CMTH/TYC Linux Cluster Overview

Éamonn Murray
1 September 2017

Cluster Documentation

Everything I'll talk about is covered in the cluster documentation

Web: http://www.cmth.ph.ic.ac.uk/computing

Local (rst – text files with markup): /common/info/doc

These links are also given in the file ~/README.firstlogin

Cluster Overview

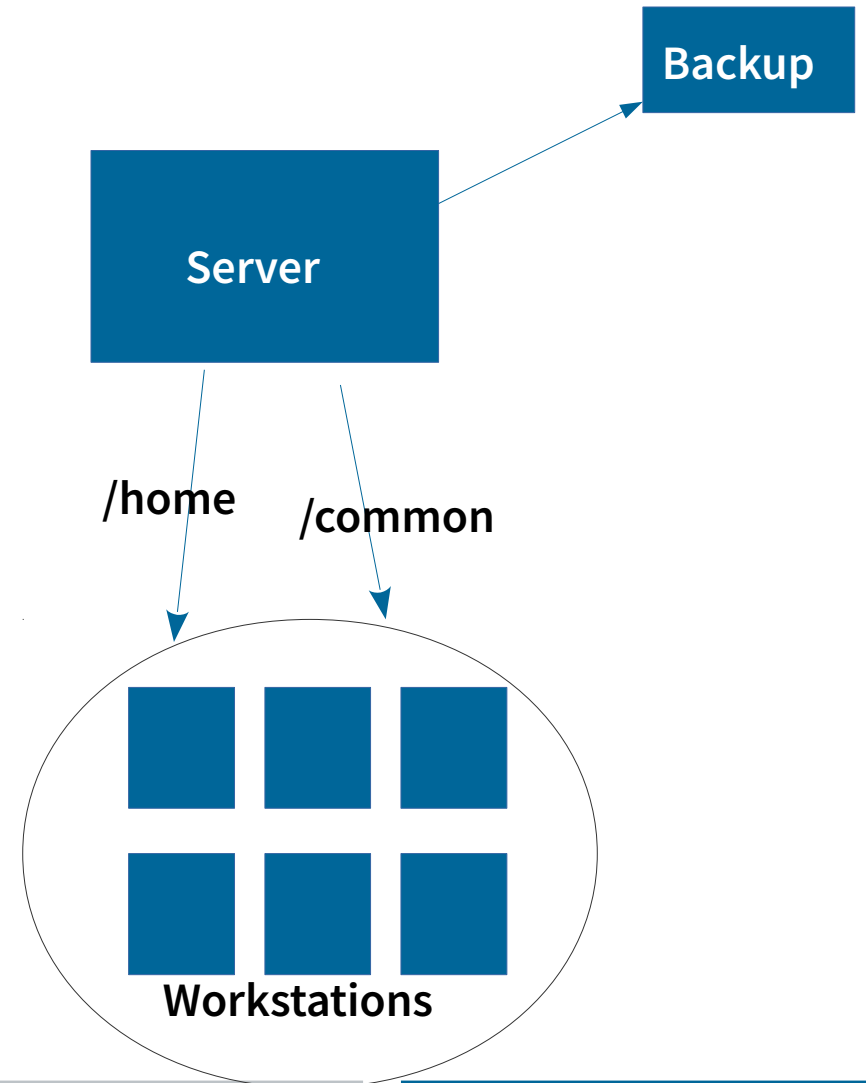Almost **70 user workstations** running Debian Jessie (v8)

Central server (+mirror): user credentials, home directories (**/home**), common compiled software (**/common**)

Webserver

VCS server

Two backup servers

3 compute servers (Hess group)

Backup

Server

/home

/common

Workstations

Home

/home

When you login you start in `/home/username`

quota of 10GB by default

NFS mounted from main server (access is across the network)

Not suitable for running I/O intensive operations/calculations

Backed up to backup server and other workstations

Keep anything important here!

Workspace

`/workspace`

Access as `/workspace/username` or `/home/username/workspace` (symbolic link)

Local harddrive on each machine – much faster access than /home

Run calculations here!

Space varies (check with df -h)

Not backed up – if the disk fails or you delete something I cannot recover it

Backups

Two backup servers

one on-site, one off-site

Daily, weekly and monthly snapshots of **/home**

Contact me to recover files

Squirrel – network backup

Nightly backup of /home to the /backup directory of a workstation

Use "`yesterday`" to find where your backup is stored

You can access this backup yourself (ssh to the machine)

Workspace is not backed up

Script in `/common/sbin/setup_local_rsnapshot.sh` that will help you set up a regular backup of your workspace to an external drive if you need to

Managing space

`quota -s`

will tell you how much of your quota you've used

`du -sch .[!.]* * | sort -h`

will list all files and directories, including hidden directories in order of size

Large system directories can sometimes be moved to your workspace and replaced with a symbolic link (take care)

Need to remember to recreate the directory if you move to another workstation

E.g. To move a Dropbox folder cache you can do

```
dropbox stop
chmod -R u+w ~/Dropbox
mv ~/Dropbox ~/workspace/
ln -s /workspace/$USER/Dropbox ~/
dropbox start
```

Security

www.cmth.ph.ic.ac.uk/computing/setup/security.html

Use strong passwords

Rules for our system are listed above

Our system is separate to the main Imperial system

Avoid recycling passwords

Password managers Keepass2 and KeepassX are both installed on the system

Be wary of any email that sends you a link asking you to reset your password or something similar

SSH Keys

No password access via ssh our systems. Keys must be used

Generate with `ssh-keygen`

The public key (e.g. ~/ssh/id_rsa.pub) can be appended to ~/ssh/authorized_keys on any system you want to access with it

If logged into XFCE, typing "`ssh-add`" in a terminal will save the decrypted key in gnome-keyring for that session

To start a shell with the decrypted key cached (e.g if connected remotely) you can do

```
ssh-agent bash
ssh-add
```

SSH Key rules

SSH key passphrase is **even more important than your password** – actually make it a phrase – several words

Do not use **passphrase-free ssh keys**

Keep private key private – don't email it, or upload it <u>anywhere</u>

Remote Access

http://www.cmth.ph.ic.ac.uk/computing/setup/remote_access.html

Remote login is via ssh using an ssh key (password login is disabled)

Only three machines are accessible from outside the cluster:

hartree.cmth.ph.ic.ac.uk, kirchhoff.cmth.ph.ic.ac.uk and
pauling.cmth.ph.ic.ac.uk

You can configure ssh to redirect your connection through one of these
machines to your workstation

Modify ~/.ssh/config to add an entry for your workstation

```
Host teller
    User emurray
    Hostname teller.cmth.ph.ic.ac.uk
    ProxyCommand ssh -W %h:%p hartree.cmth.ph.ic.ac.uk
```

Remote Desktop

VNC installed on all workstations – instructions at

www.cmth.ph.ic.ac.uk/computing/setup/remote_access.html

x2go installed on all workstations – much easier to use than VNC

Clients available for Linux, Mac, Windows: x2go.org

Forward full desktop or particular application

Using ssh proxy (connecting through hartree or kirchhoff) is straight forward

BASH

Default shell on the cluster (csh/tcsh also available)

Startup scripts:

Login shells - started when logging in – e.g. via ssh

Reads commands in .profile

Interactive shells – started when opening a new shell or terminal window

Reads commands in .bashrc

The default for new users is that .profile reads commands in .bashrc – so you can stick to using .bashrc

Modules

Environment modules simplify the process of changing shell environment variables such as $PATH and $LD_LIBRARY_PATH to make different versions of software or libraries available

For example, to make the Intel compilers available in your current shell enter

```
module load intel
```

This can also be placed in your .bashrc file

To search the full list of available modules use `module spider`

`module spider abinit` will list any available modules containing the word abinit

`module spider abinit/7.10.5` will give details of any modules that need to be loaded before this particular abinit module

More Modules

Installed packages are listed at
http://www.cmth.ph.ic.ac.uk/computing/software.html

A full list can be generated by typing

`module spider`

I'm happy to take requests for any packages you'd like to see available as a module

Sample Modules

Compilers

Intel, gcc, julia

Libraries

Mkl, armadillo, fftw, hdf5, openmpi

Electronic Structure software

Abinit, quantum espresso, castep, siesta, berkeley gw, wannier90

Desktop software

Mathematica, Matlab,

Recent versions of libreoffice, keepassX, unison

Installing Software

Available through aptitude

Usually no issue to install – just let me know

Can check with e.g. `aptitude search gnuplot`

Needs to be compiled, or more recent version needed

Contact me, and I can make available as a module

OR, you can compile & install locally to your machine

```
./configure --prefix=/workspace/username/install_dir
make && make install
```

NB compiling software to home is **not recommended – use your workspace**

Python Packages

We have many python packages installed both by default on workstations and available through the modules system.

As packages could be compiled using either python 2 or python 3, you need to have the appropriate python module loaded to see available modules for that version.

```
ml python ase
```

will give you ase as compiled with the default python module (3.5). While

```
ml python/2.7.13 ase
```

will give you ase as compiled with python 2.

(ml is a shorthand version of the module command)

Job Scheduling

A job scheduler (slurm queue system) system is set up across our cluster - similar to that on HPC resources

Create a job script and submit with sbatch

squeue to see list of jobs

Default queue: 1 process per workstation (no parallel jobs).

Overview:
http://www.cmth.ph.ic.ac.uk/computing/setup/running_calculations.html

Example scripts etc:
http://www.cmth.ph.ic.ac.uk/computing/software/slurm.html

Webserver

CMTH runs its own webserver

main CMTH website is hosted by ICT's webservers

Hosts user docs and some research pages

Contact me if you want a personal page hosted there

URL will be www.cmth.ph.ic.ac.uk/people/user_name

Content auto-updated from `/home/web/user_name` every 10 minutes

VCS Server

We also run a server to host software tracked via a version control system – tycpc15

Allows collaborative software development between groups both internal and external to Imperial

Runs both svn and git (gitolite)

Contact me for setup and access

GitHub – Imperial has enterprise access

GitLab – unlimited private repos

Imperial CM Hub Workshops

CM Hub run many courses for CDT students throughout the year:

http://www.imperial.ac.uk/computational-methods/cm-hub/

Check the website for upcoming events. Many require pre-registration.

Common Tasks

Open Microsoft office documents:

Libreoffice

office.live.com (web app – seems to work better on chromium than firefox)

Annotate pdfs

okular (save the annotated document using "save as")

Install python package locally

```
python setup.py install --user
Or use a virtual enviroment (default module in versions > 3.5)
```

Print to ICT Printservice Printers

Enter username as "ic\imperial_username" and password as your imperial password

Install a font

Put it in ~/.fonts

Example Issues

No sound

Open xfce4-mixer and check levels on both ALSA and PulseAudio mixers

Output regarding module errors when you open a shell

Check your ~/.bashrc and ~/.profile (or ~/.bash_profile if you have one left over) files for lines which load modules that no longer exist

Some characters don't print correctly in a pdf

Try using okular instead of evince

Can't log in

Please let me know

Workstations

Moving workstations

Please let me know. Within Blackett generally no issue. Others typically need network changes updated by ICT (usually < 1 day).

Please don't turn off workstations - this lets me keep them updated etc

Buying workstations

I can make up a quote for you if you like, but I won't be able to put it through

Takes a few hours to install/set up once arrived

Spares

Currently there are a few unused if you need one

Other Resources

Getting started with Linux

https://gitlab.com/eamonn.murray/IntroToLinux

Introduction to High Performance Scientific Computing

Free book by Victor Eijkhout

https://bitbucket.org/VictorEijkhout/hpc-book-and-course