

*Project-Team ScAlApplix**High Performance Schemes and Algorithms
for Complex Scientific Applications**Futurs*

THEME 4B

Activity
Report

2003

Table of contents

1. Team	1
2. Overall Objectives	2
3. Scientific Foundations	2
3.1. Numerical schemes for fluid mechanics	2
3.2. Schemes and algorithms for computational chemistry	3
3.3. Algorithms and high-performance solvers	4
3.3.1. High-performance direct solvers for distributed clusters	4
3.3.2. Fast Multipole Methods	5
3.4. Parallel algorithms for heterogeneous platforms	5
3.5. Computational steering for distributed numerical simulations	5
4. Application Domains	6
4.1. Introduction	6
4.2. Fluid mechanics	6
4.3. Molecular chemistry	7
4.4. Population Dynamics	7
5. Software	8
5.1. Introduction	8
5.2. MUMPS	8
5.3. PaStiX	8
5.4. Scotch	9
5.5. QC++	9
5.6. FluidBox	10
5.7. EPSILON	10
6. New Results	10
6.1. Numerical schemes and algorithms for fluid mechanics.	10
6.1.1. Compressible multiphase flows	10
6.1.2. Residual distribution schemes	11
6.1.3. Hamilton Jacobi equations	11
6.1.4. Porous medium flows	11
6.2. Schemes and algorithms for computational chemistry	11
6.2.1. Divide & Conquer method	11
6.2.2. Crack simulation	12
6.3. Numerical schemes and algorithmic for population dynamics	12
6.4. Algorithms and high-performance solvers	13
6.4.1. Domain decomposition and sparse matrix reordering	13
6.4.2. High-performance direct solvers on multi-platforms	13
6.4.3. High-performance hybrid direct-iterative solvers for large sparse systems	14
6.4.4. ILU factorization based on a hierarchical interface decomposition algorithm	14
6.4.5. Parallel fast multipole method	15
6.5. Parallel algorithms for heterogeneous platforms	15
6.6. Computational steering environment for distributed numerical simulations	16
7. Contracts and Grants with Industry	17
7.1. CEA contract for scientific expertise in high performance computing	17
7.2. CEA research and development contracts	17
8. Other Grants and Activities	17
8.1. Regional initiatives	17
8.1.1. “Réseau Dirigeable”	17

8.2.	National initiatives	18
8.2.1.	RedGrid: “redistribution of large data on computational grids”	18
8.2.2.	EPSN: “a computational steering environment for distributed numerical simulations”	18
8.2.3.	TLSE: “test for large systems of equations”	18
8.2.4.	“GRID5000”	18
8.3.	International initiatives	19
8.3.1.	NSF - INRIA project: “robust parallel preconditioning methods”	19
9.	Dissemination	19
9.1.	Participation to the Scientific Community	19
9.2.	Teaching	19
10.	Bibliography	19

1. Team

SCALAPPLIX is a joint project of INRIA Futurs, LaBRI (Laboratoire Bordelais de Recherche en Informatique – CNRS UMR 5800, University of Bordeaux 1 and ENSEIRB) and MAB (Laboratoire de Mathématiques Appliquées de Bordeaux – CNRS UMR 5466, Universities of Bordeaux 1 and Bordeaux 2).

SCALAPPLIX has been created on the first of november 2002 (<http://www.labri.fr/scalapplix>).

Head of project-team

Jean Roman [Professor, ENSEIRB¹, LaBRI]

Staff members from INRIA

Josy Baron [Administrative assistant]

Christophe Berthon [Research scientist (on partial secondment) since september 2002, Assistant professor at University of Bordeaux 1, MAB]

Olivier Coulaud [Research director]

Pascal Hénon [Research scientist since december 2002]

François Pellegrini [Research scientist (on secondment) since september 2002, Assistant professor at ENSEIRB, LaBRI]

Pierre Ramet [Research scientist (on partial secondment) since september 2003, Assistant professor at University of Bordeaux 1 (IUT), LaBRI]

Staff member from University of Bordeaux 1

Boniface Nkonga [Senior assistant professor, MAB]

Staff member from ENSEIRB

Olivier Beaumont [Assistant professor, LaBRI]

Post-doctoral fellow

Nikolaï Andrianov [funding from HYKE since october 2003 until end of june 2004, MAB]

Mohamed Mezine [funding from Conseil Régional d'Aquitaine until december 2003, MAB]

Ph. D. students

Guillaume Anciaux [funding from INRIA and CEA since december 2003, INRIA]

Benjamin Braconnier [funding from CEA since october 2003, MAB]

Aurélien Esnard [funding from Ministry of Research and Technology, LaBRI]

Pierre Fortin [funding from CNRS and Conseil Régional d'Aquitaine, LaBRI]

Sébastien Gaucel [funding from Ministry of Research and Technology, MAB]

Mikaël Papin [funding from CEA, MAB]

Gaël Tessier [funding from Conseil Régional d'Aquitaine since october 2003, LaBRI]

Technical staff

Michaël Dussère [Project technical staff since october 2002]

Dimitri Lecas [Junior technical staff since september 2002]

Research scientist (partners)

Rémi Abgrall [Professor, University of Bordeaux 1 and Institut Universitaire de France, MAB]

Patrick Amestoy [Senior assistant professor, ENSEEIHT², IRIT³]

Serge Chaumette [Senior assistant professor, University of Bordeaux 1, LaBRI]

Michel Langlais [Professor, University of Bordeaux 2, MAB]

Guillaume Latu [Assistant professor, University Louis Pasteur of Strasbourg, LSIIT⁴]

Gérald Monard [Assistant professor, University Henri Poincaré of Nancy, GCTN⁵]

¹École Nationale Supérieure d'Électronique, d'Informatique, et de Radiocommunications de Bordeaux.

²École Nationale Supérieure d'Électrotechnique, d'Électronique, d'Informatique, et d'Hydraulique de Toulouse.

³Institut de Recherche en Informatique de Toulouse, UMR 5505.

⁴Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection, UMR 7005.

⁵Equipe de Chimie et Biochimie Théoriques de Nancy, UMR 7565.

Student interns

Ali-Réda Benkhaldoun [funding from INRIA since february 2003 until end of december 2003, ENSEIRB]
Benjamin Braconnier [funding from CEA since february 2003 until end of august 2003, MatMéca]
Sébastien Christy [contract with CEA until end of november 2003, ENSEIRB]
Stéphane Ekobo Kidou [Internship since february 2003 until end of august 2003, INRIA]
Thomas Fouquet [since february 2003 until end of june 2003, MatMéca]
Tristan Letellier [funding from INRIA since june 2003 until end of august 2003, MatMéca]
Mayer Noukou [funding from INRIA since february 2003 until end of june 2003, ENSEIRB]
Séverin Marret [contract with Dassault Aviation since october 2002 until end of july 2003, MatMéca]
Marc-Olivier Titeux [funding from INRIA since august 2003 until end of september 2003, MatMéca]

2. Overall Objectives

The purpose of the SCALAPPLIX project is to analyze and solve scientific computation problems arising from complex research and industrial applications and involving scaling. These applications are characterized by the fact that they require enormous computing power, on the order of tens or hundreds of teraflops, and that they handle huge amounts of data. Solving these kinds of problems requires a multidisciplinary approach concerning both applied mathematics and computer science. In applied mathematics, it is essentially the field of numerical schemes that is concerned. In computer science, parallel computing and the design of high-performance codes to be executed on today's major computing platforms are concerned (parallel computers organized as a large network of SMP nodes, GRID platforms).

Through this approach, SCALAPPLIX intends to contribute to all steps in the line that goes from the design of new high-performance, more robust and more precise, numerical schemes to the optimized implementation of algorithms and codes for the simulation of *physical* (fluid mechanics, inert and reactive flows, multimaterial and multiphase flows), *biological* (molecular dynamics simulations) and *environmental* (host-parasite systems in population dynamics) phenomena that are by nature multiscale and multiphysics.

Another domain we are currently investigating is the development of distributed environments for coupling numerical codes and for steering interactively numerical simulations. The computational steering is an effort to make the typical simulation work-flow (modeling, computing, analyzing) more efficient, by providing on-line visualization and interactive steering over the on-going computational processes. On-line visualization appears very useful to monitor and detect possible errors in long-running applications, and interactive steering allows the researcher to alter simulation parameters on-the-fly and immediately receive feedback on their effects. Thus, scientists gain a better insight in the simulation regarding to the cause-effect relationship and can better grasp the complexity of the underlying models.

3. Scientific Foundations

3.1. Numerical schemes for fluid mechanics

Key words: *unstructured meshes, upwind schemes, complex physical models.*

A large number of industrial problems can be translated into fluid mechanics ones. They may be coupled with one or more physical models. An example is provided by aeroelastic problems, which are studied in details by other INRIA teams. Another example is given by flows in pipelines where the fluid (a mixture of air–water–gas) has not very well-understood physical properties. One may also consider problems in aeroacoustics, which become more and more important in everyday life. In some occasions, one needs specific numerical tools because fluids have exotic equation of states, or because the amount of computation becomes huge, as for unsteady flows.

It is a fact that there are many commercial codes. They allow one to consider some of these examples, but the quality of the solutions is far from being optimal. Moreover, the numerical tools of these codes are often

not the most recent ones. Last, the know-how that results into these codes is not available. An example is the noise generated by vortices crossing through a shock wave. It is, up to our knowledge, out of reach of the most recent technologies because the numerical resources that would necessitate such simulations are tremendous. In the same spirit, the simulation of a compressible mixing layer in a complex geometry is also out of reach because very different temporal and physical scales need to be captured.

In order to reach this goal efficient simulation of complex physical problems, we are working on some fundamental aspects of the numerical analysis of non linear hyperbolic problems. Our goal is to develop schemes that can adapt to the modern computer architectures. More precisely, we are working on a class of numerical schemes specifically tuned for unstructured meshes. They have the most possible compact stencil that is compatible with the expected order of accuracy. The order of accuracy typically ranges from two to four. Since the stencil is compact, the implementation on parallel machines becomes simple. The price to pay is that the scheme is necessarily implicit. However, the compactness of the scheme makes us expect to use the high performance parallel linear algebra tools developed by the team. Other research theme may emerge out of these themes. In parallel to these aspects of fundamental numerical analysis, we also work on adapting more classical numerical tools to complex physical problems such as those encountered in interface flows, turbulent or multiphase flows.

Within a few years, we expect to be able to consider the physical problems that are now difficult to compute thanks to the know-how coming from our research on compact distribution schemes. These problems range from aeroacoustic ones to multiphysics problems, such as the one mentioned above.

3.2. Schemes and algorithms for computational chemistry

Key words: *multiscale schemes, molecular dynamics, parallelism and distributed computing.*

Due to the increase of available computer power, new applications such as reaction paths, free energy computations, biomolecular dynamics simulations or failure material simulations are now commonly performed by chemists. These computations simulate systems up to several thousands of atoms, for large time scales up to several nanoseconds. The larger the simulation is, the smaller the computational cost of the potential driving the phenomena is, resulting in low precision results. To achieve realistic results, simulations need to include the environment surrounding the molecules, such as water and membranes, resulting in system sizes up to about several hundred thousands of atoms. Furthermore, simulating the aggregation of proteins, which is critical for biologists studying viruses, requires models of up to one million atoms, with a simulation time up to one millisecond. This implies that atomistic simulations must be speeded up by several orders of magnitude. To obtain this speed, numerical and parallel algorithms must be improved, as well as their implementations on distributed or parallel architectures.

We are currently focusing on several aspects of these problems. First, we try to improve models and algorithms. To do this, we decrease the complexity of classical algorithms by introducing new approximations in the algorithms, in the model (this is the trick of linear scaling methods like the divide-and-conquer method), and by proposing new algorithms.

Second, we apply multiscale methods to decrease the number of atoms that are considered at the finest level (electronic or atomistic). To do this, we introduce a coarser model like continuum media to take into account the electrostatic effect of the environment, or an elasticity model for crystals. The difficulty here is to build an efficient scheme which couples the two different scales without any loss of precision.

Finally, efficient implementation is necessary to reach the desired level of performance. For instance, we can rewrite our algorithms in the form of block computations, in order to use efficient computational routines such as BLAS vector-matrix operations, or to implement accurate load balancing strategies.

Another domain we are currently investigating is the development of distributed environments for coupling numerical codes and for interactively steering numerical simulations.

3.3. Algorithms and high-performance solvers

Key words: *irregular graph partitioning, sparse linear algebra, fast multipole methods, parallel algorithms, high-performance computing.*

3.3.1. High-performance direct solvers for distributed clusters

Solving large sparse systems $Ax = b$ of linear equations is a crucial and time-consuming step, arising in many scientific and engineering applications. Consequently, many parallel techniques for sparse matrix factorization have been studied and implemented.

We have started this research by working on the parallelization of an industrial code for structural mechanics, which was a 2D and 3D finite element code and non linear in time. This computational finite element code solves plasticity problems (or thermo-plasticity problems, possibly coupled with large displacements). Since the matrices of these systems are very ill-conditioned, classical iterative methods are not an issue. Therefore, to obtain an industrial software tool that must be robust and versatile, high-performance direct sparse solvers are mandatory, and parallelism is then necessary for reasons of memory capabilities and acceptable solving time. Moreover, in order to solve efficiently 3D problems with more than 10 millions of unknowns, which is now a reachable challenge with new SMP supercomputers, we must achieve a good time scalability and control memory overhead.

In the SCALAPPLIX project, we focus first on the block partitioning and scheduling problem for high performance sparse LDL^T or LL^T parallel factorization without pivoting for large sparse symmetric positive definite systems. Our strategy is suitable for general distributed heterogeneous architectures whose computation and communication performances are predictable in advance. In order to achieve efficient parallel sparse factorization, we study the following pre-processing phases:

- The *ordering* phase, which computes a symmetric permutation of the initial matrix A such that factorization process will exhibit as much concurrency as possible while incurring low fill-in. We use a tight coupling of the Nested Dissection and Approximate Minimum Degree algorithms. The partition of the original graph into supernodes is achieved by merging the partition of separators computed by the Nested Dissection algorithm and the supernodes amalgamated for each subgraph ordered by Halo Approximate Minimum Degree.
- The *block symbolic factorization* phase, which determines the block data structure of the factorized matrix L associated with the partition resulting from the ordering phase. One can efficiently perform such a block symbolic factorization in quasi-linear space and time complexities. From this block structure, one can deduce the weighted task graph that captures all dependencies between blocks, as well as the supernodal elimination tree.
- The *block repartitioning and scheduling* phase, which refines the previous partition by splitting large supernodes in order to exploit concurrency within dense block computations, and which maps the resulting blocks onto the processors of the target architecture.

These topics are also investigated for general sparse non-symmetric matrices with a more dynamic strategy due to the need for numerical pivoting during the factorization.

Research about high performance sparse direct solvers is carried on in collaboration with P. Amestoy (ENSEEIH – IRIT) and J.-Y. L'Excellent (INRIA Rhône-Alpes), and has led to software developments (see section 5.1) and to industrial contracts with CEA (Commissariat à l'Énergie Atomique).

3.3.2. Fast Multipole Methods

In most of scientific computing applications considered nowadays as computational challenges like biological systems, astrophysics or electromagnetism, the introduction of hierarchical methods based on an octree structure has dramatically reduced the amount of computation needed to simulate those systems for a given error tolerance.

Among these methods, the Fast Multipole Method (FMM) allows the computation of the interactions in, for example, a molecular dynamics system of N particles in $O(N)$ time, against $O(N^2)$ with a direct approach. The extension of these methods and their efficient implementation on current parallel architectures is still a critical issue. Moreover the use of periodic boundary conditions, or of duplications of the system in 2 out of 3 space dimensions, just as well as the use of higher approximations for integral equations are also still relevant.

In order to treat biological systems of up to several millions of atoms, these methods must be integrated in the QC++ platform (see section 5.5). They can be used in the three (quantum, molecular and continuum) models for atom-atom interactions in quantum or molecular mechanics, atom-surface interactions for the coupling between continuum and the other models, and also for fast matrix-vector products in the iterative solving of the linear system given by the integral formulation of the continuum method. Moreover, the significant experience achieved by the SCOTCH and PASTIX projects (see section 5.1) will be useful in order to develop efficient implementations of the FMM methods on parallel clusters of SMP nodes.

3.4. Parallel algorithms for heterogeneous platforms

Key words: *parallel algorithms, computational grids, scheduling.*

Recently, a lot of work has been devoted to computational grids. Such computing architectures differ from usual parallel platforms in terms of heterogeneity (of both processing and communication resources) and scale (use of large distance network links with high latencies). Such platforms are usually not dedicated to one application, and therefore, security and fault tolerance problems also arise. In our works, we do not consider the security and fault tolerance problems, but rather concentrate on additional difficulties arising from the heterogeneity and the dynamicity (in terms of performances rather than topology) of such platforms.

Our goal is to design efficient scheduling algorithms for heterogeneous and non-dedicated platforms. Scheduling computational tasks on a given set of processors is a key issue for high-performance computing. The traditional objective of scheduling algorithms is makespan minimization: given a task graph and a set of computing resources, find a mapping of the tasks onto the processors, and order the execution of the tasks so that (i) task precedence constraints are satisfied; (ii) resource constraints are satisfied; and (iii) a minimum schedule length is provided. However, makespan minimization turned out to be NP-complete problem in most practical situations and the advent of more heterogeneous architectural platforms is likely to even increase the computational complexity of the process of mapping applications to machines.

Many of the works presented in section 6.5 have been done in collaboration with the GRAAL project (INRIA Rhône-Alpes), during the PhD thesis of Arnaud Legrand (defended on 11/12/03) and Loris Marchal (begun 09/03), both co-directed by Olivier Beaumont and Yves Robert (GRAAL project).

3.5. Computational steering for distributed numerical simulations

Key words: *numerical simulation, computational steering, scientific visualization, interaction, coupling.*

Thanks to the constant evolution of computational capacity, numerical simulations are becoming more and more complex; it is not uncommon to couple different models in different distributed codes running on heterogeneous networks of parallel computers (e.g. multi-physics simulations). For years, the scientific computing community has expressed the need of new computational steering tools to better grasp the complexity of the underlying models.

The computational steering is an effort to make the typical simulation work-flow (modeling, computing, analyzing) more efficient, by providing on-line visualization and interactive steering over the on-going

computational processes. On-line visualization appears very useful to monitor and detect possible errors in long-running applications, and the interactive steering allows researchers to alter simulation parameters on-the-fly and immediately receive feedback on their effects. Thus, the scientist gains a better insight in the simulation regarding to the cause-and-effect relationship.

A computational steering environment can be defined as a communication infrastructure, coupling a simulation with a remote user interface, called steering system. This interface usually provides on-line visualization and user interaction. Over the last decade, many steering environments have been developed; they distinguish themselves by some critical features such as the simulation integration process, the communication infrastructure and the steering system design. A first solution for the integration is the problem solving environment (PSE) approach, like in SCIRun. This approach allows the scientist to construct a steering application according to a visual programming model. As an opposite, environments like CAVEStudy only interact with the application through its standard input/output. Nevertheless, the majority of the steering environments, such as the well-known CUMULVS, are based on the instrumentation of the application source-code; this approach allows fine grain steering functionalities and achieves good runtime performances. Regarding the communication infrastructure, there are many underlying issues especially when considering parallel and distributed simulations: heterogeneous data transfers, network communication protocols and data redistributions.

In EPSN project, we intend to explore the capabilities of the CORBA technology and the currently under development parallel CORBA objects for the computational steering of parallel and distributed simulations. This environment, that we are working on, allows the control, the data exploration and the data modification for numerical simulations involving an iterative process. In order to be as generic as possible, we introduce an abstract model of steerable simulations. This abstraction allows us to build steering clients independently of a given simulation. This model is described with an XML syntax and is used in the simulation by some source code annotations. EPSN takes advantage of the CORBA technology to design a communication infrastructure with portability, interoperability and network transparency. In addition, the in-progress parallel CORBA objects will give us a very attractive framework for extending the steering to parallel and distributed simulations.

4. Application Domains

4.1. Introduction

The main objective of SCALAPPLIX project is to analyze and solve scientific computing problems coming from complex research and industrial applications and involving scaling. This allows us to validate the numerical schemes, the algorithms and the associated softwares that we develop. We have today three reference application domains which are fluid mechanics, molecular dynamics and host-parasite systems in population dynamics. In these three domains, we study and simulate phenomena that are by nature multiscale and multiphysics, and that require enormous computing power. A major part of these works leads to industrial collaborations in particular with the french CEA/CESTA and CEA/Ile-de-France centers.

4.2. Fluid mechanics

Key words: *fluid mechanics, unsteady flows, multiphase flows, unstructured meshes.*

The numerical simulation of unsteady flows is still a challenge since efficient schemes and efficient implementations are needed. This challenge is even higher if large size problems are to be tackled, and if the meshes are not regular.

Among the problems to be considered, one may list the computation of mixing layers, shock–vortices interactions, the noise generated by a flow. This last item clearly needs very high order schemes, and the today best schemes use regular structured meshes. Hence, one of our objectives is to construct very high order schemes for unstructured meshes.

Another example where large computer resources are needed is the simulation of multiphase flows. In that case, several difficulties have to be faced: unsteady flows, complex geometries and a very complex physical model.

4.3. Molecular chemistry

Key words: *protein, enzyme reaction, membrane, drug design, biological simulation, parallelism, molecular dynamics, quantum method, continuum method.*

Due to the increase of available computer power, new applications such as reaction paths, free energy computations, biomolecular dynamics simulations or failure material simulations are now commonly performed by chemists. These computations simulate systems up to several thousands of atoms, for large time scales up to several nanoseconds. The larger the simulation is, the smaller the computational cost of the potential driving the phenomena is, resulting in low precision results. To achieve realistic results, simulations need to include the environment surrounding the molecules, such as water and membranes, resulting in system sizes up to about several hundred thousands of atoms. Furthermore, simulating the aggregation of proteins, which is critical for biologists studying viruses, requires models of up to one million atoms, with a simulation time up to one millisecond. This implies that atomistic simulations must be speeded up by several orders of magnitude. To obtain this speed, numerical and parallel algorithms must be improved, as well as their implementations on distributed or parallel architectures.

4.4. Population Dynamics

Key words: *Biomathematics, population dynamics, host-microparasite system, host-macroparasite system, deterministic model, individual based model, spatially explicit model, parasite aggregation, health.*

In population dynamics, systems can present very complex behaviors and can be difficult to analyse from a purely mathematical point of view. The aim of this interdisciplinary project was to develop numerical tools for population dynamics models arising in modelling complex heterogeneous host-parasite systems. Some typical heterogeneities we consider are spatial locations, age or ability to recruit macroparasites for hosts, age of macroparasites. Our main goals are: understanding the impact of a host population structure on a parasite population dynamics, developing accurate numerical simulations using parallelization, designing prophylactic methods. For many host-parasite systems different time scales between the host population (e.g. a one year period) and the virus (e.g. an infected host dies with a few weeks) require a small time step. Numerical schemes of the resulting nonlinear epidemiological model in spatially heterogeneous environment are complex to perform and reliable numerical results become difficult to get when the size of the spatial domain is increasing. In addition, many input parameters (biological and environmental factors) are taken into account to compare results of simulations and observations from field studies. Therefore, a realistic simulator has a significant computation cost and parallelization is required.

Individual-Based Models (IBM) are becoming more and more useful to describe biological systems. Interactions between individuals are simple and local, yet can lead to complex patterns at a global scale. The principle is to replicate several times the simulation program to obtain statistically meaningful results. The Individual-Based Model approach contrasts with a more aggregate population modeling approach in providing low level mechanisms to manage the population interactions. Stochastic simulations reproduce elementary processes and often lead to prohibitive computations; thus we need parallel algorithmic.

In our developments of both stochastic and deterministic models, biological processes are combined to reach a good level of realism. For host-parasite systems, it make a big difference with purely mathematical models, for which numerical results could hardly be compared to observations. Parallel numerical simulations mimic some of the dynamics observed in the fields, and supply a usable tool to validate the models. This work is a collaborative effort in an interdisciplinary approach: population dynamics with P. Silan (UMR CNRS 5000, Sète), mathematics with M. Langlais and S. Gaucel (MAB - Université Victor Segalen Bordeaux 2) and computer science.

5. Software

5.1. Introduction

We develop two kinds of software. The first one consists in generic libraries that will be used in the applications. We work on a partitioner for large irregular graphs or meshes (SCOTCH), on high performance direct or hybrid solvers for very large sparse systems of equations (MUMPS, PASTIX). The second one corresponds to dedicated software for molecular chemistry (QC++), fluid mechanics (FLUIDBOX), and to a platform for computational steering (EPSILON). For these parallel software developments, we use the message passing (MPI) paradigm, the OpenMP programming language, threads, and the Java and/or CORBA technologies.

5.2. MUMPS

Participant: Patrick Amestoy [corresponding member], François Pellegrini.

Key words: *parallel asynchronous sparse multifrontal solver.*

In the context of PARASOL (Esprit IV Long Term Project, 1996-1999), CERFACS and ENSEEIHT-IRIT teams have initiated a parallel sparse solver MUMPS (“MULTifrontal Massively Parallel Solver”). Since the first public release of MUMPS (March 2000), much research work has been done in collaboration with J. Y. L’Excellent from INRIA project Remap and Sherry Li and Esmond Ng from NERSC (Lawrence Berkeley National Laboratory). This work related to performance scalability, orderings for unsymmetric matrices, and dynamic scheduling has been incorporated in the new improved version of the package (release 4.3.2 available since November 2003 at <http://www.enseeiht.fr/apo/MUMPS>).

MUMPS is a package for solving linear systems of equations $\mathbf{Ax} = \mathbf{b}$, where the matrix \mathbf{A} is sparse and can be either unsymmetric, symmetric positive definite, or general symmetric. It uses a multifrontal technique which is a direct method based on either the LU or the LDL^T factorization of the matrix. The main features of the MUMPS package include numerical pivoting during factorization, solution of the transposed system, input of the matrix in assembled format (distributed or centralized) or elemental format, error analysis, iterative refinement, scaling of the original matrix, and return of a Schur complement matrix. It also offers several built-in ordering algorithms, a tight interface to some external ordering packages such as SCOTCH and is available in various arithmetics (real or complex, single or double).

5.3. PaStiX

Participants: Pascal Hénon, Dimitri Lecas, François Pellegrini, Pierre Ramet [corresponding member], Jean Roman.

Key words: *supernodal sparse parallel solver.*

This work is supported by the French “Commissariat à l’Energie Atomique CEA/CESTA” in the context of structural mechanics and electromagnetism applications.

PASTIX (<http://www.labri.fr/~ramet/pastix>) is a scientific library that provides a high performance solver for very large sparse linear systems, based on block direct and ILU(k) iterative methods. Many factorization algorithms are implemented in simple or double precision (real or complex): LLt (Cholesky), LDLt (Crout) and LU with static pivoting (for non symmetric matrices having a symmetric pattern). This latter version will be used in FLUIDBOX .

The PASTIX library uses the graph partitioning and sparse matrix block ordering package SCOTCH (see section 5.4). PASTIX is based on an efficient static scheduling and memory manager, in order to solve problems with more than 10 millions of unknowns. The mapping and scheduling algorithm handles a combination of 1D and 2D block distributions. This algorithm computes efficient static schedulings of the block computations for a parallel solver based on a supernodal approach with total aggregation of contribution blocks, such that the parallel solver is fully driven by this scheduling. This can be done by taking into account

very precisely the computational costs of the BLAS 3 primitives, the communication costs and the cost of local aggregations.

Another important point is that our study is suitable for any heterogeneous parallel/distributed architecture the performance of which is predictable, such as clusters of SMP nodes. We also improved our static computation and communication scheduling algorithm to anticipate the sending of partially aggregated blocks, in order to free memory dynamically. By doing this, we are able to reduce dramatically the aggregated memory overhead, while keeping good performance.

However, direct methods may fail to solve very large three-dimensional problems, due to the large amount of memory needed for these cases and despite any memory optimization. Our approach now consists in symbolically computing the block structure of the factors that would have been obtained with a complete factorization, and then deciding to drop off some blocks of this structure according to relevant criteria. This incomplete factorization induced by the new sparse pattern, called RICAR (Robust Incomplete Cholesky Algorithm), is then used within a preconditioned GMRES or Conjugate Gradient solver.

5.4. Scotch

Participant: François Pellegrini [corresponding member].

Key words: *graph partitioning, mesh partitioning, static mapping, sparse matrix block ordering.*

The initial purpose of SCOTCH was to provide an efficient software environment for partitioning and mapping statically applications modeled as valuated process graphs of arbitrary topologies. The original contribution consisted in developing a “*divide and conquer*” algorithm in which processes are recursively mapped onto processors by using graph bisection algorithms that are applied both to the process graph and to the architecture graph. This allows the mapper to take into account the topology and heterogeneity of the valuated graph which models the interconnection network and its resources (processor speed, link bandwidth). This technique allowed to compute high quality mappings with low complexity.

Based on these results, new graph partitioning algorithms have been developed, which compute vertex separators instead of edge separators, using a multi-level framework. Recursive vertex separation is used to compute reorderings of the unknowns of large sparse linear systems, which both preserve sparsity when factoring the matrix and preserve concurrency for computing and solving the factored matrix in parallel. The original contribution has been to study and implement a tight coupling between the nested dissection and the approximate minimum degree methods; this work was carried out in collaboration with Patrick Amestoy, of ENSEEIHT-IRIT.

Recently, new classes of methods have been added to the SCOTCH library, which allow it to compute efficient orderings of native meshes, resulting in increased performance compared to standard graph partitioners.

The SCOTCH software package (<http://www.labri.fr/~pelegrin/scotch/>), the version 4.0 of which is about to be released as a GPLed libre software, is quite popular and compares favorably with its most known US competitor, METIS.

5.5. QC++

Participants: Olivier Coulaud [corresponding member], Ali-Réda Benkhaldoun, François Pellegrini, Gérald Monard.

Key words: *quantum chemistry, semi-empirical, simulation, parallelism.*

QC++ is a Quantum Chemistry software written in C++. The current version of QC++ supports the semi-empirical quantum models MNDO, AM1 and PM3. It allows to calculate the energy of a molecular configuration by some Self Consistent Field (SCF) algorithms : fixed point, optimal damping and level shifting. It is also possible to optimize its geometry by using the minimization algorithms L-BFGS and BFGS.

The major new feature in the version 2.0 is the implementation of linear scaling method Divide and Conquer both in sequential and parallel to calculate the electronic energy. Several type of subdomains and strategies to partition a molecule are available, one of them being based on SCOTCH.

5.6. FluidBox

Participants: Christophe Berthon, Boniface Nkonga [corresponding member], Rémi Abgrall.

Key words: *fluid mechanics, inert and reactive flows, multimaterial and multiphase flows.*

FLUIDBOX is a software dedicated to the simulation of inert or reactive flows. It is also able to simulate multiphase and multimaterial flows. There exist 2D- and 3D dimensional versions. The 2D version is used to test new ideas that are later implemented in the 3D one. Two classes of schemes have been implemented: classical finite volume schemes and the more recent residual distribution schemes. Several low Mach preconditioning techniques are also implemented. The code has been parallelised with and without overlapping of the domains. It is used by our industrial contacts : CEA/CESTA, EADS,SNPE.

5.7. EPSILON

Participants: Serge Chaumette, Olivier Coulaud [corresponding member], Michaël Dussère, Aurélien Esnard, Guillaume Latu, Boniface Nkonga, Jean Roman.

Key words: *numerical simulation, computational steering, on-line visualization, CORBA.*

EPSILON is the prototype of EPSN computational steering environment. This plat-form allows scientists to couple their existing simulations (written in C, C++ or FORTRAN) with remote user interfaces providing on-line visualization and user interaction. Epsilon, which is based on CORBA (omniORB), provides a C API for instrumenting both simulations and client applications. It addresses sequential numerical simulations and has an extension for steering parallel simulations. Release 0.1, which is distributed among the project participants, allows the construction of multi-simulation and multi-client steering systems providing the control, the data exploration and the data modification for the iterative processes. The client examples distributed with Epsilon show the diversity of solutions permitted by the environment from the generic Java-swing client to the AVS modules for on-line advanced visualization.

6. New Results

6.1. Numerical schemes and algorithms for fluid mechanics.

Participants: Rémi Abgrall, Nikolai Andrianov, Christophe Berthon, Benjamin Braconnier, Stephane Ekobo, Thomas Fouquet, Dimitri Lecas, Tristan Letellier, Séverin Marret, Mohamed Mezine, Boniface Nkonga, Mikaël Papin, François Pellegrini, Pierre Ramet, Jean Roman, Marc-Olivier Titeux.

Key words: *Compact distribution schemes, multimaterial and multiphase flows, aeroacoustics.*

6.1.1. Compressible multiphase flows

We have extended our previous works [14] to the case of unstructured meshes in FLUIDBOX. This new work has been presented in one of the satellite conferences of ICM2002 [47]. The proposed scheme is currently being extended to a simplified equation of state in order to better take into account the quasi incompressible nature of some constituents. This work is supported by the CEA/CESTA.

In order to improve the numerical accuracy and to reduce the computational time consumption, some simplified models of the Bear – Nuziato multi-fluid model have been proposed. The reduced models are obtained after a scaling of the equations and a WKB development. These models can take into account some complex phenomena such as: turbulence, gravity, surface tension... However, it does not exist a fully conservative form of the models. In order to recover the physical behavior, strategies are defined to approximate the non-conservative products: nonlinear projection, extended consistency. The resulting schemes are accurate and robust. Four communications are associated to this topic.

The other challenge is to developed an efficient numerical strategy for complex equations of states. Some accurate relaxations schemes have been developed and validated for the isobaric and unique velocity model.

This work is under progress in the context of B. Braconnier PhD thesis, with some 3D applications proposed by the CEA/CESTA. These computations are still expensive and these schemes need to be implemented on parallel computers. The parallel solution strategy is based on a domain decomposition achieved by the mesh partitioner SCOTCH. This work is under progress and will integrate the high performance parallel solvers developed in the SCALAPPLIX project for the solution of the linear systems associated with the implicit numerical scheme.

6.1.2. Residual distribution schemes

We are still working on the theoretical understanding of these particular schemes. Several results have been obtained about the stability, the approximation of viscous terms in the Navier – Stokes case. This has been done in collaboration with the Von Karmàn Institute (in collaboration with H. Deconinck, M. Ricchiutto).

The residual distribution version of FLUIDBOX has been extended to the 3D case, and an original preconditioning technique for low Mach number has been implemented by M. Mezine (post-doctoral fellow). This work is supported by the Région Aquitaine (project about the conception of airships).

The residual distribution scheme has been also evaluated for aeroacoustic problems by S. Marret through a collaboration with Dassault Aviation (M. Ravachol). The conclusion of this work, presented in an Euromech conference in december 2003, is that these schemes are indeed much more accurate than their second order finite volume counterparts, but also more accurate than higher order versions designed by the former SINUS project. This work has been supported by the Ministère de l'Industrie. The unsteady version of the residual distribution scheme has also been studied for space and time accuracy higher than two by Th. Fouquet. This work is being extended by N. Andrianov (HYKE post-doctoral fellow).

6.1.3. Hamilton Jacobi equations

Last year we designed a scheme for Hamilton Jacobi that was inspired by our work on residual distribution scheme in their blended version. A 3D version of the scheme has been implemented for SNPE. This has motivated a better understanding of the method: a convergence proof is available for steady problems, as well as a more efficient version in term of accuracy.

6.1.4. Porous medium flows

In the context of the management of subsoil waters, a numerical model for flows in porous medium has been designed. The mathematical model, based on the Darcy's law, is a 2D diffusion equation solved by using a mixed hybrid finite volume method. This work has been done in the intership INRIA's program with the university of Yaoundé (Cameroon).

6.2. Schemes and algorithms for computational chemistry

Participants: Guillaume Anciaux, Ali-Reda Benkhaldoun, Olivier Coulaud, Pierre Fortin, Gérald Monard, François Pellegrini, Jean Roman.

Key words: *quantum chemistry, molecular dynamics, hierarchical methods, coupling model.*

6.2.1. Divide & Conquer method

In the QC++ code, the divide and conquer method has been parallelized by using the message passing paradigm. The main features of the parallelization are the following. First, several domains can be handled by each processor, which will allow us to implement load balancing methods. Second, asynchronous communications are used to overlap communications by computations, according to two different communication patterns: for data that should be shared across all subdomains, such as the diagonal blocks of the global density matrix which are used to build the global Fock matrix, a ring of communication is used; otherwise, point to point communications are performed. Third, a molecule partitioner based on the SCOTCH library, called KIMIKA,

has been developed. Several types of partitioning and clustering strategies (by atom or by fragment) are currently being evaluated. An article describing the parallelization strategies and the partitioning aspects is in preparation.

6.2.2. Crack simulation

A study of crack propagation in silica glass with a coupling method between molecular dynamics and elasticity has just begun in collaboration with the CEA. A state-of-the-art for large scale atomistic simulation for crack propagation, finite element continuum elasticity and for the different coupling methods has been done.

6.3. Numerical schemes and algorithmic for population dynamics

Participants: Sébastien Gaucel, Guillaume Latu, Michel Langlais, Jean Roman, Gaël Tessier.

Key words: *deterministic simulator, stochastic simulator, sparse matrices.*

For many host-parasite systems using matrix computations, memory storage problems and lengthy computational times occurred. To tackle this problem, a crucial observation is that the nonlinear matrix model involves sparse periodic matrices. Then, using appropriate storage techniques and sparse matrices multiplication techniques allow to simulate large size spatial domains and get robust numerical results. Methods using sparse matrix multiplications are applied to analyze rabies propagation and help designing prophylactic methods.

Fish ectoparasites interact continuously with their host populations. A model describing the demographic strategies of such fish and parasite populations has been developed for the *Diplectanum aequans*-Sea Bass system. This model is mostly deterministic with some stochastic aspects. Aggregation of macroparasites on the hosts is not assigned to the model but occurs or does not occur depending on the parasite population dynamics. Previous deterministic discrete simulations brought about too large computations and some simulations had important run-times. A high-performance simulator working on parallel machines (IBM SP3, Regatta and SGI Origin 3800, parallel machines of CINES – Montpellier) and providing more accurate computations has been implemented. The algorithmic study and a performance analysis establish the efficiency and scalability (checked up to 448 processors) of the parallel algorithm. The parallel simulator provides more accurate computations than the sequential one. The parallel efficiency reaches 77 % on 448 processors for a complete simulation. A study of memory accesses and cache utilization leads to an implementation reaching 60 % of peak performance in the computation kernel on the IBM SP3 and on the Origin 3800. A complete simulation of 1.45 PFLOP was achieved in only two hours on 256 processors (IBM SP3). A full analysis of numerical simulations has allowed us to tune the model to get a realistic qualitative behavior [27]. Then, a thorough validation of the model is actually performed. Results of numerical simulations show the effect of overdispersion, parasite and host mortality on the parasite distribution, and host regulation (occurring through cycles).

Subsequently an individual-based model has been developed. We used a Monte Carlo algorithm for this stochastic simulator for which we described three different levels of parallelism. Analysis and performance, up to 256 processors, of a hybrid MPI/OpenMP code were studied for a cluster of SMP nodes. The qualitative results of both parallel simulators were compared [10]. Improving the model leads to a deeper understanding of the processes occurring in the real biological system. Even though the method is different from the deterministic one, results are qualitatively similar for identical data sets.

A cooperation involving a biologist (Agnes Callonec - INRA UMR Santé végétale 1065 - Villenave d'Ornon) and a thesis student in computer science (Gaël Tessier) began in 2003. Using numerical methods and parallel technics, we are interested in modelling the spread of *oidium*, a disease of vineyard. Correct prediction of this type of parasite epidemics needs a realistic simulator, and would have an industrial impact. An architectural model of vine stocks is used for two purposes: the study of the growth of stocks and the influence of its structure on the dispersal of *oidium*. In this model, we consider a large number of infectious elements and several spatially heterogeneous environmental parameters. Indeed, the dispersal of *oidium* is a multiscale mechanism that takes place within vine stocks, and along and across the rows of the vineyard.

6.4. Algorithms and high-performance solvers

Participants: Patrick Amestoy, Olivier Beaumont, Olivier Coulaud, Pierre Fortin, Pascal Hénon, Dimitri Lecas, Boniface Nkonga, François Pellegrini, Pierre Ramet, Jean Roman.

Key words: *sparse matrix ordering, graphs and irregular meshes, high performance computing, direct and hybrid direct-iterative solvers, parallel fast multipole methods.*

6.4.1. Domain decomposition and sparse matrix reordering

The work carried out within the SCOTCH project (see section 5.4) focused on the development of efficient algorithms for partitioning and reordering meshes. As a matter of fact, one of the bottlenecks that have been evidenced when using the PASTIX software suite (see section 5.3) to solve large 3D problems is that the size of the adjacency graphs, which is linear in the number of elements but quadratic with respect to the number of nodes of each of the elements (since these latter are considered as cliques), resulted in memory shortage and disk swapping during the pre-processing phase.

Using native mesh structures instead of adjacency graphs makes memory occupation linear with respect to the number of nodes of the elements, but requires additional work on the algorithms that traverse these structures such that the gain in space complexity does not result in a substantial loss in time complexity. Preliminary results show that both space and time occupations are significantly reduced when using 3D meshes, because the reduction of memory usage allows for better data locality and less page faults, which compensates for the extra computing cost.

On-going work now focuses on the tuning of approximate minimum degree methods to account for the specificities of the native mesh structures, in a joint work with Patrick Amestoy, and on the analysis of the coupling of nested dissection and minimum degree methods. Preliminary studies on the parallelization of the SCOTCH library will also be continued.

6.4.2. High-performance direct solvers on multi-plateforms

In order to solve linear systems of equations coming from 3D problems and with more than 10 millions of unknowns, which is now a reachable challenge for new SMP supercomputers, the parallel solvers must keep good time scalability and must control memory overhead caused by the extra structures required to handle communications.

In our current works, we obtain with our direct solvers rather good scalability for many test problems. In most large cases, the efficiency is better than the one obtained on IBM SP2, and this shows that our techniques are suitable for SMP architectures. Some experiments were run on the TERA parallel computer at CEA (France), and the factorization times are close to the ones obtained on the IBM SP3 of CINES (Montpellier, France). For example, on our largest problem (26 millions of unknowns for a 2D 1/2 problem), we reach 500Gflops on 768 processors, that is, about 50% of the peak performance of the TERA computer.

The first major improvement has consisted in taking into account heterogeneous architectures, and more particularly the ones based on SMP nodes, such as the IBM SP3. Our communication model, used during the static scheduling and mapping step, is extended to manage both data exchanges by shared memory (less costly) and data exchanges through the network (more costly). We have proposed a mapping and scheduling algorithm for clusters of SMP nodes, and we have shown the benefits on run-time performance of such strategies.

In addition to the problem of run-time performance, another important aspect in the direct solution of very large sparse systems is the great memory requirements usually needed to factor the matrix in parallel. These memory requirements are due either to the structures required for communication, or by the matrix coefficients themselves.

To deal with these problems of memory management, our statically scheduled factorization algorithm can take advantage of the deterministic access (and allocation) pattern to all of the data stored in memory. Indeed, the data access pattern is determined by the computational task ordering that drives the access to the matrix coefficient blocks, and by the communication priorities that drive the access to the communication structures.

The results of this work has been presented in particular at the SIAM Linear Algebra 2003 conference [37], and is supported by CEA/CESTA for an extension of the PASTIX solver to unsymmetric matrices with symmetric pattern.

6.4.3. High-performance hybrid direct-iterative solvers for large sparse systems

In this work, we consider an approach which, we hope, will bridge the gap between direct and iterative methods. The goal is to provide a method which exploits the parallel blockwise algorithmic used in the framework of the high performance sparse direct solvers. We want to extend these high-performance algorithms to develop robust parallel incomplete factorization based preconditioners for iterative solvers such as GMRES or Conjugate Gradient solvers.

The idea is to define an adaptive blockwise incomplete factorization that is much more accurate (and numerically more robust) than the scalar incomplete factorizations commonly used to precondition iterative solvers. Such incomplete factorization can take advantage of the latest breakthroughs in sparse direct methods and particularly should be very competitive in CPU time (effective power used from processors and good scalability) while avoiding the memory limitation encountered by direct methods. By this way, we expect to be able to solve systems in the order of hundred millions of unknowns and even one billion of unknowns. Another goal is to analyse and justify the chosen parameters that can be used to define the block sparse pattern in our incomplete factorization.

The driving rationale for this study is that it is easier to incorporate incomplete factorization methods into direct solution software than it is to develop new incomplete factorizations. Our approach consists in computing symbolically the block structure of the factors that would have been obtained with a complete factorization, and then deciding to drop off some blocks of this structure according to relevant criteria. Our main goal at this point is to achieve a significant diminution of the memory needed to store the incomplete factors (with respect to the complete factors) while keeping enough fill-in to make the use of BLAS3 primitives profitable.

Preliminary results have been presented during the ‘‘Sparse Days’’ at Saint-Girons [8] and in the last SIAM conference on preconditioners [35]. This research is included in a NSF/INRIA project and is carried out in collaboration with Yousef Saad (University of Minneapolis, USA).

6.4.4. ILU factorization based on a hierarchical interface decomposition algorithm

In recent years, a few Incomplete LU factorization techniques were developed with the goal of combining some of the features of standard ILU preconditioners with the good scalability features of multi-level methods. The key feature of these techniques is to reorder the system in order to extract parallelism in a natural way. Often a number of ideas from domain decomposition are utilized and mixed to derive parallel factorizations.

Under this framework, we developed in collaboration with Yousef Saad (University of Minnesota) algorithms that generalize the notion of ‘‘faces’’ and ‘‘edge’’ of the ‘‘wire-basket’’ decomposition. The interface decomposition algorithm is based on defining a ‘‘hierarchical interface structure’’. The hierarchy consists of classes with the property that Class k nodes, with $k > 1$, are separators for class $k - 1$ nodes. In each class, nodes are grouped in independent sets. Class 1 nodes are simply interior nodes of a domain in the graph partitioning of the problem. These are naturally grouped in group-independent sets, in which the blocks (groups) are the interior points of each domain. Nodes that are adjacent to more subdomains will be part of the higher level classes and are ordered last.

The second part of this work is a factorization that uses dropping strategies which attempt to preserve the independent set structure. The Gaussian elimination process proceeds by classes: nodes of the first class are eliminated first, followed by those of the second class, etc. We propose two dropping rules: the first one privileges numerical robustness and the second one privilege parallelism. Generally, we apply the first rule for the factorization of the first 2 or 3 classes because in those first levels, the factorization implies communication between a few processors and we need to capture more numerical accuracy. We use the second rules in the higher classes factorization for the opposite reasons.

These algorithms have been implemented in the PHIDAL library [9][38]. Some first experiments have been realized on realistic 3D fluid dynamics test cases and they show a very promising scalability of the preconditioner (iterations, time). This work, still in progress, has been presented during the “Sparse Days” at Saint-Girons and in the last SIAM conference on preconditioners; a first version of the library should be released in 2004.

6.4.5. Parallel fast multipole method

The Fast Multipole Method (FMM) is a hierarchical method which computes interactions for the N-body problem in $O(N)$ time for any given precision. A theoretical study of this method has first determined its various aspects: 2 or 3 dimensional systems, with uniform or highly non-uniform distributions, with free-space or periodic boundary conditions. A rigorous study of the error bound has also been done succeeding in a clear error bound. Secondly a first version of this method for 2 and 3 dimensional cases of isolated systems with uniform tree has been successfully implemented, giving an efficient code.

In order to compute energy and forces on large systems, we need to improve the computation speed of the method. In this purpose, we have already used the BLAS (Basic Linear Algebra Subprograms) library to compute the near field part. This block version has improved the speed of this part of the computation of up to 40%. The far field part should also be speeded up thanks to a matrix formulation of the FMM.

The next work is the parallelization of the code: this, combined with an implementation of the periodic boundary conditions, will enable us to compare our work to available software for real study cases in molecular dynamics.

6.5. Parallel algorithms for heterogeneous platforms

Participants: Olivier Beaumont, Pierre Ramet, Jean Roman.

Key words: *parallel algorithms, heterogeneous platforms, master slave tasking, scheduling, divisible tasks, communication/computation overlapping, collective communications.*

As already mentioned in section, makespan minimization turns out to be very difficult, even for simple homogeneous processors and links. Our objective is to lower the ambition of makespan minimization in order to build efficient scheduling algorithms for more realistic platform models. In our works, we usually adopt the so-called “one-port with overlap model”, where a processor can simultaneously send one message, receive one message and process one task and contentions over communication links are taken into account. This requires a fine knowledge of the topology of the platform, but recently, some tools (like ENV and AINEM) have been designed to build such platform models. An idea to circumvent the difficulty of makespan minimization is to lower the ambition of the scheduling objective. Instead of aiming at the absolute minimization of the execution time, why not consider asymptotic optimality ? After all, the number of tasks to be executed on the computing platform is expected to be very large: otherwise why deploy the corresponding application on computational grids ? This approach has been pioneered by Bertsimas and Gamarnik. The dramatic simplification of steady-state scheduling is to concentrate on steady-state operations ! The scheduling problem is relaxed in many ways. Initialization and clean-up phases are neglected. The initial integer formulation is replaced by a continuous or rational formulation. The precise ordering and allocation of tasks and messages are not required, at least in the first step. The main idea is to characterize the activity of each resource during each time-unit: which (rational) fraction is spent computing, which is spent receiving or sending to which neighbor. Such activity variables are gathered into a linear program, which includes conservation laws that characterize the global behavior of the system.

This approach has been applied with success to many scheduling problems. We have first considered very simple application models, such as master-slave tasking, where a processor initially holds all the data [3][5], and the makespan minimization counterpart has been studied. Generalizations, when some parallelism can be extracted within tasks, has been considered [16][31], and the case of divisible tasks (perfect parallel tasks that can be arbitrarily divided) has been addressed [17][30].

We have also studied the problem of communication/computation overlapping for pipelined communications [32]. This corresponds to a do-dopar loops whereas master slave tasking corresponds to dopar-do loops. The use of automatic parallelization techniques in order to obtain a canonical representation of loops, and then to derive both mapping and scheduling of tasks is one of the perspectives of this work.

Recently, we have applied steady-state techniques to collective communication schemes, such as scatters, broadcasts, parallel prefix and multicasts. We have derived polynomial algorithms for broadcasts and scatters [29], but multicasts and parallel prefix have been shown NP-Hard. Another perspective of our work is to understand the limitations of steady state techniques, with respect to these new NP-Hardness results.

6.6. Computational steering environment for distributed numerical simulations

Participants: Serge Chaumette, Olivier Coulaud, Michaël Dussère, Aurélien Esnard, Guillaume Latu, Boniface Nkonga, Jean Roman.

Key words: *numerical simulation, computational steering, scientific visualization, interaction, coupling, data redistribution.*

EPSN is a software environment for the coupling of numerical simulations with visualization tools such as OpenInventor, VTK, AVS, etc. This environment mainly allows the control, the data exploration and the modification of parameters for numerical simulations involving an iterative process.

One can sum up the EPSN principles with three keypoints established during this year and presented at PARCO 2003 [6] and RenPar 15 [33] conferences.

- First, we propose a coupling model based on source code instrumentation of both the simulation and the visualization programs, as it allows fine grain steering functionalities and achieves runtime performances.
- In order to be as generic as possible, we introduce an abstract model of steerable simulations, described in XML. This abstraction allows us to build steering clients independently of a given simulation. We have validated this point by developing generic clients, which enable the end-user to interact with any EPSN simulation through simple numerical datasheets or by using basic visualization plug-ins for fields, particles or meshes.
- Regarding the communication infrastructure, we advocate the choice of CORBA as a coupling technology, because it allows portability, interoperability and network transparency. Preliminary experiments on sequential simulation have shown that the performances of the EPSN library on data transfers are quite as good as TCP/IP's. We have also demonstrated the overlapping capabilities provided by the EPSN infrastructure. Further experiments should be carried out with parallel simulation in order to validate the scalability of our prototype.

Finally, we have validated the use of EPSN with two real scientific simulations: the parallel fluid flow simulation software FLUIDBOX and the molecular docking application FINGERPRINT.

The developments of EPSN are now oriented to the coupling of parallel and distributed simulations with parallel visualization tools (e.g. Paraview). Until now, we have only considered the coupling with sequential clients because of the inherent problem of data redistribution. In this context, we have developed a $M \times N$ redistribution library based on CORBA, which is called RedEPSN. We have already validated RedEPSN for data fields with generic rectilinear distribution, like those classically used in HPF. On-going works will propose redistribution methods for particles and meshes. An article that will present in detail the RedEPSN library is currently under writing.

7. Contracts and Grants with Industry

7.1. CEA contract for scientific expertise in high performance computing

Participants: Olivier Coulaud, François Pellegrini, Pierre Ramet, Jean Roman.

This contract consists of expertise tasks concerning high performance aspects for some industrial CEA/CESTA parallel codes.

7.2. CEA research and development contracts

- Distributive schemes for unstructured meshes, steady and unsteady problems (Rémi Abgrall, Mohamed Mezine);
- Numerical aeroacoustic (Rémi Abgrall, Séverin Marret, Mohamed Mezine);
- Parallel resolution of multifluid flows (Benjamin Braconnier, Boniface Nkonga);
- Numerical simulation of compressible multifluid flows (Rémi Abgrall, Michaël Papin);
- Design and optimization of a high performance parallel code for dislocation dynamics (Sébastien Christy, Olivier Coulaud, François Pellegrini, Jean Roman);
- Feasibility study of new direct-iterative hybrid solvers for very large complex sparse systems (Pascal Hénon, Dimitri Lecas, François Pellegrini, Pierre Ramet, Jean Roman);
- Development of a parallel finite element assembly algorithm based on the hybrid 1D/2D distribution of blocks of the PASTIX parallel solver (Sébastien Christy, Pascal Hénon, Pierre Ramet, Jean Roman);
- Numerical simulation of crack propagation in silica glass by coupling molecular dynamics and elasticity methods (Guillaume Anciaux, Olivier Coulaud, Jean Roman).

8. Other Grants and Activities

8.1. Regional initiatives

8.1.1. “Réseau Dirigéable”

Participants: Rémi Abgrall, Christophe Berthon, Pascal Hénon, Mohamed Mézine, Boniface Nkonga, Pierre Ramet, Jean Roman.

Grant: Conseil Régional d’Aquitaine

Dates: 2003 – 2005

Partners: Université de Pau et des Pays d’Adour

Overview: The main goal of this program is to create a new type of airship in order to transport very big load. Actually, several projects are considered and numerical simulations are needed to propose a feasibility study. We have some years of experience in the framework of the usual aerodynamic problems but in the context of the airship, the situation turns out to be very distinct. Indeed, we have to simulate a compressible flow with a low speed (50m/s) in a large domain (larger than 300m). In this framework, we don’t know the main characteristic of the flow and, in a numerical point of view, we have to deal with a low Mach number flow which involves difficulties to obtain the numerical convergence of the simulation. In collaboration with the *Université de Pau et des Pays d’Adour*, this program has begun in 2003 and a first financial contribution of the *Région Aquitaine* has financed the post-doc of Mohamed Mézine. This first study is devoted to an extension of the residual distributive schemes in the framework of the airship.

8.2. National initiatives

8.2.1. *RedGrid: “redistribution of large data on computational grids”*

Participants: Olivier Beaumont, Olivier Coulaud, Aurélien Esnard, Pierre Ramet, Jean Roman.

Grant: INRIA New Investigation Grant

Dates: 2003 – 2004

Partners: GRAAL (INRIA Rhône-Alpes, leader of the project), PARIS (INRIA IRISA), Résédas (INRIA LORIA)

Overview: The objective of RedGrid action is to develop for a CORBA environment efficient redistribution algorithms for large data sets on grids. The application domain is the computational steering of distributed numerical simulations.

Web: <http://graal.ens-lyon.fr/~desprez/REDGRID/>

8.2.2. *EPSN: “a computational steering environment for distributed numerical simulations”*

Participants: Serge Chaumette, Olivier Coulaud (leader of the project), Michaël Dussère, Aurélien Esnard, Guillaume Latu, Boniface Nkonga, Jean Roman.

Grant: ACI GRID (“Action Concertée Incitative Globalisation des Ressources Informatiques et des Données” – French Ministry of Research)

Dates: 2002 – 2004

Partners: IPARLA (INRIA Futurs), SOD (LaBRI), IECB (Bordeaux 1, Bordeaux 2, CNRS, INSERM), SRSMC (CNRS), SMEL (Montpellier 2)

Overview: This project aims at conceiving a framework enabling the steering of distributed and parallel numerical simulations through visualisation or virtual reality environments. We are focusing on the design and implementation of the framework including data redistribution and parallel visualisation client.

Web: http://www.labri.fr/epsn/accueil_eng.html

8.2.3. *TLSE: “test for large systems of equations”*

Participants: Patrick Amestoy, Pascal Hénon, Dimitri Lecas, François Pellegrini, Pierre Ramet, Jean Roman.

Grant: ACI GRID (“Action Concertée Incitative Globalisation des Ressources Informatiques et des Données” – French Ministry of Research)

Dates: 2002 – 2004

Partners: CERFACS, ENSEEIHT – IRIT (leader of project), GRAAL (INRIA Rhône-Alpes), CEA, CNES, EADS, IFP

Overview: The main objective of TLSE is to design and to develop an expertise platform for sparse linear algebra by using the grid technology. There has been much joint work over many years on sparse matrix software between CERFACS, ENSEEIHT-IRIT, Rutherford Appleton Laboratory, LaBRI, LIP-ENSL, Parallab, University of Florida, Berkeley, and other collaborators. This has given rise to the production of several software packages that are available to the scientific community. The goal of the Project is to design an expert site that uses the accumulated expertise just mentioned and provides a one-stop shop for potential users of sparse codes. The user may want to interrogate our databases for information or references on sparse matrix work or may want actual statistics from runs of sparse software on his or her problem. The site will provide an easy access to the tools and will allow comparative analysis of these packages on a user-submitted problem or on particular matrices in the matrix collection also available on the site.

Web: <http://www.enseeiht.fr/lima/tlse/>

8.2.4. *“GRID5000”*

Participants: Olivier Coulaud, Michaël Dussère, Aurélien Esnard, Pascal Hénon, Pierre Ramet, Jean Roman.

Grant: ACI GRID (“Action Concertée Incitative Globalisation des Ressources Informatiques et des Données” – French Ministry of Research)

Dates: 2003 – 2005

Partners: RunTime (INRIA Futurs), IPARLA (INRIA Futurs), SOD (LaBRI)

Overview: The main objectives of GRID5000 is to deploy, to manage and to use a very large PC grid on the french territory. In this project, we focus on the algorithmic aspects and on the development of complex applications on such an heterogeneous grid. This work is complementary with the research works of EPSN project.

8.3. International initiatives

8.3.1. NSF - INRIA project: “robust parallel preconditioning methods”

Participants: Patrick Amestoy, Pascal Hénon, Dimitri Lecas, François Pellegrini, Pierre Ramet, Jean Roman.

Grant: NSF and INRIA

Dates: april 2001 – april 2004

Partners: Yousef Saad/Minneapolis, Randall Bramley/Indiana, Esmond Ng/Berkeley, Maria Sosenkina/Minneapolis, John Gilbert/Santa Barbara, Bernard Philippe (IRISA), Jocelyne Ehrel (IRISA), Jean-Yves L’Excellent (INRIA Rhône-Alpes)

Overview: This project focuses on algorithmic and experimental comparaisons between the parallel solvers developed by the participants, and on different solutions to design efficient parallel hybrid direct-iterative solvers. Yousef Saad has visited the SCALAPPLIX project in Bordeaux several times. Pascal Hénon, Pierre Ramet and Jean Roman have stayed in Minneapolis for 15 days during april 2003.

9. Dissemination

9.1. Participation to the Scientific Community

Rémi Abgrall is scientific editor of the international journals “Math. Model. Numer. Anal.” and “Comp. and Fluids”, and member of the scientific committee of the international conference ICCFD.

Olivier Coulaud is member of the scientific committee of the international conference VECPAR’04.

Jean Roman is leadership of the topic “Algorithmics and High Performance Applications” for the national initiative GRID2.

9.2. Teaching

In complement of the normal teaching activity of the university members and of ENSEIRB members, Olivier Coulaud and Pascal Hénon teach at ENSEIRB (computer science) and MatMéca (applied mathematics) engineering schools.

10. Bibliography

Major publications by the team in recent years

- [1] R. ABGRALL, B. NKONGA, R. SAUREL. *Efficient Numerical approximation of Compressible Multi-material flow for unstructured meshes.* in « Comp. and Fluids », number 4, volume 32, 2003, pages 571-605.
- [2] P. R. AMESTOY, I. S. DUFF, S. PRALET, C. VÖMEL. *Adapting a parallel sparse direct solver to architectures with clusters of SMPs.* in « Parallel Computing », number 11-12, volume 29, 2003, pages 1645-1668.
- [3] C. BANINO, O. BEAUMONT, L. CARTER, J. FERRANTE, A. LEGRAND, Y. ROBERT. *Scheduling strategies for master-slave tasking on heterogeneous processor platforms.* in « IEEE Trans. Parallel Distributed Systems », 2003, to appear.

- [4] M. BAUDIN, C. BERTHON, F. COQUEL, R. MASSON, H. TRAN. *A relaxation method for two-phase flow models with hydrodynamic closure law*. in « Numer. Math. », 2003, to appear.
- [5] O. BEAUMONT, A. LEGRAND, Y. ROBERT. *The master-slave paradigm with heterogeneous processors*. in « IEEE Trans. Parallel Distributed Systems », number 9, volume 14, 2003, pages 897-908.
- [6] O. COULAUD, M. DUSSÈRE, A. ESNARD. *Toward a Computational Steering Environment based on CORBA*. in « ParCo 2003, Dresden », North Holland / Elsevier, September, 2003, to appear.
- [7] O. COULAUD, M. DUSSÈRE, P. HÉNON, E. LEFEBVRE, J. ROMAN. *Optimization of a kinetic laser-plasma interaction code for large parallel systems*. in « Parallel Computing », number 9, volume 29, 2003, pages 1175–1190.
- [8] P. HÉNON, F. PELLEGRINI, P. RAMET, J. ROMAN. *An efficient hybrid MPI/Thread implementation on a network of SMP nodes for the parallel sparse direct solver PaStix: ordering / scheduling / memory management / out-of-core issues, and application to preconditioning*. in « Sparse Days and Grid Computing, Saint Girons, France », June, 2003.
- [9] P. HÉNON, Y. SAAD. *A Parallel ILU factorization based on a Hierarchical Interface Decomposition algorithm*. in « International SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Napa Valley, USA », October, 2003.
- [10] M. LANGLAIS, G. LATU, J. ROMAN, P. SILAN. *Performance analysis and qualitative results of an efficient parallel stochastic simulator for a marine host-parasite system*. in « Concurrency & Computation: Practice and Experience », number 11-12, volume 15, 2003, pages 1133-1150.

Articles in referred journals and book chapters

- [11] R. ABGRALL. *Numerical discretisation of boundary conditions for first order Hamilton Jacobi equations*. in « SIAM J. Numer. Anal. », 2003, to appear.
- [12] R. ABGRALL, M. MEZINE. *Construction of second order accurate monotone and stable residual distribution schemes for unsteady flow problems..* in « J. Comp. Phys. », number 1, volume 188, 2003, pages 16-55.
- [13] R. ABGRALL, P. ROE. *High order fluctuation schemes on triangular meshes*. in « J. Sci. Comput. », 2003, to appear.
- [14] R. ABGRALL, R. SAUREL. *Discrete equations for physical and numerical compressible multiphase mixtures*. in « J. Comp. Phys. », number 2, volume 186, 2003, pages 361–396.
- [15] P. R. AMESTOY, I. S. DUFF, J.-Y. L'EXCELLENT, X. S. LI. *Impact of the Implementation of MPI Point-to-Point Communications on the Performance of Two General Sparse Solvers*. in « Parallel Computing », number 7, volume 29, 2003, pages 833–847.
- [16] O. BEAUMONT, A. LEGRAND, L. MARCHAL, Y. ROBERT. *Scheduling strategies for mixed data and task parallelism on heterogeneous clusters*. in « Parallel Processing Letters », number 2, volume 13, 2003.

- [17] O. BEAUMONT, A. LEGRAND, Y. ROBERT. *Scheduling divisible workloads on heterogeneous platforms*. in « Parallel Computing », volume 29, 2003, pages 1121-1152.
- [18] C. BERTHON, D. REIGNIER. *Nonlinear projection methods to approximate combustion turbulent flows*. 2003, to appear.
- [19] C. BERTHON, D. REIGNIER. *An approximate nonlinear projection scheme for a combustion model*. in « M2AN, Math. Model. Numer. Anal. », number 3, volume 37, 2003, pages 451-478.
- [20] G. BILLET, R. ABGRALL. *An adaptive shock-capturing algorithm for solving unsteady reactive flows*. in « Comp. and Fluids », 2003, to appear.
- [21] T. BOUCHERES, A. BOURGEADE, T. COLIN, B. NKONGA, B. TEXIER. *Study of a mathematical model for stimulated Raman scattering*. in « Math. Models Methods Appl. Sci. », 2003, to appear.
- [22] A. BOURGEADE, B. NKONGA. *Dynamic load balancing computation of pulses propagation in a nonlinear medium*. in « J. of Supercomputing, Special Issue », 2003, to appear.
- [23] T. COLIN, B. NKONGA. *Numerical model for light interaction with two-level atoms medium*. in « Physica D », 2003, to appear.
- [24] S. GAUCEL, M. LANGLAIS. *Some mathematical problems arising in heterogeneous insular ecological models*. in « Rev. R. Acad. Cien. Serie A. Math. », number 3, volume 96, 2003, pages 389-400.
- [25] F. HUOT, A. GHIZZO, P. BERTRAND, E. SONNENDRUCKER, O. COULAUD. *Instability of the time splitting scheme for the one-dimensional and relativistic Vlasov-Maxwell system*. in « Journal of Computational Physics », number 2, volume 185, 2003, pages 512-531.
- [26] R. LOUBÈRE, J. OVADIA, R. ABGRALL. *A Lagrangian discontinuous Galerkin type method on unstructured meshes to solve hydrodynamic problems*. in « Int. J. Numer. Meth. in Fluids », 2003, to appear.
- [27] P. SILAN, H. CALTRAN, G. LATU. *Taxonomy, ecology and evolution of metazoan parasites*. Presses Universitaires de Perpignan, 2003, chapter - Ecologie et dynamique des populations de Monogènes, ectoparasites de Téléostéens marins : approche et contribution montpelliérainne, C. Combes & J. Jourdané editors.

Publications in Conferences and Workshops

- [28] R. ABGRALL, M. MEZINE. *A compact residual scheme for unsteady compressible flow problems*. in « Second International Conference in Computational Fluid Dynamics », series Lecture Notes in Physics, Springer verlag, S. SRINIVAS, editor, 2003.
- [29] O. BEAUMONT, A. LEGRAND, L. MARCHAL, Y. ROBERT. *Pipelining Broadcasts on Heterogeneous Platforms*. in « IPDPS'2004 », IEEE Computer Society Press, 2003, to appear.
- [30] O. BEAUMONT, A. LEGRAND, Y. ROBERT. *Optimal algorithms for scheduling divisible workloads on heterogeneous systems*. in « HCW'2003, the 12th Heterogeneous Computing Workshop », IEEE Computer

Society Press, 2003.

- [31] O. BEAUMONT, A. LEGRAND, Y. ROBERT. *Scheduling strategies for mixed data and task parallelism on heterogeneous clusters and grids*. in « PDP'2003, 11th Euromicro Workshop on Parallel, Distributed and Network-based Processing », IEEE Computer Society Press, pages 209-216, 2003.
- [32] O. BEAUMONT, P. RAMET, J. ROMAN. *Asymptotically optimal algorithm for Laplace task graphs on heterogeneous platforms*. in « Fifth International Conference on Parallel Processing and Applied Mathematics, Workshop HeteroPar, Czestochowa, Pologne », 2003.
- [33] M. DUSSÈRE, A. ESNARD. *Vers le pilotage des simulations numériques sur la grille*. in « 15ème Rencontres francophones du parallélisme », pages 196–203, 2003.
- [34] D. GOUDIN, P. HÉNON, M. MANDALLENA, K. MER, F. PELLEGRINI, P. RAMET, J. ROMAN, J.-J. PESQUÉ. *Outils numériques parallèles pour la résolution de très grands problèmes d'électromagnétisme*. in « Séminaire sur l'Algorithmique Numérique Appliquée aux Problèmes Industriels, Calais, France », May, 2003.
- [35] P. HÉNON, F. PELLEGRINI, P. RAMET, J. ROMAN. *Towards High Performance Hybrid Direct-Iterative Solvers for Large Sparse Systems*. in « International SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Napa Valley, USA », October, 2003.
- [36] P. HÉNON, F. PELLEGRINI, P. RAMET, J. ROMAN, Y. SAAD. *High Performance Complete and Incomplete Factorizations for Very Large Sparse Systems by using Scotch and PaStiX softwares*. in « Eleventh SIAM Conference on Parallel Processing for Scientific Computing, San Francisco, Californie, USA », February, 2003, to appear.
- [37] P. HÉNON, P. RAMET, J. ROMAN. *Efficient algorithms for direct resolution of large sparse system on clusters of SMP nodes*. in « SIAM Conference on Applied Linear Algebra, Williamsburg, Virginie, USA », July, 2003.
- [38] P. HÉNON, Y. SAAD. *PHIDAL: A Parallel Hierarchical Interface Decomposition Algorithm for solving sparse linear systems*. in « Sparse Days and Grid Computing, Saint Girons, France », June, 2003.
- [39] F. PELLEGRINI, D. GOUDIN. *Using the native mesh partitioning capabilities of Scotch 4.0 in a parallel industrial electromagnetics code*. in « Eleventh SIAM Conference on Parallel Processing for Scientific Computing, San Francisco, Californie, USA », February, 2003, to appear.

Internal Reports

- [40] P. R. AMESTOY, X. S. LI, E. NG. *Diagonal Markowitz Scheme With Local Symmetrization*. Technical report, number RT/APO/03/5, ENSEEIHT-IRIT, October, 2003, Also appeared as Lawrence Berkeley Lab report LBNL-53854.
- [41] T. COLIN, B. NKONGA. *Multiscale numerical method for nonlinear Maxwell equation..* Preprint, number No LRC-03.17, MAB, 2003.
- [42] P. HÉNON, D. LECAS, P. RAMET, J. ROMAN. *Amélioration et Extension du Solveur Direct Parallèle pour Grandes Matrices Creuses du CESTA*. Technical report, C.E.A. / C.E.S.T.A, 2003, Rapport Final.

Miscellaneous

- [43] R. ABGRALL, M. MEZINE. *Residual distribution schemes for steady problems*. 33rd Computational Fluid dynamics course, von Karman Institute, april, 2003.
- [44] C. BERTHON, M. BREUSS, M.-O. TITEUX. *A relaxation scheme to approximate pressureless gas dynamic*. submitted to Num. Meth. Partial Differential Equations, 2003.
- [45] C. BERTHON, B. NKONGA. *Multifluid numerical approximations based on multipressure formulation..* submittted to Comp. and Fluids, 2003.
- [46] C. BERTHON, B. NKONGA. *Numerical model of a compressible multifluid fluctuating flow*. submitted to Int. J. Finite Volume, 2003.
- [47] M. MEZINE, M. RICCHIUTO, R. ABGRALL, H. DECONINCK. *Monotone and stable residual distribution schemes on prismatic space-time elements for unsteady conservation laws*. 33rd Computational Fluid dynamics course, von Karman Institute, april, 2003.
- [48] F. PELLEGRINI. *Native mesh ordering with Scotch 4.0*. submitted to VECPAR'04, Valencia, Spain, 2003.