October 2008, IST

# Proposal: Project "Multipolar"

Consortium as per Table 1
*Miguel Casquilho* (redactor)
*Ave. Rovisco Pais, IST; 1049-001 Lisboa, Portugal*
Telephone, fax: (+351)21.841 7310, ()21.849 9242

A research and development Project, promoted by IST and INESC, is proposed for the study of a distributed computing system with direct access on the Internet, applicable to engineering calculations useful in industrial informatics. The Project is based on a set of cases existing on the proponent's web page. It is intended that the computing distribution may be made with geographical diversity, including UT Austin. Support to the industry-university interaction is underlying.

Keywords: *distributed computing; advanced computing; industrial informatics.*

## 1. Fundamentals and scope

Among the innumerable studies and works made in university to simulate the industrial processes —either for didactic or research and development purposes, many with reasonable realism—, the current ease of communication suggests the application of the results by the companies and, also, suggests the distribution of the computational load. In fact, that people moves just to do calculations is no longer justifiable, as is not the repetitive installation of software in several computers. Under these assumptions, we propose, jointly by IST and INESC, a research and development Project to establish a distributed computing system, with direct access on the Internet by its users, public or private. The system must be based on a limited and predefined set of machines, from national and foreign institutions, preferably from IST partners, namely the University of Texas at Austin (UTA). In the Project, it is supposed that the computation applications to envisage will be engineering applications, hence the mention to "industrial informatics".

The Project aims at *(a)* the interaction between industry and university, making easy the validation and improvement of the methods and studies of the university, without excluding reverse flows, and *(b)* respecting the underlying legal aspects, the exchange of software among the partners, namely scientific computing libraries. Other utilizations, not only industrial, will of course become possible, once the methodology of communication between the computational poles has been established.

The intention of starting the Project is due to the following entities and persons (addresses given further below):

IST —Prof. Miguel Casquilho (Department of Chemical and Biological Engineering, IST);
INESC — Prof. João Paulo de Carvalho (Department of Electrical and Computers Engineering, IST).
Exide Technologies — Eng. J. C. Carvalho Oliveira (Exide Techn., ex-Tudor [Tudor, 2008]).

The proponents intend to present this Project to the FCT program UT Austin [2008]. The strategy to be followed, execution details and some conclusions are now addressed.

Miguel Casquilho is Assistant Professor in the Department of Chemical and Biological Engineering, Instituto Superior Técnico, Technical University of Lisbon, Portugal. E-mail address: **mcasquilho@ist.utl.pt**

## 2. Strategy

In the field of technology or engineering, calculations nowadays rely on personal computers, sometimes with powerful software. These computers, when isolated, afford ease and independence, but are subject to risks (breakdown, loss) avoidable in a time of easy communications. The idea of networking thus comes up, an idea already consolidated but to which many common users are still aloof, in pite of the ubiquity of the Internet.

As regards software, attention must be paid to the huge heritage of half a century of programs of all the types —also of industrial origin—, which many users erroneously suppose "incompatible" with the Internet, and which certainly, in innumerous instances, have uselessly been rewritten.

One of the proponents (MC), at IST, diverted from the autonomous personal computer paradigm and, since about one decade, adapted to the Internet his former work. He would later verify that such a shift was natural, it being proposed by authors such as Locklin [1998] or Rainer and Orfanogianni [1999]. Otherwise, after having made that adaptation, he would start to lecture directly from the Internet his classes of (statistical) "Quality Control" and "Operations Research" (Management Science).

That proponent concluded, after attaining a certain stability in the construction of numerous applications —mainly didactic, made available on the Internet [Casquilho, 2008a]—, that:

*a)* It is indifferent to run programs (in the Internet) on IST system, the cluster **Sigma**, or on another system, in another geographical localization.

*b)* The access to another system, in this case the **Alfa**, of Faculdade de Ciências, University of Lisbon, permits to use other software, which may not exist on Sigma.

The access to Alfa had as motivation the use of the NAG [2008] library for scientific computing, in Fortran 90. This library existed on the then "Centro de Cálculo" (Computing Center) of IST (now "Centro of Informática", Informatics Center, CIIST) in versions successively updated; however, with the renovation of the hardware at CIIST, that is, when the system Alpha (DEC, Unix) was closed and the Sigma (Linux) was installed, the library stopped working, no means having been found to reacquire a new version.

Recently, initiatives have been launched between the IST and foreign universities: MIT, Carnegie Mellon, E. P. F. Lausanne, and the University of Texas at Austin (UTA). Now, in the collaboration with UTA, advanced computing is explicitly included, a fact that looks adequate to our purposes, with the advantage that UTA possesses the NAG library.

The Project is likely to have a budget of about 40 000 € and a duration of 2 years. The contacts already available are indicated in Table 1 Tabela 1, as well as the Proposal web page.

## 3. Execution of the Project

The former work mentioned, which can be seen on the proponent's web page, shows the feasibility of the proposal. In the terminology used there, each of the many cases presented is called a "plate". Each plate is a web page (normally written in 'php'), a simple interface that is available to the user, who introduces in an HTML *form* his data for that application. It is supposed that —without preventing other

generalizations— any of the cases to solve is the equivalent, on the Internet, of the "classical" computation, reducible to the command line shown in {1} and obvious:

$$\%\ executable\ <\ input\_file\ >\ output\_file \qquad \{1\}$$

The *form* mentioned has as *action* an executable (produced in Fortran or other language), which receives the data (from the standard input) and yields results, which are injected on the final web page, frequently including a graph.

Successively, in their historical evolution, the various plates have been based on 'scripts' (always the same in each evolution stage), made or improved along the years by Informatics students, as follows:

*a)* In 'cgi' (shell).
*b)* In 'php', which solves the issue (of the former case) of the permanence of the (possible) graph in the 'cache', displaying a preexistent graph instead of the correct one.
*c)* In 'php' supported on another in 'Perl' to run a program situated where indicated, a local system or a geographically remote one.

The motivation for the third of these stages came from the convenience to hide the location of the executable from the user. Besides —and directly in the origin of the present proposal—, it was noticed that the making of the graph, in 'gnuplot', would benefit from the more recent version existing at IST, that is: for the resolution of a given problem, two parts were needed, the NAG from Alfa (Fac. Ciências) and the 'gnuplot' of the Sigma (IST). Therefore, the missing link is a further generalization, permitting to launch several remote executions, not only one, giving rise to the production of the dynamic, final page on the Internet.

The basic description made refers to a series calculation, in this case, a calculation and a consequent graph. Once this generalization is begun, others come up, obviously: the distribution of the computation, in a topology of parallel and series, to be defined for each instance. Such a computation is exemplified in Fig. 1, where there would be four systems, as woud be the Sigma, UTA, the Alfa and the system of the Instituto Politécnico of Portalegre, with which there is a privileged relationship. The development of this research with foreign partners of IST, thus, looks to us adequate.

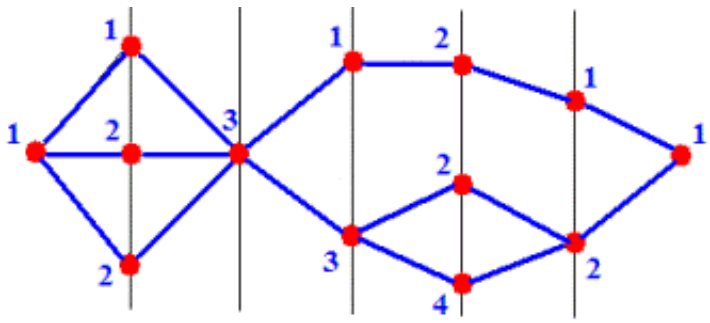


**Fig. 1** — Distributed computation in series and parallel on 4 systems.

## 4. Conclusions

From the work in the university to simulate industrial processes, much is applicable by the enterprises, through the current ease of communication, also with distribution of the computational load. It is thus proposed, by IST and INESC, to do a research Project on distributed computing. Given the recent collaboration of IST with foreign universities, the University of Texas at Austin (UTA) looked appropriate

to us for this initiative, under the support of FCT. UTA possesses the NAG library, which would be important for the calculations.

The Project would have a budget of about 40 000 € and a duration of 2 years.

## Acknowledgements

The researcher (MC) of DEQB/IST works in the «Centro de Processos Químicos» (Center for Chemical Processes). The computing will be based on the cluster Sigma of the CIIST, «Centro de Informática do IST».

## Contacts

**Table 1 Tabela 1** – Contacts of the members involved in the Project and its web page

| | Contact [phone (fax), e-mail] |
|---|---|
| Prof. Miguel Casquilho (DEQB) | (+351)21.841 7310; *mcasquilho@ist.utl.p* |
| Prof. João Paulo of Carvalho (DEEC, INESC) | (+351)21.310 0238; *joao.carvalho@inesc-id.pt* |
| Prof. Keshav Pingali (UTA) | (+001)512-232-6567; *pingali@cs.utexas.edu* |
| Eng.º José C. Carvalho Oliveira (Tudor) | *carvalho.oliveira39@hotmail.com* |

DEQB: Dept. of Chemical and Biological Eng.ing (IST); DEEC: Dept. of Electrical and Computers Eng.ing; INESC: Instituto de Engenharia de Sistemas e Computadores;
"Tudor", presently Exide Technologies, Castanheira do Ribatejo (Vila Franca de Xira, Portugal)

**Web page:** http://web.ist.utl.pt/mcasquilho/compute/utaustin/
Username, password (if requested): multipolar

## References

– CASQUILHO, Miguel, 2008a, "Calculations", http://web.ist.utl.pt/mcasquilho/compute/ (acedido em 2008-09-26).
– CASQUILHO, Miguel, 2008b, "Send by get", http://web.ist.utl.pt/mcasquilho/compute/misc/F-sendget.php (accessed 2008-09-26).
– FCT UT Austin, 2008, «Concursos 2008 para atribuição de financiamento no âmbito do Programa UT Austin Portugal» (Year 2008 funding for the Program UTA Portugal), http://alfa.fct.mctes.pt/apoios/projects/concursos/utaustin/2008/ (accessed 2008-09-26).
– LOCKLIN, Scott, 1998, "Integrating legacy software with a Web front end", in Casquilho [2008b] (accessed 2008-09-26).
– BACHER, Rainer and Tina ORFANOGIANNI, 1999, "WWW based computation services: transfer of power system applications to the WWW", Power Engineering Society 1999 Winter Meeting, IEEE, vol. 1, pp 496–501, 31-Jan to 04-Feb, New York, NY, US.
– Tudor, 2008, presently Exide Technologies, http://www.tudor.pt/ (accessed 2008-09-26).
– UT Austin ("UT Austin Portugal Program"), 2008, http://www.ist.utl.pt/en/html/austin-pt/ (acedido em 2008-09-26).
– UT Austin Advanced Computing, 2008, http://www.utaustinportugal.org/AdvancedComputing/ (accessed 2008-09-26).

❖