation you're dealing with, proper IP address architecture is a must. Your anycast address should be on its own subnet, separate from any other existing subnets. The anycast subnet must never, ever, be included in a summary.

Implementation Details

Many projects provide routing protocol dæmons for Linux, any number of which would be usable for this scenario. For this article, I use Quagga, which is a fork of GNU Zebra. Quagga is available both on the install media and from the standard package repositories of pretty much every enterprise-oriented Linux distribution.

For the following examples, I also use a network populated with Cisco routers, running OSPF version 2, for IPv4. Quagga also supports BGP, RIP, RIPng and OSPFv3. The remainder of this article assumes at least a basic familiarity with OSPF theory and configuration. (See Resources for links to basic primers.) Cisco also publishes a ton of very good reference material (again, see Resources). I cover the required configuration on the router side, but not in extensive detail.

Using a loopback interface alias instead of a physical interface alias allows you to do a number of cool things.

Now, let's get down to the good stuff: setting up Quagga on Linux. To begin, I describe how to install Quagga, set up a loopback alias to hold the anycast IP address and configure Quagga to talk to your local routers. Then, I go over a few optional configuration extras.

First, install Quagga. For example, on Red Hat Enterprise Linux (RHEL), run `yum install quagga`. Substitute the appropriate package-management command for your distribution, as needed.

Next, create a loopback interface alias on the system. Configure the anycast IP address on this loopback interface. Using a loopback interface alias instead of a physical interface alias allows you to do a number of cool things. You could segment your service traffic from your administrative traffic. You could add some redundancy by responding to the anycast address on two physical interfaces, each attached to a different router or switch (although I won't go into that kind of configuration here). You also could take down the anycast interface (and, therefore, remove that interface from the anycast scheme) without affecting your ability to administer the system remotely. On RHEL, the interface configuration files are located in /etc/sysconfig/networking-scripts/. Create a file in that directory named ifcfg-lo:0 with the following contents:

```
# cat /etc/sysconfig/networking-scripts/ifcfg-lo:0
DEVICE=lo:0
IPADDR=10.0.0.1
NETMASK=255.255.255.255
BOOTPROTO=none
ONBOOT=yes
```

That file's format is fairly self-explanatory. You can control the lo:0 interface with your normal interface control commands (ifup, ifdown, ifconfig and so on).

Some versions of Fedora use NetworkManager to control eth0 by default. This may cause strange things to happen when you try to bring up a loopback alias. If that happens to you, add the line `NM_CONTROLLED=no` to /etc/sysconfig/networking-scripts/ifcfg-eth0, and restart your network. At this point, you should be able to bring up your new interface with `ifup lo:0`.

Now, you need to configure Quagga. By default, the Quagga configuration files are in /etc/quagga and /etc/sysconfig/quagga. There are a number of example configuration files in /etc/quagga: one for each routing protocol that Quagga supports; one for zebra, the main process; and one for the vtysh configuration. We primarily are interested in the ospfd.config and zebra.config files. The syntax in those files is similar to the standard Cisco configuration syntax, but with important differences. Also note that, by default, all routing processes bind to a dæmon-specific port on 127.0.0.1. If you configure a password for that routing process and Telnet to that port, you can monitor and configure the process on the fly using the same Cisco-like syntax. In these files, ! is the comment character:

```
# cat zebra.conf
hostname Endpoint1
!
interface eth0
  ip address 10.0.1.2/24
!
interface lo:0
  ip address 10.0.0.1/32
```

The above file is pretty quick and easy. It contains the IP addresses and netmasks of the physical adapters and the loopback adapter that has the anycast address. This file is much more complex:

```
# cat ospfd.conf
hostname Endpoint1
!
interface eth0
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 foobar
  ip ospf priority 0
!
router ospf
  log-adjacency-changes
  ospf router-id 10.0.1.2
  area 10.0.1.2 authentication message-digest
  area 10.0.1.2 nssa
  network 10.0.1.0/24 area 10.0.1.2
  redistribute connected metric-type 1
  distribute-list ANYCAST out connected
!
access-list ANYCAST permit 10.0.0.1/32
```

Let's go over the above section by section, starting with the following:

```
interface eth0
  ip ospf authentication message-digest
  ip ospf message-digest-key 1 md5 foobar
```

The first thing in the file is the OSPF MD5 authentication configuration. Always configure MD5 authentication on your OSPF sessions. Replace `foobar` with the appropriate key for your environment.

Next, we have:

```
ip ospf priority 0
```

Also set the OSPF priority to 0, which prevents the server from being elected as the Designated Router on that link.

Next come the router configuration directives:

```
router ospf
  log-adjacency-changes
```

log-adjacency-changes is a great configuration directive that gives you more details when there is a change in neighbor relationships between your server and any other OSPF-speaking device.

Then:

```
ospf router-id 10.0.1.2
```

Here the router ID is set to the server's service address. Router IDs must be unique within the routing domain.

We then configure this server to be in its own Not So Stubby Area (NSSA):

```
area 10.0.1.2 authentication message-digest
area 10.0.1.2 nssa
redistribute connected metric-type 1
distribute-list 5 out connected
```

NSSA areas are a form of stub area that limits the routes sent into the area to summary routes, but still allows external routes to come from that area. We need to allow external routes because we advertise our anycast IP address by redistributing our connected interfaces and running that through a distribute list to confine our advertised interfaces to just the anycast IP address. However, we don't want this server to have to deal with all the routes in area 0.0.0.0.

The following statement selects the interfaces that will participate in OSPF:

```
network 10.0.1.0/24 area 10.0.1.2
```

We want our eth0 interface to participate in OSPF, so we specify 10.0.1.0/24, and we put those interfaces in area 10.0.1.1.

The following line defines the access list that will allow route advertisements out:

```
access-list ANYCAST permit 10.0.0.1/32
```


Advertiser Index

CHECK OUT OUR BUYER'S GUIDE ON-LINE.
Go to www.linuxjournal.com/buyersguide where you can learn more about our advertisers or link directly to their Web sites.

Thank you as always for supporting our advertisers by buying their products!

Advertiser	Page #	Advertiser	Page #
1&1 Internet, Inc. www.oneandone.com	1	Microway, Inc. www.microway.com	C4, 5
Aberdeen, LLC www.aberdeeninc.com	19	Mikro Tik www.routerboard.com	7
ApacheCon www.us.apachecon.com/us2008	23	Parallels, Inc. www.parallels.com	17, 79
Archie McPhee www.mcphee.com	79	Polywell Computers, Inc. www.polywell.com	13, 79
ASA Computers, Inc. www.asacomputers.com	31	Rackspace Managed Hosting www.rackspace.com	C3
Cari.net www.cari.net	43	Saint Arnold Brewing Company www.saintarnold.com	78
Digi-Key Corporation www.digi-key.com	78	SEE 2009 SEE2009.org	67
Emac, Inc. www.emacinc.com	73	ServerBeach www.serverbeach.com	37
EmperorLinux www.emperorlinux.com	47	Servers Direct www.serversdirect.com	9
Florida Linux Show www.floridalinuxshow.com	53	Silicon Mechanics www.siliconmechanics.com	21, 25
FSCONS fscons.org	59	StrayTats www.straytats.com	78
Gecad Technologies/Axigen www.axigen.com	79	SuperComputing SC09 sc09.supercomputing.org	75
Genstor Systems, Inc. www.genstor.com	63	Technologic Systems www.embeddedx86.com	33
Gutsy Geeks www.gutsygeeks.com	77	Ubiquiti Networks, Inc. www.ubnt.com	C2
iXsystems, Inc. www.ixsystems.com	3	USENIX LISA www.usenix.org/events/usenix07	39
Logic Supply, Inc. www.logicsupply.com	71	Utilikilts www.utilikilts.com	78
Lullabot www.lullabot.com	27, 61		

ATTENTION ADVERTISERS

February 2010 Issue #190 Deadlines
Space Close: November 19; **Material Close:** December 1

Theme: Desktop

BONUS DISTRIBUTIONS:
Black Hat DC, Southern California Linux Expo (SCALE), SxSW, Cloud Connect, USENIX FAST

Call **Joseph Krack** to reserve your space
+1-713-344-1956 ext. 118, e-mail joseph@linuxjournal.com

Now that Quagga is configured, we need to open up the proper IP protocol number on our firewall. OSPF uses protocol number 89. The details of opening that particular protocol number will vary significantly with the firewall configuration you're using.

In general, you'll use a command like this:

```
# iptables -I INPUT -p 89 -j ALLOW
```

which inserts the rule permitting IP protocol 89 at the start of the INPUT chain. That command will work with most any standard firewall configuration. After all of this, you finally can get Quagga going. Start it with `service zebra start` and `service ospfd start`. Your system now should be participating in your OSPF routing scheme.

You can confirm that with a quick look at your router's routing table:

```
R1>show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 21, type NSSA extern 1
  Last update from 10.0.1.2 on FastEthernet0/0, 00:00:14 ago
  Routing Descriptor Blocks:
  * 10.0.1.2, from 10.0.1.2, 00:00:14 ago, via FastEthernet0/0
      Route metric is 21, traffic share count is 1
```

Optional Quagga Configuration Extras

To enable remote administration, you must set a password in ospfd.conf as follows:

```
password YOUR-PASSWORD
enable password YOUR-ENABLE-PASSWORD
```

If you are feeling paranoid about your server establishing neighbor relationships with devices other than your router, you can disable OSPF automatic neighbor discovery on your server with the following additional commands in ospfd.conf:

```
interface eth0
  ip ospf network non-broadcast

router ospf
  neighbor ROUTER-ID-OF-ROUTER
```

This configuration has each endpoint in its own OSPF NSSA area. You just as easily could have the endpoints become part of whatever area is already in existence, as long as that area allows external routes. Having each server in its own area gives you a little more control over what kind of routes propagate to and from each endpoint. It is a bit more work, both initially and when you move a server to a different router. It also means your servers have to be able to connect directly to an ABR with access to area 0, which may or may not be possible in your network.

Sample Anycast Layout

Anycast with one endpoint is fairly useless, so let's take a look at a simple deployment scenario. Each endpoint is configured exactly like the endpoint we just configured, with the exception of the service address and the OSPF area number.

In this scenario, let's say we have anycast running between

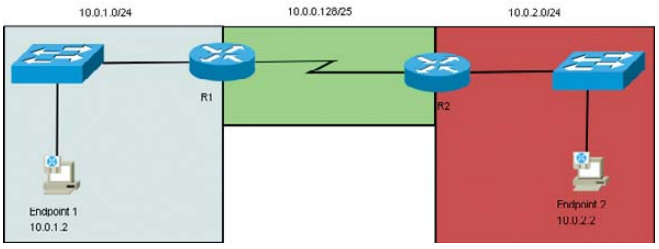


Figure 1. Two-Site, Two-Server Environment

two sites (for instance, a main office and a satellite office) connected over a WAN. There is one anycast endpoint at each site. The main office is 10.0.1.0/24, the satellite office is 10.0.2.0/24, and our anycast address is 10.0.0.1, from our anycast subnet, 10.0.0.0/25 (Figure 1).

OSPF on R1 is configured as follows:

```
router ospf 1
 log-adjacency-changes
 network 10.0.1.0 0.0.0.255 area 10.0.1.2
 network 10.0.0.128 0.0.0.128 area 0.0.0.0
 area 10.0.1.2 nssa no-summary default-information-originate
 area 10.0.1.2 authentication message-digest
 area 0.0.0.0 authentication message-digest
```

OSPF on R2 is configured as follows:

```
router ospf 1
 log-adjacency-changes
 network 10.0.2.0 0.0.0.255 area 10.0.2.2
 network 10.0.0.128 0.0.0.128 area 0.0.0.0
 area 10.0.2.2 nssa no-summary default-information-originate
 area 10.0.2.2 authentication message-digest
 area 0.0.0.0 authentication message-digest

R1>show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 21, type NSSA extern 1
  Last update from 10.0.1.2 on FastEthernet0/0, 00:00:14 ago
  Routing Descriptor Blocks:
  * 10.0.1.2, from 10.0.1.2, 00:00:14 ago, via FastEthernet0/0
      Route metric is 21, traffic share count is

R2>show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 21, type NSSA extern 1
  Last update from 10.0.2.2 on FastEthernet0/0, 00:05:07 ago
  Routing Descriptor Blocks:
  * 10.0.2.2, from 10.0.2.2, 00:05:07 ago, via FastEthernet0/0
      Route metric is 21, traffic share count is 1
```

Traffic from each of the sites is flowing to the local anycast endpoint. Here's what happens if we take out the main office endpoint:

```
Endpoint1# ifdown lo:0
Endpoint1#
```

We're meeting at The Symbian Exchange & Exposition – **SEE 2009**

- to share
- to create
- to network
- to learn
- and to have lots of fun!

..... it's open & it's FREE .....

Register now!
SEE2009.org
Booking code: CU1266 ADLJ

Headliners include:

Jimmy Wales
Founder, Wikipedia

Geoffrey Moore
Author, Crossing the Chasm

Feel free to join in.

Exchange – We're holding an ideas Exchange on each day of SEE 2009. Community members get together to talk, promote their ideas and drive development together through conversation.

Visionaries – **SEE 2009** is an informal 'unconference' in which everyone participates, but there'll also be talks from leading minds and industry experts on the main stage.

Dialogue – **SEE 2009** is the forum for people to talk – in the exhibition hall – on the online networking site – at the event lounges and during three separate discussions offering in-depth workshops:

- Application development
- Marketing & Business models
- Device Creation & Contributing to Symbian

Where?

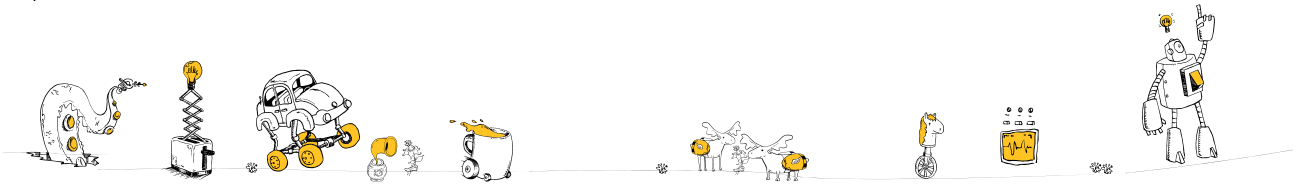
Earls Court 2
Warwick Road,
London
SW5 9TA

When?

27-28 October 2009

How?

Register: SEE2009.org



Application/Router Configuration Notes

1. Adjusting the cost of a link can be a great way to prepare an endpoint for removal gracefully. Using any other method, especially in a high-traffic environment, can result in dropped connections and other transient issues until OSPF reconverges. Setting the link cost very high before removal, on the other hand, avoids any transient problems during the brief reconvergence period. Once the endpoint in question is no longer receiving traffic, you can disable the anycast loopback and do whatever work needs to be done. Adjust the cost of a link on the router connected to your server with the following commands (in the example above that would be R1 or R2):

```
interface WHATEVER-INTERFACE-CONNECTS-THE-ROUTER-TO-QUAGGA
  ip ospf cost NUMBER
```

Replace {number} with some large number that is greater than the cost of the replacement anycast endpoint.

2. Make sure nonresponse traffic is not sourced from the anycast address. One example is in configuring DNS. You want DNS replies to come from the anycast IP address, but you do not want DNS zone transfers to come from or go to anycast IP addresses. In the case of a caching nameserver, you also don't want recursive queries originated from the server to be sourced from the anycast address.

3. Applications that maintain state in some way are not good candidates for anycast addressing, even if the underlying transport protocol is stateless. The exception to that rule would be if all the anycast endpoints got their application-level state information from the same place.

4. UDP is the de facto standard for the anycast transport-layer protocol. Use any other transport-layer protocol at your own risk. See Resources for a detailed review of issues associated with using other transport-layer protocols.

```
R1>show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 85, type extern 1
  Last update from 10.0.0.130 on Serial0/0, 00:00:21 ago
  Routing Descriptor Blocks:
  * 10.0.0.130, from 10.0.2.2, 00:00:21 ago, via Serial0/0
      Route metric is 85, traffic share count is 1

R2>show ip route 10.0.0.1
Routing entry for 10.0.0.1/32
  Known via "ospf 1", distance 110, metric 21, type NSSA extern 1
  Last update from 10.0.2.2 on FastEthernet0/0, 00:05:07 ago
  Routing Descriptor Blocks:
  * 10.0.2.2, from 10.0.2.2, 00:05:07 ago, via FastEthernet0/0
      Route metric is 21, traffic share count is 1
```

All traffic starts to flow to the remaining endpoint, as designed and desired.

Monitoring and Automatic Route Withdrawal

As I mentioned previously, the fact that a host is up does not mean that the service that host provides is up. When a host running Quagga goes down, any routes that host inserted into OSPF will be withdrawn. We need to do the same thing when a service does down. Any piece of monitoring software that can run a handler script in response to a monitoring event can be used for this task. The basic idea is to execute a test against the anycast IP from each anycast endpoint. If a test fails, you need to run `ifdown lo:0` on the failed endpoint. Quagga will detect the downed interface and withdraw the route to that interface from OSPF. Administrators then can fix the box at their leisure and place the box back into service with a simple `ifup lo`.

Conclusion

Anycast is a great technique to enhance the reliability and fault tolerance of applications and services on your network. When designing your anycast topology, keep several rules and guidelines in mind. I've shown a very basic use case and deployment of anycast here. You can take the same concepts covered in this article, along with a fair bit of networking knowledge, and scale them to a worldwide deployment. If you do it right, you can have redundancy without nearly as many idle machines sitting around.■

Philip Martin has been working and playing with Linux for about ten years and is currently a Systems Engineer for a large on-line retailer. When he is not working with computers, he spends his days trying to be more like Alton Brown and in an ongoing quest to get invited to an *Iron Chef America* filming. He can be reached at phillip.martin@gmail.com.

Resources

root-servers.org: **www.root-servers.org**

OpenBGPD: **www.openbgpd.org**

GNU Zebra: **www.zebra.org**

"IP Routing Primer, Part One": **www.networkcomputing.com/netdesign/1122ipr.html**

"Cisco administration 101: What you need to know about OSPF": **articles.techrepublic.com.com/5100-10878_11-6132046.html**

"Open Shortest Path First (OSPF)": **www.cisco.com/en/US/docs/internetworking/technology/handbook/OSPF.html**

"Architectural Considerations of IP Anycast": **tools.ietf.org/html/draft-mcpherson-anycast-arch-implications-00**

Celebrating 15 years of *Linux Journal*,
we've brought together every article
ever published in the world's #1
Linux magazine and packaged it
in one convenient CD.

With nearly 4,000 articles written by industry experts on everything from cool projects, desktop how-tos, security, embedded systems, networking, virtualization, multimedia, system administration and programming tricks and techniques—this unique collection is a must-have for every Linux enthusiast.

Get your NEW *Linux Journal* Archive CD today featuring all issues from 1994 through 2008. Just $34.95.

www.linuxjournal.com/archivecd

Host Identity Protocol for Linux

Have you ever wondered why your multimedia streams stop working after you switch to a different network with your laptop? Have you thought about why setting up a server on your home network behind a NAT is so awkward or even impossible? Host Identity Protocol for Linux (HIPL) offers a remedy to these and other problems.

ABHINAV PATHAK, MIIKA KOMU and ANDREI GURTOV

An IP address determines the name and network location of a computer on the Internet. The network stack reuses this IP address at all layers, including the application layer. As a consequence, existing network connections break when an IP address changes. For example, suppose you are streaming a video from your favorite Web site and you switch from a WLAN to LAN connection. Then, your host's IP address changes and breaks the stream. This happens because the video-streaming application and the host use different IP addresses. Even though the host uses the new IP address, the application still uses the old address.

What creates this problem in the current Internet architecture? The IP address specifies both the name and the topological location of a host on the Internet. Here's an analogy: a person named Abhinav Pathak who lives in New Delhi should still be called Abhinav Pathak when he is visiting London. As simple as it may seem, this analogy currently does not work on today's Internet.

How HIP Solves the Problem

Host Identity Protocol (HIP) assigns a permanent, location-independent name to a host. HIP names a host essentially using a public key, which is referred to as Host Identity in HIP literature. As public keys are quite long, usually it is more convenient to use a 128-bit fingerprint of the HI, which is called the Host Identity Tag (HIT). The HIT resembles an IPv6 address, and it gives the host a permanent name. The Internet Assigned Numbers Authority (IANA) allocated an IPv6 address prefix for HITs (2001:0010::/28).

The HIT is similar to an SSH fingerprint, but unlike SSH, it can be used by all applications. HIP also supports IPv4-compatible names called Local Scope Identifiers (LSIs). HITs in HIP are statistically unique and inherently secure because they are derived from public keys and, therefore, are difficult to forge.

In HIP, sockets in transport protocols, such as TCP, are bound to HITs rather than IP addresses. The networking stack translates the HITs to IP addresses before packet transmission on the wire. The reverse occurs on the host receiving HIP packets. When the host changes the network, the networking stack changes the IP address for the translation. The application doesn't notice the change because it is using the constant HIT. This is how HIP solves the problem of host mobility.

HIIT has developed an implementation of HIP for Linux (HIPL), which is available from the InfraHIP II Web pages. In this article, we describe how you can benefit from HIP and explain how to install and run HIP on your Linux system.

HIP Applications

Linux is ported to many platforms and devices, such as laptops, smartphones and PDAs. These devices are mobile but usually lack mobility support from the networking stack. Many networking applications on Linux don't provide communications privacy either. HIP solves both of these problems and also provides support for multihomed hosts. Here, we describe four practical problems that HIP solves.

1. Access Control

Access to host services usually is constrained using IP addresses. For example, consider the access control files for Linux. The hosts.allow and hosts.deny files contain the service names and hostnames (or IP addresses) of the hosts that are allowed to access certain services.

Suppose a server grants permission to a particular client to access its remote services, such as SSH, FTP and so on. It specifies its hosts files as follows:

```
$ cat /etc/hosts.deny
ALL: ALL

$ cat /etc/hosts.allow
ALL: 10.0.0.10
```

This states that only a client with an IP address of 10.0.0.10 is allowed to access services running on this host. All other IPs are blocked.

Now, what happens when the client with IP 10.0.0.10 moves to a new network and its IP address changes? Or, what happens if its DHCP lease time expires and it is granted another IP address? In such cases, the client would no longer be able to access the server. Either it has to regain its IP address or the server has to update its hosts.allow and hosts.deny files.

HIP easily solves this problem. The server's /etc/hosts.allow file contains the HIT of the client instead of the IP address. The client has the same HIT independent of its IP address and, hence, its network location. The entry in the /etc/hosts.allow file looks like this with HIP:

```
$ cat /etc/hosts.allow
ALL: [2001:15:e156:8a78:3226:dbaa:f2ff:ed06]
```

This shows that the client with the HIT (that is, name) 2001:15:e156:8a78:3226:dbaa:f2ff:ed06 is allowed to access

the services on the server.

The HIP software running on the server uses public keys to authenticate the client before the client can use a particular service. A crucial part of the authentication is for the server to check that the client's HIT (fingerprint) matches the public key. This way, the server can cryptographically verify that the client is the one it claims to be.

2. Security—Authentication and Encryption

HIP authenticates and secures communication between two hosts. HIP authenticates hosts and establishes a symmetric key between them to secure the data communication. The data flow between the end hosts is encrypted by IPsec Encapsulating Security Payload (ESP) with the symmetric key set up by HIP. HIP introduces mechanisms, such as cryptographic puzzles, that protect HIP responders (servers) against DoS attacks. Applications simply need to use HITs instead of IP addresses. Application source code does not need to be modified.

3. Mobility

HIP provides transparent mobility support for existing network applications. TCP connections are bound to HITs instead of IP addresses. HITs do not change for a given host. HITs are further mapped to IP addresses. When an IP address changes, new mappings between the HIT and the new IP address are formed. When a host moves to a new network and obtains a new IP address, the host informs its peers about its new IP address, and TCP connections are sustained.

4. NAT Traversal

WLAN access points and broadband modems employ NATs due to the lack of IPv4 addresses. However, you have to configure your NAT settings manually if you want to use P2P software or connect to your computer behind a NAT. It may even be impossible if your ISP employs a second NAT.

With HIP, hosts can address each other with HITs across private address realms of NATs. HIP makes use of two alternative NAT traversal technologies, ICE and Teredo, to traverse the NATs. Setting up a server behind a NAT using HIP does not require manual configuration of the NAT. The HIPL on-line manual **infrahip.hiit.fi/hipl/manual/ch21.html** describes the details.

Name Lookup Support

The InfraHIP site offers free services for the HIP community. For example, you can register your HIT to the DNS or Distributed Hash Table (DHT). The site also offers free HIP forwarding services to assist in NAT traversal and locating mobile nodes.

How HIP Works

The Host Identity Protocol architecture (Figure 1) defines a new namespace, the Host Identity namespace, which decouples the name and locator roles of IP addresses. With HIP, the transport layer operates on host identities instead of IP addresses as endpoint names. The host identity layer is between the transport

layer and the network layer. The responsibility of the new layer is to translate identities to routable locators before a host transmits the packet. The reverse applies to incoming packets.

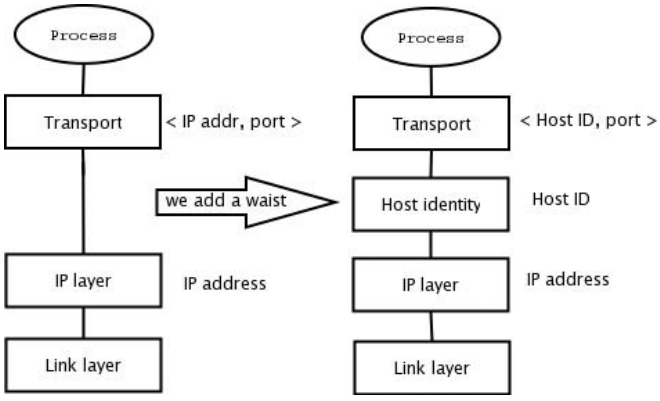


Figure 1. The Host Identity layer is located between the transport and network layers.

Protocol Overview

The actual Host Identity Protocol (HIP) is composed of a two round-trip, end-to-end Diffie-Hellman key-exchange protocol, called base exchange, mobility updates and some additional messages. The networking stack triggers the base exchange automatically when an application tries to connect to an HIT.

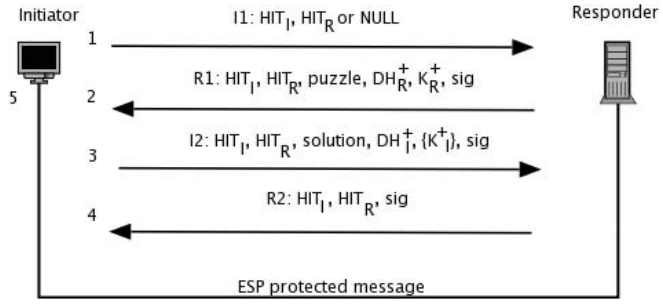


Figure 2. HIP Base Exchange

During a base exchange, a client (initiator) and a server (responder) authenticate each other with their public keys and create symmetric encryption keys for IPsec to encrypt the application's traffic. In addition, the initiator must solve a computational puzzle. The responder selects the difficulty of the puzzle according to its load. When the responder is busy or under DoS attack, the responder can increase the puzzle difficulty level to delay new connections.

We can describe this process as follows:

```
I --> DNS: lookup R
I <-- DNS: return R's address and HI/HIT
```

The initiator application connects to an HIT:

```
I1   I --> R (Hi, Here is my I1, let's talk with HIP)
R1   R --> I (OK, Here is my R1, solve this HIP puzzle)
I2   I --> R (Computing, here is my counter I2)
R2   R --> I (OK. Let's finish base exchange with my R2)

I --> R (ESP protected data)
R --> I (ESP protected data)
```

Mobility and Rendezvous

HIP provides a mechanism similar to base exchange to handle IP address changes. When a host detects a new IP address, it informs all its peers of the address change. The hosts adjust their IPsec security associations accordingly, and the applications running on the hosts continue sending data to each other as if nothing happened.

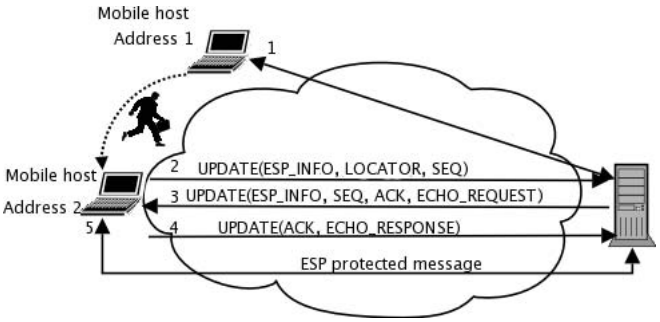


Figure 3. HIP Mobility Updates

When two hosts are connected to each other using HIP and one of them moves, the mobile host tells its current location to the other. If both hosts move at the same time, they can lose contact with each other. In this case, an HIP rendezvous server assists the hosts. The rendezvous server has a fixed IP address and, therefore, it offers a stable contact point for mobile hosts. The rendezvous server relays only the first packet, and after the contact, the hosts can communicate with each other directly. HIP includes another similar service, called HIP Relay, that forwards all HIP packets to support NAT traversal.

How to Install and Use HIPL

The HIPL software bundle consists of the following main components:

- HIPD (HIP Dæmon): HIP control, IPsec key and mobility management software.
- HIPFW (HIP firewall utility dæmon): supports HIP packet filtering to enable public key-based access control and LSI implementation. It also provides userspace IPsec support for legacy hosts running kernel versions below 2.6.27.
- DNS Proxy for HIP: translates hostname queries to DNS to HITs to applications when an HIT can be found.

Installation

You can install HIPL from the precompiled binaries or source code.

To install HIPL on Ubuntu Jaunty, add a new file, /etc/apt/sources.list.d/hipl.list, with the following contents:

```
deb http://packages.infrahip.net/ubuntu jaunty main

$ apt-get update
$ apt-get install hipl-all
```

For Fedora 9 and above, first make sure that SELinux configuration is disabled in /etc/selinux/config, and reboot your machine:

```
SELINUX=disabled
```

Next, add a new file /etc/yum.repos.d/hipl.repo:

```
[hipl]
name=HIPL
baseurl=http://packages.infrahip.net/fedora/base/$releasever/$basearch
gpgcheck=0
enabled=1
```

Then, run:

```
yum install hipl-all
```

For details on HIPL installation for other distributions, see **infrahip.hiit.fi/index.php?index=download**.

Alternatively, you can compile the HIPL software bundle manually from the sources. To do so, first download and extract the HIPL software bundle from **infrahip.hiit.fi/hipl/hipl.tar.gz**. Run `autogen.sh --help` to list the library and header dependencies. After you have installed the missing dependencies, you can compile the software by running the script without any arguments. To complete the manual installation, run `make install`.

The default installation encapsulates all HIP and IPsec traffic over UDP to support client-side NAT traversal. At minimum, you need to allow UDP port number 50500 in both directions for IPv4. The HIPL manual describes this in more detail at **infrahip.hiit.fi/hipl/manual/ch02.html**.

Once installation has been completed, you should start the HIP dæmon as follows:

```
$ sudo hipd
```

When you start the hipd the first time, it generates its configuration files and identities in the /etc/hip/ directory. Your identity is visible as an IPv6 address on the dummy0 device. To see your host's identity, run the following:

```
$ ifconfig dummy0
## OR
$ ip addr show dev dummy0
```

Correspondingly, your IPv4-based "alias" for the HIT is listed on the dummy0:1 interface.

To perform name lookups for other hosts, you also have to start the HIP DNS proxy as follows:

```
$ sudo hipdnsproxy
```

Testing HIP with Firefox

HIP can be used with many applications and protocols, including FTP, SSH, VLC, LDAP, sendmail, Pidgin and VNC. However, the easiest way to validate your HIPL software installation is to start Firefox and connect to the Web server located at **crossroads.infrahip.net**. The Web server is running HIP and displays whether HIP was used for the connection. You optionally can install a Firefox add-on (**https://addons.mozilla.org/en-US/firefox/addon/10551**), if you prefer a client-side indicator for HIP.

Streaming Multimedia and Testing Mobility with VLC

Now, let's stream some video with VLC and then try mobility. The example in this section assumes you have two computers with HIPL installations. We also assume that the computers are running in the same LAN with DHCP services. In this example, the two computers connect to LAN using the eth0 device.

First, display an HIT for the first host, and start VLC client on one computer:

```
client$ hipconf get hi default     # HIT_OF_CLIENT
client$ vlc -vvv 'rtp://@[HIT_OF_CLIENT]:50004'
```

Then, start the VLC server on the second host:

```
server$ vlc -vvv SOMEFILE.avi \
            --sout '#rtp{mux=ts,dst=[HIT_OF_CLIENT]}'
```

The string `HIT_OF_CLIENT` should not be taken literally. Instead, you can discover it from the output of the hipconf command at the client. The brackets around the HIT are mandatory for VLC to distinguish IPv4 addresses from IPv6.

Because the video stream is established directly to an HIT, the connection is guaranteed to use HIP; otherwise, the stream just fails. In this case, we did not use a hostname, and the server learns the client's IP address by broadcasting the first HIP packet to the LAN. The use of hostnames also is possible, and the HIPL software bundle publishes your hostname on InfraHIP's free name lookup servers by default.

Finally, let's test mobility. Type the following on the command line to obtain a new IP address from your network:

```
$ sudo dhclient eth0
```

You may see a small glitch during the dhclient run caused by a short disconnectivity period from the network. If you also have wireless connectivity, feel free to experiment with handovers from the wired network to wireless and vice versa.

Resources

HIP Architecture RFC: **www.rfc-editor.org/rfc/rfc4423.txt**

HIP Base RFC: **www.rfc-editor.org/rfc/rfc5201.txt**

InfraHIP Project: **infrahip.hiit.fi**

Freshmeat Page for HIPL: **freshmeat.net/projects/hipl/?branch_id=64825&release_id=228615**

Host Identity Protocol (HIP): Towards the Secure Mobile Internet by Andrei Gurtov, Wiley, June 2008

M. Komu, S. Tarkoma, J. Kangasharju and A. Gurtov, "Applying a Cryptographic Namespace to Applications", in Proc. of First International ACM Workshop on Dynamic Interconnection of Networks, September 2005: **www.niksula.cs.hut.fi/~mkomu/docs/f17-komu.pdf**

OpenHIP: **www.openhip.org**

HIP for inter.net Project: **hip4inter.net**

IETF: **www.ietf.org**

Miredo: **www.remlab.net/miredo/**

Teredo: **technet.microsoft.com/en-us/library/bb457011.aspx**

ICE: **tools.ietf.org/html/draft-ietf-hip-nat-traversal**

The HIPL Community

HIPL is open-source software for Linux. We are actively improving the software according to feedback from user mailing lists (**www.freelists.org/list/hipl-users**). We welcome all Linux enthusiasts to the HIPL community, and we are looking for more users and developers.

Conclusion

Host Identity Protocol brings communications privacy and mobility support for existing applications by introducing a new cryptographic namespace. It also allows you to set up servers behind NATs easily. In this article, we discussed how HIP works and how you can install it on your Linux box. We have shown how you can use HIP with Firefox and how to stream video with VLC successfully during network IP address change.■

Abhinav Pathak (pathaka@purdue.edu) is a PhD student at Purdue University. He completed his Bachelor's degree in computer science from IIT Kanpur. He worked as a research assistant at HIIT.

Miika Komu (miika@iki.fi) is a researcher at HIIT. He does HIP standardization and is one of the developers for the InfraHIP Project. He also practices martial arts at a Takado club.

Andrei Gurtov (gurtov@hiit.fi) is a principal scientist and group leader at HIIT. He received his PhD degree from the University of Helsinki in 2004. He has written a book on HIP as well as more than 70 other publications. Andrei is a fan of sailing.

Statement of Ownership, Management, and Circulation

1. Publication Title: *Linux Journal*
2. Publication Number: 1075-3583
3. Filing Date: September 10, 2009
4. Issue Frequency: Monthly
5. Number of Issues Published Annually: 12
6. Annual Subscription Price: $29.50
7. Complete Mailing Address of Known Office of Publication: PO Box 980985 Houston, TX 77098 Contact Person: Mark Irgang 713-344-1956 x111
8. Complete Mailing Address of Headquarters of General Business Office of Publisher: PO Box 980985 Houston, TX 77098
9. Full Names and Complete Addresses of Publisher, Editor, and Managing Editor: *Publisher.* Carlie Fairchild PO Box 980985 Houston, TX 77098 *Editor.* Doc Searls PO Box 980985 Houston, TX 77098 *Managing Editor.* Jill Franklin PO Box 980985 Houston, TX 77098
10. Owner(s): Carlie Fairchild PO Box 980985 Houston, TX 77098 Joyce Searls PO Box 980985 Houston, TX 77098 Adele Soffa PO Box 980985 Houston, TX 77098
11. Known Bondholders, Mortagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortages, or Other Securities: None
12. Tax Status: Has not Changed During Preceding 12 Months

13. Publication Title: *Linux Journal* — 14. Issue Date: October 2009

15. Extent and Nature of Circulation	Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies: (Net press run)	47,116	47,861
b. Paid and/or Requsted Circulation		
(1) Paid/Requested Outside-County Mail Subscriptions on Form 3541.	17,275	16,430
(2) Paid In-County Subscriptions Stated on Form 3541	0	0
(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	27,007	28,005
c. Total Paid and/or Requested Circulation	44,282	44,435
d. Free Distribution Outside the Mail		
(1) Outside-County as Stated on Form 3541	651	668
(2) In-County as Stated on Form 3541	0	0
(3) Other Classes Mailed Through the USPS	0	0
(4) Free Distribution Outside the Mail	2,183	2,758
e. Total Free Distribution	2,834	3,426
f. Total Distribution	47,116	47,861
g. Copies Not Distributed	10,175	8,910
h. Total	36,941	38,951
i. Percent Paid and/or Requested Circulation	94%	93%

PS Form 3526

SC Portland, OR 2009

Computing For A Changing World.

November 14-20, 2009
Oregon Convention Center
Portland, Oregon
http://SC09.supercomputing.org

Sponsors:
ACM SIGARCH/
IEEE Computer Society

Association for Computing Machinery

IEEE computer society

Conference Dates
November 14-20, 2009

Exhibition Dates
November 16-19, 2009

The International Conference for High Performance Computing, Networking, Storage, and Analysis

KYLE RANKIN

Ext3 vs. XFS

After focusing on the Web the past couple months, for this round, Kyle and Bill turn to something a little more local—Linux filesystems.

BILL CHILDERS

Unlike the Twitter-fest that was the previous column, this time Bill will be able to talk in full sentences—even if they are more than 140 characters! Stay tuned for Kyle's love of (and Bill's angst toward) XFS.

KYLE: Okay, I admit I use ext3 in plenty of systems. It's an all-around good filesystem, but when I need high performance, especially with large files, I always turn to XFS.

BILL: That's fine and dandy when you're running Debian or Ubuntu, but what about a Red Hat box? And, no, CentOS, although cool, doesn't count here. I'm talking the actual Crimson Fedora here—Red Hat. It doesn't support XFS in its kernel at all, does it? So, you're forced to build your own stuff, and then you're in for an admin headache.

KYLE: Well, you already ruled out CentOS, but if you are stuck in a Red Hat-like environment and want high performance, like XFS, you might have to stray from the list of supported packages and either use CentOS or a custom repository. In any case, it wouldn't be the first time an admin had to make up for the limited set of supported packages in Red Hat.

BILL: Pop quiz, hotshot: does that break you out of the support matrix for Red Hat?

KYLE: My understanding is that you would pop out of support only for problems that are directly related to the filesystem or kernel. I'll be honest though, in all the years I've had Red Hat support, I can't think of a time I legitimately needed it. I have, however, had plenty of situations where a developer wanted to use the filesystem like a database and store millions of files in huge nested directories—something XFS handles quite well.

BILL: Shhh. You'll anger a possible advertiser. Didn't we engage them once on a dm-mapper or ocfs issue or something? But yeah, that's a tangential thing—just because you don't use the support doesn't mean you don't need it. There's a reason you continue to pay that fee. And with respect to developers using the filesystem as a database, we've both seen that. XFS helps here, but redesigning your whole filesystem isn't necessarily a fix for poor software architectural design. (I used the architectural word in there—bonus points for me!)

KYLE: Yeah, yeah, put away your drafting table Mr Architect. We both know how rarely a sysadmin can dictate how a developer solves a problem. In any case, there isn't much of a redesign. In a Red Hat-based system, it's a matter of a different kernel package (included with CentOS) and reformatting a filesystem. With a Debian-based system, the ability is already there. In any case, you don't have to format everything with XFS, only the filesystems where it would benefit.

Speaking of ability, what about the ability to recover a system when things go pear-shaped?

BILL: True. Speaking of ability, what about the ability to recover a system when things go pear-shaped? I've never had any luck *fixing* an XFS filesystem—they've always gobbled my data. Ext3 may be slow, but I've always been able to save something off a damaged filesystem.

KYLE: See, I've had the exact opposite experience. The one big advantage in my mind to XFS is how well the recovery tools work. I've lost data on basically every major filesystem out there from ReiserFS (let's not go there) to ext2 and ext3 and yes, even XFS, but whenever I've needed to do a recovery, the XFS recovery tools always have been successful, even when the problem was related to a bad hard drive controller.

BILL: Well, I may be biased, as I formatted my laptop with XFS at *your* behest some years ago and watched as a fun bug gobbled all my data. I never did recover that, if you recall.

KYLE: Like I said, I've had data gobbled by every filesystem. I'll also note that I never was really bitten by that bug, but I do remember a pretty nasty ext3 bug from a few years back that was so bad they actually labeled the kernel as defective after the fact. I've used XFS on my personal systems both for large file storage and even as the filesystem for my own /home directory on my personal laptop now, without issue. The fact is, I noticed a tangible difference on the speed of my system when I moved to XFS.

BILL: I know you have. You have a halo effect about you with things like that though. I have...the opposite effect. If it can break, I will break it. I've always been able to recover from an ext3 explosion. For me, it's not about the speed. The ext3 filesystem and tools are well known and have been shipped in *everything*. I know if my machine dies, I can move the disk to another Linux box and be able to read the filesystem, or I can use a "standard" recovery disk. Heck, there are even plugins for other OSes that can read ext3. Try that with XFS.

KYLE: My standard recovery disc always has been Knoppix [insert plug here], and it always has worked just fine with XFS filesystems. I'm not saying that XFS should be used for everything. There's a reason ext3 is the default filesystem for most distributions. After all, it offers good all-around performance for all kinds of filesystems. When you need high performance for terabytes of large files or millions of small files though, it's hard to beat XFS. Even formatting an XFS filesystem is substantially faster than ext3.

The fact is, I noticed a tangible difference on the speed of my system when I moved to XFS.

BILL: I'm not arguing that XFS isn't faster—it is, and by a large margin. I don't think it's as safe, however.

KYLE: I think these days the biggest risk on any system isn't from filesystem corruption. It's either from hard drive failure or from user error. In either case, if you are that worried, you should have a solid, tested backup system in place.

BILL: True, no filesystem or RAID is a substitute for a solid backup method. That's something we *do* agree on.

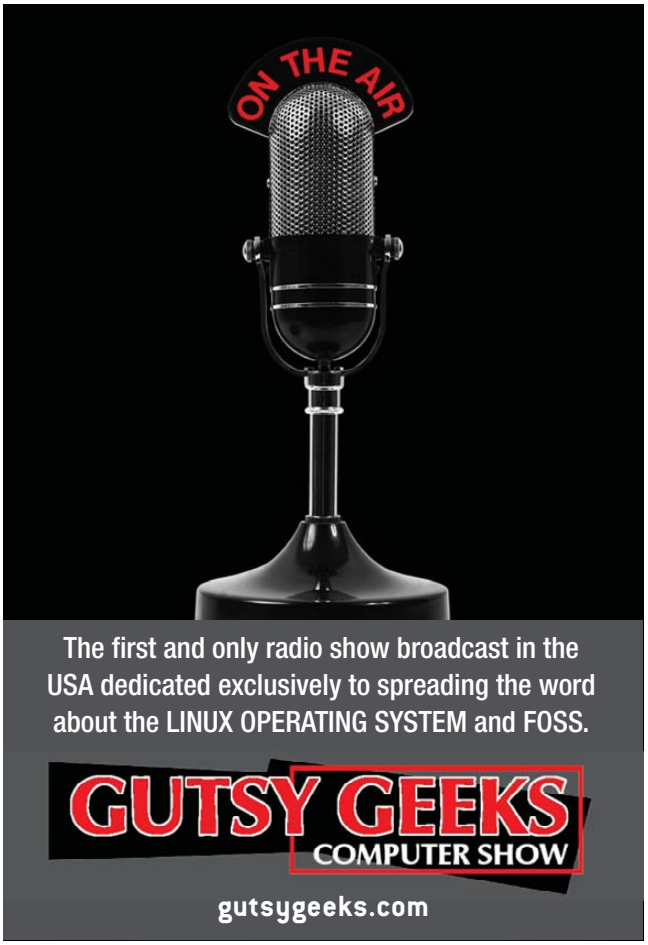
KYLE: The bottom line for me is that whenever I reach the limits of ext3, I know I have a solid, fast alternative in XFS. The XFS recovery tools are excellent, and in my experience, they work well. Plus, it's been available in Linux long enough to iron out any major bugs and is available in anything from CentOS to Debian to Ubuntu. When I want an ordinary filesystem, I choose ext3, but when I need speed, I choose XFS.

BILL: Except that XFS isn't available from Red Hat, and that's a considerable installed base. Just because something is faster doesn't necessarily mean it's the best long-term solution. XFS may be more capable, but the downside of possibly falling out of a supportable configuration at the enterprise level keeps me from deploying XFS on anything but nonessential gear.

Last minute note from Bill: Just as we were readying this for print, Red Hat announced that XFS will be in the Red Hat Enterprise 5.4 beta. My arguments will be sent to /dev/null after 5.4 releases.■

Kyle Rankin is a Senior Systems Administrator in the San Francisco Bay Area and the author of a number of books, including *Knoppix Hacks* and *Ubuntu Hacks* for O'Reilly Media. He is currently the president of the North Bay Linux Users' Group.

Bill Childers is an IT Manager in Silicon Valley, where he lives with his wife and two children. He enjoys Linux far too much, and he probably should get more sun from time to time. In his spare time, he does work with the Gilroy Garlic Festival, but he does not smell like garlic.



INNOVATION ON THE GO
ORDER YOUR BEAGLE BOARD FROM DIGIKEY.COM

AVAILABLE EXCLUSIVELY AT DIGI-KEY
beagleboard

only $149.00

LOW-COST, NO FAN, SINGLE-BOARD COMPUTER

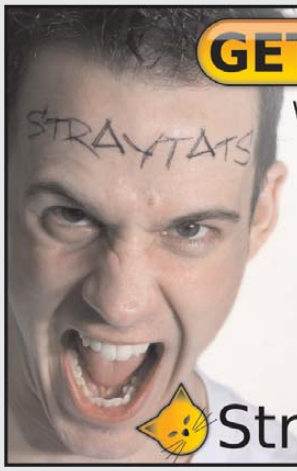
Digi-Key CORPORATION

www.digikey.com

TEXAS' OLDEST CRAFT BREWERY

WWW.SAINTARNOLD.COM

GET NOTICED!



With custom, temporary tattoos! Any image!

5% Off With This Code:
linuxjournal

Great for conventions!

StrayTats.com

American made Utility Kilts for Everyday Wear

UTILIKILTS.com

2010
A LINUX ODYSSEY

I'M SORRY MR. GATES,
I'M AFRAID I CAN'T
DO THAT

Get your Linux Odyssey t-shirt today.
LinuxJournal.com/LinuxOdyssey

Archie McPhee

Supplying the world with impractical weirdness for over 25 years!

mcphee.com

axigen MAIL SERVER

The Mail Server
for IT Professionals

www.axigen.com

STAY CONNECTED

- flickr.com/groups/linuxjournal/pool/
- twitter.com/linuxjournal
- linuxjournal.com/rss_feeds
- facebook.com/linuxjournal
- identi.ca/linuxjournal

YOUR TECHNOLOGY SHOULD WORK AS FAST AND SMART AS YOU DO.

And choosing Linux should never limit your technology options.

Parallels
Optimized Computing

Talk to the people who know Linux; talk to Parallels.

E: offer@parallels.com
P: 425.282.6448
www.parallels.com

Polywell Mini-ITX PCs

NVIDIA® ION™

The World's Small, Greenest, Fanless PC with Blu-Ray Ready

NVIDIA ION

ITX-40A with NVDIA® ION™ Graphics
4GB RAM, 1.6GHz Intel® 4W Processor
$399 with DVD-RW, 320G HD
$599 with Blu-ray, 500G HD

ITX-30A with PCI Riser

ITX-10A 1.4" ThinPC

Fanless Slim PC with Intel® 1.6GHz 1W Mobile Atom™ Processor, DC12V Power-in, 1GB DDR2 RAM starts at **$199**
Supports SATA Hard Drive or Solid State Drive (SSD), Optional PCI RISER Slot for TV Tuner or other Add-on Device on ITX-30A

Over 50 Mini-ITX Models Available with
- NVIDAI® GeForce 8200/8100 with AMD® Processor
- NVIDIA® GeForce 9300/7100/7050 with Intel® Processor
- PCI, PCIe, MiniPCIe Slot for TV Tuner or Industrial Add-on
- Custom Design Chassis for Small to Mid Size OEM Project

888.765.9686
linuxsales@polywell.com
www.polywell.com/us/Lx

■ 23 Years of Customer Satisfaction ■ 5-Year Warranty, Industry's Longest ■ First Class Customer Service

Polywell Computers, Inc 1461 San Mateo Ave. South San Francisco, CA 94080 650.583.7222 Fax: 650.583.1974

NVIDIA, ION, nForce, GeForce and combinations thereof are trademarks of NVIDIA Corporation. Other names are for informational purposes only and may be trademarks of their respective owners.

POLYWELL

EOF

The Hacking of Infrastructure. And Vice Versa.

Because open source is the nature of infrastructure itself.

DOC SEARLS

At the MIT end of Massachusetts Avenue in Cambridge, for a few blocks on either side of Central Square, the sidewalks are wider than two traffic lanes. The widest part, alongside the street, is paved with nice red brick. Much of the sidewalk is shaded by small trees planted in squares covered with four iron grates, each with a round bite off the inside corner, to make room for the tree. On the sidewalk and the road, the brick and asphalt are covered with spray-painted markings, in blue, red, orange, yellow, green and white. Roots under some of the trees are lifting and spreading the grates.

The graffiti is official. The markings are made by professionals who identify different forms of underground utility infrastructure, with color coding on the "Dig Safe" standard: blue for potable water, red for electric wiring, yellow for gas lines, orange for communications cabling (mostly phone and TV), green for waste water and white for planned digging perimeters. The colored markings say what lies beneath. Think of Dig Safe as the open sourcing of utility infrastructure.

What Dig Safe recognizes and codifies is the innate hackability of infrastructure. It's all temporary, all improvable, all replaceable. Bricks can be turned over to erase markings. Grates can be removed when trees outgrow them. Wires draped on poles can be buried, dug up and buried again. In some towns, buried service requires a trench the depth and width of a grave, with minimum spacings of three feet each between electric, cable television and telephone wiring. Some towns now are getting ready to eliminate all that deep digging and are requiring that communications utilities use fiber-optic cabling, which can run through conduits as narrow as a half-inch across, right next to electric wiring. Trenches then will be shallower and cheaper, but on the surface, the markings still will be red and orange. The simple necessities of construction and reconstruction outweigh those of aesthetics.

Many years ago Ron Wilson, then the public voice of SFO (San Francisco's big airport) was a guest on a radio talk show, explaining the airport's improvements. A caller complained that the airport always seemed to be under construction. His response: all major airports are going to be under construction for as long as we have aviation.

The same thing goes for operating systems. With durable ones, such as Linux, all the parts are improvable and replaceable, while the architecture persists. In fact, improvability is an architectural imperative.

When most of us think about architecture, what usually comes to mind is the idealized sort. "The mother art is architecture", Frank Lloyd Wright said. Wright was perhaps the greatest architect of the 20th century. Wright also said the job of the architect was to bankrupt the builder.

At the other end of the aesthetic scale is the practical architecture politely called vernacular: informal, common and arising out of local or regional usage. The word was borrowed from linguistics, where it means the same thing, but with one important difference: vernacular architecture looks to the future. It anticipates changes and uses that might come along. It is built to adjust and adapt.

In *How Buildings Learn* (Penguin, 1995), Stewart Brand said the best example of vernacular architecture was MIT's Building 20. Called "the magical incubator", it lived from 1943 to 1998. Wrote Fred Hapgood, "The edifice is so ugly...that it is impossible not to admire it, if that makes sense; it has ten times the righteous nerdly swagger of any other building on campus, and at MIT any building holding that title has a natural constituency." Among the nerds who swaggered there was the Tech Model Railroad Club, which coined the label "hackers" (also foo, mung, cruft and much more), while also spawning countless hacks, including the first video game, *Spacewar*. Among Building 20's other credits are radar, microwave, spectroscopy, quantum mechanics, atomic and molecular beams, masers and lasers, atomic clocks, radio astronomy, linear particle acceleration, magnetron phasing, fiber optics and digital data transmission.

Think of those as things that happened in user space, made possible by Building 20's kernel space. With Linux, user space exists because kernel space is there to support it. And, user space expands as kernel space becomes progressively more supportive of more uses.

The fact that Linux is practical, however, does not diminish the need for an aesthetic sense. On a *Linux Journal* Geek Cruise in 2003, Linus gave a "State of the Kernel" talk in which he presented a slide titled "People". The first bullet read, "Calm, rational, non-flaming—and good technical tastes too! In a word: rare."

Since then, Linux has become far more infrastructural, because far more of the world depends on it. For example, Netcraft.com reports that Microsoft's new Bing.com runs on Linux in Akamai netblocks. And why wouldn't it? Microsoft wants Bing to work reliably, just like Google has always done—on Linux. Hey, we all have room for improvement. Here, Linux gave some to Microsoft.■

Doc Searls is Senior Editor of *Linux Journal*. He is also a fellow with the Berkman Center for Internet and Society at Harvard University and the Center for Information Technology and Society at UC Santa Barbara.

WEB HOSTING BACKED BY FANATICAL SUPPORT®

Hosting is all we do.
And we do it better than anyone else in the world. We've set the standard for **Managed**, **Cloud** and **Email Hosting** with our support, expertise and diverse solutions for customers of all sizes, kinds and needs.

MANAGED HOSTING · CLOUD HOSTING · EMAIL HOSTING

experience ***fanatical support***®

- Unlimited live support available 24x7x365 from your dedicated team
- Hundreds of IT experts including Red Hat Certified Engineers
- Scalable solutions to efficiently meet growth needs
- On-demand, low-cost computing from The Rackspace Cloud™
- Highest level of protection with an array of security services

rackspace.com/linuxjournal | 1.888.571.8976

rackspace® HOSTING

More GFLOPS, Less WATTS

Intel® Nehalem is here!

Higher Memory Bandwidth with DDR3 and QPI
Clusters and Servers Consume Less Power

Four Servers in a 2U Chassis with all Hot-Swap:

- 1200 Watt 1+1 supply, 12 Drives, and Server Modules!

FasTree™ ConnectX® QDR and DDR InfiniBand Switches and HCAs

Intel Professional Compiler Suite and Cluster Toolkit

- Version 11 with Nehalem Enhancements
- Academic Pricing Available

Configure your next Cluster today!
www.microway.com/quickquote

intel Channel Partner Premier Member 2009

GPU Computing

WhisperStation™

With 1 to 4 Tesla GPUs

Tesla C1060 GPU Performance:

- 1 TFLOPS per GPU
- 4 GB DDR3 per GPU
- 102 GB/Sec Bandwidth
- CUDA SDK

Run MATLAB® on Tesla with "Jacket"

Clusters With Tesla™

S1070 - 4 GPU Servers

- 36 GPUs + 36 CPUs + 24 TB in 24U
- 40 Gbps FasTree™ InfiniBand
- InfiniScope™ Network Monitoring

FREE 15-day trial available at **microway.com**

Microway®
Technology you can count on℠

508-746-7341
microway.com

GSA Schedule Contract Number: GS-35F-0431N