MAGRID: THE MOROCCAN GRID COMPUTING INITIATIVE

CHAKER EL AMRANI

Faculté des Sciences et Techniques de Tanger, Université Abdelmalek Essaadi,
Tanger, Morocco
http://www.uae.ma/uaegrid/

OTHMANE BOUHALI

Inter-University Institute for High Energies, Univesité Libre de Bruxelles,
Brussels, Belgium
http://www.ulb.be/

REDOUANE MERROUCH

Centre National pour la Recherche Scientifique et Technique,
Rabat, Morocco
http://www.cnrst.ma/

Grid computing has recently emerged as a response to the growing demand for computing resources exhibited by the scientific and researcher communities.

The Grid has witnessed rapid growth in the past few years. Several countries have adopted this technology as a basic computing infrastructure for scientific and research disciplines.

We present a report on Grid activities in Morocco in the last three years as a consequence of an established collaboration between our three institutions.

A national grid initiative, called MaGrid, has been setup which aims at deploying a robust, flexible and sustainable Grid infrastructure.

This initiative proved to be a catalyst for other Moroccan universities and research centers in the country with a growing interest to collaborate in Grid activities. Moreover the involvement of Morocco in the EumedGrid project, in the Mediterranean region, will open new collaboration efforts in terms of the Grid.

Keywords: Grid computing; EumedGrid; MaGrid; UAEGrid.

1. Introduction

Grid Computing; e.g. Foster and Kesselman (1999); has recently emerged as a response to the growing demand for computing resources exhibited by the scientific and researcher communities. It is an innovative approach aiming at integrating and using computing resources available in different places as a single geographically distributed computing systems.

At the core of the grid is the middleware. This latter is a series of software components that enables the interface between the distributed resources on one side and

the applications on the other side. These components comprise resource discovery, authentication and authorization, job scheduling, events and data logging and replication etc.

The Grid provides a wide spectrum of benefits, among them:

- Transparent access to and optimized utilization of resources
- Large (almost infinite) computing and storage capacity
- Enhancement of national and international collaborations.

In order to sign onto the grid, the user first needs to acquire a grid user certificate from a trusted Certification Authority (CA) and apply for a Virtual Organization (VO) membership. A Virtual Organization is a group of individuals and/or institutions that come together to share resources with a common purpose.

Over the last years European, North-American and several Asian countries have setup their own Grid and develop e-Science and e-Business projects to enable resource sharing and collaboration from different parts of the globe. As an example the European Union has taken an active role in encouraging and sponsoring grid projects like the EGEE (Enabling Grids for E-scienceS) project. It started in 2004 as a successor to several successful European projects (in particular the European DataGrid and the LHC Computing Grid). Its main goal is to gather all existing national and regional grid infrastructure in Europe into a common grid infrastructure for scientific research. Built on the pan-European research network GEANT2, it will also provide interoperability with other Grids to establish a world-wide grid infrastructure.

As an attempt to ease the digital divide between the less-resourced countries and the EU, the EGEE collaboration has set up the EumedGrid (Empowering e-Science across the Mediterranean) initiative, e.g. Fig. 1. It aims at building, in the Mediterranean region, a pilot grid infrastructure for research, which will be interoperable and compatible with the EGEE platform as well as with other analogous initiatives in Balkans, North Europe,

Fig. 1.   EumedGrid participating sites.

Latin America and Far-East Asia.

Morocco was one of the first countries to answer actively and positively to this initiative. Indeed there were already several groups interested in this technology in the country.

2. Grid computing activities in morocco

2.1. *Motivations*

There are a quite large number of reasons to start a grid activity in the country:

- Researchers are very often facing complex problems that require intensive data processing
- There is a substantial increase of scientific production in the last ten years. This places Morocco in the 3rd place at the African level
- Involvement in international collaborations. An example is the participation of several Moroccan groups in the LHC project at CERN
- Improvement of network infrastructure through the Moroccan Wide Area Network (Marwan). It consists of providing, deploying and maintaining the necessary infrastructure for universities and research centers. The MARWAN network is connected to the European network GEANT2 through a 155 Mbps connection link
- Existence of expatriate Moroccan scientists, working on the subject, who can give a valuable support.

2.2. *First grid activities*

Grid activities in the country started already in 2003. Our collaborative work led to several proof of concepts achievements; e.g. Bouhali and El Amrani (2005); Bouhali and El Amrani (2006a); among them:

- Prototype of a computing cluster: It consists of one master and two worker nodes running Scientific Linux. The job manager and resource allocation software is based of Condor, a software system that creates High Throughput Computing environment. Moreover a user-friendly graphical interface has been developed to enable job submission and monitoring; e.g. El Mhouti, A.; *et al.* (2005).
- Prototype of a certification authority: the CA is an entity which provides digital user and/or machine certificates to be used in the grid. There should be one CA per country. A prototype model of a CA has been installed, and certificates were issued to users and several machines; e.g. Bendahmane (2005).
- Basic grid computing prototype: Two machines were installed with the basic grid middleware. Using the certificates issued by the above mentioned CA, user jobs could successfully submitted to the test bed machines; e.g. Younes (2005).
  In addition two PhD theses have started in terms of grid.

In order to proactively raise awareness of this important technology, a substantial dissemination campaign took place, nation-wide, through seminars and presentations.

3. The MaGrid initiative

With the above successful achievements it became necessary to pursue this activity and to expand it to a national grid platform. Indeed creating such an infrastructure will bring novel approaches to scientific computing and will improve the conditions for

scientific activities and applications of economic and social interest that involve complex computations or large quantities of data.

Thus a proposal of a national grid platform has been submitted to the director of the CNRST; e.g. Bouhali (2005); Bouhali and El Amrani (2006b). The proposal has been warmly welcomed as it fitted in the government strategy to build a national computing center for researchers.

A national committee has been composed in order to set a roadmap for the installation of a grid computing infrastructure in the country. Its major recommendations are:

- The creation of a national grid management committee
- Designating the CNRST as a natural and a unique CA for the country. CNRST is also responsible for the MARWAN research network in the country
- Launching the grid infrastructure in two phases: the pilot phase and the final phase.

The pilot phase, which is a proof of concept phase, started in 2006 with the vision to create a network of 2 or three connected computing sites. This would form the basis of the future national infrastructure in terms of grid computing. Two sites are already connected: the CNRST and the UAE.

In the CNRST site, which is the central management site, the following machines have been installed:

- User Interface: the machine used by the user to submit and monitor the job and to retrieve its output. This is done through a set of user-level client tools installed on the machine.
- Resource Broker: its task is to dynamically identify the resources on the grid and allocate the necessary resources for a given job.
- Storage Element: is the interface between the grid clients and the storage system. The raw storage capacity is 1TB.
- Computing Element: provides an interface to the local batch queuing system and manages the local worker nodes.
- Worker node: is a local farm computing machine where the jobs are executed. There are ten dual processor worker nodes at the CNRST site.

At the UAE site, UAEGrid platform is made of a computing element, a storage element, a user interface and two worker nodes. For the other machines use is made of the central CNRST machines.

Both sites are running the up-to-date version (3.0) of the EGEE middleware called gLite.

Several applications have been localized as candidate to run on the Grid.

The final phase will see an expansion of the existing platform to include other universities and research centers. To achieve this goal a continuous training policy is foreseen in order to enable users and future administrators to maintain their sites and applications.

Another important issue for the future is to expand the use of Grid beyond the research world. The private sector should be engaged starting by sectors where this technology offers the most potential.

4. MaGrid within EumedGrid

The primary objective of the EumedGrid initiative is to bring the less-experienced countries of the Mediterranean region to the level of the European countries in terms of e-Infrastructure, in particular the grid infrastructure. Funded by the EU within the sixth framework Program for Research and Development, it started in January 2006 for a period of two years. All EumedGrid countries are connected to the pan-European GEANT network thanks to the Eumedconnect project: an initiative to establish an IP-based network that serves the research and education communities in the Mediterranean countries. Morocco, represented by the CNRST, has been involved in both projects since the beginning.

Presently the EumedGrid platform comprises a total of 381 CPUs and 30 TB of storage capacity, e.g. Fig. 2. Several applications have been selected to run on this platform, from different disciplines: High Energy Physics, Biology and Biomedical, Earth Science.

The platform is actually participating to worldwide test of WISDOM again malaria which started since October 2006. WISDOM is an international initiative whose goal is to demonstrate the ability of the grid approach to address drug discoveries for neglected and emergent diseases.

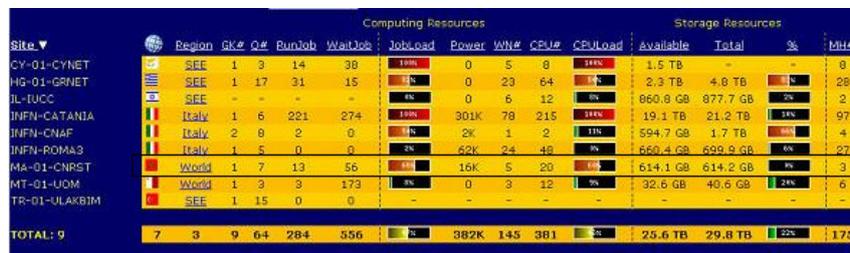Another application for hydrology, called CODESA, has been ported to this application.

Fig. 2.   EumedGrid monitoring showing MaGrid activity.

5. Tutorials and dissemination

Since 2004, several seminars and talks about grid computing have been organised. They mainly aim at presenting this new technology to Moroccan researchers from Universities and scientific Institutes, and showing its fruitful impact on research and development; e.g. El Amrani and Bouhali (2006a); El Amrani and Bouhali (2006b).

Recently, within the context of its collaboration with EumedGrid, MaGrid team has organised a successful workshop in Marrakech on the 4th of December 2006. It was the first large EumedGrid dissemination event held outside Europe.

Many delegates from Universities all around Morocco and from other Partner Countries have attended this workshop, and discussed the new opportunities of use and

cooperation provided by Grid computing technologies and applications in various scientific areas.

This workshop was followed by a user tutorial for beginners and another tutorial dedicated to the new site administrators, as several other Moroccan sites are expected to join shortly the EumedGrid platform.

6. Applications on MaGrid

MaGrid site is operational, and highly solicited by international users running computational and data intensive applications, under the tow Virtual Organisations eumed and biomed, e.g. Fig. 3. For local users, disseminations and seminars are organised, as mentioned above, to help them running their applications and taking benefits from the shared local and Mediterranean grid resources.
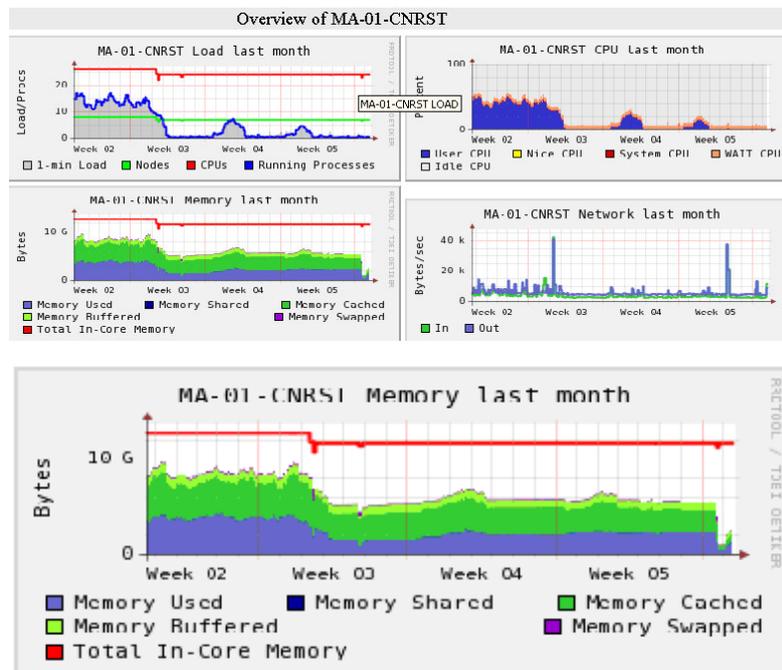

Fig. 3.   MaGrid cluster activities (during December 2006).

7. Future of MaGrid

After the pilot phase, the final phase of MaGrid project will start and will aim at achieving a number of issues:
- Connecting all Moroccan Universities to MaGrid
- Defining the rules for an appropriate use of MaGrid resources

- Being able to use National MaGrid resources as well as those available all over the world
- Reaching a reasonable level of capacity, reliability and ease of use
- Sharing experiences between Scientific Groups involved in the MaGrid concerns, by organizing workshops, seminars, etc.

8. Conclusion

The Grid computing presents a new trend to distributed computing. Its applications have the potential to integrate geographically dispersed compute resources, data repositories, scientific instruments, and human users.

In this paper we present the MaGrid project, funded by CNRST, which is a successful attempt to achieve above goals by exploring various resources on the existing network infrastructure Marwan2. Activities held under this project, mission and future of MaGrid are mainly discussed.

MaGrid is currently used by international community. To encourage Moroccan researchers to exploit this platform, several tutorials and seminars are organised.

UAEGrid is the first University site already connected to MaGrid and currently operational. The next goal, set up by the MaGrid managers, is to help Universities to build up their local grid, and link all of them to MaGrid, to share and use all available resources of the national and the Mediterranean grid infrastructures.

Acknowledgments

References

Bendahmane, A. (2005): Installing a Grid Certification Authority at Abdelmalek Essaadi University: Final Master Project, Master of Communications and Information Systems. Promoters: Bouhali, O. and El Amrani, C.

Bouhali, O. (2005): Proposition d'installation d'une grille de calcul nationale : A technical proposal submitted to CNRST.

Bouhali, O.; El Amrani, C. (2005): Installing and operating a Grid Infrastructure at the Abdelmalek Essaadi University, Proceedings of the Information and Communication Technologies International Symposium, ISBN: 9954-8577-0-2, Tetouan, Morocco, pp. 391–393.

Bouhali, O.; El Amrani, C. (2006): UAE-Grid project, connecting the UAE to the future National Grid Computing facility MaGrid.

Bouhali, O.; El Amrani, C. (2006): MaGrid, Brief Technical proposal.

Foster, I.; Kesselman, C. (1999): *The Grid: Blueprint for new Computing Infrastructure*, 1st edn. Morgan Kaufmann.

El Amrani, C.; Bouhali, O. (2006): Building up a Grid Computing at Abdelmalek Essaadi University, IEEE Morocco Section Workshop on Wireless Telecom and Information Systems, USMBA, Fez, Morocco.

El Amrani, C.; Bouhali, O. (2006): Improving research productivity with grid computing, IEEE Morocco Section E-Week'06, INPT, Rabat, Morocco.

El Mhouti, A.; Alla, J. A.; Stitou, A. (2005): Cluster deployment at the Faculty of Sciences and Techniques of Tangier, Final Bachelor Project, Bachelor of Computer Science. Promoters: Bouhali, O. and El Amrani, C.

Younes, A. (2005): Installing a basic Grid Computing prototype at Abdelmalek Essaadi University. Final Master Project, Master of Communications and Information Systems. Promoters: Bouhali, O. and El Amrani, C.