

# CHPC Accounts Policy Document

Note: the CHPC reserves the right to amend and/or extend this document. Notice will be given to all users of any proposed amendments to allow for feedback. The CHPC reserves the right to make final determination on the amendment. Any user who is unable to accept the amendments may voluntarily withdraw as a user.

## 1 Introduction

### 1.1 Scope of the document

User policies (UP) at the CHPC are divided into the following areas:

1. User account policies, including accountability, restrictions and responsibilities;
2. Resource limits;
3. User job scheduling policies, which covers the scheduling and priority of users' job on CHPC's computing systems

This user policy document (UPD) covers the following items:

1. duration of accounts: every account expires on 28 February and has to be renewed for the new annual cycle
2. renewal is conditional on progress (measured in publications, student graduations, dissertations, etc.) and appropriate usage (measured in efficiency of resources used)
3. scheduling priority can be allocated to important Research Programmes as identified by the scientific advisory committee (SAC)
4. the user's responsibilities in terms of security, privacy and proper use of their account
5. misuse, use for work other than in application, sharing one account, etc. and other abuses will result in suspension of account
6. storage (quotas) and cpu limits (allocations)
7. attendance at CHPC training or groups responsible for own training
8. supervisors are the principal investigators (PI) for student accounts
9. users to provide essential information for CHPC user database (UDB): pertinent personal information, contact details, status (staff/student), publication history, workshop attendance, &c.
10. student users are expected to present a poster at CHPC conference\* (\*subject to funding and space)

## 2 User Accounts

### 2.1 Research Programmes and Principal Investigators

Computational and storage resources at the CHPC are only provided on a per Research Programme basis. Each Research Programme is lead and managed by a principal investigator (PI). The PI is primarily responsible for the the use of the allocations provided by the CHPC and is entirely responsible for the users within his/her Research Programme.

1. Full time permanent staff members of an academic research institution in the following countries may apply to be a PI: Republic of South Africa, Namibia, Botswana, Ghana, Kenya, Madagascar, Mauritius, Mozambique and Zambia. An academic researcher from any other country who wishes to use the CHPC may apply to be a user of a Research Programme lead by a qualifying PI and must be a bona fide collaborator of the PI.
2. In the case of a post-graduate student, his/her thesis supervisor must be a PI at the CHPC, even in the case where the student is the sole member of the Research Programme.
3. The PI for a non-academic or commercial Research Programme will be identified in the memorandum of agreement (MoA) or contract between the CHPC (CSIR) and the non-academic or commercial entity, respectively.
4. A qualifying PI must apply to the CHPC user database web site. Once all requested details are provided online the Research Programme proposal will be evaluated and the PI confirmed.
5. Each Research Programme will receive an allocation of computational resources (measured in cpu·hours).
6. The PI will then identify and allocate users to the Research Programme. Users within a Research Programme will draw upon the Research Programme's allocation for all their job-runs. Once an allocation is exhausted the PI may apply and motivate for an additional allocation.
7. Each PI will be required to provide the CHPC with regular reports (quarterly or semi-annually as needed) on the progress of the Research Programme and on all outputs (journal articles, conference proceedings or slides, student dissertations, student graduations, technical reports, etc.) produced as a result of the computational work performed on the CHPC's systems.
8. All users are required to give full acknowledgement to the CHPC in all public presentations of their work and the PI is required to confirm this for all users within their Research Programme.

### 2.2 User Accountability

All users are expected to use their CHPC account, storage, network and computational resources in a responsible, ethical and professional manner.

1. Any user must apply online to the CHPC User Database and provide all requested details.
2. The user must join an existing CHPC Research Programme and apply to the Research Programme's PI. The PI must then add that user as a member to the Research Programme on the CHPC User Database web site. Once that is verified by the CHPC, the user's account will be created.
3. Users may not allow other people to use their account login.
4. Users must apply appropriate security best practices to secure their account, including keeping their password secret.
5. Any files, results or data in a user account will be considered to belong to the Research Programme and the PI. Copyright and/or other applicable intellectual property laws will be followed by the CHPC. In the case of student users, it is required that the PI ensure that all rules and regulations of the student's university with the regards to copyrights, et al., are followed.
6. In particular it is expected that all proprietary and open source code licenses are fully complied with. Users who write their own program code must ensure the appropriate licencing is respected and all licensing conditions followed.
7. Users in an academic Research Programme may not use their CHPC account for any commercial or paying work, nor undertake any work for a third party who is not their PI or member of their Research Programme.
8. Users of a commercial or non-academic Research Programme are required to follow the parameters of the contract or MoA with the CHPC (CSIR).
9. Any abuse of the CHPC account or facilities, or violation of these policies (including the associated scheduler, storage and network policies) will result in termination of the user account.

## 2.3 Obligations of Academic PIs and Users

PIs and members of academic Research Programmes are not charged for their use of CHPC resources as these are funded by the DST and provided to support academic research and teaching in South Africa. However, there is a *quid pro quo* obligation on academic users of this national facility to provide the CHPC with:

1. Full acknowledgement of the CHPC in all public presentations and publications of their work.
2. Regular reports to the CHPC (either two semi-annual reports, or three quarterly briefings and one annual report).
3. A short synopsis of the user's research with suitable illustrations of major results along with permission for the CHPC to use these in our newsletters, website, pamphlets, slides, and any other outlet or publication where we publicise and promote the work done at the CHPC.

4. Student users are required to submit an abstract for a poster on their work to the CHPC's National Meeting and Conference in December of each year.
5. Non-student users are required to submit an abstract for a short talk to the above conference each year.

## 3 Systems Policies

### 3.1 File Systems and Storage

The primary file systems on the CHPC cluster are:

Mount point*	File System	Size*	Quota	Backup	Access
/home	NFS	80 TB	15 GB	Yes	Yes
/mnt/lustre	Lustre	4 PB	none	NO	Yes
/lustre/SCRATCH5	Lustre	1 PB	none	NO	2015 users only <sup>†</sup>
/apps	NFS	20 TB	none	Yes	On request
/lustre/data	Lustre	1 PB	none	NO	On request only <sup>†</sup>

\* the mount point and size are subject to change: use `df` to determine the actual mount point and sizes for the above file systems.

<sup>†</sup> SCRATCH5/ will be decommissioned after all 2015 users have transitioned to `lustre/`. The hardware will then be reused for `data/`.

**home/** contains users' home directories.

**scratch/** is used as work space for running jobs. There is a strict 90 day deletion policy on files.

**apps/** is used for *shared* installations of software tools and applications. Code installed under here may be supported to some extent by the CHPC.

**apps/user/** is provided to Research Programmes to store shared libraries and compiled codes. The CHPC will not provide technical support for these codes. This space is provided to users as an alternative to using `/home` which is intended for source files and input data. The CHPC may contact users that are storing these kinds of files in their home directories and request that they be moved to `apps/user/` or be purged from `home/`

The tree structure under `apps/` consists of:

<code>compilers/</code>	Compilers, other programming languages and development tools: GCC (C/C++/Fortran), GDB, AutoTools, cmake, Intel (C/C++/Fortran), Python (2.7.x & 3.x), Java, Octave, Open64 (C/C++/Fortran), LLVM These are installed in subdirectories using a lower case name and version, e.g., <code>gcc/4.8.2/</code> . Each subdirectory follows the LSB: <code>bin/ doc/ etc/ include/ lib/ man/ share/</code> and so on.
<code>libs/</code>	Libraries: OpenMPI, MPICH3, Atlas (BLAS), LAPACK, ScaLAPACK, FFTW, Boost,

GSL, Intel (MKL/MPI), ACML

These are installed in subdirectories by name and version, under the compiler or architecture top-level identifier, e.g.:

gcc/4.8.2/openmpi/1.6.5/; intel/2015/atlas/3.10.2/;  
cuda/7/amber/0.9.2/; phi/2016/fftw/1.5.2/; and so on.

The naming convention is:

`<comp_arch>/<comp_ver>/<code_name>/<code_version>/`

`<comp_arch>` = compiler name (gcc, intel, llvm, open64) or architecture name (cuda, phi) or a combined compiler-architecture for cross-compilers (gcc-powerpc)

`<comp_ver>` = compiler version (eg, 2016.1.150 for version 2016.1.150 of Intel Parallel Studio XE)

`<code_name>` = Name of the library, application or software (eg, atlas for ATLAS BLAS library)

`<code_version>` = Version of the library, application or software (eg, 3.10.2 for version 3.10.2 of ATLAS). Note: there may be a stable symbolic link to the latest stable version.

chpc/	Application codes supported by CHPC: Installed as name/version/ as above. (the version may be tagged with compiler-compiler_version/ (for example, proot/2.0.3-gcc520/). See below for domains.
scripts/	Modules and other environment setup scripts
tools/	Miscellaneous software tools, (file format convertors, image processing, AV tools, &c.)
user/	Any code installed by a user and not supported by us. This is by Research Programme or PI.

**data/** is used for shared data storage and comprises files or databases. This is for long term storage and is not subject to automatic deletion.

## Access Permissions

Each subdirectory `<Research_Programme_PI>/` under `/apps/user/` will have the UID of the technical expert (TE) and GID of the programme name or PI. The TE must set their umask to 0027. Then users in that Research Programme or group under that PI will have read access to the files but only the TE can alter them, and that directory is private to the Research Programme. If desired, the umask can be 0022 to make it publicly readable.

## Support Levels

0. **Unsupported:** Code installed in `/apps/user/<Research_Programme>`. This code is installed by the technical expert (TE) designated by the Research Programme PI. The TE is solely responsible for the installation, updating, use and documentation of this code. The TE

will be given an account on the CHPC wiki to provide full documentation on the use of this code at the CHPC. The TE is required to follow CHPC standards and practices for installation of software to work with the CHPC scheduler and to assist the CHPC technical staff in monitoring use of the software.

1. **Limited support:** Code installed in central shared directory (/apps/chpc/). Minimal documentation on wiki describing how to run code on CHPC. Code only updated on user request, or when major version released (1.X → 2.0). Only basic user questions answered. CHPC recognises that more than one user requires code, but it is outside of CHPC staff expertise. At the discretion of the CHPC support may be provided by a 3rd party.
2. **Standard support:** Documentation on wiki describing how to run code on CHPC, along with example scripts. Code updated when new stable version released (X.2 → X.4). User questions on running code answered. Technical queries supported. Introduction to code included in induction course. CHPC recognises this is an important code for SIG or community.

### CHPC provided codes: /apps/chpc/

The table below lists the subdirectories below /apps/chpc/ that contain application codes maintained by the CHPC. Each directory covers a broad scientific domain and consists of sub-directories for each code.

<b>/apps/chpc/</b>	<b>Scientific Domain</b>
astro/	Astrophysics & Cosmology
bio/	BioInformatics
chem/	Chemistry
compmech/	Mechanics
cs/	Computer Science
earth/	Earth
image/	Image Processing
material	Material Science
phys/	Physics
space/	Space

Other domains may be added in the future, but note that many are already included under the above, for example:

compmech/	CFD, structural mechanics, electromagnetics, DEM, FEM, FVM, BEM
cs/	Graph algorithms, computational mathematics, numerical analysis, CSA, GA
image/	Graphics (raster) algorithms, CGI, tomography, computer vision
phys/	Space physics, particle physics, QED, QCD

Commercial codes appear under the above, but they have their own GID to restrict access to users with permission to use that license. For example, Ansys (Fluent) has GID of ansys and appears under compmech/.

## 3.2 Resource Limits

By default, users on the CHPC systems are limited to:

1. Maximum 240 cores total simultaneous running and queued jobs.
2. Maximum of 20 simultaneous running and queued jobs.
3. Standard priority for jobs as determined by job size and wall clock time.
4. Maximum of 15GB storage in home/.
5. Unlimited storage size in scratch/ but maximum file lifetime of 90 days.
6. No access to data/ or apps/user/.

### Application for Additional Resources

An application for the use of additional resources is required for any modification, departure or exemption from the standard CHPC user policies and limits. This application can be made to request:

1. access to the large queue (i.e., use more than 240 cores per job)
2. higher priority
3. larger storage quota
4. extended file lifetime in scratch/
5. access to data/
6. access to install codes in apps/user/

## 3.3 Standards

User are required to follow the conventions of the Linux Standards Base (LSB) and other open

source software.

In your home directory you should install software in `local/` or `opt/` subdirectories.

`$HOME/local/`

is used to store software according to the LSB for software that shares the standard common directory tree of:

`bin/`   `doc/`   `etc/`   `include/`   `lib/`   `man/`   `share/`

and so on.

`$HOME/opt/`

is used to store applications that require their own directory tree. Subdirectories under `opt/` will be the name of the software and version, for example

`openfoam-2.3.1/`

Other directories to use include:

`src/` for source code (unpack source code tarballs here)

`build/` for compiling code (where supported)

## 4 Scheduler Policy

The new scheduler policy is to be applied on the new cluster. The implementation is subject to change: we will monitor and evaluate the queuing system and adjust parameters where necessary.

1. There are **allocations**: after a Research Programme has been created by a PI and has been associated with a CHPC research scientist it is granted an initial allocation of 100 000 cpu-hours. This allocation can be further extended within a financial year upon agreement with the CHPC researcher associated with the Research Programme. Allocations are reduced for each job using the formula:

$$24 \times \text{number\_of\_nodes} \times \text{wallclock\_time}$$

Note: a user job that uses less than the full number of cores on a node is still charged for the full amount.

2. Fairshare is applied to each user individually:
  - the Fairshare measurement unit is **cpu-hours**
  - this may be changed to **node-hours** or **hours** to encourage more efficient use
3. *Single* node jobs to have lower priority (60) and have their own queue (smp).
4. *Standard* jobs to have standard priority (80).
5. Users must apply a group related tag (`-l software=[software title from a pre-determined list]`) to their job submission scripts. The CHPC wiki will list these tags.
6. Priority is increased (100) for important Research Programmes and *large* jobs — To be classified as such and receive higher priority a PI should contact the CHPC researcher associated with his/her Research Programme.
7. Reservations are exceptional and rare: reservations are restricted to crucial and time-critical

Research Programmes. For example, a reservation may be allocated to a workshop where the participants will need access to the cluster only during a limited and specific calendar period.

8. Users are required to ensure that their jobs exit with a correct job-exit-status. An automated notification will be sent to CHPC staff for users who, in any given month, have more than 5% of their wall time hours listed as unsuccessful jobs.
9. Monthly limits are monitored:
  - a. Total hours of all jobs run by a user
  - b. Total number of jobs submitted by user and by Research Programme
  - c. Median and maximum job size
10. Specific monthly limits are enforced:
  - a. Total hours of all jobs run by user:  $\text{total\_hours} < 15\% \text{ of } 24 \text{ hours} \times 30 \text{ days}$
  - b. Total number of jobs submitted:  $\text{total\_jobs} < 20 \text{ jobs} \times 30 \text{ days}$

## Queues

Queue name	[Min-]Max. size of job	Max. duration	Max. Queued / Running Jobs	Comment
test	1 node: 24 cores	3 hours*	1 / 1	For testing, development (compile, profile or debug) and other 1 node jobs
serial	1~24 cores	48 hours	20 / 10	For very small core count jobs; shared nodes.
smp	1 node: 24 cores	96 hours	20 / 10	For single node SMP jobs
normal	2~10 nodes: 240 cores	48 hours	20 / 10	Standard and default queue
large	11~100 nodes: 2400 cores	96 hours	10 / 5	By application only
internal	–	240 hours	–	Restricted to CHPC staff
bigmem	5 fat nodes: 280 cores	48 hours	4 / 1	For big memory nodes
vistest	1 VNC node	3 hours*	1 / 1	For testing visualisation
vis	1 VNC node	12 hours*	1 / 1	Visualisation node

\*The test and vis queues are used for interactive jobs.

## Regulations

The following principles are followed:

1. All Research Programmes have allocations and all jobs draw from this. A user who has exhausted their Research Programme's allocation can not submit jobs.
2. Most jobs are only limited by allocations and scheduled by PBS Pro using the backfill scheduler according to resources requested and available. They have standard priority, except *single-node (any job using 24 or fewer cores)* which have reduced priority: users are

encouraged to use *job arrays* for multiple small jobs.

3. There is a global limit on *simultaneous queued and running jobs* for job size (cores or nodes) and total number of jobs per user. Both maximum total cores (240) and maximum jobs (20) be enforced for standard jobs. Jobs in the large queue have a limit of 2400 cores.
4. A monthly usage soft limit is implemented for all jobs (24 000 cpu-hours per month) and Principal Investigators are alerted when they or their users exceed this.
5. Only *special* Research Programmes and commercial jobs have higher priority. Both classes require approval of an exemption from CHPC standard policies.
6. Interactive jobs are limited to 1 node, and have the highest priority (120). They are limited to 3 hours wall time and will be monitored for abuse. A 300 s idle timeout will also be imposed. Their intended purpose is for pre-processing, post-processing, debugging and compiling. For longer compilation jobs a request can be submitted to helpdesk to extend wall time to 8 hours or more.
7. File transfers require a node with external network connectivity and can not be scheduled as an interactive job. This will need a dedicated login node for long and large file transfers: scp/sftp, rsync and wget/curl.

## Monitoring

In order to accurately monitor usage we will require that users must be allocated to Research Programmes in order to run jobs on the CHPC clusters.

The following thresholds are monitored and notices sent to CHPC and PI if they are exceeded:

1. soft quota
2. software=other
3. non-zero exit status
4. low efficiency: load < 0.6

## Special Circumstances:

At the Director's discretion, in case of emergency or highly unusual circumstances, where deemed necessary to fulfil the mandate of the CHPC, exceptions may be made to the above conditions.