

# Master Document

## High Performance Computing

### Universidad de los Andes

20 Junio 2014

## 1 Governability

### 1.1 Guiding Principles

ICAR is a joint venture of the Sciences Faculties in partnership with the Direction of IT Services (DSIT) and open to any other university instance interested in HPC. The guiding principles are flexibility and decentralization.

Oriented towards Academia/Research/Consulting. Respond to the HPC education and training needs at Uniandes by providing support and participating actively in the process of research, teaching, and academic consulting activities.

Flexibility for the user Minimize the need for complex administrative processes in the operational decision-making. Respond immediately to the different requirements and technological needs that may appear during research, teaching, and consulting activities. Be able to adapt to the different academic needs in a speedy fashion for the benefit of the user.

Responsiveness to user needs. Commitment to continued improvement in the quality of services offered responding to the feedback of users.

Interdependence. The responsibility for the administration and operation of the Center depends jointly on all partners (initially represented by DSIT and Sciences as represented in the Central Committee).

Privacy and security. Protect user data according to current privacy and security standards.

### 1.2 Organizational Scheme

In order to reflect the Guiding Principles stated above we propose an Organizational Scheme as presented in the following diagram.

The Advisory committee (AC). Composed of external members from different computational centers around the world that guide and advise the steering committee with respect to different strategies and projects in research, teaching, and consulting to help reach the mission and objectives of the center.

\* The main function of the AC is to guide the efforts in high performance computing at Uniandes, promoting growth, sustainability, state-of-the-art technology, and a focus on research.

\* The AC is composed of 3 members proposed by the steering committee. The AC is elected every two years with a possibility to re-elect their members.

\* The AC receives a yearly report of activities (prepared by the coordinator and reviewed by the steering committee) that presents usage statistics, active projects, scientific production, teaching and consulting services activity, a financial report, and a projection for future years.

\* The AC meets once a year in the month of November at Uniandes with the steering committee. In this meeting the coordinator of ICAR makes a presentation summarizing the yearly report. This is followed by an oral presentation of the Advisory Committee with their observations regarding the annual activities of ICAR. Afterwards, this is followed by an open discussion on the future directions of ICAR. The coordinator integrates the comments into a general guideline for future directions.

The Steering Committee (SC) is the highest governing authority of ICAR and is responsible for defining its vision, mission and strategies. It works in conjunction with the ICAR coordinator. It is

composed by two members from each of the participating instances (initially Faculty of Sciences and DSIT) and the ICAR Coordinator. The members from the Faculties are full-time employees and will hold to their terms between 1 and 2 years; they are proposed by the Deans of their respective Faculties among the ICAR users on campus. The committee meets regularly once a month, and extraordinary meetings can be organized in case of special circumstances. The AC will meet every January with the DSIT to discuss the hardware needs for the following year and for renewal of the Service Level Agreement. The SC will actively participate in the annual meeting with the Advisory Committee.

The Steering Committee is committed to:

\* Supervise and make decisions for improvements of all activities at ICAR based on monthly reports presented by the ICAR coordinator.

\* The SC will serve as hiring committee for the Coordinator of ICAR and will present the candidates to the participating deans.

\* Define the policies and strategic plans to fulfill the mission and vision of the ICAR. The focus will be to ensure sustainable, high quality research output, high percentage of usage, and a high level of user satisfaction. The policies and strategies are written in a yearly report based on the suggestions from the Advisory Committee, the monthly reports by the ICAR Coordinator, comments from the users, and suggestions from DSIT. The report must:

- a) Define policies for the efficient use of computational resources.
- b) Define policies for the correct operation of the machines in the center. Such as approving the purchase of software and hardware through a discussion with DSIT; making new hires of technicians and coordinators.
- c) Direct the HPC Staff by communicating its decisions to the ICAR Coordinator.

d) Define the Service Level Agreement between DSIT and ICAR and between ICAR and end users.

e) Suggest the strategies for attracting external funding, and the policies for ensuring sustainability and managed growth.

The Coordinator reports to and follows the recommendations set out by the Steering Committee. The coordinator is hired through the Deans offices of participating Faculties on equal proportions (initially 100% covered by Sciences), and its yearly evaluation is conducted jointly by participating Deans. This role is central for the success of the Center for two reasons:

Firstly, the coordinator is responsible for maintaining fluid communication between the users, DSIT and the technical staff with the purpose of guaranteeing an efficient use of ICARs facilities.

Secondly, the coordinator acts as a leader to foster ICARs success. This involves maintaining a high level of excellence for all the research, teaching, and consulting activities carried out by ICAR; while maintaining a state-of-the-art infrastructure. Additionally, the coordinator will actively seek adequate funding level through external grant applications, industrial relationships, and donations.

The Coordinator is required to fulfill the following functions:

- \* Promote the usage of ICARs facilities within the Faculty members at Uniandes to support research, teaching and consulting activities. This is done by directly engaging Faculty members and increasing membership and usage in ICAR (indicator 20 Faculty members actively engaged in ICAR and 70% average usage).

- \* Be in charge of the hiring process and supervise the activities of the Technical Staff.

- \* Ensure that the service agreements (ICAR-DSIT and USER-ICAR) are implemented properly. Respond to written requests by the Users and DSIT.

- \* Interact with other HPC centers in the region and around the world to maintain and improve ICARs computing efficiency and scope (indicator 2 regional and 2 international collaborations)

- \* Advise researchers on how to improve their research productivity by using new software or computational tools available at ICAR (indicator satisfaction survey of users).

- \* Participate in joint research proposals with researchers and engage with the external sector to attract funds with the purpose of increasing the infrastructure capacity of ICAR.

- \* Precede the SC and ensure the proper record keeping for the committee meetings.

- \* Present a monthly oral report with performance indicators such as: usage, research projects in-course, teaching activities, scientific production, consulting activities, which will be presented to the SC (indicator 1 report per month).

- \* Write an annual report summarizing the results of the monthly reports and presenting a future perspective (indicator 1 report per year).

- \* Implement the recommendations by the SC and AC, through modifications in the service agreements, and user interactions.

- \* Propose the yearly budget for investment and operation.

The profile for such coordinator is someone with research experience, a publication record, in the Sciences/Engineering and working knowledge of the technical aspects of running a HPC cluster. Additionally, this person must have great interpersonal skills allowing efficient communication with researchers, technicians and managers.

The Technical Staff are technicians coming from each one of the ICARs partners. They will be dedicated 100% of their time to ICAR. They respond to the directives by the Coordinator. Their main responsibility is to implement technical solutions according to the decisions taken by the HPCC committee.

They are committed to

- \* Support teaching programs in HPC at graduate and post-graduate level.
- \* Provide short tutorials and workshops to illustrate the use of ICARs facilities.
- \* Administer all the computing platforms installed at ICAR.
- \* Implement new HPC platforms at ICAR.
- \* Advise users on HPC performance optimizations adequate to their research projects.
- \* Keep and active research stance towards new HPC options and trends that could be implemented at ICAR.

Descriptions of Users and DSIT are based on our expectations, and their role and responsibilities will be explicitly stated on their respective Service Level Agreements that will be renewed yearly.

Users. Their main role is to efficiently use the services provided by ICAR with the goal to reach specific goals in their learning and research processes. They are expected to:

- \* Clearly state their goals in using the ICARs facilities.
- \* Follow the guidelines established by the Steering Committee for the users at ICARs facilities.
- \* Provide information to the Coordinator about their projects (difficulties, successes, obtained results).
- \* Provide feedback to the Coordinator about their user experience.
- \* Give credit to ICAR in all the mediums used to present the results of their research.

DSIT is an active and essential partner of ICAR, with representation on the SC to help coordinate and guide to successfully reach our objectives. DSIT will also be involved through a Service Level Agreement that, in general, will include the following aspects:

- \* Provide physical space, electricity, cooling and connectivity for all the ICAR machines.
- \* Process buying orders following the directions of the Steering Committee.
- \* Provide insurance and technical support for software and hardware.
- \* Regularly perform machine maintenance.
- \* Facilitate the performance of different activities by the technical staff associated with the Faculties of Science and Engineering.

\* Suggest and participate on the selection and implementation of the technical architecture of ICAR.

## 2 Sustainability

The Faculty of Sciences and DSIT will provide the start-up source of funding for ICAR.

Starting from the second year of activities we expect to have a larger input from Principal Investigators (PI) of projects running at ICAR. The following paragraphs describe in detail how this PI investment will be channeled.

Our University is characterized by having a percentage of Principal Investigators (PIs) with substantial external funding (>100K USD), and a percentage of PIs with limited internal and external funding (<15K USD/year). We need to develop a mixed model to support the activities of those professors with extensive funding and the activities of PIs with limited funding or who only require occasional usage of the ICAR.

Is very important to take into account that the HPC should be purchased, built and set up in a single initial investment (to maintain homogeneity and optimal connectivity among the parts), which will be renewed every 5 years. For this reason, there will be a call for 6 months, every 5 years, where the potential users (defined as: PIs, Departments or Faculties) will determine how many servers they will buy, thus determining the size of the HPCC. Addition of equipment during the following 5 years should be kept to a minimum and under authorization of the SC. Because of this restriction it is very important to plan ahead the needs not only of individual investigators but of the department or faculty as a whole.

The following models explain how the user may gain access to ICAR, the cost associated with the access is calculated only on the cost of the 5-year investment on equipment. It does not consider costs of salaries, physical space, maintenance, etc, which should be covered by the participating partners (Departments, Faculties and DSIT). In principle, relieving the administrative and technical costs from the users should become a principal motivation for individual investigators to invest in the HPCC.

ICAR will be composed by servers bought by individual PIs, servers bought by Departments or Faculties (for late buy in model) and a 5% of the servers will be bought by DSIT for the subscription model.

Initial investment model: PIs invest money from their grants to build ICAR. The PIs invest in purchasing servers, and the University provides the technical support and the basic infrastructure to sustain the cluster. The PI has priority access to the servers they purchase, and a first come, first served access to all of the other servers in ICAR (for up to 4 hr runs in individual jobs). The servers bought by individual PIs may be used by other users (for up to 4h jobs) when they are idle, implying that in the moment the PI want to use his/her own servers, there should be a maximum waiting time of 4 hours.

Late buy in model: In this model, the individual Departments and Faculties will make the initial investment, taking into account the needs of their researchers. This model will benefit PIs that want to have their own machines but do not have the initial funding or the need to acquire a whole server (24 cores). They will be able to buy individual cores from the ones acquired by the Departments or Faculties at any point. This will allow the user to use the number of core(s) acquired with priority and no time restriction and still have access for the 4h runs in the other available nodes. This model could also take into account cases where the funding of a PI is not available at the moment of the initial investment or available all at once (e.g. when an assistant professor wants to buy 1 server out of their FAPA grant but divided throughout the first 3 years of tenure). The model should also be an opportunity for users that join the University after the initial purchasing period. Each Department or Faculty will have the freedom to define the criteria on how the nodes or resources will be distributed among potential users. Could be a first come, first serve mode or an internal call where the use of the computational resources must be justified as part of a research project, priorities could be made based on providing opportunities to young faculty with limited startup funding.

Subscription model: For PIs with limited funding or occasional need of computing resources. In this case DSIT makes an initial investment equivalent to 5% of the total number of servers in ICAR. The user will pay for a yearly or monthly (for occasional users) subscription for these resources. The subscription will allow the user to submit jobs of up to 4hrs at a time using any of the available DSIT purchased servers or any other idle core on the whole ICAR, thus maximizing the usage and minimizing idle times in ICAR. The money obtained with this subscription fee will be used to reimburse for the initial payment of the 5% of shared servers (initially paid by DSIT) and other HPCC costs. It should be considered to waive this subscription fee for first time users as a further motivation to engage them in the use of the ICAR.

#### Other HPCC Funding

In order to fund all other HPCC associated costs, an important commitment has to be made by the participating Departments, Faculties and DSIT. The different partners through their representatives in the SC should decide how the cost of personnel, technical support and maintenance should be distributed. Additional ideas for sustainable funding are:

- \* Apply for collaborative research grants either as a multidisciplinary center (HPCC) or by collaboration among individual users from different departments. These projects could be proposed for external or internal (University) funding. The HPCC coordinator will have an important role in promoting this activity and looking for funding opportunities and collaborations.

- \* Academic uses of the HPCC and ICAR. Shared nodes and the HPCC personnel could be used in the context of the academic life in the University (see below Academic aspects), where faculty could use the computational resources as part of their teaching tools with a minimal cost or the technical staff could use their expertise to create and offer lectures in high performance computing related topics.

### 3 Technical Aspects

The requirements for ICAR impose different requirements from those of other Campus Services used in Teaching (i.e. SICUA) and Communications (i.e. email). The reason is two-fold. First, the diversity and types of scientific data and projects implies that there is no general service that could be offered to all users; second, the technical aspects of HPC solutions require a close and frequent interaction of users and the hardware and software administrators. This goes from the very first instance of software installation to their performance optimization.

This is translated into the need of users to interact with the system at its most basic level (users must be, for instance, capable of installing and modifying software), and the need to prioritize efficiency and flexibility over friendliness of the system. In particular this implies that there is no need or benefit on the usage of virtual machines. Instead just accounts on a mounted shared file system are ideal.

To do this, the system will need to be based on robust software for job administration where several different queues should be created. While some jobs will be very processor intensive which will benefit of parallelization, others have high demand on RAM, while other jobs are very intense in I/O, reading and writing to disk (up to several hundred Gb in a single job), so the speed of the network is essential. Due to these particular needs and the fact that it will be a computing environment as no other offered currently by the University, there is a need to have both dedicated machines and dedicated staff to take care of ICAR.

A description of the basic needs, the goals, and the current standing of resources is given in the following table:

Resource	Minimum Required	Goal	Current availability
No of Accounts	/Users	25-50	
CPU hours	/ year	200K	
Max RAM per job		500G	
Storage per account		1-2Tb	
Storage under backup			
Cores	180 (Physics + Biol)		
Speed of the cores			
Speed of the connectivity		10G	

### 4 Academic aspects

A metric for the ICAR success is a high level of usage with a minimum of idle times. We believe that a central strategy to reach that goal is to be able to quickly integrate the students at Uniandes (both undergraduate and graduate) into projects running with ICAR facilities.

This asks for an academic strategy in order to:

- \* Regularly offer semester-long courses on HPC scientific computing, ideally joint between Sciences and Engineering.
- \* Regularly offer training workshops to use the ICAR facilities.
- \* Maintain an updated web presence with all the results, publications and congress participations.
- \* Encourage the regular acknowledgement of the usage of ICAR facilities in all publications and theses.
- \* Organize a social gathering at least once per year to get to know all the students, researchers and staff that use and maintain ICAR facilities.

## 5 Additional Suggestions

The success of the HPCC partnership also depends on nurturing a culture that incentivizes the use of the resources and the communication among partners. We believe that the following points must be kept in mind at the moment of deploying the structures and policies described in this document.

- \* Run for free for first time users. The Community Cluster must run using the initial contributions from all the PIs, Faculties and Departments. For first time users there must be the possibility to use the resources for free, subsidized by other contributors.
- \* Standard software must be provided to everyone for free.
- \* Negotiate with Departments to pay for specialized software to be installed in the cluster. More specialized commercial software can be bought and installed by the user, but multiple users who want the same software can share the cost.
- \* Negotiate with Vicerrectora de Investigaciones to apply for a programa de investigacin in order to build and extend the cluster.
- \* Scientific Visibility. Select a series of flagship programs in Sciences and Engineering that must produce interesting results and publications using ICAR facilities. This must be used to raise awareness about the existence of ICAR and its potential.