

dice NEWSLETTER

Winter 2007

Welcome to the winter edition of the DICE newsletter. In this issue you will find details of the SL5 rollout, advice on using external services, the move to staffmail, and other useful snippets.

Contents

- New Building Update
- Scientific Linux 5 on Servers
- SL5 Dual-boot machines in Open Access Labs
- Eddie
- GPFS
- Guidelines for using External Services
- Staffmail
- Visitors and Staffmail
- Extra External Access Machine
- Web Project News

New Building Update

With the move to the Forum only months away, detailed planning on how we migrate the School's computing facilities to our new home is now underway.

The expectation is that, in order to expedite the move, ALL computing staff will be heavily involved in preparing the building for occupation, assisting in configuring desktops for their new location, and moving School servers from their existing locations at Kings Buildings, Forrest Hill and Buccleuch Place. We will retain a presence in the Kings Buildings server room to give us disaster recovery capability.

This work will have to be performed in addition to our existing commitments, so it is inevitable that progress on development projects will be affected and response to

user queries and requests be slower than usual.

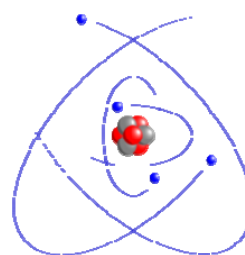
Particularly affected will be our ability to respond to last minute requests.

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

Scientific Linux 5 on Servers

We generally aim to upgrade all DICE machines to the latest release of our distribution of Linux (currently Fedora Core) once every year or so, in order to obtain support for new hardware, to obtain updated software, and - importantly - to ensure that we have continued access to security fixes.

However, the fact that each release of Fedora Core has a relatively short support window, combined with the effort and disruption involved in upgrading all of our server machines in this way, has made us think again. So, for our servers only, we have decided to move to using Scientific Linux.



Scientific Linux - currently at release 5 - is a freely-available distribution of Linux originally developed at CERN and Fermilab.

It's closely based on Red Hat Enterprise

Linux, and is becoming increasingly popular within the scientific community. Importantly for us, official support for Scientific Linux 5 runs until at least March 2010.

Since SL5 is closely related to Red Hat Enterprise Linux 5, which itself is based on Fedora Core 6, in practice SL5 is not very different from FC6.

Over the summer we ported our system configuration software (LCFG) to SL5 in a joint project with EPCC. Over the next few months, we plan to move our servers to SL5. End users should notice no changes in service as a result of these upgrades.

SL5 has been adopted by some of the Science & Engineering schools as their first agreed common Linux platform, and we hope that the commonality will benefit everyone. The ECDF will move to SL5 in the Spring, and our use of the same platform will make it easier to grid our servers with ECDF.

For more information on Scientific Linux, visit <https://www.scientificlinux.org/>

Ian Durkacz <idurkacz@inf.ed.ac.uk>
Infrastructure Unit

Chris Cooke <cc@inf.ed.ac.uk>
Managed Platform Unit

SL5 Dual-boot Machines in Open Access Labs

There are three Open Access Labs locations providing dual boot XP and SL5 Linux: Kings Buildings Centre Level 2; Hugh Robson Building Lab A; and Appleton Tower Cafe and Mezzanine floor.

At the XP logon there should be an option to restart the machine - do this and watch for the GRUB menu to select Linux.

On logon a stub native home directory of 10Mb size is mounted using loopback mount. Your dot files can happily reside in here and be preserved across sessions, as it is stored back to the server (compressed) on logout.

In this stub directory your standard Windows home directory is accessible at My Documents. This is an SMB type mount so support for dot files and fifo etc. is limited.

Informatics users may access their AFS home directories by browsing to /afs as normal.

The system monitors your session and will auto-logout on idle. Back at the Linux logon prompt, a timeout of 5 minutes can elapse before the machine will reboot back to Windows XP.

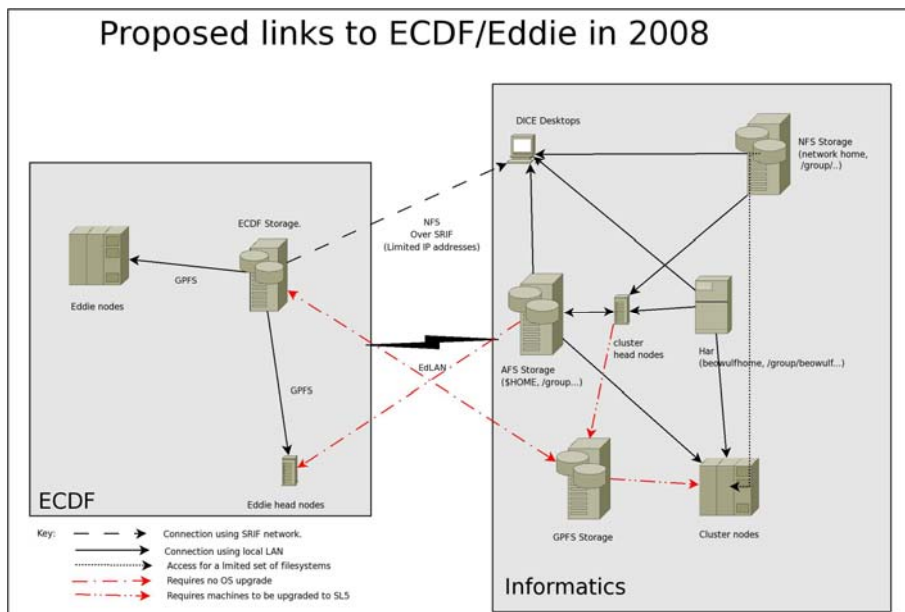
Printing is not fully available at the current time.

Information Services

Eddie

Eddie is a beowulf cluster run by the Edinburgh Compute and Data Facility (ECDF) as a central service for researchers across the University. It consists of 512 processors made up of 128 nodes, each with 8 cores. It is backed by a large datastore (275Tb) which is accessible via a fast parallel network filesystem called GPFS (for more details see GPFS article in this issue).

Jobs are submitted to the cluster using SUN's gridengine batching system which also allows scheduled interactive sessions on a select number of nodes. The nodes themselves run Scientific Linux version 4 (SL4) which is a rebuilt, rebadged version of Redhat Enterprise Linux 4 and is based on FC3.



Other operating systems can be installed on the cluster nodes and it's expected that the default OS will be upgraded to SL5 sometime in the New Year.

Eddie is not just used for parallel jobs requiring interprocess communication, it is ideally suited to problems which can be broken down into a large number of independent tasks, or where there is a wish to run a program a large number of times with different parameters or against different datasets.

The use of GPFS gives access to terabytes of data at higher transfer rates than NFS or AFS. If you have a problem that requires large amounts of CPU time or handling large amounts of data you may want to look at using Eddie.

There are two levels of service on Eddie: 25% of compute resources are offered on a "baseline" service which is available equally to all research groups based at the University; the remaining 75% is available on a charged guaranteed basis, which allows researchers to buy time on the clusters at a flat rate (charges are listed at: <http://www.is.ed.ac.uk/ecdf/ecdf-charges.shtml>).

The School is currently a heavy user of Eddie - and we have dedicated disk storage attached, based at ECDF.

We are currently working to improve accessibility of filestore between ECDF and School systems and the School's AFS filesystem has recently been made available on Eddie's login nodes and we are planning to allow NFS/Samba access to the School's filesystem at ECDF. Finally, we are working with ECDF to set up shared GPFS filesystem between Eddie and the School's beowulf clusters.

Accounts for Eddie can be obtained via ECDF and will use your EASE account details.

(<http://www.is.ed.ac.uk/ecdf/accessform2.shtml>)

More information about Eddie/ECDF can be found at <http://www.ecdf.ed.ac.uk>

External links:

GPFS:
<http://www-03.ibm.com/systems/clusters/software/gpfs/index.html>

Gridengine:
<http://gridengine.sunsource.net/>

Iain Rae <iainr@inf.ed.ac.uk>
Research and Teaching Unit

GPFS

We have been considering a parallel filesystem for the School clusters for some time. Parallel or clustered filesystems increase performance by "striping" data in a filesystem across multiple servers. Unlike traditional network filesystems like NFS or AFS, it is possible to increase the filesystem's performance by adding more servers.

In operation, a parallel filesystem is not unlike a RAID system on an individual server. Equally, like RAID, it is possible to store multiple copies of data for high availability.

Given the University's investment in Eddie and knowing that our cluster users were likely to be heavy users of Eddie, we decided to use the same system, IBM's General Parallel File System (GPFS). Unfortunately, GPFS was not available on our platforms (FC5 or Solaris).

However, given that a later version would be available to run on Scientific Linux 5 - which we would be using for servers - it was decided to adopt SL5 for the clusters.

We are aiming to provide the existing cluster filespace under GPFS on the School clusters whilst still providing NFS access across DICE. We are also working with ECDF to share GPFS filesystems between Eddie and the School clusters, this would remove the need to copy data from location to location for people using both systems.

We are currently running a small test cluster (4 gx240s) to gain experience of using GPFS, and in particular to test sharing filesystems between the School and Eddie. Based on the results from this we will have a better idea of what hardware may be required and how it would be best deployed.

Independently of this GPFS work we will be upgrading the clusters to run SL5 rather than FC6, users of the cluster should be aware that the software installation on the nodes may drift away from that installed on their Desktops.

Because of the move to the new building, and the work likely to be involved, it's not possible to give a firm timetable for this.

GPFS is currently available on AIX5L, RHEL4 and RHEL5. It will not be possible to run GPFS on Fedora. Because of network bandwidth and security implications we will not be offering GPFS access to desktops.

For more information:

IBM Page: <http://www-03.ibm.com/systems/clusters/software/gpfs/index.html>

Wikipedia page:
<http://en.wikipedia.org/wiki/GPFS>

Local test cluster (available soon):
http://www.dice.inf.ed.ac.uk/units/research_and_teaching/projects/GPFS/index.html

Iain Rae <iainr@inf.ed.ac.uk>
Research and Teaching Unit

Guidelines for Using External Services

There is an increasing use across the University of externally-hosted applications and services for teaching and research purposes. Examples of such services include blogs, wikis, project and source code hosting, social space and virtual worlds.

The paper below has been written by Information Services and provides useful guidance on many of the issues which should be considered before using such services for University purposes.

While there are plenty of perfectly valid reasons for using externally hosted services there are also downsides to be aware of:

- Security of Service Provider
- Confidentiality
- Ownership of Data
- Security of Data
- Performance
- Reliability
- Support
- Single sign-on
- Lock-in
- Longevity
- Functional Stability
- Legislative and Regulatory Issues, including Accessibility Legislation and University Regulations

The paper includes a very useful checklist of things to look for when signing up for external services. It also provides a comprehensive template for risk analysis - which can either be used casually or more formally depending on the criticality of the service you are considering hosting externally.

<http://www.vp.is.ed.ac.uk/pub/jsp/polopoly.jsp?d=290&a=1245>

Tim Colles <timc@inf.ed.ac.uk>
Research and Teaching Unit

Staffmail

With the University's current drive to make more use of central services, rather than School/Colleges implementing their own, we are considering ceasing to provide our own mail service for staff and visitors (as we have already done for students), and make use of the central staffmail.ed.ac.uk service instead.

The University already provides all members of staff with a staffmail.ed account, but generally Informatics staff have not been making use of this.

With a view to the day when we may be all asked to use staffmail.ed, rather than our own mail.inf, all computing staff (and a few other volunteers) have been using staffmail.ed for months, and have not had any problems following their move. Recently all new members of staff have also started using staffmail.ed, rather than mail.inf, to store their mail.

Using staffmail.ed does not mean changing your email address, all that it means is that your mail is stored on central servers (imap.staffmail.ed.ac.uk) instead of our own servers (mail.inf.ed.ac.uk). Mail sent to your username@inf.ed.ac.uk address is forwarded to your INBOX on staffmail.ed.ac.uk, and all you have to do is update your mail client to access your mail folders from the staffmail server instead.

The staffmail.ed service supports features our mail service does not, such as:

- the ability to share specific folders with other users
- mail folders can contain messages, as well as further mail folders
- better support for accessing your mail via mobile devices
- much more disk space

For current plans, more information, and for those wishing to become early adopters of staffmail, please see the following web pages for instructions on the steps

required, and contact the services unit (via the support form) if you need help/have any questions.

www.inf.ed.ac.uk/systems/email/staffmail/

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

Visitors and Staffmail

The University's Information Services (IS) do not provide mail accounts for all visitors. The categories of visitor that WILL be given staffmail accounts by IS are:

- University Honoraries
- University Fellows
- University Staff not on main payroll
- Temporary Staff
- Postgraduate Workers
- Unregistered Students (credit)

Many of our visitors do not fall into any of these categories so Institute Administrators will need to ask such visitors for a suitable e-mail address that they are already using so that we can forward mail sent to username@inf to them. Hopefully this will be of benefit to them as they'll no longer have to remember to check their @inf INBOX as well as their normal INBOX while visiting Informatics.

Alison Downie <alisond@inf.ed.ac.uk>
User Support Unit

Extra External Access Machine

There is an additional external-access ssh machine. The host name is ssh2.inf.ed.ac.uk and it is located in Kings Buildings. Should there be further power problems in the central area, users may be able to log in remotely to it and, if their home directory is not affected by the loss of power in the central area, should be able to do something productive.

Web Project News

As most of you will hopefully be aware, the University is working towards a unified website, a "website without boundaries", that will offer a consistent "look & feel" and editorial style. This move is being managed by the Website Development Project, which includes in its aims:

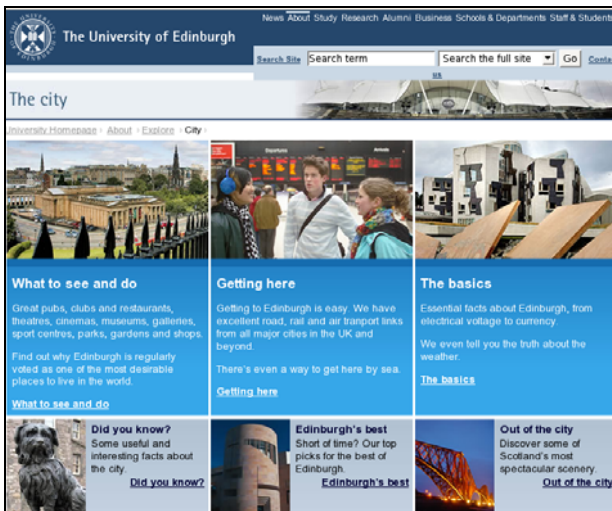
- To design from the user's perspective: with particular concern paid to external audiences, especially new or infrequent visitors to the site and those unused to the languages and structures of Higher Education.
- To satisfy our corporate values, the expectations of our target audiences and the technical and "look and feel" requirements

The centralisation of web content will also cut down on duplication of effort, and the re-use of content (referencing the same content from different pages) will make sure that information doesn't get out of sync.

In addition to delivering a new University website, the Project will develop a comprehensive staff contact details database. The University Contacts Directory (UCD) project and the Group Membership System (GMS) project will deliver a reliable 'people finder' service and dynamically-generated staff contact web pages.

The first phase (proof of concept) University web site is now up and running, and sample pages are available for inspection at

<http://www.uws.mis.ed.ac.uk/UoE/about/explore/city> (it will not go live for some while yet, as we have to wait for web content to be ported from all participating Schools).



To help with the preparations for migrating web pages to the new service, the Web Development Project are offering "Writing

for the Web" courses that will outline the 'Look & Feel' and the editorial style favoured by the University - everyone is eligible to attend such courses, and is encouraged to do so. To find out more, go to the "Event Booking" page on MyEd (<https://www.myed.ed.ac.uk/>) and look for the next "Writing for the Web" course.

For further information on all aspects of the Website Development Project, see <http://www.ed.ac.uk/websiteproject>

Roger Burroughes <roger@inf.ed.ac.uk>
 Rollout Co-ordinator
 (Web Development Project)
 College of Science & Engineering