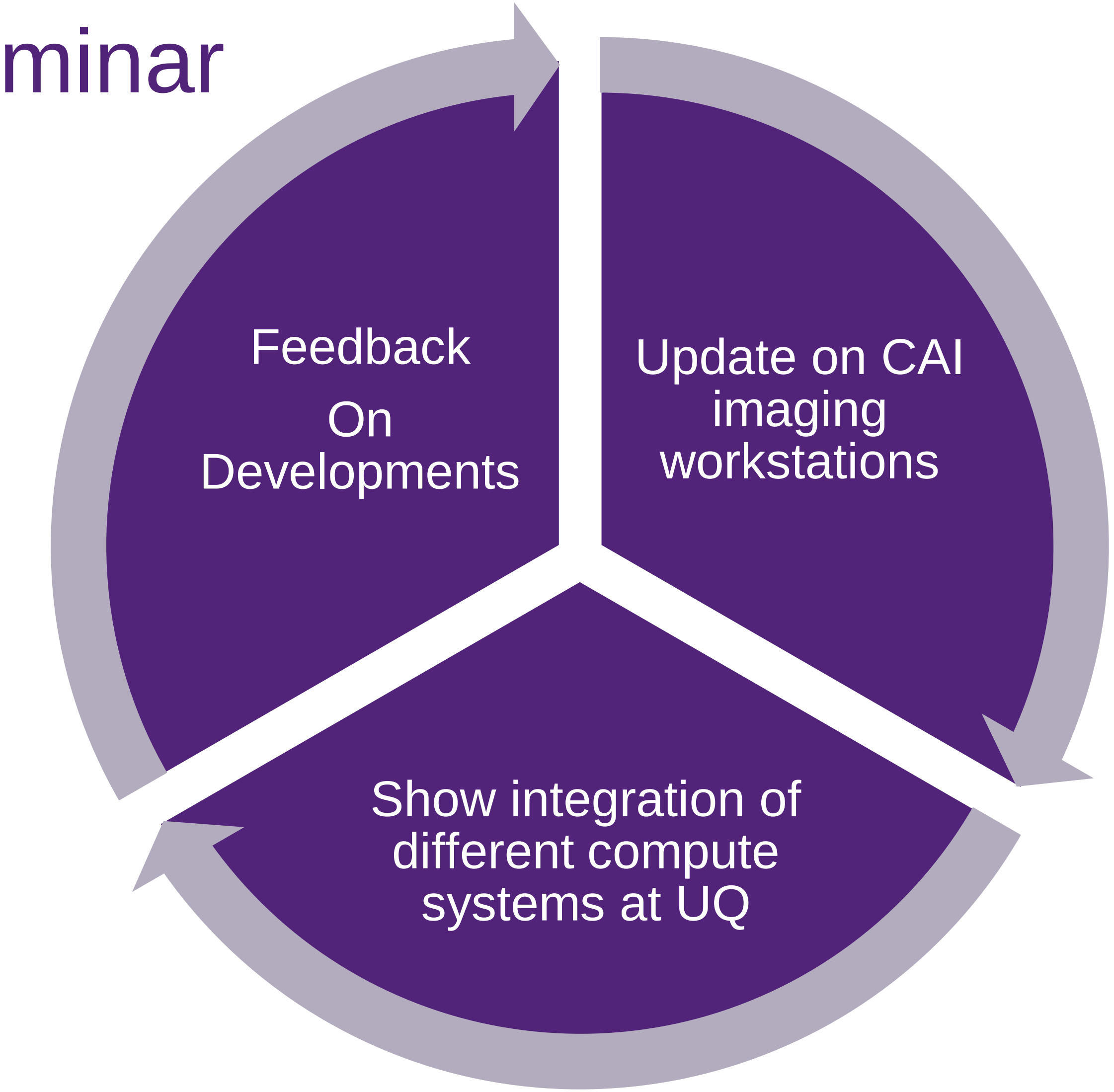
The University of Queensland
Australia
CREATE CHANGE

Large scale image processing and high performance computing

Steffen Bollmann
Research Fellow
School of Information Technology and Electrical Engineering

David Butler, Quang Tieng, Alan Hockings, Jake Carroll
Oren Civier, Aswin Narayanan, Markus Barth, Tom Johnstone
Jakub Kaczmarzyk, Martin Grignard

Goals of this seminar



Feedback

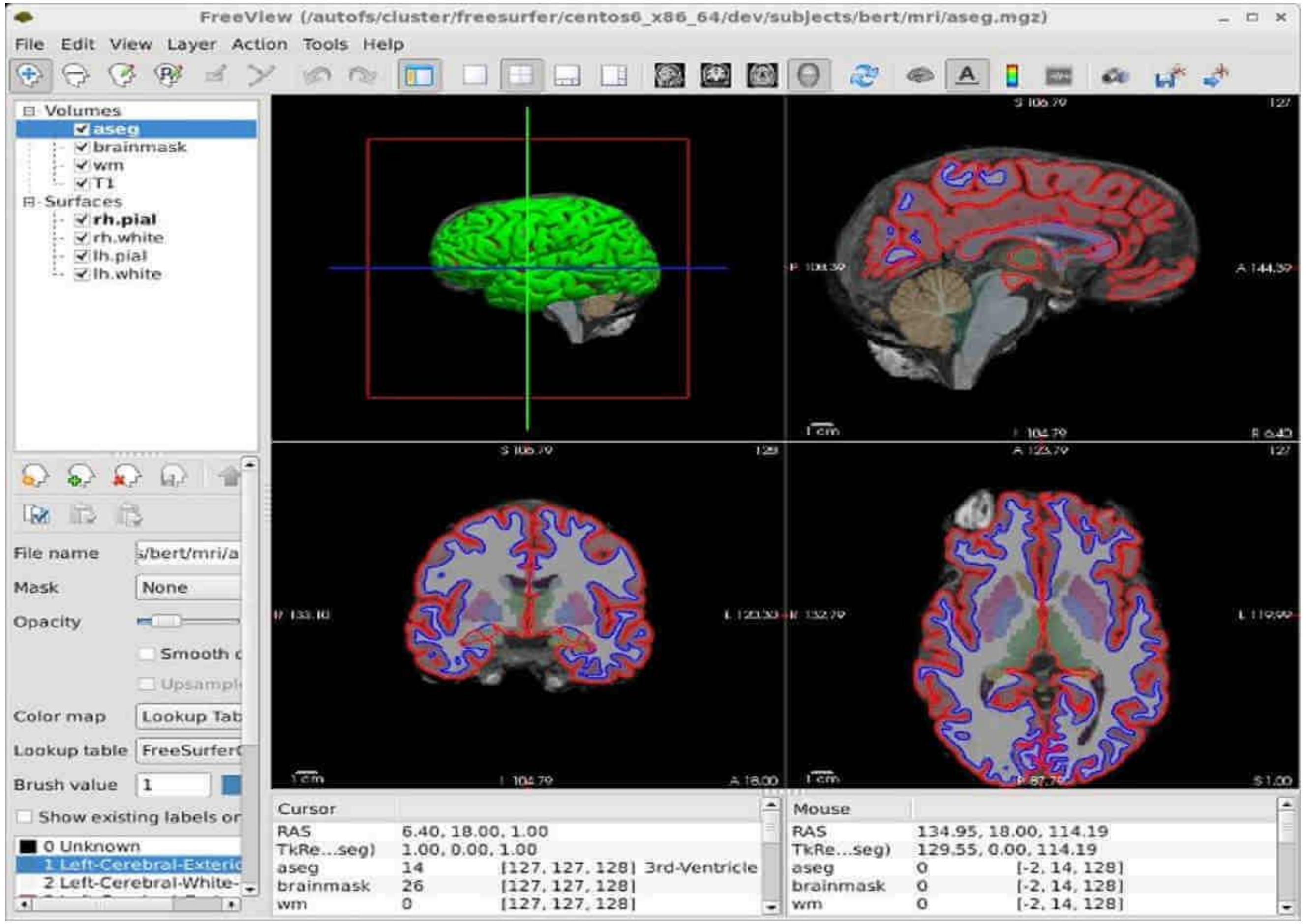
On Developments

Update on CAI imaging workstations

Show integration of different compute systems at UQ

Large ecosystem of scientific software …

FreeSurfer



FreeView (/autofs/cluster/freesurfer/centos6_x86_64/dev/subjects/bert/mri/aseg.mgz)

File name s/bert/mri/a

Mask None

Color map Lookup Tab

Lookup table FreeSurferC

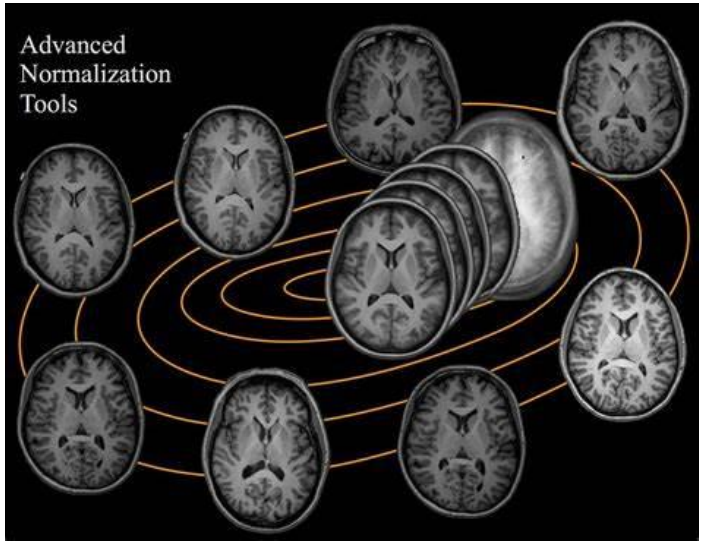
Brush value 1

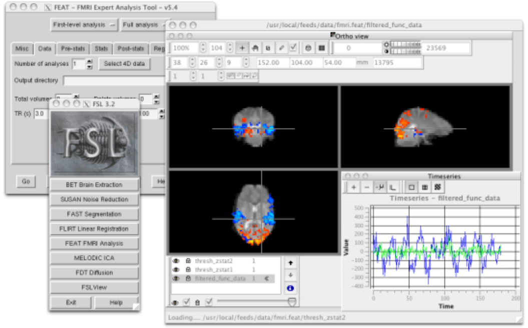
Cursor			
RAS	6.40, 18.00, 1.00		
TkRe...seg)	1.00, 0.00, 1.00		
aseg	14	[127, 127, 128]	3rd-Ventricle
brainmask	26	[127, 127, 128]	
wm	0	[127, 127, 128]	

Mouse		
RAS	134.95, 18.00, 114.19	
TkRe...seg)	129.55, 0.00, 114.19	
aseg	0	[-2, 14, 128]
brainmask	0	[-2, 14, 128]
wm	0	[-2, 14, 128]

Large ecosystem of scientific software …

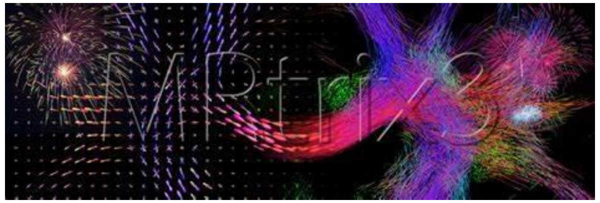
FreeSurfer





python™

julia

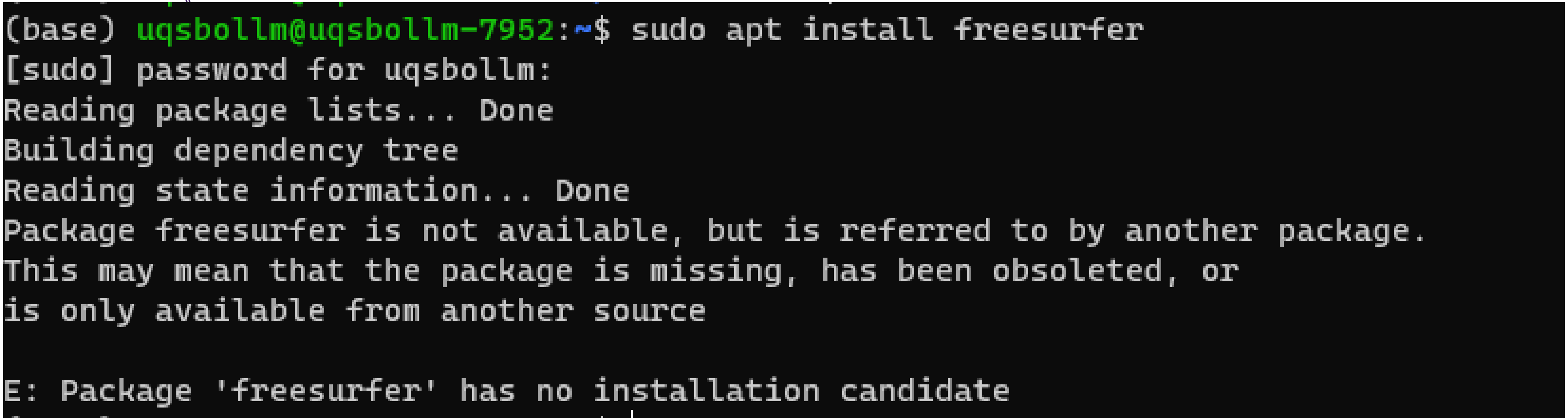


… creating problems for researchers:

- Most tools require Linux

… creating problems for researchers:

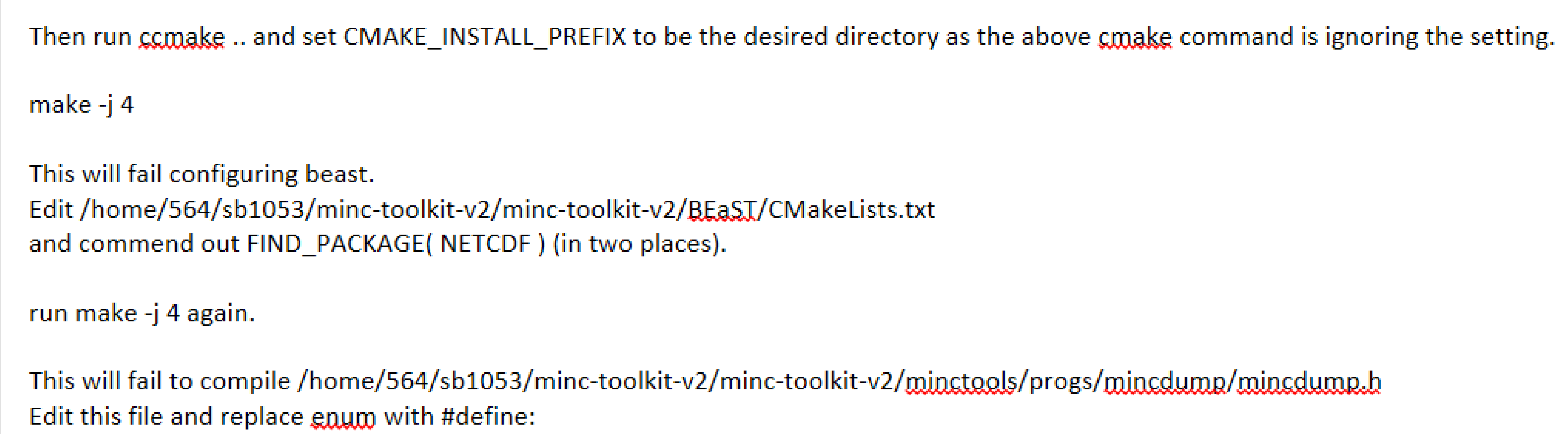
- Most tools require Linux
- Tools are not available in standard package systems



```
(base) uqsbollm@uqsbollm-7952:~$ sudo apt install freesurfer
[sudo] password for uqsbollm:
Reading package lists... Done
Building dependency tree
Reading state information... Done
Package freesurfer is not available, but is referred to by another package.
This may mean that the package is missing, has been obsoleted, or
is only available from another source

E: Package 'freesurfer' has no installation candidate
```

… creating problems for researchers:

- Most tools require Linux
- Tools are not available in standard package systems
- Compiling from source often a nightmare



Then run ccmake .. and set CMAKE_INSTALL_PREFIX to be the desired directory as the above cmake command is ignoring the setting.

make -j 4

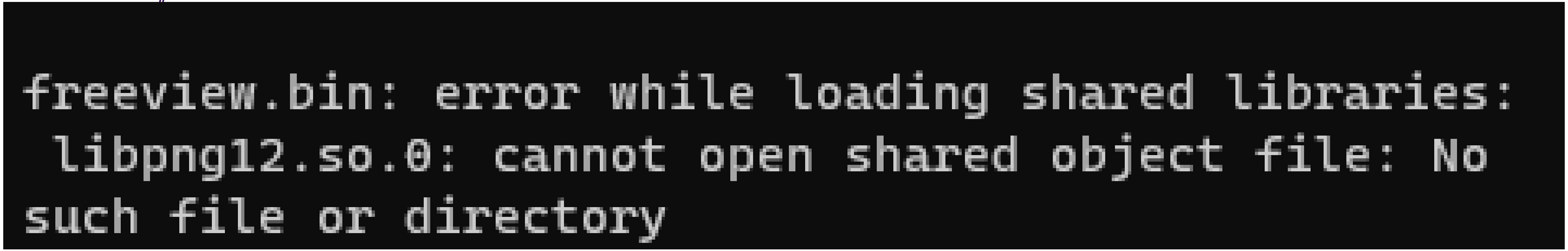
This will fail configuring beast.
Edit /home/564/sb1053/minc-toolkit-v2/minc-toolkit-v2/BEaST/CMakeLists.txt
and commend out FIND_PACKAGE(NETCDF) (in two places).

run make -j 4 again.

This will fail to compile /home/564/sb1053/minc-toolkit-v2/minc-toolkit-v2/minctools/progs/mincdump/mincdump.h
Edit this file and replace enum with #define:

… creating problems for researchers:

- Most tools require Linux
- Tools are not available in standard package systems
- Compiling from source often a nightmare
- Conflicting dependencies



```
freeview.bin: error while loading shared libraries:
 libpng12.so.0: cannot open shared object file: No
such file or directory
```

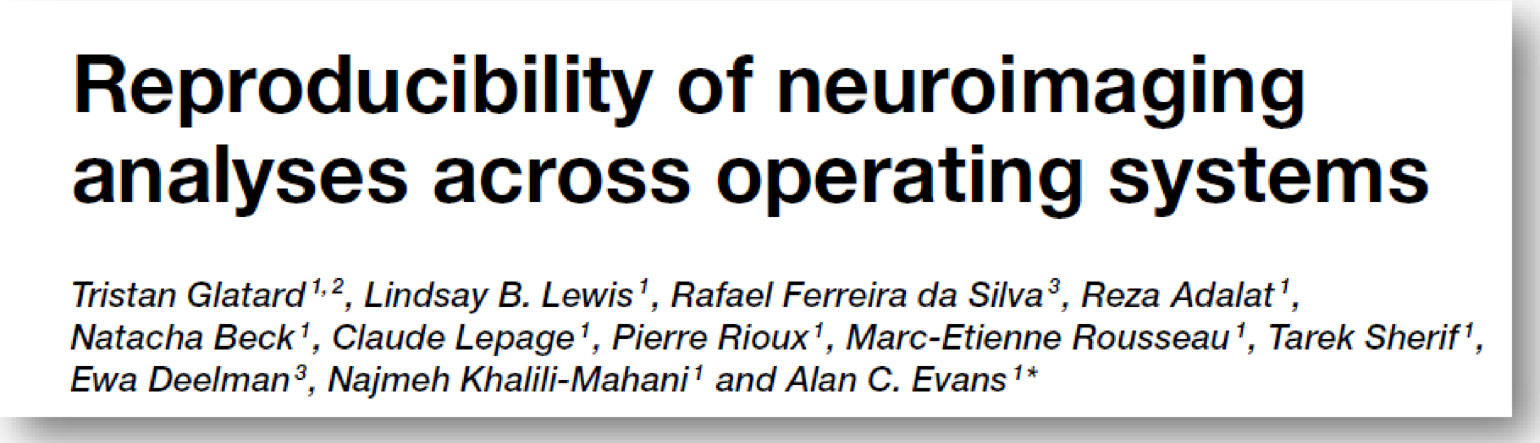
… creating problems for researchers:

- Most tools require Linux
- Tools are not available in standard package systems
- Compiling from source often a nightmare
- Conflicting dependencies
- Reinstalling tools on different platforms takes time

… creating problems for researchers:

- Most tools require Linux
- Tools are not available in standard package systems
- Compiling from source often a nightmare
- Conflicting dependencies
- Reinstalling tools on different platforms takes time
- Differing results between software versions

GLIBC 2.5 vs 2.18



Reproducibility of neuroimaging analyses across operating systems

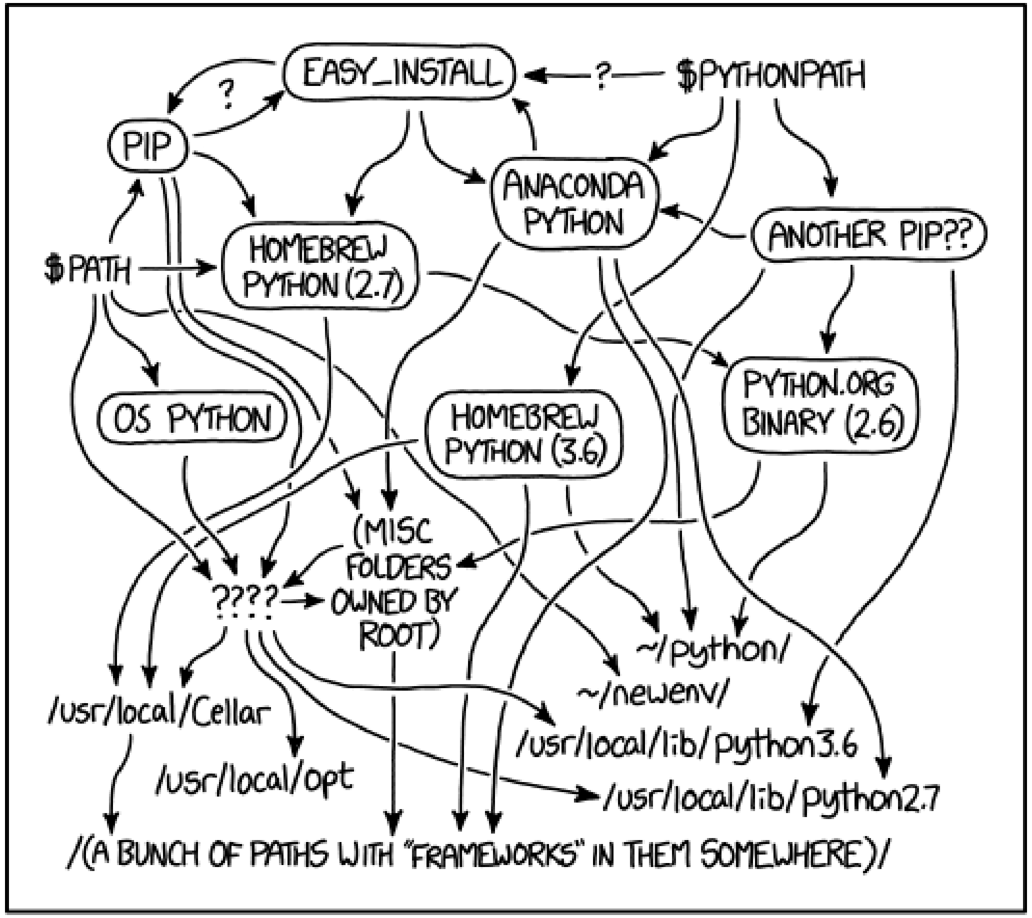
Tristan Glatard[1,2], Lindsay B. Lewis[1], Rafael Ferreira da Silva[3], Reza Adalat[1], Natacha Beck[1], Claude Lepage[1], Pierre Rioux[1], Marc-Etienne Rousseau[1], Tarek Sherif[1], Ewa Deelman[3], Najmeh Khalili-Mahani[1] and Alan C. Evans[1]*

```
expf(1.540518522262573242187500000000)
          =4.667009353637695312500
```

```
expf(1.540518522262573242187500000000)
     =4.667009830474853515625 0
```

- glibc 2.5 vs 2.18 deliver different floating-point results
- leads to significant differences in long pipelines

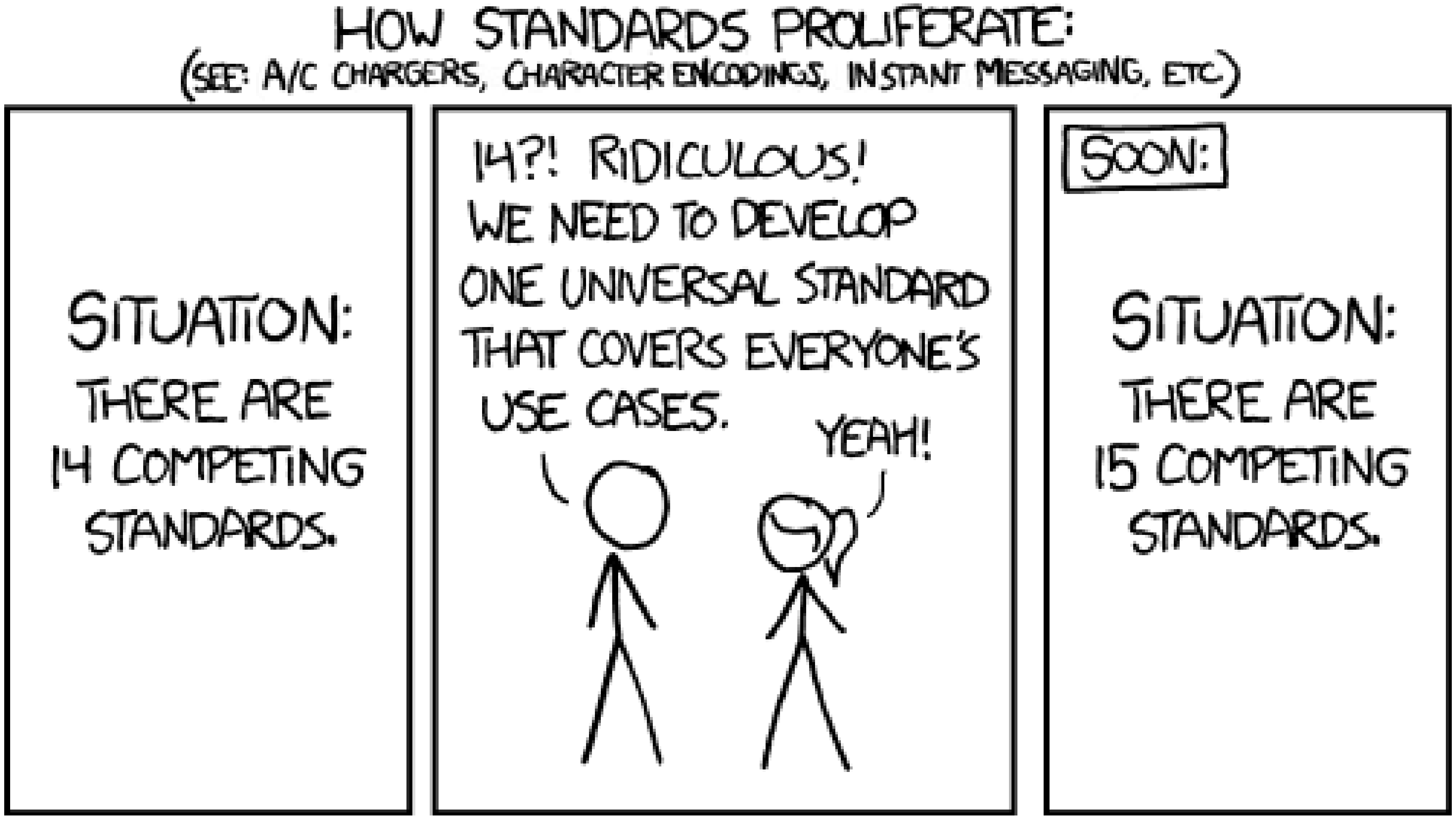
How to help with this…



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

https://xkcd.com/1987/

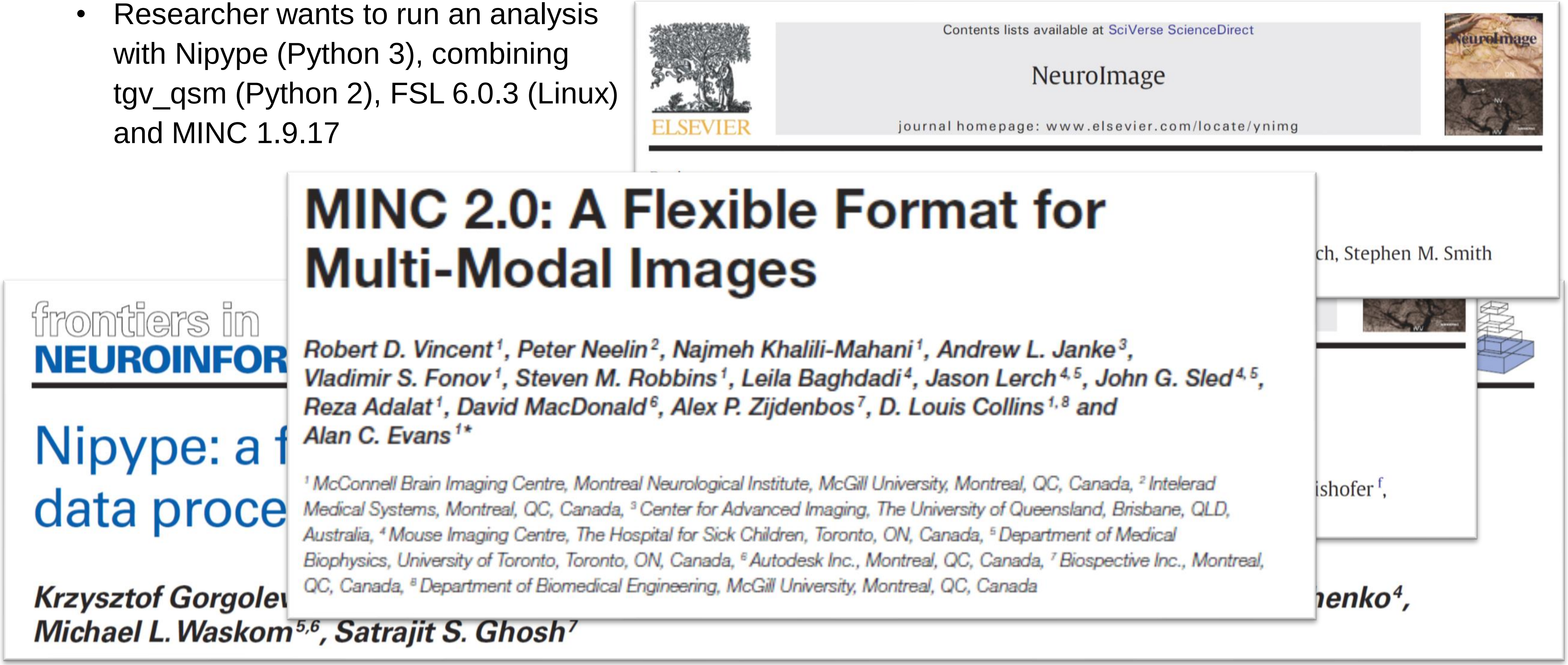
… but avoid …



https://xkcd.com/927/

Let’s start with a use case

- Researcher wants to run an analysis with Nipype (Python 3), combining tgv_qsm (Python 2), FSL 6.0.3 (Linux) and MINC 1.9.17



Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

ELSEVIER

ch, Stephen M. Smith

MINC 2.0: A Flexible Format for Multi-Modal Images

*Robert D. Vincent[1], Peter Neelin[2], Najmeh Khalili-Mahani[1], Andrew L. Janke[3], Vladimir S. Fonov[1], Steven M. Robbins[1], Leila Baghdadi[4], Jason Lerch[4,5], John G. Sled[4,5], Reza Adalat[1], David MacDonald[6], Alex P. Zijdenbos[7], D. Louis Collins[1,8] and Alan C. Evans[1]**

[1] McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, QC, Canada, [2] Intelerad Medical Systems, Montreal, QC, Canada, [3] Center for Advanced Imaging, The University of Queensland, Brisbane, QLD, Australia, [4] Mouse Imaging Centre, The Hospital for Sick Children, Toronto, ON, Canada, [5] Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada, [6] Autodesk Inc., Montreal, QC, Canada, [7] Biospective Inc., Montreal, QC, Canada, [8] Department of Biomedical Engineering, McGill University, Montreal, QC, Canada

ishofer[f],

frontiers in
NEUROINFOR

Nipype: a f
data proce

Krzysztof Gorgolev ienko[4],
Michael L. Waskom[5,6], Satrajit S. Ghosh[7]

Let’s start with a use case

- Researcher wants to run an analysis with Nipype (Python 3), combining tgv_qsm (Python 2), FSL 6.0.3 (Linux) and MINC 1.9.17 (Prebuilt packages only available for Ubuntu)
- Develop pipeline interactively on Windows 10 notebook
- Test analysis on pilot data on a Linux workstation running Ubuntu 18.04
- Analyse all data on a cluster running ROCKS Centos
- Visualize results interactively on Windows 10 notebook and prepare for publication
- Share analysis pipeline with readers of paper

Outdated libraries on old Centos cluster

FSL/MINC do not run on Windows

Mixing Python versions causes trouble

Compiling MINC fails due to missing libraries

Software setup not reproducible

What exists already and how can we combine efforts?

CONDA®

NeuroDebian

ReproNim

nectar

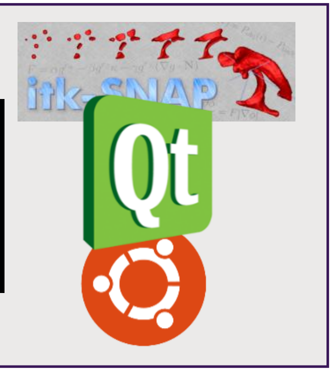
docker

CVL | Characterisation Virtual Laboratory

Funded by
ARDC
Australian Research Data Commons

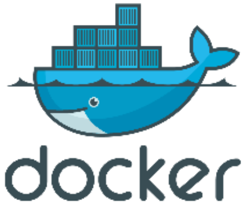
Virtual Machines VS Containers

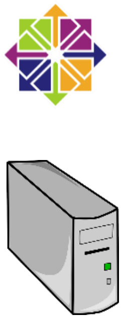
Storage: 10 GB
Startup: 15s
RAM: 4GB



Application	• e.g. itksnap
Libraries	• e.g. QT4
Guest OS	• e.g. Ubuntu 16.04
Hypervisor	• e.g. Virtualbox
Host OS	• e.g. Centos 6
Hardware	• e.g. Dell Precision

Storage: 0.1 GB
Startup: 0.2 s
RAM: 0.1 GB





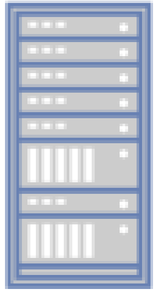
Where to Run analyses?

PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful



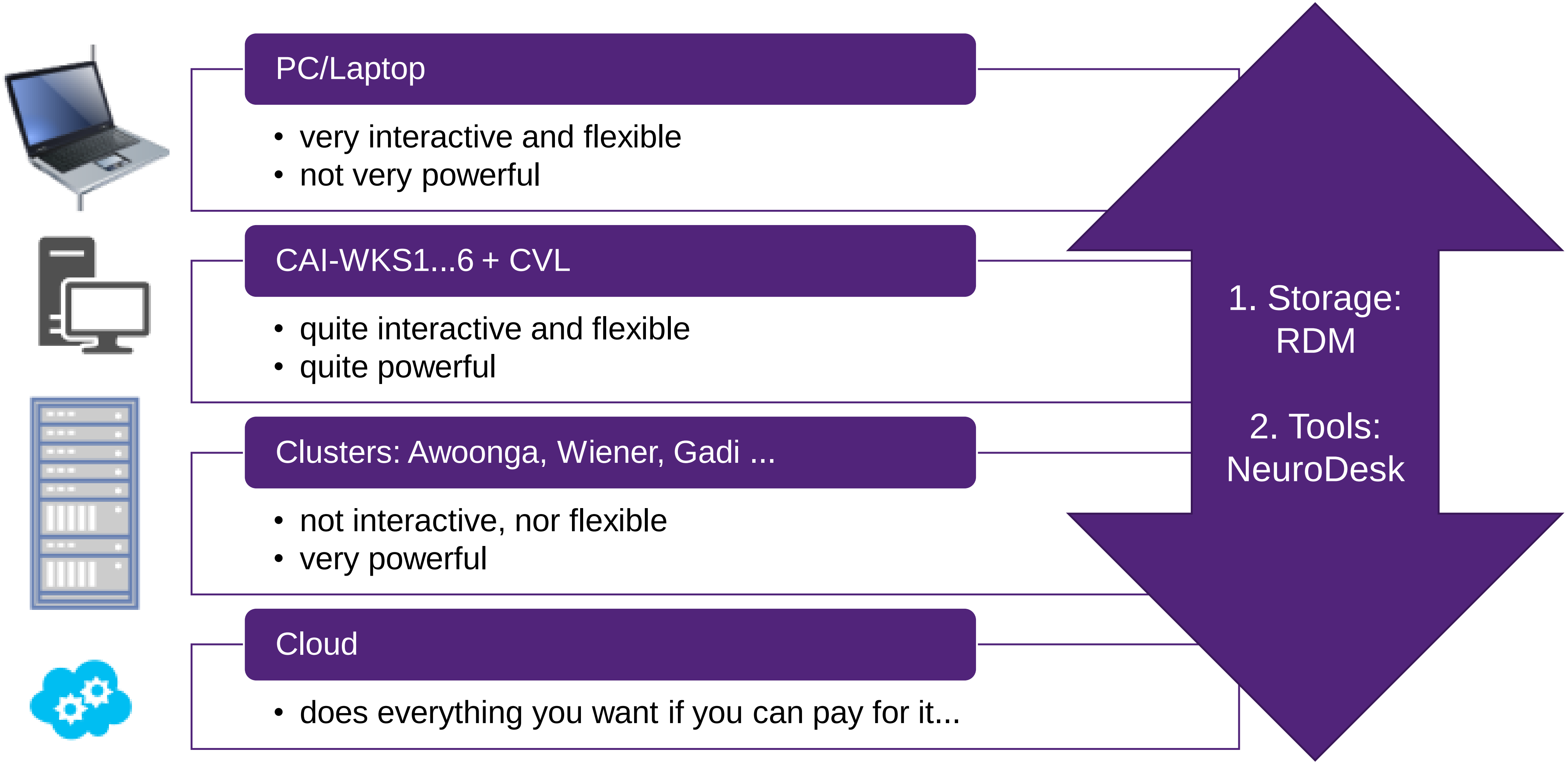
Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

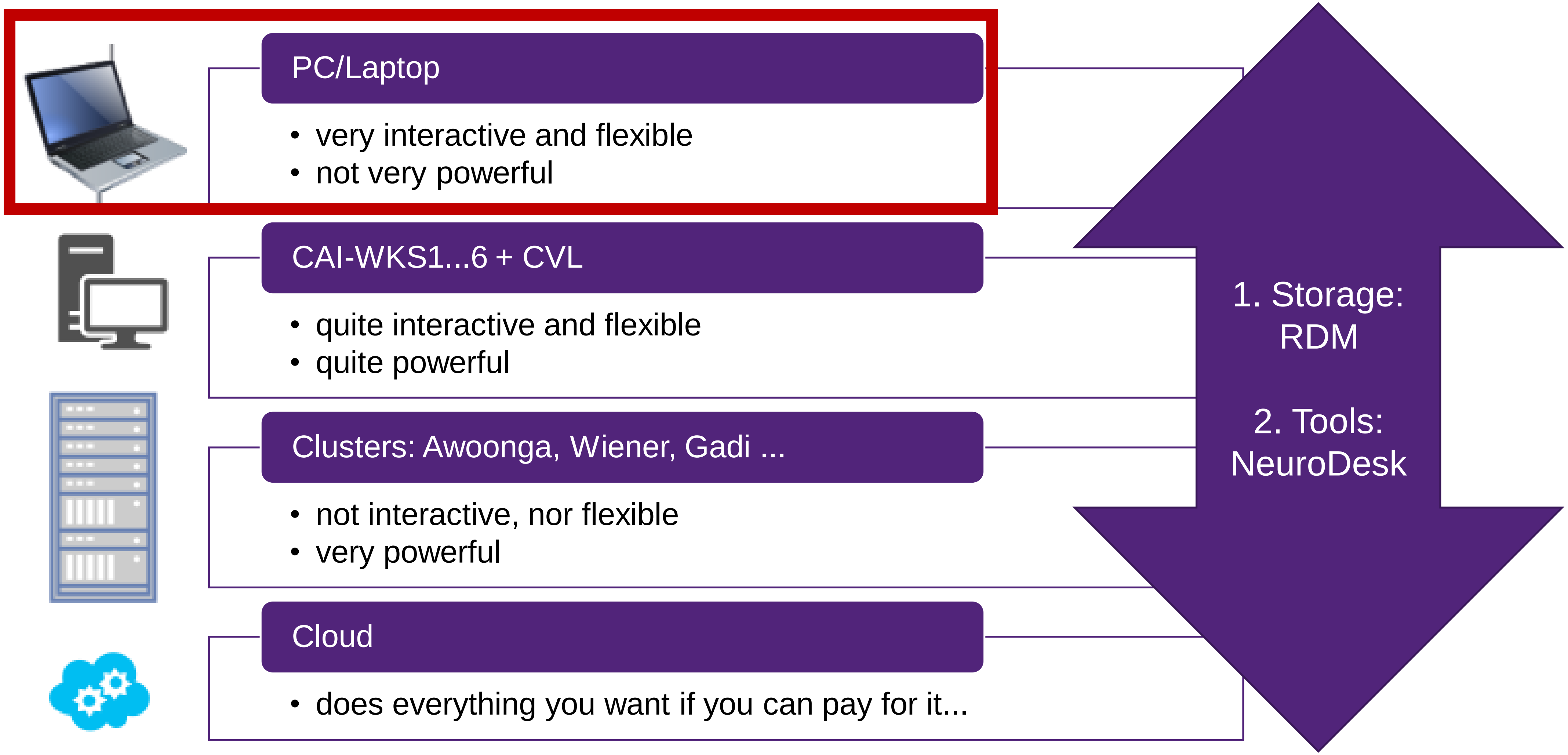
Cloud

- does everything you want if you can pay for it...

1. Storage: RDM

2. Tools: NeuroDesk

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

1. Storage: RDM
2. Tools: NeuroDesk

PC/Laptop:

4 cores
16 GB RAM

Use case

- very interactive and flexible

Access

- helpdesk@cai.uq.edu.au

RDM Storage

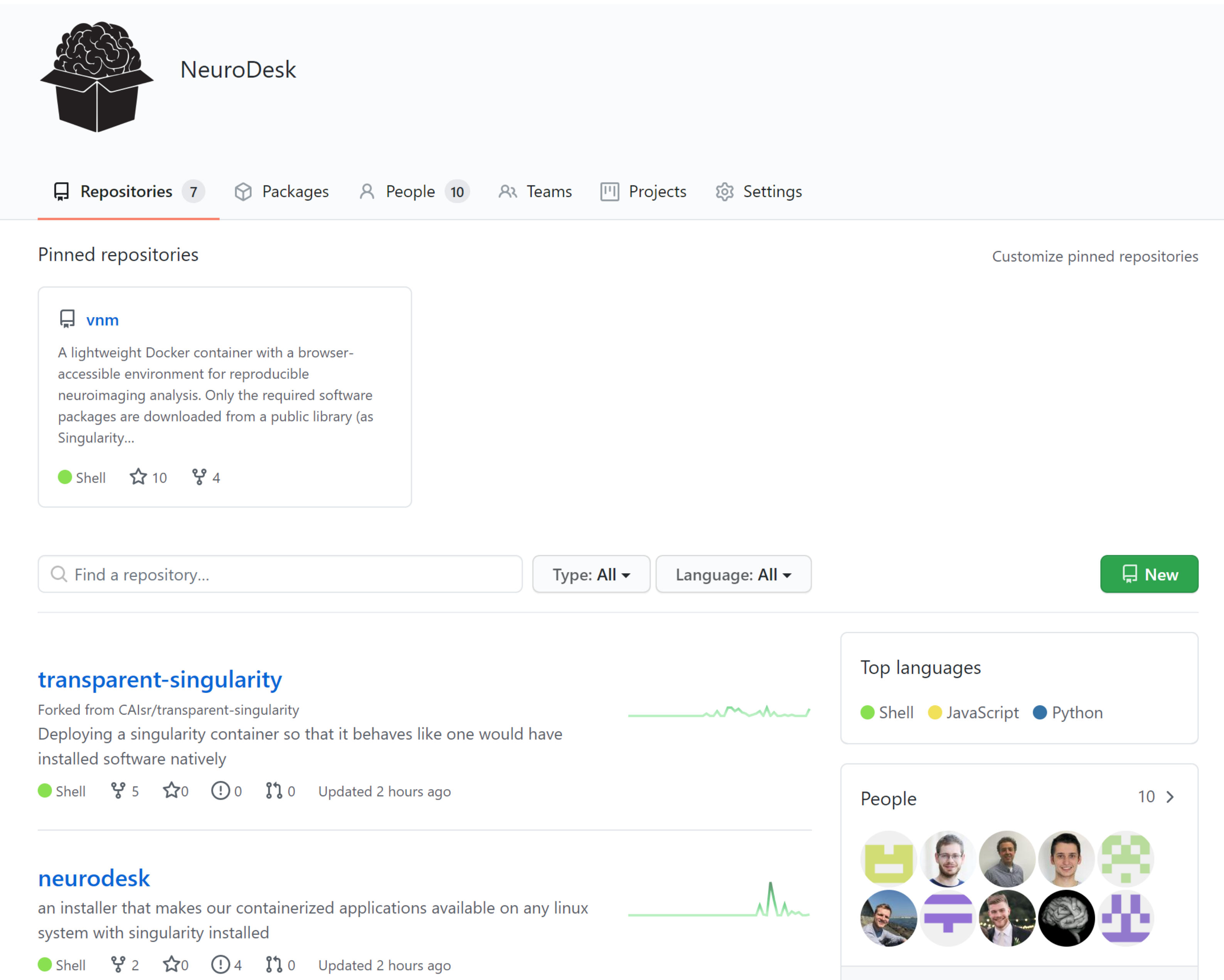
- R-drive
- UQ-InstGateway data.cai.uq.edu.au

Linux-Applications

- https://github.com/NeuroDesk/vnm/ (requires Docker)

NeuroDesk

- Community project
- Started at Organisation for Human Brain Mapping Hackathon



Pinned repositories

vnm

A lightweight Docker container with a browser-accessible environment for reproducible neuroimaging analysis. Only the required software packages are downloaded from a public library (as Singularity...

● Shell ☆ 10 Forks 4

transparent-singularity

Forked from CAIsr/transparent-singularity

Deploying a singularity container so that it behaves like one would have installed software natively

● Shell Forks 5 ☆ 0 Issues 0 Pull requests 0 Updated 2 hours ago

neurodesk

an installer that makes our containerized applications available on any linux system with singularity installed

● Shell Forks 2 ☆ 0 Issues 4 Pull requests 0 Updated 2 hours ago

Top languages

● Shell ● JavaScript ● Python

People 10

Design principles for NeuroDesk

Linux, Mac, Windows
- Docker

Scale to HPC
- Singularity

Interactive
- Full Linux desktop interface

Lightweight
- Tools are installed on demand

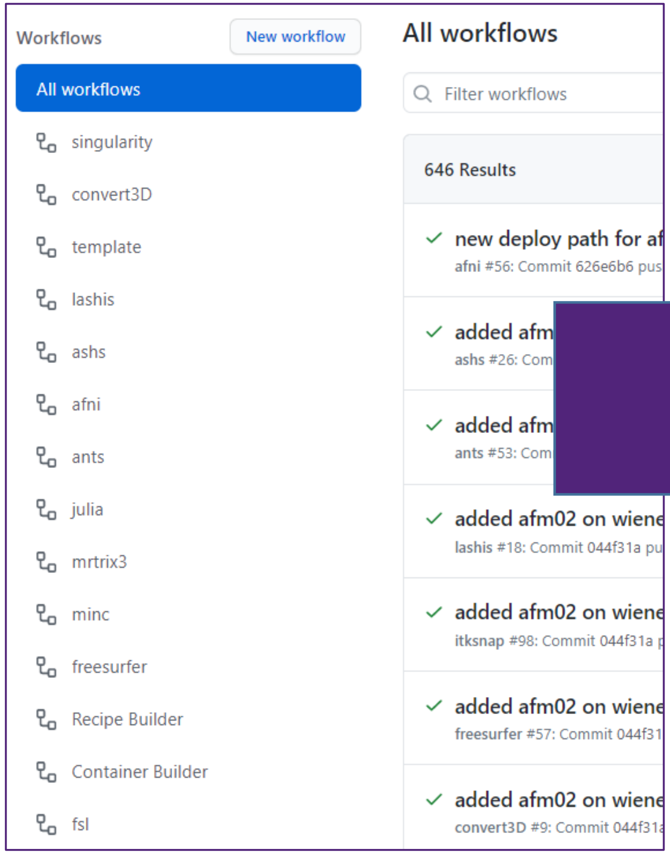
Re-use existing repositories
- NeuroDebian, conda, NeuroDocker

Architecture

Community developing recipes using conda, neurodebian, neurodocker ….

CAID – Automated Container building

Automated container building using github actions

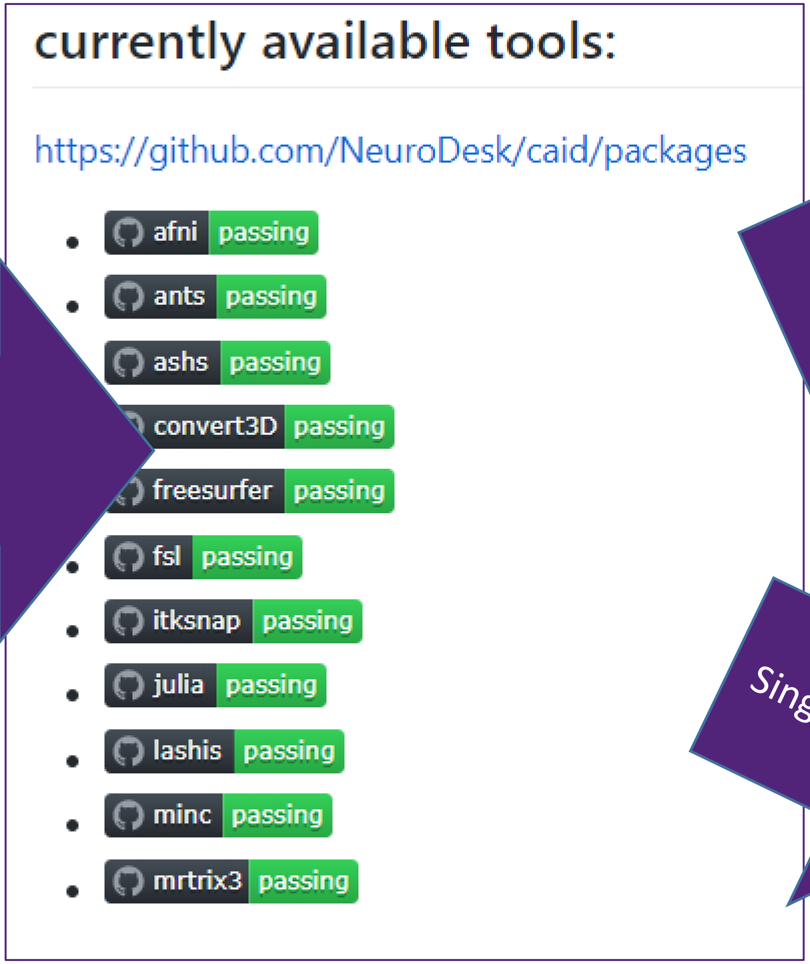


- singularity
- convert3D
- template
- lashis
- ashs
- afni
- ants
- julia
- mrtrix3
- minc
- freesurfer
- Recipe Builder
- Container Builder
- fsl

All workflows

646 Results

- ✓ new deploy path for af
  afni #56: Commit 626e6b6 pu
- ✓ added afm
  ashs #26: Com
- ✓ added afm
  ants #53: Com
- ✓ added afm02 on wiene
  lashis #18: Commit 044f31a pu
- ✓ added afm02 on wiene
  itksnap #98: Commit 044f31a p
- ✓ added afm02 on wiene
  freesurfer #57: Commit 044f31
- ✓ added afm02 on wiene
  convert3D #9: Commit 044f31a

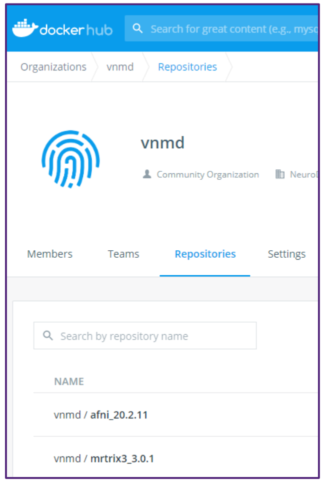


currently available tools:

https://github.com/NeuroDesk/caid/packages

- afni passing
- ants passing
- ashs passing
- convert3D passing
- freesurfer passing
- fsl passing
- itksnap passing
- julia passing
- lashis passing
- minc passing
- mrtrix3 passing

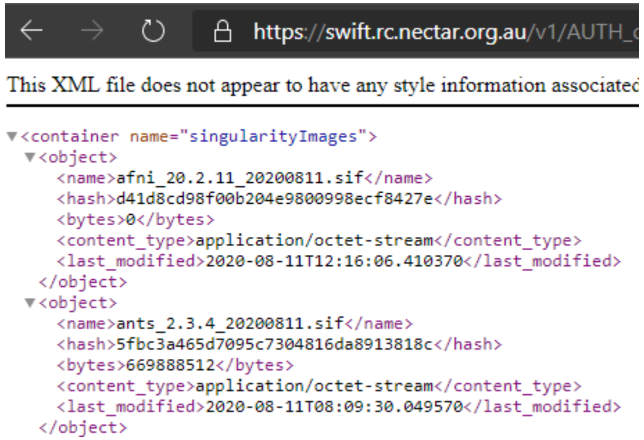
Docker



vnmd

NAME
vnmd / afni_20.2.11
vnmd / mrtrix3_3.0.1

Singularity

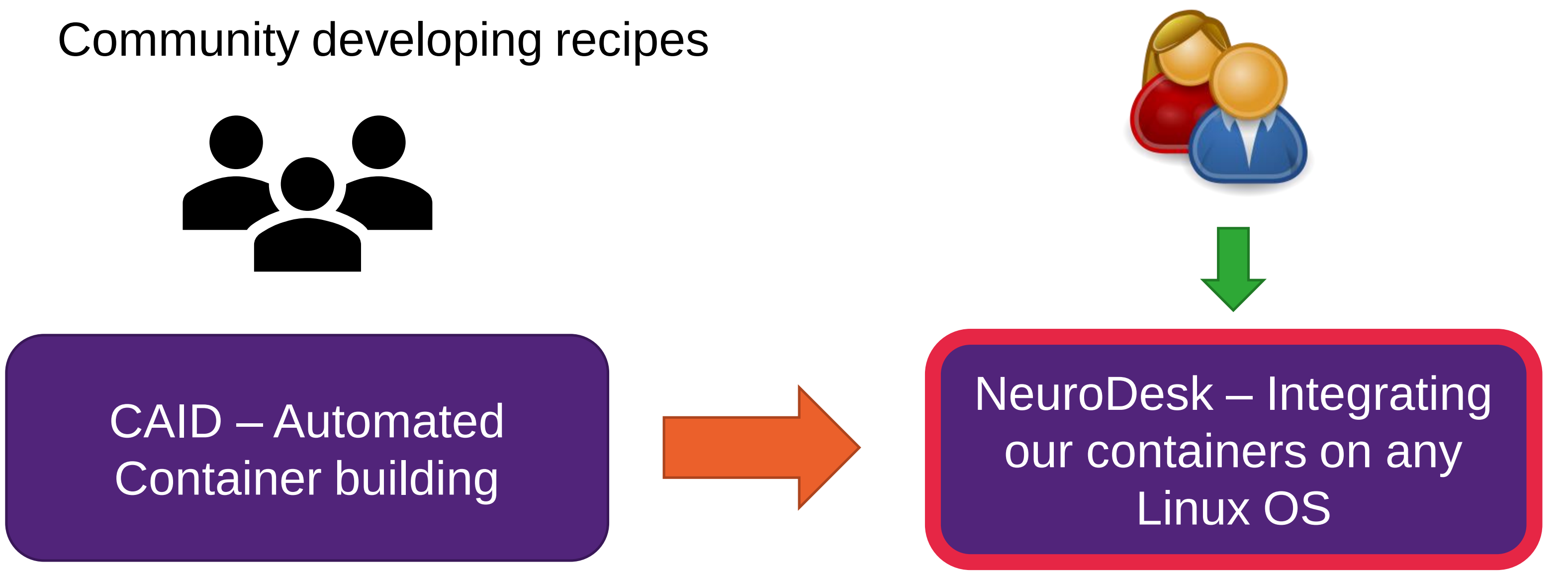


https://swift.rc.nectar.org.au/v1/AUTH_

This XML file does not appear to have any style information associated

```
▼<container name="singularityImages">
  ▼<object>
     <name>afni_20.2.11_20200811.sif</name>
     <hash>d41d8cd98f00b204e9800998ecf8427e</hash>
     <bytes>0</bytes>
     <content_type>application/octet-stream</content_type>
     <last_modified>2020-08-11T12:16:06.410370</last_modified>
   </object>
  ▼<object>
     <name>ants_2.3.4_20200811.sif</name>
     <hash>5fbc3a465d7095c7304816da8913818c</hash>
     <bytes>669888512</bytes>
     <content_type>application/octet-stream</content_type>
     <last_modified>2020-08-11T08:09:30.049570</last_modified>
   </object>
```

Architecture



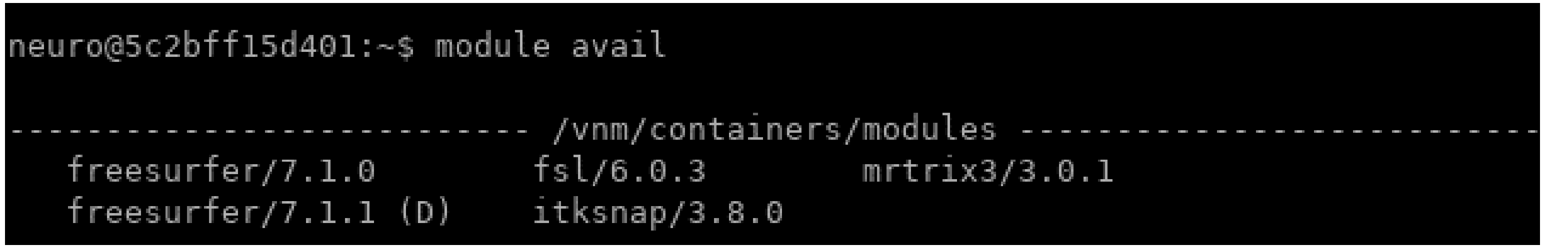
PowerUsers on Linux, HPC, CVL

Community developing recipes

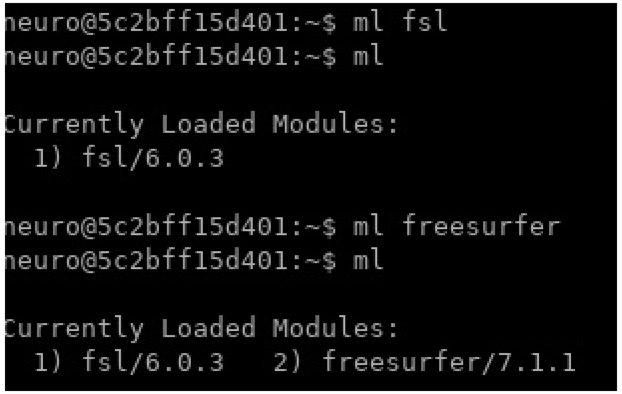
CAID – Automated Container building

NeuroDesk – Integrating our containers on any Linux OS

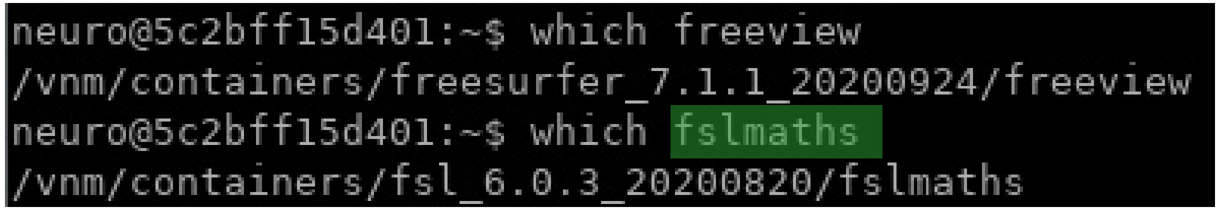
Combining tools from different Containers using modules



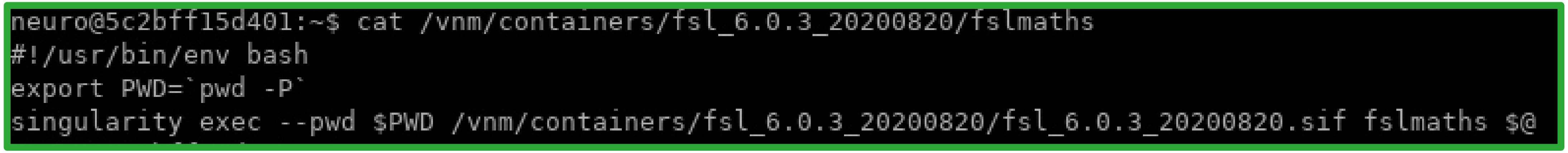
```
neuro@5c2bff15d401:~$ module avail

--------------------------- /vnm/containers/modules ---------------------------
   freesurfer/7.1.0        fsl/6.0.3        mrtrix3/3.0.1
   freesurfer/7.1.1 (D)    itksnap/3.8.0
```



```
neuro@5c2bff15d401:~$ ml fsl
neuro@5c2bff15d401:~$ ml

Currently Loaded Modules:
  1) fsl/6.0.3

neuro@5c2bff15d401:~$ ml freesurfer
neuro@5c2bff15d401:~$ ml

Currently Loaded Modules:
  1) fsl/6.0.3   2) freesurfer/7.1.1
```

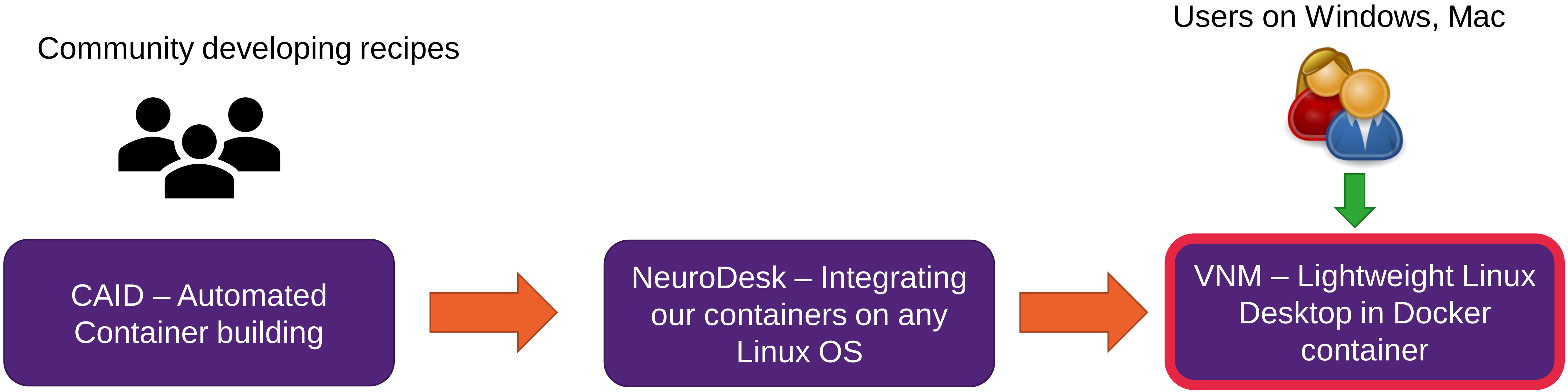


```
neuro@5c2bff15d401:~$ which freeview
/vnm/containers/freesurfer_7.1.1_20200924/freeview
neuro@5c2bff15d401:~$ which fslmaths
/vnm/containers/fsl_6.0.3_20200820/fslmaths
```



```
neuro@5c2bff15d401:~$ cat /vnm/containers/fsl_6.0.3_20200820/fslmaths
#!/usr/bin/env bash
export PWD=`pwd -P`
singularity exec --pwd $PWD /vnm/containers/fsl_6.0.3_20200820/fsl_6.0.3_20200820.sif fslmaths $@
```

Architecture



Community developing recipes

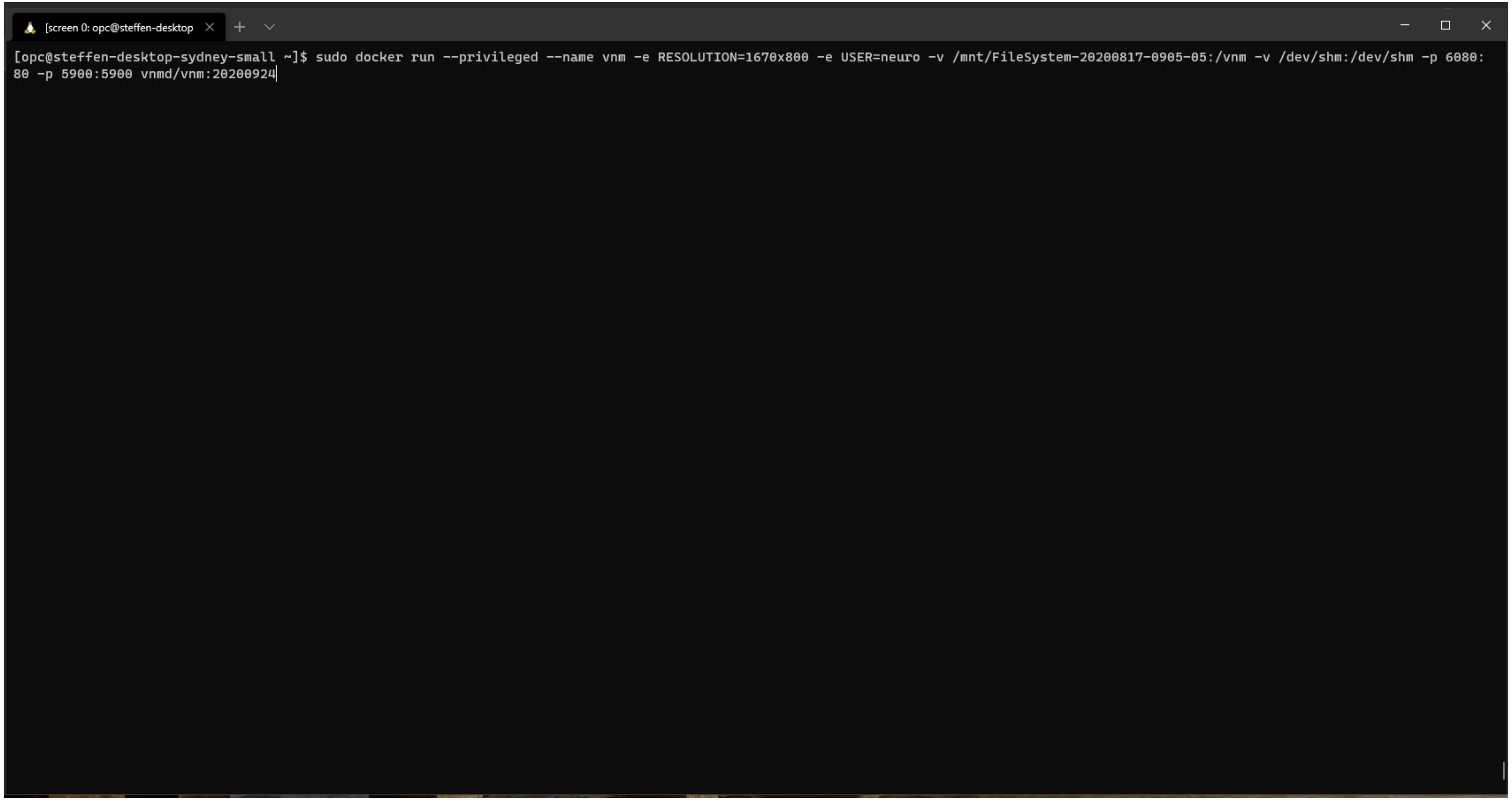
Users on Windows, Mac

CAID – Automated Container building

NeuroDesk – Integrating our containers on any Linux OS

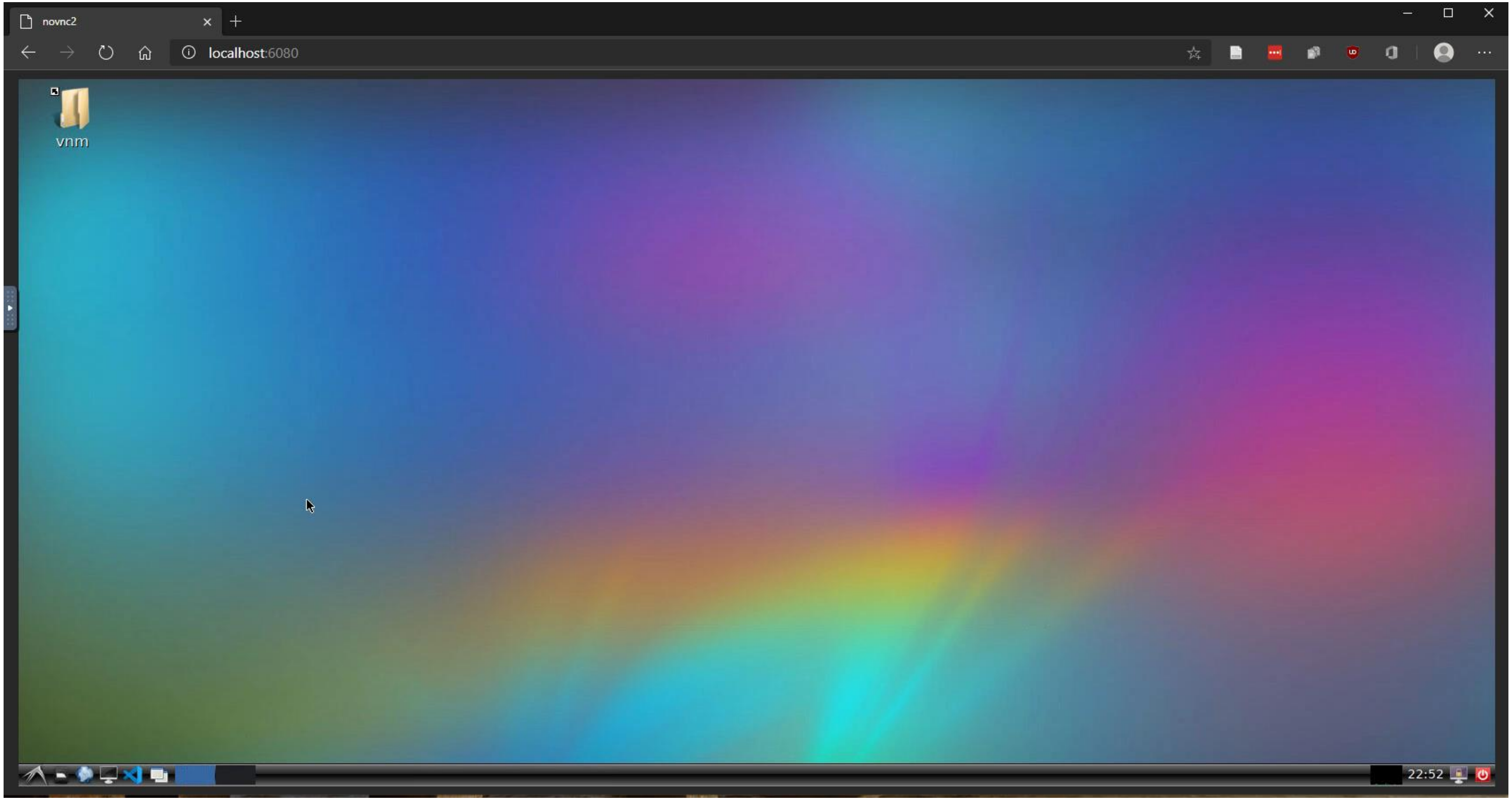
VNM – Lightweight Linux Desktop in Docker container

VNM – Interface accessible from any browser ☺

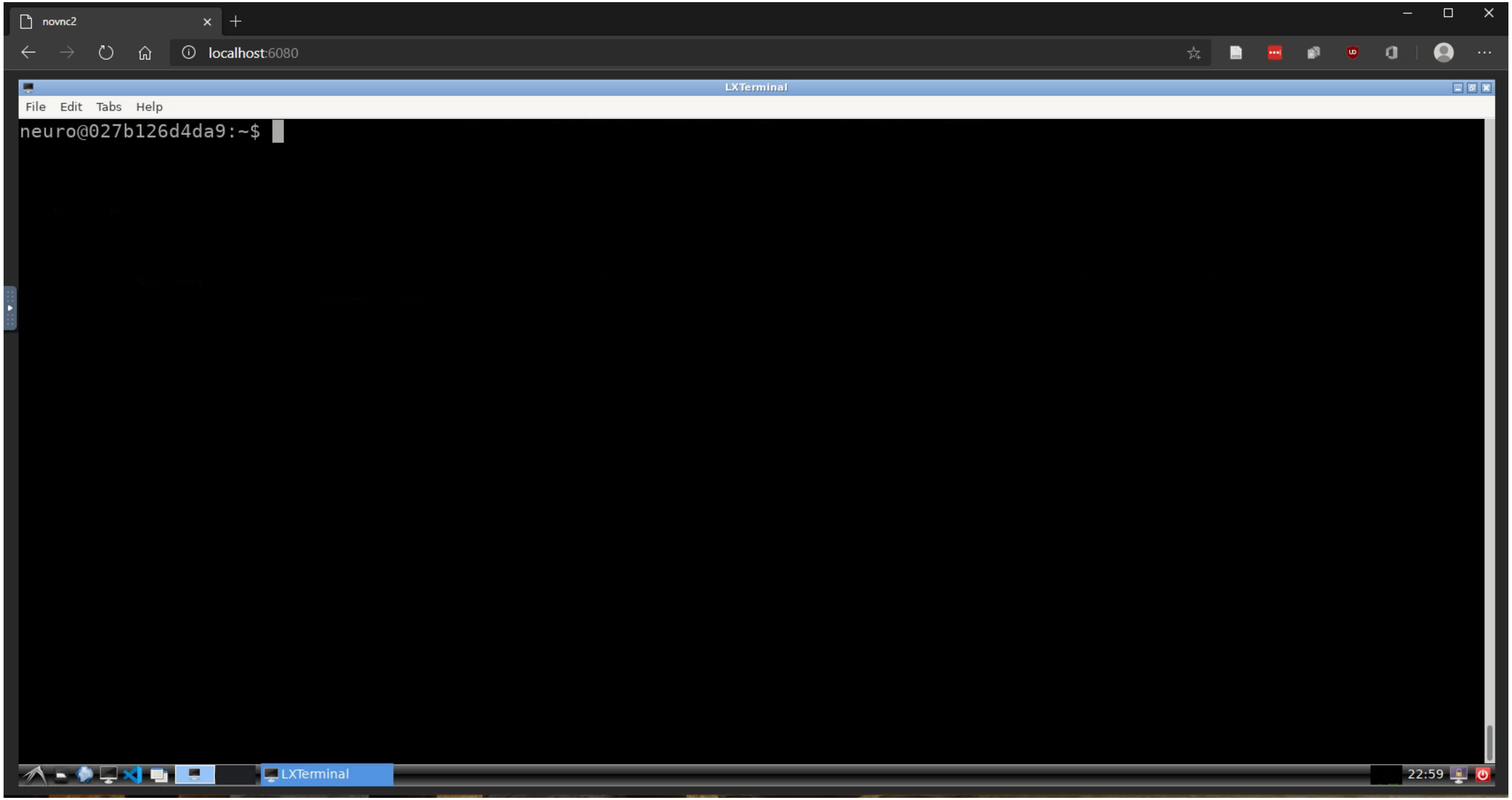


```
[opc@steffen-desktop-sydney-small ~]$ sudo docker run --privileged --name vnm -e RESOLUTION=1670x800 -e USER=neuro -v /mnt/FileSystem-20200817-0905-05:/vnm -v /dev/shm:/dev/shm -p 6080:
80 -p 5900:5900 vnmd/vnm:20200924
```

VNM – Containers are installed when needed ☺

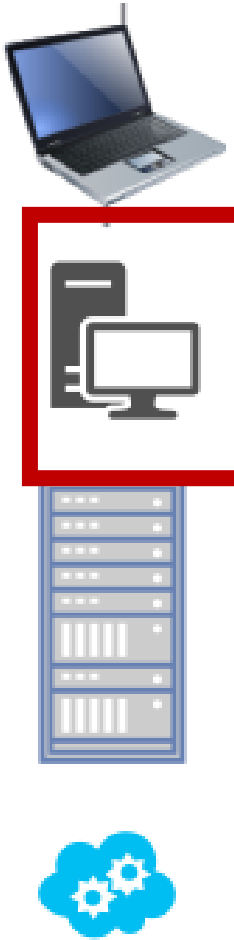


VNM – Reproducible/Scriptable via lmod module system ☺



```
neuro@027b126d4da9:~$
```

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

1. Storage: RDM

2. Tools: NeuroDesk

CAI-WKS1..6:

28 cores
192 GB RAM

Use case

- quite interactive and flexible

Access

- helpdesk@cai.uq.edu.au

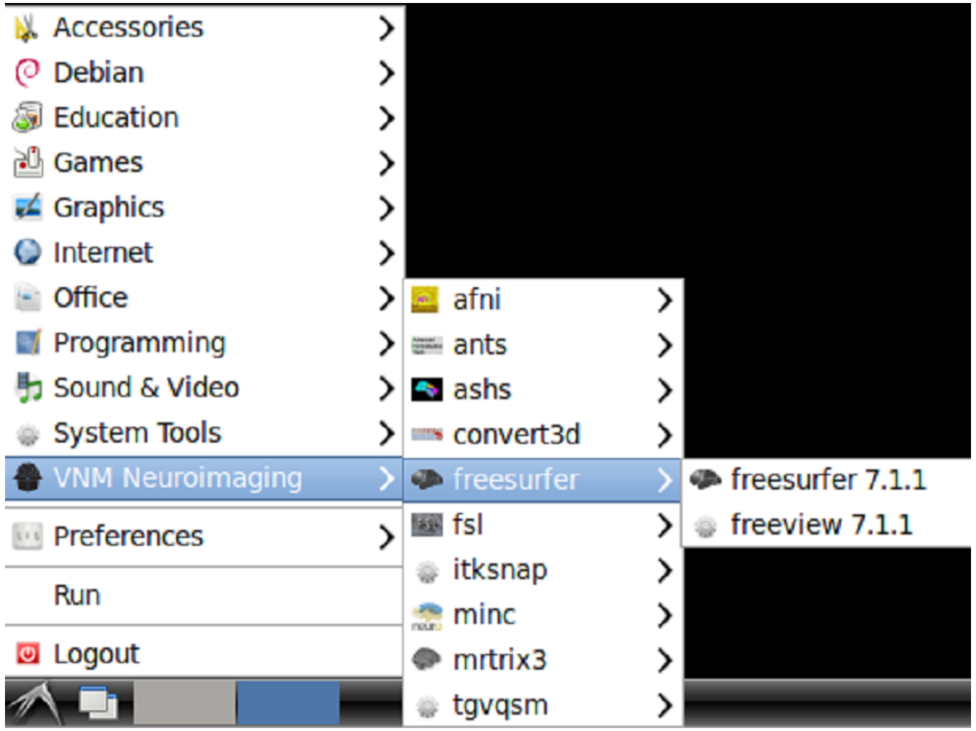
RDM Storage

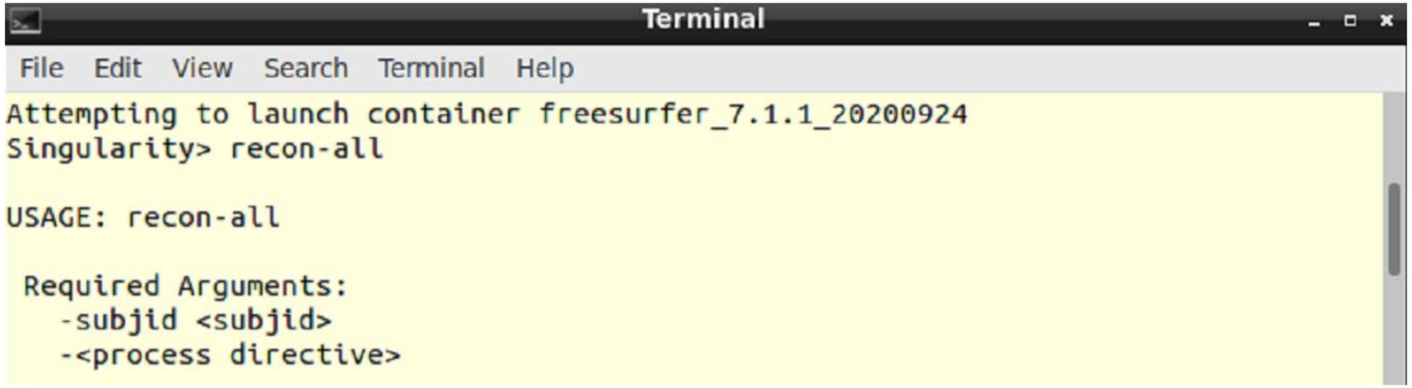
- /winmounts/uqusername/data.cai.uq.edu.au/CollectionName-Qxxxx
- /winmounts/uqusername/uq-research/CollectionName-Qxxxx

Applications

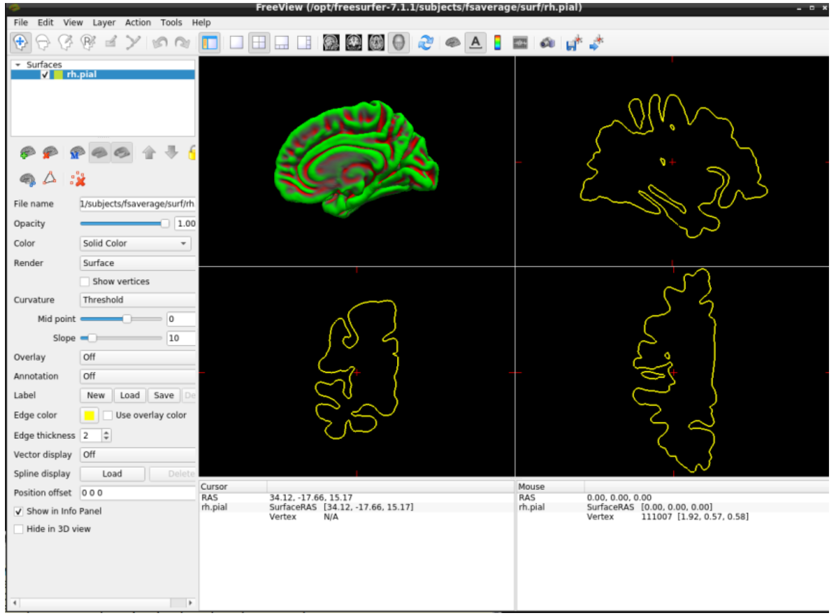
- VNM menu
- Module system

NeuroDesk on CAI-WKS1 – Menu:





```
Attempting to launch container freesurfer_7.1.1_20200924
Singularity> recon-all

USAGE: recon-all

 Required Arguments:
   -subjid <subjid>
   -<process directive>
```

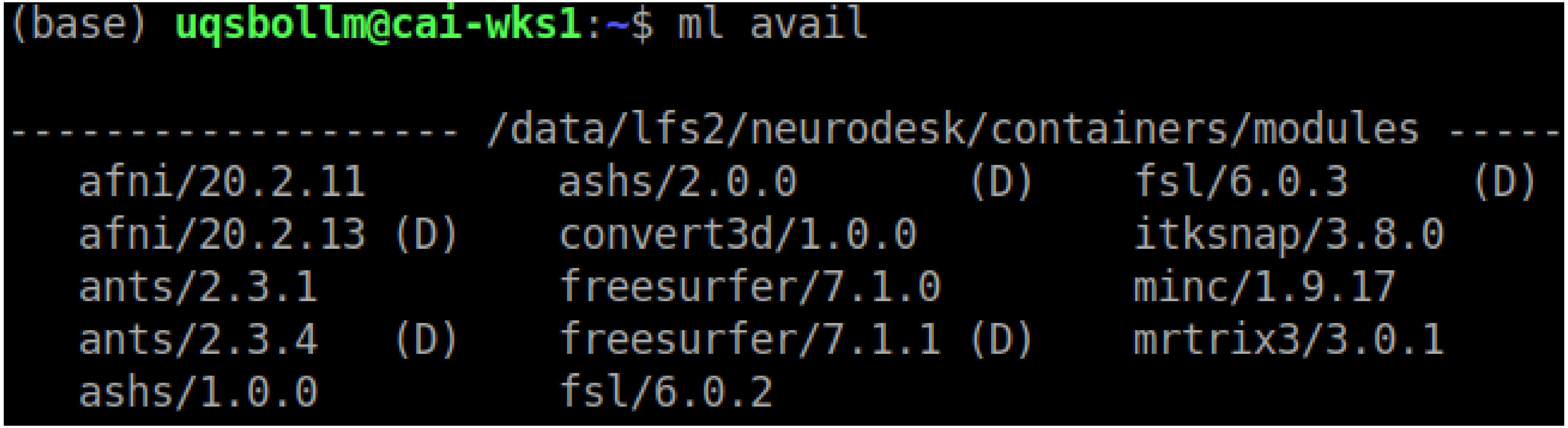


File name 1/subjects/fsaverage/surf/rh
Opacity 1.00
Color Solid Color
Render Surface
☐ Show vertices
Curvature Threshold
Mid point 0
Slope 10
Overlay Off
Annotation Off
Label New Load Save
Edge color ☐ Use overlay color
Edge thickness 2
Vector display Off
Spline display Load Delete
Position offset 0 0 0
☑ Show in Info Panel
☐ Hide in 3D view

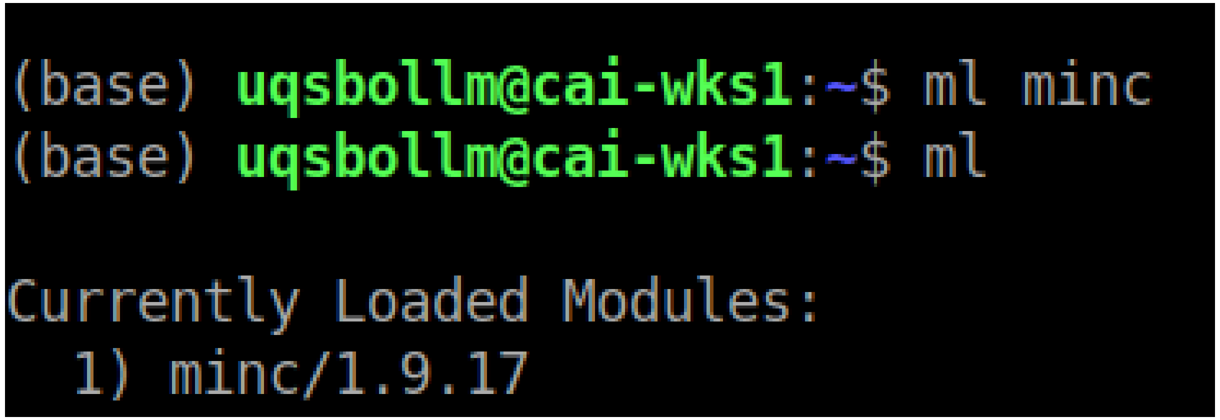
```
Cursor
RAS         34.12, -17.66, 15.17
rh.pial     SurfaceRAS  [34.12, -17.66, 15.17]
            Vertex      N/A
```

```
Mouse
RAS         0.00, 0.00, 0.00
rh.pial     SurfaceRAS  [0.00, 0.00, 0.00]
            Vertex      111007  [1.92, 0.57, 0.58]
```

NeuroDesk on CAI-WKS1..6 – Module System:

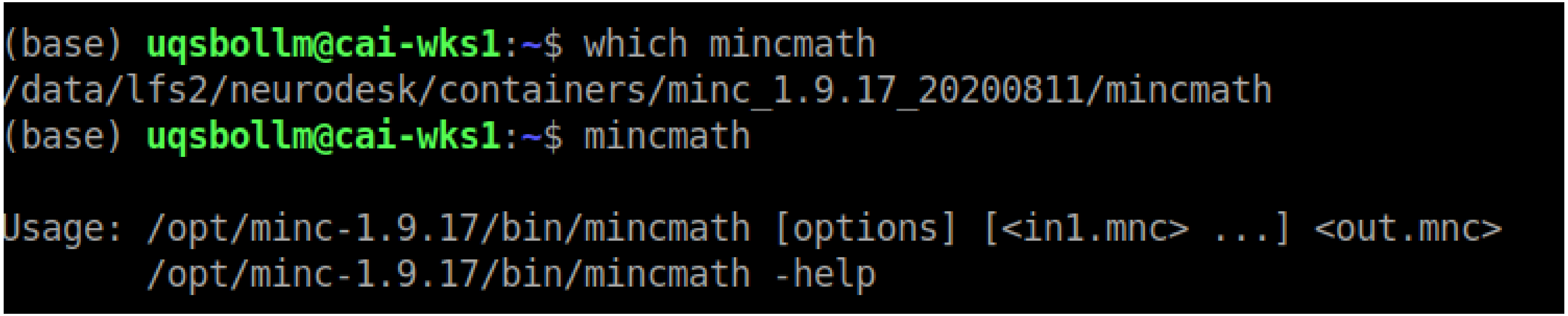


```
(base) uqsbollm@cai-wks1:~$ ml avail

------------------- /data/lfs2/neurodesk/containers/modules -----
   afni/20.2.11        ashs/2.0.0       (D)    fsl/6.0.3     (D)
   afni/20.2.13 (D)    convert3d/1.0.0         itksnap/3.8.0
   ants/2.3.1          freesurfer/7.1.0        minc/1.9.17
   ants/2.3.4   (D)    freesurfer/7.1.1 (D)    mrtrix3/3.0.1
   ashs/1.0.0          fsl/6.0.2
```

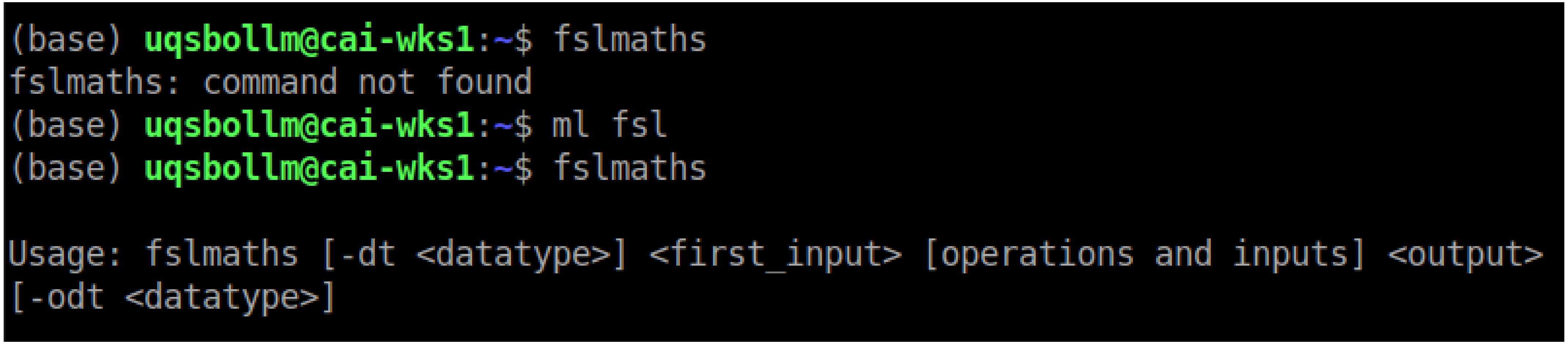


```
(base) uqsbollm@cai-wks1:~$ ml minc
(base) uqsbollm@cai-wks1:~$ ml

Currently Loaded Modules:
  1) minc/1.9.17
```

NeuroDesk on CAI-WKS1..6 – Module System:

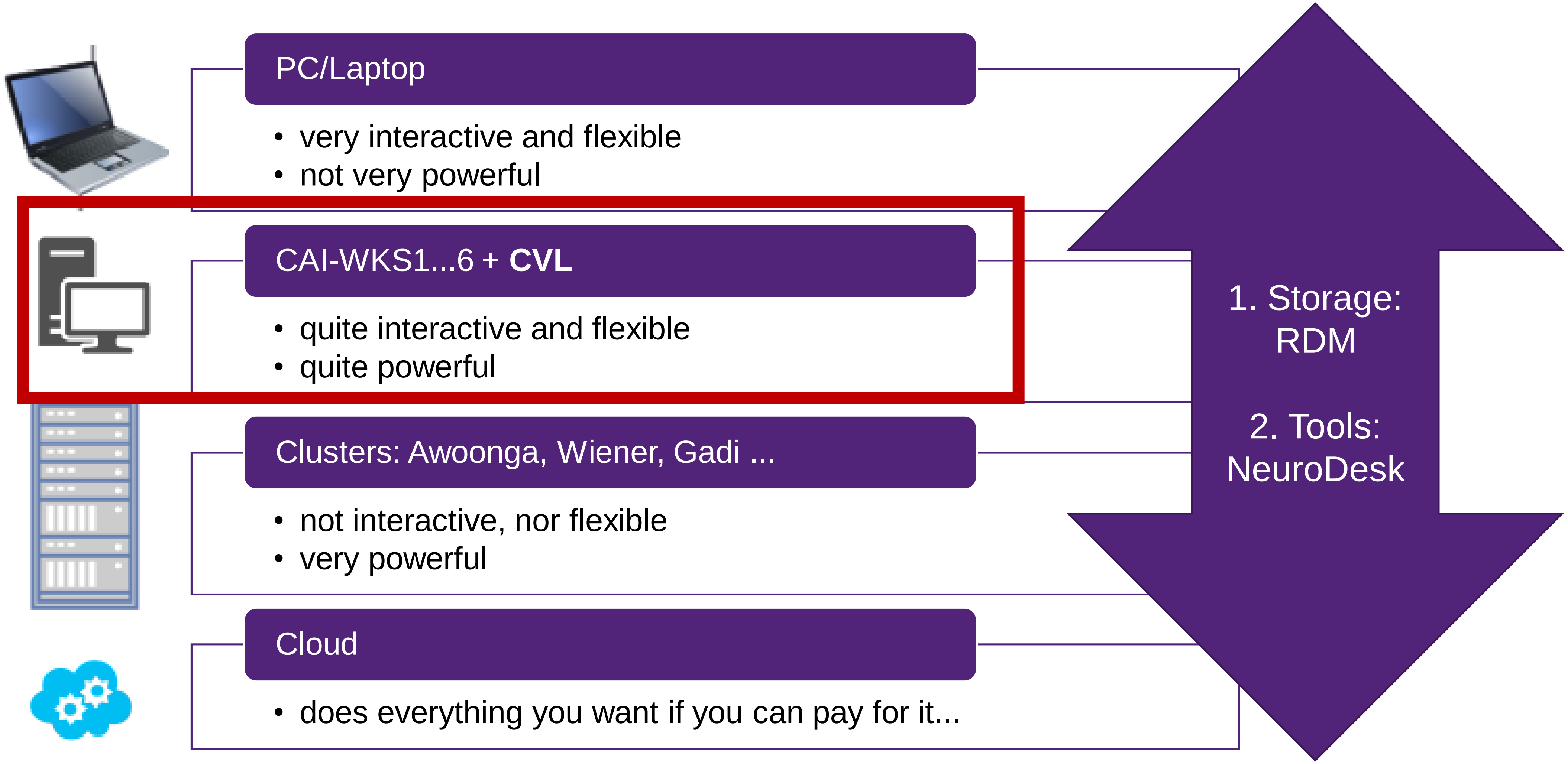


```
(base) uqsbollm@cai-wks1:~$ which mincmath
/data/lfs2/neurodesk/containers/minc_1.9.17_20200811/mincmath
(base) uqsbollm@cai-wks1:~$ mincmath

Usage: /opt/minc-1.9.17/bin/mincmath [options] [<in1.mnc> ...] <out.mnc>
       /opt/minc-1.9.17/bin/mincmath -help
```



```
(base) uqsbollm@cai-wks1:~$ fslmaths
fslmaths: command not found
(base) uqsbollm@cai-wks1:~$ ml fsl
(base) uqsbollm@cai-wks1:~$ fslmaths

Usage: fslmaths [-dt <datatype>] <first_input> [operations and inputs] <output>
[-odt <datatype>]
```

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + **CVL**

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

1. Storage: RDM
2. Tools: NeuroDesk

CVL@Wiener:

12 cores
120 GB RAM
1 GPU

Use case

- quite interactive and flexible

Access

- helpdesk@qbi.uq.edu.au

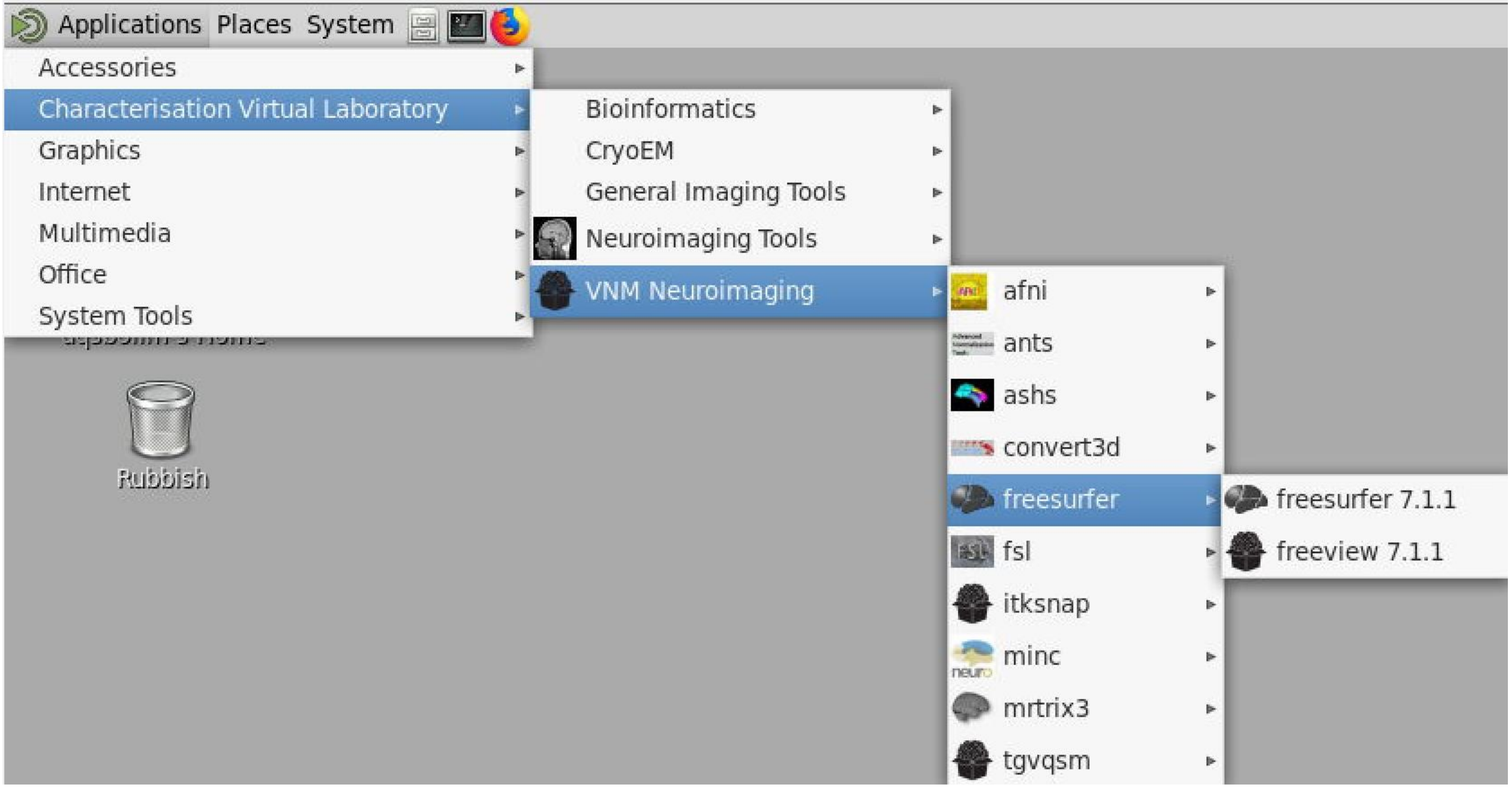
RDM Storage

- /afm02/Q[0,1,2,3]/Qxxxx

Applications

- VNM menu
- Module system

NeuroDesk on CVL@Wiener – Menu:



NeuroDesk on CVL@Wiener – Module System:

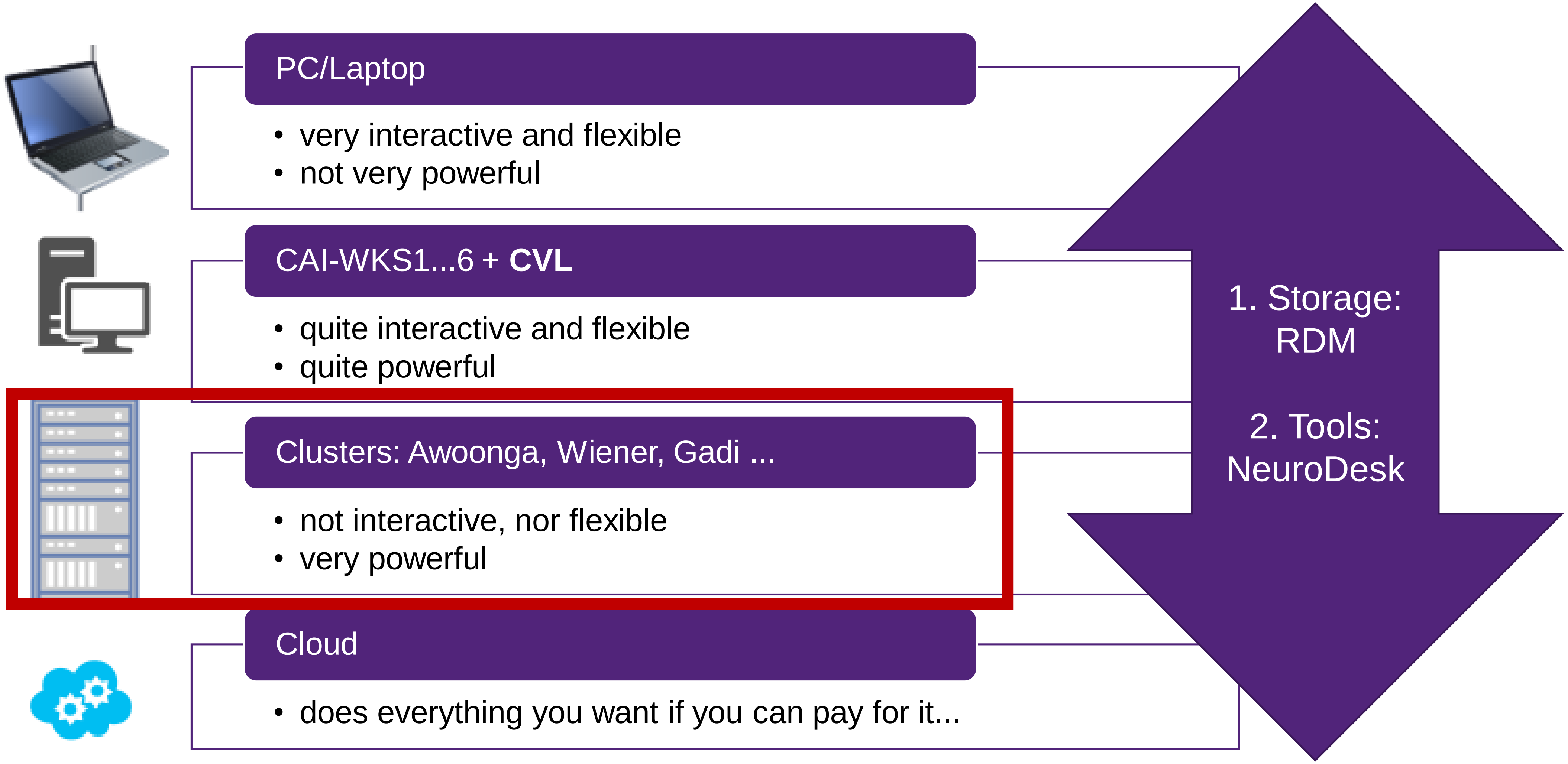
```
module use /scratch/cvl-admin/neurodesk/local/containers/modules/
```

```
module avail
```

```
module load itksnap
```

```
[uqsbollm@gpunode-2-1 containers]$ which itksnap
/scratch/cvl-admin/neurodesk/local/containers/itksnap_3.8.0_20200811/itksnap
```

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

1. Storage: RDM

2. Tools: NeuroDesk

Awoonga:

1920 cores
256 GB RAM

Use case

- not interactive, nor flexible, but high performance

Access

- rcc-support@uq.edu.au

RDM Storage

- /QRISdata/Qxxxx

Applications

- https://github.com/NeuroDesk/neurodesk
- Module system

NeuroDesk on Awoonga – Module System:

Installation:

```
uqsbollm@awoonga1:/90days/uqsbollm> git clone https://github.com/NeuroDesk/neurodesk.git
```

```
uqsbollm@awoonga1:/90days/uqsbollm> cd neurodesk/
```

```
echo "module use `pwd`/menus/containers/modules" >> ~/.bashrc
```

```
ml singularity/3.5.0
```

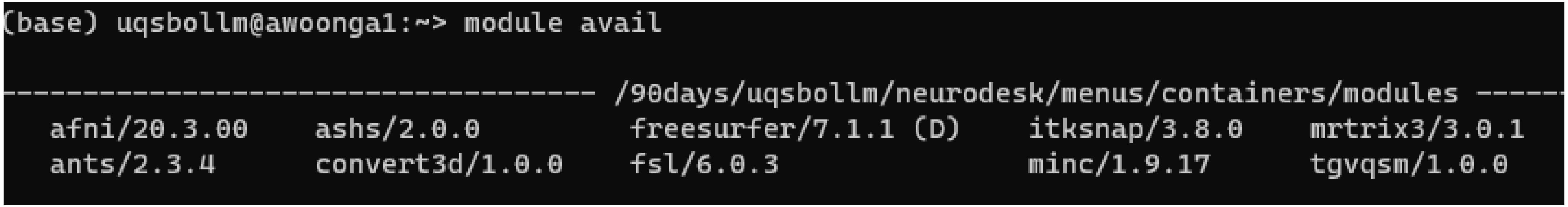
```
./neurodesk.sh --install_all_containers true
```

```
echo "module use `pwd`/menus/containers/modules" >> ~/.bashrc
```

Then logout and back in

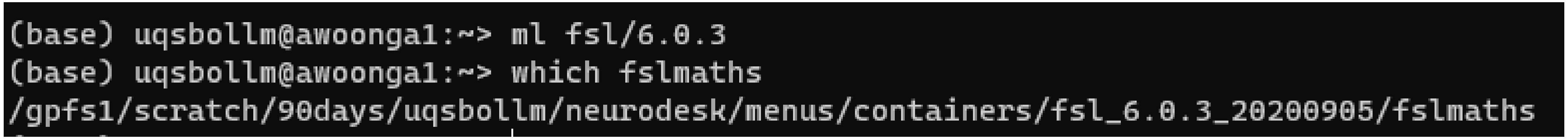
NeuroDesk on Awoonga – Module System:

Use:



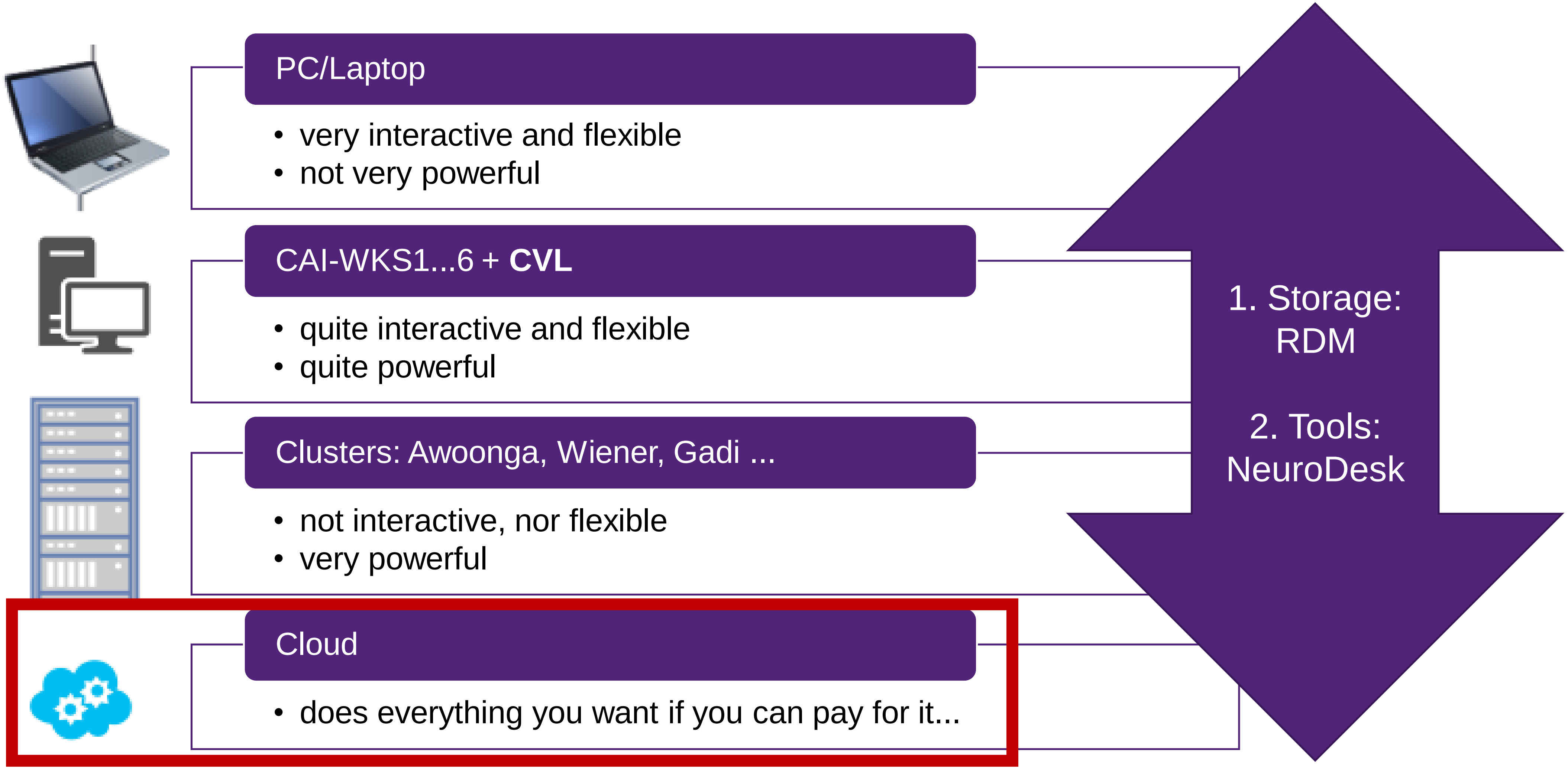
```
(base) uqsbollm@awoonga1:~> module avail

----------------------------------- /90days/uqsbollm/neurodesk/menus/containers/modules ------
   afni/20.3.00    ashs/2.0.0         freesurfer/7.1.1 (D)    itksnap/3.8.0    mrtrix3/3.0.1
   ants/2.3.4      convert3d/1.0.0    fsl/6.0.3               minc/1.9.17      tgvqsm/1.0.0
```

```
ml singularity/3.5.0
```



```
(base) uqsbollm@awoonga1:~> ml fsl/6.0.3
(base) uqsbollm@awoonga1:~> which fslmaths
/gpfs1/scratch/90days/uqsbollm/neurodesk/menus/containers/fsl_6.0.3_20200905/fslmaths
```

Goal: Run on optimal hardware for job at hand



PC/Laptop

- very interactive and flexible
- not very powerful

CAI-WKS1...6 + CVL

- quite interactive and flexible
- quite powerful

Clusters: Awoonga, Wiener, Gadi ...

- not interactive, nor flexible
- very powerful

Cloud

- does everything you want if you can pay for it...

1. Storage: RDM

2. Tools: NeuroDesk

Cloud:

Infinite cores
512 GB RAM

Use case

- interactive, flexible, but expensive

Access

- Nectar, Amazon, Google, Microsoft, Oracle …

RDM Storage

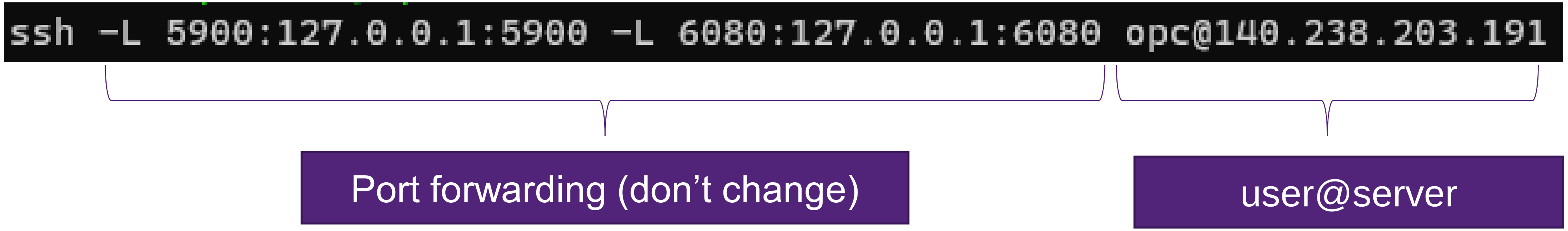
- Nextcloud client

Applications

- https://github.com/NeuroDesk/vnm

Neurodesk VNM in the cloud:

Start a compute instance on your provider of choice, then:



```
ssh -L 5900:127.0.0.1:5900 -L 6080:127.0.0.1:6080 opc@140.238.203.191
```

Port forwarding (don’t change)

user@server

```
[opc@steffen-desktop-sydney-small ~]$ sudo docker run --privileged --name vnm -e RESOLUTION=1670x800 -e USER=neuro
 -v /mnt/FileSystem-20200817-0905-05:/vnm -v /dev/shm:/dev/shm -p 6080:80 -p 5900:5900 vnmd/vnm:20201012
```

Then open browser: http://localhost:6080/ (or any vnc client on localhost:5900)

Our use case

- Researcher wants to run an analysis with Nipype (Python 3), combining tgv_qsm (Python 2), FSL 6.0.3 (Linux) and MINC 1.9.17 (Prebuilt packages only available for Ubuntu)
- Develop pipeline interactively on Windows 10 notebook
- Test analysis on pilot data on a Linux workstation running Ubuntu 18.04
- Analyse all data on a cluster running ROCKS Centos
- Visualize results interactively on Windows 10 notebook and prepare for publication
- Share analysis pipeline with readers of paper

✓ Patient data can now stay local

✓ Outdated libraries on old Centos cluster don’t matter

✓ FSL/MINC now usable on Windows

✓ Python versions now separated in containers

✓ Compiling MINC done in CI/CD pipeline

✓ Software setup reproducible

Acknowledgements

NEURODESK GITHUB CONTRIBUTORS

@aswinnarayanan	Aswin Narayanan
@civier	Oren Civier
@crnolan	Christopher Nolan
@geetah05	Geeta Hariharaputran
@kaczmarj	Jakub Kaczmarzyk
@MartinGrignard	Martin Grignard
@meganEJcampbell	MEJ Campbell
@stebo85	Steffen Bollmann
@thomshaw92	Thom Shaw

National Imaging Facility

The University Of Queensland
AUSTRALIA

SWIN BUR NE
SWINBURNE UNIVERSITY OF TECHNOLOGY