Planning a statewide Grid infrastructure for the São Paulo scientific community

Eduardo Bach

Núcleo de Computação Científica
Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP)
São Paulo, SP - Brazil
<bach@ncc.unesp.br>

Abstract

This paper discusses the initial planning actions needed for the sharing of high-end computing resources across research institutions of the State of São Paulo in a more reasonable way, using the successful example of the GridUNESP project and the OSG software stack, in order to deliver high-performance and high-throughput computing resources for research scientists.
GridUNESP is a long term project that intends to leverage the scientific research at UNESP in the coming years by means of a production-quality Grid infrastructure aimed to meet present and future requirements of UNESP researchers, and could be used as a successful example. The São Paulo Grid will depend upon the establishment of a local consortium, based on US Open Science Grid model. By means of a local Grid Operations Center, it will provide the necessary IT expertise to develop and operate the shared and scalable infrastructure. The GOC will provide operation and system engineering services, monitoring/information services, as well as consulting and training for administrators and end users.

I. Introduction

The state of São Paulo is the major industrial and economic pole of South America. With an area of approximately 250,000 km^2 and over 40 million inhabitants (almost 22% of the population of Brazil), it is the most populous country subdivision in the Western Hemisphere. The state excels not only in population and economic production, but also in the number and quality of its research institutions, with more than 500 universities, colleges, and research centers spread over the state. There are three state universities: USP; UNESP; and UNICAMP. They are among the best R&E institutions in Brazil.
The support for scientific and technological research is provided by FAPESP, the State of São Paulo Research Foundation considered one of the most important institutions for leveraging Brazilian science. Its recent investment on priority projects, such as genomic science, industrial innovation, and digital information technologies, has resulted in a big international visibility for Brazilian science and technology. The Academic Network of São Paulo (ANSP), which is the academic Internet provider in the state, and SciElo, one of the largest projects of scientific digital libraries in the world, are both examples of projects funded by FAPESP.

UNESP, acronym for Universidade Estadual Paulista (São Paulo State University), is the second largest university in Brazil, with distinguished achievements in teaching, research and extension services. UNESP is part of the State of São Paulo public higher education system, which also includes the University of São Paulo (USP) and the State University of Campinas (UNICAMP). A prominent distinction of UNESP is its multi-campus structure, with 33 faculties and institutes on 23 campuses distributed throughout the state of São Paulo.

UNESP has built a distributed computational system that is considered one of the largest Campus Grid initiatives in Latin America, with computing resources widely dispersed on seven campuses throughout the state. It comprises a network of clusters located in seven distinct campuses of the University and meets the demand of a variety of scientific fields. GridUNESP, as the project is known, is a statewide, multipurpose Grid infrastructure that is providing a reliable and high-performing computational infrastructure to around 60 research groups from several areas of science and engineering, allowing them to access state-of-art data processing and storage systems. The whole infrastructure, with computing resources dispersed in seven different cities over the State of São Paulo, includes almost 400 servers, over 200 terabytes of storage space and a dedicated wide area optical network for inter-cluster connection. As the central resource of computational power in the whole university, GridUNESP's goal is to leverage the scientific research at UNESP [1].

The project is based on a hierarchical 2-tiered architecture, with a central cluster located in São Paulo capital and seven secondary clusters at distinct UNESP campuses located on different cities of the state of São Paulo. The central cluster is most appropriate for running tightly coupled parallel applications using MPI and proprietary parallel libraries, whereas the secondary clusters are more suitable for running loosely coupled applications or near embarrassingly parallel computations.

A formal partnership between UNESP and the Open Science Grid Consortium was signed in 2009 [6], which enabled GridUNESP to use the OSG middleware stack to integrate its computational resources and share them with other R&E institutions worldwide. The GridUNESP VO was created in December 2009 as the first Virtual Organization outside U.S. All the eight sites are interconnected through the KyaTera network, a R&E experimental optical network provided by FAPESP in association with Telefónica Brasil, a subsidiary of the Spanish company that provides fixed-line telecommunications services in the state of São Paulo.

II. GridUNESP as a model to be followed

GridUNESP is a multipurpose VO, with almost 60 projects from many diferent fields, from medical biology to weather prediction, and supports more than 180 users. Considering Gratia information from 2010, GridUNESP is the forth larger VO in number of active users (behind cms, usatlas and osgedu).

A considerable number of GridUNESP users have previous experience on HPC resources and applications. GridUNESP support team efforts have been focused on adapting HPC software to a grid friendly workflow. GridUNESP successfully runs applications like Gromacs, Gaussian, GAMESS, and POY with MrBayes parallel (mpi-like) support.

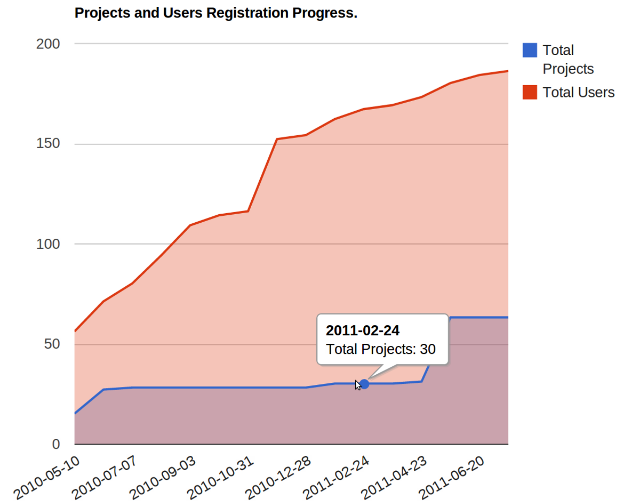

Figure 1: Progress of GridUnesp projects and users

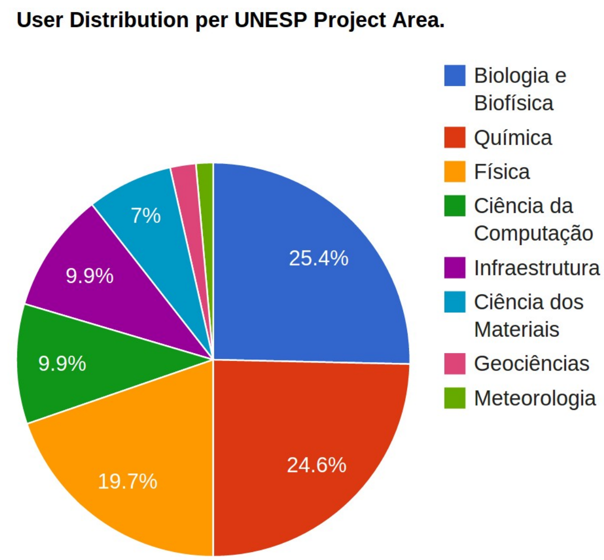

Figure 2: Distribution of users per scientific discipline

The number of researchers interested in using GridUnesp infrastructure has been growing systematically since the start of operations, showing that there is a real demand for high-end computational resources and for qualified technical support.

III. Research institutions of the state

It is a valid assumption to consider the users of the GridUnesp, reference for the users of the other R&E institutions of state of São Paulo, lead us to conclude that there is a real demand for resources capable of running HPC and HTC jobs. A similar situation that occurs in many other regions of the world - the University of Nebraska is an interesting example that comes to mind [2].

Now it is time to evaluate the size of the institutions ´ resources. The universities and research centers of the state of São Paulo have very heterogeneous distributed computational resources. The common baseline seems to be researchers that work on small laboratories and research groups. In almost all institutions there are researchers with small clusters often installed in inadequate conditions. Often those researchers are not properly supported by experienced sysadmins that can handle day-to-day cluster management activities, and usually have to allocate students to these task, resulting in inefficient use of computing resources and a decrease of the resources lifetime. This happens either through lack of maintenance and proper infrastructure conditions, or lack of knowledge about how to maintain or reuse those resources.

In some top-class universities it is also common to find clusters of medium size that belong to large research groups that manage these resources entirely alone or in a partnership with a few other research groups located in other institutions. It is common to find more than a cluster with these characteristics, and in many cases, several ‘isolated’ clusters within the same university. In all situations the resources are heavily used for just short periods of time and then become idle, or, ever worse, resources are insufficient for some members of research groups.

There is a lack of major centers built and sustained by the universities themselves, and we are not witnessing coordination efforts on this direction. GridUnesp has been built with this purpose, to provide high-end computational resources that can be shared by the scientific community. The rapid growing number of users interested in using the resources shows that this is the right way to go. So far, GridUnesp is the only multipurpose distributed infrastructure that uses the OSG stack in the whole country.

IV. A NATURAL STEP FORWARD

Considering the knowledge and practical experience accumulated by the GridUNESP support team, a natural next step is to plan and design a grid infrastructure that could pool computational resources from different institutions of the state of São Paulo. Another goal is to enable users to access heterogeneous resources owned by different organizational units (research groups, laboratories, computer centers, and so on) [4,7]. It is necessary to bring together all interested researchers and stimulate them to work collaboratively. This ‘consortium’ should involve researchers, research institutions and funding agencies.

A. Advantages for everyone

If research institutions begin to offer more computing resources to their users, or even more important, decide to start using a standard software stack, allowing the exchange of more information between system administrators, researchers would benefit immediately, whatever the size of their projects or clusters. They would be able to immediately access more idle computing resources available at other research groups. Moreover, if they would have the opportunity to deliver their resources to the care of experienced sysadmins, they would be able to receive proper support and maintenance for their equipment, leaving them free to work on what they know best: doing research. The university contribution would be to provide hosting facilities for the computing resources on demand for its own researchers and stimulate the exchange of resources and technical knowledge with other research institutions.

The funding agencies, with minor modifications to their fund concession protocols for computational resources, could contribute to the dissemination of this culture in a more rational way, rather than paying for all necessary infrastructures to store the clusters and for the infrastructure management to each research group. The agencies could simply act under the demand of the researchers, e.g. providing more resources to existing large data centers, offering more resources to their researchers for the same amount of money. As in those large centers the reuse of hardware polices are common, computing resources purchased by the development agencies would also have a longer lifetime.

B. Basic services to be delivered

The final users of such a distributed infrastructure need to be properly trained to use it. Those training sessions should be offered on a regular basis and should cover not only basic concepts such as submitting jobs but also advanced ones such as parallelization and porting of applications.

The system administrators should also be trained to build and keep the structure running smoothly. Training activities should have practical sessions (hands-on) and cover installation, administration, monitoring, and security. Technical documentation should be developed and made available through web pages or wikis.

Consulting services on how to build a new site to be part of an existing grid infrastructure should also be ready to use for the members of the consortium. Proper documentation should be available.

Technical support specialized in software should also be provided in such a way that researchers could ask for help on porting their applications to the grid infrastructure.

C. Organizing users

Users should be part of and interact by means of virtual organizations. A virtual organization, or VO, is a collection of people that work collaboratively to achieve a goal by sharing their resources. New users should be assigned to existing VOs, or new VOs should be defined according to the researchers´ needs.

Consulting services on how to organize users on VOs, how to integrate new users with existing VOs, and how and when to establish new VOs should also be provided by the consortium.

Heavy users, those who need access to vast amounts of computing power, should receive special attention. It is important to show them that their needs could be fully satisfied by a shared infrastructure.

D. Organizing operations

To keep a statewide distributed infrastructure built with highly specialized resources shared by hundreds of researchers operating smoothly is not a trivial task. In order to keep such a production infrastructure running, operations services should be offered to the participant sites. A 24/7 support service for the administrators and end users would be desirable, although a 9/5 should work reasonably well in a research environment. A remotely accessible monitoring system that can take care of the whole infrastructure should be built to help administrators´ monitoring activities. An interaction protocol should be established and hardware and software tools (call center, ticket system) should be provided to make it easy for end users and local support teams to interact with the central support group.

Coordination efforts for such an infrastructure are definitely non-trivial. The coordination of support activities should be done by an experienced and well trained technical team installed in a regional Grid Operations Center. The regional GOC should be the coordination point for user and technical support and for training activities.

V. Choosing the sites´ size accordingly

Each university campus that brings together geographically close computing resources such as clusters should be considered a site. Some sites may consist of only one cluster or, as it is the case of some of the large research institutions, may be composed of several clusters located throughout the campus.

For the first situation, a small OSG site should be prefered, for which the middleware stack is less complex. As an example we should consider a storage elements based on Bestman over Hadoop.

For the second situation, each cluster could be configured to use Condor as the local scheduler and accept and do flocking with the other clusters. One of the clusters could be chosen to be the gatekeeper of the site, or one extra server could be installed just to offer this service.

On the beginning of the deployment some institutions or research groups might not be comfortable on sharing its resources with others. For those situations, an experienced consulting team should be available in order to show them the advantages of being part of such an infrastructure. The consulting team should be able to tech concepts like the Campus Grid, mainly based on Condor, and show them that their resources would be fully dedicated for their own use unless they are idle. For clusters that already run dedicated schedulers, such as PBS, a Campus Factory structure could be installed instead, in order to use the maximum number of available resources.

For the submission between sites an infrastructure based on pilot jobs should be deployed, probably using GlideinWMS, the same software being used by OSG. A front-end should be installed at each participant site.

VI. The Open Science Grid Model

The U.S. Open Science Grid model should be applied, and its middleware stack should be used, as it has been successfully applied on the GridUnesp project. According to References [3,5], the Open Science Grid is a consortium of institutions and organizations that operates a large-scale shared distributed infrastructure, which includes computing, storage, networks, software, and institutional support for leveraging science and engineering in U.S. and abroad. Its fundamental goal is to support the scientific discovery process for consortium members and partners who utilize the distributed facility. The OSG

Consortium operates the OSG infrastructure. The consortium itself is not directly engaged in software development efforts; rather, it only tests and integrates new technologies and services according to the consortium needs.

OSG resources are owned and operated by consortium members, which contribute to the common infrastructure, aiming to provide scientists and engineers from many fields with access to shared resources worldwide. As a consequence, OSG computing and storage resources are not all available at any time. In spite of that, it is not an easy task to assign a simple measure for the actual size of the OSG.

VII. Conclusions

The prominent participation and the expertise on grid computing acquired by our group in operating a Tier-2 site of the hierarchical processing structure for the LHC, has generated an important by-product: the deployment of GridUNESP, a grid infrastructure that comprises 7 sites distributed in different campuses of UNESP in the State of São Paulo.

As a natural next step, we intend to promote a partnership for scientific computing involving several universities and research centers of the state, aiming at deploying a statewide grid infrastructure: the São Paulo State Grid. We intend to reproduce a structure similar to the New York State Grid (NYSGrid) that provides computational power for several intuitions of the New York State.

Coordination efforts for dynamically pooling together heterogeneous resources from real research institutions to form virtual organizations are non-trivial. The U.S. Open Science Grid model, based on the establishment of a Consortium that operates de infrastructure, is definitely the right way to go.

The establishment of a Grid Operations Center will be the first step. The GOC will be the coordination point for user and technical support and for training activities such as site installation and administration and end-users´ help on the usage of the shared infrastructure and porting of applications. The technical support team should be focused on end-users' needs, teaching them step by step how to join the Grid community, adapt their scientific applications to the new technology and use it efficiently.

VIII. How my participation in the OSG School helped me in these activities?

The OSG school I attended has provided me with the key design ideas I have written in this work.

The most important aspect related to my attendance was the contact I had with users of other scientific areas. Only in OSG School I would have a chance to see so clearly that most, if not all, the problems faced by users in Brazil are exactly the same of those faced by the rest of the world's users. The selected users in the OSG School reflected quite well the reality of users we have in Brazil. Another important aspect of this interaction with the users is that I was also attending the school as a user, so I learned about the details of users difficulties - not only about the use of distributed computing tools, but also some more basic issues. This experience surely will guide me at the time that we start outlining the learning initiatives.

Another important aspect I have retained during the school was the interaction with the Condor developers. I cannot imagine any better learning context in which one could learn so much from the developer's experience, their perceptions on how the problems must be attacked to reduce the users´ learning curve. We can not forget that at the end the scientific production will be shaped by the ability of end users to make more effective use of these distributed computing tools.

References

[1] R. L. Iope, N. Lemke, G. A. von Winckler, "GridUNESP: a multi-campus Grid infrastructure for Scientific Computing", in Proc. of the 3rd Latin American Conference on High Performance Computing (CLCAR 2010), 25-28 August 2010, Gramado, Brazil, pp. 76-84.

[2] D. Weitzel, "Campus Grids: A Framework to Facilitate Resource Sharing". May 2011.
http://osg-docdb.opensciencegrid.org/cgi-bin/ ShowDocument?docid=1052

[3] P. Avery, "Open science grid: Building and sustaining general cyberinfrastructure using a collaborative approach", First Monday - Peer Reviewed Journal on the Internet, Vol. 12, Number 6, June 4, 2007. Available online at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/issue/view/240

[4] R. L. Iope presentation: "Grid efforts in Brazil: The GridUnesp projet and planning towards a statewide Grid infrastructure in São Paulo". OSG Sites Coordination Meeting, February 10, 2011.
https://twiki.grid.iu.edu/bin/view/SiteCoordination/SitesCoord110210

[5] OSG Document 18-v12: A Blueprint for the Open Science Grid
http://osg-docdb.opensciencegrid.org/cgi-bin/ ShowDocument?docid=18

[6] OSG Document 827-v1: GridUNESP and OSG Agreement
http://osg-docdb.opensciencegrid.org/cgi-bin/ ShowDocument?docid=827

[7] OSG Document 1020-v1: GridUnesp and the São Paulo State Grid initiative
http://osg-docdb.opensciencegrid.org/cgi-bin/ ShowDocument?docid=1020