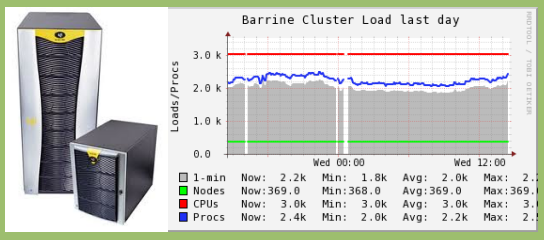
Research Computing @ UQ



1 August 2013 Issue 1

The Research Computing Centre supports collaboration to facilitate discoveries in science and engineering, humanities, and social sciences, through advanced computation, data analysis and other digital research tools.

Message from the Director

I am delighted to have joined the University of Queensland as Director of the Research Computing Centre (RCC), and look forward to building on its achievements over the past few years.

RCC plays a crucial role in UQ's research infrastructure by working with researchers to better leverage new computing technologies. Many of these systems are at the leading edge but are not always easy to use. On the other hand, they promise significant benefits for UQ researchers. Our job is to work closely with you to realise these benefits.

Over the past 4 months I have been moving around the University to meet people and discuss how we can work more closely. In response to the feedback, I have been reformulating the RCC structure and engagement model, and have been testing this on some key stakeholders. Over coming months we will be releasing the details and seek further input. Overall, the model develops and fosters a unique combination of deep technical IT knowledge with domain expertise, supporting collaboration and support with our clients. We look forward to working with you.

David Abramson
Director

A new home

RCC has a new home in Axon Bld, Level 5.

We are co-located with the Queensland Cyber-Infrastructure Foundation (QCIF) and the Research Data Storage Infrastructure project (RDSI). Together these organisations offer a unique portfolio of expertise to UQ, Brisbane and Queensland.

We look forward to seeing you in Axon 5 in the near future.

CLOUD COMPUTING IS IT HOT AIR?

Just what is it?

Cloud computing is a hot topic - everyone is talking about it and it seems to be a panacea for all ills. Let's take a quick look at the key ideas and outline a little about local plans at UQ.

No single technology in Cloud computing is new. High performance compute clusters with multiple computing cores underpin modern computing infrastructure. High speed networks have been deployed for many years, particularly in the research environment. So called "thin clients", mobile devices and enormous data archives and databases have also been around for some time. Virtual machine technology was invented in the 1960s and even pay-as-you-go computing has been attempted over the years.

However, Clouds leverage all of these things, and the combination makes them particularly attractive for both research and business. Putting it simply, a Cloud is a place where you can put your data, perform your computations and interact with others. You may not even be aware where the Cloud is located - it is available 24/7 from anywhere that is networked. It means that organisations can shift Capital Expenditure (so called CAPEX) to Operational Expenditure (OPEX), and this makes it easier to budget for infrastructure needs over the long term. Some Clouds support "pay-as-you-go" meaning you can scale your infrastructure needs as demand changes.

Unfortunately, Cloud computing is also the latest hype, and has been promised as a solution to every ailment known to mankind. This makes it difficult for researchers to decide what's in it for them, and what the appropriate choices are. At RCC, we have built a team of experts in Cloud computing, and are working with our partners (QCIF, NeCTAR and RDSI) to provide simple and effective solutions in research.

Cloud computing really embodies a few ideas:

- On demand and scalable access
- Application services
- Location transparency
- Virtual machines

On demand and scalable access means that users should be able to scale resources dynamically without the need to purchase hardware. If you need 1 compute core today, but 1000 tomorrow, it should be as simple as filling in a web form and requesting these. Amazon calls this "elastic computing" and promise unbounded resources, although clearly it is ultimately limited by the size of their cloud. But, this can be very large.

Application services means that the cloud provider might expose high level services that meets some business (or research) need. This currently works well for tools like e-mail and calendar, but the

number of applications of value to researchers is growing all the time; some of these will be very domain specific.

Location transparency means just that; you don't know where your services are being sourced. Whilst this is liberating and simple for many users, there are clearly applications where physical location is important. For example, whilst you may not care where your e-mail is handled, you might want data to be held in a particular data centre. Thus, most cloud providers allow you to indicate a *locale* that restricts the actual domain. (A lot has been said about legal issues such as the US Patriot Act and censorship laws. This is a very tricky area that suggests the legal profession will ultimately have to deal with virtual locations rather than laws that govern physical domains. But that won't happen overnight!).

Virtual machines are a technical trick that simplify the deployment of new services, and are part of providing rapid scale up and scale down. Putting it simply it means you can build a software stack locally and deploy it on a many physical machines as required.

Commercial clouds also include pay-as-you-go economic models. This means that you pay only for the infrastructure you use - whether that's machine time or data storage space. Companies such as Amazon are happy to take your credit card and provide services according to a charging schedule. This is incredibly useful for business applications, but not commonly used in research circles.

What's happening at UQ?

UQ researchers have access to data storage via the Research Data Storage Infrastructure (RDSI) project to compute services via the National eResearch Collaboration Tools and Resources (NeCTAR) project. Researchers can get virtual machines for their own use.

RDSI and NeCTAR are Australian Government projects conducted as part of the Super Science initiative and financed by the Education Investment Fund. The Queensland node is called QCloud, and its main data centre is currently located at The University of Queensland's St Lucia campus and operated by QCIF and the RCC.

QCloud is an integral part of new, national research infrastructure, leveraging data collections stored in the local and national RDSI nodes, and integrating with access to Queensland-based HPC facilities and specialised cloud data services. It will provide large scale storage for UQ researchers and computational capacity suited to Cloud computing.

Virtual Machines/Research Cloud

NeCTAR offers eResearch infrastructure in four key areas:

Research Cloud
Virtual Laboratories
eResearch Tools
A secure and robust hosting service – the National Servers Program

The research cloud is a computing resource for all Australian researchers. If offers virtual machines where researchers can develop and deploy applications and collaborate in a uniform environment with controlled sharing of data..

Research Data Collections

QCloud makes available the RDSI Research Data Services (ReDS) data collections. Researchers are welcome to apply to QCloud for RDSI ReDS storage. The aim of ReDS is to fund infrastructure to store research data collections of national merit and collections of interest for future research. Collections that do not qualify for ReDS storage will be considered for an allocation of locally approved data storage for their research data collection, including collections still in development.

Want to know more?

See the QCIF Cloud information page.

Need High Performance Computing?

The Research Computing Centre has access to a number of high performance computers - both at UQ and nationally. We would be happy to discuss your HPC needs and help secure access on the most appropriate resource.

Specifically, UQ has access to Barrine (at UQ) and the National Computational Infrastructure (NCI) at ANU.

Barrine is a High-Performance Computing cluster located on The University of Queensland's (UQ) St. Lucia campus. It consists of 384 compute nodes with over 3000 cpu cores connected via an Infiniband network. Data storage facilities include a 92TB parallel network filesystem (Panasas) and archived storage featuring a Hierarchical Storage Management (HSM) system and remote file mirroring. Job submission and execution is handled by PBS Professional, a commercial-grade job scheduling software.

The cluster has been funded by a variety of groups and represents a major increase in capability for UQ and its partners. Its usage is closely monitored to ensure that the entitlements of all stakeholders are satisfied. The maintenance and management of Barrine including system upgrades and configuration is handled by UQ's Research Computing Centre (RCC).

Barrine is used by researchers from all over Australia, and also hosts Bioinformatics Resource Australia-EMBL web services & EBI FTP Mirror. Whilst bioinformatics represents a substantial amount of the work done on Barrine, there are a number of other groups including physics and mechanical engineering using the cluster.

NCI operates two large-scale peak systems, Raijin (a Fujitsu Primergy cluster, commissioned in 2013, and which entered production in June 2013), and Vayu (a Sun/Oracle Constellation cluster, commissioned in 2009, and due to be decommissioned later in 2013), and a small specialised system, Fujin (a Fujitsu PrimeHPC FX10 system, commissioned in March 2013).

A key characteristic of both of the production systems, Vayu and Raijin, is their balance. System balance is that important characteristic that provides hardware and applications scalability by requiring that crucial system parameters such as performance of memory, the interconnect fabric, and the filesystem all scale with node and aggregate processor performance. Such balance, which is displayed in the hardware solution architected by the HPC systems team at NCI, provides the facility with strong and scalable performance, and delivers a quality user experience.

RCC SEMINARS

RCC launches the Semester II e-Research seminars

On 2 August, RCC launched an exciting range of seminars in Axon 5. These feature high profile keynote speakers on an exciting range of topics from e-Research infrastructure to advanced applications. We invite you to come along and see what's happening in some of the most important e-research centres. The seminars

02/08/2013	Prof Michael Norman, SDSC	Application Experiences with Gordon - A Flash-based HPC System
09/08/2013	Dr Sameer Tilak, CalIT2, UCSD	Storage and Analysis of Big Data from Sensor Networks: Challenges and Opportunities
16/08/2013	Dr Wayne Pfeiffer, SDSC	Bioinformatics Meets Big Data
23/08/2013	Prof Geoffrey Fox, U. Indiana	Data Science and the Cloud
30/08/2013	Dr Shonali Krishnaswamy, I2R, A-star, Singapore	Big Data Analytics: When Research Meets Reality
06/09/2013	Prof Jack Dongarra, U. Tennessee	Emerging Heterogeneous Technologies for High Performance Computing
13/09/2013	Dr Nancy Wilkins-Diehr, SDSC	Science in the Digital Age: The Science Gateways Institute and NSF Software Investments
20/09/2013	Dr Arun Konagurthu, Monash U.	
27/09/2013	Dr William Gropp, NCSA	
11/10/2013	Dr Mark Miller, SDSC	Embedding CIPRES Science Gateway Capabilities in Phylogenetics Software Environments
18/10/2013	Dr Ilkay Altintas, SDSC	On Workflow-Driven Science Using Scientific Workflows and Provenance
25/10/2013	Prof Peter Fox, RPI	How Environmental Informatics is Preparing Us

CONTACTS

Name	Position	e-Mail
David Abramson	Director	David.Abramson@uq.edu.au
Brendon Carter	QCloud systems programmer	brendan.carter@uq.edu.au
Justin Clark	eResearch analyst	j.clark4@uq.edu.au
Sharon Cook	Executive Officer	s.cook8@uq.edu.au
Dr David Green	Deputy Director	David.green@uq.edu.au
Ron Horst	Project Manager - GVL	r.horst@uq.edu.au
Ryurick Hristev	Systems Administrator	r.hristev@uq.edu.au
William Hsu	Senior IT Officer - BRAEMBL	w.hsu1@uq.edu.au
Minh Huynh	QCloud Systems Programmer	m.huynh1@uq.edu.au
Adrian Lee	QCloud Systems Programmer	ajlee@uq.edu.au
Dr Igor Makunin	Research Fellow - GVL	i.makunin@uq.edu.au
Michael Mallon	Technical Assistant	m.mallon@uq.edu.au
Fran Moore	Administration Officer	f.moore1@uq.edu.au
Martin Nicholls	Specialist Systems Programmer	m.nicholls@uq.edu.au
Dr Edan Scriven	QCloud Systems Programmer	e.scriven@uq.edu.au