

The Computational Shared Facility Service Level Definition

1. Background

The Computational Shared Facility (CSF) is a high-throughput and high-performance computing (HTC & HPC) cluster procured as a response to a key recommendation of the Manchester Informatics White Paper on Computational Science. The paper recommended that the University adopt a more strategic approach to the procurement and support of HTC & HPC computing systems. The initial procurement, in December 2010, comprised an initial University seed investment of 90k and four academic contributions totalling approximately the same value. Since then the Facility has grown considerably through further academic contributions. The total investment to date (Nov 2012) is almost one million pounds. It is intended that the CSF continues to grow in this sustainable manner through contributions from academics as funds become available.

2. Description of Service

The CSF is hosted and maintained by IT Services. Application support is provided by IT Services (central and faculty IS teams).

Unless stated otherwise, decisions regarding the service will be made jointly by the *service provider* and the *service user group* (see *Definition of Terms* for definition of all italicized terms).

2.1. The Contribution Model

The *service provider*, in consultation with the *service user group*, will:

1. Define a small set of *contribution unit* hardware configurations, including minimum possible contributions.
2. Fix the set of configuration units for a period of time to reduce unnecessary heterogeneity of the Facility. The configuration units will be changed in response to advances in technology after this period. Any revision of the contribution configurations will consider factors such as performance, energy use and price.
3. Define a mutually acceptable definition of a share of available computational resources which will reflect the size of a contribution.
4. Define and maintain mutually acceptable system operating policies (including those for job scheduling and the scratch filesystem).
5. Provide system accounting data in order that the *service user* can monitor the utilisation of the Facility by their project.
6. Provide a web page indicating service news/status and utilisation.

2.2. Responsibilities of Service

Responsibilities for costs and other items associated with the Facility are attributed in the following manner.

The Computational Shared Facility Service Level Definition

The *service provider* will:

1. Provide the initial infrastructure investment and a limited compute contribution to enable service flexibility.
2. Fund staffing costs associated with system administration and application support of the service.
3. Fund licensing costs for system software (not application software).
4. Provide back-up of home-directories.
5. Pay hosting and energy costs incurred by operation of the contribution units.
6. Support and maintain the contributed hardware for three years from initial contribution (i.e., while on vendor support).
7. Support and maintain hardware which is out of vendor support (three years), on a best efforts basis only, for a further two years.
8. Operate contributed hardware as part of the Facility for five years where possible, after which hardware is considered end-of-life and will be decommissioned.
9. Provide limited access (i.e. soft-landing) at this five-year transition point, *via* access to unused resources and the limited university contribution, only, for a further two years.
10. Provide free evaluation of the system for potential contributors who have funds available, for approximately one month.
11. Provide pump-priming access to potential *service users via* unused resources and the limited university compute contribution, only. (Such access will require a short case from the potential *service user*.)

The *service user* will:

1. Fund the compute contribution units.
2. Make an additional sustainability payment when contributing to help maintain Facility infrastructure hardware, i.e., to help sustain the system and its contribution model.
3. Fund additional home or other required storage not covered by the sustainability payment, though not the cost of back-ups. (The scratch filesystem is considered to be part of the system infrastructure).
4. Fund any costs associated with licensed software required for their research (system software excluded).
5. Fund any share of required shared hardware (e.g., high-performance network or GPU-hosting chassis).

3. Hardware

3.1. Infrastructure

The infrastructure hardware comprises the following components:

- Login nodes;
- head (management) nodes;
- fileserver nodes and disk-arrays for both home-directories and scratch space;
- racks & PDUs;
- network switches and cabling (1 GbE and 10 GbE).

3.2. Contributed

The contributed hardware comprises the following components:

- HTC compute nodes, including some high-memory nodes.
- HPC compute nodes with fast interconnect (Infiniband switches, cards and cabling).
- GPU-hosting nodes.

4. Software

A full description of currently supported applications and the supported environment (operating system, batch system, compilers, tools, libraries) can be found in the CSF User Documentation¹.

The service provider will:

1. Install and support software with campus-wide or floating network licenses, or software that is required by multiple contributing groups.
2. The service provider will install and support other application software on a best efforts basis only.

5. Resource Management

Computational resources on the CSF are managed by the batch scheduling system, SGE. Jobs are not run on a first-in first-out basis; they are prioritised by factors including job size and waiting time, and in particular, the share associated with the project to which the job owner belongs.

The full detail of the scheduling policy, for example available job environments suitable for different types of job (e.g., long/short, serial/parallel) is beyond the scope of this document.

5.1. Shares and Share-Based Scheduling

1. An initial number of CSF-Shares will be allocated to a contributor based on each new contribution to the CSF. Each contribution will be used to purchase up-to-date equipment as described above.

¹ <http://wiki.rcs.manchester.ac.uk/community/csf/>

2. CSF-Shares will not be allocated for contributions which will not be of value or use to other contributors, such as additional user filestore.
3. Early contributors to the CSF will be allocated one CSF-Share for each 1000 pounds (excl. VAT) worth of computational resource contributed. (For example, early contributors will be allocated 12 shares for each C6100 chassis contributed.)
4. All subsequent share allocations will be made relative to the above.
5. The number of shares associated with a project is used as the primary parameter in the SGE “fairshare” algorithm to determine the priority when scheduling the compute jobs of a contributor in the system. This will be the normal method of access to CSF compute resources.
6. It is acknowledged that as vendors introduce new CPU architectures, and make other changes, newer hardware is likely to provide greater compute power/resource than previous older contributions, for a given cost. Hence later contributions will be allocated more shares for each 1000 pounds contributed.
7. The scaling factor used will be determined by using those applications most commonly run on the system to compare like-for-like nodes (e.g., new HTC node to old HTC node, or new HPC node to old HPC node).

5.2. Key Features assumed by the contribution mechanism

1. The contribution/share mechanism assumes that each contributor will principally utilise the specific category of compute-equipment corresponding to their initial contribution. For example, a contributor of GPU nodes would principally utilise these nodes rather than HTC nodes.
2. The SGE fairshare scheduler was agreed as the principal mechanism to provide contributors with access to their compute contribution of the CSF. This was deemed to be necessary to provide a workable service on what will inevitably be a developing and heterogeneous system.

5.3. Equipment Associated with a Contributor

1. For novel contributions, such as very high memory nodes or GPU-hosting nodes, the contributor will always get priority over other users.
2. For all other contributions, the CSF fairshare system moves away from the concept of direct association of a contributor with specific hardware.
3. The nature of the scheduling described above means that it is not always possible for a group to have immediate access to computational resource equivalent to the group’s share.
4. However within the constraints of the SLD and the CSF scheduling system, it was agreed that ITS would endeavour to meet specific research demands of a contributor to access equipment according to their contribution, at short notice, when necessary and where possible.

5.4. Advanced Reservation

A *service user* may on occasion require short term access to a share of the Facility which is much greater than that determined by their contribution. Given a minimum of four weeks notice, ITS will endeavour to schedule such requirements.

6. Sustainability of Service

In order to maintain an academic case for the facility the *service provider* will:

1. Provide suitable systems for accounting for resources on a project level.
2. Request, in all publications derived from the use of the service, use acknowledgement *via* a citation or suitable alternative.
3. Request copy and images for the production of case studies to promote the service to other potential service users and stakeholders.
4. Produce any publicity material including each of the above.

The *service user* will:

1. Provide brief yearly reporting on research carried out, publication and funding history associated with the service.
2. Provide copy and images for case studies.

7. Service News/Status, Developments and Utilisation

7.1. News/Status and Developments

Service news/status and developments, will be will be communicated to users *via*:

1. the Facility “MOTD”;
2. dedicated pages in the CSF User Documentation for both news/status² and upcoming developments³;
3. the CSF Users’ Group;
4. and in appropriate circumstances, use of the CSF Users’ email list.

7.2. Utilisation Dashboard

A web-based “dashboard” will be developed which will show:

1. CPU utilisation each month by each *service user*;
2. a CPU utilisatition history for each *service user*;

² <http://wiki.rcs.manchester.ac.uk/community/csf/news>

³ <http://wiki.rcs.manchester.ac.uk/community/csf/developments>

3. a summary of existing and upcoming contributions;
4. statistics relating to the each users' files on both home storage and scratch space.

8. Support

[[This section is subject to change owing to the imminent changes in research support structures within IT Services.]]

IT Services staff will answer users' queries during core hours only. User queries will be tracked using the Action Request System. An initial response will be given before the end of core hours of the next working day and should include:

- either a solution to the query,
- a request for more information,
- or a brief description of work being done to address the query.

All major changes to the CSF will be managed via the standard RFC procedures.

9. Maintenance Schedule

Sometimes it will be necessary to reduce or withdraw service for system maintenance. For scheduled maintenance two weeks' notice will be given.

For exceptional maintenance and unplanned disruptions, where the system may become unavailable at short notice, as much notice will be given as possible, for example *via* the CSF Users' email list. Such exceptions may occur for environmental reasons, such as air conditioning or power failure, or in an emergency where immediate shutdown is required to save equipment or data. The *service provider* will endeavour to keep these to a minimum, but should they occur service will be restored as quickly as possible with appropriate status reporting to *service users*.

10. Additional Policies

Over and above the service description described in this document, the standard University regulations⁴ (including IT) apply.

⁴ <http://www.manchester.ac.uk/medialibrary/governance/generalregulations.pdf>

11. Definition of Terms

Service Provider

IT Services provides the service to contributing members of the University of Manchester.

Service User

For the purposes of this document, a service user is the PI of a project at the University of Manchester who has contributed research funds to the Facility.

Service User Group

The Computational Shared Facility User Group represents this body of stakeholders and will be consulted by the *service provider* regarding any proposed change to the service as defined by this document.

Contribution Unit

Hardware bought using research funding and contributed to the Facility.

System Software

Software stack required by the *service provider* to effectively operate a multi-user shared system.