news@UK

The Newsletter of UKUUG, the UK's Unix and Open Systems Users Group
Published electronically at `http://www.ukuug.org/newsletter/`

Volume 14, Number 1 ISSN 0965-9412 March 2005

Contents

News from the Secretariat

Jane Morrison

Thank you to everyone who kindly sent in their subscription payments so promptly. We have received a good number of early payments. Those remaining outstanding will be chased this month and any not paid at the end of March will not receive the next issue (June) Newsletter.

The UKUUG Winter Conference and Perl workshop held in Birmingham on 24th and 25th February was very successful. We achieved good attendance numbers which was due to the strong programme put together by Ray Miller. The UKUUG part of the event was preceded by the Exim Tutorial and Conference: this was well attended and the two events may be jointly organised again in the future.

There is an article about the Winter Conference elsewhere in this issue, and you should find a CD enclosed with this Newsletter containing the papers from the conference.

The Linux 2005 Conference will be held in Swansea between the 4th and 7th August. Details, including a web form for potential speakers' submissions are available at
`http://www.ukuug.org/events/linux2005/`

As in previous years we are hoping to be able to achieve sponsorship for the event which will allow us to keep delegate fees at a minimum. If you know of any company who may be interested in sponsorship please let me know.

The next Newsletter will be the June issue and the copy date is Friday 20th May.

Any comments about past or future events, or if you have something to say about the UKUUG or this Newsletter please contact `ukuug@ukuug.org`.

Linux 2005 – Conference and Tutorials

Jane Morrison

Thursday 4th to Sunday 7th August

University of Wales, Swansea

Planning for the 2005 Linux Technical Conference is already underway and again we are seeking ideas, speakers and sponsors.

We invite speakers on all types of Linux development to contribute. The programme will cover a variety of subjects, including kernel and desktop development, tools, applications, and networking. Any topic likely to be of interest to Linux developers and enthusiasts will be considered.

The topics presented in recent years have included: ARM Linux, Benchmarking, Clustering, CORBA, Debian Package Management, Enterprise Filesystems, Exim, Flightgear, GNOME, Heartbeat, I20, JFFS, KDE, Mail Servers, Memory Management, Performance Programming, Powertweak, Pymmetry, Samba, Security, SMP, Vmware, Zerocopy and Zope.

More information about previous events can be found at:
`http://www.ukuug.org/events/`

Abstracts for the main conference or tutorial programme should be accompanied by a short biography, and, ideally, should be about 250-500 words long. Final papers should normally last about 45 minutes, including 10 minutes for questions and answers. If you need more time for your presentation, please tell us when you submit your abstract.

We shall acknowledge all submissions.

Significant Dates

Closing date for abstracts: 24 March 2005

Accepted authors notified by: 18 April 2005

Final papers due by: 17 June 2005

Particular queries should be sent either to the UKUUG office, or to the Linux2005 programme committee, `linux2005@ukuug.org`.

Sponsors and Exhibitors

To keep the conference fees low, we are seeking sponsors and exhibitors. For further information about sponsoring, exhibiting, or attending the event please contact the UKUUG office.

UKUUG Award 2005

UKUUG Council

The UKUUG Award (previously known as the UKUUG Open Source Award) is given annually (if submissions of sufficient merit are received) for a significant contribution to free and open source software; this might be in the form of an article or paper, software product, or other contribution.

The purpose of the Award is to encourage and foster developers and other contributors in or connected with the British Isles. We are looking for a work or project that is, or might be, significant in the world of free and open source software. The Award is not limited to (recent) UK students although special encouragement is given to entries therefrom and the best of those entries will be recognized.

Initially, a one-page summary (accompanied by an abstract and a short biography of about 250-500 words) of the work should be submitted via
`http://www.ukuug.org/osa/submit`

If the work is part of a joint project, the personal contribution should be stated clearly. The judging panel will include representatives from UKUUG, UK computer science departments, and the wider community. If the judges are unable to distinguish between the merits of the best candidates, the prize shall be divided accordingly. No person may be awarded the prize more than once.

UKUUG are giving a prize of £ 500, to which O'Reilly UK Ltd are generously adding both a pass to the Open Source Convention and monies towards travel and accommodation.

In 2004 the award winner was Julian Field for MailScanner, and Jake Stride's submission of an Enterprise Groupware system and study of Open Source development techniques was highly commended.

UKUUG Winter Conference 2005

Jono Bacon

Of all the list of flourishing organisations with an interest in Open Systems, Open Source and Linux, the UK UNIX Users Group is possibly the oldest and most established. Boasting a large membership, many of whom attended the Winter 2005 conference, the UKUUG has traditionally provided those with an interest in UNIX-like Operating Systems with a community and associated events to continue and discuss UNIX and its culture. As a writer who is not part of the UKUUG and a relative newcomer to UKUUG events, I was intrigued to see how the conference would fare. Based at the elegant Paragon Hotel in the centre of Birmingham, the event lined up a number of talks and presentations with a focus on networking and security.

The Winter conference took place immediately after the Exim conference, and the organisation of the two events was combined. The Exim event also attracted a good number of delegates and offered a number of interesting in-depth talks relating to this popular MTA. With Exim as the the default MTA in the popular Debian distribution, it was no surprise to see a number of Debian T-shirts at the Winter Conference.

Alongside the last sessions of the Exim conference, but before the start of the UKUUG conference sessions, a tutorial for UKUUG members on Perl 6 was given by Allison Randall of the Perl Foundation.

Day 1

Kicking off the Winter conference were two presentations, both looking at different aspects of security. The first, 'Open Source Security Lessons' by Wietse Venema told the the tale of how an intruder broke into a number of computers in the Netherlands in the mid 1990s and also potentially breached a number of US military systems. Wietse shared with everyone the story of how he and his colleagues explored how the intruder was getting in, and how, partly in response, they developed SATAN: a tool which was embroiled in furore and media controversy when it was first released. The second presentation, 'Worry about security later: separation of concerns in distributed services' by Christian Fernau characterised security as a bolted-on problem that existing developers are often not experienced in. This presentation focused on the development of an abstracted security process that gives application developers the opportunity to concentrate on their code, and not have to up-skill themselves in cutting edge security development.

Following a short break in which everyone mingled to discuss the plethora of security related information just digested, the proceedings continued with a discussion of the Border Gateway Protocol, also known as BGP. This protocol is used to exchange routing information, and Henning Brauer provided an interesting discussion about the protocol, its design and his implementation of bgpd for OpenBSD. After another short coffee break, Tony Finch offered a discussion of Bounce Address Protection for Email; a technique that can be used in the war against spam. Following this presentation, a discussion of LDAP Schema Design was given by Andrew Findlay. With LDAP providing a solution for large-scale business problems, the presentation filled in a number of gaps in developing a sensible topological LDAP schema. The day was then wrapped up with a PGP key-signing session and dinner.

Day 2

Day two kicked off at 9.10am on a ferociously cold Birmingham day. After the introduction and announcements, Daniel Serain of Oracle presented a talk on 'The Ultimate of Networking: the Business Process Layer'. In this presentation, concepts such as RosettaNet and SOAP/XML were touched on in a presentation that looked at business processes depending on Web services and related high level technologies. Following Daniel's talk, Wietse Venema gave a presentation on Secure Programming. This was a highlight of the conference, using as a case study the reasons why an apparently useful piece of code intended to shred files actually fails at a number of different levels, and drawing some very useful morals from the story.

Following a brief coffee break, a discussion on digital certificates was given by Mark Norman. As a member of the DCOCE Project, Mark gave a compelling presentation on the much discussed issue of digital certificates and how they can scale in different environments. Henning Brauer then returned to educate the audience on how to write signal handlers; a topic commonly riddled in complexity, made more manageable by Henning.

After lunch and the discussion of the morning's content by the audience, Henning returned to complete his hat trick and gave a talk about ntpd; the network time daemon. He described the design and philosophy behind his re-implementation of ntpd for the OpenBSD project, with particular reference to security and simplicity.

Following this talk, a presentation on the insecurities of contemporary GRID middleware was given by Ivaylo Kostadinov. With Globus being a common middleware toolkit, Ivaylo discussed

how Globus suffers from some major security issues. The final presentation of the day was given by Simon Biles, who talked about a SPADE plugin for the Snort intrusion detection system. Snort has gained huge popularity in recent years and this talk was sure to capture the minds of many of the delegates.

Following the end of the conference proper, a meeting was held to discuss UKUUG strategy.

In addition to the two days packed with talks, there was an array of books and other items on sale in the nearby room. The inimitable Josette was there providing a large range of O'Reilly books, and on a separate table, many of the GNU Press documentation books were available to buy. As a first timer at a UKUUG event, it was great to see such a large and impressive turn-out, and it was fantastic to see such a wealth of talks. If you are a member of the UKUUG, it is highly recommended that you get along to the next event.

UKUUG membership survey

Roger Whittaker

UKUUG recently carried out a membership survey. The purpose of the survey was to obtain more information about UKUUG's membership for two purposes. Firstly if Council has better information about who the organisation's members are, we can serve the membership more effectively. Secondly, having a profile of members' professional positions and influence within their organisations is very useful in persuading vendors to offer sponsorship for UKUUG events. Sponsorship from vendors is a very welcome way of keeping down the costs of events, and provides a mutually beneficial way of improving our contacts with vendor companies.

The participation rate in the survey very good, relative to comparable surveys in similar organisations. In many ways the survey confirmed what we already suspected about our membership: the main job roles were system administration and programming, and the membership was biased towards academic institutions.

A number of questions were included to gauge support for certain new initiatives which UKUUG is now actively considering, following the strategy meeting at the end of the Winter Conference. For example a majority of survey participants were interested in a training and certification initiative, in conjuction with local groups.

There were also majorities in favour of UKUUG taking on some kind of lobbying and/or activism role.

A summary of the survey results is available at
`http://www.ukuug.org/survey2005/`

Strategy Meeting: 25th February 2005

Roger Whittaker

At the end of the Winter Conference, delegates were invited to stay on for a strategy meeting to discuss aspects of UKUUG's organisation. About 23 people attended the meeting, and there was an interesting and lively discussion.

Ray Miller chaired the meeting and began by referring to the results of the recent membership survey. He said that this had been remarkably successful in that the participation rate had been considerably higher than is often achieved in such surveys.

He summarised the results of the survey (which can be found at

`http://www.ukuug.org/survey2005/`

There was some discussion about the possibility of UKUUG initiating some kind of certification system for individual skills, possibly in conjunction with the Linux User Groups. This might be done in conjunction with Open Forum Europe's OSCoP Competency Framework.

Mike Banahan agreed to chair a new training working group in which this would be explored further. He was also willing to liaise with Open Forum Europe on this and other developments, as he is CTO of OFE.

Members also agreed that it would be useful to try to cultivate closer links with organisations such as AFFS and FFII. This would allow coordination between organisations, prevent wasted duplication of effort and allow each organisation to work at what it does best.

Ray explained the agreement which had recently been made with Leslie Fletcher who will be attending a variety of meetings and organisations, and following and reacting to press articles and other news on behalf of UKUUG. He explained that Leslie's main purpose in all this would be to raise the profile of UKUUG and to present arguments in furtherance of its aims and objectives. He noted that Leslie would attend various meetings and introduce himself as a member of UKUUG, but that he would not be there as a representative of UKUUG in the sense of making or expressing the organisation's official policy.

It was agreed that it would be useful if members could pass on news articles to Leslie, particularly if there was a possibility of replying or reacting to them with a statement on behalf of UKUUG. In this connection, it was also agreed that Council would formulate official policies and produce informational leaflets on certain matters which are central to the aims and objectives of UKUUG, but only where if was felt that there would be overwhelming support for these from members.

Various members expressed their willingness to serve on a "list of experts" to be called on when necessary to provide a quote on matters of topical interest.

There was some discussion about the newsletter. Almost all of those present wanted a paper edition of the newsletter to continue, but it was agreed that there was no objection to an on-line version being posted on the web site shortly after the printed version was published. It was agreed that corporate members would receive multiple copies for their organisations: in some cases this had not been happening lately.

There was general agreement that more in-depth technical articles would be welcome: members were invited to contribute more material. Book reviews were also welcome: any member could volunteer to review books or contribute articles by joining the relevant mailing lists.

The subscription address for the books list is: `book-subscribe@ukuug.org`

For the newsletter: `newsletter-wg-subscribe@ukuug.org`.

UKUUG Treasurer

UKUUG Council

At the UKUUG AGM in the autumn, James Youngman will be stepping down as treasurer since by then he will have served the maximum two consecutive terms on Council. James is currently the UKUUG's treasurer and so Council is seeking a replacement.

The duties of the treasurer are not onerous but need to be done regularly (for example, there are some cheques to sign every month). To ease the transition, Council is now looking for a volunteer to take over from James when he steps down. The idea would be for the new treasurer to work with James before James retires, in order to get up to speed before taking over. This will ensure that James is put out to pasture with a minimum of disruption.

So, if you would consider helping the UKUUG and are willing to spend a few hours every month, your contribution would be very much welcomed. Similarly, if you know someone else who doesn't mind a little light financial work (it's not complex) then please encourage them to get in touch with us.

As this point you will be wondering what this really entails. We're certainly not talking about involved book-keeping (the Secretariat does that). The duties of the Treasurer typically include:

1. Take responsibility for ensuring the continued financial viability of the UKUUG (formally, all directors share this responsibility but it is principally the Treasurer's concern).

2. Authorise and regulate expenses including Council expense claims, and budgets for events, in conjunction with the organiser of each event.

3. Ongoing financial control including:

- Review monthly expenditure (and sign the cheques)
- Authorise funds transfers (e.g. for speakers' expenses)
- With council, review membership rates
- With council, plan the disposal of sponsorship income
- Review of other things, for example book discounts

4. Financial planning: periodically review the provision of financial services (e.g. deposit accounts and merchant services).

You don't need to have previous experience of financial management or bookkeeping to do this, and it's not that complicated. James is a Unix buff, not an accountant, after all.

As a member of the UKUUG Council you would also have a voice in the running of the organisation, along with the rest of the members of the Council.

UKUUG press watch

UKUUG Council

Following various discussions recently, including the recent public strategy meeting, UKUUG hopes to be more reactive to news items and events. Members who have a paricular area of expertise are invited to write to `ukuug@ukuug.org` if they are willing to be on an "experts list" and are willing to speak to journalists or others on relevant matters. Note that in doing this you will speaking on your own behalf, but as a member of UKUUG, not making or representing UKUUG policy as such.

Examples of topics which might be relevant in this regard are:

- Software Licensing
- Free Software
- Software Patents
- Software Copyrights
- Free BSD
- Linux

- Unix
- Apple OS X
- Viruses
- Microsoft
- Open Source Desktop
- Computer and Network Security
- Spam
- Internet
- e-government
- Education, Training and Certification
- Databases

UKUUG Diary

Roger Whittaker

We now maintain an on-line diary of events which may be of interest to members at
`http://www.ukuug.org/diary/`

The following is a small selection of summaries of some interesting forthcoming events (with starting dates) taken from the diary.

We hope that the UKUUG diary is already a useful resource: we are always happy to receive information about other events of possible interest to members which will be added to the diary.

CREATIVE CAPITAL: Culture, Innovation and the Public Domain in the Knowledge Economy

17th March 2005: Amsterdam, The Netherlands
`http://www.creativecapital.nl/`

The Shock of the Old 2005: Implementing Innovation

7th April 2005: Oxford
`http://www.oucs.ox.ac.uk/ltg/events/shock2005/`

USENIX '05 Annual Technical Conference

10th April 2005: Anaheim, California, USA
`http://www.usenix.org/events/usenix05/`

ACCU Conference 2005

20th April 2005: Oxford
`http://www.accu.org/conference/`

Debconf5

10th July 2005: Helsinki, Finland
`http://www.debconf.org/debconf5/`

O'Reilly Open Source Convention

1st August 2005: Portland, Oregon, USA
`http://conferences.oreillynet.com/os2005/`

UKUUG Linux Technical Conference

4th August 2005: Swansea
`http://www.ukuug.org/events/linux2005/`

FAVE - Free Audio and Video Event

13th August 2005: Bristol
`http://fave.org.uk/`

LinuxWorld Conference and Expo

5th October 2005
`http://linuxworldexpo.co.uk/`

UKUUG member benefits reminder

James Youngman

This is a reminder that as a UKUUG member you can have an email alias of the form

`Firstname.Lastname@ukuug.org`

The use of such an address is also the key to taking advantage of some of the other benefits, for example discounts on training from GBDirect and hosting from Bytemark Hosting, because these companies will take your possession of such an email address as evidence that you qualify for discounts. For further details, please see
`http://www.ukuug.org/membership/`

and
`http://www.ukuug.org/discounts/`

As well as our traditional discounts on O'Reilly books, we also offer discounts on books from GNU Press, Pearson Education, Wiley and UIT Cambridge. There is more information at
`http://www.ukuug.org/books/`

Consultants' referral scheme

Alain Williams

UKUUG receives requests for technical help from the public: corporate and individual. To be better able to deal with these and to do so in a fair manner it has been decided to create a consultants' listing area on the UKUUG website where UKUUG members will be able to advertise their services to visitors.

The listing will be split into two parts: commercial and individual members. For fairness the ordering of entries within the two parts will be randomised.

Corporate members will be allowed two paragraphs (100 words).

Individual members will have their name and a URL and/or mail address.

There will be no initial verification by UKUUG that a consultant is qualified for the services claimed - but feedback will be accepted from clients. There will be no extra charge to UKUUG members for listing, but, as with other member services, verification of membership status will require a valid UKUUG email address of the form `Firstname.Lastname@ukuug.org`.

The list of consultants will be at:
`http://www.ukuug.org/consultants/`

Announcement: Gnome-UK

Ray Miller

GNOME-UK is about organising and promoting GNOME awareness in the United Kingdom. This includes, among other things, organising stands and Linux Events, such as the Linux Expos in London.

If you'd like to be involved with GNOME activities and meet GNOME users and developers in the UK, then sign up to the mailing list, or pop in to the gnome-uk channel on IRC.

Mailing List:
`http://lists.linux.org.uk/mailman/listinfo/gnome-uk/`

Website:
`http://www.uk.gnome.org/`

There is also an IRC channel:

#**`gnome-uk`** on **`irc.gnome.org`**.

Announcement: Free Software Magazine

Ray Miller

Free Software Magazine is a free magazine for the free software world, available on paper and in electronic format. All the articles published in Free Software Magazine (with some necessary exceptions) are released under a free license six weeks after they are published. This means that only subscribers get the latest articles; it also means that over time more and more articles will be available online. See
`http://www.freesoftwaremagazine.com/`

for more information and subscription details.

Issue 2 is avalilable for free download now:
`http://www.freesoftwaremagazine.com/free_issues/issue_02/`

Announcement: Open Source Observatory

Ray Miller

We have recieved the following announcement.

The European Commission's Open Source Observatory (OSO)
`http://europa.eu.int/ida/oso/`

is a clearinghouse of information related to free/libre/open source software in the public sector, and is intended to promote and spread the use of best practices in Europe. The OSO is part of the EC's IDA (Interchange of Data between Administrations) programme, and ultimately aims to provide a comprehensive overview of open source software policies and activities in the public sector, especially in current and future EU Member States.

The OSO is working on an inventory of free and open source software. This OSO Software Inventory is a catalogue of replicable free and open source software solutions for eGovernment. The aim of the Inventory is to classify and briefly describe these solutions, to provide contact information for the software solution developers, to permit providers and potential users of eGovernment solutions to specify and search for relevant applications, and to thereby allow public administrations to estimate, if possible, whether and how these solutions might be replicable.

We invite you to submit any free or open source software that your project or organisation has produced for inclusion in the OSO Software Inventory. Submission of your software project is easy, you only need to fill in an online form
`http://europa.eu.int/ida/en/chapter/5649`

with a few questions regarding the nature of the software. In a few days thereafter, you will see your software description online on the Inventory website.

We would also appreciate it if you could pass this email on to any software projects you might know of, or post this message to any free or open source software mailing lists.

If you have any questions regarding the OSO or the OSO Software Inventory, please feel free to send an email to: `gposs@cec.eu.int`

The IDA-Open Source Observatory Team

Links: IDA-OSO:
`http://europa.eu.int/ida/oso/`

OSO Software Inventory – submit your software at:
`http://europa.eu.int/ida/en/chapter/5649`

Open Source related News on the OSO:
`http://europa.eu.int/ida/en/chapter/469`

Usenix: LISA 2005 call for papers

Ray Miller

We have received the following announcement of the Usenix LISA 2005 event to be held in San Diego, California, USA.

The event will take place between the 4th and 9th of December 2005.

The LISA '05 organizers invite you to contribute proposals for refereed papers, invited talks, and workshops, plus any ideas you have Guru Is In sessions, Work-in-Progress reports, and training sessions.

The Call for Participation with submission guidelines and sample topics can be found on the USENIX Web site at
`http://www.usenix.org/lisa05/cfpa`

The annual LISA conference is the meeting place of choice for system, network, security, and other computing administrators. Administrators of all specialties and levels of expertise meet at LISA to exchange ideas, sharpen skills, learn new techniques, debate current issues, and meet colleagues and friends.

People representing every work assignment from the full-time position at a large site to the part-time one at a small shop come to LISA from over 30 countries, bringing divergent backgrounds and experience levels to the conference dedicated to them.

System and network administrators from environments as diverse as academia, large corporations and small businesses, government organizations, and research sites find LISA to be The Place to go for training, education, networking, and interacting with their peers.

Submissions are due by May 10th 2005: authors will be notified during June.

Final papers are required by September 27th 2005.

See also:
`http://usenix.org/lisa05/`

Where am I? The security problems of chdir()

James Youngman

After buffer overflows, race conditions are probably the most widely known form of security problem. The essence of a race condition is that it is a circumstance where somebody else can negatively affect what your program does by performing some other action at just the wrong moment.

Many race conditions are normal and are harmless. For example, it is normal for people to create or delete directory entries while **find** is running. All implementations of **find** are expected to cope successfully with that. However, not all race conditions are so benign.

From a security point of view, one can think of **find** as a tool for converting an untrusted set of data (the contents of the filesystem) into a set of data with known properties (i.e. a list of files which at some point matched a set of criteria). This means that almost anything that an outsider can do that produces a mismatch between the files the user intended **find** to match and the list of files that **find** actually matched would represent a potential security problem. This might include for example, ensuring that **find** matches files that you didn't intend, or fails to match files that you did intend it to match. A range of security problems of this sort arise if an attacker can persuade **find** to search a part of the directory tree that you didn't want it going into. For example it would be bad if the command

```
find /tmp -mindepth 1 -mtime +30 -delete
```

could be persuaded to search **/etc** as well as **/tmp** since this could well result in the deletion of many useful system configuration files. Similarly it would be bad if an attacker could defeat the intent of **-xdev** or **-prune**. Unfortunately a naive implementation of **find** is open to exactly this type of attack, and race conditions are easy ways to make this happen. This article discusses how such a thing can happen and how GNU **find** protects itself against such things.

As find searches the file system, it finds subdirectories and then searches within them by changing its working directory. First, **find** notices a subdirectory when its name is returned by **readdir()**. It then decides if that subdirectory meets the criteria for being searched; that is, any **-xdev**, **-maxdepth** or **-prune** expressions are taken into account. The **find** program will then change working directory and proceed to search the directory.

A race condition attack might take the form that once **find** has called **lstat** to collect information about a subdirecory (which we'll call **foo**) and the checks relevant to **-xdev** and **-prune** have been done, an attacker might rename the directory that was being considered, and put in its place a symbolic link that actually points to **/etc**, for example. The next thing that **find** will do is use **chdir** to change working directory into **foo**. However, **foo** is now a symbolic link to **/etc**. While **find** expected to end up in the subdirectory **foo** it has actually ended up in **/etc**.

The idea behind this attack is to fool **find** into going into the wrong directory. This can have all the nasty consequences we saw above. This form of attack is particularly problematic if the attacker can predict when the **find** command will be run, as is the case with **cron** tasks for example.

GNU **find** has specific safeguards to prevent this general class of problem. The exact form of these safeguards depends on the properties of your system.

O_NOFOLLOW

If your system supports the **O_NOFOLLOW** flag to the **open(2)** system call, GNU **find** uses it when safely changing directory. This flag was introduced in FreeBSD 3.0-CURRENT and in version 2.1.126 of the Linux kernel.

The target subdirectory is first opened and then **find** changes working directory with the **fchdir()** system call. This ensures that symbolic links are not followed, preventing the sort of race condition attack in which use is made of symbolic links.

You can tell if your system supports **O_NOFOLLOW** by running

```
find --version
```

This will tell you the version number and which features are enabled. For example, if I run this on my system now, this gives:

```
GNU find version 4.2.18 Features enabled:
D_TYPE O_NOFOLLOW(enabled)
```

Here, you can see that the **D_TYPE** and **O_NOFOLLOW** features are present. **O_NOFOLLOW** is qualified with "enabled". This means that the current system seems to support **O_NOFOLLOW**. This check is needed because it is possible to build **find** on a system that defines **O_NOFOLLOW** and then run it on a system that ignores the **O_NOFOLLOW** flag. We try to detect such cases at startup by checking the results of the **uname(2)** system call; when this determines that your system is too old to support **O_NOFOLLOW**, you will see "**O_NOFOLLOW(disabled)**" instead.

Systems without O_NOFOLLOW

The strategy for preventing this type of problem on systems not supporting the **O_NOFOLLOW** flag is more complex. Each time **find** changes directory, it examines the directory it is about to move to, issues the **chdir()** system call, and then checks that it has ended up in the subdirectory it expected. This check is performed by examining the device number and inode number returned by **lstat()**.

If we have ended up in a directory other than the one we expected, an error message is issued and **find** exits immediately. This method prevents filesystem manipulation attacks from persuading **find** to search parts of the filesystem it did not intend. However, we have to take special steps in order not to unnecessarily conclude that there is a problem with any "automount" mount points.

Working with automounters

Where an automounter is in use it can be the case that the use of the **chdir()** system call can itself cause a new filesystem to be mounted at that point. On systems that do not support **O_NOFOLLOW**, this will cause **find**'s security check to fail.

However, this does not normally represent a security problem (since the automounter configuration is normally set up by the system administrator). Therefore, if the **chdir()** sanity check fails, **find** will check to see if a new filesystem has been mounted at the current directory; if so, **find** will issue a warning message and continue.

To make this solution work, **find** reads the list of mounted filesystems at startup, and again when the sanity check fails. It compares the two lists to find out if the directory it has moved into has just been mounted. The lists are compared by using device numbers and inode numbers, in case there are symbolic links within the paths to the mount points as they are specified in **/etc/fstab** (or its equivalent on the current system) or the working directory from which **find** was started includes a component which is a symbolic link.

Problems with dead NFS servers

Examining every mount point on the system has a downside too. In general, **find** will be used to search just part of the filesystem. However, **find** examines every mount point. If the system has a filesystem mounted on an unresponsive NFS server, **find** will hang, waiting for the NFS server to respond. Worse, it does this even if the affected mount point is not within the directory tree that find would have searched anyway.

This is very unfortunate. However, this problem only affects systems that have no support for **O_NOFOLLOW**. As far as I can tell, it is not possible on such systems to fix all three problems (the race condition security problem, the false-alarm at automount mount points, and the hang at startup if there is a dead NFS server) at once. If you have some ideas about how **find** could do this better, please send email to the `bug-findutils@gnu.org` mailing list.

Summary

The symbolic link is a very useful feature which makes many system administration tasks easier. However, there is a trade-off between security and usability. Symbolic links bring with them a wide range of potential security problems. Many of these problems would not exist if it were possible to require that symbolic links be ignored some of the time. For example it would be easier to write and maintain a secure version of **`find`** if this feature existed in the POSIX standard. Some versions of Unix do provide a simple version of such a feature, and it proves to be very useful. The workarounds for not having **`O_NOFOLLOW`** are very complex and far from perfect. Please ask your Unix vendor to implement and support the widespread adoption of the **`O_NOFOLLOW`** flag.

James Youngman is the maintainer of GNU findutils (http://www.gnu.org/software/findutils/) and GNU CSSC (http://directory.fsf.org/CSSC.html). James is also currently the treasurer of the UK Unix Users' Group.

High Performance Linux Clusters with OSCAR, Rocks Open-MOSIX and MPI

Joseph Sloan

O'Reilly and Associates

ISBN 0-596-00570-9
367 pp.
£ **28.50**

reviewed by John Hearns

Anyone reviewing this book should be of course aware that this is the second foray by Oreilly into Beowulf clustering. The first, "Building Linux Clusters" by David Spector was not well received by the community. By coincidence, just when I was completing this review a rather uncomplimentary review of Sloan's book was posted to the Beowulf list.

This book does exactly what it says on the tin – it gets you onto the path of constructing clusters using the above-mentioned packages. The book is divided into four parts: Introduction, Getting Started Quickly, Building Custom Clusters and Cluster Programming.

In the section regarding choice of hardware for a Beowulf system, there is a recommendation to make sure that a video adapter is included in the purchase – in my experience any suitable motherboard these days has an inbuilt adapter, which is perfectly adequate for diagnostic use. The author also quite correctly recommends motherboards with PXE network booting capability, however then talks about using PXE ROMS in sockets on the board. PXE capability is part of the list of 'must haves' for any motherboard suitable for a Beowulf these days, and you should not be concerning yourself with inserting boot ROMS these days. This section refers to motherboards which may refuse to boot when a keyboard is not detected – in my opinion if you cannot set 'ignore keyboard' in the BIOS then such motherboards are again not suitable for a Beowulf. Better to make an informed choice of motherboard based on these requirements before starting your project. As regards serial console access, we enable this on all our systems and find it a highly convenient way of accessing multiple rack-mounted systems. Using Cat5 cable for the serial connections makes for neat cabling, and a terminal server in the rack gives you direct access to all systems. The author does discuss this, however making the claim that this is "a fair amount of work on most Linux systems", which I don't agree with.

This section belies the book's bias towards the homebrew, "build a cluster from a pile of donated machines" philosophy. Things have moved on dramatically from those days, and Linux clusters are now an integral part of many scientific and engineering department's research tools, and critical to large business compute resources. The author too quickly dismisses rackmount cases

as being expensive and "for the high end". On the contrary, the LOBOS (Lots of Boxes on Shelves) approach is an invitation to a snake's wedding, unless one person rules the cabling infrastructure with a rod of iron, and desktop PSUs are really not intended to take the 24 hour high loads of a cluster on full song. A proper rack mount is the professional way to do things, you'll get server-grade nodes with good cooling flows and future addition of nodes when you inevitably come to grow will be as easy as slotting them in the rack. The author certainly does discuss the issues of cooling and power requirements, but extracting heat from (say) a 128-node dual Opteron cluster takes a proper machine room infrastructure.

There is a short mention in Chapter Three concerning high performance networks for Beowulf clusters. Less than half a page does not do this topic justice, as it is an integral part of specifying and tuning a high performance cluster. Many clusters certainly do perform very well using gigabit ethernet. Using a motherboard with twin inbuilt gigabit, segmenting parallel traffic and general cluster/NFS as discussed here is a good technique. We also provide low-latency drivers which work on commodity ethernet. The author fails to flag up that the choice of gigabit switch is important for performance – you can't expect a cheap office-grade eight-port switch to perform in a heavily loaded network.

However, many applications will benefit from a high-speed low-latency interconnect such as Myrinet, Quadrics QSNet or Infiniband. The choices of these interconnects, and benchmarking against your own applications is a fascinating part of clustering, and the subject of much debate, it deserves a more space in the book. Dismissing Quadrics and Infiniband as "competitive technologies that are emerging or are available" certainly does them no justice. QSNet is based on a mature technology (Meiko).

Quadrics is used in some of the biggest clusters in the world, and developed in Bristol to boot. Myrinet is commonly deployed in many clusters, both research and industrial. Quadrics and Myrinet continue to innovate, e.g. with future plans to utilise 10gig ethernet switches. This means they can use as many commodity components as possible. Infiniband is the newcomer to the market, but is gaining ground, and is moving from the R+D stage to position itself as a useful cluster interconnect.

The choice of an interconnect becomes increasingly important as node count increases. Scaling a network which can handle thousands of processors across a large fabric, without introducing bottlenecks, is the forte of these high performance interconnects. If you want to achieve the best performance, and scaling, you need one of these. The book's comment that "these highly expensive technologies are no longer needed for most applications" is misleading. They certainly do cost more than onboard gigabit interfaces and ethernet switches. However, if you are planning a new HPC cluster you should do the figures, and get the benchmark results. Your budget should aim to get the most computing performance for your budget – and setting aside part of the budget for a good interconnect rather than piling on more nodes will achieve that, depending on the nature of your application. HPC is about network/memory performance, balance and tuning as well as adding nodes.

Another topic which the author doesn't deal with in depth is the choice of commercial compilers. Many high performance applications run significantly faster using an optimised compiler. It again makes sense to reserve some of your budget for a compiler. You can easily obtain an evaluation license for most compilers, and if you (say) prove a 30% speedup in your application it will be more cost effective than buying more nodes.

Chapter 4 discusses the choice of Linux distribution and how to configure it. It is certainly important to use a stable distribution, which supports your chosen hardware well. This probably won't be the latest and greatest test kernel. The book shows its slant here, again discussing using recycled hardware and having to use an older distribution to cope with this. This chapter is a useful high-level overview of the services which a typical cluster depends on – DHCP, NFS, SSH, NTP and security, though it is really no substitute for some proper systems admin

knowledge. When it comes to debugging problems with these services, there's no substitute for experience.

The section on cloning systems was interesting to me. I had not heard of g4u – which is an equivalent to Norton Ghost for Linux systems. g4u has the nice feature of storing compressed images, but the downsides of having to reboot machines to make any updates and difficulty coping with different disk geometries make me think it is not that useful for HPC cluster installs. It certainly should be in your mind if you have to roll out many systems, eg. in a classroom or office environment. Kickstarting or image based install tools are the way to go with HPC clusters.

The next three chapters give details on openMOSIX, Oscar and Rocks. The first is a set of kernel patches which provide process checkpointing and migration. This gives the behaviour of a Single System Image machine, but is not truly an SSI mach aines the author implies. Oscar and Rocks are both frameworks for installing and deploying clusters, though with different features. One useful feature of Rocks is the inclusion of Rolls (geddit?) for the easy inclusion of additional software, eg. commercial compilers or batch systems. For my taste, both these chapters are a little too heavy on the "this is how to download, and here are the exact steps to take" approach. This is common in many O'Reilly books, and you should be prepared to download and follow the latest documentation for any package you implement, not solely depend on the book.

The section on batch schedulers deals with OpenPBS adequately, and is certainly enough to get you up and running. My only personal quibble is that little mention is made of one of the main alternatives, Sun Gridengine:
`http://gridengine.sunsource.net`

We configure SGE on the majority of our clusters, and find that the free (as in beer) version and the commercially supported N1 Grid Engine version do what our customers want. The support on the SGE mailing list is excellent, and should you decide to give it a try – see you on the list! However one can't expect an author to become an expert in all alternative batch systems just to write a book.

In such a fast-moving field, a book such as this will inevitably be slightly out of date. For instance, in the section on filesystems, there is a link to the OpenGFS project, and a comment that: "RedHat markets a commercial, enterprise version of OpenGFS". RedHat have now released GFS in open source again, and of course provide a supported version with their Enterprise Linux. Anyone considering GFS would be unwise to go with OpenGFS now.

The author discusses parallel command tools (C3 toolkit) and the Ganglia monitoring framework, both of which are important for the effective control of a whole set of machines. You want your cluster management to scale well – clusters will only get bigger, and having the tools on hand to manage them will continue to be a topic for people to work on, and there are plenty of interesting problems to handle in that area.

The final section of the book is an introduction to parallel programming. After all, now that you've chosen the hardware, installed the distribution, compiled up your libraries you want the thing to DO something. And hopefully the skills you learn here, using standards such as MPI will be transferable to other installations and large setups.

This book is a good resource for anyone wanting to get started building a homebrew cluster, or a cluster as a learning project at a high school or university level. There are in addition lots of other sources of information, many included in the appendices. If you are working as a scientist or engineer, or working for a company which needs reliable, well-managed high performance computing I would urge you to also consult the online resources such as the original Beowulf site or Clusterworld magazine. One resource which stands out is Robert Brown's online book, released under the Open Publication License, based on his experiences of building clusters at Duke University. Another excellent treatment of clusters, by one of the original Beowulf team, is Thomas Sterling's "Beowulf Cluster Computing with Linux".

Linux clusters have come of age in the last five or six years, scaling on the one hand to be the leading systems in the Top500, with thousands of processors, and down to turnkey rack mounted clusters suitable for a workgroup or research group, which can be delivered and working the same day.

If you are specifying or purchasing a new cluster for your department or company then the 'roll your own' approach isn't the best these days. Consult with the Beowulf community, ask your campus IT services for advice and the regional E-sciences centres. And any clustering company worth its salt will be happy to spend a good deal of time with you, finding out about your applications, giving informed advice about hardware choices, networking, high performance interconnects, system software and running benchmarks.

References:

`http://www.beowulf.org`
`http://www.clusterworld.com`
`http://www.phy.duke.edu/resources/computing/brahma/Resources/beowulf_book.php`

"Beowulf Cluster Computing with Linux", ed. Thomas Sterling. MIT Press.

My thanks to Ashley Pittman of Quadrics for valuable input.

Contact: `john.hearns@streamline-computing.com`

Dancing Barefoot

Wil Wheaton

O'Reilly and Associates

ISBN 0-596-00674-8
118 pp.
£ **9.95**

reviewed by Lindsay Marshall

GarageBand the Missing Manual

David Pogue

O'Reilly and Associates

ISBN 0596-00695-0
304 pp.
£ **13.95**

reviewed by Lindsay Marshall

See the combined review below.

Hackers and Painters

Paul Graham

O'Reilly and Associates

ISBN 0596-00662-4
225 pp.
£ **15.95**

reviewed by Lindsay Marshall

It seems that I am supposed to have heard of Wil Wheaton but I haven't. I have no idea who he is. Why he has a book of five short stories published by O'Reilly is beyond me. I can only

presume that he is a friend of the Boss. There can be no other reason why a technical book publisher would publish these rather dull short stories. They are workman like, but nothing more. They entirely failed to communicate any emotion to me, even though they were clearly being written with that purpose in mind. I just read the "about the author" section. Seems he has a weblog that was voted "best" in some award thing and is read by lots of people. Never seen it mentioned on any of the logs that I read. He also loves the band Cake. I saw Cake supporting the Counting Crows several years ago. They were one of the worst bands I have ever seen. 'nuff said, I think.

Which brings me on to GarageBand. I'm not a huge fan of the Missing Manual series, mostly they seem to be rather pointless - the manual is missing for a reason, you don't really need one. But GB is fairly complicated and it's good to have something that walks you through setting up some tunes. A big problem is that the illustrations are in black and white and the text keeps referring to blue and green loops, though I suppose if you have the program running when reading the book this will be obvious. The trouble I have is that I only have a G3 iBook and it really isn't beefy enough to run GB well so I can't get the best out of it (anyone want to give me a nice dual-processor G5 so that I can demonstrate my musical creativity?) Garageband is a lot like the Dance E-Jay series of programs, though the set of samples that comes with it for free is not as good. However it is undoubtedly a much more powerful program as this book shows. It is however firmly directed at the novice. Users with experience of other programs of this kind or with some musical knowledge will find that the writing style is a little condescending in places and downright grating in others. I got particularly annoyed by the repeated digs at people who don't like to use keyboard shortcuts. This is a shame because if you can get past this, there is much solid information in here that can definitely improve your GB experience. I know that I shall now make a more series go at using the program even if my machine is underpowered.

And then Hackers and Painters - literature, music and now the decorative arts. How cultured! For years now I have been banging on about how programming and painting are like each other. And now someone has written about the same thing in a chapter of a book. Damn, I should have written about it myself years ago. But that's life. Mostly I like this book, it is something to dip in to and find interesting nuggets. Partly of course this is because I agree with a lot of what Paul Graham is saying about the present state of computing and its future state. I'm not going to attempt to summarise any of it here. You should read this book. If you are a programmer and love programming, you'll enjoy it. Quite how you'd justify it as a business expense I don't know though. (Oh, and it's got a nice cover (Bruegel) and it's a hardback too).

Unixx Roxx

Digby Tarvin

Digby Tarvin writes:

A bit of trivia if you need to fill a corner of a newletter - while I was in Hong Kong in early December, on Dec 11th there was a debut album launch by a new rock band called... 'Unixx'.

From a South China Morning Post article:

Their name comes from a more classically Hong Kong perspective, says Prudent. "Unix is the name of an alternative operating system. Since we are an alternative band, it seemed to work well. I added an extra X for balance, and it looks cool."

(Sean Prudent, aka Yung Chun-yin, is lead vocalist). It apparently "combines elements of shoegazing, noisenik, post-punk, Brit-pop and grunge", but no, I haven't heard them..... So apparently Unix is cool..

Shell script puzzle

James Youngman

The shell puzzle has returned! In this issue it has been set by Pod, pod@herald.ox.ac.uk

Write a shell script which determines the endianness of the host on which it is running and sizeof(long).

The result should be produced in the form “Xnnn” where X is one of L or B, denoting little-endian or big-endian respectively. The ‘nnn’ should be the (decimal) number of bytes occupied by a ‘long’ on the system where the script is running.

For example, on the system I’m writing the article the correct answer would be ‘L4’. Your script must produce no other output on stdout, apart from white space.

The winning entry will receive an O’Reilly book. The winning entry will be decided as follows:

1. A compiler must not be used.

2. Entries which work on the greatest variety of POSIX-compliant systems will score best.

3. If rule (2) doesn’t produce a clear winner, the winning entry will be the one which occupies the fewest bytes.

4. Style, readability, obfuscation, and interesting technique are not criteria for winning, but may get you a (dis)honourable mention.

You can assume that on all systems the judges use, **/bin/sh** is a POSIX compliant shell (that is, you won’t be disqualified for assuming that Solaris has a POSIX-compliant **/bin/sh**). Your script may produce anything it likes on a PDP11-endian machine without risk of being marked down (the judges do not have access to a PDP-11 running a POSIX-compliant shell).

Entries should be submitted by email to puzzle@ukuug.org.

Contacts

Ray Miller
Council Chairman; Events; Newsletter
Oxford
Tel: 01865 273 200
ray.miller@ukuug.org

Mike Banahan
Ely
mike.banahan@ukuug.org

James Youngman
UKUUG Treasurer
Manchester
james.youngman@ukuug.org

Sam Smith
Website
Manchester
sam.smith@ukuug.org

Alasdair Kergon
Events
Reading
alasdair.kergon@ukuug.org

Alain Williams
Watford
alain.williams@ukuug.org

Roger Whittaker
Schools; Newsletter
London
roger.whittaker@ukuug.org

Newsletter
newsletter@ukuug.org

Jane Morrison
UKUUG Secretariat
PO Box 37
Buntingford
Herts
SG9 9UQ
Tel: 01763 273 475
Fax: 01763 273 255
office@ukuug.org