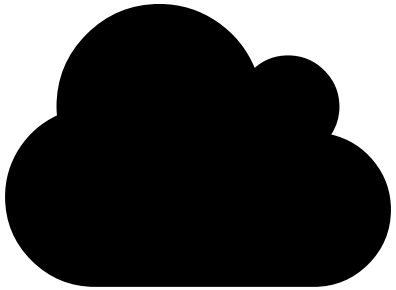
Getting Started with the Melbourne Research Cloud



Dr. David Perry

perry.d@unimelb.edu.au / @dwjperry

Pre-Requisites

- Laptop
- Working SSH Client
- University Email Address
- Basic Linux Commands

Training Outcomes

- Understand what cloud computing is about
- Submit an allocation request
- Create and operate a computer in the cloud
- Understand cloud data storage
- Know where to go for help

Part 1: Background

A boring cloud taxonomy:

Infrastructure as a Service (IaaS)	**Platform as a Service (PaaS)**	**Software as a Service (SaaS)**
• Amazon Web Services • Microsoft Azure • Nectar	• Heroku • Google App Engine	• Dropbox • Google Docs

You've probably heard of Dropbox, but what about Amazon Web Services (AWS)? They provide IaaS, which is what we mean when we talk about 'cloud' in todays session. Some years back, Amazon figured they were good at running data centres, and should rent out spare capacity to others. Users could simply rent a computer for an hour at a time, rather than buy and maintain their own. Amazon actually make more money from AWS these days than retailing! IaaS is what we're talking about today.

No really, what is it?

Computing on tap...

Do these apply to you?

- I want a *faster* computer
- I want *lots* of computers
- I want a *special* kind of computer
- I want to *share* my computer with collaborators
- I want my computer to *host* a web service

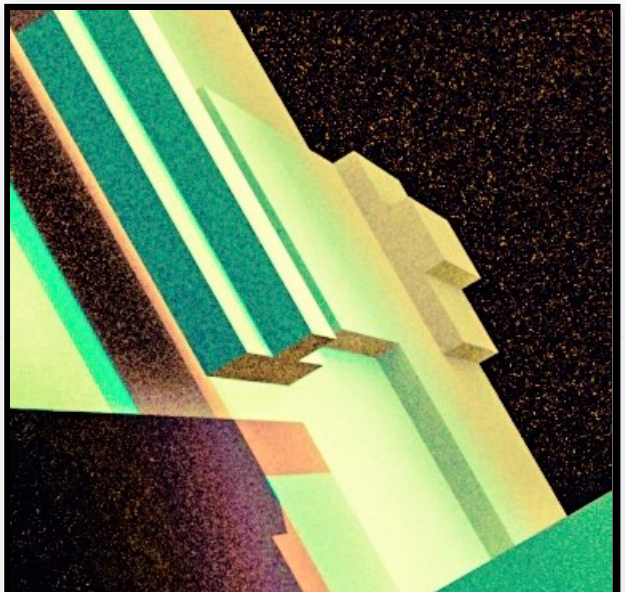
...and you want them all right now.

Then maybe you need cloud computing!

Some (Brief) Case Studies

Rendering Artwork

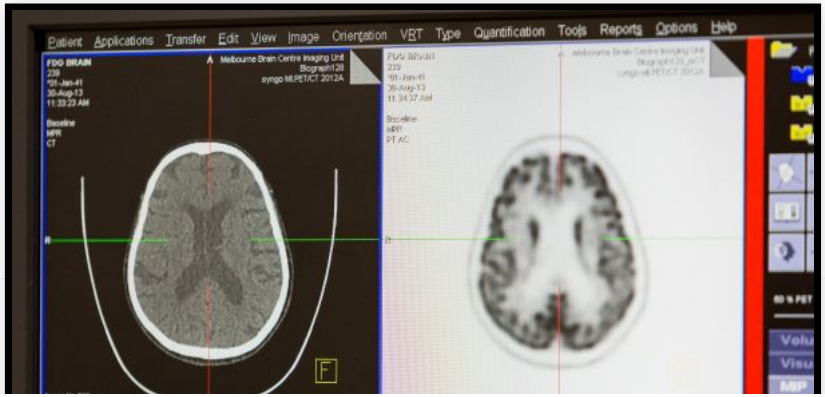
- Rendering animation works; takes too long on desktop computer!
- Run on a large computer in the cloud instead.



Brendan Lee is doing a Masters of Fine Art, and his works are generated in a 3D rendering package called Carrera. This was getting slow! We helped him provision a powerful computer in the cloud, and install Windows and Carrera. The work was able to be completed more quickly, could be left running unattended, and freed up his computer for other work.

Analyse Brain Function

- A desktop environment, with MRI analysis software pre-installed.
- Easy for new students to get started.



Scott Kolbe and his group at the Florey use neuroimaging to study the brain. They do much of their work on cloud-based computers, using a customised environment with all the necessary specialist software pre-installed. As you might imagine, this software is fiddly to setup and so this saves a bunch of time, makes sure everyone can work consistently, and access powerful computation when needed (they also use HPC on occasion, which I'll talk about in a bit). Since they are using computers in the university data centre, this also means they can access their data stores much more quickly than if they had to go through their local network to upload/download.

What about high-performance computing (HPC)?

i.e. Use a supercomputer!

Another service we provide is HPC via our local cluster, called Spartan. HPC is suited to big analysis jobs, especially for software that can run in Linux. In the cloud you can get your own private resources, whereas in HPC you submit work to a shared system. That could be a good thing or a bad thing depending on your needs.

High-Performance vs. Cloud Computing

Need complete control? Don't like sharing/waiting? Hosting a web service? Need Windows? ➡️ Cloud

Don't want to manage a system yourself? Need enormous scale? ➡️ HPC

Here are some guidelines on which to choose. Some researchers use both. We run 1-day training on HPC if you're interested in finding out more, or visit our documentation page.

For more info, visit: http://dashboard.hpc.unimelb.edu.au

Virtual Laboratories

Domain-specific environments built upon the research cloud.

- Genomics Virtual Laboratory
- All-Sky Virtual Observatory
- Virtual Geophysics Lab
- ... and more: https://nectar.org.au/labs-and-tools/

So, in the research cloud we're mostly creating our own environment from scratch. But several groups have built their own specialist platforms ontop, and perhaps there is one to suit you? These include pre-built software, datasets and an established user community.

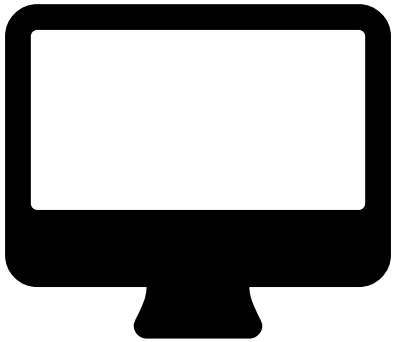
Part 1: Introduction to the Research Cloud

Oh no, jargon!

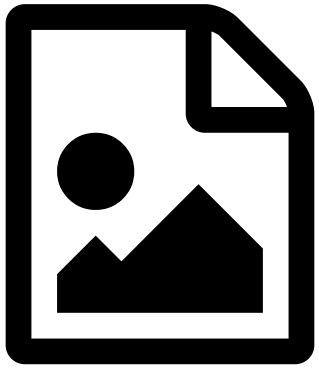
- Instance
- Image
- Volume Storage
- Object Storage
- Security Groups
- Flavor
- Availability Zone
- Keypairs

Unfortunately there is some jargon we have to get through. Let's work through these concepts one at a time, and see how they link together. I don't expect it will make sense right away; we'll sketch out these concepts and then work through some concrete examples with some repetition.

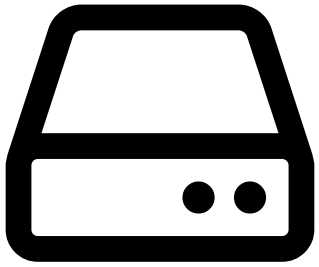
Instance



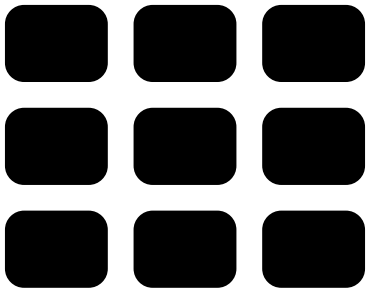
Image



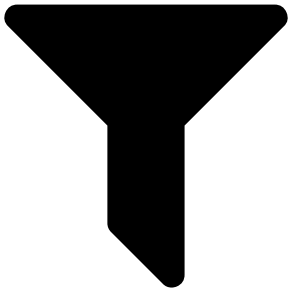
Volume Storage



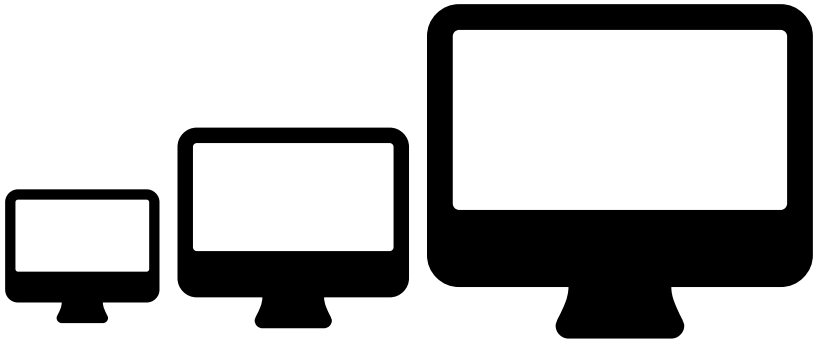
Object Storage



Security Group

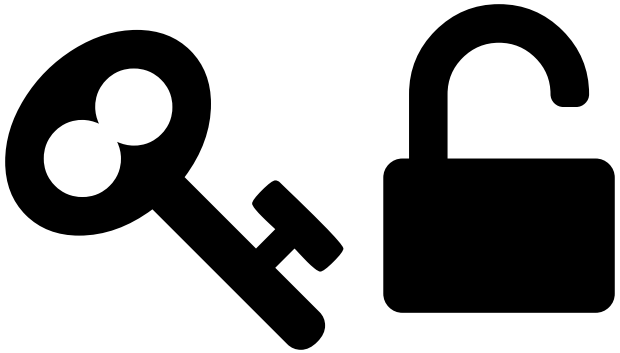


Flavor

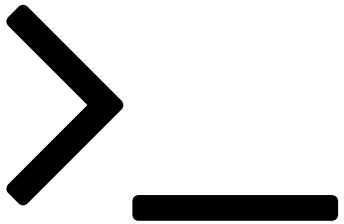


Availability Zone

Key Pair



SSH (Secure SHell)



Nectar

Part 2: Create an Instance

Demo: Let's create an instance.

Recap:

- Trial project with 2 CPU-Months of compute time available immediately.
- Create an **Instance**, with a particular **Flavor**.
- Use an **Image** to load an operating system (and perhaps software) onto it.
- Select a **Keypair** to use with the **Instance**, allowing us to login.
- Set some **Security Groups** to allow network access in/out of the **Instance**.
- Choose an **Availability Zone** (data centre) within the Nectar network where our **Instance** will reside.
- We connect to our **Instance** via **SSH**, which authenticates us using our **Keypair**.

Your turn!

Challenge 1: Create an instance and connect to it.

Instructions: http://go.unimelb.edu.au/j7d6

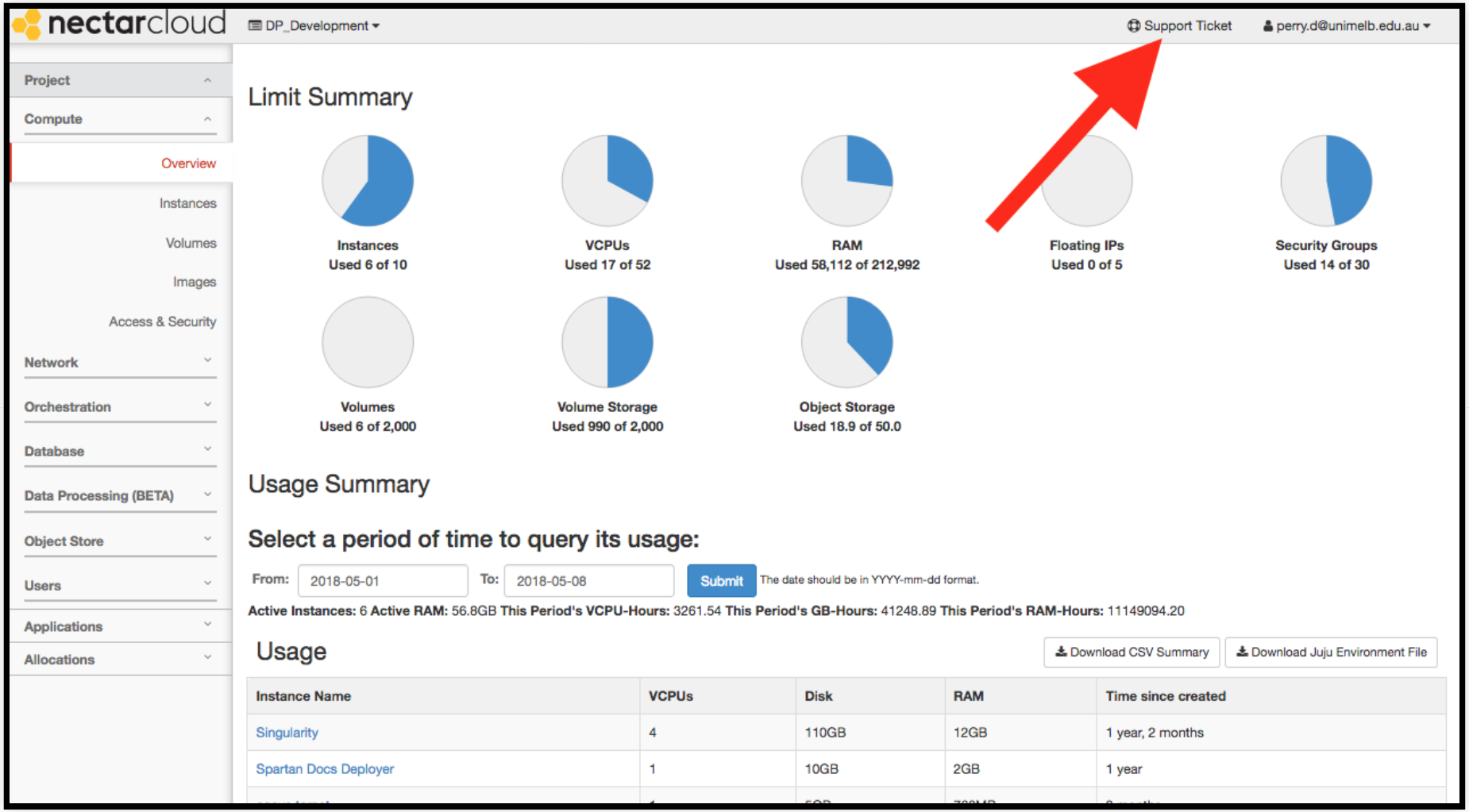
Now it's your turn to have a go! Work in groups; I suggest one person has the instructions open on their computer, while another drives the dashboard and SSH client through their computer.

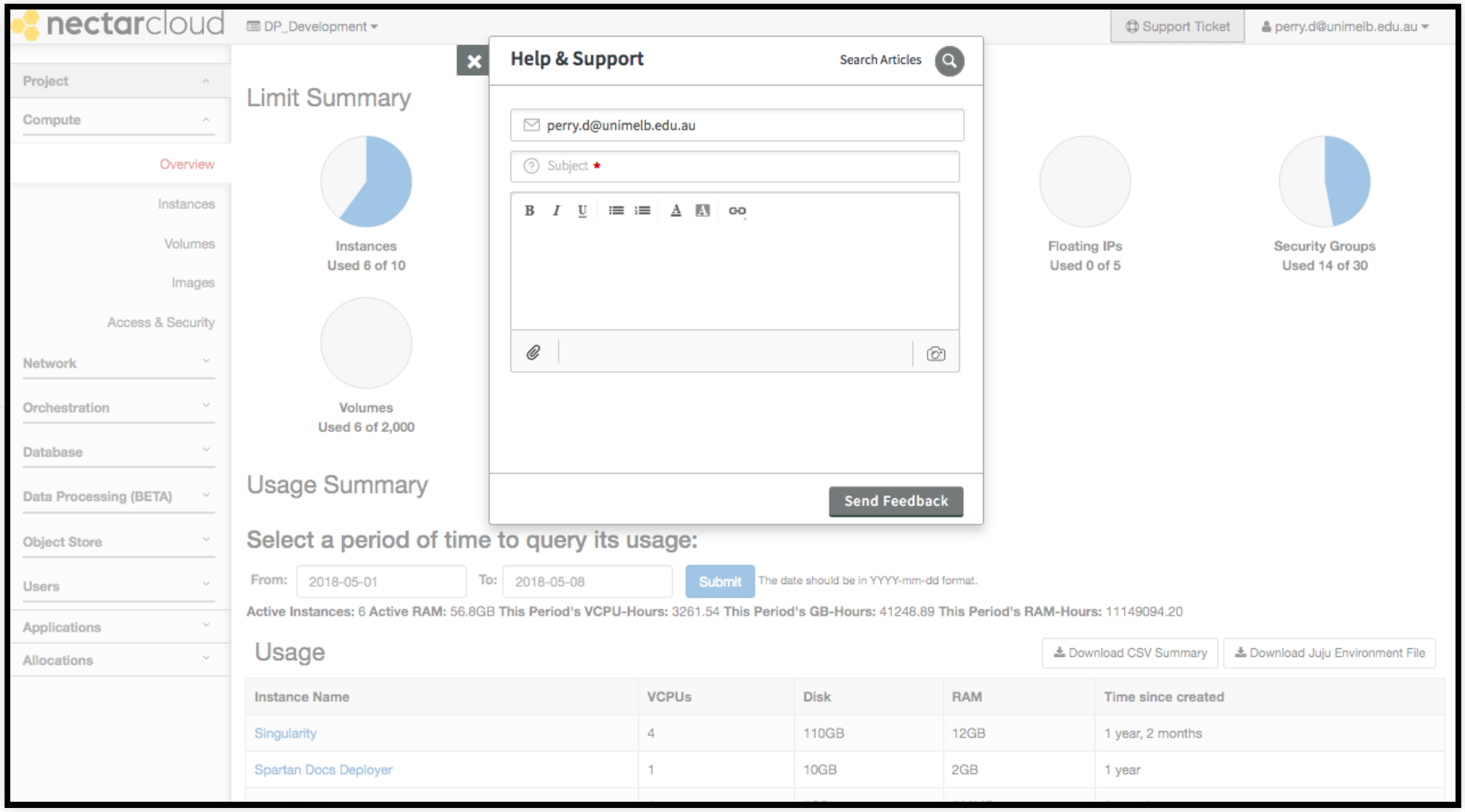
Where to go for help?

Self-Paced Training: http://training.nectar.org.au/

General Documentation: http://support.nectar.org.au/

Helpdesk: support@nectar.org.au or call 1300 080 431 (9am-7pm).


nectarcloud
DP_Development
Support Ticket
perry.d@unimelb.edu.au
Project
Compute
Overview
Instances
Volumes
Images
Access & Security
Network
Orchestration
Database
Data Processing (BETA)
Object Store
Users
Applications
Allocations
Limit Summary
Instances
Used 6 of 10
VCPUs
Used 17 of 52
RAM
Used 58,112 of 212,992
Floating IPs
Used 0 of 5
Security Groups
Used 14 of 30
Volumes
Used 6 of 2,000
Volume Storage
Used 990 of 2,000
Object Storage
Used 18.9 of 50.0
Usage Summary
Select a period of time to query its usage:
From: 2018-05-01
To: 2018-05-08
Submit
The date should be in YYYY-mm-dd format.
Active Instances: 6 Active RAM: 56.8GB This Period's VCPU-Hours: 3261.54 This Period's GB-Hours: 41248.89 This Period's RAM-Hours: 11149094.20
Usage
Download CSV Summary
Download Juju Environment File
Instance Name
VCPUs
Disk
RAM
Time since created
Singularity
4
110GB
12GB
1 year, 2 months
Spartan Docs Deployer
1
10GB
2GB
1 year


nectarcloud
DP_Development
Support Ticket
perry.d@unimelb.edu.au
Project
Compute
Overview
Instances
Volumes
Images
Access & Security
Network
Orchestration
Database
Data Processing (BETA)
Object Store
Users
Applications
Allocations
Limit Summary
Instances
Used 6 of 10
Volumes
Used 6 of 2,000
Floating IPs
Used 0 of 5
Security Groups
Used 14 of 30
Help & Support
Search Articles
perry.d@unimelb.edu.au
Subject *
Send Feedback
Usage Summary
Select a period of time to query its usage:
From: 2018-05-01
To: 2018-05-08
Submit
The date should be in YYYY-mm-dd format.
Active Instances: 6 Active RAM: 56.8GB This Period's VCPU-Hours: 3261.54 This Period's GB-Hours: 41248.89 This Period's RAM-Hours: 11149094.20
Usage
Download CSV Summary
Download Juju Environment File
Instance Name
VCPUs
Disk
RAM
Time since created
Singularity
4
110GB
12GB
1 year, 2 months
Spartan Docs Deployer
1
10GB
2GB
1 year

Challenge 2: Find out about Object Storage

- Use the training/documentation/Google to figure out:
  - How to upload a file to object storage from the dashboard?
  - How many copies of your data does the object store keep to ensure integrity?

Hint: http://training.nectar.org.au/package07/sections/objectStorage.html

Recap: Object Store

Part 3: Security

Joint responsibility

- We protect the cloud backend
- But you are responsible for what runs on it

We have joint responsibility; the research cloud has a team working on keeping the infrastructure operational and secure; but we don't control (or necessarily know) what you're running inside it. For example, if you run some very old blogging software with known vulnerabilities, it's possible your instances will be taken over, and we won't notice until someone else on the web makes a complaint. This doesn't happen often (perhaps once a month), but does occur. The hackers aren't out to get you exactly, there are some just bad actors that scan the internet for every single machine that exhibits a particular known vulnerability.

Basic Security Principles

It's not possible to eliminate all security risks, but there are some simple steps you can take which don't require a bunch of specialist know-how.

Minimise Means of Access

- Only allow access from necessary ports/addresses.
- Power off or terminate resources you don't need.
- Don't run services you don't need.

The first step is to provide only the bare minimum means of access to do your work. Ideally, power off or suspend your machine when you're not using it (can't hack a machine that's turned off!) Don't add security groups you don't need (ideally just stick to SSH, and keep your private key safe), and if possible lock down access to particular IP address ranges. For example, you could limit access to addresses used by the University virtual private network (VPN; staff: 10.1.128.0/22, student: 10.9.128.0/22). In reality though, you might to run a web server or similar, in which case you have to take a more active role.

Keep OS and Software Up-To-Date

Subscribe to relevant updates/advisorsies for your software and operating system.

If you're running something like a web app, database or other service that's possible for others to access, you have to be a bit more diligent. This includes keeping both your operating system (OS) up to date, as well as the software running on it. Over time security flaws are discovered in various packages, and if you hold onto an old version for too long, it's possible it could be compromised.


WordPress.org
Search
Showcase
Themes
Plugins
Mobile
Support
Get Involved
About
Blog
Hosting
WordPress 4.9.2 Security and Maintenance Release
Posted January 16, 2018 by Ian Dunn. Filed under Releases, Security.
WordPress 4.9.2 is now available. This is a security and maintenance release for all versions since WordPress 3.7. We strongly encourage you to update your sites immediately.
An XSS vulnerability was discovered in the Flash fallback files in MediaElement, a library that is included with WordPress. Because the Flash files are no longer needed for most use cases, they have been removed from WordPress.

Sensitive Data?

- Be aware of ethics, funding and institution requirements.
- Encrypt sensitive data, and secure the decryption key/password.

Let's say your instance has been compromised somehow, even after being very careful. Unfortunately it still happens despite best intentions, for example through social engineering, physical theft or human error. How can you limit the resulting damage? If you're working with potentially sensitive data, it's worth considering an extra layer of protection, including encrypting files.

Backup

Things can still go wrong ☹️

Say there has been a security issue, hardware failure, or you have just goofed somehow, and you need to recover. The Research Cloud makes it easy to create backups so you can do so.

Part 4: Snapshots

Introducing Snapshots

- Like an image, but copied from an existing instance.
- Includes: Everything - OS, Software, Data
- Use for backup, duplicating instances, and more!

Don't forget the object store

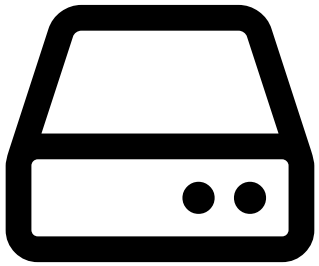
Automatically keeps three copies of your data ➡️ Free backup!

Challenge 3: Clone your instance using a snapshot

Instructions: http://go.unimelb.edu.au/43v6

Part 5: Storage

Volume Storage



Demo: Create a volume, connect it to instance, format it, and write to it.

```
sudo mkfs.ext4 /dev/vdc # Format the volume
sudo mkdir /my_volume # Create a mount point
sudo mount /dev/vdc /my_volume -t auto # Mount the volume
sudo chown ubuntu /volume_name # Make sure our ubuntu user can access it
sudo du -sh /my_volume # Show how much of the volume is used
```

Catch all of that?

Don't worry, instructions for doing this at:
https://support.ehelp.edu.au/support/solutions/articles/6000055382-introduction-to-cloud-storage

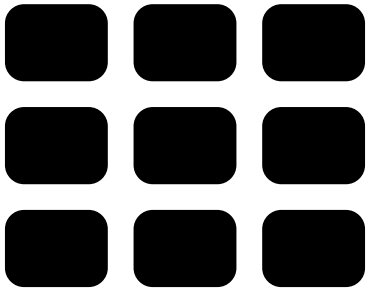
I know that's a bit to take in, but don't worry, you can refer to the documentation as needed (I certainly don't remember all of this of the top of my head!)

Ephemeral Storage

- Special type of storage included with some instance flavours.
- Not included in snapshots, and lost on reboot.
- Handy for 'scratch space': Temporary storage for files you don't need to keep.

Another type of storage I'll mention is ephermal storage -- you might notice this is available for some of the larger flavours. Say you need to save some intermediate results while you're doing some data analysis, but don't need to keep them -- ephemeral storage is ideal for this. Just take care that you don't storage anything crucial there, everything will disappear when you reboot your instance!

Object Storage



- Files, stored in 'containers'
- Can organise into (pseudo) folders
- Can keep container private, or make public (readable by everyone)
- Resilient: At least 3 copies of everything retained.

Need managed data services?

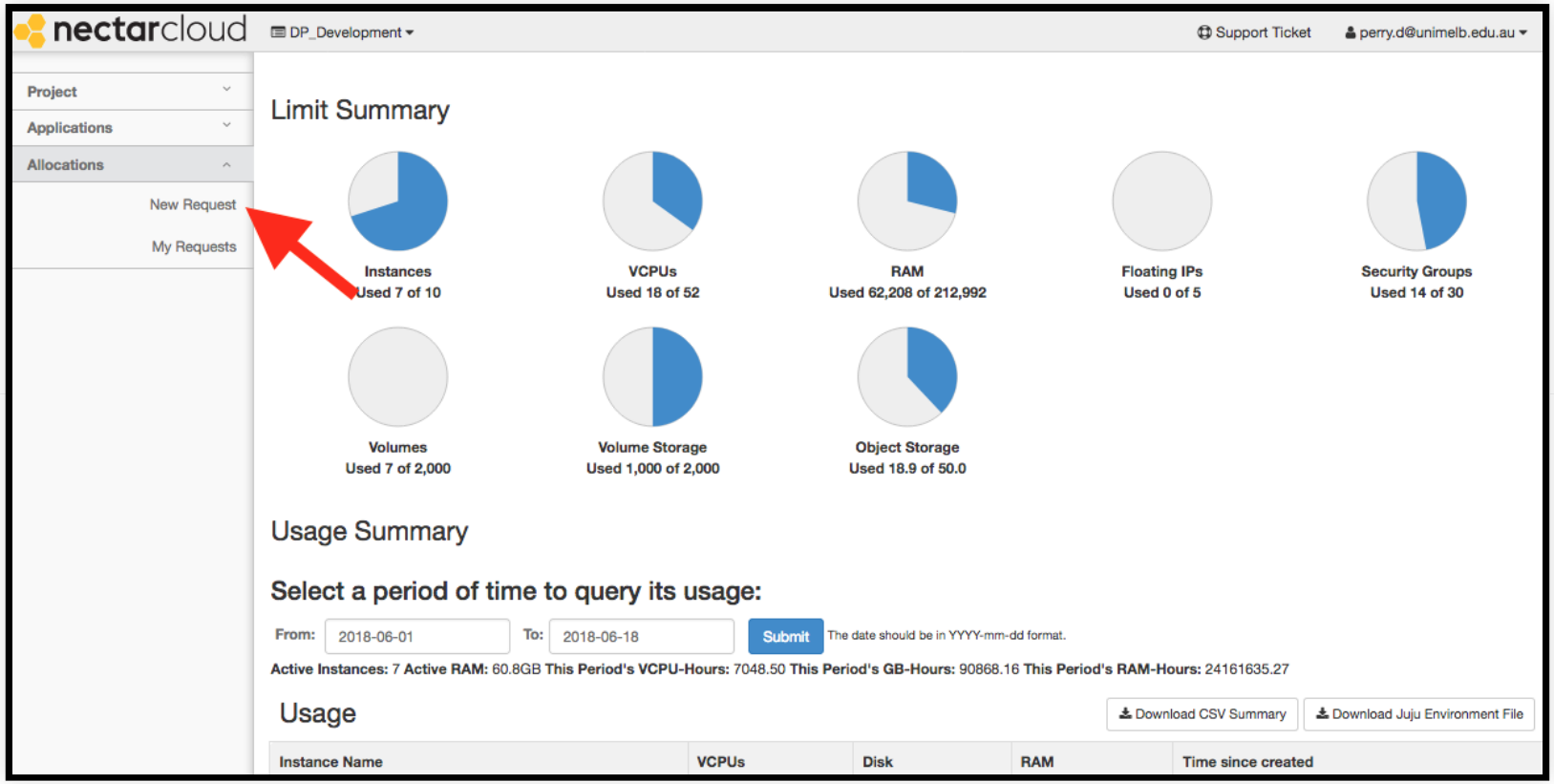
Visit: https://research.unimelb.edu.au/infrastructure/research-platform-services/services/data-storage-management

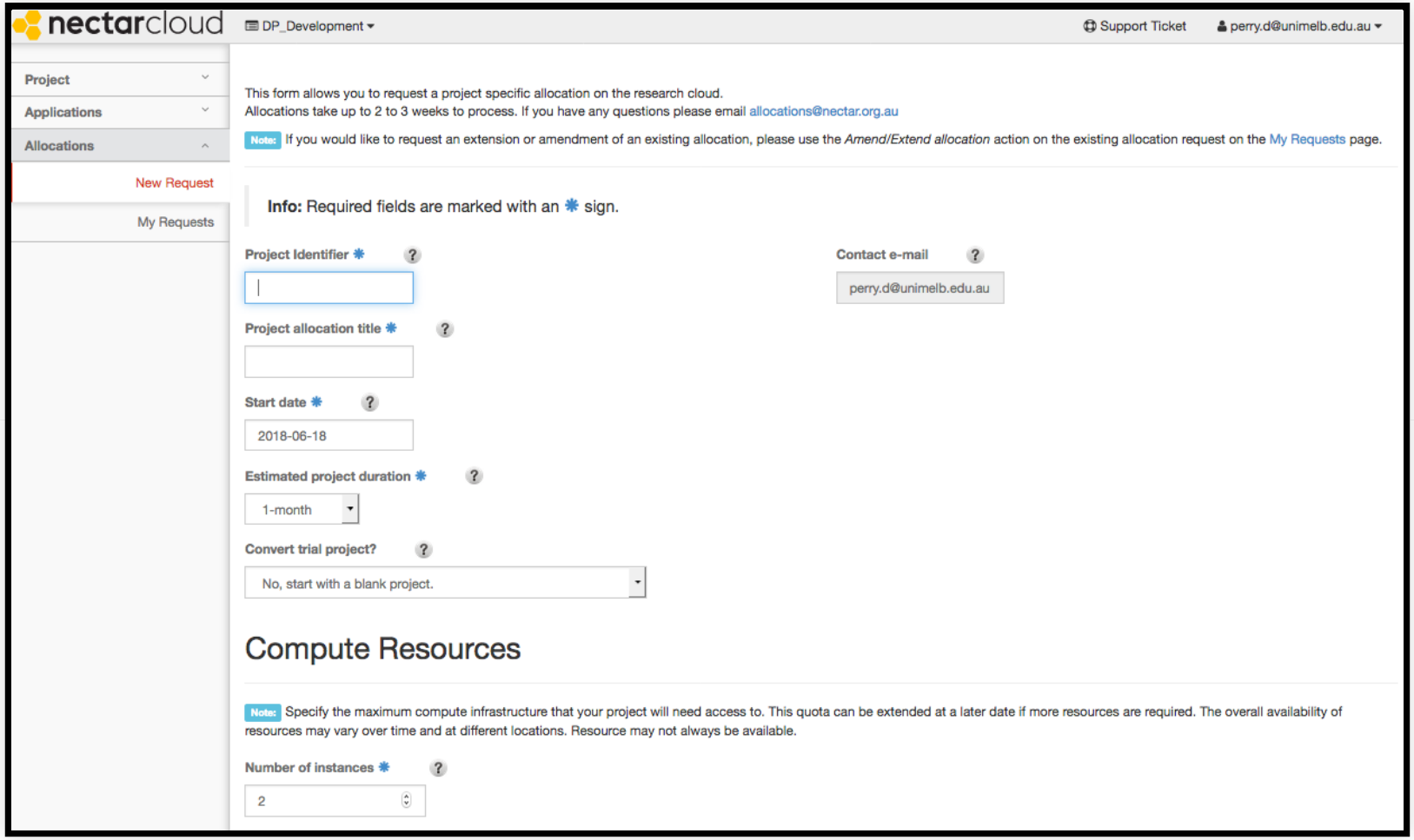
I've talked about quite low-level storage infrastructure, which you can use to build more complex systems like databases yourself. But our department (ResPlat) also offers some more specialised managed services, such as systems for managing instrument data. Get in touch and we can help you decide if they might work for you.

Part 6: Resource Allocations

Automatic trial allocation: ~2 small instances for 2 months.

Requesting More Resources

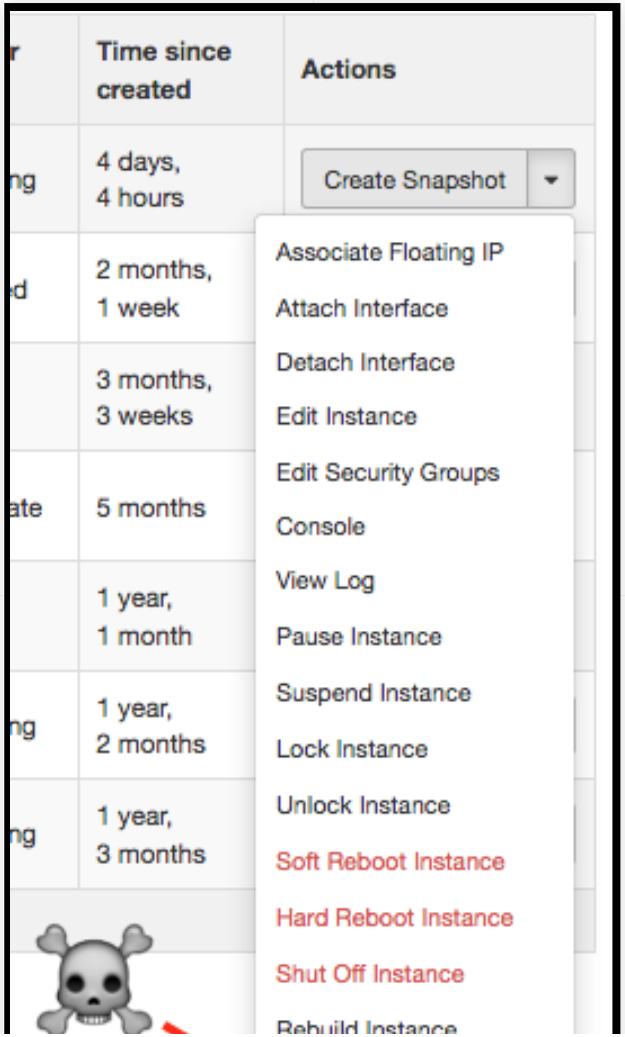



nectarcloud
DP_Development
Support Ticket
perry.d@unimelb.edu.au
Project
Applications
Allocations
New Request
My Requests
This form allows you to request a project specific allocation on the research cloud.
Allocations take up to 2 to 3 weeks to process. If you have any questions please email allocations@nectar.org.au
Note: If you would like to request an extension or amendment of an existing allocation, please use the Amend/Extend allocation action on the existing allocation request on the My Requests page.
Info: Required fields are marked with an * sign.
Project Identifier *
Contact e-mail
perry.d@unimelb.edu.au
Project allocation title *
Start date *
2018-06-18
Estimated project duration *
1-month
Convert trial project?
No, start with a blank project.
Compute Resources
Note: Specify the maximum compute infrastructure that your project will need access to. This quota can be extended at a later date if more resources are required. The overall availability of resources may vary over time and at different locations. Resource may not always be available.
Number of instances *
2

Part 7: Wrapping Up

We're getting near the end, just a few little items to wrap up.

Terminating Instances



So, when we are done with an instance, we can terminate it. Note that this will destroy everything on the instance, permanently. However, any snapshots will be preserved, as will anything stored in a separate storage volume.

Final Challenge: Terminate your instance!

Special Bonus Topics

Windows Instances

Possible, but a bit complicated for now.

See: https://github.com/resbaz/HOWTO-Nectar-Windows-VM

Software Licensing

- You can BYO license
- Or use a University site license, provided you select a Melbourne availability zone.

You might be wondering about commercial software. Well, just about anything you can run on your own computer will work in the cloud too, and most software licenses allow for that. You're welcome to bring your own license for a piece of software if you already have it, or the University has an extensive range of licensed software already available, including things like MATLAB, and almost every Microsoft product. Keep in mind is that University-licensed software generally has to run on University hardware, in which case you must select a Melbourne Availability Zone for your instance.

GPU Instances

NVIDA M10 GPUs available for Graphics Work

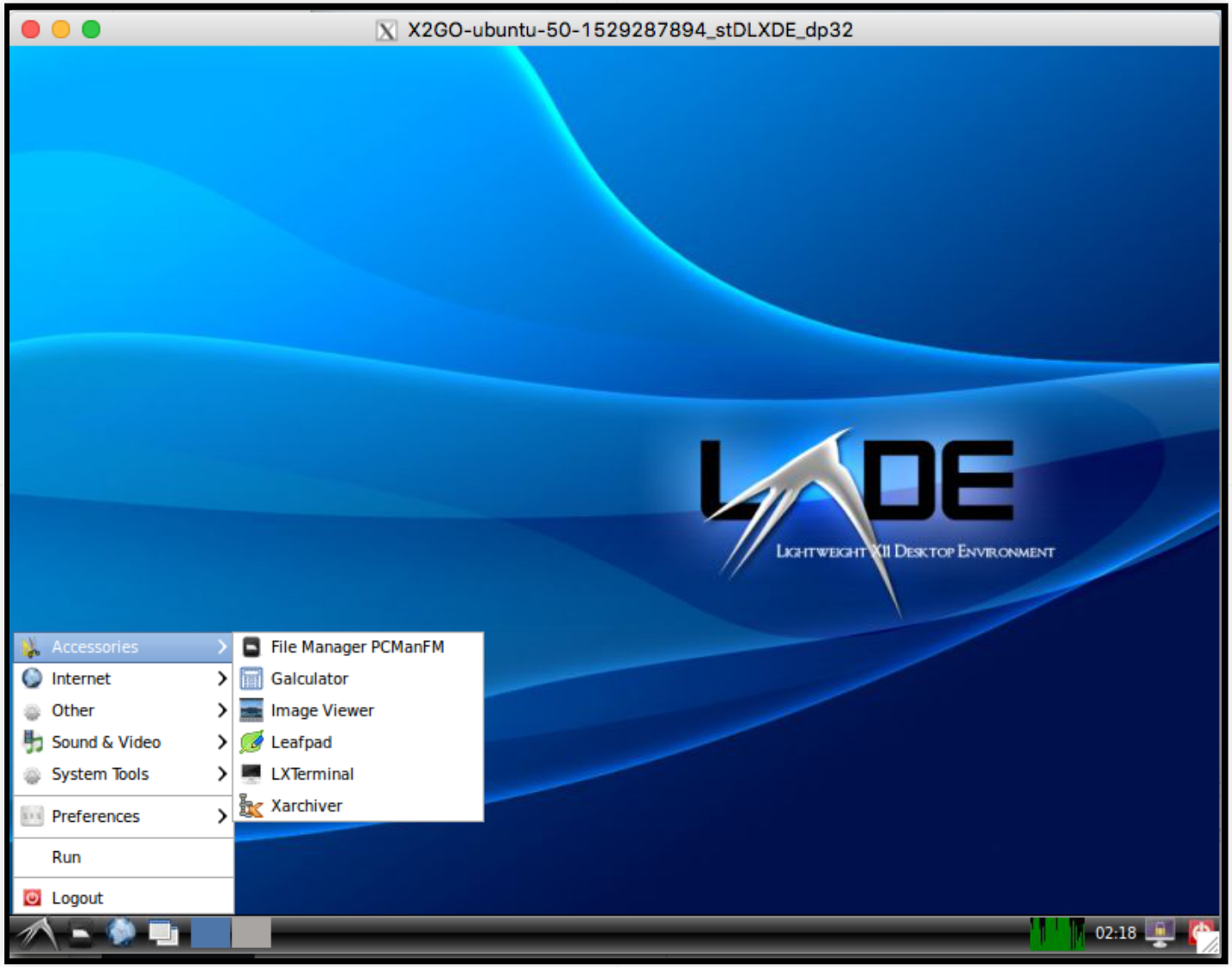
e.g. CAD, neuroimaging

Contact us, or checkout this for more: https://github.com/resbaz/HOWTO-Nectar-GPU

Need general-purpose GPU resources (e.g. for deep learning)?

Try our HPC cluster Spartan:
https://dashboard.hpc.unimelb.edu.au/gpu/

Linux Desktop Access



To get started with a Linux desktop:
http://training.nectar.org.au/package07/sections/connectWithX2Go.html

The End

Questions?