

# University of Sussex

## IT Services

# Proposed High Performance Computing Service

---

We are proposing to develop a new service to support researchers using high performance computing systems. The main reason for the development of such a service is to allow the research teams to focus on using the HPC facility to advance their research programmes by providing a professionally managed support service.

It is expected that at least a part of the funding for the service will be generated from the research activities the facility supports.

This *discussion* document provides an outline of the service that is being proposed and describes some of the aspects of how it might be funded. Comments are most welcome.

### The proposed service

1. IT Services would establish a high performance computing environment into which researchers could add clusters of high performance compute-intensive nodes. A centrally funded front-end system would provide local file storage that could be accessed from any node in any of the attached clusters and would control and oversee the operations of each of these.
2. Note: The front-end system would incorporate clustering facilities to offer a high level of resilience against environmental and front-end hardware failures. The computing clusters would not normally have this level of resilience.
3. IT Services would provide systems support (i.e. install operating system and middleware patches and upgrades) and administration services for the front-end system and the attached clusters.
4. In partnership with the research groups who would use the system IT Services will select a default system and middleware stack that will be supported within the system. Where a cluster requires that a different set of operating system and middleware software be used, IT Services will work with the research team to enable a third party (perhaps the system supplier) to apply software updates – any costs charged by the third party will have to be met in full by the research team.
5. IT Services would also provide appropriate identity services and, in particular, deal with issues associated with verification of identity for access to major national and international resources. (Please note there are some HPC communities where this has to be carried out within the community – here IT Services would not be able to assist.)
6. The HPC support team would arrange appropriate training events and seminars to help research teams make appropriate and optimal use of HPC systems. Where appropriate the team would also liaise with other HPC teams in the sector and particularly with the National Grid Service.

7. The IT Services HPC team would not be able to provide programming services to the research groups using the facilities, though the team would be able to provide general advice on making good use of the available HPC facilities.
8. It would be an objective that this team also worked with other IT Services colleagues to provide desk-top support for LINUX systems. IT Services would look for some commonality between these services.

### Management of the service

1. The management and provision of the service would be the responsibility of IT Services.
2. A user group would be established for users of the facility. This would meet when necessary to discuss use of the service and make suggestions on how the service could be developed.
3. An advisory group would be established as a subcommittee of the University's Information Systems Committee to advise IT Services and the ISC on all aspects of the facility (including usage policies, planning and development). It would be expected that this group would also act as a resource allocation group when required.
4. All research groups who provided a cluster of nodes which was attached to the service would be represented on the advisory group. A senior member of the University with responsibilities for research would be invited to chair the group. Appropriate IT Services staff would be in attendance.

### Access to the resource

1. Access to the resources would formally be subject to the approval of the advisory group. Normally IT Services would make the approvals following the agreed policy set out by the advisory group and approved by the University's ISC.
2. Research groups would have priority access to the cluster(s) they provide; however, they would normally be expected to make about 10% of their cluster capacity available to others within the University of Sussex. Exceptions to this would be subject to the agreement of the advisory group.

### Resources required

#### *Staff*

In order to deliver this service IT Services would require:

- A service manager with HPC, UNIX, good user liaison and management skills;
- A systems programmer with UNIX and HPC skills.

This small team would also require support from within parts of the IT Services infrastructure team (UNIX/LINUX, Networking and storage).

The additional cost of the two dedicated staff members would be of the order of £95Kpa.

#### *Computing equipment*

Funds would be required to establish the "hosting" front-end facility and probably a small University- funded cluster to provide initial facilities for the service. An initial investment of around £120K would provide a suitable system with a resilient front-end system and local file store (10TB) and a 20 node (2 x 4 core,64GB, 200GB local disk) (total of 160 cores) InfiniBand connected cluster.

## Capacity

1. The initial (very small) team will have limited capacity. From experience elsewhere, we would expect the team to be able to handle the front-end and university cluster plus up to a further 5 or 6 small clusters (<80 nodes) providing these were either configured in the “default” way or configured and maintained by a third party.
2. HPC systems are very high in their demand for power (and cooling); the full costs of all power used (including cooling) and a contribution towards the cost of the running of the data centre would also need to be paid by research groups for “hosting” their cluster(s). There will be limited capacity within the IT Services data centres for hosting such systems and research groups will have to seek agreement that any system they hope to purchase can be accommodated before proceeding to purchase / bid for the system. The architecture of the service would allow a large HPC cluster to still operate via the university front-end but be housed outside the IT services data centres.
3. When a cluster becomes “old” and is no longer being maintained (by an external maintenance company), then it would *normally* be removed from the set of clusters being hosted by the service. Any exception to this would need the agreement of the advisory group and the Director of IT Services.

## Funding

1. The University would be asked to fund the initial set up of the service (£120K), the ongoing cost of the team (£95Kpa), additional electricity for the front-end and “university” cluster (£5Kpa), annual software licences and staff development and training (£10Kpa).
2. Research Groups using the service would be expected to contribute to the cost of the hosting service and to the cost of the replacement of the equipment providing the service. Assuming a three year life for the equipment, the service would need to generate an annual income of £150Kpa. (£95K + £5K + £10K + £120K/3). This could come from either a direct line charge within research grants, or from the overheads charged under FEC on the grants by the University.
3. Research Groups would be required to meet the full maintenance and licence costs associated with their cluster.
4. Advice will be sought from the Director of Research on how best to apply charges to the research grants providing the services.

## Next steps

Members of the University who are interested in HPC facilities are invited to comment on the proposal set out in this paper and to make suggestions to improve the proposed service.

If there is general support for the proposal (after improvements), then this will be taken forward by IT Services for consideration by the Information Services Committee and the University’s senior officers and funding sought to establish the service.

**Iain Stinson**

**March 27, 2009**