Cluster-UY: collaborative scientific High Performance Computing in Uruguay

Sergio Nesmachnow and Santiago Iturriaga

Universidad de la República, Uruguay
{sergion,siturria}@fing.edu.uy

Abstract. This article describes the national initiative for installing and operating a collaborative scientific HPC infrastructure in Uruguay (Cluster-UY). The project was conceived as a mean to foster research and innovation projects that face complex problems with high computing demands. The main ideas and motivations of the Cluster-UY project are described. The technological decisions to install the platform are explained and the collaborative operation model to guarantee sustainability is introduced. In addition, the perspectives of the national scientific HPC initiative are highlighted and sample current projects are presented.

Keywords: HPC, scientific instrumentation systems

1 Introduction

Scientific instrumentation systems play a major role in the development of high-quality research, especially in those areas of science and technology that heavily rely and emphasize on experimental studies. In this context, having platforms that implement scientific instrumentation services allows researchers to strengthen proposals, broaden the horizons of researches, and obtain high-impact results that are validated in practice [13, 24].

Nowadays, a specific and relevant case of scientific instrumentation systems are Supercomputing or high performance computing (HPC) systems. These systems allow solving complex problems that have large computing demands (e.g., because they deal with complex mathematical models, address very complex problem instances, and/or manages very large volumes of data). Parallel computing techniques [9] help researchers to face the aforementioned complex problems and solve them in reasonable computing times. However, specific high performance hardware is required to apply HPC and parallel computing techniques.

The reality of the scientific instrumentation systems in the world is diverse. There is a clear differentiation between developed countries and Latin American countries, from both the conceptual point of view and also from the methods applied for its implementation and operation. Regarding HPC systems, even within Latin America different situations are identified (e.g., between Brazil and other countries). HPC platforms are expensive and significant efforts are required to acquire, maintain, and operate the facilities.

This article describes a proposal for an HPC scientific instrumentation system to be installed and operated within the current reality of Latin American countries. The proposal was conceived as a mean to foster research and innovation projects with high computing demands, in a context where monetary resources are scarce. An approach based on ideas from collaborative economy is proposed for operation. The impact on science and technology is described and sample applications are briefly described.

The article is organized as follows. Next section describes scientific instrumentation systems and HPC facilities. Section 3 introduces the Cluster-UY project. Section 4 presents the details about the proposed self-managed business model. The impact and sample applications are described in Section 5. Finally, Section 6 formulates the main conclusions and current lines of work.

2   Scientific instrumentation systems and HPC facilities

This section describes the main concepts about scientific instrumentation systems and HPC facilities.

2.1   Scientific instrumentation systems

Nowadays, most of scientific research, and all of experimental research, require using various equipment to address experimental analysis, simulation development, validation of prototypes, and other activities [24].

Scientific instrumentation services are structures for research and development that provide access to specialized and up-to-date equipment, and also to expert personnel in the methodologies for using the equipment. A specific objective of scientific instrumentation services is to diversify and expand the use of equipment by an increasing number of researchers, providing in turn the mechanisms, assistance, and training needed to take advantage of the systems. In universities, having sophisticated scientific instrumentation is vital to develop quality research. Scientific instrumentation services complement formation and research activities developed in groups and institutes [13].

A specific feature of modern scientific instrumentation systems, which is very relevant for the proposal introduced in this article, is their high economic cost and their very fast obsolescence. Having a well-planned and efficient service allows optimizing the resources allocated to the purchase of equipment. A centralized management of the infrastructure also allows reducing operating costs and improving the efficiency of operation/maintenance tasks by specialized technicians.

Due to the aforementioned reasons, the implementation of a scientific instrumentation service that provides various advisory services in the use of equipment is an efficient and rational alternative for research groups to make use of large and highly qualified equipment. In those institutions where there is no scientific instrumentation service, authorities must raise awareness in research groups that require the use of specific equipment to initiate the formation of multidisciplinary groups to gather critical mass and interest on a scientific instrumentation service.

Scientific instrumentation services also provides support to researchers for operating the equipment, advise on the design and planning of research methodologies and experimental analysis based on the computing infrastructure to use, and assistance on methodologies for processing results. The training of highly specialized technical personnel is fundamental for the correct operation of a scientific instrumentation service. Institutions must develop and strengthen a group trained to operate the service, including new personnel from those groups that gradually begin using the service. The most important lines for training include development and maintenance of tools, techniques to improve practice, scientific dissemination, design, analysis and adaptation of protocols for using the equipment, training in attention to users, etc. Likewise, institutions must provide or seek to obtain resources that complement the specific contributions of the research groups that use the infrastructure, to bring from other universities or foreign centers professors to train technicians, researchers, and students in situ, and facilitate the mobility of technicians to train and learn in other institutions.

2.2 Management models of scientific instrumentation services

A scientific instrumentation service requires applying management techniques to guarantee the effectiveness and efficiency of the service model offered.

The main aspects to be managed in a scientific instrumentation service correspond to: i) the *availability of equipment, technicians, and methodologies* for using the infrastructure, considering the different models in which researchers can use the services according to your concrete needs; ii) the *location and access to the infrastructure and services by the researchers*, considering the geographical location of the equipment, which can be centralized in a single installation of large dimensions or have different degrees of distribution in smaller and more dispersed facilities, and contemplating different policies and mechanisms for accessing equipment and services through local or remote interfaces; and iii) the *maintenance and permanent updating of the equipment and the knowledge required for its efficient use*, including updating the infrastructure and equipment, considering the technological advances in products related to the discipline, and technical staff and researchers, contemplating the training in the methodological advances in operation and use and the possible future lines of work and application of the equipment for the scientific advance [13].

All the aforementioned aspects must be taken into account for a correct sizing, operation, operation and updating of the services, so that the scientific instrumentation system is an effective, agile, and dynamic tool adapted to the needs of researchers, which can provide an advantage to the development of quality scientific research. The management model must consider all these considerations to guarantee the success of the proposed system.

2.3 Large computer systems as scientific instrumentation services

One of the technological transfer products recognized as the most spectacular, has been the development of computer systems.

In the last 50 years, large computer systems have become essential tools for scientific research, allowing calculations that would be impossible with other techniques.

Most of the modern scientific advances have been produced by applying systems that combine scientific instruments (microscopes, telescopes, etc.) and large computational systems that control the instruments, and process and analyze huge volumes of data from experiments. Through the application of new methodologies and techniques in the areas of mathematical methods, numerical simulations, computer science, software engineering, high-performance computing, distributed computing in grid and cloud environments, visualization, etc., computer platforms have been capable of providing researchers with the computational power necessary to solve problems involving complex systems, large-scale mathematical models, experimental analyzes that demand the processing of large volumes of data, and other problems of great difficulty [5, 9].

Turnkey-purchased computer systems (e.g., integrated supercomputers) have an extremely high cost that makes them prohibitive for small and medium-sized institutions. These systems are closed, practically do not allow modifications, and updating/maintenance is restricted to products and applications of a certain type. In addition, closed systems cover general needs that are sometimes different from the specific demands imposed by scientific research. These limitations are of especial relevance in our Latin American, and especially in universities. Usually, our institutions must develop their own computational systems, built with low cost components and applying knowledge of electronics and computers. These systems allow using the paradigm of open source software, based on freely developed and distributed products, which allow modifications and therefore the development of custom solutions, having a direct impact on the applicability of the instrumented system and expanding the target audience. Likewise, free software products allow reducing the operating costs of the platform, making possible the implementation of an inclusive model for a scientific HPC platform.

The proposal presented in this article refers to a system of scientific instrumentation that is built on the ideas of an open platform built through the integration of low-cost components and the paradigm of open source software, as support for egalitarian research under the Developmental University model [3].

3 The Cluster-UY project

This section presents the National Supercomputing Center of Uruguay and its technical details.

3.1 The previous project: Cluster-FING

Cluster FING [14] was the HPC infrastructure of Engineering Faculty, Universidad de la República, Uruguay, that provided support for solving complex problems that demand great computing power in science and engineering. The

initial infrastructure for Cluster FING was funded by Sectoral Commission for Scientific Research, Universidad de la República, Uruguay, in 2008.

Cluster-FING began operating in March 2009. Previously, there were no centralized resources for scientific computing in the country with reliable, sustained, and growing capacity over time. Until that date, scientists acquired their own computing resources at their laboratories. Acquisitions were limited to the very limited resources available in specific projects (a few thousand dollars, which in general were only enough to acquire a conventional computer, not a powerful server). In addition, the acquired infrastructure did not include any type of update. Rapid obsolescence meant that at the end of the research project, the use of the computer was very limited. Researchers also faced the problem of the great inefficiency of the human resources required for the administration of the equipment and the energy costs (operation and thermal conditioning).

The limited capacity of computer equipment at laboratories limited the challenges that researchers could face, effectively cutting off the creativity of projects. Cluster-FING proposed a first step to change the paradigm of non-cooperative use of computing resources. Collaboration between different research groups was proposed to acquire a more powerful and shared computing platform. Instead of buying their own computing resources, i.e., spending valuable research funds, researchers had the opportunity of contributing to Cluster-FING and use their services. The contributions were significantly lower than in a non-cooperative model, as a result of savings due to centralized management of computing resources, equipment acquisition, operation, and energy costs.

Cluster-FING started to break down the barriers that had the computing capacity available individually in a laboratory as a limitation of the creativity of researchers. Gradually, the new paradigm had a significant impact on research in basic and applied sciences. As a result, researchers enhanced the quality of research and set new horizons and objectives, with application to the national reality. These more ambitious objectives have in turn implied a greater need for scientific computing resources. Likewise, the effectiveness of the Cluster-FING model caused that a greater number of research groups were interested in using the infrastructure. Cluster-FING progressively increased its computing capacity (from 8 servers with 64 computing cores in 2009, to 20 servers with more than 500 cores in 2017) and reached ten million hours of effective computing core in July, 2017, executing more than one million computing hours per semester.

3.2   National Supercomputing Center: Cluster-UY

The previous experience of the Cluster-FING project showed the viability of creating a HPC platform for HPC by applying a collaborative model. The new initiative described in this article proposed implementing a scientific HPC service at national level, based on an aggregation architecture (cluster) type, to create the National Supercomputing Center, Cluster-UY.

The main motivations of the Cluster-UY project include the success of the paradigm of use and centralized management of computing resources initiated by the Cluster-FING and its significant impact on national research in various

areas of science, the need for greater resources of scientific computation, and the increase in research groups that use scientific and HPC techniques in Uruguay.

The Cluster-FING project increased the demand of scientific applications, which exceeded the available computing capacities. To significantly improve the ability to address relevant scientific problems in the country, allow the resolution of larger problems, and promote better quality of research, the Cluster-UY project proposed scaling the dimension of the computing infrastructure and extending the service at national level. Cluster-UY provides support to research, development, and innovation activities of all scientific, technological, industrial, business and social communities, providing free and unrestricted access, focusing on the egalitarian paradigm proposed by the Developmental University model [3].

The main features of the proposed model are presented in Section 4, linking them with the installation, management, and administration of scientific instrumentation systems [13] and with relevant ideas of collaborative economy [1, 4].

3.3 Technical description

Cluster-UY is comprised of 31 nodes: 28 computing nodes, two file-server nodes and one service node. Each computing node is comprised of two Intel Xeon Gold 6138 CPU with 20 cores each, 128 GB of RAM, a solid-state drive with a capacity of 400 GB, and one Nvidia Tesla P100 GPU. In total, Cluster-UY is comprised of 560 Xeon Gold computing cores delivering 35 trillion double precision floating-point operations per second (TFLOPS) and 28 Tesla P100 devices delivering 131 double precision TFLOPS. Hence, its peak computing power is 166 TFLOPS.

Storage is served by two dedicated nodes: FS1 and FS2. FS1 hosts a RAID 1+0 storage with 60 TB of raw capacity (30 TB usable) and FS2 hosts a RAID 6 storage with 80 TB of raw capacity (60 TB usable). FS1 is used exclusively for storing user data since RAID 1+0 provides adequate fault tolerance levels while excelling at read and write performance, and it allows a rapid recovery upon failure. FS2 is dedicated for storing non-critical data such as intermediate computations, unprocessed raw data, initial conditions for experiments, etc. RAID 6 was chosen for FS2 because it increases the usable/raw capacity ratio when compared to FS1. However, RAID 6 provides much lower performance than RAID 1+0 for write operations and it may require heavy rebuilding operations when a failure is detected, so it is not advised to store critical data in FS2. On top FS1 and FS2, the high-speed SSD local storage of each computing node (11 TB) is used as scratch space for short-lived data. The available raw storage space was expanded by incorporating 144 additional TB in 2019.

Finally, the service node hosts three critical services isolated from each other in different virtual machines: a distributed resource manager, a front-end access to Cluster-UY and an application development environment for users. Virtualizing these services allows improving their provisioning, availability, and recovery. The Simple Linux Utility for Resource Management (SLURM) is used as distributed resource manager. SLURM is a well-known open-source utility for managing resources in Linux-based clusters [25]. The distributed resource manager is a key component of every HPC cluster. Its role consists in tracking which

resources are in use and which are available, scheduling which job to execute first and in which resources, starting the actual processes of the scheduled job, and finally cleaning resources when the job finished its execution. SLURM also enables the accounting and report of resource usage for each user. The front-end access or *login VM* provides the entry point to the cluster to all users. This login VM is a stripped-down system with no development tools. However, the development environment or *dev VM* is a separate virtual machine with a full set of development libraries.

All nodes are interconnected by two separate networks. The administrative network is supported by a 1 GbE network and is used primarily for lights-out management of the nodes and for monitoring tasks. The general purpose network is supported by a 10 GbE network and is used for all other network traffic. User's homes are shared among all nodes using NFS over the general purpose network. Fig. 1 shows the overall schema of Cluster-UY.
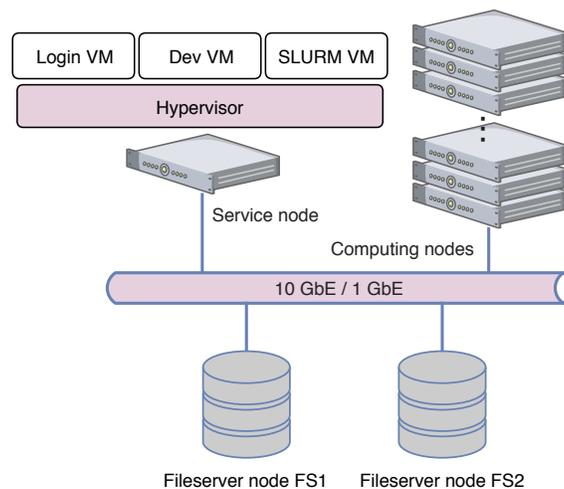

Fig. 1. Schema of the Cluster-UY infrastructure.

4   Installation, operation, and business model

This section describes the main details about the operation and services of the National Supercomputing Center, Cluster-UY.

4.1   Geographical location and access to services

A flexible centralized model is applied for Cluster-UY. The infrastructure is hosted in a large centralized installation, in datacenter "Ing. José Luis Massera" of the National Telecommunications Administration (ANTEL), a site that provides the necessary functionalities for the correct access, management and operation of the equipment, and to optimize the necessary energy resources.

A specific agreement between ANTEL and Universidad de la República allow both institutions to promote research in the related areas of knowledge (datacenters, mass data processing, data networks, etc.) and provide the services of the new platform in a reliable and efficient manner. Likewise, an agreement with the National Electricity Administration (UTE) was signed to receive support with operation costs (e.g., electric power for servers and thermal system). Cluster-UY services are provided as a counterpart for research and initiatives carried out by both public companies. These partnerships are pillars of the proposed self-sustaining model. Interconnection with smaller computational infrastructures available in the country are also considered, as well as interconnection with computing platforms existing in the region, to share resources through the mechanism of voluntary peer networks [2]. This model establishes a paradigm of collaborative consumption, while promoting sustainable development and energy efficiency, as proposed by collaborative economy.

Access to the platform and its services is guaranteed 24 hours a day, 365 days a year, from anywhere in the country and abroad. Cluster-UY services are accessible through Infrastructure as a Service, Software as a Service, and Platform as a Service models, operating at different levels according to the needs of each problem. Remote access and proper training are provided to allow users to gain a level of expertise to work autonomously and perform simple tasks without the assistance of experts. Two modalities are used: technicians in charge of solving specific problems in the infrastructure provide services on a permanent basis, as established in the hosting agreement of the infrastructure subscribed with ANTEL, and assistants scientists who support researchers using the infrastructure are available on a continuous schedule in working hours.

4.2   Services provision and potential users

Services provision. A mixed mechanism is applied for service provision, allowing two modalities. On the one hand, self-service utilization of the instrumentation by researchers is encouraged, mainly through assisted works, in which researchers with more experience contribute to guide less experienced researchers and students. On the other hand, continuous training is implemented by academic technicians to expand the base of users with knowledge of HPC techniques. New users will be requested, once they acquire the necessary knowledge to autonomously operate the platform and take advantage of their services, to collaborate with training/mentoring of new users.

This way, group work is encouraged, users are linked to each other and to the platform, and a cooperative training mechanism is established to overcome the lack of funds to develop formal training strategies through courses. A model of open knowledge and mutual assistance, two relevant postulates of collaborative economy, is then configured.

Cluster-UY services reach to a wide range of users, both internal and external to the research groups that proposed the project. Internal users are students, professors, and researchers of Universidad de la República, and profession-

als/technicians of companies that support the project. All internal users have guaranteed free access to the services, disregarding money or work contributions.

External users include technicians and researchers from institutions that use the Center through specific agreements to develop scientific-technological or social research activities. Peer networks are encouraged for technical/scientific assistance and to rationalize contributions. Services offered to external users are not for obtaining an income, but to allow the largest use of the platform and solidarily collaborate to the maintenance, administration, and updating of the infrastructure. Contributions from external groups are via agreements that establish the commitment of collaboration and clearly defines the contributions to be made (economic, if research funds are available, or work, in case of having experience for guiding/training other researchers, etc.) and their counterparts.

Users. All researchers can use Cluster-UY, even if they do not have funds. Foreign users also can use the Center when they work in collaboration with national groups, have research agreements, or propose a socially interesting research for our country. The agreements must establish the commitment of collaboration and clearly define the contributions to be made and their counterparts.

Assisted works, through the solidary collaboration of experienced researchers, are strongly promoted. These models can be applied due to the simplicity of use of computer services, i.e., delegating to users basic operative tasks that can be performed in a restricted environment, while technicians carry out heavier tasks (e.g., installation of libraries and frequent used software packages). Restricting the tasks enabled to users allows minimizing problems associated with installation/configuration of software products, especially considering the number of potential users of the Center (500-1000), thus impacting on the quality of service. Technicians work jointly with researchers and students and courses are organized to train users in the use of HPC techniques. The service focuses on frequent and very frequent users, with different levels of knowledge of the equipment, but knowing the specific techniques to be used in their research.

4.3 Organizational structure

An autonomous ad-hoc Commission, including representatives of scientists, researchers, agencies, and companies that support the Center, is responsible for making relevant decisions. The commission also advise in the preparation of bids for purchasing equipment, to plan a useful platform, according to the needs of researchers. The Commission meet periodically with authorities from Universidad de la República, National Research and Innovation Agency, and other institutions to guarantee a proper relationship with the research community.

A Manager is in charge of coordinating the daily activities of the technical work team, of analyzing technological alternatives for the expansion and updating of the infrastructure, and of interacting with the engineers of the data center where the platform is located. The Manager is the visible face of the Center acting as main promoter of the services offer in order to attract a greater number of users. The Manager position is occupied by a researcher with high technical

profile, appropriate knowledge of HPC, and experience in infrastructure management. The position is financed with part of the contributions of the research groups and the companies associated with the initiative.

4.4 Benefits, contributions and responsibilities

The Center applies an egalitarian model for access to services. Identical features are offered to all users, who in turn have the same responsibilities regarding the correct use, maintenance, and updating of the platform and services.

Contributions for operation. Voluntary collaborations of research groups, organizations, and companies that use the Center contribute to self-finance the constant updating of equipment and knowledge (maintenance, administration, expansion, etc.). The model is based on the previous experience of the Cluster-FING project, but including other universities, institutions, and companies.

Taking into account the financing difficulties for science in Uruguay, contributions are not mandatory, but voluntary and according to the projects. A low amount is suggested, allowing the Center to be attractive to researchers without affecting the development of other project activities. This way, a more equitable work mechanism is established, without requiring the economic contribution of smaller research groups. Furthermore, funding agencies can appreciate the advantages of the cooperative model to rationalize the use of resources.

Training courses, seminars, and self-study groups are used to gain expertise for the correct use of the Center, taking advantage of peer networks to acquire and share knowledge at national, regional, and international level. Members of the scientific/technical group that operates the Center took courses on management of HPC systems at Universidad de Buenos Aires, Argentina. Through a specific agreement, the participants did not pay for the courses registration.

Contributions for maintenance and updating are not obligatory, but users are provided with guides on the amounts to be quoted in their project proposals. The proposed business model considers the costs of amortization, maintenance, management, and updating of the infrastructure, to obtain a reference value of 0.01 USD for computing resource/hour. This value is *ten times lower* than the cost of computing services in the cloud. The contribution comparable to acquiring an isolated server, without HPC capabilities, allows using more than 600,000 computing hours in the Center. Even with the cost of buying a desktop computer, not useful for solving complex problems, a researcher can use 100,000 computing hours. Researchers must not pay additional costs for server administration, software management, network use, power consumption, or cooling. Furthermore, by providing a minimum contribution, researchers access to a number of computing resources a thousand times greater than on an isolated server. These advantages show the convenience of the proposed collaborative and self-financing model.

A greater reference value is considered for agreements signed with organizations and companies that obtain an economic benefit from their activities. Projects that are validated as social interest and non-profit can use the infrastructure without requiring financial contributions.

Undergraduate and graduate students do not contribute for using the platform. Specific agreements are established with higher education institutions and the Center in included into those systems available for use and qualifiable for the applications for competitive funds for maintenance and expansion.

Agreements. Work and support agreements are signed with institutions, organizations, and companies. Partnerships are established to consolidate the collaborative model and open the Center to a wide variety of users, to promote research, innovation, and development. The Center makes special emphasis on promoting inclusive development in areas with social impact (health, education, research for development, etc.), under the University of Development model.

Agreements allow consolidating a clear self-financing and self-management model, and a work method with real social impact, allowing organizations and companies to justify investments to support of the Center. All information, indicators, agreements, contributions, and utilization from the Center are publicly open. The management is also open to external audits by agencies that support the initiative. This approach encourages and consolidate an open data model, closely linked to the ideal of collaborative systems.

Responsabilities. The project is based on clearly defining the relationships between the participants, defining an egalitarian collaborative model. Participants are considered as peers with the same rights and obligations, building a network for collaboration to improve and make equitable access to the infrastructure and services of the Center. Work is proposed in a mutual assistance regime to access self-sustaining resources and services, which would be very difficult to access in a non-collaborative model. The impact of the proposed model is of great relevance for national research and helps develop and consolidate incipient research groups, promoting equity in access to resources, diversity and equal opportunities, following the ideals of University for Development.

5 Impact on science and innovation, and sample projects

This section describes the impact of Cluster-UY and sample lines of research.

5.1 Impact on science and innovation

Cluster-UY has had a direct impact on science and innovation, significantly increasing the goals and horizons of research activities within the country. The list of application areas include Astronomy, Bioinformatics, Biology, Computer graphics, Computer Sciences, Data analysis, Energy, Engineering, Geoinformatics, Mathematics, Optimization, Physics, Social Sciences, Statistics, and others.

Industry and public organizations are also using Cluster-UY. UTE is developing research related to the analysis of domestic power consumption patterns, load curve classification, energy efficiency, and other subjects. ANTEL is developing research on big data analysis, datacenter performance analysis, mobility of users, and other subjects. Other organizations and administrations are using Cluster-UY too, including: the National Administration for the Electric Market,

for supporting the power generation investment and energy export planning for Uruguay, the Ministry of Industry and Mining, for energy-related research, and the Pasteur Institute, for research on bioinformatics and biotechnology.

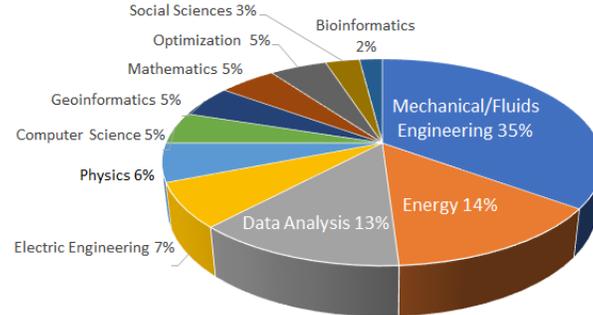Figure 2 summarizes the main areas of research using Cluster-UY.

Fig. 2. Main areas of research using Cluster-UY.

5.2 Sample projects and researches

Many research and innovation projects have used Cluster-UY to improve its capabilities and results. Three relevant examples are presented next.

Weather prediction models applied to wind turbines. This project studies wind velocity focusing on measurements at wind turbine heights (around 100 meters). Using experimental measurements and through extensive simulations numerical simulations using different planetary boundary layer schemes and mesoscale grid resolutions, a gust parametrization was proposed for wind forecasting in Uruguay. This gust parametrization provides gust factors to be applied to predicted turbine-level winds, achieving higher accuracy at coarser resolution than an algorithm based on surface layer data alone [11]. Figure 3 shows a forecast example of wind velocity values at wind turbine height. Red dots show considered the wind farms and each subfigure shows a possible atmospheric state, that is, a possible wind generation scenario in the forecast.

Analysis of mobility data from Intelligent Transportation Systems (ITS) in smart cities. ITS allow collecting large volumes of data that can be processed to extract valuable information for understanding mobility in smart cities [8]. The information can be offered to citizens planners, and decision makers, in order to improve the quality of service and user experience. This is a very important issue for Latin American cities, where the information can help improving public services. Our research group at Universidad de la República has applied data analysis, data processing, and computational intelligence for improving transportation systems and other public services [15, 16, 18–20, 17].
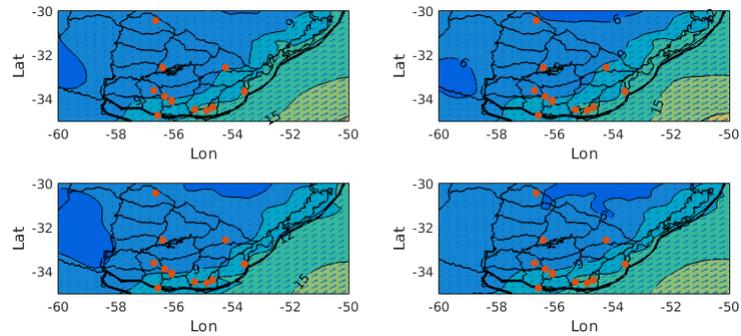
Fig. 3. Forecast example of wind velocity values at wind turbine height.

By applying a parallel-distributed approach for massive data analysis, the computing power of Cluster-UY has been applied to analyze data from the ITS in Montevideo (e.g., GPS location data from buses, ticket sales and smart card transaction data) and obtain valuable information to improve the access to the transportation system, quality of service, socio-economic implicances, etc. A general diagram of the proposed approach is presented in Figure 4. The approach has proven to be very efficient, achieving significantly large speedup values [7, 15], and also very valuable for citizens and administrator. A sample analysis is presented in Figure 5, showing a heatmap of ticket sales (smart card transactions) in the center of Montevideo in May 2015. Bright (white) pixels in indicate high concentration of ticket sales while dark (red) areas indicate low ticket sales.
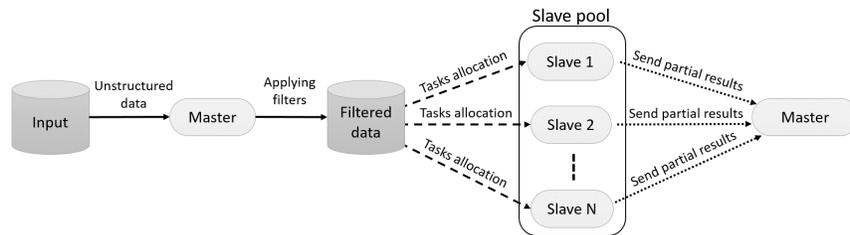

Fig. 4. Parallel-distributed model for mobility data analysis from ITS [15].

Other projects. Cluster-UY has also been applied to many other research efforts in several areas, including Astronomy [10], Biomedicine [6], Energy [21], Fluid dynamics [23], Statistics [12], Telecommunications [22], and others.

6 Summary and conclusions

This article presented the National Supercomputing Center, Uruguay (Cluster-UY), a national initiative for installing and operating a scientific HPC infrastructure following a collaborative operation model.

Fig. 5. Heatmap of ticket sales (smart card transactions) in the center of Montevideo

The Cluster-UY project was described and the self-funded collaborative operation model involving scientific institutions, academia, and public/private companies, to guarantee sustainability was clearly explained.

The perspectives of Cluster-UY as a mean to foster research and innovation projects that face complex problems with high computing demands were highlighted and sample projects developed in the Center were briefly presented.

The main lines for future work are related to continue developing and improving the infrastructure and services of Cluster-UY, as a key tool for improving scientific research in the country.

References

1. Algar, R.: Collaborative consumption (2007), leisure Report, april 2007
2. Andrade, N., Brasileiro, F., Cirne, W., Mowbray, M.: Automatic grid assembly by promoting collaboration in peer-to-peer grids. Journal of Parallel and Distributed Computing 67(8), 957–966 (2007)
3. Arocena, R., Göransson, B., Sutz, J.: Knowledge policies and universities in developing countries: Inclusive development and the developmental university. Technology in Society 41, 10–20 (2015)
4. Botsman, R., Rogers, R.: What's Mine Is Yours: The Rise of Collaborative Consumption. Collins Business, Nueva York, USA (2010)
5. Chapman, D., Joshi, K., Yesha, Y., Halem, M., Yesha, Y., Nguyen, P.: Scientific services on the cloud. In: Furht, B., Escalante, A. (eds.) Handbook of Cloud Computing, pp. 379–405. Springer (2010)
6. Curi, M., Carozzi, L., Massobrio, R., Nesmachnow, S., Danoy, G., Ostaszewski, M., Bouvry, P.: Single and multiobjective evolutionary algorithms for clustering biomedical information with unknown number of clusters. In: International Conference on Bioinspired Methods and Their Applications. pp. 100–112 (2018)
7. Denis, J., Massobrio, R., Nesmachnow, S., Cristóbal, A., Tchernykh, A., Meneses, E.: Parallel computing for processing data from intelligent transportation systems. In: International Supercomuting Conference in Mexico (2018)
8. Figueiredo, L., Jesus, I., Tenreiro, J., Rui Ferreira, J., Carvalho, J.: Towards the Development of Intelligent Transportation Systems. In: IEEE Intelligent Transportation Systems. pp. 1206–1211 (2001)
9. Foster, I.: Designing and building parallel programs. Addison-Wesley (1995)

10. Frascarelli, D., Nesmachnow, S., Tancredi, G.: High-performance computing of self-gravity for small solar system bodies. IEEE Computer 47(9), 34–39 (2014)
11. Gutiérrez, A., Fovell, R.: A new gust parameterization for weather prediction models. Journal of Wind Engineering and Industrial Aerodynamics 177, 45–59 (2018)
12. López-Vázquez, C., Hochsztain, E.: Extended and updated tables for the friedman rank test. Communications in Statistics - Theory and Methods pp. 1–14 (2017)
13. Navarro, C.: Los servicios de instrumentación científica. In: Master de Política y Gestión Universitaria, Universidad de Barcelona. pp. 73–127 (2017)
14. Nesmachnow, S.: Computación científica de alto desempeño en la Facultad de Ingeniería, Universidad de la República. Revista de la Asociación de Ingenieros del Uruguay 61(1), 12–15 (2010)
15. Nesmachnow, S., Baña, S., Massobrio, R.: A distributed platform for big data analysis in smart cities: combining Intelligent Transportation Systems and socioeconomic data for Montevideo, Uruguay. EAI Endorsed Transactions on Smart Cities 2(5), 1–18 (2017)
16. Nesmachnow, S., Massobrio, R., Arreche, E., Mumford, C., Olivera, A., Vidal, P., Tchernykh, A.: Traffic lights synchronization for bus rapid transit using a parallel evolutionary algorithm. International Journal of Transportation Science and Technology (2018), online first, `https://www.sciencedirect.com/science/article/pii/S2046043018300339` [12/2018]
17. Nesmachnow, S., Massobrio, R., Cristóbal, A., Tchernykh, A.: Planificación de transporte urbano en ciudades inteligentes. In: I Ibero-American Congress of Smart Cities. pp. 204–218 (2018)
18. Nesmachnow, S., Rossit, D., Toutouh, J.: Comparison of multiobjective evolutionary algorithms for prioritized urban waste collection in Montevideo, Uruguay. Electronic Notes in Discrete Mathematics 69, 93–100 (2018)
19. Pena, D., Tchernykh, A., Nesmachnow, S., Massobrio, R., Feoktistov, A., Bychkov, I., Radchenko, G., Drozdov, A., Garichev, S.: Operating cost and quality of service optimization for multi-vehicle-type timetabling for urban bus systems. Journal of Parallel and Distributed Computing (2018), online first, `https://www.sciencedirect.com/science/article/pii/S0743731518300297` [02/2018]
20. Péres, M., Ruiz, G., Nesmachnow, S., Olivera, C.: Multiobjective evolutionary optimization of traffic flow and pollution in Montevideo, Uruguay. Applied Soft Computing 70, 472–485 (2018)
21. Porteiro, R., Garabedián, G., Nesmachnow, S.: Generation and classification of energy load curves using a distributed MapReduce approach. In: International Supercomputing in México. pp. 1–15 (2019)
22. Risso, C., Nesmachnow, S., Robledo, F.: Metaheuristic approaches for IP/MPLS network design. International Trans. in Operational Research 25(2), 599–625 (2018)
23. Sassi, P., Freiria, J., Paz, P.L., Mendina, M., Draper, M., Usera, G.: Coupled discrete element and finite volume methods for simulating loaded elastic fishnets in interaction with fluid. Computers & Fluids 156, 200–208 (2017)
24. Wilson, E.: An Introduction to Scientific Research. Dover Publications, USA (1956)
25. Yoo, A., Jette, M., Grondona, M.: SLURM: Simple Linux Utility for Resource Management. In: Job Scheduling Strategies for Parallel Processing. pp. 44–60 (2003)