ITIM Distributed Grid System applied in high energy, biomolecular and nanotehnology physics

F. Farcas, G. Popeneciu, B. Attila, C. Morari,
INCDTIM, Cluj-Napoca, Romania
S. Belov
JINR, Dubna, Russia
L. Miclea,
Technical University Cluj-Napoca, Romania

Abstract

Grid Technology is the next phase in computer infrastructure that will fundamentally change the way we think about and use computing power. The word Grid is used in analogy with the railway infrastructure and electric power grid, which, as the computer has had a dramatic impact on human capabilities and society. In this paper we present the new GRID site of INCDTIM that will offers his facilities to the ATLAS GRID project, and other inland and world wide research. In this paper there are discussed the projects in which we intend to apply our grid site for now.

Supported by the Romanian R&D Agency through CEEX, C program this project combines particle bimolecular and nanotechnology physics and Grid computing for distributed analysis

I. Introduction

The grid computing vision promises to provide the needed platform for a new and more demanding range of applications. For this promise to become true, a number of hurdles, including the design and deployment of adequate resource management and information services, need to be overcome. In this context we would present the structure of grid computing that applies to our need at the INCDTIM institute. This structure is based on many projects that are under development and need a huge computing power. In our institute we are interested to apply the grid system in the following research areas:

a. One of them is the fundamental research in particle physics that uses advanced methods in computing and data analysis. Particle physics is a branch of physics that studies the elementary constituents of matter and radiation, and the interactions between them. It is also called "high energy physics", because many elementary particles do not occur under normal circumstances in nature, but can be created and detected during energetic collisions of other particles, as is done in particle accelerators. [1]

b. During the last years computational simulations based on the atomic description of biological molecules have been resulted in significant advances on the comprehension of biological processes.

For example a molecular system has a great number of conformations due to the great number of degrees of freedom in the rotations around chemical bonds, leading to several local minima in the molecular energy hyper-surface. It has been proposed that proteins, among the great number of possible conformations, express their biological function when their structure is close to the conformation of global minimum energy [2, 3]. This type of research involves a large amount of computing power and fills to be a very suitable application for grid technology.

c. Over the past few years, a sizable body of experimental data on charge transport in nanoscopic structures has been accumulating.

We face the birth of a whole new technological area: the molecule-based and molecule-controlled electronic device research, often termed “molecular electronics” (ME). The simplest molecular electronic device that can be imagined consists of a molecule connected with two metallic nanoelectrodes. There is now a variety of approaches to form such nanojunctions (NJ) [4] that differs in the types of electrode, materials employed, the way in which the molecule-electrode contacts are established and the number of molecules contacted. Recently the realization of the first molecular memory was reported [5].

II. Discussion

A. Atlas and LHcb context

The Large Hadron Collider (LHC) is a particle accelerator and hadron collider located at CERN, near Geneva, in Switzerland. The LHC is scheduled to begin operation in May 2008 The LHC is expected to become the world's largest and highest-energy particle accelerator.[6] At LHC are four experimental programs prepared: ATLAS, ALICE, LHCb and CMS.

Through a contract with IFIN Bucharest our institute is involved in the ATLAS and LHCb project, and one of our tasks is to offer computing power facilities.

The ATLAS program at the LHC is searching for evidence of the existence of the god particle, named the HIGGS bosson. When the particle collide, at the highest energy $\sqrt{s} = 14TeV$ and luminosity $L=10^{34}$ $cm^{-2}s^{-1}$, they’ll produce shower of debris as their energy gets transformed into mass. The physicist won’t see the Higgs itself in the shower. The Higgs particle, probably

1-4244-2577-8/08/$20.00 ©2008 IEEE

100 to 200 times the mass of a proton, is unstable. It will last less then a millionth of a billionth of a billionth of a second before decaying into a spray of other particle. Two to of the four major experiments (ATLAS and CMS) are capable to record the detritus of the disintegrating Higgs. And the assumption is that only few collisions (one among many trillions) will produce a Higgs. The particle will show up in a detector's computers, found by sorting through massive amounts of data measured in petabytes (PB).[7]

The LHCb physicist will run their physics analysis jobs, processing the DST output of the stripping on events with physics-analysis event-tags of interest and processing algorithms to reconstruct the β decay channel being studied. This generates quasi-private data, which are analyzed further to produce the final physics steps.

During the first stage, of the ATLAS project, at a trigger rate of 160 Hz and for an event size equal to 2MB the result / day would be of 20TB, in one year 20 PB. That is for what we need a grid system to store and than to analyze this amount of information

B. Biomolecular context

To make possible the global optimization of a protein, a GRID environment is being developed for the execution of THOR processes through the middleware MYGRID [8]. MYGRID is a user-centric grid middleware that provides a global execution environment which permits the remote execution of a large amount of parallel tasks through the machines that the user has access. Massive calculations for the prediction of protein structure and protein-protein interaction (protein docking [9]) in a GRID may be an important step in order to fill in the huge gap between the knowledge of the aminoacids sequence that compose proteins and the determination of three-dimensional structures through experimental methods, such as X-ray Crystallography and Nuclear Magnetic Resonance (NMR).

C. Nanostructural context

From theoretical point of view, the first studies on molecular junctions have focused on benzene dithiolate adsorbed on gold electrodes [10]. Typically, theoretical simulations are based on Landauer's formula (or its generalization, the Landauer-Buttiker formula) and the theory of non-equilibrium Green function (NEGF) [11]. The electronic structures of the nanojunctions are computed in the frame of the Density Functional Theory (DFT) [11]. Theoretical simulation is expected to play an important role in the development of ME. There are two arguments supporting this affirmation: (i) from the experimental point of view it is difficult to characterize exhaustively the NJ's; therefore, any information (in particular theoretical ones) are of major importance (ii) from fundamental point of view, a clear understanding of the electronic transport at quantum scale shall open the road to an intelligent design of the future nanodevices with practical applications.

III. GRID TECHNOLOGY

A. The concept

Distributed or "grid" computing in general is a special type of parallel computing which relies on complete computers (with CPU, storage, power supply, network interface, etc.) connected to a network (private, public or the Internet) by a conventional network interface. One feature of distributed grids is that they can be formed from computing resources belonging to multiple individuals or organizations (known as multiple administrative domains). This can facilitate commercial transactions, as in utility computing, or make it easier to assemble volunteer computing networks. [12]

The idea about Grid is collaboration, working together, solving problems. As the new Grid technologies come into widespread use, we must shift out social patterns of interaction and our reward structure so that we can realize the potential gains from the nonlinear advancements that collaboration will create. Collaboration can be an almost magical amplifier for human work. The success of the Grid will both enable and depend on this amplification. If we speak about collaboration, we think for the large amount of data acquire in the particle physic domain. For mining al this data, we have to collaborate with each other. The collaboration on the software basis and the architecture one is done by a so called machine: *The middleware*. The middleware is a service need to support a common set of application in a distributed network environment.

Why is collaboration such a fundamental concern? It is our need to ensure that sharing relationship can be initiated among arbitrary parties, accommodating new participants dynamically, across different platforms, languages, and programming environments.

To be helpful this we introduce the key concept of GRID – the Virtual Organization (VO). Without collaboration and interoperability, VO applications and participants are forced to enter into bilateral sharing arrangements, as there is no assurance that the mechanism used between and to collaborators will extend to any others. Without such assurance, dynamic VO formation is all but impossible, and the types of VOs are severely limited. Just as the World Wide Web showed us that we require a universal protocol, and syntax (http and Html), the necessity for a standard protocol and syntax arise in the Grid domain too.[12]

Another key concept in the Grid distribution system is "On demand Computing", which means virtualizing machines, servers and indeed the entire datacenter and being able to focus instead on business needs. The ensuing flexibility will allow an e-business to own its own infrastructure or subscribe to the service of an external service provider, or both. In fact, it can decide to make and buy, choosing to build its own infrastructure for the efficient delivery of mission critical application, while drawing on a service provider for more common business processes such as accounts receivable; for variable infrastructure needs such as storage; for peak load computing power or for advanced capabilities for which it lacks the needed skill. In the near future in on demand computing a company might respond to the increasing need for computing resources first with spillover services and if it's not sufficient with capacity from a service provider, purchasing only the needs for the moment.

The resources that all this system needs is a high speed connectivity, both local and wide area; a huge computer throughput, a huge amount of data storage; human resources to manage local configuration and to implement GRID specific activities.

B. LHC, biomolecular and nanostructure Computing GRID

LCG project was launched in 2002 and it has the goal to make a distributed GRID to be used by ALICE, ATLAS, CMS and LHCb experiments at CERN. Till now we have in Romania 2 sites that are working for the ALICE project and 3 sites that are working for the ATLAS project. The amount of data that CERN has to process is big enough so that only one site in the world could not work for. The computing resource estimated to be necessary are presented below [13],

CPU(MS12k)	2007	2008	2009	2010
CERN Tier0	0.91	4.06	4.06	4.61
CERN Tier1	0.49	2.65	4.18	7.00
All tier-1s	4.07	23.97	42.93	72.03
All tier-2s	3.65	19.94	31.77	53.01
Total	9.12	50.61	82.94	138.45

Where 1 Athlon 3Ghz ~= 1 kSI2k.

Beside LCG project we intend to use our grid site for other inner projects. Both of them are connected with the numerical simulation of computational physics with direct application in biophysics and nanostructures

As a general rule, the computational techniques used by numerical simulations of computational physics are based on the use of Unix-like operating systems. The codes used are mostly developed in Fortran. A vast majority of these codes are developed under GNU license (such as ABINIT) or academic license (SIESTA, GAMESS etc). The commercial ones are represented by GAUSSIAN or MOLPRO.

The systems under investigation may include up to hundreds of atoms. As a consequence, the computing time needed by current applications can easily exceed several weeks or even months. For example, the computation of the current-voltage characteristic of a realistic nanodevice needs about 6-10 GB of memory and one week computing time on the last generation of Opteron processors. Therefore, the development of parallel computational techniques is nowadays in the forefront of development of the methods used by computational physics.

The MPI (Message Passing Interface) has proved to be a very good candidate for this task. The modern developments of the above mentioned codes are using the MPICH2 implementation of MPI in order to reach a good scaling of the computing time with the number of processors.

C. Structure and testbeds

In the frame of our proposed Grid model, a hierarchical structure has been proposed for data processing and storage by introducing regional centers nominated as Tier 0, 1, 2, which will take and process the primary data and will store the necessary information from the projects.

Data coming from the experiments are written on tape at the Tier-0 centers, and a second copy of raw data is provided to the Tier-1 sites. Raw data at the Tier-0 will be reconstructed according to the scheme of the experiment, and the resulting datasets also distributed to Tier-1 sites. The primary Tier-2 site role is the producing and processing of data and end-user analysis.[13]

A functional testbed needs the following components to be full functional: the infrastructure, the middleware and the specific application.

The infrastructure is a tiny, medium or large datacenter.

The middleware is the key component that hold together the whole grid computing infrastructure dedicate to resolve a task.

The specific software resolves the tasks that are launched by each user.

Each task is launched by a user through the user interface (UI) which is the only point where the user has access in the Virtual organization through a valid digital certificate.

If one user is launching a task, the task is passing through the following steps: First it will encounter the resource broker (RB). The RB is reading the information that the task need to be fulfilled and is redirecting the task to the optimum node for execution. A copy of the task is passed to the Central element (CE). The CE redirects the task to the appropriate Work Node (WN). The result and the input file are stored in the Storage Element after the job is finished.

IV. INCDTIM GRID - Site

Since 1994, when the project with CERN begun, in Romania a lot have changed. In Bucharest nowadays are 7 GRID sites, in Iasi there is one and in Timisoara one. In Cluj we have begun the development of a Grid site that would function independent from Bucharest. The first task was to have a stable 1 Gbps link with Romanian Education Network (RoEduNet), successfully obtained in November 2007. The second objective that was finalized was the building of a structure that could host all servers and work nodes, a modern datacenter after the TIER2 standards. The third and final step is the configuration and certification of a GRID site for hosting and helping for calculating the jobs of the first three projects that we have developed in collaboration with our partners.

A. Testbed Itim – Cluj

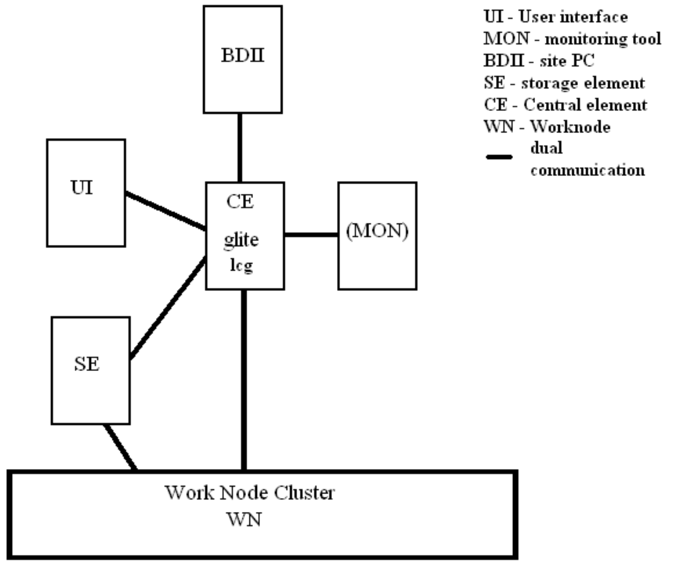
The Cluj Site would be an EGEE type, running GLite middleware and Scientific Linux operating system for the CERN cluster and SunGrid engine 6.1 as batch job system on OpenSuse Linux for the other project.

Our site consists of two clusters and 5 different servers that are controlling the whole site. The server will have one Central Element (CE), a user interface (UI), a database site server (DBII), a storage element (SE) with a capacity of 28 Tbytes and finally a monitoring tool (MON box). Each of them is on different server installed. This is an advantage because, after some studies on the local grid sites, we won't have problems with the incompatibility of software and opening firewall ports on each server.

For security reasons we have separate the DBII server from CE and UI server. The UI is a host to provide running some interactive grid commands for users. Having UI on our site is a good reason not to push our users in Cluj or from other projects to setup personal UI on their own desktops/portables.

Commands running from UI are communicating mainly with RB and other grid services all over the Grid, not especially with services on our site. If we send a request to resources (like SE) or resource broker (RB) it will send the task to the most suitable place in the grid. The VoBox is acting like a RB in a grid site. His job is to redirect the job to the optimum cluster if the site is interconnecting the web. The VoBox is not necessary on all grid sites and so we have decided not to have one on our. If the site will have a necessity for a VoBox we will install one in the future.

The monitoring tool inside the site, MON box supports some monitoring facilities, including R-GMA. CE-s puts accounting data on finished jobs to MON Box, and then MON Box publishes such data to accessible worldwide.



The WN cluster for ATLAS consist now of 14 servers a part are Dual Core dual Processor Intel 3Ghz with 8 GbRam and the rest are Quad Core, dual Processor Intel 2,66 Ghz with 16 GbRam. The Storage Capacity has 12 Tbytes but it will expand in the near future to 28 Tbytes. The WN cluster is dedicated to data processing. Our servers are dual core dual processor machines, installed with SLC 4.5 and GLite. We have installed all WN stations and are now in the research of setting up the best possible way for our site.

V. Conclusions

Our Grid distributed site system is sustaining our research area. On IT basis we would improve in our near future the WN and SE area. We look to collaborate with new partners in international or national programs or projects. That will give us the opportunity to develop our site and specialists in grid system, hoping to become one of the greatest datacenter in Transylvania and sustain the NORD – WEST research and economical area of Romania.

Acknowledgment

This work is supported by the Romanian R&D Agency through CEEX05-D11-58, ET-30, ET-60 and CORINT 3262

References

[1] http://en.wikipedia.org/wiki/Particle_physics

[2]. S. P. Brown, S. W. Muchmore, *J. Chem. Inf. Model.*, **46**, 999 (**2006**).

[3]. R. Bonneau, I. Ruczinski, J. Tsai, D. Baker, *Protein Science*, **11**, 1937 (**2002**)

[4] M.A. Reed, C. Zhou, C.J. Muller, T.P. Burgin, J.M. Tour: Conductance of a Molecular Junction, Science 278(5336), 252-254 (1997); A. Salomon, D. Cahen, S. Lindsay, J.Tomfohr, V.B. Engelkes, C.D. Frisbie: Comparison of Electronic Transport Measurements on Organic Molecules, Adv. Mater. 15(22), 1881-1890 **(2003)**

[5] J.E. Green, J.W. Choi, A. Boukai, Y. Bunimovich, E. Johnston-Halperin, E. DeIonno, Y. Luo, B.A. Sheriff, K. Xu, Y.S. Shin, H.-R. Tseng, J. Fraser Stoddart, J.R. Heath: A 160-kilobit molecular electronic memory patterned at 1011 bits per square centimetre, Nature 445(7126), 414 **(2007).**

[6] http://en.wikipedia.org/wiki/Large_Hadron_Collider

[7] National Geographic March2008, 90-105

[8]. Cirne, W. and Marzullo, K., Open Grid: A User-Centric Approach for Grid Computing. 13th Symposium on Computer Architecture and High Performance Computing (2006).

[9]. C. S. de Magalhães, H. J.C. Barbosa, L. E. Dardenne, *Genetics and Molecular Biology*, **27**, 4, 605-610 (**2004**).

[10] M. Di Ventra, S.T. Pantelides, N.D. Lang; First-Principles Calculation of Transport Properties of a Molecular Device, Phys. Rev. Lett. 84(5), 979-982 (2000); S. Piccinin, A. Selloni, S. Scandolo, R. Car, G. Scoles: Electronic properties of metal–molecule–metal systems at zero bias: A periodic density functional study, J. Chem. Phys. 119(13), 6729-6735 **(2003)**; Y. Xue, M.A. Ratner: End group effect on electrical transport through individual molecules: A microscopic study, Phys. Rev. B 69(8), 085403 **(2004).**

[11] S. Datta, "Quantum Transport: Atom to Tranzistor", Cambridge University Press, **(2005),** 418 p.

[12] The Grid 2, Blueprint for a new Computing Infrastructure, Ian Foster Carl Kesselman

[13] LHC Computing Grid, Technical Design Report, J. Knobloch and L. Robertson