



## JORNADAS DE SEGUIMIENTO

# PROYECTOS EN TECNOLOGÍAS DE LA INFORMACIÓN

### DESCRIPCIÓN DE RESULTADOS

**Referencia del proyecto:** TIC1999-0754-C03

**Título:** ENTORNOS GEOGRÁFICAMENTE DISTRIBUIDOS

**Investigador principal:** José María Díaz Cort

**Dirección de contacto:** Departament de Llenguatges i Sistemes Informàtics.  
Universitat Politècnica de Catalunya. Edificio C5, Campus Nord. C/ Jordi Girona 1-3. E-08034 Barcelona.

**Datos sobre el grupo investigador:**

¿Se trata de un proyecto coordinado? SI

**Referencia del proyecto:** TIC1999-0754-C03-02

**Investigador principal:** José María Díaz Cort

**Dirección de contacto:** Departament de Llenguatges i Sistemes Informàtics.  
Universitat Politècnica de Catalunya. Edificio C5, Campus Nord. C/ Jordi Girona 1-3. E-08034 Barcelona.

**Referencia del proyecto:** TIC1999-0754-C03-01

**Investigador principal:** Casiano Rodríguez León

**Dirección de contacto:** Departamento de Estadística, Investigación Operativa y Computación. Universidad de La Laguna. Edificio de Física y Matemáticas. Calle Astrofísico Fco. Sanchez s/n. E-38271 La Laguna. Tenerife.

**Referencia del proyecto:** TIC1999-0754-C03-03

**Investigador principal:** Manuel Díaz Rodríguez

**Dirección de contacto:** Departamento de Lenguajes y Ciencias de la Computación. Universidad de Málaga. Complejo Politécnico Campus de Teatinos. E-29079 Málaga.

## **1. PROJECT OBJECTIVES**

The goal of this project is to design, to implement and to evaluate a library of algorithmic skeletons for solving combinatorial optimisation problems. A skeleton is a generic tool that allows to define a concrete optimisation problem by creating instances of the general optimisation method that it implements. Three methods for solving problems of this area will be supported: exact methods, heuristic methods and hybrid methods. The library will offer three different implementations: sequential, locally distributed (LAN) and geographically distributed (WAN). The geographic platform that will be used to develop and evaluate this library will include the three heterogeneous networks of personal computers located at the three participant sites: Barcelona, La Laguna (Tenerife Island) and Málaga.

The library will offer to the users a unique interface for every algorithmic skeleton, independently from the platform. The final user will just chose one of the three methods of resolution and he will particularize the skeleton with the features that define his problem, but without taking care of those aspects related to the method itself or to platform-dependent implementation (e.g, users do not need to have any knowledge about parallelism, but they will obtain all the benefits of distributed computation — LAN and WAN — just from their sequential instantiations of skeletons).

The WAN computation will allow to solve problems in a geographically distributed environment. That represents an important computation power because computers that are in other sites (not only in our own local network) can be used to compute part of a problem. This remote computers could be accessed because of the platform provided by Internet.

The different resolution methods that we plan to cover in this project are divided into the three participant sites: the research team from La Laguna works on the Exact methods, the team from Málaga works on Hybrid methods and the research team from Barcelona works on Heuristic methods. The initial schedule for the three years of the project was done according to the different implementations of the library: the sequential implementation at the first year, the locally distributed (LAN) implementation at the second year, and the geographically distributed (WAN) implementation is planned for the third year.

## **2. LEVEL OF SUCCESS**

We are now at the end of the second year of the project. The sequential implementation of the library was done last year. This sequential implementation became the version 1.0 of the MALLBA library after the coordination method that took place on February of this year. This version is not public yet, but some research groups from some foreign Universities have shown to be interested on it. Some of them are already using this first version of the MALLBA library for their research interests.

During this year we have developed most of the LAN implementation of the library. Many measures of the network were needed before starting the development of this locally distributed part. These measures helped us to know what was the behaviour of the networks and we could implement some of the LAN implementations of our skeletons in a more intelligent manner. Some measures of the WAN environment have been also done. These measures allow to obtain a very useful knowledge about how does the communication over Internet behaves and what are the real levels of traffic congestion in it. Depending on these results, a good equilibrium between computation and communication has to be reached when designing the WAN implementation of the skeletons.

### **2.1. Internet site**

At the first year the website of the project was also designed. The URL of the website is <http://www.lsi.upc.es/~mallba>. The website is a place for announcing the MALLBA project all

over the world, but also a place where all the members of the project can exchange useful information, merge the work we develop separately and discuss about it. This later part is private and just the members of the project can access it. Between one and two people per day have visited the MALLBA website this year.

## **2.2. External Users**

The version 1.0 of the MALLBA library is not public yet, but some research groups from some foreign Universities have shown to be interested on it. They knew about the library because of the Internet site but also because of the research articles we have presented on conferences and invited talks. The aim is to obtain some feedback from them in order to check the real usefulness of the library. Before getting in touch with these customers we were the only users of the library and that is not good for a real evaluation of the product.

Some of the people actually using the MALLBA library are: Mr. Eddy Parkinson (City University, London, UK), Mr. Rong Qu (University of Nottingham, UK), Mr. Likaj Aida (Tirana, Albania), Mr. Sbihi Abdelkader (University of Paris, France), Mrs. Sidi Ykhlef Soraya (University of Baghdad, Algeria), Mr. YoonHo-Joung (Seoul, Republic of Korea), Mr. Kai-Hui Lee (Taiwan University of Science and Technology, Taipei, Taiwan), Mr. Marcos Roberto Silva (Universidade de Sao Paulo, Sao Paulo, Brazil), Mr. Pablo Moscato (Universidade Estadual de Campinas, Campinas, Brazil).

Customers have to fill up a license agreement before getting any code of the library. The purpose of this license is to enable the use of MALLBA code for academic, research, development and demonstration purposes only. It is not permitted the use of MALLBA for any commercial or military purpose. Other constraints for the right use of the MALLBA code are specified in this license, such as the responsibility of using MALLBA library personally and the obligation of referencing it where it is used. The license form can be found in the web page of the project. We are waiting now for six more possible customers to send us back the license agreement.

## **3. RESULTS**

Following we describe some other results of the MALLBA project related to publications, teaching and educational aspects, collaborations with other research groups, etc.

### **3.1. Research articles, technical reports and book chapters**

During these two years of research in the MALLBA project, a total amount of 61 research articles, 11 technical reports, 1 PhD and 2 book chapters have been published. The detailed bibliographic information of all these publications can be found at the appendix attached to this document.

### **3.2. Graduation essays**

Some undergraduate students from the Computer Science Faculties of the three participant sites have been interested in the MALLBA project. These students have been able to develop their graduation essays in the context of the project, either in the algorithmic part (e.g. by developing some part of a skeleton) or at the network part (e.g. by developing tools to monitor the network platform). A total amount of 9 graduation essays have been developed during this two years of research in the MALLBA project. These graduation essays are explained in more detail at the appendix attached to this document. Some memories of these graduation essays can be obtained at the website of the project.

### 3.3. PhD courses

PhD courses in the context of the MALLBA project have been organized in the PhD programs of our respective Universities. These courses have been organized from a theoretical but also from a practical point of view. They have been very useful specially to establish an initial contact with scientific users of different areas. After using the version 1.0 of the library to solve some well known problems, we discussed about some practical features that are difficult to value if the product is only tested by the same people that develops it, for instance how difficult is to use the library, how good is the documentation, etc. The work developed at these courses can be obtained at the website of the project, but just by the members of the project.

### 3.4. Seminars, talks and internal meetings

#### Meetings

Internal meetings have been organized during all the development period. Most of these meetings are internal and there the members of the different groups discuss about the work that is being developed by their site. These internal meetings have been also used to teach ourselves on some technical and new areas. Usually internal meetings are organized every two months.

Two coordination meetings are also organized every year. At these coordination meetings some members of each site participating in the project meet at any of the sites and discuss about the project. Mainly these meetings allow to present the work developed, to merge it, to discuss about next things to do, to study new alternatives for the development, etc.

#### Invited talks

People inside and outside the MALLBA project have been invited to give talks at our respective departments. The topics covered at these talks are considered to be interesting from the point of view of the MALLBA project. Some of these talks were: *Computational Biology* (by X. Messeguer, Computational Biology and Bioinformatics team at the UPC), *BSP-style computation: a semantic approach* (by A. Stewart, University of Belfast), *Entendiendo la Web* (by Ricardo Baeza-Yates, University of Chile), *On the performance of three classic approximation algorithms for the Steiner problem on graphs* (by Pere Guitart, Combinatorics and Digital Communication team at the UAB).

#### Seminar on Object Oriented Programming with C++

Theoretical and practical training on C++ language was done for some members of the project.

#### Seminar on Fringe analysis of synchronized parallel insertion algorithms on 2—3 trees

At the Department of Computer Science, The Queen's University of Belfast (Ireland).

#### Seminar on BSP Library

The focus of this seminar was on how to develop *Middleware for the BSP Library*. BSP is a communication standard library, that can be adapted to a great number of platforms for SPMD (Single-Program Multiple-Data) parallelism type. The BSP Library is based on the BSP (Bulk Synchronous Parallel) model.

#### Seminar on Wide Area Computation

Requirements and challenges of computation on wide area networks were exposed and discussed in this seminar. All the work done by Luca Cardelli and other people of the area was studied in order to know what is the current "state-of-art" at these topics.

### Seminar on informatics Unix tools for research development

Topics covered: Unix operating system, LaTeX, HTML, SPSS, MATLAB, Mathematica, Gaussian, C programming language, FORTRAN 90 programming language.  
Online information on <http://nereida.deioc.ull.es/~pcgull/ihiu01/>

### Seminar on Last advances in Computer Science 2001: Internet and Data Bases

Topics covered: Systematic development of WEB tools, use of ontologies to access to heterogeneous data deposits distributed over the web, federated data bases, XML data bases and information systems distributed via Web.  
Online information on [http://nereida.deioc.ull.es/~paco/uai2001/uai\\_2001.html](http://nereida.deioc.ull.es/~paco/uai2001/uai_2001.html)

### Seminar on Last advances in Computer Science: High-Performance Computing

Topics covered: Programming for shared memory architectures, international efforts in global computing (the Grid), basic support for Web based cooperation in distributed heterogeneous environments, efficient scheduling for cluster computing, scalable cluster Fortran.  
Online information on [http://nereida.deioc.ull.es/~paco/uai2001/uai\\_2001.html](http://nereida.deioc.ull.es/~paco/uai2001/uai_2001.html)

## **3.5. Research stays**

### UNSL (Argentina) and University of Valenciennes (France)

Research on embedded parallelism, network performance prediction, Divide & Conquer, Branch & Bound and Dynamic Programming methods.

EU's transnational access to research infrastructures programme: At the Edinburgh Parallel Computing Center.

### Centralized and decentralized parallel algorithms

Dr. Sami Khuri (University of San José State, USA) stayed at the Universidad de Málaga for 3 months to collaborate on the resolution of combinatorial optimisation problems using centralized and decentralized parallel algorithms, with possible extensions for Hybrid methods.

Técnicas de programación con Java para Internet / Intranet: At the Institut Català de Tecnologia.

Foundations on Wide Area Network Programming: At the 13<sup>th</sup> International Lipari Summer School.

Extreme Scientific Computing: In Manchester.

### Métodos formales para el desarrollo y verificación de programas BSP

The project *Métodos formales para el desarrollo y verificación de programas BSP* belongs to the program *Acciones Integradas entre España y Reino Unido*. Participants of this project are Alan Stewart and Maurice Clint, from the School of Computer Science at the Queen's University of Belfast, and Joaquim Gabarró and María José Serna, from the Departament de Llenguatges i Sistemes Informàtics at the Universitat Politècnica de Catalunya.

1<sup>st</sup> MALLBA coordination meeting: At the Universidad de Málaga.

2<sup>nd</sup> MALLBA coordination meeting: At the Universidad de La Laguna.

3<sup>rd</sup> MALLBA coordination meeting: At the Universitat Politècnica de Catalunya.

Research on BSP semantics: At the Department of Computer Science. The Queen's University of Belfast (Ireland).

### 3.6. Relations with other projects and companies

#### ALGGEN

The Algorithmic and Genetics Group of the LSI Department (UPC) is dedicated to research and teaching in the area of Computational Biology and Bioinformatics. The research projects are developed jointly in collaboration with "in lab" research groups in genetics of other centers. The aim is the design of software tools that allow them to work "in silico" and also the development of software tools to analyse the genome "in silico" in collaboration with "in lab" research groups in genetics.

The Internet site (<http://www.lsi.upc.es/~algggen>) is the virtual laboratory and contains the tools that have been designed both for teach and research purposes. The on-line Bioinformatics tools provided at the Internet site run over the machines of the MALLBA cluster and they allow to solve any problem introduced at the web page. Until now ALGGEN have collaborative projects with:

- Project *ESSEM*: Dr. Ramon Messeguer (Merck-Farma y Química, Barcelona)
- Project *PROMO*: Dr. Mar Alba (Bioinformatics Group of the Wohl Virion Center, UCL London), Dr. Roderic Guigó (Genome Informatics Group of IMIM and UPF, Barcelona), Dr. Julio Rozas (Department de Genètica of UB, Barcelona)
- Project *TRANSP0*: Dr. Josep Casacuberta (Dept. de Genètica Molecular of CID, CSIC Barcelona), Dr. Joan Rigau (Dept. de Genètica Molecular of CID, CSIC Barcelona)
- Project *PRIM*: Dr. Victor Moreno (Institut de Recerca Oncològica), Dr. Benjami Piña (Dept. de Genètica Molecular of CID, CSIC Barcelona)

#### GRAFCAN

Cartográfica de Canarias (<http://www.grafcan.rcanaria.es>). We are working on the problems that they exposed to the Consejería de Medio Ambiente of the Canary government about the study and protection of local species. The Islands are divided into a grid of 500 m<sup>2</sup> squares and the richness and expenses of each square are well known by biologists. An assignment is wanted that maximizes the number of protected species constrained to the money available.

#### ICID

Instituto Canario de Investigación y Desarrollo (<http://www.rcanaria.es/icid>). We have collaborated together to measure network features among the canary islands (see [http://nereida.deioc.ull.es/~cicyt/analisis\\_de\\_rendimiento](http://nereida.deioc.ull.es/~cicyt/analisis_de_rendimiento)).

#### IAC

Instituto Astrofísico de Canarias (<http://www.iac.es>). The tool developed for the two polarimeters Stokes of the solar telescope from the observatories of Tenerife and La Palma is being adapted, in order to allow the computation of every point in the beowulf associated to this project.

#### TECNATOM, S.A.

This company wants to use the MALLBA library to solve their combinatorial optimisation problems. They have helped to measure the parameters of the WAN connection with the Universidad de Málaga, in order to obtain statistics of this area of the network.

#### AUREN

This company is interested in using the Genetic Algorithms developed in the MALLBA project to solve vehicle routing problems. The aim of the work performed is to decide the routes that vehicles must follow taking into account some load and time constraints. This kind of problems can be modelled and solved in the context of the MALLBA project. This work has been described as a PROFIT proposal.