



Introduction to the CBSU computing system

Tibor Auer

MRC Cognition and Brain Sciences Unit, Methods group

With thanks to the IT

Overview

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

Overview

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



Resources – Network Storage Spaces

Home space:

- 50GB quota
- Snapshot backups

- Default

- Personal

- Things that can't be recreated
 - Scripts
 - Figures
 - Documents

Imaging space:

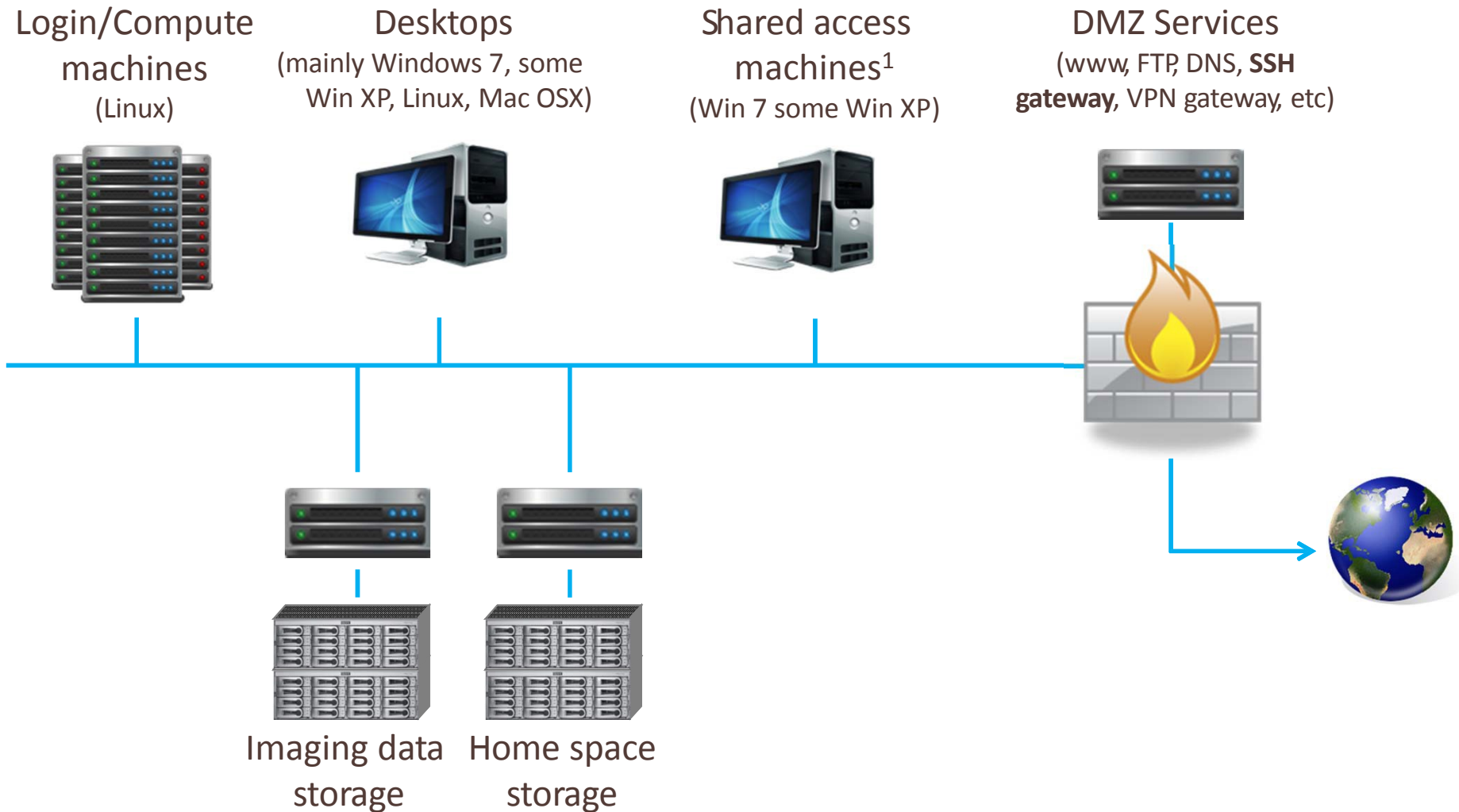
- No quotas
- No backup

- Not default
 - Available on request for people doing imaging analysis

- Shared

- Analysis

Resources – Computing





Resources – Computing

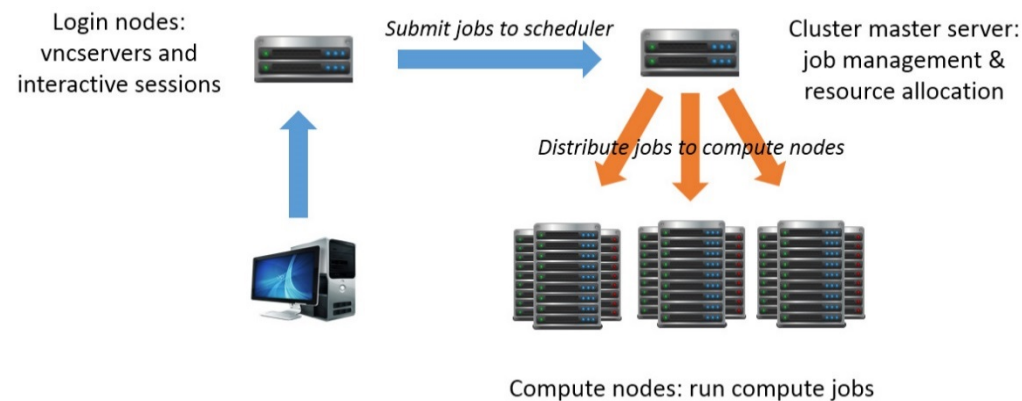
- **Login machines:**
 - Names:
 - login01 – login22
 - login-gpu01 – login-gpu03 (with Tesla/Quadro)
 - Provides
 - Access (login): SSH
 - Interactive session: VNC
 - **Develop/debug codes**
 - **Check results**
 - **Create graphics**
 - N.B.:
 - Load balancing: e.g. ssh login(-gpu)
 - **NOT for running full analysis on a whole study**
 - **CAVEAT: If you break it down, you may harm other users!**

Resources – Computing

- **Computing machines:**

- Names: no need to know
 - node-c01 – node-f08
 - node-gpu01 – node-gpu04 (with Tesla)
- Provides
 - **Computation power for jobs and analyses**
- N.B.:

- No direct access

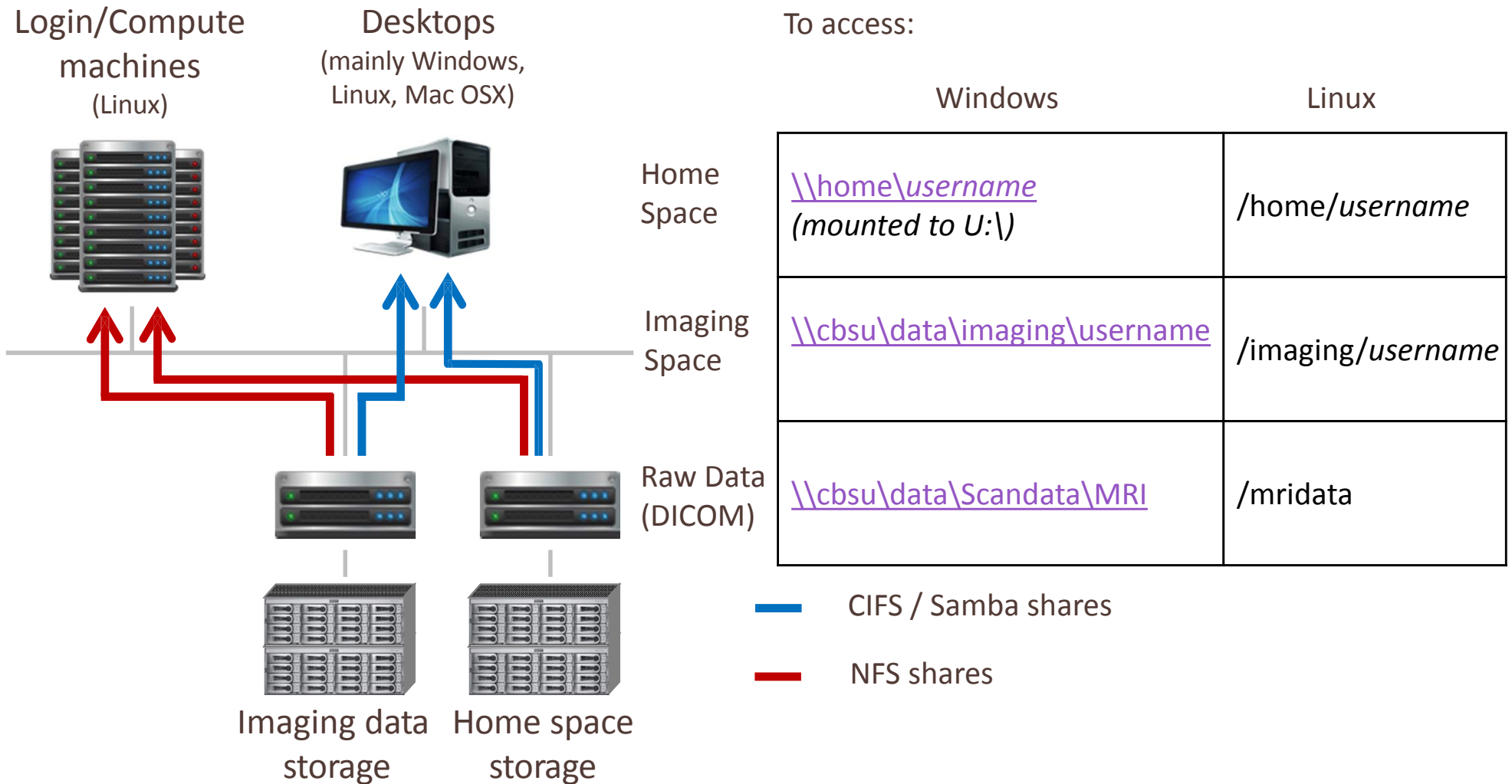


- **Non-interactive (or batch)**
- **If you break it down, it will not affect others!**

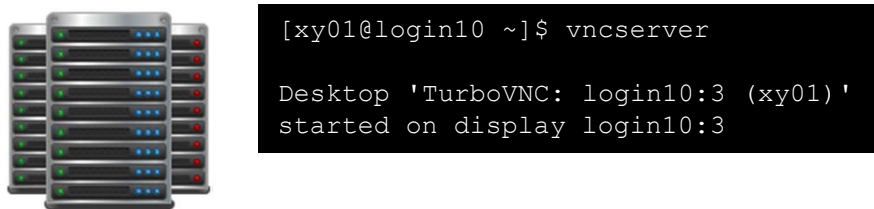
Overview

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

Accessing Resources – Network Storage Spaces



Accessing Resources – Computing



1. Pick a login machine (e.g. login10)¹
2. Log in using SSH² (text only)
 - Linux: *ssh*
 - Win: PuTTY
3. Graphical sessions via VNC³
 - *vncserver* → desktop number (e.g. 3)
4. Win: TurboVNC Viewer
 - login10:3
5. Close when finished!⁴
 1. Close TurboVNC Viewer
 2. *vncserver -kill :3*
 3. Close SSH



Accessing Resources – Computing

- **vncserver parameters**
 - *:DesktopNo*: specify desktop number
 - *-geometry XXXXxYYYY*: specify geometry (for “fit-to-win” or multi-monitor access)
 - e.g. for our regular 1280x1024 monitor *-geometry 1280x920* gives the best fit¹
- **showallvncs**
- **OpenGL via VirtualGL**
 - Only on login-gpu machines
 - *vglrun “command”*

Accessing Resources – Computing

- **Remote access (VPN Portal):** <https://portal.mrc-cbu.cam.ac.uk>
 - <http://intranet.mrc-cbu.cam.ac.uk/computing/Remote-Access>
 - eduroam is also “remote”
- Once you are “in”, you can use resources as you were in the Unit.
<http://intranet.mrc-cbu.cam.ac.uk/computing/accessing-resources-remotely/>
 - SSH
 - VNC
 - Remote desktop
 - Data



Overview

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



Data

- **/mridata/cbu**
 - Raw data (DICOM) archived
 - Ready for SPM and aa
 - FSL requires initial conversion¹
 - Format: /mridata/cbu/CBU{subjectID}_{projectID}²/{date}_{time}
 - E.g.: /mridata/cbu/CBU080705_CBU080705/20080826_152343
- **Series_{seriesNo}_{seriesID}**
 - E.g.: *Series_001_CBU_Localiser*
 - Stores DICOM files separate folders

Data

- **Series_{seriesNo}_{seriesID}**
 - Most frequent seriesIDs:
 - CBU_Localiser - quick pilot scan
 - Plan for the CBU_MPRAGE
 - CBU_MPRAGE - high resolution, T₁-weighted, structural scan
 - Plan for the upcoming scans
 - Registration
 - CBU_FieldMapping - map of magnetic inhomogeneities
 - Correcting for distortions
 - CBU_EPI_BOLD_{volNo} - fMRI with {volNo} number of volumes

Overview

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

Scientific Software

- **/imaging/local**
 - Readable by everyone, writeable by members of imagers_devel
 - /imaging/local → /imaging/local/software
- **Matlab**
 - Available on /hpc-software/matlab
- **aa**
 - Available on /imaging/local/software/AA/release-latest
 - GitHub: <https://github.com/rhodricusack/automaticanalysis/tree/devel-share>

Scientific Software

- Packages are managed by the **Methods group** and specific scientists:

Stimulus generation	E-Prime	Nadja Tschentscher
	PsychToolBox	Tibor Auer
	ScannerSynch, MEGSynch	Tibor Auer
Real-time (f)MRI		Marta Correia, Tibor Auer
Image analysis	SPM	Rik Henson, Tibor Auer
	aa	Danny Mitchel, Tibor Auer
	FSL	Marta Correia, Tibor Auer
	FreeSurfer	Kristjan Kalm, Marta Correia, Olaf Hauk, Tibor Auer
	Camino (DTI)	Marta Correia
	Fieldtrip (EEG/MEG)	Lucy MacGregor, Olaf Hauk
	Neuromag (MEG)	Olaf Hauk
Statistics	SPSS	Peter Watson
	R	Dennis Norris
Programming	Matlab	Tibor Auer, Rik Henson, Russell Thompson
	Python	Dennis Norris
	C++	Dennis Norris

<http://imaging.mrc-cbu.cam.ac.uk/imaging/AvailableSoftware>

Scientific Software – Wrapper Scripts

- **Convenience**
 - Parse any options (or use default values)
 - Set any necessary paths
 - Some: load balancing
- **The paths to the wrapper scripts are configured in login scripts**
 - Some available right from the start
 - Matlab, SPM
 - Some you can specify for yourself (*/imaging/local/linux/config*)
 - FSL, FreeSurfer

Scientific Software –Wrapper Scripts

- **Matlab**

- Links to different versions are in

- */hpc-software/bin/matlab(_version)*

- For all available: *ls /hpc-software/bin/matlab* -la*

- *matlab(_version)*

- E.g.: *matlab_2012b*

- *matlab*: default version (2009a) on a particular machine

- **N.B.:**

- You can run as many Matlab 2009a (default) wherever you want, but newer version are allowed only on one machine at once

- Some parallel computing option are available only on newer Matlab

- aa runs on any Matlab in “localsingle” mode but “qsub” mode requires Matlab 2012b

Scientific Software –Wrapper Scripts

- **SPM**
 - Optional arguments: version, machine, matlab version
 - *spm*: SPM8 fmri matlab r2013a on the lowest load machine
 - *spm 12 fmri matlab2009a*: SPM12 in fmri mode using matlab r2009a
- **N.B.:**
 - **aa requires SPM12 for maximum functionality**



Scientific Software –Wrapper Scripts

- **FSL, FreeSurfer**

- Run configurator before calling¹

- *source /imaging/local/linux/config/fsl_csh [v<version>]*
- *source /imaging/local/linux/config/freesurfer_csh [<version>]*

- Use environmental variables to customise²

- FSL: FSLOUTPUTTYPE
- FreeSurfer: FSF_OUTPUT_FORMAT, SUBJECTS_DIR, SUBJECT

Overview

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



Best Practice – Storage

- **Home space is backed up – hard drives on desktops aren't!**¹
- **Imaging space has no quotas – Home space has!**
 - Genuine documents you can't easily recreate (documents, scripts) → Home space
 - Derived data / images → Imaging space
- **Raw imaging data is archived off-site**
 - 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!**¹
- **Spare with computing:**
 - VNC sessions will persist until you close them or the host machine is rebooted
 - Limit the number of running VNC sessions
 - Re-use old VNC sessions or
 - Kill unnecessary VNC sessions
 - Limited number of MATLAB licences
 - Close SPM/MATLAB when finished
 - Run demanding jobs at quiet times (e.g. overnight, at weekends)
- **If your job crashes a machine, SPEAK TO IT**²

Further Information

wiki: <http://imaging.mrc-cbu.cam.ac.uk>

Computing intranet page: <http://intranet.mrc-cbu.cam.ac.uk/computing>

CBSU Forum: <http://forum.mrc-cbu.cam.ac.uk/qa>