

## HPCC Allocation Policy Guidelines

### Committee Members:

Nikos Chrisochoides, Computer Science and Engineering  
Sharon Hammes-Schiffer, Chemistry and Biochemistry  
Edward Maginn, Chemical Engineering  
Dale Southard, Office of Information Technology

### Background

The above committee members were asked by the CTC to come up with recommendations regarding seven particular issues dealing with the allocation of high performance computing resources. The committee was charged with developing general guidelines in response to these issues (not necessarily specific implementation details). Below are the recommendations of the committee, as slightly revised by the CTC, in response to the seven issues raised by the CTC.

These recommended guidelines were adopted by the CTC on Dec. 6, 1999. CTC will review these allocation policy guidelines on a regular basis.

### Issues and Recommendations

1. *How do we deal with a shared facility that is established using research grant funds, but then also supported from MUR funding? What is the fraction of campus use and what is the fraction of use by the group that obtained the original research grant? The Hydra and IBM SUR grants were used as examples.*

Recommendation: The committee felt this was the key issue in developing a coherent allocation policy. It was unanimously agreed upon that access to a computational resource should scale directly with the level of monetary support that is put into the facility. That is, if a computational resource is obtained by a PI or group of PIs from an external research proposal, and the University does not provide any financial support of the facility (beyond standard, required matching funds), then the facility should only be available to the PIs. In this way, computational resources are treated in the same manner as any other piece of research equipment. However, if the University contributes additional funds *over and above* standard matching funds or if the University contributes funds for maintenance, then a portion of the computational facility should be accessible to the general University community. The “amount” of the resource available to the general community would be proportional to the fraction of the total cost of the facility (including maintenance, software, etc.) borne by the University. The amount of the resource available to each PI is governed by the relative fraction of the facility to which they made a contribution (either through writing of grants, or added support through capitalization funds, etc.) Whatever queuing system is used, reasonable accommodations should be made to PIs and the general public for performing large-scale benchmarks or when pending deadlines require dedicated use.

Example 1: if a PI obtains an NSF grant for \$45,000 with the University contributing a mandatory (for example) \$5,000 in matching funds, the total \$50,000 resource would not be eligible for use by anyone but the PI. If, on the other hand, the University contributed an additional \$50,000 to the project so that a much larger \$100,000 machine could be purchased, then one-half ( $\$50,000 / \$100,000$ ) of the resource would have to be made available to the general University community and one-half would be available only to the PI.

Example 2: The same PI is awarded \$45,000 with Notre Dame contributing a required \$5,000. After one year, the PI cannot provide maintenance on the machines and the University agrees to pay maintenance costs of \$10,000 in year 2. Now the total “cost” of the facility is \$60,000, and the general University community should have access to 1/6 ( $\$10,000 / \$60,000$ ) of the resource. If an additional \$10,000 are provided for maintenance in years 3, then 2/7 ( $\$20,000 / \$70,000$ ) of the resource must be made available to the general public.

**Comments:** The committee felt this was a fair way to allocate scarce resources and will encourage a number of positive actions:

- a) By allowing people who actually obtain external funding to control a resource, it will encourage people to seek funding, thus enhancing the University’s computational resource base.
- b) By allowing the University to “buy into” research facilities, it enables Notre Dame to provide all students and faculty access to the latest high performance computing technology without having to purchase standalone dedicated facilities.
- c) Allowing the University to pick up maintenance costs and thereby “open up” a portion of a facility to the greater University community will alleviate the dilemma PIs face with using funds to maintain computational resources that are perhaps not “state-of-the-art” anymore but that can be of great benefit to the general University population. This will enable advanced users to acquire the latest technology while still keeping in operation older but useful machines. Note that the University would not be required to pay this maintenance, but would have the option of doing so.
- d) Limiting usage by individuals (see Issue 3 below) to fractions of the cost of the facility to which they are responsible will discourage “free riders” or the domination of a resource by a single individual or group. Groups who are hamstrung by computer resources will therefore be compelled to seek additional resources themselves.

The determination of what constitutes a University contribution to a facility should be done in a manner consistent with the principles applied to experimental equipment. For example, physical space, heating/cooling are normally considered as overhead and would ordinarily not be considered a University contribution. System administrators and specialized hardware unique to the facility and necessary for its operation would generally be considered as equivalent to a monetary University contribution.

A written contract-like Letter of Understanding should be prepared which clearly states the details of both the contribution to be made by the University and the facility to be made available by the PI or group of PIs. If executed, then the details of this agreement should be made available to the campus community.

*2. How do we deal with multi-platform issues in the HPCC? Should everyone have equal access to each platform? Who decides who goes on what platform?*

This is an issue that is already upon us. The future of computing is multi-platform, and Notre Dame must embrace this fact. This issue takes care of itself if recommendation #1 is followed. That is, users who go out and acquire resources will base their platform decisions on which platform best meets their needs. If a given application doesn't run on a platform available to a user, the user then can either acquire a platform that does meet their needs, or re-write the application to take advantage of the platform that is available to them. Users who don't care about a platform or have small computing needs can use public time available on any resource they like.

*3. Should an allocation be done by individual user or by research group (that is by a PI)? A database that ties users to PIs would have to be established if allocation is done by research group.*

In the HPCC allocation model we are proposing, allocations must be based on PIs, not on individual users. This is because resources are acquired by PIs, not users. If a PI has a large group that requires extensive computing resources, they will be compelled to support their large group's needs by obtaining external funding and thereby guaranteeing that their group has a large share of a computational resource. This again treats computational facilities in much the same way as regular experimental equipment.

In the database of PIs, provision should be made for usage in courses. Generally a course allocation would be equivalent to a PI allocation. Courtesy dictates that, prior to the beginning of the semester in which the course is offered, the instructor discuss with the appropriate system administrator the demands of the course on the facility. Instructors who receive a course allocation should be required to explicitly inform students that the course allocation is to be used for course purposes only and not for other work. Similarly the course instructor and teaching assistants are expected to use the course allocation for course purposes only and not for their other work.

*4. When faculty and staff leave campus, how long should they be allowed to access computing at ND?*

Following recommendation #1, they should have access to any resource in the relative proportion that they supported the resource. That is, if a PI provides 50% of the cost of a computer, they should be allowed 50% access as long as that facility is up and running. As the facility ages and the equipment is gradually retired or upgraded by the University

or other researchers, the person who left campus will see their share of the resources gradually decline.

*5. Graduate accounts expire after graduation. Can they be extended if necessary?*

If the PI responsible for this graduate student wants to allow this student to use his/her allocation, that is up to them. The PI maintains the same total allocation level, however, regardless of how many students he or she has currently or how many graduate.

*6. Guest vs. Ghost accounts and the processes and procedures for approval and removal. Guest accounts are created when a ND faculty collaborates with faculty from another institution on a research project. Can the "guest" faculty bring \$\$ with the collaboration? Ghost account belong to people outside of ND that have left the University for one reason or another and still have active usage on ND's HPCC.*

This can be handled in a similar fashion to Item #5. If a PI wants to allow a guest to use a portion of his or her allocation, that is the PI's option. If a guest brings money with them, then this can be used to "buy into" a facility. The guest would then be allowed to use a fraction of the facility in proportion to the amount of money they bring to the facility.

*7. What is the mechanism for allocating time across resources? Is this some kind of "bank points" systems, some kind of fair share system, or what? Can't we give a certain amount to each person (or group) who can then appeal if more is needed?*

Resource allocation to each PI should be proportional to their contribution to a given facility. The method for implementation/enforcement of this policy is to be the responsibility of the primary administrator for that facility, with oversight for shared PI/University resources provided by an appropriate body of the University (i.e. OIT, the CTC, or an appointed committee).

CTC will review allocation policy guidelines stated above on a regular basis.