

HPC “Hands On”

David Green
UQ Research Computing Centre

Introductions

- ❑ Who we are
 - ❑ What we are here for .. **LINUX HPC**
 - ❑ Other training is available when there is demand (see rcc.uq.edu.au website)
-

Expectations

□ Core HPC Competencies

- How to get data on and off the HPC.
 - Command line proficiency.
 - How to use software packages on HPC.
 - How to submit and monitor batch jobs.
 - How to archive your data and why you should.
-

HPC at UQ

☐ Compute Resources

- Linux Clusters
- Windows Clusters
- hpcu.uq.edu.au/ganglia

☐ License Services

- Network licenses for research software

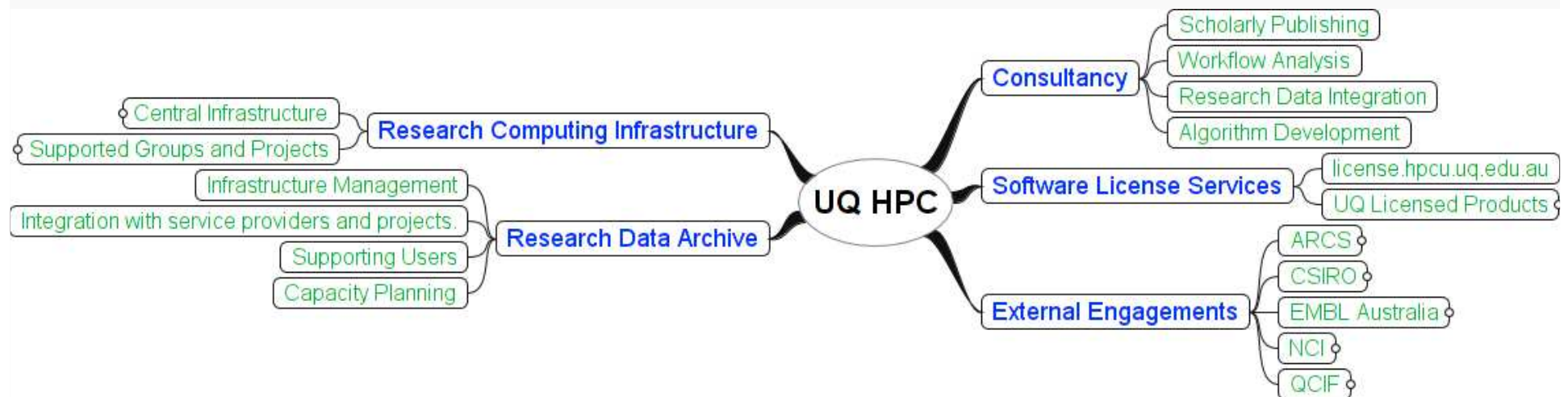
☐ Storage Resources

- Local Disk
- Network Storage
- Hierarchical Storage Management (HSM)

☐ Project Resources

- EBI Mirror
 - NCI SF
-

HPC at UQ



Barrine Cluster

- 384 Compute Nodes (3144 cores)
 - Most are 8 CPUs per node and 24GB RAM per node
 - Some with 8 CPUs and extra RAM/disk/GHz
 - 3 nodes have 32 CPUs and 1024 GB RAM per node
 - Attached Storage
 - 250 GB local disk per node
 - 90 TB shared disk (parallel file server by Panasas)
 - 2.5 PB scale (hierarchical storage management by SGI)
 - Operating Environment
 - Linux OS (SuSE SLES VII)
 - Batch system (Altair PBS Pro VII)
 - Software development (Intel & TotalView tools)
-

All nodes should see the same disk

- ❑ Local Disk
 - /scratch (but better to use \$TMPDIR)
 - ❑ Parallel Filesystem (“Panasas”)
 - /sw
 - /home
 - /work1, /work2
 - /other project specific
 - ❑ Hierarchical Storage (“DMF”)
 - /HPC/home
 - /PROJ
-



This is just login + medium nodes and Infiniband ... and SGI's Gerald!

Who and What?

Partner	%	Cores
NCI SF	32.3	1000
EBI	24.8	768
UQ	21.4	664
QCIF	21.4	664
Totals	100	3096



NCI SF Shares	%	Cores
UQ	32.787	328
QCIF	23.419	234
NCI	18.735	187
CSIRO	18.735	187
QFAB	6.324	63

Login Nodes are for ...

- ❑ Login nodes are all reached as `barrine.hpcu.uq.edu.au`
 - ❑ Organising, submitting & monitoring batch jobs
 - ❑ Accessing storage
 - ❑ Limited pre/post processing
 - ❑ Please don't treat a headnode like a desktop PC.
 - Primum non nocere – first, do no harm
 - Think about what you are trying to achieve
 - Test automation on a limited scale
 - Read documentation
 - Ask (If in **any** doubt, please ask us!)
-

Lab I

□ To Learn

- How to get data on and off the HPC
 - Basic command line proficiency
-

HPC Connection Tools

- ❑ Barrine Users Toolkit
 - PuttySSH
(command line client)
 - WinSCP
(file transfer program)
 - Xming
(X11 server for local display of HPC graphics)
 - All are free software for Windows desktops.
 - ❑ There are Mac and Linux utilities available.
-

Lab 1: The Basics (15-20 Minutes)

- Copy the ZIP file from
/sw/doc ... on
barrine.hpcu.uq.edu.au
 - Unzip the file onto your PC desktop
 - Is there a text file?
 - Is it readable under Windows?
 - Push a file onto barrine
 - Edit a text file
 - Push your text file to your home directory.
 - Create a profile in putty for connecting yourself to barrine
 - Login to barrine
 - Try some of the basic commands in the notes
 - Use the man pages to find out some more about how to use them.
 - Login to barrine with X11 enabled
 - Try the xlogo demo for yourself!
 - If it didn't work, why didn't it ?
-

Scenario	Command(s)
Help!	man
Where am I?	pwd
Take me home! Change to some where different	cd cd someWhere
Show me a list of files in /tmp with different amounts of information and sort order	ls /tmp ls -salS /tmp ls -salrt /tmp
Show me the contents of a text file Show me the file one screen at a time	cat /etc/fstab less /etc/hosts
I want to zip a directory called "it" I want to unzip a ZIP file I uploaded to barrine	zip -r it.zip it unzip file.zip
I want to copy a file A to a new copy B I want to move a file A to rename it to B I want to create a symbolic link to something	cp A B mv A B ln -s something
Create a new directory called New	mkdir New
Connect to a remote host with X11 tunnel	ssh -X user@hostname

Software? ... yes there's some

- ❑ Modules is the way to find out “what” and “how” about software
 - ❑ If you are manually typing/setting full paths to installed software then you are not getting the most out of modules.
-

Software

- ❑ Linux comes with a great many useful things!
 - ❑ Non-operating system applications are in /sw
 - All nodes can see /sw
 - ❑ Named packages should always have an environmental module associated with them.
 - Software modules are also on /sw
 - The module command
 - ❑ module avail
 - ❑ module load/display/help
 - ❑ module unload
 - ❑ module purge
-

Batch Mode Computing

□ Batch Computing for Beginners

- Understanding qstat output

- Job types and behaviour

- Interactive batch jobs

 - `qsub -I`

- Exporting DISPLAY variable

 - `-v DISPLAY`

- Chunks

- Resource requests

 - `-l select=1:ncpus=2:mem=5GB:NodeType=medium`

 - `-l select=2:ncpus=8:mpiprocs=4:NodeType=medium`

Lab 2 (15 minutes)

- ❑ How to find and use packages on HPC
 - ❑ Environment Modules
 - ❑ Why interactive use of packages should always be done via the batch system using `qsub -l`
 - You can launch X11 apps on compute nodes via the batch system with `-v DISPLAY` option
 - Users of Matlab software, and X11 based apps, should ask for assistance with how to run those technically/legally/"socially" on the cluster.
-

Lab 2: Packages

- Login to barrine
 - Acquire an interactive session on a medium compute node using PBS
 - Take a whole node
`-l select=1:ncpus=8:NodeType=medium`
 - -A uq-Training
 - Try a command line package that might be relevant to your work
 - R
 - math (Mathematica)
 - matlab –nodisplay
 - Quit the batch job at the end
 - Repeat but this time include –v DISPLAY for X11-ness
 - Make sure you are X11 enabled
 - X11 set in puttySSH ?
 - Xming started ?
 - Xlogo test works ?
 - try running an X11 based application such as
 - display (ImageMagick)
 - matlab
 - mathematica
-

Half Time Show

- ☐ Will resume in 10 minutes

Batch Mode Computing

- ❑ Batch script anatomy
 - Why setting walltime is so important.
 - Why the Account String (-A) is a requirement.
 - ❑ Using the batch system “carefully”
 - ❑ What a “job array” is and why you might care.
-

Anatomy of a PBS Script

```
#PBS -N myJobName
#PBS -l select=1:ncpus=4:mem=16gb
#PBS -l walltime=1:00:00
#PBS -A uq-Training
```

PBS job controls and resource requests

```
module load someSoftware
```

Set up environment by loading module(s)

```
echo 'Working directory: '$PBS_O_WORKDIR
echo 'Temporary / scratch directory: '$TMPDIR
```

PBS environment variables

```
cd $TMPDIR
cp $PBS_O_WORKDIR/myInputFile ./
```

Moving data into TMPDIR on the node

```
echo 'Working on node: '$(hostname)
someSoftware < myInputFile
```

Carry out computations on the node

```
cp -p myOutputFile $PBS_O_WORKDIR
```

Moving results back off node

Lab 3 (20-30 minutes)

- Discover

- How to submit and monitor batch jobs

Lab 3: Batch Computing

- ❑ Determine your account string options
 - ❑ Modify the template PBS script to create job scripts that perform the following
 - A listing of \$TMPDIR contents
 - A listing of /etc/hosts file
 - A job that will fail and generate error messages
 - ❑ Modify the PBS template to create jobs that
 - Run on Large nodes
 - Run using an entire medium node
 - Request specific values for
 - ❑ walltime
 - ❑ mem
 - ❑ nodetype
-

Retaining your Research Data

- ❑ A HSM system gives you access to very large amounts of storage without it all being on disk.
 - ❑ /HPC/home and /PROJ/ filesystems
 - ❑ Data on tape is held at St Lucia and Ipswich.
-

Retaining your Research Data

□ The dm commands

- dmls
- dmput
- dmget
- dmdu

□ Better to work with bigger slabs of data

- Create zip or tgz archive files when you have multiple files.
 - Arrange your workflow so you pre-fetch your data.
 - We don't “do” metadata, but we can help you.
-

Lab 4 (15 minutes)

- Become aware of
 - How to archive your data.
 - How to maximise your productivity with HPC
-

Lab 4: Research Data Archive

- Create a large ZIP or tar file in your /HPC/home directory
 - Use DMF commands to force it off to tape.
 - Verify the files offline state.
 - Retrieve the file from tape.
 - Think about how you can “ramp up” your research by making use of the capacity and performance of HPC
 - We can discuss your ideas at the end of the session.
-

Please Do

- Do visit hpcu.uq.edu.au and www.ncisf.org
- Do read PBS User Guide
- Do use `qsub -I`
- Do use `-A` account string
- Do use wall time settings
- Do use select statements
- Do use `$TMPDIR`
- Do use `/work1` if `$TMPDIR` is not suitable.
- Do lump results and copy to `/HPC/home`
- Do see `/sw/doc/Support`

Please Don't

- Don't forget that barrine cluster is a shared resource
 - Don't run heavy or long tasks on the headnode.
 - Avoid jobs out of `/home`
 - Don't perform operations in `/home` or `/work1` that involve large numbers of filesystem queries.
 - Don't forget that `/home` is not backed up.
-