

New Director VeRSI

On Friday 2nd May the e-Research Office organised an afternoon tea to welcome Dr. Ann Borda, the newly appointed Executive Director of VeRSI. Ann has extensive experience in the field of further education and research, having worked in senior roles in both the UK and Canada.

In the UK Ann was responsible for identifying requirements and facilitating the adoption of a national support infrastructure for researchers, including Grid services, data facilities and other collaborative tools for research.

Ann's previous roles have included managing the Open Source Software Advisory Service in the UK and conducting academic research on informatics, human-computer interaction and collaborative technologies.



Contact us:
e-Research Office, Suite 24
RD2, R&D Park
La Trobe University
Victoria 3086
AUSTRALIA
Tel: +61 3 9479 5140
Fax: +61 3 9479 1237
<http://www.latrobe.edu.au/e-research>
Email: eresearch@latrobe.edu.au

Melbourne University - VeRSI Coffee Network

These events are held weekly on Tuesdays at 11am, in the VeRSI Melbourne office, ground floor of the Brownless Biomedical Library, University of Melbourne. The Coffee Network provides an opportunity for collaboration between parties with an interest in e-Research and a chance to network and discuss topics with colleagues.

The invitation is open to all at La Trobe; please refer to the link below for topics that may be of interest:

<http://www.versi.edu.au/versi/melbcoffee.html>

Remote visualisation

The demand for high graphics capability from the gaming industry has completely changed the research graphics landscape, with relatively cheap and very high performance graphics processing units easily available. Very large and detailed screens are also tumbling in price.

However, a growing problem in the e-Research space is how to analyse and visualise results from extremely large data sets. With databases and data files growing from gigabytes to terabytes it is not viable to move them around ad hoc from central storage to workstations. Transferring a 1 terabyte file over an unloaded gigabit Ethernet link would still take

over 1 hour. With visualisation of simulations and models from large data sources now becoming standard practice, how do multiple people at their workstations do their work without lots of sitting around?

A solution made available from Silicon Graphics Inc. (SGI) is to put the graphics processing cards in the analytical clusters that are next to the data, and sending the graphics to the remote workstation over the network. Even the highest resolution graphics screen running at 25 frames per second is viable over high speed links, unlike moving the data sets.

(<http://www.sgi.com/products/visualization/remote/index.html>)

Another solution is shown by the Optiportal, recently installed at University of Melbourne, which provides for a 96 million pixel video wall linked to overseas processing (i.e. the pixels are pushed over the network, rather than the data that must be analysed for the visualisation, which maybe much larger, thus difficult to move, and require computing power not available locally).

A further example that anyone can try is Microsoft's HD View. This is a way for a remote user on the web to view massive images of about 4 gigapixels. Zooming in almost endlessly on an image of that density is fascinating, as the whole versus the atomic parts can be viewed almost at once.

(<http://research.microsoft.com/ivm/HDView.htm>)

Recent Seminars highlight need for e-Research techniques

Two seminars held recently have highlighted a combined application of different "omics": Systems Biology. This field brings together

genomics, transcriptomics, proteomics, metabolomics and so on. It is an attempt to reverse the reductionism often seen in experimental and biochemical analysis of living systems by bringing together different analysis of the cell's biochemistry into a larger picture of cell biology. The two seminars, one on the value of systems biology and techniques for drug development and testing and the other on mathematical models of systems biology.

Both seminars showed that modern analysis of cell genetics and biochemistry can involve thousands of analyses through micro-arrays, mass spectrometry, robotic and automated comparisons with existing libraries of DNA and RNA fragments and so on. The modelling of the results, as well as the analytical methods, now involve massive databases, comparison amongst databases and huge results files. Visualisation of these results is also critical, because of the high complexity of the protein and other structures.

Large data storage, access to databases, database creation, high performance computing and high performance visualisation are now required, and at a scale not affordable by any single group. Fortunately, much of the equipment is generic across all disciplines, so central infrastructure with the required security, access, scale and reliability is perfectly suitable.

Central Data Storage

As will be known by most of you now La Trobe does not have dedicated central research storage. There is increasing difficulty in funding major infrastructure, and yet e-Research, a key method of the University increasing its research capacity and output, requires it. In the absence of this provision,

departments/schools are purchasing small high-capacity drives to use as shared and/or networked storage. This is a false economy and dangerous in the long-term.

Cheap NAS or larger capacity desktop drives are not adequate for a shared, high capacity service. They are unreliable (mean time to failure tests show high chance failure within a year of operation under server conditions), have low performance and associated facilities for backup/recovery are rarely provided.

A white paper on costs and appropriate uses of storage at various levels of the University hierarchy (University, Faculty, School) is available on the e-Research web-site (<http://www.latrobe.edu.au/eresearch/reference.html>).

The paper clearly shows that high capacity and secure data requirements are best handled in properly managed and environmentally secure central facilities, and immediate, scratch data is best handled on the desktop, with little in-between making much sense.

So, next time you need a few terabytes of space think about where the money to pay for that should go.