

January 2009, IST

Proposal: Project *Multipolar*

Miguel Casquilho (proposer)

Instituto Superior Técnico

Ave. Rovisco Pais, IST; 1049-001 Lisboa, Portugal

Telephone, fax: (+351)21.841 7310, (=)21.849 9242

SUMMARY: A research and development Project, promoted by IST with MIT, is proposed for the study of a distributed computing system with access on the Internet, applicable to engineering calculations useful in industry. The Project is essentially a structure for the interchange of applied research physically based on a computer network. The academic partners wish to link both with industry, and with one another to share software resources.

Key concepts in the Project are: communication, made easy through the Internet; distributed computing and parallelization, an outcome of the previous point; sharing of programs created by the members and directed to industrial use; recovery of existing programs, making them compatible with the Internet; and wider use of scientific libraries.

The computing distribution will be made with geographical diversity, including the academic and enterprise partners. Funding for the Project is 50 000 € and will be sought from FCT. The Project can be based on a set of simple illustrative academic problems solved in the proposer's web site.

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

1. Fundamentals and scope

Innumerable studies and works, many of them realistic, are done in the university to simulate industrial processes, for either didactic or research and development purposes. The current ease of communication permits sharing some programs thus generated with the companies, without the need to move just to do calculations and avoiding the repeated installation of a program in several computers, with distributed computing as a natural extension.

Under the above assumptions, we propose, promoted jointly by IST and MIT and in the context of the MIT Portugal [2009] Program, the research and development Project “Multipolar” to establish a distributed computing system, with access on the Internet. The system, of small size at least initially, should be based on a predefined set of machines, from IST, MIT, and, desirably, a few of their academic or, in a perspective of “industrial informatics”, enterprise partners. The focus is on the production of software (programs) applicable to engineering (or more broadly, technological) calculations useful in industry. The members are supposed to bring contributions in the form of both programs to solve issues believed to be of industrial necessity, and reversely, problems to be solved. The Project is, therefore, essentially a structure for interchange of free or paid research activities.

The Project aims at: (a) linking the university and industry, facilitating the validation of the academic research and results and promoting the technological transfer of these to industry, while fomenting the reverse flow, and (b) using software in common among the academic partners, namely scientific computing libraries, complying with the due legal terms. Other utilizations, not only industrial, may become possible, once the methodology of communication between the computational poles will have been established, in this “multipolar” concept.

The Project is different from many other web sites and surpasses their capabilities, as it is directed to the whole range of calculations (not only small), hence the importance of the presence of the academic partners, and of possible

Miguel Casquilho (Chem. Eng., PhD) 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

parallelization, thus allowing lengthy computations (minutes, hours). Small and medium enterprises (SME, or SMB, businesses) as users may also benefit from this information technology tool, typically above their financial capacity.

This proposal mentions as starters the persons listed in Table 1 (where the proposal’s web address is given). We intend to submit the Project to the FCT—Fundação para a Ciência e Tecnologia (Portuguese “Foundation for Science and Technology”) for funding. The foreseen budget is 50 000 € for a duration of 2 years.

In the next sections, we describe the strategy to be followed, and the contents of the Project, and we present some conclusions.

Table 1 – Members involved in the Project proposal, and its web page^a

Member ^b	Contact
Prof. Joaquim Sampaio Cabral ^c (DEQB)	(+351) 21.841 9063; joaquim.cabral@ist.utl.pt
Prof. Miguel Casquilho (DEQB)	(+351) 21.841 7310; mcasquilho@ist.utl.pt
Prof. João Paulo de Carvalho (DEEC, INESC)	(+351) 21.310 0238; joao.carvalho@inesc-id.pt
Prof. ... (MIT)	(+1-617) 254-...; x@mit.edu
Prof. João Luís de Miranda (ESTG)	(+351) 245.300 284; jlmiranda@estgp.pt
Prof. João Miranda Reis ^d (DEQB)	(+351) 96.8563636; jmreis@newext.biz
Eng.º José C. Carvalho Oliveira (Tudor)	(+351) 263 200 800 carvalho.oliveira39@hotmail.com
...	...

^a <http://web.ist.utl.pt/mcasquilho/compute/MIT/>, where this text can be obtained.

Username, password (same), if requested: *Multipolar* ^b Abbreviations, acronyms: at the end.

^c Head of Department

^d Invited professor, with industry relations, advisor.

2. Strategy

In this Project, the following are the strategic topics: communication; distributed computing and parallelization; sharing; recovery of existing programs; and use of scientific computing libraries.

Communication In engineering (or technology), calculations now frequently rely on small, relatively powerful computers, affording ease and independence. If isolated, however, they are subject to risks and do not fully benefit from the current effective communications. The initial step is that the Project members of each academic (only) partner have accounts in the other’s systems. This is offered free by IST to registered researchers from academic partners.

Distributed computing and parallelization This feature is consolidated in the Internet era, but may be further explored. After communication is established between all the partners (at least IST and MIT), any combination of distribution of computing load may be made, including parallelization, once intensive computing is expected in realistic applications.

Sharing Some software created in or assigned to the Project will be shared for utilization by members, namely the industrial ones. If adequate, these non-academic members may become also providers (reverse flow).

Recovery of existing programs Attention must be paid to a quantity of half a century of existing programs, namely from industry, wrongly supposed “incompatible” with the Internet, and often uselessly rewritten. Also, making them compatible can make them more usable.

Scientific libraries The development of programs in technology benefits from the use of commercial scientific libraries (or “packages”) of quality, such as NAG, available at MIT, or IMSL. Although much free software can be found, the more delicate calculations cannot depend on it (usually without “any warranty”).

We (proposer) departed from the personal computer paradigm and since long have adapted our former work to the Internet. We later attested that such a shift should be natural (as proposed, e.g., by Locklin [1998] or Rainer and Orfanogianni

[1999]). We have since long lectured directly from the Internet our classes of statistical "Quality Control" and "Operations Research" (Management Science).

With programs of ours, mainly didactic, made available for execution on the Internet [Casquilho, 2009], we have made simple experiments with two geographically remote systems in Lisbon: Sigma, at IST (Opteron cluster, Linux) and Alfa, at FCUL* (Digital Alpha cluster, Unix), where an old (*ca.* 1996) version of NAG [2008] exists. *This experimentation steers the current Project.* As expected, a user addressing a web page on the Sigma has no means to notice any difference in running an executable on the local system (Sigma) or on the remote (Alfa).

The above basic experimentation showed how the small scripting effort involved yielded the important benefit of accessing software elsewhere. The reason to use the Alfa was to access its NAG library, which used to exist at CIIST, but could not be reacquired upon the renovation from the old Alpha (DEC, Unix) hardware to the new Sigma (Linux). In the recent months, similar but lesser initiatives have been unfruitfully tried to establish connections with other institutions.

3. Contents of the Project

As an illustration of some basic features of the future Project, our mentioned work on the Internet points to the usefulness of the Project. Each case (or "plate") shown there, a simple web page interface (written in 'php'), is available to anyone introducing his own data in the underlying HTML *form*. Every plate is the equivalent, on the Internet, of the "classical" computation, reducible to the obvious command line shown in {1} ('\$' standing for the system prompt):

$$\$ executable < input_file > output_file \quad \{1\}$$

The *form* has as its *action* an *executable* file (typically from Fortran 90) that receives the data (via standard input), runs the program, and sends the results (via standard output) to the final web page, frequently including a graph.

Technically and historically, the plates have been based on 'scripts' improved by IST Computer Science students: *i*) in 'cgi' (shell); *ii*) in 'php', which solved the issue of the permanence of the (possible) graph in the 'cache', displaying an old one instead; and *iii*) in 'php' linked to a 'Perl' script to run a (local or remote) program wherever selected. The last stage finally managed to hide the location of the executable from the user.

As all of our web examples (plates) refer to series computing (calculation possibly followed by a graph), an obvious generalization is distributed computing, in a certain parallel-series topology for each program. As exemplified in Fig. 1, there might be four systems (such as Sigma, MIT, ESTG[†] and a company).

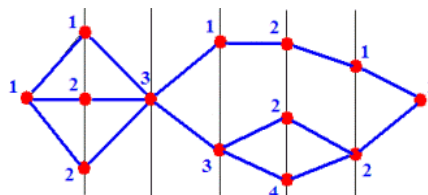


Fig. 1 — Example of distributed computation in parallel-series on 4 systems.

The Project attempts to provide the various missing links, namely, as enumerated: communication, to build the programs using several packages collaboratively; distributed computing with possible parallelization for the extensive runs; and sharing of the utilization of the programs among the partners.

* FCUL: Faculdade de Ciências (Faculty of Sciences), University of Lisbon

† Escola Superior de Engenharia e Gestão, Polytechnic Institute of Portalegre (see abbreviations).

4. Conclusions

From the academic work to simulate industrial processes, much is applicable by the enterprises, given the current ease of communication, also with distribution of the computational work load. It is thus proposed by IST to do a research Project on distributed computing with MIT, with which there is a recent institutional agreement, an organization that we find fully adequate for this initiative, and under the financial support of FCT. MIT has several means, namely scientific libraries, as NAG, which are crucial for applicable computations. The Project is launched for a small number of partners, but due to its nature can be enlarged according to the founding partners.

The Project would have a budget of 50 000 € and a duration of 2 years.

Acknowledgements

The proposer, from DEQB/IST, does research in the «Centro de Processos Químicos» (Center for Chemical Processes). The computing mentioned has been based on the cluster Sigma of the CIIST, «Centro de Informática do IST» (Computing Center of IST).

Abbreviations and acronyms

CIIST	Centro de Informática do IST [Computing (Informatics) Center of IST]
DEQB	Dept. of Chemical and Biological Engineering (IST)
DEEC	Dept. of Electrical and Computers Engineering (IST)
INESC	Instituto de Engenharia de Sistemas e Computadores Rua Alves Redol, 9 — 1000-029 Lisboa; Portugal
ESTG	Escola Superior de Tecnologia e Gestão, of IPP (Polytechnic Institute of Portalegre, Portugal) Lugar da Abadessa — 1000-029 Lisboa; Portugal
“Tudor”	Exide Technologies (formerly, Tudor [2009]; near Vila Franca de Xira) Ave. Dr. Carlos Leal — 2600-619 Castanheira do Ribatejo; Portugal

References

- BACHER, Rainer and Tina ORFANOIANNI, 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.
- CASQUILHO, Miguel, 2009, “Calculations”, <http://web.ist.utl.pt/mcasquilho/compute/> (accessed 2009-01-15).
- FCT, 2009, «Projectos de I&D» (R&D Projects), <http://alfa.fct.mctes.pt/apoios/projectos/> (accessed 2009-01-15).
- LOCKLIN, Scott, 1998, "Integrating legacy software with a Web front end", in Casquilho [2008b] (accessed 2008-09-26).
- MIT / Portugal, 2009, <http://www.mitportugal.org/> (accessed 2009-01-15).
- Tudor, 2009, now Exide Technologies, <http://www.tudor.pt/> (accessed 2009-01-15).

