

# Introduction to the RHUL-Psychology computing system

**Tibor Auer**

RHUL Department of Psychology  
Research Fellow in MRI

# Overview

---

- **Resources**
- **Accessing resources**
- **Data**
- **Scientific software**
- **Best practices**

# Overview

---

- Resources
- Accessing resources
- Data
- Scientific software
- Best practices

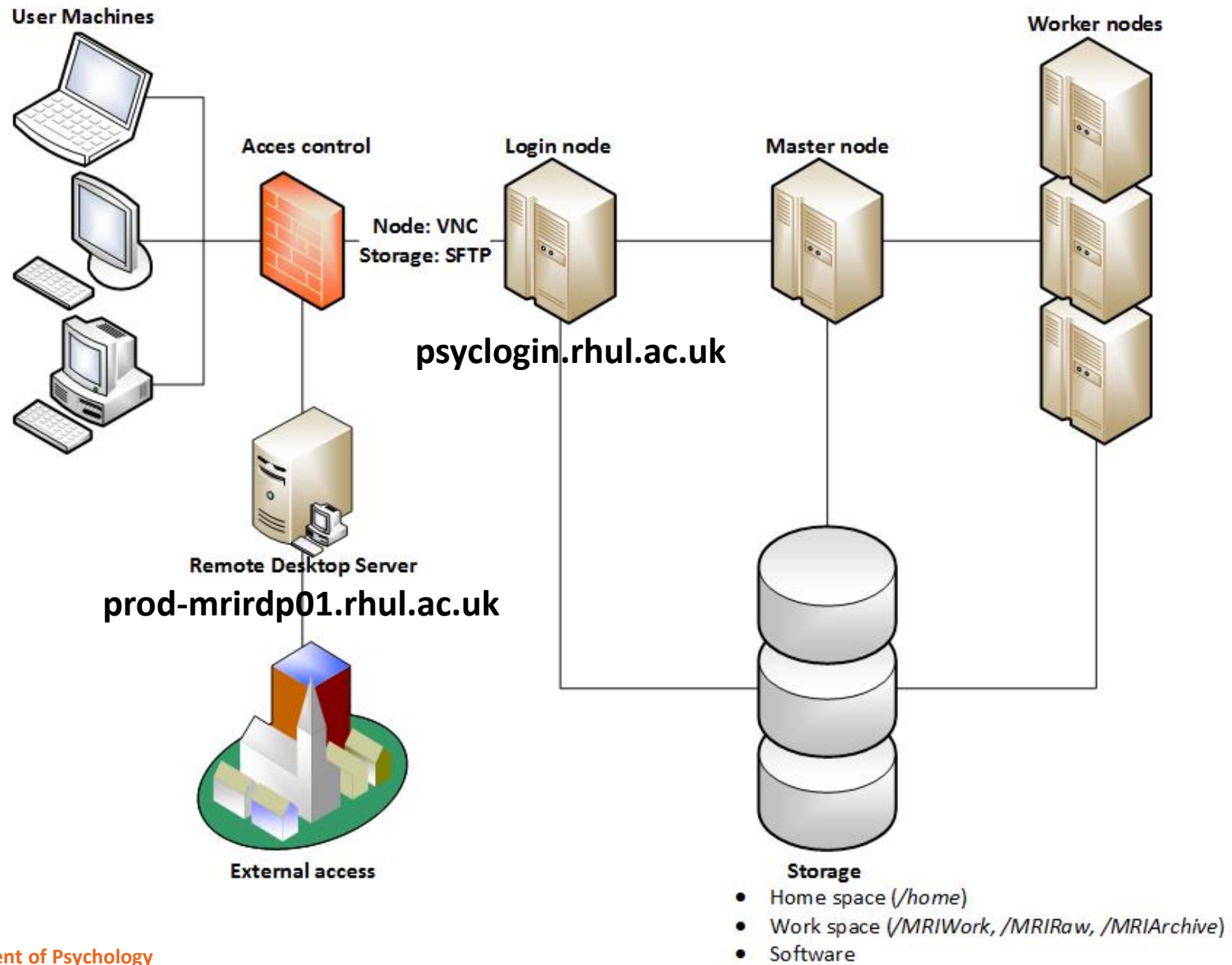


# Resources – Network Storage Spaces

---

- **Home space:**
  - 1GB quota per user
  - Personal
  - Configurations
  - Scripts
  - Figures
  - Documents
- **Work space:**
  - 1TB quota per PI<sup>1</sup>
  - Shared within the group  
Managed at <https://psycmatlab.rhul.ac.uk:10000>
  - Analysis
- **All**
  - Backed up
    - Daily (12.10am)
    - 30 days retention
- **Archive space:**
  - 1TB quota per PI<sup>1</sup>
  - Shared within the group  
Managed at <https://psycmatlab.rhul.ac.uk:10000>
  - Archiving

# Resources – Computing



# Resources – Computing

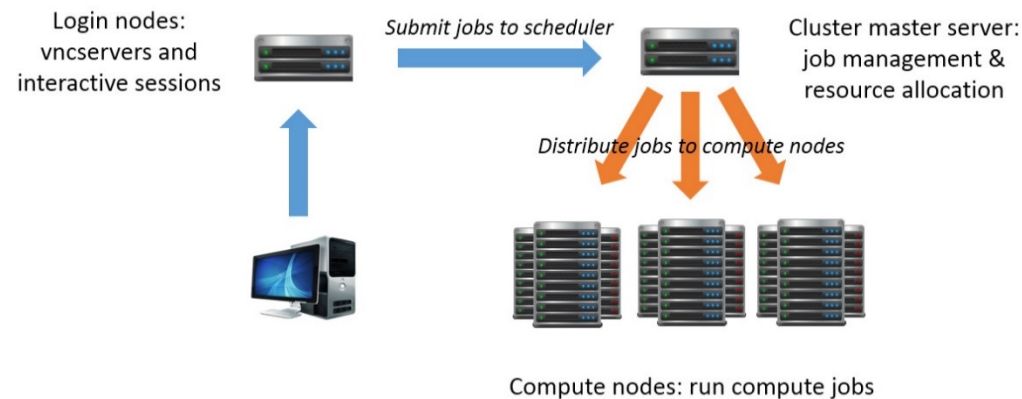
---

- **Login nodes:**
  - Name:
    - psyclogin (with Quadro M5000 – OpenGL)
  - Provides
    - Access (login): SSH
    - Interactive session: VNC
      - **Develop/debug codes**
      - **Check results**
      - **Create graphics**
  - N.B.:
    - **NOT** for running full analysis *locally* on a whole study
    - **CAVEAT:** If you clog it, you may hinder other users!

# Resources – Computing

- **Worker nodes:**

- Names: no need to know
  - psycmri01-02
- Provides
  - **Computation power for jobs and analyses** (total 20 cores and ~3GB RAM per core)
- N.B.:
  - No direct access



- **Non-interactive**
- **If you clot it, it will not affect others!**

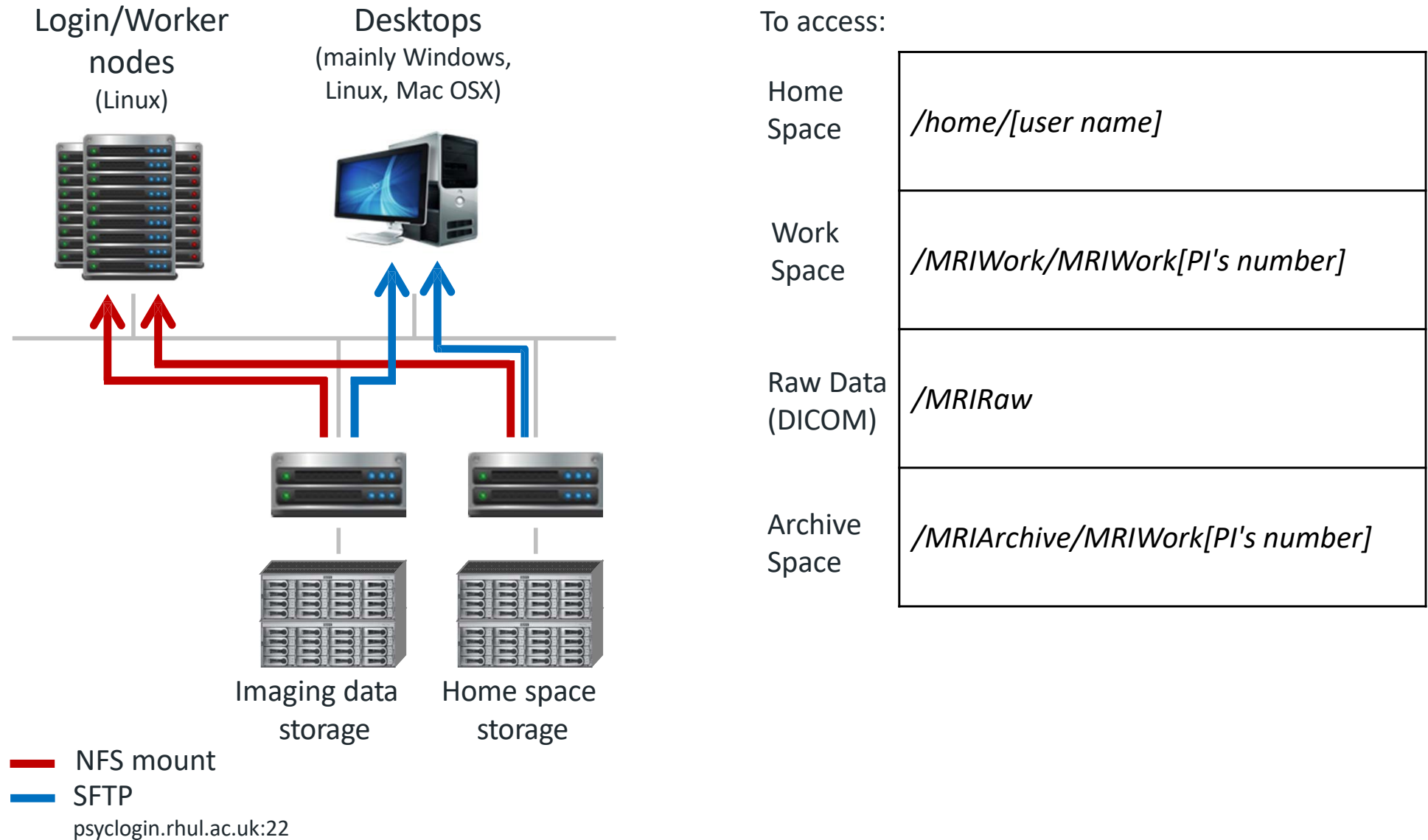
# Overview

---

- Resources
- Accessing resources
- Data
- Scientific software
- Best practices



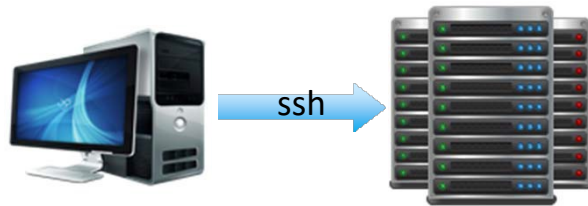
# Accessing Resources – Network Storage Spaces



# Accessing Resources – Computing

## 1. Log in using SSH<sup>1</sup> (text only)

- Linux: `ssh`
- Win: PuTTY



```
login as: vwxy123
vwxz123@psyclogin.rhul.ac.uk's password:
[vwxz123@psyclogin ~]$
```

## 2. Graphical sessions via VNC<sup>2</sup>

- Check: `ps -ef | grep <user name>.*Xvnc | grep -v grep | awk '{print $9}'`

```
[vwxy123@psyclogin ~]$ ps -ef | grep vwxy123.*Xvnc | grep -v grep | awk '{print $9}'
:7
```

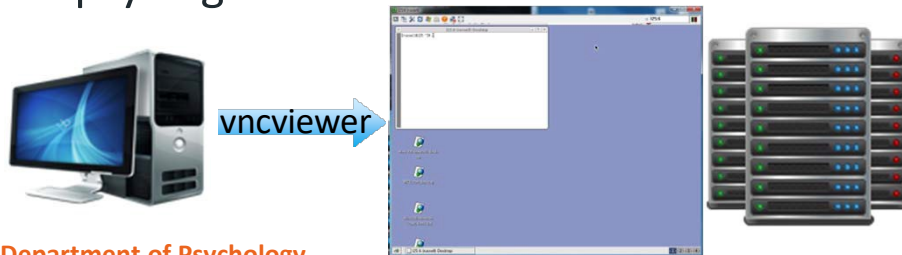
- Open: `vncserver :<desktop number>`

```
[vwxy123@psyclogin ~]$ vncserver :7
```

```
Desktop 'TurboVNC: psyclogin:7 (vwxz123)' started on display psyclogin:7
```

## 3. Win: TurboVNC Viewer

- `psyclogin.rhul.ac.uk:7`





# Accessing Resources – Computing

---

- **vncserver parameters**
  - *-kill* *:[session number]*: stop VNC session (N.B.: also kills all processes in that session)<sup>1</sup>
  - *-geometry* *XXXXxYYYY*: specify geometry (for “fit-to-win” or multi-monitor access)
    - e.g. for our regular 1920x1080 monitor *-geometry 1920x1010* gives the best fit<sup>2</sup>
- **OpenGL via VirtualGL**
  - *vglrun* “*command*”
  - Programs requiring OpenGL<sup>3</sup>:
    - *fsleyes*
    - *freeview*
    - *matlab -nosoftwareopengl*

# Accessing Resources – Computing

---

- **Remote access (RDP via VPN):**
  - Remote Desktop Connection:
    - Computer: prod-mrirdp01.rhul.ac.uk (134.219.205.209 )
    - User: CC\[user name]
    - Password: usual
  - eduroam is also “remote”
  - Once you are “in”, you can use resources via the RD:
    - Home folder
    - filezilla (SFTP)
    - putty (SSH)
    - vncviewer (VNC)



# Overview

---

- Resources
- Accessing resources
- Data
- Scientific software
- Best practices



# Data

---

- **/MRIRaw**
  - Raw data (DICOM) archived
    - Ready for SPM, FreeSurfer and aa
    - FSL and others require initial conversion<sup>1</sup>
      - mri\_convert (FreeSurfer) – linux, to any format
      - dcm2niix (<psycapps>/dcm2niix) – linux, generates BIDS JSON sidecar
      - spm\_dicom\_convert (SPM) – MATLAB
- **Structure and format (still under consideration...):**
  - **/MRIRaw/<DaT of acquisition>\_<subject ID>**: e.g. “201706131237\_19810218EIJO”
    - DaT of acquisition in YYYYMMDDhhmm format: e.g. “201706131237”
    - Subject ID as <DOB in YYYYMMDD format><1<sup>st</sup> and 3<sup>rd</sup> letter of mother’s maiden first- and surnames>> e.g. “19810218EIJO”
  - ↳ **Series\_<Series number>\_<Protocol name>**: e.g. „Series\_002\_MPRAGE”
    - Separate folders for each series
      - More efficient to manage
      - Easier to overview

# Overview

---

- Resources
- Accessing resources
- Data
- Scientific software
- Best practices

# Scientific Software

---

- **/usr/local/apps**
  - Readable by everyone, writeable by
    - matlab: user *matlab* (available for Can, Tibor)
    - psycapps: members of *psycadmin* (Can, Tibor)
- **Matlab R2015b**
  - Available on /usr/local/apps/matlab
- **SPM, FSL, Freesurfer, automaticanalysis, etc.**
  - Available on /usr/local/apps/psycapps
  - [http://www.cubic.rhul.ac.uk/wiki/doku.php?id=cluster:cluster\\_architecture#software](http://www.cubic.rhul.ac.uk/wiki/doku.php?id=cluster:cluster_architecture#software)



# Scientific Software – Wrapper Scripts

---

- **Convenience**
  - Parse any options (or use default values)
  - Set any necessary paths
- **The paths to the wrapper scripts can be configured in login scripts**
  - Some available right from the start
    - Matlab
  - Some you can specify for yourself (*/usr/local/apps/psycapps/config*)
    - FSL, FreeSurfer



# Scientific Software – Wrapper Scripts

---

- **FSL, FreeSurfer**
  - Run configurator before calling<sup>1</sup>
    - *source /usr/local/apps/psycapps/config/freesurfer\_bash [<version>]*
    - *source /usr/local/apps/psycapps/config/fsl\_bash [<version>]*
  - Use environmental variables to customise<sup>2</sup>
    - FSL: FSLOUTPUTTYPE
    - FreeSurfer: FSF\_OUTPUT\_FORMAT, SUBJECTS\_DIR, SUBJECT

# Overview

---

- **Resources**
- **Accessing resources**
- **Data**
- **Scientific software**
- **Best practices**



# Best Practice – Storage

---

- **Cluster storage spaces are backed up on daily basis – hard drives on desktops aren't!<sup>1</sup>**
- **Work space has much larger space and quotas than Home space has!**
  - Genuine documents you can't easily recreate (documents, scripts) → Home space
  - Derived data / images → Work space
- **Raw imaging data is also archived/backed up**
  - No need to make copies of the raw data.
- **Spare with space:**
  - Clean up after your analyses
  - Avoid multiple copies of the same files (aa: *aas\_garbagecollection*)
  - Single data directory for all analyses
  - Symbolic links



# Best Practice – Compute Machines

---

- **Shared resources!**<sup>1</sup>
- **Spare with computing:**
  - VNC sessions will persist until you close them or the login machine is rebooted
  - MATLAB consumes memory
    - *clear all*
    - *java.lang.Runtime.getRuntime.gc*
  - Run demanding jobs at quiet times (e.g. overnight, at weekends)
- **If your job crashes a machine, SPEAK TO IT**<sup>2</sup>

# Further Information

---

**wiki:** [http://www.cubic.rhul.ac.uk/wiki/doku.php?id=cluster:cluster\\_root](http://www.cubic.rhul.ac.uk/wiki/doku.php?id=cluster:cluster_root)