ISSN 1824 – 2049 http://jcom.sissa.it/

Comment

Grid computing and e-science: a view from inside

Stefano Cozzini

Setting the scene: terms and definitions

There are many terms in vogue (buzzwords like e-infrastructure/cyber-infrastructure e-science) which aim at describing the transformation of scientific research enabled by the grid computing paradigm. It is therefore important to clarify and define them in order to frame our discussion.

Firstly, let me define what is grid computing generally means. The term Grid was coined by Ian Foster and Carl Kesselman in the mid 1990's to indicate "*An infrastructure that enables flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions and resources*".[1]

It is important to recognize that resource in this context includes computational systems and data storage and specialized experimental facilities. A typical comparison is made between grid computing and the Power Grid. Within the framework of the grid computing paradigm, computing power should be as ubiquitous and as easily and seamlessly accessible as electricity.

The goal of Grid Computing is therefore "*to provide a service-oriented infrastructure that leverages standardized protocols and services to enable pervasive access to, and coordinated sharing of geographically distributed hardware, software, and information resources*."[2]

One of the key activities of Grid Computing is the standardization of grid protocols in order to facilitate and speed up access to storage systems, data resources, supercomputers, instruments and devices. Grid computing is thus a key computing paradigm that enables the creation and management of Internet-based utility computing infrastructure generally called Cyberinfrastructure and/or e-infrastructure.

The term Cyberinfrastructure was first introduced in the United States in 2003 in what has become known as the Atkins Report,[3] entitled 'Revolutionizing Science and Engineering Through Cyberinfrastructure.' The term e-infrastructure is instead preferred in Europe and I will use it as synonymousto the American term.

An e-infrastructure/cyberinfrastructure can be then defined as an infrastructure of distributed computer, information, and communication technologies which is available to all as a commodity. The development is seen as parallel to the infrastructures that already permeate modern societies: roads and railways for transportation, water, gas, and power networks for basic services and resources. In the words of the Atkins Report, *'If infrastructure is required for an industrial economy, then ... cyberinfrastructure is required for a knowledge economy'*. (Atkins, 2003, p. 5).

The term e-science was first introduced in United Kingdom by John Taylor, then Director General of the Office of Science and Technology in the U.K. In Taylor's words: "*In the future, e-Science will refer to the large scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically, a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large scale computing resources and high performance visualisation back to the individual user scientists*".

Euphoric statements about the transformation of the scientific enterprise quite similar to those permeating cyberinfrastructure discourse marked the launch and subsequent promotion of e-science: "*e-Science is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.*"

It's worth reiterating what according to the Atkins report is the impact of the concepts/structures introduced here:

'The conduct of science and engineering is a social activity, pursued by individuals, collaborations, and formal organizations (...) Advanced cyberinfrastructure offers the potential to conduct new types of research in new ways. Doing this effectively requires holistic attention to mission, organization, processes, and technology. It creates the need to involve social scientists as well as natural scientists and technologists in a joint quest for better ways to conduct research' (Atkins, 2003, p. 14/15)

Where are we on the road toward e-science?

Although is too early to examine the overall impact of the grid technologies and especially the vision it promotes, it is worth indicating the progress made after more then ten years of activities. The aim here is to identify where and when grid computing really enabled or better started to enable the paradigmatic shift in the way research is done envisioned ten years ago.

Many national and international projects around the world are now carrying out research and innovation activities that transform the vision of e-science/ e-infrastructure and Grid computing into reality. The first waves of such initiatives came mainly from the natural sciences, where large volumes of data are involved in research and modern simulation approaches require huge amount of raw computing power. High energy physics, astronomy, meteorology, and computational biology are just a few areas where the new paradigm has been applied with considerable success.

Recently a few other interesting e-science projects have been started to appear in social and economic area as well. E-infrastructure developments offer an opportunity to answer fundamental scientific questions about the complexity of human interactions because of the new potential to collect link, access and analyse data on a scale transcending what is now possible. The approach is still in first phase with many projects predominantly technology oriented, particularly with regard to application of grid architecture.

All these efforts allowed to reach numerous achievements and these have led more and more research institutes, governmental organizations and enterprises to adopt Grid technologies in their activities for three main reasons:

- the cost-effective use of a given amount of computer resources;
- the on-demand availability of an enormous amount of computer power;
- the possibility to collaborate with virtual organizations from different sectors in order to allow a multidisciplinary approach to the solutions of complex problems.

There are however some difficulties and issues for a wider and pervasive adoption of grid computing in scientific research.

From one side the vision depicted in the definitions reported above is still to come at infrastructural level: technical problems limit the usability of the e-infrastructure presently in production: grid technology is still far to offer to a research group a “seamless integration” of its scientific applications and/or computational experiments. From my own experience as both end-user and grid developer the level of detail that needs to be controlled for the successful development and deployment of scientific applications on the grid remains too high.

On the other side it is also quite evident that many scientific communities simply do not know what to do with the amount of computational resources made available by grid. Better said: they are not able to identify or define computational challenges large enough to saturate what is now at their disposal. This is mainly due to their inability to change the way research can be done once large computational resources are made available. It seems that for instance the cooperative and interdisciplinary approach (that will require complex workflows of the computational research) is not taken into account in many communities.

The above considerations can be confirmed by reporting the recent experience within the EGRID[1] project.

A case study: the EGRID project

The EGRID project was conceived upon a call from the national Ministry of Research for projects that addressed the modeling of complex systems. Its original aim was to use the Grid paradigm for research on complex systems in economics and finance. The development intended to produce an Italian national facility for economic and financial data based on Grid technology, and supply the user community with the needed tools to interact with such an infrastructure. This should enable researchers to study problems that could otherwise not be addressed given the computing resources locally at their disposal.

The MIUR contract required to use/collaborate with existing Grid e-infrastructures and thus a collaboration with INFN, contracting partner in the Enabling Grids for e-Science (EGEE) project was established. EGRID therefore implemented and used the EGEE/ gLite infrastructure trying to adopt it to satisfy the requests from our finance user community.

This adoption was quite problematic in several regards, in evident contrast with "the seamless approach" that an e-infrastructure like EGEE should provide. EGRID user community complained about the limited usability of the grid software from the end user point of view, the complex installation and maintenance of a gLite Grid site and the poor documentation.

The developers' team I coordinate on the other side experienced many difficulties in adapting the e-infrastructure to implement data security required by the financial community. As developers we argued that at that it would have been simpler to redo many things without using what was available at that time from EGEE. We recognized that there is a large effort to convert the EGEE e-infrastructure, conceived as a grid to satisfy the Large Hadron Collider (LHC) data needs, into a tool that was of interest to science in general. However, the core technology and the developer was driven by LHC requests and it took a lot of time before needs that were extraneous to them became just listened to. This is another clear indication how far we still are in "general purpose e-infrastructure approach".

On the other end the finance community we addressed was completely unprepared and not really ready to exploit the e-infrastructure proposed: essentially what was needed by them to do research in finance is an information infrastructure and not only computational power.

EGRID advanced in the direction of creating such an information infrastructure for finance in some important aspects, in particular in regard to data management and enhancing usability. However, the problems listed above affect usability and demand solutions at other levels.

A mayor problem was also related to the communication difficulties among different scientific communities: the financial users from one-side and grid providers from the others. The problem of different languages of provider and users was crucial and not simple to solve. "Translators", interface figures, were proposed as one possible solution to avoid these misunderstandings and bridge the language barriers. They would stand between the technical developer and the user communities, be familiar with the working practice of the users as well as with e-Infrastructure. They would need to translate from one side to the other, in disciplinary terms and in terms of the vision that each side has, what it is offering, and what it wants to see on the other side. They would act as mediators, look into the needs of the research community, formalize these needs and have an impact on the developers, as they can express themselves in a language understood by them.

What we experienced in the EGRID project is common to other project as well: there is some agreement in different research communities that computational Grids have not been able to deliver on the promise of better applications and usage scenarios. Although the reasons are often context dependent and resist over-simplified explanations, if there is a single factor that stands out it is probably the complexity associated with Grids – both from a programmatic point of view as well as from a technology, management and deployment. This factor was clearly present in the EGRID project.

There is therefore a need for higher levels of abstractions and provide functionality in a simplified way especially for communities where IT skills and competence are not so high in order to benefit from Grids.

The future: cloud computing?

The need of higher levels of abstraction could be addressed by the concept of "cloud computing", a buzzword that seems to play out the grid computing, at least on specialized media.

Wikipedia provides the following definition for Cloud computing:

"Cloud computing gained prominence in 2007 as a term used to describe computing that is made generally available on a publicly available IP basis (i.e. the Internet) -- "in the cloud". The term derives from the fact that most technology architecture diagrams depict the Internet or IP availability by using a drawing of a cloud. The compute resources being accessed are typically owned and operated by a third-party on a consolidated basis in [Data Center] locations. Consumers of the "cloud" are concerned with services it can perform rather than the underlying technologies used to achieve the requested function."

There is nothing really new in the term, as already pointed out by many experts. What is new is that is that companies like Amazon Google and others aim at playing a mayor role as providers of the "cloud".

Cloud computing is therefore perceived as the next "big thing" in computing. In contrast grid has passed the peak of the new technology hype and is now, even with all the problems discussed above, in production usage for many scientific communities. Due to the commonality between the technologies underlying grid and cloud, a question might be "will cloud computing make grid obsolete?" Yet, a better question should be "how can users benefit from the developments around cloud computing to extend and simplify their grid utilisation." There are on-going discussions on such and other similar questions.

One important issue is also to understand the advantages (if any) to have commercial players as resource providers within existing scientific grid infrastructure. Scientific institutions maybe can consider outsourcing their computing need to commercial providers. Comparing costs between custom construction of data centre like what we are doing at Sissa/eLab and the Amazon retail price list, a gap still exists in favour of our custom solution. This comparison might not, however, be so clear cut in the case of institutes with perhaps less in- house expertise in running large computing infrastructure. The answer to the question "is commercial cloud offering a real competitor to grid resource providers?", is probably "maybe".

There is therefore the opportunity that the next generation e-Infrastructure is as inclusive as possible, federating both resources from academic organisations as well as commercial providers to ensure it is as pervasive, accessible, performing and cost effective as possible.

The benefits of simplicity that cloud technologies propose and envision can create as well opportunities for present grid technologies and infrastructure to better serve their current research communities and hopefully attract new research communities and thus accelerate the transition toward a truly e-science research.

Notes and references

[1] I.Foster, K.Kesselman, and S.Tucke, *The anatomy of the GRID: Enabling scalable virtual organization*, *International Journal of High Computing Applications* **15** (2001) 200. Available at http://dx.doi.org/ 10.1177/109434200101500302.

[2].M. Parashar and C. A.Lee, *Grid Computing: introduction and overview*. Available at http://www.caip.rutgers.edu/TASSL/Papers/proc-ieee-intro-04.pdf.

[3].D.E. Atkins, K.K. Drogemaier, S.I. Feldman, H. Garcia-Molina, M.L.Klein and P.Messina, *Revoluzionasing Science and Engineering through CyberInfrastructure*, www.nsf.gov/cise/sci/reports/atkins.pdf.

[4] www.egrid.it

Author

Stefano Cozzini is a Developing Scientist of CNR/INFM working at National Simulation Center DEMOCRITOS hosted at Sissa (Trieste, Italy). He is presently technical director of SISSA/eLab, a joint SISSA/Democritos laboratory for e-science, which coordinates and maintain all the HPC and Grid resources of the two institutions. He recently coordinated the Italian EGRID project and is now technical coordinator of the EU-IndiaGrid project. His main professional interests are in the field of High Performance Computing and Grid Computing applied to computational physics.

E-mail: cozzini@democritos.it.