

Introduction to the iCSF (Incline) & CSF (Computational Shared Facility)

-
- WELCOME!
 - Please take a seat near the front.
 - Course materials can be found at
<http://ri.itservices.manchester.ac.uk/course/icsf-csf/>
 - There is no sign-sheet. Attendance is recorded using a feedback form (see above link)
 - <https://ri.itservices.manchester.ac.uk/hpc-help>

Chris Grave, Research IT, IT Services

its-ri-team@manchester.ac.uk

Who we are - Research IT

- Central Group in IT Services offering a variety of researcher specific services
 - HPC/Compute
 - Applications support
 - Software developers for research projects
 - Storage for Research Data
 - Virtual Machine Service
 - Advice regarding external services Training courses
 - Many other topics of relevance to researchers
- Blog and newsletter
- Events: Research IT Club, Drop-in sessions
- More Info - <https://research-it.manchester.ac.uk/services/>

Presentation Overview

- Introduction to the iCSF (incline) & CSF
- Research Virtual Desktop Service
 - Exercise 1
- Research Data Storage (RDS) & CIR Filesystems
 - Exercise 2
- Using the iCSF (incline)
 - Exercise 3
- Using the CSF
 - Exercise 4
- Other Research IT Infrastructure Provided Services

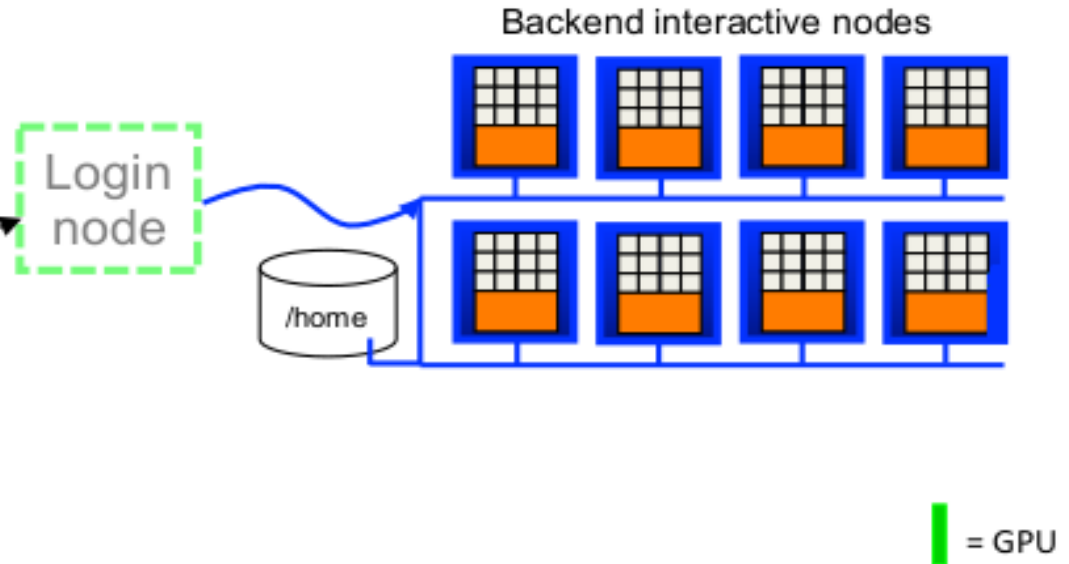
Introduction to the iCSF (Incline) & CSF

Why use the iCSF & CSF?

- Offload computational work from your PC
 - Do larger simulations/analyses that wouldn't fit on your PC
 - Work in an interactive manner using a GUI (iCSF)
- Variety of software applications
 - MATLAB, R, Stata, SAS, many more
- 1-core (serial) jobs or many-core (parallel) jobs = speedup
 - A core is a processing unit in a CPU/processor
 - Modern CPUs have several cores: your PC/laptop may have a single CPU with 4,6,8 cores; the iCSF (32 cores per node), CSF (up to 168 cores per node), a lot more in total.
- Tackle larger problems involving larger datasets
 - Lots of memory 1.5TB, 2TB, 4TB nodes

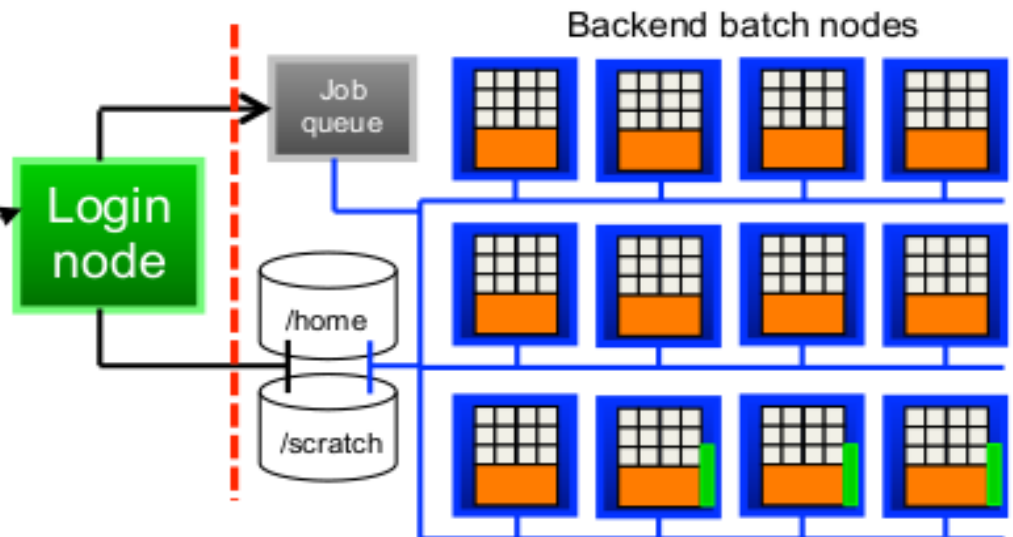
iCSF

- Run interactively
- A backend node becomes your workstation
- Start application GUI



CSF

- No direct access to backend nodes
- Batch system runs jobs on backends without GUIs



iCSF vs CSF Differences

iCSF - run interactively

- After login a backend node becomes your 'shared' workstation
- Load and interact with applications via a GUI inc., MATLAB, Rstudio, Stata, SAS
- Easier to use then CSF
- Some command line use
- Instant access to resources

CSF - run in batch

- No direct access to backend nodes
- All jobs are submitted via a batch system jobs run on backends without GUIs
- Steeper learning curve
- Command line heavy
- Greater computational Resource
- Large Scratch Filesystem

iCSF vs CSF Similarities

- Both are a collection of servers (compute nodes) linked together to form two large systems
- Operating system on both is the same – Linux (Red Hat Based)
- They share the same home directory (similar to your P Drive - a private area for your files)
- Both platforms are built based on a contribution model, i.e. research groups cost in cost of hardware in research grant proposals.
- Limited Free at the Point of use offering (FATPOU) on both platforms.

Is the iCSF suitable for my work?

Yes if

- You need to work interactively
- Have low to moderate computational requirements
- Are unfamiliar with the Command Line.
- Are developing or testing code

It also

- Complements other CIR platforms
- Provides a first step into computational work

Unsure

email its-ri-team@manchester.ac.uk we're here to help.

Is the CSF suitable for my work?

Yes, if

- You have moderate to high computational needs
- You need to run many 'Jobs' e.g. Simulations/Analyses at once.
- You need GPUs

You will need to

- Be a little more familiar with a Linux Command Line.
- use a batch system
- Don't worry, we will teach you some of the basics today
- We have 1 day training course specifically for the CSF
- Lots of docs, and friendly support team
(its-ri-team@manchester.ac.uk)

Networking & Security

- The iCSF/CSF has a private (campus only) IP address & firewall
 - Controls connections to and from the system
- **On campus**
 - Connect from any PC/laptop with a wired connection or University wireless and Eduroam
 - No need for VPN but will need to authenticate using your 2FA device, e.g.

```
myssing@i-ulit6368381 ~ $ ssh -X myssing@csf.itservices.manchester.ac.uk
Password:
Duo two-factor login for myssing@
Enter a passcode or select one of the following options:

1. Duo Push to +XX XXXX XX6353

Passcode or option (1-1): 1
```

- **Off campus:** <http://ri.itservices.manchester.ac.uk/wfh/>
 - University Global Protect VPN Gateway with 2FA
 - **Must** be running
 - Allows you to the CSF & iCSF
 - Can be some issues with some ISPs, see above WFH webpage
- Further documentation:
<http://ri.itservices.manchester.ac.uk/csf3/getting-started/connecting/>

How do I Connect to the iCSF & CSF?

- Via the **Remote Virtual Desktop Service (RVDS)** - focus of today's exercises.
- Directly via **SSH** using a Terminal or suitable Terminal Emulator such as e.g. MobaXterm, Putty - NOT RECOMMENDED for iCSF
- What operating system are you connecting from, Windows, Linux, OSX?
 - **Windows** - users need to install a free terminal app called MobaXterm <http://mobaxterm.mobatek.net/download-home-edition.html>
 - **Linux and Mac** - both have a terminal application by default
 - **Mac** – depending on the version of OSX you may need to install X-Quartz in order to view GUI
<http://xquartz.macosforge.org/>

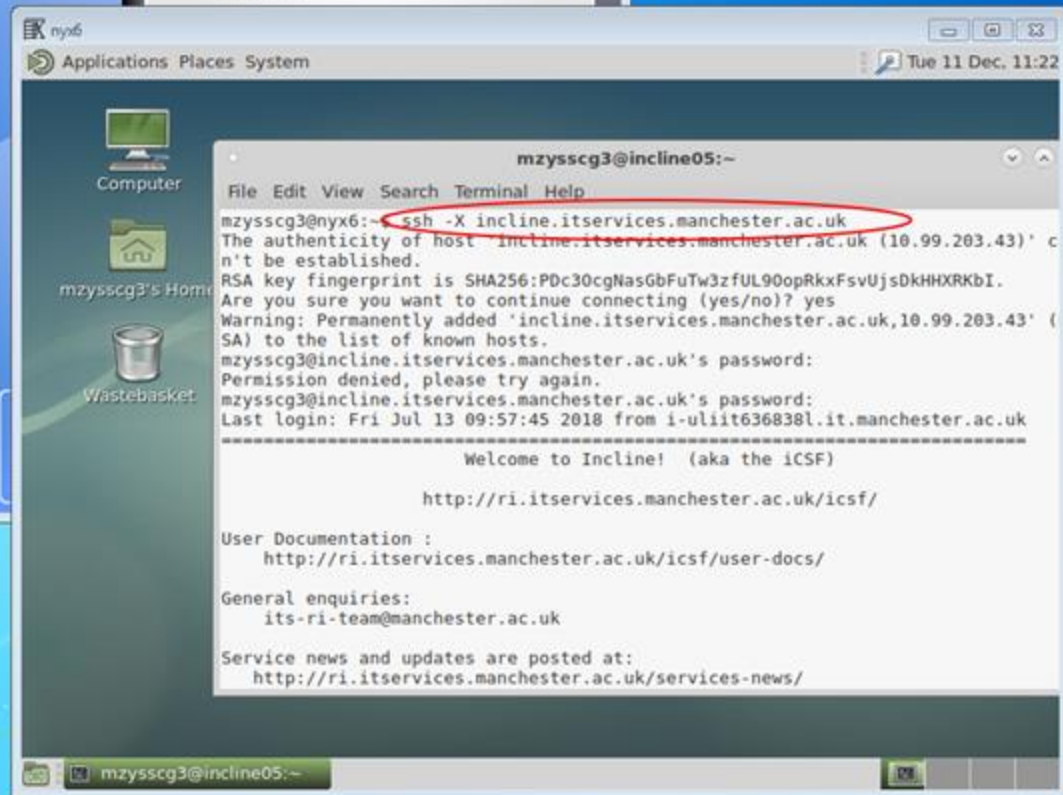
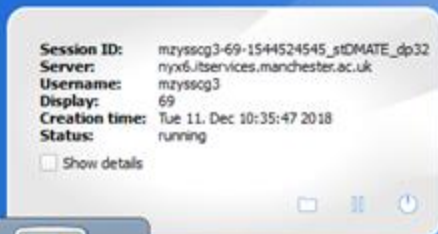
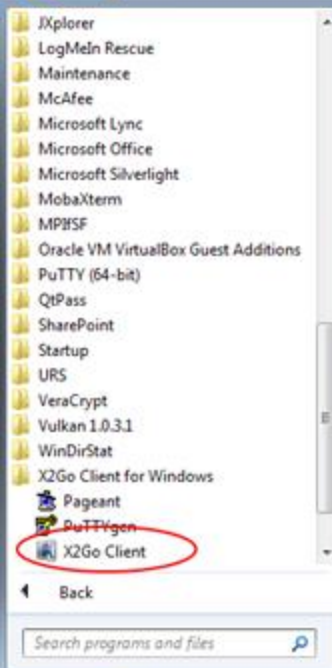
RESEARCH VIRTUAL DESKTOP SERVICE

Research Virtual Desktop Service & X2Go

- Recommended method for connecting to the iCSF can be used to connect to other platforms within the CIR ecosystem
- Accessible from off/on-campus
- Requires a small piece of Software called X2Go (open-source) - MAC users will also have to download Xquartz
- Setup Guide - <http://ri.itservices.manchester.ac.uk/virtual-desktop-service/x2go/>
- PLEASE USE either nyx5, nyx6, nyx7 - these provide the virtual desktop
- Login with central username & password

Benefits of Research Virtual Desktop Service

- Accessible from off/on-campus
- Provides a remote desktop enabling users to connect to the platforms
- Start a session at work power off your work PC at the end of the day, go home and **reconnect** to the same session continue from where you left off.
 - Connecting directly to CIR platforms without going via the RVDS does **NOT** provide the reconnect functionality e.g. Closing Mobaxterm which is directly connected to the iCSF will also log you out of iCSF and kill any running jobs. :(
- X2go by design uses compression to make GUI's run smoothly even over slow connections



RVDS Exercise

- Connect to incline using the RVDS

Research Data Storage (RDS) & CIR Filesystems

Research Data Storage Service

- Also commonly known as isilon (the technology on which it relies)
- Data stored on RDS can be considered secure e.g. files can be recovered if accidentally deleted
 - Automatically backed up once an hour (changes kept for 24 hours)
 - Automatically backed up daily and copy kept for 35 days
- Does not provide Encryption at Rest therefore if handling restrictive/highly restrictive data please speak to us first.
- Can be accessed from your PC both on and off campus
- Groups can share files easily

Can I have some RDS

- Each funded Research group entitled to 8TB free of charge
 - Over 8TB costs money
- Must be requested by academic responsible for the group /project e.g. PI
- All users' home directories reside on RDS
 - Project data should reside in dedicated Project folders not home
 - Projects should request further storage if required
- If requesting RDS
 - Research Data Management – Research Council requirement that you have a Research Data Management plan – Outside scope of this presentation for more info
 - <http://www.library.manchester.ac.uk/services-and-support/staff/research/services/research-data-management/>

Filesystems cont....

- After logging onto the iCSF/CSF you will be automatically placed into your home directory.

```
[username@incline31 ~]$ pwd (tells you where you are)  
  
/mnt/hum01-home01/username
```

- Home space is shared by everyone in the group and space is limited. It is not designed for long term storage of data - use project space
- Only you can access your home directory. However sharing is possible using file permissions. Please contact Research IT regarding this.

Filesystems cont....

- You can see how much space you are using with the command

```
$ du -sh
```

- How big is that file? -

```
$ ls -lh filename
```

- How much space is used/free overall?

```
$ df -h . (don't forget the .)
```

Uploading and Downloading Files

- Various ways this can be achieved. MobaXterm, WinSCP, RDS-SSH, SSHFS & command line utilities e.g. scp, rsync.
- How large are the files you need to transfer?
- What operating system are you using?
- Documentation can be found at the following location

<http://ri.itservices.manchester.ac.uk/userdocs/file-transfer/>



The University of Manchester

WinSCP

Login

New Site

Session

File protocol:
SCP

Host name:
incline.itservices.manchester.ac.uk

Port number:
22

User name:
mzysscg3

Password:

Advanced...

Login Close Help

Documents - mzysscg3@incline.itservices.manchester.ac.uk - WinSCP

Local Mark Files Commands Session Options Remote Help

Synchronize Queue Transfer Settings Default

mzysscg3@incline.itservices.manchester.ac.uk New Session

My documents

Upload Edit Properties New

C:\Users\mzysscg3\Documents\

Name	Size	Type	Changed
..		Parent directory	28/08/2018 15:38:3
Corel		File folder	07/06/2017 10:28:1
MobaXterm		File folder	04/09/2017 12:11:0
My Palettes		File folder	28/09/2017 14:55:2
Outlook Files		File folder	07/07/2017 15:56:4
email addresses.txt	2 KB	Text Document	06/04/2017 12:02:0
RDS Support Matrix.d...	19 KB	Microsoft Word D...	20/09/2017 10:15:0
test	0 KB	File	17/08/2017 09:23:0

/mnt/iusers01/support/its-sct/mzysscg3/

Name	Size	Changed	Rights
..		09/11/2017 12:03:16	rwxr-xr-x
00007_M_acc2_22ph_...		03/12/2018 13:41:31	rwxrwxr-x
20180925_e00868_PET...		04/12/2018 09:13:58	rwxrwxr-x
PET-MR		10/10/2018 15:34:52	rwxrwxr-x
R		19/06/2018 10:17:17	rwxr-xr-x
scratch		14/03/2016 13:20:09	rwxrwxrwx
test_transfer		11/10/2018 09:50:44	rwxr-xr-x
airline.dta	3 KB	27/01/2008 17:34:46	rw-r--r--
airline.tar.gz	2 KB	10/10/2018 15:43:44	rw-r--r--
alces-login.log	1 KB	14/03/2016 15:57:34	rw-r--r--
create_multiple_direc...	1 KB	25/05/2017 11:14:05	rwxr-xr-x
hash.txt	101 KB	11/10/2018 09:43:29	rw-r--r--
id_gorg.pub	1 KB	04/10/2018 14:22:36	rw-r----
md5sum_results.txt	1 KB	11/10/2018 09:51:14	rw-r--r--
out.txt	0 KB	19/11/2018 15:16:21	rw-r--r--
scp-speed-test.sh	2 KB	17/05/2017 10:27:59	rwxr-xr-x
Stata15Linux64.tar.gz	390,865 KB	16/01/2018 13:25:35	rw-r--r--
test20181005.txt	0 KB	05/10/2018 17:06:43	rwxrwxrwx
transfer.tar.gz	4 KB	11/10/2018 09:58:58	rw-r--r--

0 B of 19.7 KB in 0 of 7

5 hidden 0 B of 381 MB in 0 of 18

41 hidden

SCP

0:00:25

Uploading & Downloading Files via the Command Line e.g. Linux or OSX

- Upload a file from your computer to your iCSF home dir

`scp file1.txt username@incline2000.itservices.manchester.ac.uk:`

Secure Copy Protocol

No destination after the : means “use your home directory”

- Download a file from your iCSF home dir to your computer

`scp username@icsf.itservices.manchester.ac.uk:results.out.`

The . translates to “the current directory” on your computer

- Upload a file from your computer to an CSF folder in home

`scp file2.txt username@csf.itservices.manchester.ac.uk:mydata/`

- Download a file from your CSF folder in home to your computer

`scp username@csf.itservices.manchester.ac.uk:mydata/results.out results.copy`

Will rename the copied file to “results.copy”

File Transfer Exercise

- Upload a file to the iCSF using WinSCP
- Manipulate a file using Linux commands

USING THE iCSF

(logging in, setting up and running applications)

- The iCSF is a shared resource you will likely be sharing the compute nodes with other people
- The login node assesses the load on them before deciding which one to place you on.
- 7x Super-High-Memory nodes - 32core 2TB (2000GB) of memory

`ssh -X`

centralusername@incline2000.itservices.manchester.ac.uk

Connecting to the iCSF

- Open X2go application and connect to to Virtual Desktop Service i.e. nyx 5,6,7.
 - From Menu - ***Applications > System Tools > MATE Terminal***
- In Terminal log into the iCSF
 - `ssh -X username@incline2000.itservices.manchester.ac.uk`
 - Perform 2FA when prompted
- You will end up on a backend node e.g. `incline31`
- The node is indicated by the 2 digit number in the above example
 - `31-37` = Super High Memory 2TB nodes

Starting Apps

- All software is controlled and loaded using module files
- List available software

```
module avail
```

Not all software will be available due to license restrictions

- All software is loaded via modulefiles e.g.

```
module load apps/binapps/rstudio/0.98.1103-any-r
```

```
module load apps/binapps/matlab/R2017a
```

- List currently loaded modules i.e.

```
module list
```

- Once module file is loaded open the software (the `&` lets you keep using the terminal)

```
rstudio &
```

```
matlab &
```

- Voila! GUI/window should open
- More information on available software, and module files please follow this link

<http://r-it-services-manchester.co.uk/icsf/software/>

iCSF Exercise

- Load and Run Rstudio
- Logout of the iCSF

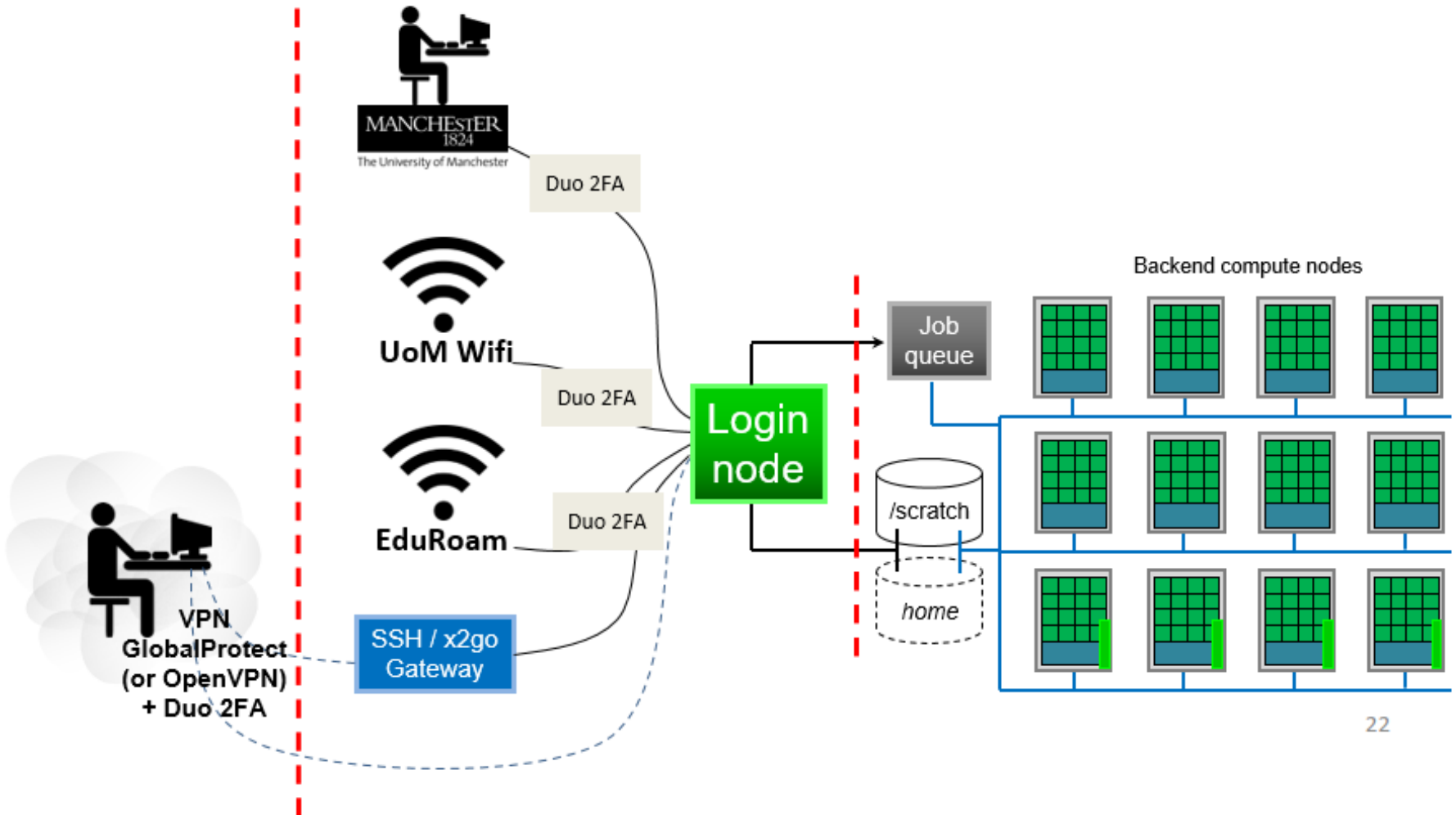
Using the CSF

What we'll be using today - the CSF

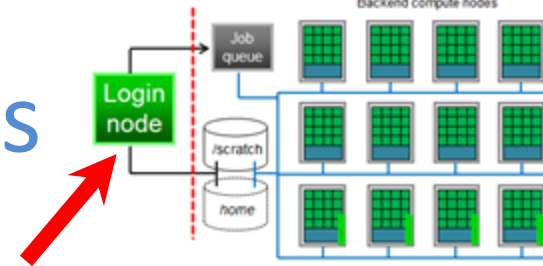
- CSF3 current config:
 - A large Linux cluster system
 - 14,016 CPU cores (Intel “Xeon” or AMD “EPYC” CPUs)
 - 152 Nvidia GPUs (68 x v100, 72 x A100, 12 x L40s)
 - Got big datasets to process? Can run large-memory jobs
 - (we'll cover all of these details throughout the course)

You do not need to be running huge parallel jobs, or be a Linux / HPC expert, to use our systems and to benefit from the CSF

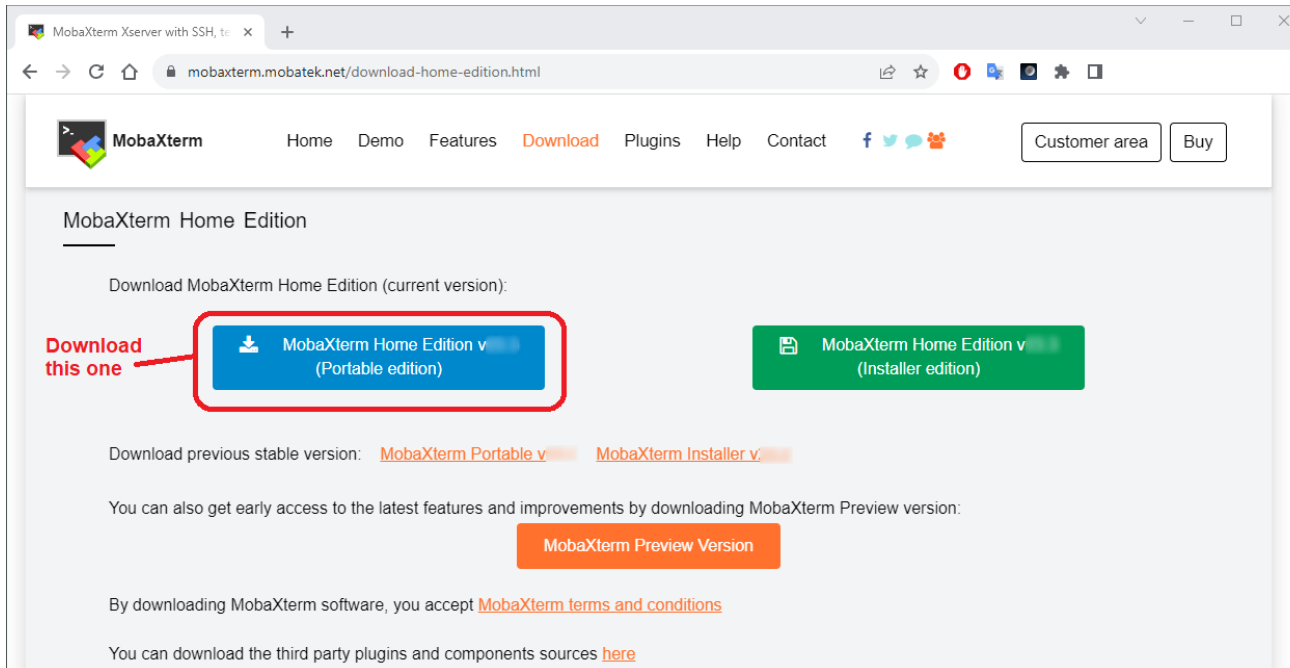
How can I access the CSF



Connect to CSF from Windows



- Access the CSF from a PC / laptop using an SSH (Secure **Shell**) app
 - Sometimes called a "terminal".
 - There's no web-site or other fancy GUI on the CSF – use the "command-line".
- **Windows users** need to install a free *terminal* app called **MobaXterm**
- <https://mobaxterm.mobatek.net/download-home-edition.html>
the **Home edition (portable edition)** does *not require* Administrator rights - just *extract* the small .zip file in your P-Drive or USB stick for example.



1. Download using the **blue** box.
2. Once downloaded, *right-click* on the .zip file and select:

"Extract all ..."

This will *unpack* the .zip file to a folder.

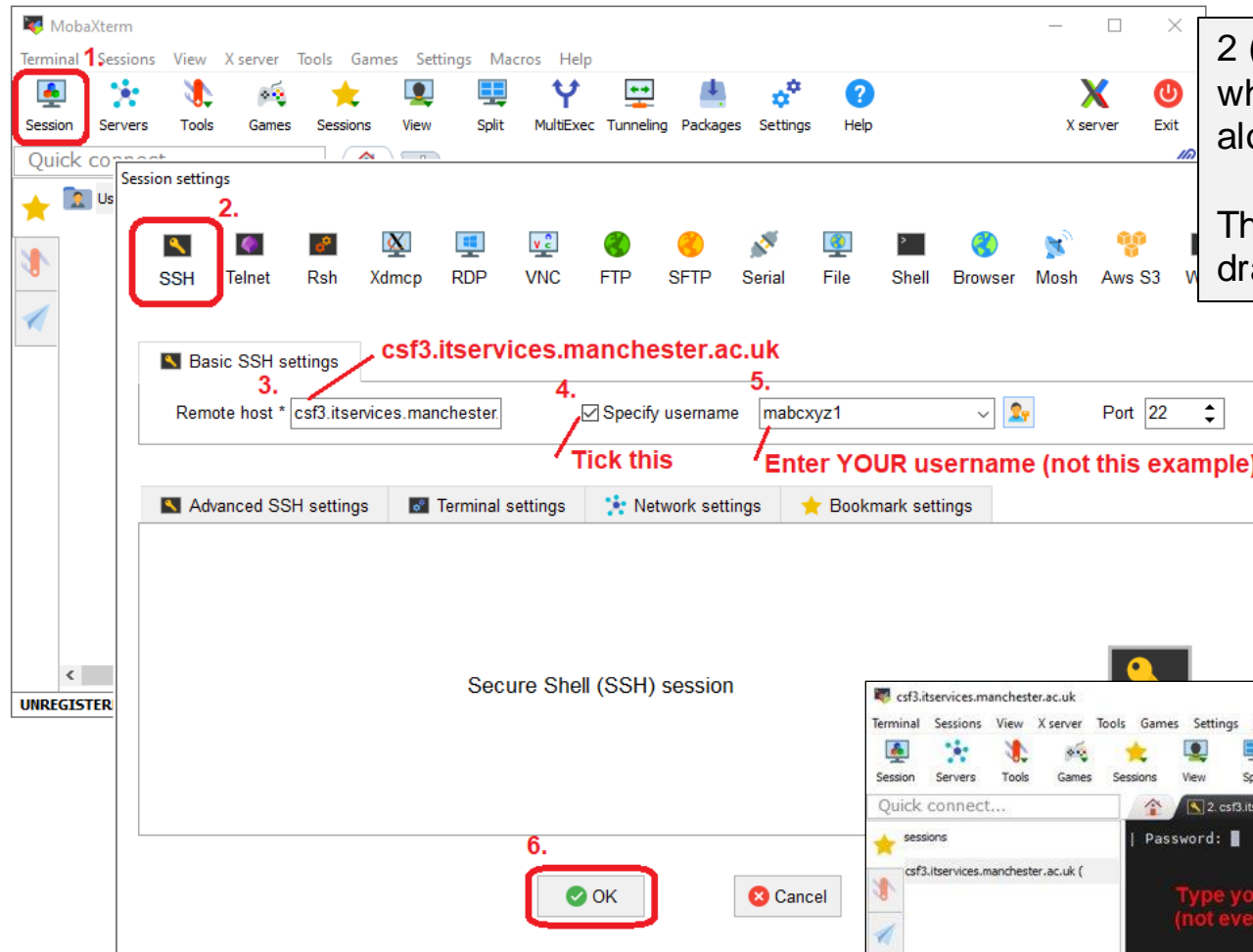
MobaXterm "Session"

(username saved in the session setup)

1. After **extracting** the .zip file, start MobaXterm_Personal_xy.z (double-click on the icon)

- 2 (1-6). Create a "Session" which saves the CSF's details along with **your username**.

This is needed to make file drag-n-drop work (see later.)



3. This will then start to log you into the CSF – it will ask for your password. Type carefully!

Type your password carefully! It won't show any characters (not even *****) but it IS noticing what you type.

4. See slide about 2FA – you may be asked for DUO after your password

Do you want to save password for [redacted]@csf3.itsservices.manchester.ac.uk?



Yes

No

If you want maximum security for your stored password, you can define a "master password" by going to ["Settings" -> "Misc" tab -> "MobaXterm passwords settings"](#)

☒ Do not show this message again

If asked to save your password, we recommend you say "No", for security.

Drag-n-drop file browser for upload / download

(new users won't have as many items in the list!)

We're on (one of) the CSF login nodes. Any commands you use will be typed "at the prompt", which shows your username and current directory (folder.)

csf3.itsservices.manchester.ac.uk ([redacted])

Terminal Sessions View X server Tools Games Settings Macros Help

Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help

Quick connect...

/mnt/users01/support/ [redacted]

Name

- ..
- .alces
- .ansible
- .ansys
- .apptainer
- .aspera
- .cache
- .cfx
- .chainer
- .checkm
- .chimera
- .cmake
- .compucell3d_py3
- .comsol
- .comsol_old_173471
- .conda
- .conda.csf3
- .conda.old
- .config
- .continuum
- .cpan
- .cpan-ignore2
- .cpanm
- .cst-workdir
- .cst2012
- .cupy
- .cytoscape
- .dart
- .dbus
- .dirm2nii

Welcome to CSF3

Docs: <https://ri.itsservices.manchester.ac.uk/csf3/getting-started>

Help: its-ri-team@manchester.ac.uk

*** REMINDER: Scratch Tidy In Operation ***

Reminder that scratch **cannot** be used for long term storage. Files not used (not read or written by you or your jobs) for 3 months or longer will be removed. However, please **note** that if you have recent jobs reading very old datasets, those datasets will NOT be deleted.

!!! You may have files at risk !!!

Use your CSF 'home' dir or RDS for keeping **important** files long term. These areas are backed up. Scratch is NOT backed up.

Jan 2023: New - check your scratch usage (space consumed and number of files) by running the following command on the login node: `scrusage`

19th June 2023: Due to the cyber-incident the University has turned off the web-proxy. Therefore, users will NOT be able to access external websites, repositories etc. via the web-proxy. If external access is required, please use a batch job or interactive session on a compute node (via `qssh`.) For more **info** on doing this please see: <https://ri.itsservices.manchester.ac.uk/csf3/batch/qssh/>

22nd Aug 2023: All access to scratch and RDS (isilon) storage has been restored and batch jobs are running normally. Please check the outputs of any recent jobs to ensure they completed as expected.

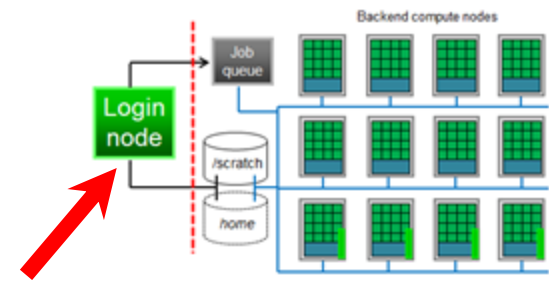
6th Sept 2023: scratch performance issues have been resolved.

Please read all notices above

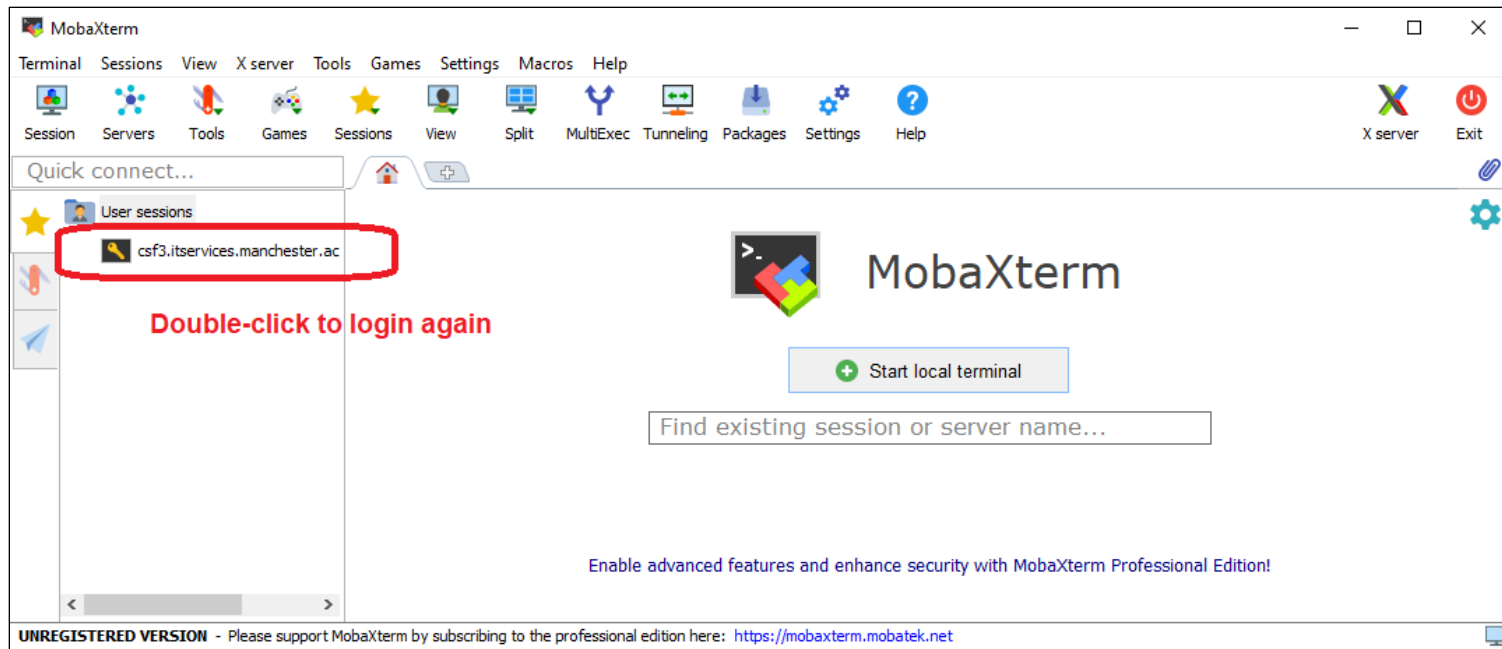
[redacted]@login1 [csf3] ~]\$

UNREGISTERED VERSION - Please support MobaXterm by subscribing to the professional edition here: <https://mobaxterm.mobatek.net>

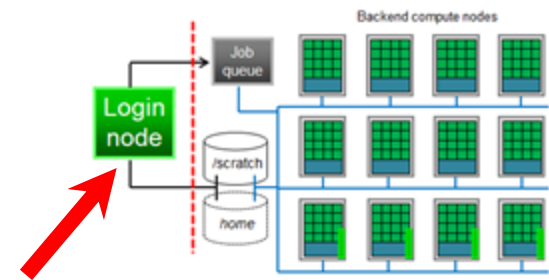
Next time you want to login to CSF from Windows



- Just double-click the csf3 "session" in the list of "User sessions"
- The CSF details are saved in the "session"
- (this also makes the file browser work, for drag-n-drop file transfers.)



Connect to CSF from a Mac



- Access the CSF from a PC / laptop using an SSH (Secure Shell) app, eg a 'terminal'.
 - There's no web-site or other fancy GUI on the CSF - command-line for now.
- **Mac** users - have a *terminal* application by default
 - You will need to install X-Quartz first
<http://xquartz.macosforge.org/> (install then you should reboot your Mac)
 - Start a Terminal and type the following command:

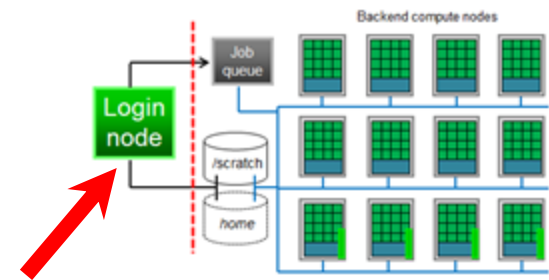
```
ssh -Y username@csf3.itservices.manchester.ac.uk
```

UPPERcase Y

Central IT Services username.
Answer 'Yes' to continue *if* asked.
Enter central IT password when asked (same as for email)

- Finished using CSF? Log out with: **logout** or **exit**

Connect to CSF from Linux



- Access the CSF from a PC / laptop using an SSH (Secure Shell) app, eg a 'terminal'.
 - There's no web-site or other fancy GUI on the CSF - command-line for now.
- **Linux** users - have a *terminal* application by default
 - Start a Terminal and type the following command:

```
ssh -X username@csf3.itservices.manchester.ac.uk
```

UPPERcase X

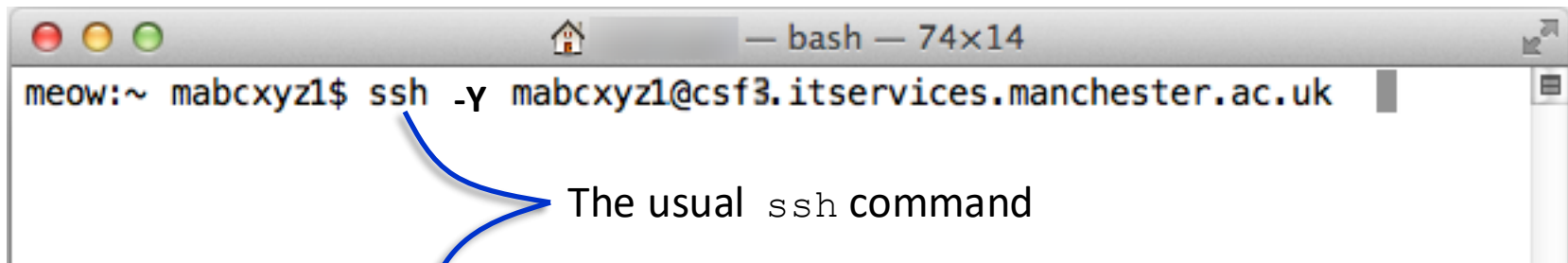
Central IT Services username.

Answer 'Yes' to continue *if* asked.

Enter central IT password when asked (same as for email)

- Finished using CSF? Log out with: **logout** or **exit**

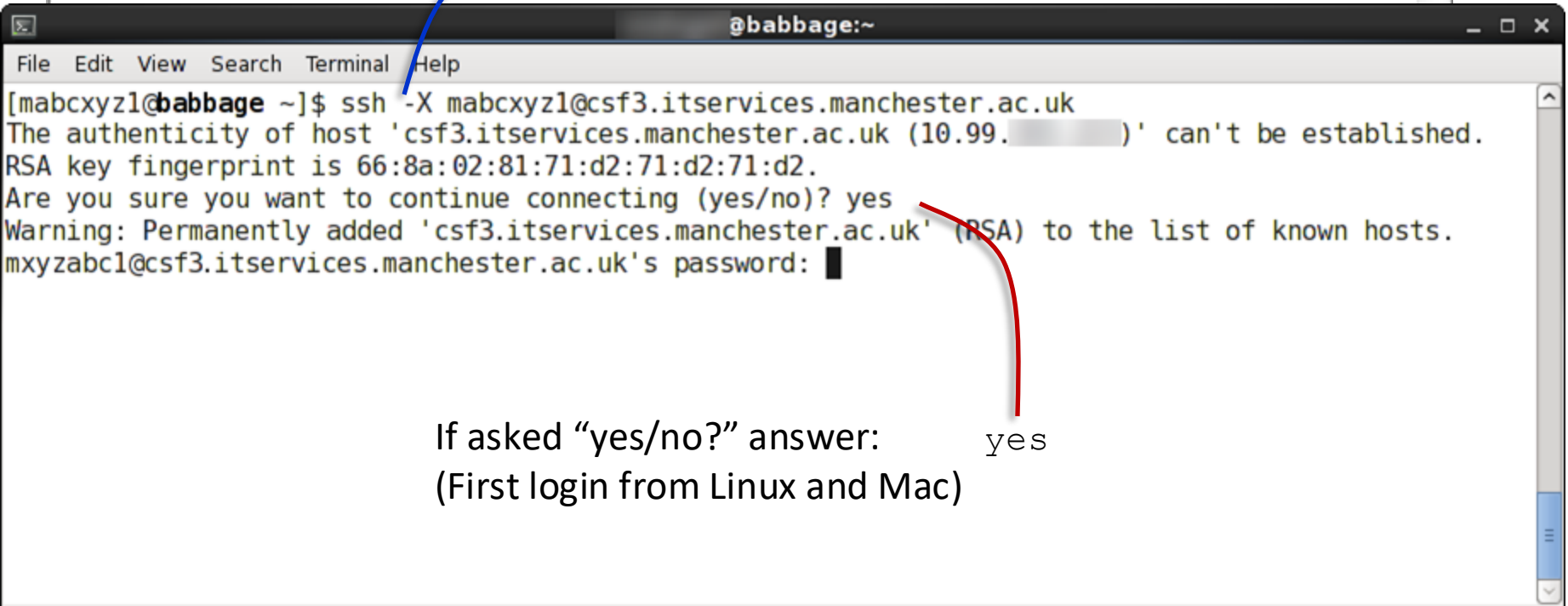
Linux / Mac Terminals



A terminal window with a grey title bar containing window control buttons and the text "— bash — 74x14". The terminal content shows the prompt "meow:~ mabcxyz1\$" followed by the command "ssh -y mabcxyz1@csf3.itservices.manchester.ac.uk". A blue arrow points from the text "The usual ssh command" to the "ssh" command in the terminal.

```
meow:~ mabcxyz1$ ssh -y mabcxyz1@csf3.itservices.manchester.ac.uk
```

The usual `ssh` command



A terminal window with a dark title bar containing the text "@babbage:~". The terminal content shows the command "[mabcxyz1@babbage ~]\$ ssh -X mabcxyz1@csf3.itservices.manchester.ac.uk" and its output. A red arrow points from the text "If asked 'yes/no?' answer: yes" to the "yes" response in the terminal.

```
[mabcxyz1@babbage ~]$ ssh -X mabcxyz1@csf3.itservices.manchester.ac.uk
The authenticity of host 'csf3.itservices.manchester.ac.uk (10.99.10.10)' can't be established.
RSA key fingerprint is 66:8a:02:81:71:d2:71:d2:71:d2.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'csf3.itservices.manchester.ac.uk' (RSA) to the list of known hosts.
mxyzabcl@csf3.itservices.manchester.ac.uk's password:
```


If asked "yes/no?" answer: yes
(First login from Linux and Mac)

What you see when you log in

- CSF uses CentOS Linux (c.f. Red Hat EL)
 - Command line – **requires the input of commands**, can be a little scary at first to new users
 - A welcome *message of the day* - announcements
 - The system awaits input/commands from you at a *prompt* (after you've logged in):

```
[username@login1 [csf3] ~]$  
or [username@login2 [csf3] ~]$
```

- Learning linux commands (more later):
 - <http://www.ee.surrey.ac.uk/Teaching/Unix>



Type your
commands
at the prompt

Warning: The login nodes

- Do *not* run computational work here:
 - Not enough cores
 - Not enough memory
 - 100+ users connected, so running work causes serious problems
- You *can* do the following:
 - Transfer files on and off the CSF
 - Set up and submit your jobs (covered in next few slides)
 - Basic data processing/viewing
- Computational work running on the login node will be killed without warning!

CSF Scratch filesystem

- An extra and very large filesystem available on the CSF
- Shared by all CSF users
- Run jobs here - it is fast storage, sometimes home can fill up causing jobs to fail
- You should have a 'symlink' to it in your home directory (if not ask us to create one)

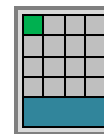
```
cd ~/scratch
```

- Not backed up
 - When job finishes move important files to home or other RDS area
 - Not considered safe for long term storage – hardware failure could cause loss
 - Files that have not been accessed in the past 3 months will be deleted

A simple batch job

- We recommend you have a directory for every job
 - All files relevant to the job need to be in the folder
 - Be tidy, it makes managing results easier
- We need a jobscript – Create it on the CSF using this command
`gedit &`
this avoids a problem with linux/batch not being able to read text files created on windows.
- Like iCSF software is accessed by modulefiles
eg: `module load apps/R/3.4.2`
 - Unlike iCSF, **most** of the time we only use `module load` in a jobscript

A simple Jobscript – *Serial* (1 core)



#! on first line only (a special line)

First line indicates we use the *bash* script language to write our jobscript.

#\$ indicates a **batch system parameter** to specify our job requirements. We'll use various combinations of these.

lines are just comments - anything on the line after it will be ignored.

Actual Linux commands we run in our job. They will execute on a compute node.

myjob.txt

```
#!/bin/bash --login
#$ -cwd
#$ -N myjob
#$ -l resource

# Let's do work
date
hostname
sleep 120
date
```

-cwd indicates we'll run from our current (working) directory. Input / output files will usually be found here.

-N (optional). Set the *jobname*. Otherwise will use name of your jobscript as the name.

-l (optional) used to add extra resource requirements e.g. memory, time limits
#\$ -l course only works on the day of a course.

Submit and monitor the job

- To give the job to the system:

```
qsub jobscript
```

Replacing 'jobscript' with the name of the text file you have created.

- To see the job in the batch system

```
qstat
```

- You should see the job either queued (qw) or (r)
- Your job will have a unique numerical JobID
- If you can't see the job it has finished. :)

Where did my results go?

- Two possibilities
 - A file called **jobname.oJobID**
Where jobname is the name of your submission script
JobID is the number of your job
 - An output file specific to your job
- If your job had a problem or failed it may be indicated in one of the above or **jobname.eJobID**
- Various options to view the file:

`cat filename`

`more filename` (allows you to page through with spacebar)

`gedit filename` (not recommended if it is large)

- Depending on the application and type of files produced you might want to use the iCSF to view them, postprocess, visualise, etc.

Jobs do not always start quickly

- Your job will wait until the resources requested become available.
- You can log off, switch off your PC and your job will stay on the CSF. Log in again to collect the results.
- You can submit more than one job at once, more than one may run at the sametime (good to have different folder for each)

What does Eqw mean? How do I delete a job?

- `qstat` reports your job as `Eqw`
 - CSF tried to start it, but something went wrong
 - Usually no output file so to indicate what happened use
`qstat -j JobID | grep error`
- Error can be cryptic
- Most common causes
 - Missing directory
 - You created the jobscript on windows
 - Unusual characters or spaces in file and directory names
 - No disk space on the filesystem – highly recommend you use scratch to avoid this
- To delete a problem job or one you simply no longer want:
 - `qdel jobID`

Parallel Jobs on the CSF

- CSF excels at doing this
- Each compute node has multiple processors, each of which has multiple cores, each of which can run a single task
- A parallel job uses more than one of these cores
 - Specify how many cores are required
 - Software splits your job into chunks and runs them all at the same time
 - Usually means jobs run **faster**
 - Allows access to **more memory**
 - At the end pulls everything together into your results
- Requires a parallel environment (pe) -

https://ri.itservices.manchester.ac.uk/csf3/batch/parallel-jobs/#Intel_parallel_environments

There's another way to run in parallel!

High Throughput Computing (HTC)

- Lots of *independent* computations
 - Processing lots of data files (e.g., image files)
 - Running the same simulation many times over with different parameters ("parameter sweeps")
- Run many copies of your program
 - Programs may be serial (single core) but running lots of them at once. They **don't** communicate.

Parallel Jobscript – Multi-core (single-node)

`#!` and `#$` see serial jobscript earlier.

myparajob.txt

`smp.pe` is the **parallel environment** name. This one means: app will use a single compute node (up to 32 cores.)

`-pe` indicates we'll run a parallel job in a particular **parallel environment**.

`#` indicates line is a comment so does nothing.

Any commands we run in our job. They will execute on a backend node that has required number of cores free. `mdrun_d` is gromacs.

```
#!/bin/bash --login
#$ -cwd
#$ -pe smp.pe 4

# Set up to use a chemistry app
module load apps/binapps/anaconda2/2021.11

# Inform app how many cores to use
export OMP_NUM_THREADS=$NSLOTS

# This job runs a python script
python3 name_of_python_script.py
```

4 is the number of **cores** we want to *reserve* in the system. Each **PE** has a maximum allowed.

Key concept!

Must *somehow* inform the app how many cores we *reserved*. `$NSLOTS` is automatically set to number (**4**) given on `-pe` line. **Our app** wants `OMP_NUM_THREADS` environment variable setting. **Your app** might use a different method!

What else can the CSF do?

- Access more cores across multiple-nodes. Up to 1024 via the HPC Pool
- Job Arrays – Multiple Similar Jobs -
<https://ri.itservices.manchester.ac.uk/csf3/batch/job-arrays/>
- High Memory jobs, access up to 4TB ideal for processing large DataSets -<https://ri.itservices.manchester.ac.uk/csf3/batch/high-memory-jobs/>
- Over 150 GPUs, ideal for machine learning type workloads
- One day CSF course covers the above

CSF Exercise 1:

- Connect to CSF via RVDS
- Run a serial & parallel python jobs
- View the results on the CSF

Other Research IT Infrastructure Provided Services

Dedicated pool for “true” HPC jobs

- 4096 cores
- Minimum 128-core job size, maximum 1024
- Frontend shared with CSF3
- Lightweight application process – must be made by PI
- Currently free
- <http://ri.itservices.manchester.ac.uk/csf3/hpc-pool>

Research Virtual Machine Service (RVMS)

- Another component of the CIR ecosystem
- Centrally-hosted virtual machines (VMs) are provided for University research staff and PG students
- RVMS is suitable for VMs which are expected to generate large amounts of network traffic, may be moderately CPU-intensive or require tight integration with the computational platforms (CSF, iCSF, etc.) or the Research Data Service (RDS).
- Administrative control of VMs provided maybe handed over to research groups.
- More info <http://ri.itservices.manchester.ac.uk/rvms/>

- Another component of CIR ecosystem
- Utilises 100s of Desktop PCs which are idle on campus
 - Hence typically only available overnight, weekends, holidays.
- Suitable for jobs which require the same computation i.e. repeated many times over different data sets/ranges. Also known as High Throughput Computing.
- Burst into the Cloud using Amazon Web Services
- More Information can be found by following this link

<http://ri.itservices.manchester.ac.uk/htccondor/>

Research IT

- Thank you for attending!
- Course Materials - <http://ri.itservices.manchester.ac.uk/course/icsf-csf/>
- Feedback PLEASE COMPLETE - <https://goo.gl/forms/zfZyTLw4DDaySnCF3>
- Support/Queries <https://ri.itservices.manchester.ac.uk/hpc-help>

Chris Grave, Research IT,
its-ri-team@manchester.ac.uk