

Demystifying the NeCTAR Research Cloud



Bernard Meade, Tom Fifield,
Clint Walsh, Steven Manos

Abstract



The NeCTAR Research Cloud (nectar.org.au/research-cloud) is a free service provided to Australian researchers who may benefit from access to cloud computing resources.

The Research Cloud allows any researcher who is authenticated via the Australian Access Federation service, access to a self-service “server-on-demand”. These servers can be used to host web services, customized e-research tools and even provide an ad-hoc high performance compute cluster.

This presentation will introduce the NeCTAR Research Cloud, the design philosophy and architecture that underpins the service and the many ways it is currently being used. We will also look at the benefits to research institutions and discuss practical issues, including security and allocation requests.

This presentation will complement the NeCTAR Research Cloud Workshop.

Aim



The purpose of this talk is to have
EVERYONE in this room walk
away thinking

“I can do that!”

Overview



- ❧ What is Cloud Computing?
- ❧ What is the NeCTAR Research Cloud?
- ❧ Who is the RC for?
- ❧ What are the benefits of the RC?
- ❧ How can I use the RC?
- ❧ What is the Dashboard and how do I connect?
- ❧ How do I start a “computer in the cloud”?
- ❧ How do I connect to my Virtual Machine?
- ❧ What is the Object Store?
- ❧ Please Sir, I want some more!
- ❧ NeCTAR Virtual Laboratories and e-Research Tools

What is Cloud Computing?



☞ From Dictionary.com:

☞ **cloud computing**

☞ noun Internet-based computing in which large groups of remote servers are networked so as to allow sharing of data-processing tasks, centralized data storage, and online access to computer services or resources.

☞ A more practical definition:

☞ Cloud computing allows us to create virtual computers that we connect to remotely. These Virtual Machines (VM) can be treated like a real computer, providing everything from web services to number crunching. They can be operated individually or in clusters. They exist within a cloud of computing resources that allows transparent upgrading and efficient sharing of these resources. They can also be backup up, transferred and cloned many times.

This cloud.. why build it ourselves?



1. **Proximity – the honeypot** – infrastructure attracts community
2. **Local infrastructure** is more responsive to research needs
3. **Service offering and usage modes** suitable for research
4. **Locality** to instruments, research networks and other infrastructure.
5. **Data sovereignty**

What NeCTAR services are on offer?



☞ Research Cloud

- ☞ @ nectar.org.au/research-cloud
- ☞ Instant access, scalable, self-service
- ☞ for all researchers



☞ National Servers program

- ☞ @ nectar.org.au/national-servers-program
- ☞ Robust, highly available, backup up, small scale,
- ☞ for core research services



Who is the RC for?



- ☞ Researchers!
- ☞ Teams of researchers
- ☞ Developers working for researchers
- ☞ Researchers providing web access to their research
- ☞ Researchers needing lots of compute resources
- ☞ Researchers who collaborate with other researchers
- ☞ Researchers who don't collaborate with other researchers

What are the benefits of the RC?



- ❧ “Instant” computer – just add water! (and coffee – it takes a minute to start so you might as well have a cuppa).
- ❧ Core usage: cloud vs. desktop
- ❧ Need more grunt? More horse power – less paperwork!
- ❧ Managed underlying infrastructure means a highly reliable and available service
- ❧ Server-grade computers no longer require dedicated space within your department’s server rooms
- ❧ Did I mention it’s free?

How can I use the RC?



- ❧ Anyone from an Australian Access Federation institution can log in and immediately access 2 cores (with 8 GB of RAM and 60GB additional storage) – for ~3 months
- ❧ Images are provided to get you started
- ❧ You can create your own snapshots of any VM you set up for backup, transfer or cloning
- ❧ You can apply to NeCTAR for a larger, ongoing allocation – allocations are merit-based
- ❧ Your local institution can purchase nodes to attach to the Research Cloud infrastructure, which can then be allocated at the discretion of the institution

What is the Dashboard and how do I connect?



- ❧ The Dashboard allows you to:
 - ❧ Monitor and manage your VMs
 - ❧ Manage Access controls such as keypairs and security groups
 - ❧ Create, delete and modify snapshots of your VMs
 - ❧ Submit allocation requests to NeCTAR

Demo

How do I start a “computer in the cloud”?



1. Connect to the Dashboard
2. Create a keypair and save it locally
3. Create a security group allowing necessary ports
4. Select an image or snapshot and click “Launch”
5. Choose a name, a size, a keypair and a security group
6. Launch your VM

Demo

How do I connect to my Virtual Machine?



- ❧ SSH (Secure Shell)
 - ❧ Terminal (Mac), Xterm (Linux), Putty (Windows)
- ❧ VNC Console
 - ❧ Issues on Windows at the moment
- ❧ Remote Desktop (VNC or NX Client)
 - ❧ NX Client seems to be the best performer
- ❧ Web services (for snapshots like Web2GO)
 - ❧ Still need SSH or something else to secure and maintain the system

Demo

What is the Object Store?



- ❧ Redundant, scalable, API-accessible storage platform
- ❧ Simple: buckets of objects, not a filesystem
- ❧ Good for: virtual machine images, backups, stage-in-stage-out style datasets

Please sir, I want some more!



NeCTAR Resource Allocation Request Form

- ❧ Project name
- ❧ Start and end dates
- ❧ Primary instance type
- ❧ Number of Cores
- ❧ Number of core hours
- ❧ Number of instances
- ❧ Object storage size
- ❧ Use case
- ❧ Usage patterns
- ❧ Geographic requirements

NeCTAR Virtual Laboratories



- ❧ Virtual Geophysics Laboratory
 - ❧ Workflows and access to geophysical tools, data and resources
- ❧ Genomics Virtual Laboratory
 - ❧ Collaborative genomics workflows in the cloud
- ❧ Marine Virtual Laboratory (MARVL)
 - ❧ Virtual environments to unify marine modelling and observation
- ❧ All Sky Virtual Observatory
 - ❧ Federating astronomy data: SkyMapper and the Theoretical Astronomical Observatory
- ❧ Climate and Weather Science Laboratory
 - ❧ Integrated environment for climate and weather science modelling and data
- ❧ Humanities Networked Infrastructure (HUNI)
 - ❧ Unlocking and uniting Australia's cultural data
- ❧ Characterisation Virtual Laboratory
 - ❧ Research environments for exploring inner space
 - ❧ And more to come....

NeCTAR eResearch Tools



- ❧ **UniCarbKB** - e-infrastructure for glycomics (Macquarie University)
- ❧ **OzTrack** - Tools for the storage, analysis and visualisation of animal tracking data (University of Queensland)
- ❧ **Australian Synchrotron** - tools for the Australian Synchrotron community
- ❧ **Federated Archaeological information management system** - A comprehensive information system for archaeology (University of NSW)
- ❧ **Geology from Geodynamics** - Access geophysical data to improve accuracy when using 3D modelling technology (Monash University)
- ❧ **High throughput computing for globally connected science** - Support participation in research CERN's Large Hadron Collider (Centre of Excellence for Particle Physics)
- ❧ **Cloud-based bioinformatics tools** - Analyse how underlying genetic variations/mutations interact with environmental factors to cause disease (University of Western Australia)
- ❧ **Cloud based image analysis and processing toolbox** - Provide improved access to biomedical image processing and analysis software packages (CSIRO)

NeCTAR eResearch Tools



- ❧ **Bioscience data platform (TARDIS in the cloud)** - bring existing computational systems together to allow scientists to seamlessly work with data from capture through to publication (Monash University)
- ❧ **Drishti and Voluminous** - Volume visualisation tools (Australian National University)
- ❧ **CATAMI** - Collaborative and Automated Tools for the analysis of marine imagery and video (Curtin University)
- ❧ **Quadrant** - Cloud-based research project management and data collection tool (Queensland Cyber Infrastructure)
- ❧ **The Aust-ESE project** - Tools to support the collaborative authoring and management of electronic scholarly editions (University of Queensland)
- ❧ **Human Variome Project** - A national repository of molecular data generated by Australian diagnostic laboratories (University of Melbourne)
- ❧ **SHaRED** - Submission, harmonisation and retrieval of ecological data (University of Adelaide)

Other uses of the Research Cloud

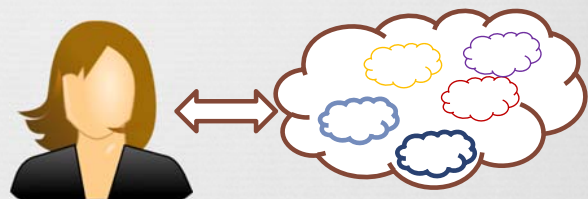


- œ Running a genomics tools training workshop
- œ Running Monte Carlo simulations
- œ Hosting and publishing census, map and voting data
- œ Hosting of an online training environment for doctors
- œ Characterisation of mammograms
- œ Hosting of new databases for protein design
- œ Rendering of animations

A Federated Cloud



- œ A single national cloud
 - œ Up to seven nodes
 - œ \$1.5M per node
- œ Supporting inward and outward federation
 - œ EC2/S3 API, OpenStack API



Specifications

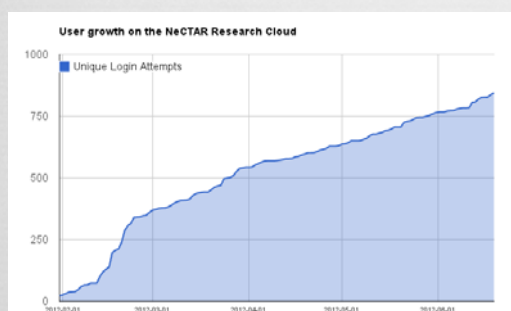


- OpenStack Essex/Stable (+~5% trunk backport)
 - Ubuntu 12.04 LTS, Puppet
- Hardware
 - 336 cores – 48 Core Dell R815s
 - 3840 cores – 160* 24 core, 128GB, 10Gbit/s Xenon Quad2U
 - 195TB – HP DL180G w/ DL2000 @ 24TB/node
 - 146TB – Dell R715 w/ MD1200 @ 24TB/node
 - 10Gbit/s – Cisco Nexus (2232, 5596, ...)
 - Hitachi HNAS/BluARC – 100TB for running VMs

A growing community



- ✧ Almost 1000 researchers across every state and territory, and more every day
 - ✧ Running events in most states, aiming to create a self-supporting, sustainable community
- ✧ Built on OpenStack – Australia has the largest per-capita OpenStack group in the world
 - ✧ We're confident in the middleware choice and encourage you to find out more



Further Information



www.nectar.org.au/research-cloud
support.rc.nectar.org.au
support.rc.nectar.org.au/forum