

A Quick Start CSD3 Guide for UKRI CDT Biomedical AI Students

A brief overview for CSD3

CSD3 stands for Cambridge Service for Data Driven Discovery (CSD3) and is one of the EPSRC Tier2 National HPC Services hosted by the Research Computing Services at the University of Cambridge.

CSD3's history can be found here: <https://www.hpc.cam.ac.uk/history>

CSD3 is a combination of three platform types:

- Intel Scalable processors (Intel Xeon Gold - Skylake)
- Intel Xeon Phi (KNL)
- Nvidia P100 GPUs (Cambridge is currently upgrading these to A100)

Important: At time of writing CDT BMAI have access only to GPU Cluster (Wilkes2)

Pascal partition (GPU CSD3 partition)

- Each compute node:
- 4 × NVIDIA P100 GPUs
- 1x12 cores, Intel Broadwell 2.2 GHz
- 96 GB RAM
- 100 Gb/sec (4X EDR) Infiniband. 10 GB/sec (for MPI and storage)
- 90 compute nodes.
- 8 login nodes (login-gpu.hpc.cam.ac.uk)

CSD3 Cluster Storage

- Lustre cluster filesystem:
- Very scalable, high bandwidth.
- Multiple RAID6 back-end disk volumes.
- Multiple object storage servers.
- Single metadata server.
- Tape-backed HSM on newest filesystems.
- 12 GB/sec overall read or write.
- Prefers big read/writes over small

Signing up for a CSD3 account

- You will need an SSH public key.
- You will need to complete the online registration form:
<https://www.hpc.cam.ac.uk/external-application>
- Approval for your account is manual and usually takes a few hours.
- Upon approval, you will receive an email with your username.

1. Check for SSH public key Public key is usually stored in:

~/.ssh/id_rsa.pub (Linux/UNIX)

\Users\\.ssh\id_rsa.pub (Windows)

Simply check with a ls (or dir command for Windows)

```
nikos@mosfet-vm:~$ ls -la ~/.ssh
total 16
drwx----- 2 nikos nikos 4096 Jul 18 20:19 .
drwxr-xr-x  4 nikos nikos 4096 Jul 18 20:19 ..
-rw-----  1 nikos nikos 3381 Jul 18 20:19 id_rsa
-rw-r--r--  1 nikos nikos  741 Jul 18 20:19 id_rsa.pub
```

2. Obtaining SSH public key

Same command for all platforms.

```
ssh-keygen -t rsa -b 4096
```

You will be prompted to choose the file to save the key (you can keep the default by pressing Enter)

You will be prompted to enter a passphrase. (Highly recommended)

SSH public key is generated. Copy the contents of the file (default: ~/.ssh/id_rsa.pub)

Terminal Output

```
nikos@mosfet-vm:~$ ssh-keygen -t rsa -b 4096
Generating public/private rsa key pair.
Enter file in which to save the key (/home/nikos/.ssh/id_rsa):
Created directory '/home/nikos/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/nikos/.ssh/id_rsa
Your public key has been saved in /home/nikos/.ssh/id_rsa.pub
The key fingerprint is:
SHA256: [REDACTED] nikos@mosfet-vm
The key's randomart image is:
+---[RSA 4096]-----+
|
|
|
|
|
|
|
|
|
|
+---[SHA256]-----+
nikos@mosfet-vm:~$
```

3. Complete the online registration form

Visit <https://www.hpc.cam.ac.uk/external-application>

Fill in the fields as described below:

Field	Value
Account Type	CORE
Name of Organisation	University of Edinburgh
Primary Contact	Ian Simpson
Contact Position	Director, CDT Biomedical AI
Contact Email	Ian.Simpson@ed.ac.uk
Contact Phone	<Leave blank>
Organisation Address	10 Chrichton Street, Edinburgh
User's Name	<Your full name>
User's Position	CDT BMAI student
User's Public Key	<Copy&paste SSH public key>
University of Cambridge ID	<Leave blank>
Project Title	CDT-BMAI

4. Receiving activation email

You will receive an email from CSD3 administrators with your username and useful information. Automatically you will be added to CSD3 mailing list. You will periodically receive news, updates and maintenance warnings.

Accessing CSD3

Logging in

At time of writing, CDT BMAI students have only access to the Wilkes2-GPU (GPU cluster) of CSD3. If there is a need for accessing CPU partitions, please let us know.

Once you have received your username you can access CSD3 by simply:

```
ssh <username>@login-gpu.hpc.cam.ac.uk
```

You will be logged in one of the login nodes. GPU logins are balanced over login-e-[1-8].

The first time you login, you will be asked to check that the host key fingerprints are correct. Please check that the fingerprints reported match those on activation e-mail you received before responding yes (NB not y).

Filesystems

Directory	Explanation	Backed up?	Disk quotas	Files removed?
/home/<username>/	User home directory*	Yes (hourly snapshots)	Yes (40GB)	No
/rds/user/<username>/hpc-work	HPC work user directories*	No	Yes (1TB and 1 million files)	No automatic deletion currently
/local/	Per-node local disk**	No	Limited by partition size (57GB or 131GB)	Deleted after job completion
/ramdisks/	Per-node tmpfs (in-memory)**	No	Limited to 50% node RAM	Deleted after job completion
/usr/local/	Cluster-wide software	Yes	N/A (not for user storage)	N/A (not for user storage)

*These are shared Lustre filesystems - files created on these are immediately visible to all nodes.

**These are private for each compute node - files created on these are destroyed after completion of the current job.

File transfers

Any method of file transfer that operates over SSH (e.g. scp, sftp, rsync) should work to or from CSD3, provided SSH access works in the same direction.

Copying a local file to CSD3 home folder.

```
scp /local/script.sh <username>@login-gpu.hpc.cam.ac.uk:/home/<username>
```

Copying a file from CSD3 home folder to local folder

```
scp <username>@login-gpu.hpc.cam.ac.uk:/home/<username>/script.sh /local
```

Documentation

CSD3 has very thorough instructions and documentation at:

<https://docs.hpc.cam.ac.uk/hpc/index.html>

Instructions available for popular implementations and software, like:

- SLURM workload management
- Tensorflow
- Jupyter Notebooks
- PyTorch
- Using Python and Conda/Anaconda

Check what resources are available to you.

```
[co-dand1@login-e-7 ~]$ mybalance
```

User	Usage	Account	Usage	Account Limit	Available (hours)
-----	-----	-----	-----	-----	-----
	+		+		
co-dand1	0	BMAI-CDT-SL2-GPU	2,152	41,667	39,516

Jobs must be submitted using BMAI-CDT-SL2-GPU account.

Modules

Software on CSD3 is primarily controlled through the modules environment. By loading and switching modules you control the compilers, libraries and software available.

When compiling or running programs on CSD3 you will need to set up the correct modules, to load your compiler and any libraries that are required (e.g. numerical libraries, IO format libraries).

By default, the `rhel/default-{system-name}` module is loaded, depending on the login node you used, which loads a set of other modules.

Basic commands:

```
module avail - Finding out which modules (and hence which compilers, libraries and software)
are available on the system
module list - List what is already loaded
module load - Load a module
module remove - Unload a loaded module
module purge - Unload all loaded modules
```

Running jobs

CSD3 is using SLURM workload management and job scheduling system. Useful information about SLURM can be found:

- SLURM QuickStart User Guide (<https://slurm.schedmd.com/quickstart.html>)
- SLURM Printable Cheatsheet (<https://slurm.schedmd.com/pdfs/summary.pdf>)
- SLURM Documentation links (<https://slurm.schedmd.com/documentation.html>)
- SLURM FAQ (<https://slurm.schedmd.com/faq.html>)
- SLURM Tutorials (<https://slurm.schedmd.com/tutorials.html>)
- SLURM Presentations (<https://slurm.schedmd.com/publications.html>)

VERY basic SLURM usage

- Starting interactive jobs (srun) – Popular for testing and short running time jobs.
- Submitting passive jobs (sbatch) – For long time running jobs.

Interactive jobs

Basic srun arguments:

`--partition` or `-p` Choose cluster partition. As CDT BMAI have only access to GPUs this must be set to pascal.

`-N` Number of nodes

`--ntasks-per-node=X` Number of tasks per node

`--mem <memory in MB>` Allocate RAM

`--x11` enables X11 forwarding (needs initial connection with CSD3 to have X11 forwarding too)

`--time=DAY-HRS:MIN:SEC` Set how long the job will run for.

`--account BMAI-CDT-SL2-GPU` Use the CDT BMAI account.

`command` or `--pty bash` Command/script to execute or initiate a bash shell

We are initiating interactive jobs with srun. For CSD3 the arguments of choosing the pascal partition (`-p pascal`) and GPU resources (`--gres=gpu:X`) are mandatory. If more than one node is chosen, `--gres=gpu:X` must be set to 4.

Examples:

```
srun -p pascal --gres=gpu:1 --pty bash
```

The command above will allocate one node with 1 GPU and will start a bash shell.

```
srun --x11 -p pascal -N 4 --ntasks-per-node=4 --gres=gpu:4 hostname
```

This command will run the `hostname` command on 4 nodes 4 times (in parallel) on every node.

Passive jobs

We submit passive jobs by using SLURM batch scripts.

A simple example of such script

```
#!/bin/bash
#SBATCH -N 1
#SBATCH --ntasks-per-node=1
#SBATCH -p pascal
#SBATCH --gres=gpu:1

echo "Hello from the batch queue on node ${SLURM_NODELIST}"
```

Basic sbatch directives inside a SLURM script:

```
#SBATCH --partition or -p Choose cluster partition. As CDT BMAI have only access to GPUs this must be set to pascal.
#SBATCH -N Number of nodes
#SBATCH --ntasks-per-node=X Number of tasks per node
#SBATCH mem <memory in MB> Allocate RAM
#SBATCH --x11 enable X11 forwarding (needs initial connection to CSD3 with X11 forwarding too)
#SBATCH --time=DAY-HRS:MIN:SEC Set how long the job will run for.
#SBATCH --account BMAI-CDT-SL2-GPU Use the CDT BMAI account.
```

sbatch accepts arguments from the command line like srun does. If an argument is given at the command line and there is a directive defined already inside the SLURM script, the command line argument precedes.

Examples:

```
sbatch test.sh
```

This will submit script test.sh

```
sbatch -N 4 test.sh
```

This will submit script test.sh using 4 nodes even if inside the script exists the directive e.g.

```
#SBATCH -N 1
```

Partitions, nodes and job status.

sinfo – shows cluster/partition/nodes information

sinfo -p pascal – shows pascal partition nodes status

squeue – shows status of jobs on the cluster

squeue -u <username> - shows status of jobs of user username

scontrol - view or modify Slurm configuration and state

scontrol show node <nodename> - view information and status of node

scontrol show job <job_id> - view information and status of job

scancel <job_id> - Cancels a job

Getting support

- The University of Cambridge provides support for users on CSD3 and can be contacted via email at support@hpc.cam.ac.uk. This will open a ticket with the support team.
- Raise a request with us at cdt-it-support@inf.ed.ac.uk