
 **NEWSLETTER**

School of Informatics, University of Edinburgh

Issue 2

Contents

| | |
|--|---|
| Support News | 1 |
| Storage Area Network for Researchers | 1 |
| The School Database | 2 |
| File Services News | 3 |
| Wireless Network | 4 |
| “Lightweight” DICE Machines | 4 |
| Redhat 9 laptop - a new model | 4 |
| Mail Team News | 5 |
| KB server area network upgrade | 5 |
| Web Team News | 6 |
| Renewing Credentials | 6 |
| Coming Soon | 6 |
| Newsletters on the Web | 6 |
| Users Hints and Tips | 7 |

► Support News

Support Office at BP Each of the sites has a Support service to help with computing queries and difficulties. At most sites you can find it in the Support Office. Buccleuch Place is the last site to get a Support Office. It'll be opening in BP2/1R19 when that room's free from its General Office duties. You can find the other offices in AT-5.06, FH-A05 and JCMB-2421. Incidentally, if you find an office unmanned you can always phone one of the other offices — all of the phone numbers are on all the support office doors — or use [the support form](#)¹.

Frontline Support Manager Less dramatically, I am standing down from the post of Frontline Support Manager, a job I've done for some years now. I'll be concentrating on my other duties from now on. Alison Downie is taking over the role: I wish her the best of luck, and I hope she enjoys it as much as I have done!

¹<http://www.inf.ed.ac.uk/systems/support/form/>

Thank You and Farewell to Andrew Woods

We are sorry to have to say goodbye to Andrew Woods, who will soon be leaving us to move to Yorkshire to join his partner. Andrew has very capably handled the user support at Buccleuch Place for a number of years now, getting it down to a fine art. He's coped very professionally, creatively and cheerfully with a greatly increased workload as Informatics has expanded and DICE has evolved. We would all like to thank him very much and wish him all the best for the future.

Chris Cooke <cc@inf.ed.ac.uk>

► Storage Area Network for Researchers

The School now has access to two SANs (Storage Area Network) for research-only use :-

- ❑ our own School SAN, based at KB. This has 7TB disk space, a 16 port fibre channel switch and a Sun NFS server.
- ❑ a substantial share in the SRIF SAN (managed by EPCC), which has over 150TB disk space, a 35Tb tape library and a SUN NFS server.

The SANs can be accessed by direct fibre connection or over the SRIF high speed network using either block level (iSCSI) or file level (NFS) protocols.

The School's Beowulfs will be connected directly to the School SAN, as can other research machines that need high speed local disk storage. Further details are available at <http://www.dice.inf.ed.ac.uk/groups/infrastructure/srif/>

If you have large scale data storage requirements, or a requirement for high speed data communications between School sites please contact support.

Alastair Scobie <ascobie@inf.ed.ac.uk>

► The School Database

What do the following things have in common?

- cd-students@inf.ed.ac.uk** mailing list for students taking Computer Design
- cisa-members@inf.ed.ac.uk** mailing list for members of the Centre for Intelligent Systems and their Applications
- <http://www.inf.ed.ac.uk/admin/itodb/mgroups/stus/ads.html> (tutorial group web page for Automated Reasoning)
- <http://www.inf.ed.ac.uk/people/Telephone.html> (the School telephone directory web page)
- The reports to all Boards of Examiners.
- Transcripts for students.

Well, as you may easily have guessed from the title of this article, the contents of each of these is derived from information stored in the School of Informatics database. They are of course only a tiny selection from the 350 or so email lists and approximately 14,000 web pages that are generated automatically from the database each day using *gurgle* (the Gnu Report Generation Language) written by Tim Colles.

Information Held

Many different types of information are held in the School Database; data are held on:

- Applicants for postgraduate degrees;
- Staff and visitors;
- Students, both undergraduate and postgraduate;
- Courses taught by the School (formally modules and classes);
- Course assessment structure;
- Student assessment results;
- Informatics Research Series and some other publications.

Much of this information is entered and maintained by the administrative and secretarial staff within the institutes and the teaching offices. They use a locally developed graphical user interface (written by Tim Colles) that is an X-client that can display on any workstation that runs an X-server. It allows them to maintain the

data through a collection of views of the database tables (in which a set of tables concerned with a particular aspect are presented in a more user-friendly way).

This ability to display a view composed of fields from several tables is extremely useful when one bears in mind that there are over 150 different tables that are currently used in the School Database.

Brief History of the School Database

The School Database had its origins in a departmental database established by Tim Colles in the former Department of Artificial Intelligence back in about 1994. The database engine then, as now, was Ingres. It gradually replaced even earlier standalone databases used for various administrative tasks. In 1997 the database service was hosted on *minerva.dai.ed.ac.uk*, a Sun Ultra 1 Model 140 with 448MB of memory and two 4GB external disks.

When the Division of Informatics was formed in August 1998 the DAI database became the Informatics database and enabled staff on each of the four sites of the Division to work together on a single database rather than relying on private local PC-based databases. I think that it can be argued that having this central multi-user networked database for the Division and now School of Informatics has been a very strong unifying force. In the summer of 2001 the service was moved to *hedwig.dai.ed.ac.uk*, a Sun Blade 1000 with 1GB of memory and three external 18GB disks attached to a UPS (uninterruptible power supply).

Recent Changes

In early August this year, after several months of preparation, the School of Informatics database service was moved to *libra.inf.ed.ac.uk*, a rackmounted dual 2.4GHz Xeon processor Dell PowerEdge 2650 with 2GB of memory, hardware RAID mirrored 36GB system and 73GB data disks, and dual power supply attached to a UPS. Being a DICE server all configuration is now controlled via LCFG so that its management becomes similar to that of the majority of machines and servers that we have in the School.

Further changes to some of the tables and many of the database reports have recently been or are in the process of being made because of the move to semesterisation and the Curriculum Project.

Ken Dawson <ktd@inf.ed.ac.uk>

► File Services News

The weekend of the 20th of August 2004 was an important one for the File Services team: it marked the transfer of the last home directories from the old legacy file servers to the new Informatics servers. It seems appropriate to provide Informatics users with a brief description of the new home directories setup.

The wide variety of ageing Sun hardware in use at the various Informatics sites has been replaced by 6 Sun-Fire 280R rack mounted servers, phoenix and roc at Appleton tower, sphinx and wyvern at Kings Buildings and pegasus and hippocampus at Buccleuch Place. These servers are connected to the network by a Gigabit ethernet interface and feature dual processors, dual system disks and dual power supplies connected to a UPS for maximum redundancy. These machines are configured via LCFG.

The storage attached to these machines provides an interesting illustration of the direction storage technology is taking. Each server has two 72G Fibre channel internal disks. Half of each disk is used for the operating system (one disk contains a mirror of the other disk constantly updated so that in the event of a disk failure, the server can continue to run using the remaining disk) and the other half is used for user data. It will be seen therefore that even without any external storage attached, each of the new servers has 72G of user disk space at its disposal.

Impressive though this is, it isn't enough to satisfy the voracious appetite of Informatics users for disk space and so each of the servers has some form of external storage attached.

Phoenix, roc and pegasus have what might be termed an older style of storage. Each has a SCSI3 JBOD (a great acronym standing for 'just a bunch of disks') containing 6 36G disks attached meaning that each of these servers has 288G available for user files.

Many of you will have heard of RAID in the context of disk storage. RAID (which stands for Redundant Array Of Inexpensive Disks) potentially offers many advantages in terms of reliability and performance but until relatively recently, the 'I' in RAID was somewhat relative since the disks used within the array were usually SCSI devices restricting large scale use of RAID to high end (and expensive!) applications. Within the last two years or so, a new class of storage device has appeared, the ATA based storage array which uses cheap ATA disks, normally found within PCs, rather than SCSI disks to make up the RAID array. ATA disks have traditionally been regarded as being slower and less reliable than SCSI disks but clever circuitry and the use of redundant disks goes some way to overcom-

ing this disadvantage.

We have two such devices hosting home directories within Informatics, a Nexsan ATAbay at Kings Buildings and a Nexsan ATAbest at Buccleuch Place. Each of these contains 14 250G ATA disks for home directory use which would seem to imply that each device has a total of 3500G available but in fact one disk is reserved as a hot spare and the equivalent of another is used by RAID for parity checks in each array so we only have 3000G available.

The other point of interest about these arrays is that they are fibre channel devices. Traditionally, storage devices have been attached to a single host over a distance of a few meters at most. Fibre channel allows multiple hosts to access the same storage device over far greater distances so that a server at say Appleton Tower could directly access the Kings Buildings ATAbay. At present, we are using this capability to share each array between the two servers located at the same site. There will be more about Fibre Channel in a forthcoming newsletter article about the SAN (Storage Area Network) Informatics is currently setting up.

This new storage allows us to substantially increase the default disk quotas allocated to the various classes of informatics users. The new quotas are:

- ❑ ug1, ug2 100M
- ❑ ug3 150M
- ❑ ug4, MSc 200M
- ❑ PhD, Staff No Quota

Note that although in theory members of staff and PhD students have no restrictions on the amount of disk space they can use, each staff and PhD account has in reality a 5G quota associated with it. This is done to prevent 'rogue' programs run by users in these categories filling up the entire disk as has sometimes happened in the past. Users who wish to have their quota expanded should contact support.

Please do not regard these quotas as targets to be met! It seems to be a universal law of computing that the disk space used will always expand to match the space available. We hope that this will not be the case in Informatics for several years to come and would encourage you to keep deleting unwanted files and archiving material for which immediate access is not required to media such as CD. Support will be happy to help you with any questions you might have about how to go about this.

Craig Strachan <cms@inf.ed.ac.uk>

► Wireless Network

As announced in the previous newsletter, the “informatics” wireless pilot service, which has operated at BP and KB, is to be phased out. This will now take place by Christmas, to allow time for any remaining Linux laptops to be upgraded.

We are instead joining the University “central” service, which as well as being much more secure will provide both staff and students with uniform access across many more areas of the University. Already 2/3 of our access points operate on the “central” network, and we intend to add to these to improve coverage of those Informatics areas where reception is currently poor.

DICE managed laptops which have been upgraded to RedHat 9 have a simple way to connect to the rest of Informatics: once the machine has attached to the “central” network, the command “renc” should be used to renew your authentication credentials in the usual way; and once that has been done the command “infvpn” will bring up an encrypted tunnel, making the machine appear as though connected directly to an Informatics network port. There is no need to register with EUCS to use this facility.

Users of self-managed laptops currently do have to register with EUCS, and either use the web-based authentication mechanism or install the VPN software. Users of the central VPN service are able to access their Informatics files and print to Informatics printers by connecting to our samba service. See <http://www.inf.ed.ac.uk/systems/network/Wireless.html> for more information.

As an alternative, we are evaluating the OpenVPN package which we already use on DICE with a view to providing simple instructions for its use on self-managed Linux and Windows machines.

There are special provisions for visitors to use the “central” wireless network. The University is participating in a pilot “location independent networking” project by which visitors from participating institutions can authenticate via their home University. For those not covered by this scheme, the support team can issue temporary logins as necessary.

Please note that if your laptop is set up for any kind of peer-peer wireless networking you **MUST TURN IT OFF** before operating around the University. Failure to do so may break the wireless network for other users, even if you yourself are not using it.

See <http://www.inf.ed.ac.uk/systems/network/Wireless.html> for more details and links to wireless-related pages.

George Ross <gdmr@inf.ed.ac.uk>

► “Lightweight” DICE Machines

Feedback on the recent DICE strategy review confirmed the existence of a strong demand from many users for greater control over the configurations of their desktop machines, without the difficulties and effort required to maintain a completely “self-managed” system. It had always been our intention to provide this, but we have raised the priority in response to the received comments.

We now have a proposal for a compromise configuration which provides much of the freedom of a self-managed machine, with many of the benefits of a “centrally-managed” one. We have a number of prototype machines running, and we believe that most of the technical issues can be overcome without undue effort. However, there are significant issues concerning the level of support that would be available, and the amount of effort that could potentially be diverted into supporting a very small number of users. There are also issues concerning some restrictions that are desirable for security reasons.

We definitely want to ensure that the proposed solution meets people’s requirements, before committing to the work involved – this includes an agreement on the support and security issues which are likely to be more important than the technical details. A paper describing the proposal is currently being discussed by the COs and will be made available for comments very soon. If the feedback is generally positive, and the support issues can be resolved, then this will be presented to Computing Committee for approval and should be available in a trial form early next year.

Paul Anderson <dcspaul@inf.ed.ac.uk>

► Redhat 9 laptop - a new model

For some time now, we have been unable to buy any new laptops to run DICE as we couldn’t find a model that would run DICE satisfactorily. I’m glad to report that we have now found a couple of models that we are happy to support - the IBM Thinkpad T41 and T42. We are hopeful that we can also support the IBM Thinkpad X31 (a lightweight model) shortly.

Alastair Scobie <ascobie@inf.ed.ac.uk>

► Mail Team News

The long promised new mail server has been installed and has replaced the old Redhat 7 based machine. The new machine has two 2.8GHz Xeon Processors; 2GB of RAM; two redundant power supplies powered through a UPS; and two hardware RAID1 containers, one for the Operating System, the other for the mail data. All this should help ensure a high availability of the service.

IMP

The new machine also has a new look to the web interface at <http://mail.inf.ed.ac.uk/>, which is more configurable and has better search and address book functions.

Spam

The University's Computing Services are constantly improving their anti-spam measures, which means we benefit, as they scan all our incoming external mail for us. For more details on what they do, see their page at: <http://www.ucs.ed.ac.uk/fmd/unix/docs/mail/junk/index.html>

Throwing Spam Away: One thing it appears people are not doing is checking, and clearing out their "suspect_spam" folders. These are beginning to take up a significant amount of disk space. We will be implementing mail quotas to stop the mail disk filling up with junk mail.

Legacy DCS Mail

The legacy DCS mail server is the last of the old department servers in use. The old DAI and Cogsci mail services are now running on the Informatics Virtual Mail Relay, and the DCS mail service will be moving to it soon.

All the Virtual Mail Relay service does is redirect mail addressed to "oldusername@dai.ed.ac.uk" to (normally) "newusername@inf.ed.ac.uk", but it may also redirect to external email addresses, otherwise it just rejects the mail with "user unknown".

The process of identifying all existing @dcs.ed.ac.uk addresses and where the mail should be forwarded to has begun. This will take a couple of weeks, while any old DCS .forward or .procmailrc files, that are still in effect (most are being ignored), are checked for a suitable forwarding address. Once this has been done, the mail service will be switched to the VMR and the last DCS service can be turned off.

This should not affect any current DICE users. However, any users who have managed to bypass the existing redirection will find their .forward and .procmailrc files will become ineffective after this point.

Mail Logs and Privacy

In case users are not already aware, every piece of mail that travels through the mail servers is logged. Not the actual content, but the date, time, sender, recipient and in most cases the subject of the mail are logged to a file. This helps us trace problems when they occur, and trace any abuse should we need to. Logs are kept for 2 months, and are only available to computing staff.

To be able to carry out their duties, the mail team members may also occasionally see the contents of mail messages and mail folders, particularly if requested by a user to sort a specific mailbox problem (in which case their permission will be sought), or to stop and track down mail loops. In this second case it usually isn't possible to contact the user and so we just have to use our best judgement.

At all times users' privacy is paramount. This is not a responsibility the mail team members take lightly.
Neil Brown <neilb@inf.ed.ac.uk>

► KB server area network upgrade

Over the weekend of 21st and 22nd August a major upgrade of the networking equipment in our KB server area was performed. The purpose of this was twofold:

- ❑ to replace the shelving carrying the network terminations and switches, first installed around 1992, with larger standard racks, to provide space for expansion and to allow for the installation of more modern equipment; and
- ❑ to upgrade the switches with new Gigabit-capable ones, to provide faster connections for many KB-based services.

The operation went off remarkably smoothly, due in no small measure to the efforts of the KB technicians, both in their meticulous preparatory work as well as over the weekend itself. As a result, almost all services were restored by the Sunday afternoon, with only a very small number being left over till Monday.

The network team will shortly begin reviewing the requirements for Gigabit connections at Appleton Tower, Buccleuch Place and Forrest Hill.
George Ross <gdmr@inf.ed.ac.uk>

► Web Team News

The switch to hosting the legacy DCS web pages, www.dcs.ed.ac.uk, on DICE managed hardware went relatively smoothly. There are still some broken pages, mainly Perl CGI scripts, that just need their path to Perl updated to `/usr/bin/perl`. Please check your pages and scripts, if you still expect them to be accessed by the outside world.

Since the South Bridge fire, the old DAI web pages have been hosted on DICE hardware, so this just leaves the legacy Cogsci web services to be moved. They are currently hosted on an old Solaris server, but a new DICE RH9 server has been allocated for the task. Roger Burroughes (roger@inf.ed.ac.uk) is currently running a pilot version of these services on the new machine, and those affected should have been contacted.

Remember that the goal is to eventually freeze the content of the legacy www.dcs.ed.ac.uk, www.cogsci.ed.ac.uk and www.dai.ed.ac.uk sites. And that all active web content should be hosted on a [.inf.ed.ac.uk](http://inf.ed.ac.uk) URL.

Web Logs

One common complaint is the lack of access to web server logs. If we could just let people see them, then we would. It would relieve the web team of the task of acting as intermediary when responding to requests as to why someone's CGI script is not working. We are working to [the JaNET guidelines²](#), and unfortunately the Data Protection Act and Regulation of Investigatory Powers Act which means that we cannot allow unrestrained access to the logs because:

- the logs contain personal information
- the URL contents are not all traffic data, and so constitute interception.

We are looking at ways of providing a sanitised version of the logs which people will be able to access. For the record, we keep HTTP access logs for 6 months and record the following information:

- the IP address of the machine accessing the URL
- username used to access the page (if provided)
- date and time of the request
- the request issued, eg "GET /~neilb/private-picture.jpg"

²http://www.ja.net/documents/gn_logfiles.pdf

- The URL of the page the request came from, eg `http://www.otsosecret.com/`
- And the browser the person is using, usually revealing their OS too.

We also log any errors that may be generated by the web server and user CGI scripts, eg

- "Unable to open file: /public/homepages/mr.prof/secret.passwords"

The logs are kept so we can trace abuses of the system (if required), debug problems and monitor the overall performance of the web services.

Neil Brown <neilb@inf.ed.ac.uk>

► Renewing Credentials

All users should now use the "renc" command to renew their Kerberos credentials and not "kinit". This is primarily because "renc" also renews **KX509** credentials which we are increasingly using to provide seamless access to authenticated web services. There will be more on kx509 in the next newsletter.

Tim Colles <timc@inf.ed.ac.uk>

► Coming Soon

The next newsletter will feature an article on RT version 3. This new version of RT will enable users to follow the progress on their own new "tickets".

Do let us know if there are any subjects you would like to see covered in the next issue.

Morna Findlay <morna@inf.ed.ac.uk>

► Newsletters on the Web

Dice newsletters are available on the Informatics web pages at: <http://www.inf.ed.ac.uk/systems/newsletters/>

Morna Findlay <morna@inf.ed.ac.uk>

► Users Hints and Tips

In which fellow users present tricks of the trades that others may find useful.

To contribute hints or tips to the next newsletter, please contact the docs-team@inf.ed.ac.uk.

USB Memory Stick Under Linux

If you use a memory stick frequently, then there is an alternative to the mount/unmount method in the FAQ. If you create a file called `~/ .mtoolsrc` with the following content:

```
drive d: file="/dev/sda1" exclusive
mtools_skip_check=1
```

then you can access the stick directly as the `d:` drive using the `mtools` commands, e.g.

```
mcopy myfile d:
Chris Walton <cdw@inf.ed.ac.uk>
```

Indenting in xemacs

If you find the xemacs default indent size of 4 to be too big, then you can change it to 2 spaces by adding the following to your `~/ .xemacs/init.el` file. This affects any language mode based on the xemacs c-mode (e.g. Java):

```
;; C-mode Indenting
(setq c-default-style "user"
      c-basic-offset 2
      c-syntactic-indentation nil)
```

Chris Walton <cdw@inf.ed.ac.uk>

Large `~/ .xsession-errors` Files

Depending on the X11 setup, some programs may generate a very large `~/ .xsession-errors` file. In order not to fill the home directory, link this file to `/dev/null`.

Of course, the real solution is to check the source of errors but as things work well otherwise this may be a waste of time.

Yuval Krymolowski <ykrymolo@inf.ed.ac.uk>

Image Extraction

- `pdfimages` is a useful program for extracting images out of pdf files, available as `/usr/bin/pdfimages`

- `psrip`³ is a useful perl script for extracting images from postscript files.

(Thanks to Stephen Eglen for these tips)

Chris Williams <ckiw@inf.ed.ac.uk>

LaTeX Tips

- Problems with funny spacing around ligatures (ff, fi, ffi &c) when converting LaTeX documents to PDF? Try `ps2pdf14` or `ps2pdf13` rather than `ps2pdf`.
- Want to (La)TeX an extra line into a conference paper? `\looseness=-1` may be your friend.

David Sterratt <david.c.sterratt@ed.ac.uk>

³available from www.ctan.org/tex-archive/support/psrip/?action=tex-archive/support/