INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Team HiePACS

High-End Parallel Algorithms for Challenging Numerical Simulations

Bordeaux - Sud-Ouest

Theme : Distributed and High Performance Computing

Table of contents

HIEPACS is an INRIA Team joint with University of Bordeaux and CNRS (LaBRI, UMR 5800) and is a Research Initiative of the joint Laboratory INRIA-CERFACS on High-Performance Computing (http://hiepacs.bordeaux.inria.fr/).

1. Team

Research Scientists

Jean Roman [INRIA, Team Leader, Senior Researcher (DR on secondment), Professor at IPB, HdR]
Emmanuel Agullo [INRIA, Junior Researcher (CR)]
Olivier Coulaud [INRIA, Senior Researcher (DR)]
Luc Giraud [INRIA, Senior Researcher (DR), HdR]
Xavier Vasseur [CERFACS, ALGO Project, Senior Scientist]

Faculty Members

Aurélien Esnard [University of Bordeaux, Associate Professor (MdC)]
Abdou Guermouche [University of Bordeaux, Associate Professor (MdC)]

External Collaborators

Iain Duff [CERFACS, ALGO Project, Senior Scientist , HIEPACS **Scientific Advisor**]
Pierre Fortin [Paris 6 University, LIP6, Assistant Professor (MdC)]
Guillaume Latu [Strasbourg University, LSIIT, Assistant Professor (MdC), on secondment at CEA Cadarache]

Technical Staff

Mohamed Abdoul Asize [CNRS, funding from ANR ARA CIS NOSSI, since Nov. 2010]
Damien Genet [CNRS, funding from ANR ARA CIS NOSSI, until Sep. 2010]
Yohan Lee-Tin-Yien [INRIA, funding from ADT ParScaLi, since Dec. 2009]

PhD Students

Rached Abdelkhalek [TOTAL, CIFRE, since Jan. 2008]
Mathieu Chanaud [INRIA, funding from INRIA and CEA, since Dec. 2007]
Yohann Dudouit [CERFACS, funding from TOTAL, since Oct. 2010]
Fabrice Dupros [BRGM, funding from BRGM and ANR CIGC NUMASIS, since Feb. 2007]
Pablo Salas Medina [INRIA, funding from Europe FP7/ITN (Marie Curie) MyPlanet, since Jun. 2010]
Jérôme Soumagne [CSCS, funding from Europe FP7/ICT/FET NextMuSE STREP since Apr. 2009]
Clément Vuchener [University of Bordeaux, funding from French Research Ministry, since Sep. 2010]

Post-Doctoral Fellows

Mikko Byckling [CERFACS, ALGO Project, since Nov. 2010]
Pavel Jiranek [CERFACS, ALGO Project, until Sep. 2010]
Peter Koval [CNRS, funding from ANR ARA CIS NOSSI, until Jul. 2010]

Administrative Assistant

Barta Beneddine [INRIA]

Other

Clément Vuchener [INRIA, Internship from Feb. to Sep. 2010]

2. Overall Objectives

2.1. Introduction

Over the last few decades, there have been innumerable science, engineering and societal breakthroughs enabled by the development of high performance computing (HPC) applications, algorithms and architectures. These powerful tools have provided researchers with the ability to computationally find efficient solutions for some of the most challenging scientific questions and problems in medicine and biology, climatology, nanotechnology, energy and environment. It is admitted today that *numerical simulation is the third pillar for*

the development of scientific discovery at the same level as theory and experimentation. Numerous reports and papers also confirmed that very high performance simulation will open new opportunities not only for research but also for a large spectrum of industrial sectors (see for example the documents available on the web link http://www.sc.doe.gov/ascr/ProgramDocuments/ProgDocs.html).

An important force which has continued to drive HPC has been to focus on frontier milestones which consist in technical goals that symbolize the next stage of progress in the field. In the 1990s, the HPC community sought to achieve computing at a teraflop rate and currently we are able to compute on the first leading architectures at a petaflop rate. Generalist petaflop supercomputers are likely to be available in 2010-2012 and some communities are already in the early stages of thinking about what computing at the exaflop level would be like.

For application codes to sustain a petaflop and more in the next few years, hundreds of thousands of processor cores or more will be needed, regardless of processor technology. Currently, a few HPC simulation codes easily scale to this regime and major code development efforts are critical to achieve the potential of these new systems. Scaling to a petaflop and more will involve improving physical models, mathematical modelling, super scalable algorithms that will require paying particular attention to acquisition, management and vizualization of huge amounts of scientific data.

In this context, the purpose of the `HiePACS` project is to perform efficiently frontier simulations arising from challenging research and industrial *multiscale* applications. The solution of these challenging problems require a multidisciplinary approach involving applied mathematics, computational and computer sciences. In applied mathematics, it essentially involves advanced numerical schemes. In computational science, it involves massively parallel computing and the design of highly scalable algorithms and codes to be executed on future petaflop (and beyond) platforms. Through this approach, `HiePACS` intends to contribute to all steps that go from the design of new high-performance more scalable, robust and more accurate numerical schemes to the optimized implementations of the associated algorithms and codes on very high performance supercomputers. This research will be conduced on close collaboration in particular with European and US initiatives or projects such as PRACE (Partnership for Advanced Computing in Europe – http://www.prace-project.eu/) or IESP (International Exascale Software Project – http://icl.cs.utk.edu/iesp/Main_Page).

In order to address these research challenges, some of the researchers of the former `ScAlApplix` INRIA Project-Team and some researchers of the Parallel Algorithms Project from CERFACS have joined `HiePACS` in the framework of the joint INRIA-CERFACS Laboratory on High Performance Computing. The director of the joint laboratory is J. Roman while I.S. Duff is the senior scientific advisor. `HiePACS` is the first research initiative of this joint Laboratory. Because of his strong involvement in RAL and his oustanding action in other main initiatives in UK and wordwide, I.S. Duff appears as an external collaborator of the `HiePACS` project while his contribution will be significant. There are two other external collaborators. Namely, P. Fortin who will be mainly involved in the activities related to the parallel fast multipole development and G. Latu who will contribute to research actions related to the emerging new computing facilities.

The methodological part of `HiePACS` covers several topics. First, we address generic studies concerning massively parallel computing, the design of high-end performance algorithms and software to be executed on future petaflop (and beyond) platforms. Next, several research prospectives in scalable parallel linear algebra techniques are adressed, in particular hybrid approaches for large linear systems. Then we consider research plans for N-body interaction computations based on efficient parallel fast multipole methods and finally, we adress research tracks related to the algorithmic challenges for complex code couplings in multiscale simulations.

Currently, we have one major multiscale application that is in *material physics*. We contribute to all steps of the design of the parallel simulation tool. More precisely, our applied mathematics skill will contribute to the modelling and our advanced numerical schemes will help in the design and efficient software implementation for very large parallel multiscale simulations. Moreover, the robustness and efficiency of our algorithmic research in linear algebra are validated through industrial and academic collaborations with different partners involved in various application fields.

Our high performance software packages are integrated in several academic or industrial complex codes and are validated on very large scale simulations. For all our software developments, we use first the various (very) large parallel platforms available through CERFACS and GENCI in France (CCRT, CINES and IDRIS Computational Centers), and next the high-end parallel platforms that will be available via European and US initiatives or projects such that PRACE.

2.2. Highlights of the year

- A France-Berkeley fund has been granted jointly with the Computational Research Division, Lawrence Berkeley National Laboratory. It is entitled "Scalable Hybrid Solvers for Large Sparse Linear Systems of Equations on Petascale Computing Architectures" (http://fbf.berkeley.edu/Grantee2010.html).
- The `HiePACS` members are strongly and actively involved in the organization of two forthcoming international conferences to be held in Bordeaux next year, namely Preconditiong 2011 and EuroPar 2011, both in the organizing and scientific committees. For this latter large event, the members are involved in the local organization committee and as co-chair and local chairs for two topics (i.e., high performance and scientific applications, parallel numerical algorithms).

3. Scientific Foundations

3.1. Introduction

The methodological component of `HiePACS` concerns the expertise for the design as well as the efficient and scalable implementation of highly parallel numerical algorithms to perform frontier simulations. In order to address these computational challenges a hierarchical organization of the research is considered. In this bottom-up approach, we first consider in Section 3.2 generic topics concerning high performance computational science. The activities described in this section are transversal to the overall project and its outcome will support all the other research activities at various levels in order to ensure the parallel scalability of the algorithms. The aim of this activity is not to study general purpose solution but rather to address these problems in close relation with specialists of the field in order to adapt and tune advanced approaches in our algorithmic designs. The next activity, described in Section 3.3, is related to the study of parallel linear algebra techniques that currently appear as promising approaches to tackle huge problems on millions of cores. We highlight the linear problems (linear systems or eigenproblems) because they are in many large scale applications the main computational intensive numerical kernels and often the main performance bottleneck. These parallel numerical techniques will be the basis of both academic and industrial collaborations described in Section 4.2 and Section 4.3, but will also be closely related to some functionalities developed in the parallel fast multipole activity described in Section 3.4. Finally, as the accuracy of the physical models increases, there is a real need to go for parallel efficient algorithm implementation for multiphysics and multiscale modelling in particular in the context of code coupling. The challenges associated with this activity will be addressed in the framework of the activity described in Section 3.5.

3.2. High-performance computing on next generation architectures

Participants: Rached Abdelkhalek, Emmanuel Agullo, Olivier Coulaud, Yohann Dudouit, Iain Duff, Pierre Fortin, Luc Giraud, Abdou Guermouche, Guillaume Latu, Jean Roman.

The research directions proposed in `HiePACS` are strongly influenced by both the applications we are studying and the architectures that we target (i.e., massively parallel architectures, ...). Our main goal is to study the methodology needed to efficiently exploit the new generation of high-performance computers with all the constraints that it induces. To achieve this high-performance with complex applications we have to study both algorithmic problems and the impact of the architectures on the algorithm design.

From the application point of view, the project will be interested in multiresolution, multiscale and hierarchical approaches which lead to multi-level parallelism schemes. This hierarchical parallelism approach is necessary to achieve good performance and high-scalability on modern massively parallel platforms. In this context, more specific algorithmic problems are very important to obtain high performance. Indeed, the kind of applications we are interested in are often based on data redistribution for example (e.g. code coupling applications). This well-known issue becomes very challenging with the increase of both the number of computational nodes and the amount of data. Thus, we have both to study new algorithms and to adapt the existing ones. In addition, some issues like task scheduling have to be restudied in this new context. It is important to note that the work done in this area will be applied for example in the context of code coupling (see Section 3.5).

Considering the complexity of modern architectures like massively parallel architectures (i.e., Blue Gene-like platforms) or new generation heterogeneous multicore architectures, task scheduling becomes a challenging problem which is central to obtain a high efficiency. Of course, this work requires the use/design of scheduling algorithms and models specifically to tackle our target problem. This has to be done in collaboration with our colleagues from the scheduling community like for example O. Beaumont (INRIA CEPAGE Project-Team). It is important to note that this topic is strongly linked to the underlying programming model. Indeed, considering multicore architectures, it has appeared, in the last five years, that the best programming model is an approach mixing multi-threading within computational nodes and message passing between them. In the last five years, a lot of work has been developed in the high-performance computing community to understand what is critic to efficiently exploit massively multicore platforms that will appear in the near future. It appeared that the key for the performance is firstly the grain of computations. Indeed, in such platforms the grain of the parallelism must be small so that we can feed all the processors with a sufficient amount of work. It is thus very crucial for us to design new high performance tools for scientific computing in this new context. This will be done in the context of our solvers, for example, to adapt to this new parallel scheme. Secondly, the larger the number of cores inside a node, the more complex the memory hierarchy. This remark impacts the behaviour of the algorithms within the node. Indeed, on this kind of platforms, NUMA effects will be more and more problematic. Thus, it is very important to study and design data-aware algorithms which take into account the affinity between computational threads and the data they access. This is particularly important in the context of our high-performance tools. Note that this work has to be based on an intelligent cooperative underlying runtime (like the `marcel` thread library developed by the INRIA RUNTIME Project-Team) which allows a fine management of data distribution within a node.

Another very important issue concerns high-performance computing using “heterogeneous” resources within a computational node. Indeed, with the emergence of the `GPU` and the use of more specific co-processors (like clearspeed cards, ...), it is important for our algorithms to efficiently exploit these new kind of architectures. To adapt our algorithms and tools to these accelerators, we need to identify what can be done on the `GPU` for example and what cannot. Note that recent results in the field have shown the interest of using both regular cores and `GPU` to perform computations. Note also that in opposition to the case of the parallelism granularity needed by regular multicore architectures, `GPU` requires coarser grain parallelism. Thus, making both `GPU` and regular cores work all together will lead to two types of tasks in terms of granularity. This represents a challenging problem especially in terms of scheduling. Our final goal would be to have high performance solvers and tools which can efficiently run on all these types of complex architectures by exploiting all the resources of the platform (even if they are heterogeneous).

In order to achieve an advanced knowledge concerning the design of efficient computational kernels to be used on our high performance algorithms and codes, we will develop research activities first on regular frameworks before extending them to more irregular and complex situations. In particular, we will work first on optimized dense linear algebra kernels and we will use them in our more complicated hybrid solvers for sparse linear algebra and in our fast multipole algorithms for interaction computations. In this context, we will participate to the development of those kernels in collaboration with groups specialized in dense linear algebra. In particular, we intend develop a strong collaboration with the group of Jack Dongarra at the University Of Tennessee. The objectives will be to develop dense linear algebra algorithms and libraries for multicore architectures

in the context the PLASMA project (http://icl.cs.utk.edu/plasma/) and for GPU and hybrid multicore/GPU architectures in the context of the MAGMA project (http://icl.cs.utk.edu/magma/).

The applications targeting massively parallel architectures are very sensitive to communication or I/O management schemes. This observation becomes particularly true, when we consider applications dealing with a huge amount of data like very large scale simulations that may produce petaBytes of data. Thus, in the continuation of the work we did around out-of-core extensions of our former sparse linear solvers, we will study how we can efficiently deal with this huge amount of data. Obtaining performance when relying on I/O operations or on data transfers is mainly constrained by the capacity to overlap as much as much possible these operations with computations. Another key feature is prefetching in the context of I/O intensive applications. Even, if the problem is a well-known issue which has been studied in the past decade, it remains very complex regarding the complexity of our target platforms were we already need prefetching and asynchronism to efficiently exploit the platform (this is particularly true in the case of GPU).

A more prospective objective is to study the fault tolerance in the context of large-scale scientific applications for massively parallel architectures. Indeed, with the increase of the number of computational cores per node, the probability of a hardware crash on a core is dramatically increased. This represents a crucial problem that needs to be addressed. However, we will only study it at the algorithmic/application level even if it needed lower-level mechanisms (at OS level or even hardware level). Of course, this work can be done at lower levels (at operating system) level for example but we do believe that handling faults at the application level provides more knowledge about what has to be done (at application level we know what is critical and what is not). The approach that we will follow will be based on the use of a combination of fault-tolerant implementations of the run-time environments we use (like for example FT-MPI) and an adaptation of our algorithms to try to manage this kind of faults. This topic represents a very long range objective which needs to be addressed to guaranty the robustness of our solvers and applications. In that respect, we are involved in a ANR-Blanc project entitles RESCUE jointly with two other INRIA EPI, namely GRAAL and GRAND-LARGE. The main objective of the RESCUE project is to develop new algorithmic techniques and software tools to solve the exascale resilience problem. Solving this problem implies a departure from current approaches, and calls for yet-to-be- discovered algorithms, protocols and software tools.

In the framework of an industrial collaboration with TOTAL (that funds the PhD of Yohann Dudouit), we study new scalable parallel simulation schemes and efficient parallel implementations for the solution of the elastodynamic system with local refinements.

Finally, it is important to note that the main goal of HiePACS is to design tools and algorithms that will be used within complex simulation frameworks on next-generation parallel machines. Thus, we intend with our partners to use the proposed approach in complex scientific codes and to validate them within very large scale simulations.

3.3. High performance solvers for large linear algebra problems

Participants: Emmanuel Agullo, Mikko Byckling, Mathieu Chanaud, Olivier Coulaud, Iain Duff, Fabrice Dupros, Luc Giraud, Abdou Guermouche, Pavel Jiranek, Yohan Lee-Tin-Yien, Jean Roman, Pablo Salas Medina, Xavier Vasseur.

Starting with the developments of basic linear algebra kernels tuned for various classes of computers, a significant knowledge on the basic concepts for implementations on high-performance scientific computers has been accumulated. Further knowledge has been acquired through the design of more sophisticated linear algebra algorithms fully exploiting those basic intensive computational kernels. In that context, we still look at the development of new computing platforms and their associated programming tools. This enables us to identify the possible bottlenecks of new computer architectures (memory path, various level of caches, inter processor or node network) and to propose ways to overcome them in algorithmic design. With the goal of designing efficient scalable linear algebra solvers for large scale applications, various tracks will be followed in order to investigate different complementary approaches. Sparse direct solvers have been for years the methods of choice for solving linear systems of equations, it is nowadays admitted that such approaches are

not scalable neither from a computational complexity nor from a memory view point for large problems such as those arising from the discretization of large 3D PDE problems. Although we will not contribute directly to this activity, we will use parallel sparse direct solvers as building boxes for the design of some of our parallel algorithms such as the hybrid solvers described in the sequel of this section. Our activities in that context will mainly address preconditioned Krylov subspace methods; both components, preconditioner and Krylov solvers, will be investigated.

3.3.1. Hybrid direct/iterative solvers based on algebraic domain decomposition techniques

One route to the parallel scalable solution of large sparse linear systems in parallel scientific computing is the use of hybrid methods that combine direct and iterative methods. These techniques inherit the advantages of each approach, namely the limited amount of memory and natural parallelization for the iterative component and the numerical robustness of the direct part. The general underlying ideas are not new since they have been intensively used to design domain decomposition techniques; those approaches cover a fairly large range of computing techniques for the numerical solution of partial differential equations (PDEs) in time and space. Generally speaking, it refers to the splitting of the computational domain into sub-domains with or without overlap. The splitting strategy is generally governed by various constraints/objectives but the main one is to express parallelism. The numerical properties of the PDEs to be solved are usually intensively exploited at the continuous or discrete levels to design the numerical algorithms so that the resulting specialized technique will only work for the class of linear systems associated with the targeted PDE.

In that context, we attempt to apply to general unstructured linear systems domain decomposition ideas. More precisely, we will consider numerical techniques based on a non-overlapping decomposition of the graph associated with the sparse matrices. The vertex separator, built by a graph partitioner, will define the interface variables that will be solved iteratively using a Schur complement techniques, while the variables associated with the internal sub-graphs will be handled by a sparse direct solver. Although the Schur complement system is usually more tractable than the original problem by an iterative technique, preconditioning treatment is still required. For that purpose, the algebraic additive Schwarz technique initially developed for the solution of linear systems arising from the discretization of elliptic and parabolic PDE's will be extended. Linear systems where the associated matrices are symmetric in pattern will be first studied but extension to unsymmetric matrices will be latter considered. The main focus will be on difficult problems (including non-symmetric and indefinite ones) where it is harder to prevent growth in the number of iterations with the number of subdomains when considering massively parallel platforms. In that respect, we will consider algorithms that exploit several sources and grains of parallelism to achieve high computational throughput. This activity may involve collaborations with developers of sparse direct solvers as well as with developers of run-time systems and will lead to the development to the library `MaPHyS` (see Section 5.2). Some specific aspects, such as mixed MPI-thread implementation for the computer science aspects and techniques for indefinite system for the numerical aspects will be investigated in the framework of a France Berkeley Fund project granted starting this year.

3.3.2. Full geometric multigrid method for 3D Maxwell equations

The multigrid methods are among the most promising numerical techniques to solve large linear system of equations arising from the discretization of PDE's. Their ideal scalabilities, linear growth of memory and floating-point operations with the number of unknowns, for solving elliptic equations make them very appealing for petascale computing and a lot of research works in the recent years has been devoted to the extension to other types of PDE.

In this work (Ph. D. of Mathieu Chanaud in collaboration with CEA/CESTA), we consider a full geometric multigrid solver for the solution of methodology for solving large linear systems arising from Maxwell equations discretized with unstructured first-order Nédelec elements. This solver combines a parallel sparse direct solver and full multigrid cycles. The goal of this method is to compute the solution for problems defined on fine irregular meshes with minimal overhead costs when compared to the cost of applying a classical direct solver on the coarse mesh.

The direct solver can handle linear systems with up to 100 million unknowns, but this size is limited by the computer memory, so that finer problem resolutions that often occur in practice cannot be handled by this direct solver. The aim of the new method is to provide a way to solve problems with up to 1 billion unknowns, given an input coarse mesh with up to 100 million unknowns. The input mesh defines the coarsest level. This mesh is further refined to defined the grid hierarchy, where matrix free smoothers are considered to reduce the memory consumption.

3.3.3. Linear Krylov solvers

Preconditioning is the main focus of the two activities described above. They aim at speeding up the convergence of a Krylov subspace method that is the complementary component involved in the solvers of interest for us. In that framework, we believe that various aspects deserve to be investigated; we will consider the following ones:

Preconditioned block Krylov solvers for multiple right-hand sides. In many large scientific and industrial applications, one has to solve a sequence of linear systems with several right-hand sides given simultaneously or in sequence (radar cross section calculation in electromagnetism, various source locations in seismic, parametric studies in general, ...). For “simultaneous" right-hand sides, the solvers of choice have been for years based on matrix factorizations as the factorization is performed once and simple and cheap block forward/backward substitutions are then performed. In order to effectively propose alternative to such solvers, we need to have efficient preconditioned Krylov subspace solvers. In that framework, block Krylov approaches, where the Krylov spaces associated with each right-hand sides are shared to enlarge the search space will be considered. They are not only attractive because of this numerical feature (larger search space), but also from an implementation point of view. Their block-structures exhibit nice features with respect to data locality and re-usability that comply with the memory constraint of multicore architectures. For right-hand sides available one after each other, various strategies that exploit the information available in the sequence of Krylov spaces (e.g. spectral information) will be considered that include for instance technique to perform incremental update of the preconditioner or to built augmented Krylov subspaces. Julien Langou, associated professor in the Department of Mathematical and Statistical Sciences at University of Colorado Denver, was a visiting professor for a month thanks to an INRIA grant. During his visiting period, Julien contributed to this research activity.

Flexible Krylov subspace methods with recycling techniques. In many situations, it has been observed that significant convergence improvements can be achieved in preconditioned Krylov subspace methods by enriching them with some spectral information. On the other hand effective preconditioning strategies are often designed where the preconditioner varies from one step to the next (e.g. in domain decomposition methods, when approximate solvers are considered for the interior problems, or more generally for block preconditioning technique where approximate block solution are used) so that a flexible Krylov solver is required. In that context, we intend to investigate how numerical techniques implementing subspace recycling and/or incremental preconditioning can be extended and adapted to cope with this situation of flexible preconditioning; that is, how can we numerically benefit from the preconditioning implementation flexibility.

Krylov solver for complex symmetric non-Hermitian matrices. In material physics when the absorption spectrum of a molecule due to an exterior field is computed, we have to solve for each frequency a dense linear system where the matrix depends on the frequency. The sequence of matrices are complex symmetric non-Hermitian. While a direct approach can be used for small molecules, a Krylov subspace solver must be considered for larger molecules. Typically, Lanczos-type methods are used to solve these systems but the convergence is often slow. Based on our earlier experience on preconditioning techniques for dense complex symmetric non-Hermitian linear system in electromagnetism, we are interested in designing new preconditioners for this class of material physics applications. A first track will consist in building preconditioners on sparsified approximation of the matrix as well as computing incremental updates, eg. Sherman-Morrison type, of the preconditioner when the frequency varies. This action will be developed in the framework of the research activity described in Section 4.2.

Approximate factoring of the inverse. When the matrix of a given sparse linear system of equations is known to be nonsingular, the computation of approximate factors for the inverse constitutes an algebraic approach to preconditioning. The main aim is to combine standard preconditioning ideas with sparse approximate inverse approximation to have implicitly dense approximate inverse approximations. Theory has been developed and encouraging numerical experiments have been obtained on a set of sparse matrices of small to medium size. We plan to propose a parallel implementation of the construction of the preconditioner and to investigate its efficiency on real-life problems.

Extension or modification of Krylov subspace algorithms for multicore architectures. Finally to match as much as possible to the computer architecture evolution and get as much as possible performance out of the computer, a particular attention will be paid to adapt, extend or develop numerical schemes that comply with the efficiency constraints associated with the available computers. Nowadays, multicore architectures seem to become widely used, where memory latency and bandwidth are the main bottlenecks; investigations on communication avoiding techniques will be undertaken in the framework of preconditioned Krylov subspace solvers as a general guideline for all the items mentioned above.

Eigensolvers. Many eigensolvers also rely on Krylov subspace techniques. Naturally some links exist between the Krylov subspace linear solvers and the Krylov subspace eigensolvers. We plan to study the computation of eigenvalue problems with respect to the following three different axes:

- Exploiting the link between Krylov subspace methods for linear system solution and eigensolvers, we intend to develop advanced iterative linear methods based on Krylov subspace methods that use some spectral information to build part of a subspace to be recycled, either though space augmentation or through preconditioner update. This spectral information may correspond to a certain part of the spectrum of the original large matrix or to some approximations of the eigenvalues obtained by solving a reduced eigenproblem. This technique will also be investigated in the framework of block Krylov subspace methods.
- In the framework of an FP7 Marie project (MyPlanet), we intend to study parallel robust nonlinear quadratic eigensolvers. It is a crucial question in numerous technologies like the stability and vibration analysis in classical structural mechanics. The first research action consists in enhancing the robustness of the linear eigensolver and to consider shift invert technique to tackle difficult problems out of reach with the current technique. One of the main constraint in that framework is to design matrix-free technique to limit the memory consumption of the complete solver. For the nonlinear part different approaches ranging from simple nonlinear stationary iterations to Newton's type approaches will be considered.
- In the context of the calculation of the ground state of an atomistic system, eigenvalue computation is a critical step; more accurate and more efficient parallel and scalable eigensolvers are required (see Section 4.2).

3.4. High performance Fast Multipole Method for N-body problems

Participants: Olivier Coulaud, Pierre Fortin, Luc Giraud, Jean Roman.

In most scientific computing applications considered nowadays as computational challenges (like biological and material 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. For instance, in the N-body problem arising from these application fields, we must compute all pairwise interactions among N objects (particles, lines, ...) at every timestep. Among these methods, the Fast Multipole Method (FMM) developed for gravitational potentials in astrophysics and for electrostatic (coulombic) potentials in molecular simulations solves this N-body problem for any given precision with $O(N)$ runtime complexity against $O(N^2)$ for the direct computation.

The potential field is decomposed in a near field part, directly computed, and a far field part approximated thanks to multipole and local expansions. In the former `ScAlApplix` project, we introduced a matrix formulation of the FMM that exploits the cache hierarchy on a processor through the Basic Linear Algebra Subprograms (BLAS). Moreover, we developed a parallel adaptive version of the FMM algorithm for heterogeneous particle distributions, which is very efficient on parallel clusters of SMP nodes. Finally on such computers, we developed the first hybrid MPI-thread algorithm, which enables to reach better parallel efficiency and better memory scalability. We plan to work on the following points in `HiePACS` .

3.4.1. Improvement of calculation efficiency

Nowadays, the high performance computing community is examining alternative architectures that address the limitations of modern cache-based designs. `GPU` (Graphics Processing Units) and the Cell processor have thus already been used in astrophysics and in molecular dynamics. The Fast Mutipole Method has also been implemented on `GPU`. We intend to examine the potential of using these forthcoming processors as a building block for high-end parallel computing in N-body calculations. More precisely, we want to take advantage of our specific underlying BLAS routines to obtain an efficient and easily portable FMM for these new architectures. Algorithmic issues such as dynamic load balancing among heterogeneous cores will also have to be solved in order to gather all the available computation power. This research action will be conduced on close connection with the activity described in Section 3.2.

3.4.2. Non uniform distributions

In many applications arising from material physics or astrophysics, the distribution of the data is highly non uniform and the data can grow between two time steps. As mentioned previously, we have proposed a hybrid MPI-thread algorithm to exploit the data locality within each node. We plan to further improve the load balancing for highly non uniform particle distributions with small computation grain thanks to dynamic load balancing at the thread level and thanks to a load balancing correction over several simulation time steps at the process level.

3.4.3. Fast Multipole Method for dislocation operators

The engine that we develop will be extended to new potentials arising from material physics such as those used in dislocation simulations. The interaction between dislocations is long ranged ($O(1/r)$) and anisotropic, leading to severe computational challenges for large-scale simulations. Several approaches based on the FMM or based on spatial decomposition in boxes are proposed to speed-up the computation. In dislocation codes, the calculation of the interaction forces between dislocations is still the most CPU time consuming. This computation has to be improved to obtain faster and more accurate simulations. Moreover, in such simulations, the number of dislocations grows while the phenomenon occurs and these dislocations are not uniformly distributed in the domain. This means that strategies to dynamically balance the computational load are crucial to acheive high performance.

3.4.4. Fast Multipole Method for boundary element methods

The boundary element method (BEM) is a well known solution of boundary value problems appearing in various fields of physics. With this approach, we only have to solve an integral equation on the boundary. This implies an interaction that decreases in space, but results in the solution of a dense linear system with $O(N^3)$ complexity. The FMM calculation that performs the matrix-vector product enables the use of Krylov subspace methods. Based on the parallel data distribution of the underlying octree implemented to perform the FMM, parallel preconditioners can be designed that exploit the local interaction matrices computed at the finest level of the octree. This research action will be conduced on close connection with the activity described in Section 3.3. Following our earlier experience, we plan to first consider approximate inverse preconditionners that can efficiently exploit these data structures.

3.5. Efficient algorithmics for code coupling in complex simulations

Participants: Mohamed Abdoul Asize, Olivier Coulaud, Aurélien Esnard, Damien Genet, Jean Roman, Jérôme Soumagne, Clément Vuchener.

Many important physical phenomena in material physics and climatology are inherently complex applications. They often use multi-physics or multi-scale approaches, that couple different models and codes. The key idea is to reuse available legacy codes through a coupling framework instead of merging them into a standalone application. There is typically one model per different scale or physics; and each model is implemented by a parallel code. For instance, to model a crack propagation, one uses a molecular dynamic code to represent the atomistic scale and an elasticity code using a finite element method to represent the continuum scale. Indeed, fully microscopic simulations of most domains of interest are not computationally feasible. Combining such different scales or physics are still a challenge to reach high performance and scalability. If the model aspects are often well studied, there are several open algorithmic problems, that we plan to investigate in the HIEPACS project-team.

The experience that we have acquired in the `ScAlApplix` project through the activities in crack propagation simulations with LibMultiScale and in M-by-N computational steering (coupling simulation with parallel visualization tools) with `EPSN` shows us that if the model aspect was well studied, several problems in parallel or distributed algorithms are still open and not well studied. In the context of code coupling in `HiePACS` , we want to contribute more precisely to the following points.

3.5.1. Efficient schemes for multiscale simulations

As mentioned previously, many important physical phenomena, such as material deformation and failure (see Section 4.2), are inherently multiscale processes that cannot always be modeled via continuum model. Fully microspcopic simulations of most domains of interest are not computationally feasible. Therefore, researchers must look at multiscale methods that couple micro models and macro models. Combining different scales such as quantum-atomistic or atomistic, mesoscale and continuum, are still a challenge to obtain efficient and accurate schemes that efficiently and effectively exchange information between the different scales. We are currently involved in two national research projects (ANR), that focus on multiscale schemes. More precisely, the models that we start to study are the quantum to atomic coupling (QM/MM coupling) in the NOSSI ANR and the atomic to dislocation coupling in the OPTIDIS ANR (proposal for the 2010 COSINUS call of the French ANR).

3.5.2. Load-balancing of complex coupled simulations based on the hypergraph model

One most important issue is undoubtedly the problem of load-balancing of the whole coupled simulation. Indeed, the naive balancing of each code on its own can lead to important imbalance in the coupling area. Another connected problem we plan to investigate is the problem of resource allocation. This is particularly important for the global coupling efficiency, because each code involved in the coupling can be more or less computationally intensive, and there is a good trade-off to find between resources assigned to codes to avoid that one of them wait for the others.

The performance of the coupled codes depends on how the data are well distributed on the processors. Generally, the data distributions of each code are built independently from each other to obtain the best load-balancing. But once the codes are coupled, the naive use of these decompositions can lead to important imbalance in the coupling area. Therefore, the modeling of the whole coupling is crucial to improve the performance and to ensure a good scalability. The goal is to find the best data distribution for the whole coupled codes and not only for each standalone code. One idea is to use an hypergraph model that will incorporate information about the coupling itself. Then, we expect the greater expressiveness of hypergraph will enable us to perform a coupling-aware partitioning in order to improve the load-balancing of the whole coupled simulation.

Another connected problem we plan to investigate is the problem of resource allocation. This is particularly important for the global coupling efficiency and scalability, because each code involved in the coupling can be more or less computationally intensive, and there is a good trade-off to find between resources assigned to codes to avoid that one of them wait for the others. Typically, if we have a given number of processors and two coupled codes, how to split the processors among each code?

Moreover, the load-balancing of modern parallel adaptive simulations raises a crucial issue when the problem size varies during execution. In such cases, it could be convenient to dynamically adapt the number of resources used at runtime. However, most of previous works on repartitioning only consider a constant number of resources. We plan to design new repartitioning algorithm based on an hypergraph model that can handle a variable number of processors. Furthermore, this kind of algorithms could be used for the dynamic balancing of a coupled simulation, in the case where the whole number of resources is fixed but can change for each code.

3.5.3. Steering and interacting with complex coupled simulations

The computational steering is an effort to make the typical simulation work-flow (modelling, computing, analyzing) more efficient, by providing online visualization and interactive steering over the on-going computational processes. The online visualization appears very useful to monitor and to detect possible errors in long-running applications, and the interactive steering allows the researcher to alter simulation parameters on-the-fly and to immediately receive feedback on their effects. Thus, the scientist gains an additional insight in the simulation regarding to the cause-and-effect relationship.

In the `ScAlApplix` project, we have studied this problem in the case where both the simulation and the visualization can be parallel, what we call M-by-N computational steering, and we have developed a software environment called `EPSN` (see Section 5.3). More recently, we have proposed a model for the steering of complex coupled simulations and one important conclusion we have from these previous works is that the steering problem can be conveniently modeled as a coupling problem between one or more parallel simulation codes and one visualization code, that can be parallel as well. We propose in `HiePACS` to revisit the steering problem as a coupling problem and we expect to reuse the new redistribution algorithms developped in the context of code coupling for the purpose of M-by-N steering. We expect such an approach will enable to steer massively-parallel simulations. Another point we plan to study is the monitoring and interaction with resources, in order to perform user-directed checkpoint/restart or user-directed load-balancing at runtime.

In several applications, it is often very useful either to visualize the results of the ongoing simulation before writing it to disk, or to steer the simulation by modifying some parameters and visualize the impact of these modifications interactively. Nowadays, high performance computing simulations use many computing nodes, that perform I/O using the widely used HDF5 file format. One of the problems is now to use real-time visualization using high performance computing. In that respect we need to efficiently combine very large parallel simulation systems with parallel visualization systems. The originality of this approach is the use of the HDF5 file format to write in a distributed shared memory (DSM); so that the data can be read from the upper part of the visualization pipeline. This leads to define a relevant steering model based on a DSM. It implies finding a way to write/read data efficiently in this DSM, and steer the simulation. This work is developed in collaboration with the Swiss National Supercomputing Centre (CSCS).

As concerns the interaction aspect, we are interested in providing new mechanisms to interact with the simulation directly through the visualization. For instance in the ANR NOSSI, in order to speed up the computation we are interested in rotating a molecule in a cavity or in moving it from one cavity to another within the crystal latice. To perform safely such interactions a model of the interaction in our steering framework is necessary to keep the data coherency in the simulation. Another point we plan to study is the monitoring and interaction with ressources, in order to perform user-directed checkpoint/restart or user-directed load balancing at runtime.

4. Application Domains

4.1. Introduction

Currently, we have one major application which is material physics, and for which we contribute to all steps that go from modelling aspects to the design and the implementation of very efficient algorithms and codes for very large multi-scale simulations. Moreover, we apply our algorithmic research about linear algebra (see

Section 3) in the context of several collaborations with industrial and academic partners. Our high performance libraries are or will be integrated in several complex codes and will be used and validated for very large simulations.

4.2. Material physics

Due to the increase of available computer power, new applications in nano science and physics appear such as study of properties of new materials (photovoltaic materials, bio- and environmental sensors, ...), failure in materials, nano-indentation. Chemists, physicists now commonly perform simulations in these fields. These computations simulate systems up to billion of atoms in materials, for large time scales up to several nanoseconds. The larger the simulation, the smaller the computational cost of the potential driving the phenomena, resulting in low precision results. So, if we need to increase the precision, there is two ways to decrease the computational cost. In the first approach, we improve classical methods and algorithms and in the second way, we will consider a multiscale approach.

Many applications in material physics need to couple several models like quantum mechanic and molecular mechanic models, or molecular and mesoscopic or continuum models. These couplings allow scientists to treat larger solids or molecules in their environment. Many of macroscopic phenomena in science depend on phenomena at smaller scales. Full simulations at the finest level are not computationally feasible in the whole material. Most of the time, the finest level is only necessary where the phenomenon of interest occurs; for example in a crack propagation simulation, far from the tip, we have a macroscopic behavior of the material and then we can use a coarser model. The idea is to limit the more expensive level simulation to a subset of the domain and to combine it with a macroscopic level. This implies that atomistic simulations must be speeded up by several orders of magnitude.

We will focus on two applications; the first one concerns the computation of optical spectra of molecules or solids in their environment. In the second application, we will develop faster algorithms to obtain a better understanding of the metal plasticity, phenomenon governing by dislocation behavior. Moreover, we will focus on the improvement of the algorithms and the methods to build faster and more accurate simulations on modern massively parallel architectures.

4.2.1. Hybrid materials

There is current interest in hybrid pigments for cosmetics, phototherapy and paints. Hybrid materials, combining the properties of an inorganic host and the tailorable properties of organic guests, particularly dyes, are also of wide interest for environmental detection (oxygen sensors) and remediation (trapping and elimination of dyes in effluents, photosensitised production of reactive oxygen species for reduction of air and water borne contaminants). A thorough understanding of the factors determining the photo and chemical stability of hybrid pigments is thus mandated by health, environmental concerns and economic viability.

Many applications of hybrid materials in the field of optics exploit combinations of properties such as transparency, adhesion, barrier effect, corrosion, protection, easy tuning of the colour and refractive index, adjustable mechanical properties and decorative properties. It is remarkable that ancient pigments, such as Maya Blue and lacquers, fulfill a number of these properties. This is a key to the attractiveness of such materials. These materials are not simply physical mixtures, but should be thought of as either miscible organic and inorganic components, or as a heterogeneous system where at least one of the component exhibits a hierarchical order at the nanometer scale. The properties of such materials no longer derive from the sum of the individual contributions of both phases, since the organic/inorganic interface plays a major role. Either organic and inorganic components are embedded and only weak bonds (hydrogen, van der Waals, ionic bonds) give the structure its cohesion (class I) or covalent and iono-covalent bonds govern the stability of the whole (class II).

These simulations are complex and costly and may involve several length scales, quantum effects, components of different kinds (mineral-organic, hydro-philic and -phobic parts). Computer simulation already contributes widely to the design of these materials, but current simulation packages do not provide several crucial functions, which would greatly enhance the scope and power of computer simulation in this field.

The computation of optical spectra of molecules and solids is the greatest use of the Time Dependent Density Functional Theory (TDDFT). We compute the ground state of the given system as the solution of the Kohn-Sham equations (DFT). Then, we compute the excited states of the quantum system under an external perturbation - electrical field of the environment - or thanks to the linear theory, we compute only the response function of the system. In fact, physicists are not only interesting by the spectra for one conformation of the molecule, but by an average on its available configurations. To do that, they sample the trajectory of the system and then compute several hundred of optical spectra in one simulation. But, due to the size of interesting systems (several thousands of atoms) and even if we consider linear methods to solve the Kohn-Sham equations arising from the Density Functional Theory, we cannot compute all the system at this scale. In fact, such simulations are performed by coupling Quantum mechanics (QM) and Molecular mechanic (MM). A lot of works are done on the way to couple these two scales, but a lot of work remains in order to build efficient methods and efficient parallel couplings.

The most consuming time in such coupling is to compute optical spectra is the TDDFT. Unfortunately, examining optical excitations based on contemporary quantum mechanical methods can be especially challenging because accurate methods for structural energies, such as DFT, are often not well suited for excited state properties. This requires new methods designed for predicting excited states and new algorithms for implementing them. Several tracks will be investigated in the project:

- Typically physicists or chemists consider spectral functions to build a basis (orbital functions) and all the computations are performed in a spectral way. Due to our background, we want to develop new methods to solve the system in the real space by finite differences or by wavelets methods. The main expectation is to construct error estimates based on for instance the grid-size h parameter.
- For a given frequency in the optical spectra, we have to solve a symmetric non Hermitian system. With our knowledge on linear solvers, we think that we can improve the methods commonly used (Lanczos like) to solve the system (see Section 3.3).
- Improving the parallel coupling is crucial for large systems because the computational cost of the atomic and quantum models are really different. In parallel we have the following order of magnitude: one second or less per time step for the molecular dynamics, several minutes or more for the DFT and the TDDFT. The challenge to find the best distribution in order to have the same CPU time per time step is really important to reach high performance. Another aspect in the coupling is the coupling with the visualization to obtain online visualization or steerable simulations. Such steerable simulations help the physicists to construct the system during the simulation process by moving one or a set of molecules. This kind of interaction is very challenging in terms of algorithmic and this is a good field for our software platform `EPSN` .

4.2.2. Material failures

Another domain of interest is the material aging for the nuclear industry. The materials are exposed to complex conditions due to the combination of thermo-mechanical loading, the effects of irradiation and the harsh operating environment. This operating regime makes experimentation extremely difficult and we must rely on multi-physics and multi-scale modelling for our understanding of how these materials behave in service. This fundamental understanding helps not only to ensure the longevity of existing nuclear reactors, but also to guide the development of new materials for 4th generation reactor programs and dedicated fusion reactors. For the study of crystalline materials, an important tool is dislocation dynamics (DD) modelling. This multiscale simulation method predicts the plastic response of a material from the underlying physics of dislocation motion. DD serves as a crucial link between the scale of molecular dynamics and macroscopic methods based on finite elements; it can be used to accurately describe the interactions of a small handful of dislocations, or equally well to investigate the global behavior of a massive collection of interacting defects.

To explore, i.e., to simulate these new areas, we need to develop and/or to improve significantly models, schemes and solvers used in the classical codes. In the project, we want to accelerate algorithms arising in those fields. We will focus on the following topics (in particular in the starting OPTIDIS ANR-COSINUS project

in collaboration with CEA Saclay, CEA Ile-de-france and SIMaP Laboratory in Grenoble) in connection with research described at Sections 3.4 and 3.5.

- The interaction between dislocations is long ranged ($O(1/r)$) and anisotropic, leading to severe computational challenges for large-scale simulations. In dislocation codes, the computation of interaction forces between dislocations is still the most CPU time consuming and has to be improved to obtain faster and more accurate simulations.
- In such simulations, the number of dislocations grows while the phenomenon occurs and these dislocations are not uniformly distributed in the domain. This means that strategies to dynamically construct a good load balancing are crucial to acheive high performance.
- From a physical and a simulation point of view, it will be interesting to couple a molecular dynamics model (atomistic model) with a dislocation one (mesoscale model). In such three-dimensional coupling, the main difficulties are firstly to find and characterize a dislocation in the atomistic region, secondly to understand how we can transmit with consistency the information between the two micro and meso scales.

4.3. Application framework customers of high performance linear algebra solvers

We are currenlty collaborating with various research groups involved in geophysics, electromagnetics and structural mechanics. For all these application areas, the current bottleneck is the solution of huge sparse linear systems often involving multiple right-hand sides either available simultaneously or given in sequence. The robustness, efficiency and scalability of the numerical tools designed in Section 3.3 will be preliminary investigated in the parallel simulation codes of these partners.

More precisely, BRGM simulations require the solutions of huge linear systems with many right-hand sides given simultaneously. We notice that the collaborative work with TOTAL address the use of `GPU` for intensive numerical kernels in the Reverse Time Migration process for seismic imaging.

The CEA-CESTA simulation codes need the solution with simultaneous right-hand sides but also with right-hand sides given in sequence. The first situation arises in RCS calculations, but is generic in many parametric studies, while the second one comes from the nature of the solver that is based on a multiplicative Schwarz approach. The subproblems are solved several times in sequence. Many of the numerical approaches and possible outcoming software are well suited to tackle these challenging problems.

Research activities related to EDF and developed in the framework of the ANR SOLSTICE project have already stimulated interactions between members of the former `ScAlApplix` INRIA project team and members of the Parallel Algorithms team of CERFACS. These research activities have concerned direct and iterative solution methods for linear systems and eigenvalue computations.

On more academic sides, some ongoing collaborations with other INRIA EPIs will be continued and others will be started. In collaboration with the NACHOS INRIA project team, we will continue to investigate the use of efficient linear solvers for the solution of the Maxwell equations in the time and frequency domains where discontinuous Galerkin discretizations are considered. Additional funding will be sought out in order to foster this research activity in connection with actions described in Section 3.3.

The efficient solution of linear systems strongly relies on the activities described in Section 3.2 (e.g. complex load balancing problem) and in Section 3.3 (for the various parallel linear algebra kernels).

5. Software

5.1. Introduction

We describe in this section the software that we are developing. The first two (MaPHyS and EPSN) will be the main milestones of our project. The other software developments will be conducted in collaboration with academic partners or in collaboration with some industrial partners in the context of their private R&D or production activities. For all these software developments, we will use first the various (very) large parallel platforms available through CERFACS and GENCI in France (CCRT, CINES and IDRIS Computational Centers), and next the high-end parallel platforms that will be available via European and US initiatives or projects such that PRACE.

5.2. MaPHyS

MaPHyS (Massivelly Parallel Hybrid Solver) is a software package whose proptotype was initially developed in the framework of the PhD thesis of Azzam Haidar (CERFACS) and futher consolidated thanks to the ANR-CIS Solstice funding. This parallel linear solver couples direct and iterative approaches. The underlying idea is to apply to general unstructured linear systems domain decomposition ideas developed for the solution of linear systems arising from PDEs. The interface problem, associated with the so called Schur complement system, is solved using a block preconditioner with overlap between the blocks that is referred to as Algebraic Additive Schwarz.

In the framework of the INRIA technologic development actions; 24 man-month engineer (Yohan Lee-Tin-Yien) have been allocated to this software activity started in December 2009. The initial software prototype has been completly redesigned in order to enable us to easily interface any sparse direct solvers and develop new preconditioning technique. The same software effort has been undertaken for interfacing any graph partitioning tools.

The MaPHyS package is very much a first outcome of the research activity described in Section 3.3. Finally, MaPHyS is a preconditioner that can be used to speed-up the convergence of any Krylov subspace method. We forsee to either embed in MaPHyS some Krylov solvers or to release them as standalone packages, in particular for the block variants that will be some outcome of the studies discussed in Section 3.3.

5.3. EPSN

EPSN (Environement for Computational Steering) is a software environment for the steering of parallel numerical simulations with visualization programs that can be parallel as well (see Figure 1). Moreover, it provides a library, called RedGRID, dedicated to the coupling of parallel codes, and more precisely to the redistribution of complex parallel data objects such as structured grids, particles and unstructured meshes.

EPSN is a distributed computational steering environment which allows the steering of remote parallel simulations with sequential or parallel visualization tools or graphics user interface. It is a distributed environment based on a simple client/server relationship between user interfaces (clients) and simulations (servers). The user interfaces can dynamically be connected to or disconnected from the simulation during its execution. Once a client is connected, it interacts with the simulation component through an asynchronous and concurrent request system. We distinguish three kinds of steering request. Firstly, the "control" requests (play, step, stop) allow to steer the execution flow of the simulation. Secondly, the "data access" requests (get, put) allow to read/write parameters and data from the memory of the remote simulation. Finally, the "action" requests enable to invoke user-defined routines in the simulation. In order to make a legacy simulation steerable, the end-user annotates its simulation source-code with the EPSN API. These annotations provide the EPSN environment with two kinds of information: the description of the program structure according to a Hierarchical Task Model (HTM) and the description of the distributed data that will be accessible by the remote clients.

Concerning the development of client applications, we also provide a front-end API that enables the integration of EPSN in a high-level visualization system such as *VTK* or *ParaView*. We also provide a lightweight user interface, called *SiMonE* (Simulation Monitoring for EPSN), that enables us to easily connect any simulations and interact with them, by controlling the computational flow, viewing the current parameters or data on a simple data-sheet and modifying them optionally. *SiMonE* also includes simple visualization plug-ins to

online display the intermediate results. Moreover, the EPSN framework offers the ability to exploit parallel visualization and rendering techniques thanks to the Visualization ToolKit (VTK). This approach allows us to reduce the steering overhead of the EPSN platform and allows us to process efficiently large dataset. It is also possible to exploit tiled-display wall with EPSN in order to reach high resolution image.

As both the simulation and the visualization can be parallel applications, EPSN is based on the M×N redistribution library called RedGRID. This library is in charge of computing all the messages that will be exchanged between the two parallel components, and is also in charge of performing the data transfer in parallel. Thus, RedGRID is able to aggregate the bandwidth and to achieve high performance. Moreover, it is designed to consider a wide variety of distributed data structures usually found in the numerical simulations, such as structured grids, points or unstructured meshes.

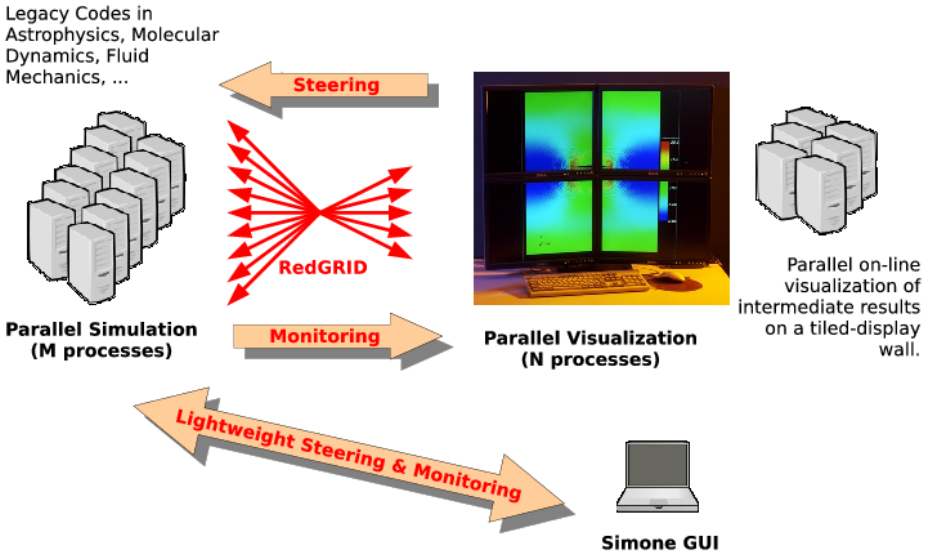

Figure 1. EPSN: software environment for $M \times N$ computational steering.

Both EPSN and RedGRID use a communication infrastructure based on CORBA which provides our platform with portability, interoperability and network transparency. EPSN has been supported by the ACI-GRID program (grant number PPL02-03), the ARC RedGRID , and more recently by the ANR program called MASSIM (grant number ANR-05-MMSA-0008-03). It is now involved in the ANR CIS NOSSI (2007). More informations are available on our web site: http://www.labri.fr/projet/epsn. This software is publicy available at INRIA Gforge (http://gforge.inria.fr/projects/epsn).

5.4. MPICPL

In the context of dynamic load-balancing and code coupling, we have started the development of a new framework, called MPICPL. This framework has an experimental purpose, that will make easier the research, the development and the experimentation of the new algorithms we are looking for, as described at section 3.5.

This framework will be based on the well-known MPI standard to obtain performance, and will fully exploit the new facilities provided by MPI-2. Indeed, the dynamic process management allowed by MPI-2 offers interesting possibility for the design of code coupling.

The framework is available at INRIA Gforge: http://mpicpl.gforge.inria.fr.

5.5. Other software

These software packages are or will be developed in collaboration with some academic partners (LIP6, LaBRI, CPMOH, IPREM, EPFL) or in collaboration with industrial partners (CEA, TOTAL, EDF) in the context of their private R&D or production activities.

- Fast Multipole with BLAS (FMB), developed in collaboration with P. Fortin (LIP6), is a high performance parallel implementation of the Fast Multipole Method for the Laplace equation. It is based on BLAS routines and on an hybrid MPI-Thread parallelization for both shared and distributed memory architectures (see Section 3.4).
- For the materials physics applications, a lot of development will be done in the context of ANR projects (NOSSI and proposal OPTIDIS, see Section 4.2) in collaboration with LaBRI, CPMOH, IPREM, EPFL and with CEA Saclay and Bruyère-le-Châtel.
- In the context of the PhD thesis of Mathieu Chanaud (collaboration with CEA/CESTA), we develop a new parallel plateform based on a combination of a multigrid solver and a direct solver (the PaStiX solver developped in the previous `ScAlApplix` project-team) to solve huge linear systems arising from Maxwell equations discretized with first-order Nédelec elements (see Section 3.3).
- Finally, we contribute to software developments for seismic analysis and imaging and for wave propagation in collaboration with TOTAL (use of `GPU` technology with CUDA in the context of the PhD thesis of Rached Abdelkhalek) and with BRGM (use of `PaStiX` and `MaPHyS` solvers in the context of the PhD of Fabrice Dupros in collaboration with Dimitri Komatitsch of MAGIQUE3D project team).

6. New Results

6.1. Algorithms and high-performance solvers

6.1.1. Hybrid direct/iterative solvers based on algebraic domain decomposition techniques

We have studied numerical variants to approximate the Schur complement based on incomplete factorization in order to reduce the memory consumption of the solver. The numerical scalability of this variant of the solver has been studied on the the solution of large linear systems resulting arising from the discretisation of three dimensional convection diffusion problems. The robustness and the scalability of the preconditioners are investigated through extensive parallel experiments on up to two thousand processors. Their efficiency from a numerical and parallel performance view point are investigated. ore detailed on this work can be found in [6].

Parallel numerical experiments with a variant of the solver exploiting two levels of parallelism (MPI-MPI) has also been investigated of large 3D strcutural mechanic problems, the design of the software and numerical experiments are reported in [18].

6.1.2. Dense linear algebra solvers for multicore processors accelerated with multiple GPUs

In collaboration with the INRIA Runtime team and the University of Tennessee, we have designed dense linear algebra solvers that can fully exploit a node composed of a multicore processor accelerated with multiple GPUs [39], [21].

6.1.3. Full geometric multigrid method for 3D Maxwell equations

In the context of a collaboration with the CEA/CESTA center, Mathieu Chanaud continues a Ph.D. concerning a tight combination between multigrid methods and direct methods for the efficient solution of challenging 3D irregular finite element problems arising from the discretization of Maxwell equations. A parallel solver dedicated to the ODYSSEE challenge (electromagnetism) of CEA/CESTA has been implemented and integrated. The robustness of the numerical scheme and its parallel scalability are ongoing activities.

6.1.4. Domain decomposition methods to solve neutron transport equations

This work, started with a collaboration between the EDF/SINETICS team and the former ScAlApplix project, intended to design and develop techniques to optimize the efficiency of the codes used to simulate the physics of nuclear reactors. In the context of Bruno Lathuilière PhD (in collaboration with Pierre Ramet from BACCHUS), we have completed a study to parallelize a SPn simulation code by using a domain decomposition method applied for the solution of the neutron transport equations (Boltzmann equations). The defense took place early February 2010.

6.1.5. High performance algorithms for wave propagation

A first work has been initiated during the ANR CIGC-05 NUMASIS project. The overall objective is the adaptation and the optimization of numerical methods in geophysics for large scale simulations on hierarchical and multicores architectures. Fabrice Dupros (BRGM) has started a PhD on these topics in February 2007 in the former ScAlApplix project. This work is also carried out in the framework of a collaboration with the INRIA MAGIQUE3D team (Dimitri Komatitsch) and BRGM. Several contributions can be underlined, for example the impact of the memory hierarchy for this class of simulations. Large scale finite-elements computations for site effects in the French Riviera urban area have also been performed on the JADE GENCI/Cines platform using the PaStiX sparse parallel direct solver. An ongoing topic is the evaluation of a spacetime decomposition for the time-domain finite-differences method (FDTD) and its application to the classical staggered-grid scheme. The defense of this PhD took place in december.

A second work is currently carried on with TOTAL (Rached Abdelkhalek PhD). The extraordinary challenge that the oil and gas industry must face for hydrocarbon exploration requires the development of leading edge technologies to recover an accurate representation of the subsurface. Seismic modeling and Reverse Time Migration (RTM) based on the full wave equation discretization, are tools of major importance since they give an accurate representation of complex wave propagation areas. Unfortunately, they are highly compute intensive. The recent development in GPU technologies with unified architecture and general-purpose languages coupled with the high and rapidly increasing performance throughput of these components made General Purpose Processing on Graphics Processing Units an attractive solution to speed up diverse applications. We have designed a fast parallel simulator that solves the acoustic wave equation on a GPU cluster. Solving the acoustic wave equation in an oil exploration industrial context aims at speeding up seismic modeling and Reverse Time Migration. We consider a finite difference approach on a regular mesh, in both 2D and 3D cases. The acoustic wave equation is solved in a constant density or a variable density domain. All the computations are done in single precision, since double precision is not required in our context. We use nvidia CUDA to take advantage of the GPU computational power. We study different implementations and their impact on the application performance. We obtain a speed up of 16 for Reverse Time Migration and up to 43 for the modeling application over a sequential code running on general purpose CPU. The defense of this thesis is planned for the first 2011 semester.

6.2. Efficient algorithmics for code coupling in complex simulations

The performance of the coupled codes depends on how the data are well distributed on the processors. Generally, the data distributions of each code are built independently from each other to obtain the best load-balancing. But once the codes are coupled, the naive use of these decompositions can lead to important imbalance in the coupling area. Therefore, the modeling of the whole coupling is crucial to improve the performance and to ensure a good scalability. The goal is to find the best data distribution for the whole

coupled codes and not only for each standalone code. One idea is to use an hypergraph model that will incorporate information about the coupling itself. Then, we expect the greater expressiveness of hypergraph will enable us to perform a coupling-aware partitioning in order to improve the load-balancing of the whole coupled simulation.

As the data handled by two coupled codes often have different structures and resolutions, we model them in a generic way with two distinct hypergraphs - each hypergraph representing data of one code - that are connected by inter-edges which represents spatial intersection bewteen geometric elements of the two codes. For instance, the fluid-structure coupling often uses a structured grid for CFD and an unstructuresh mesh for solid mechanics. Another classical example is the ocean-atmosphere coupling that typically uses two structured grids with different resolutions.

Based upon this model, we propose a new partitioning approach that is aware of the code coupling. Let us consider two codes A and B, modeled by two hypergraphs Ha and Hb, connected by inter-edges $I(Ha, Hb)$. Formally, the problem consists in partitioning Ha in M and Hb in N with accounting for coupling communications that depends on $I(Ha, Hb)$. Our strategy is divided in three steps: 1) first, we freely partition Ha in M, that give us the partition $Pa(M)$; 2) then, we projects this partition to Hb according to $I(Ha, Hb)$, that provides the partition $Pb(M)$; 3) finally, we obtain the partition $Pb(N)$ by repartitioning Hb from M existing parts into N.

However, most of works on repartitioning only consider a constant number of processors. To overcome this issue, we have proposed a new repartitioning algorithm based on an hypergraph model that can handle a variable number of processors. This algorithm is inspired from recent works in Zoltan, based on hypergraph partitioning technics with fixed vertices. Moreover, our algorithm uses a linear communication pattern, that we have proved to minimize the total number of messages between the former and newer parts.

We currently investigate how to reuse this algorithm in the context of dynamic load-balancing of parallel adaptive simulations when the problem size varies during execution. In such cases, it would be convenient to dynamically adapt the number of resources used at runtime, while minimizing the migration cost.

These preliminary works has been realized during the Master internship of Clément Vuchener and is curently followed in PhD.

6.3. Computational steering environment for distributed numerical simulations

6.3.1. Model for the steering of parallel-distributed simulations

The model that we have proposed in the EPSN framework can only steer efficiently SPMD simulations. A natural evolving is to consider more complex simulations such as coupled SPMD codes called M-SPMD (Multiple SPMD like multiscale simulation for “crack-propagation”) and client/server simulation codes. In order to steer these kinds of simulation, we have designed an extension to the Hierarchical Task Model (HTM), which affords to solve the coherency problem for such complex applications. The EPSN framework has been extended to handle this new kind of simulations. In the context of the ANR MASSIM and ANR NOSSI, we have recently validated our works with a multi-scale simulation for “crack-propagation” (LibMultiScale). In this case-study, EPSN is able to pause/resume the whole coupled simulation, to coherently get and visualize the complex distributed data: a distributed unstructured mesh at the continuum scale, mixed with distributed atoms at the atomic scale. This work has been defended in the PhD of Nicolas Richart the 20th of January, 2010.

6.3.2. Distributed Shared Memory approach for the steering of parallel simulations

As a different approach of the in-situ and steering framework of EPSN, we conceived and developed a light push-driven architecture for in-situ visualization. The architecture, part of ICARUS, is intended to address three principal objectives: Require little or no modification to the simulation code in order to allow a live visualization. Allow the simulation to be run on one parallel machine whilst the visualization is run on a

separate (or the same) parallel machine. Provide good performance to ensure that massive simulations may be handled as easily as small test cases. The interface developed is built around the HDF5 file I/O library used commonly in HPC applications. The HDF5 API allows the derivation of custom virtual file drivers (VFDs) which may be instantiated at run-time on a per file basis to control how data is written to the file system. We have made use of this facility to create a specialized MPI based VFD which allows the simulation to write data in parallel to a file, which is actually redirected over the network to a visualization cluster which in turn stores the file in a Distributed Shared Memory (DSM) buffer - or in effect a virtual file system. The ParaView application acts as a server/host for this DSM and can read the file contents directly using the HDF5 API as if reading from disk. The transfer of data between simulation and visualization machines may be done using either an MPI based communicator shared between the applications, or using a socket based communication. The management of both ends of the network transfer is transparently handled by our DSM VFD layer, meaning that an application using HDF5 can make use of in-situ visualization without any code changes. It is only necessary to re-link the application against a modified version of the HDF library which contains our driver. This work (Jérôme Soumagne PhD) has been made and is currently carrying on at CSCS - Swiss National Supercomputing Centre, under the co-supervision of Mr. John Biddiscombe, within the NextMuSE European project 7th FWP/ICT-2007.8.0 FET Open.

6.4. Material physics

6.4.1. Hybrid materials

The study of hybrid materials with a coupling method between molecular dynamics (MD) and quantum mechanism (QM) has begun in collaboration with IPREM (Pau) in the ANR CIS 2007 NOSSI. These simulations are complex and costly and may involve several length scales, quantum effects, components of different kinds (mineral-organic, hydro-philic and -phobic parts). Our goal is to compute dynamical properties of hybrid materials like optical spectra. The computation of optical spectra of molecules and solids is the most consuming time in such coupling. This requires new methods designed for predicting excited states and new algorithms for implementing them. Several tracks are investigated in the project and new results obtained as described bellow.

Optical spectra. We have improved our TDDFT method based on the LCAO method to densities and excited states in order to compute electronic excitation spectra. Firstly, we have developed an iterative method (bi-orthogonal Lanczos and GMRES) to compute the spectra. For each point we need to solve three systems. In [20] we illustrated the performance of our method for computing spectra using benzene, indigo, and fullerene. These results confirmed the $O(N^2 N_\omega)$ complexity scaling of our method where N is the number of atoms and N_ω is the number of frequency points. A drawback of our construction is that it required the full matrix of the linear reponse, χ_0; this has a high memory requirement. Typically even for a small molecule like Indigo we need more than 10 Gb to store the matrix. To overcome this limitation, we have developed in [19] a matrix free method and parallelize our algorithm. The speed of our code is roughly comparable to commercial TDDFT codes. In general, we expect our method to be faster than the solution of Casidaâs equation for systems with dense spectra, when the target range contains many allowed transitions. The algorithm was parallelized and demonstrated to be suitable for treating molecules with more than 100 atoms on large current heterogeneous architectures using the OpenMP/MPI paradigms.

QM/MM algorithm. For structure studies or dynamical properties, we intend to couple QM model based on pseudo-potentials (SIESTA code) with dynamic molecular (DL-POLY code). Therefore we have first developed a new algorithm to avoid counting twice the quantum electric field in the molecular model. Then, we have introduced an algorithm to compute faster the electric field which polarizes the quantum atoms. We are currently implementing our algorithm in SIESTA and DL-POLY codes.

7. Contracts and Grants with Industry

7.1. Research and development contracts

CEA research and development contracts:

- Conception of an hybrid solver combining multigrid and direct methods (Mathieu Chanaud (PhD); David Goudin and Jean-Jacques Pesqué from CEA-CESTA; Luc Giraud, Jean Roman).

EDF research and development contract:

- Application of a domain decomposition method to the neutronic SPn equations (Bruno Lathuilière (PhD); Pierre Ramet from BACCHUS INRIA project team; Jean Roman).

TOTAL research and development contract:

- Massive parallelism and use of `GPU` devices for seismic depth imaging problems (Rached Abdelkhalek (PhD); Olivier Coulaud, Guillaume Latu, Jean Roman).
- Parallel elastodynamic solver for 3D models with local mesh refinment (Yohann Dudouit (PhD); Luc Giraud and Sébastien Pernet from EMA-CERFACS).

8. Other Grants and Activities

8.1. National initiatives

8.1.1. NOSSI: New platform for parallel, hybrid quantum/classical simulations

Participants: Olivier Coulaud, Aurélien Esnard, Damien Genet.

Grant: ANR 2007 – CIS
Dates: 2008 – 2010
Partners: CPMOH (Bordeaux, UMR 5098), DRIMM, IMPREM (leader of the project, Pau, UMR 5254), Institut Néel (Grenoble, UPR2940)
Overview: Physicists, chemists and computer scientists join forces in this project to further design high performance numerical simulation of materials, by developing and deploying a new platform for parallel, hybrid quantum/classical simulations. The platform synthesizes established functions and performances of two major European codes, SIESTA and DL-POLY, with new techniques for the calculation of the excited states of materials, and a graphical user interface allowing steering, visualization and analysis of running, complex, parallel computer simulations.

The platform couples a novel, fast TDDFT (Time dependent density functional theory) route for calculating electronic spectra with electronic structure and molecular dynamics methods particularly well suited to simulation of the solid state and interfaces.

The software will be capable of calculating the electronic spectra of localized excited states in solids and at interfaces. Applications of the platform include hybrid organic-inorganic materials for sustainable development, such as photovoltaic materials, bio- and environmental sensors, photocatalytic decontamination of indoor air and stable, non-toxic pigments.
Web: http://nossi.gforge.inria.fr/index.html

8.1.2. SOLSTICE: High performance solvers for frontier simulations

Participants: Abdou Guermouche, Luc Giraud, Jean Roman.

Grant: ANR-06-CIS
Dates: 2006 – 2010
Partners: CERFACS, EADS IW, EDF R&D SINETICS, INRIA Rhône-Alpes and LIP, INPT/IRIT, CEA/CESTA, CNRS/GAME/CNRM

Overview: New advances in high-performance numerical simulation require the continuing development of new algorithms and numerical methods. These technologies must then be implemented and integrated into real-life parallel simulation codes in order to address critical applications that are at the frontier of our know-how. The solution of sparse systems of linear equations of (very) large size is one of the most critical computational kernel in terms of both memory and time requirements. Three-dimensional partial differential equations (3D-PDE) are particularly concerned by the availability of efficient sparse linear algorithms since the numerical simulation process often leads to linear systems of 10 to 100 million variables that need to be solved many times. In a competitive environment where numerical simulation becomes extremely critical compared to physical experimentation, very precise models involving a very accurate discretisation are more and more critical. The objective of our project is thus both to design and develop high-performance parallel linear solvers that will be efficient to solve complex multiphysic and multiscale problems of very large size. To demonstrate the impact of our research, the work produced in the project will be integrated in real simulation codes to perform simulations that could not be considered with today's technologies.
Web: http://solstice.gforge.inria.fr/

8.1.3. RESCUE: Résilience des applications scientifiques

Participants: Emmanuel Agullo, Luc Giraud, Abdou Guermouche, Jean Roman.

Grant: ANR-Blanc (computer science theme)
Dates: 2010 – 2014
Partners: INRIA EPI GRAAL (leader) and Grand Large.
Overview: The advent of exascale machines will help solve new scientific challenges only if the resilience of large scientific applications deployed on these machines can be guaranteed. With 10,000,000 core processors, or more, the time interval between two consecutive failures is anticipated to be smaller than the typical duration of a checkpoint, i.e., the time needed to save all necessary application and system data. No actual progress can then be expected for a large-scale parallel application. Current fault-tolerant techniques and tools can no longer be used. The main objective of the RESCUE project is to develop new algorithmic techniques and software tools to solve the exascale resilience problem. Solving this problem implies a departure from current approaches, and calls for yet-to-be-discovered algorithms, protocols and software tools.

This proposed research follows three main research thrusts. The first thrust deals with novel checkpoint protocols. This thrust will include the classification of relevant fault categories and the development of a software package for fault injection into application execution at runtime. The main research activity will be the design and development of scalable and light-weight checkpoint and migration protocols, with on-the-fly storing of key data, distributed but coordinated decisions, etc. These protocols will be validated via a prototype implementation integrated with the public-domain MPICH project. The second thrust entails the development of novel execution models, i.e., accurate stochastic models to predict (and, in turn, optimize) the expected performance (execution time or throughput) of large-scale parallel scientific applications. In the third thrust, we will develop novel parallel algorithms for scientific numerical kernels. We will profile a representative set of key large-scale applications to assess their resilience characteristics (e.g., identify specific patterns to reduce checkpoint overhead). We will also analyze execution trade-offs based on the replication of crucial kernels and on decentralized ABFT (Algorithm-Based Fault Tolerant) techniques. Finally, we will develop new numerical methods and robust algorithms that still converge in the presence of multiple failures. These algorithms will be implemented as part of a software prototype, which will be evaluated when confronted with realistic faults generated via our fault injection techniques.

We firmly believe that only the combination of these three thrusts (new checkpoint protocols, new execution models, and new parallel algorithms) can solve the exascale resilience problem. We hope to contribute to the solution of this critical problem by providing the community with new protocols, models and algorithms, as well as with a set of freely available public-domain software prototypes.

We notice that a natural extension of this project was our involvement in a G8-proposal (ECS : Enabling Climate Simulation at Extreme Scale) that went through the first round of selection for which the final decison whould be known early 2011.

8.1.4. OPTIDIS: Optimisation d'un code de dynamique des dislocations

Participants: Olivier Coulaud, Aurélien Esnard, Luc Giraud, Jean Roman.

Grant: ANR-COSINUS
Dates: 2010 – 2014
Partners: CEA/DEN/DMN/SRMA (leader), SIMaP Grenoble INP and ICMPE / Paris-Est.
Overview: Plastic deformation is mainly accommodated by dislocations glide in the case of crystalline materials. The behaviour of a single dislocation segment is perfectly understood since 1960 and analytical formulations are available in the literature. However, to understand the behaviour of a large population of dislocations (inducing complex dislocations interactions) and its effect on plastic deformation, massive numerical computation is necessary. Since 1990, simulation codes have been developed by French researchers. Among these codes, the code TRIDIS developed by the SIMAP laboratory in Grenoble is the pioneer dynamic dislocation code. In 2007, the project called NUMODIS had been set up as team collaboration between the SIMAP and the SRMA CEA Saclay in order to develop a new dynamics dislocation code using modern computer architecture and advanced numerical methods. The objective was to overcome the numerical and physical limits of the previous code TRIDIS. The version NUMODIS 1.0 came out in December 2009, which confirms the feasibility of the project. The project OPTIDIS is initiated when the code NUMODIS is mature enough to consider parallel computiation. The objective of the project in to develop and validate the algorithms in order to optimise the numerical and performance efficiencies of the NUMODIS code. We are aiming at developing a code able to tackle realistic material problems such as the interaction between dislocations and irradiation defects in a grain plastical deformation after irradiation. These kinds of studies where "local mechanisms" are correlated with macroscopic behaviour is a key issue for nuclear industry in order to understand material ageing under irradiation, and hence predict power plant secured service life. To carry out such studies, massive numerical optimisations of NUMODIS are required. They involve complex algorithms lying on advanced computational science methods. The project OPTIDIS will develop through joint collaborative studies involving researchers specialized in dynamics dislocations and in numerical methods. This project is divided in 8 tasks over 4 years. Two PhD thesis will be directly funded by the project. One will be dedicated to numerical development, validation of complex algorithms and comparison with the performance of existing dynamics dislocation codes. The objective of the second is to carry out large scale simulations to validate the performance of the numerical developments made in OPTIDIS. In both cases, these simulations will be compared with experimental data obtained by experimentalists.

8.2. International initiatives

8.2.1. MyPlanet

Participants: Emmanuel Agullo, Luc Giraud, Jean Roman, Pablo Salas Medina, Xavier Vasseur.

Grant: European Commission : FP7 Marie-Curie ITN
Dates: 2010-2012
Partners: CERFACS (leader), Allinea software, Alstom Power Switzerland, Czestochowa University of Technology, Genias Graphics, Rolls Royce PLC UK, Technical Univ. Munich, Turbomeca, University of Cambridge, University Carlos III Madrid and University of Cyprus.

Overview: The present MYPLANET project responds to the first FP7-call "PEOPLE-INITIAL-TRAINING-ITN-2007-1" published by the European Commission. This collaborative initial training network represents a European initiative to train a new generation of engineers in the field of high performance computing applied to the numerical combustion simulation, energy conversion processes and related atmospheric pollution issues. Indeed, the project is based on the recognised lack on the European level of highly skilled engineers who are equally well-trained in both combustion technologies and high-performance computing (HPC) techniques. Thus the MYPLANET project will clearly contribute to the structuring of existing high-quality initial research training capacities in fluid mechanics and the HPC field through combining both public and private (industrial) sectors. The participation of industrial partners in the training of the researchers will directly expose these industries to high performance computing, which will have a very favourable impact on the quality and efficiency of their activities. Reciprocally, the research community will learn more about the mid and long term industrial challenges which will enable the research partners to initiate new activities in order to anticipate and address these industrial requirements
Web: http://www.cerfacs.fr/myplanet/

8.2.2. Scalable Hybrid Solvers for Large Sparse Linear Systems of Equations on Petascale Computing Architectures

Participants: Emmanuel Agullo, Luc Giraud, Abdou Guermouche, Jean Roman, Xavier Vasseur.

Grant: France Berkeley Fund
Dates: 2010-2012
Partners: Lawrence Berkeley National Laboratory.
Overview: Our approach to high-performance, scalable solution of large sparse linear systems in parallel scientific computing is to combine direct and iterative methods. Such a hybrid approach exploits the advantages of both direct and iterative methods. The iterative component allows us to use a small amount of memory and provides a natural way for parallelization. The direct part provides its favorable numerical properties. In the framework of this joint research action we intend to address the problems related to exploiting hybrid programming models on NUMA clusters and the solution of indefinite/augmented systems.

9. Dissemination

9.1. Participation to the Scientific Community

Olivier Coulaud has been member of the scientific committee of the international conference VECPAR'10 and of the INRIA COST GTAI committe (in charge of incentive actions).

Luc Giraud has been member of the scientific committee of the international conferences HiPC 2010, PDSEC-10, PMAA-10 and VecPar 2010 and one of the program chairs of CSE-2010. He was member of the selection committee for the ANR COSINUS programme.

Jean Roman is president of the Project Committee of INRIA Bordeaux - Sud-Ouest and member of the National Evaluation Committee of INRIA. He has been member of the scientific committee of the international conference EuroMicro PDP'10 (IEEE) and of the national conference Renpar'10. He is member of the "Strategic Comity for Intensive Computation" of the French Research Ministry and is member of the "Scientific Board" of the CEA-DAM.

Finally, the HiePACS members have contributed to the reviewing process of international journal (Applied Numerical Mathematics, BIT Numerical Mathematics, Computer Physics Communications, Concurrency and Computation: Practice and Experience, Journal of Computational and Applied Mathematics, Linear Algebra and its Applications, Neural Processing Letters, Parallel Computing) as well as experts for research agencies (ANR COSINUS, STIC-Amsud).

9.2. Teaching

In complement of the normal teaching activity of the university members and of ENSEIRB-MATMECA members, Emmanuel Agullo and Olivier Coulaud teach at ENSEIRB-MATMECA and Luc Giraud teaches at ENSEEIHT and ISAE-ENSICA.

10. Bibliography

Major publications by the team in recent years

[1] R. ABDELKHALEK, H. CALANDRA, O. COULAUD, J. ROMAN, G. LATU. *Fast Seismic Modeling and Reverse Time Migration on a GPU Cluster*, in "The 2009 High Performance Computing & Simulation - HPCS'09", Allemagne Leipzig, 2009.

[2] E. AGULLO, J. DONGARRA, B. HADRI, H. LTAIEF. *Tile QR factorization with parallel panel processing for multicore architectures*, in "SIAM Conference on Parallel Processing for Scientific Computing (PP10)", United States Seattle, 2010, http://hal.inria.fr/inria-00548907/en.

[3] O. COULAUD, P. FORTIN, J. ROMAN. *High-performance BLAS formulation of the Adaptive Fast Multipole Method*, in "Mathematical and Computer Modelling", 2009, vol. 51, n^{o} 3-4, p. 177-188.

[4] F. DUPROS, F. DE MARTIN, E. FOERSTER, D. KOMATITSCH, J. ROMAN. *High-performance finite-element simulations of seismic wave propagation in three-dimensional non linear inelastic geological media*, in "Parallel Computing", January 2010, vol. 36, p. 308-325 [*DOI :* 10.1016/J.PARCO.2009.12.011], http://hal.inria.fr/inria-00528479.

[5] L. GIRAUD, S. GRATTON, X. PINEL, X. VASSEUR. *Flexible GMRES with deflated restarting*, in "SIAM Journal on Scientific Computing", October 2010, vol. 32, n^{o} 4, p. 1858–1878 [*DOI :* 10.1137/080741847], http://hal.inria.fr/inria-00542426/en.

[6] L. GIRAUD, A. HAIDAR, Y. SAAD. *Sparse approximations of the Schur complement for parallel algebraic hybrid linear solvers in 3D*, in "Numerical Mathematics: Theory, Methods and Applications", August 2010, vol. 3, n^{o} 3, p. 276-294, http://hal.inria.fr/inria-00542450/en.

[7] P. KOVAL, D. FOERSTER, O. COULAUD. *A Parallel Iterative Method for Computing Molecular Absorption Spectra*, in "Journal of Chemical Theory and Computation", 2010, vol. 6, n^{o} 9, p. 2654–2668 [*DOI :* 10.1021/CT100280X], http://hal.inria.fr/inria-00488048/en.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] F. DUPROS. *Contribution à la modélisation numérique de la propagation des ondes sismiques sur architectures multicoeurs et hiérarchiques*, Université Sciences et Technologies - Bordeaux I, December 2010.

[9] B. LATHUILIÈRE. *Méthode de décomposition de domaine pour les équations du transport simplifié en neutronique*, Université Sciences et Technologies - Bordeaux I, February 2010.

[10] N. RICHART. *Conception et mise en oeuvre d'une plate-forme de pilotage de simulations numériques parallèles et distribuées*, Université Sciences et Technologies - Bordeaux I, January 2010, http://hal.inria.fr/tel-00464406/en.

Articles in International Peer-Reviewed Journal

[11] E. AGULLO, C. COTI, T. HERAULT, J. LANGOU, S. PEYRONNET, A. REZMERITA, F. CAPPELLO, J. DONGARRA. *QCG-OMPI: MPI Applications on Grids*, in "Future Generation Computer Systems", 2010 [*DOI :* 10.1016/J.FUTURE.2010.11.015], http://hal.inria.fr/inria-00548895/en.

[12] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *Parallel hierarchical hybrid linear solvers for emerging computing platforms*, in "Compte Rendu de l'Académie des Sciences : Série Mécanique", 2010 [*DOI :* 10.1016/J.CRME.2010.11.005], http://hal.inria.fr/inria-00548916/en.

[13] E. AGULLO, A. GUERMOUCHE, J.-Y. L'EXCELLENT. *Reducing the I/O Volume in an Out-of-core Sparse Multifrontal Solver*, in "SIAM Journal on Scientific Computing", 2010, vol. 31, n^{o} 6, p. 4774-4794, http://hal.inria.fr/inria-00548902/en.

[14] M. BARRAULT, B. LATHUILIÈRE, P. RAMET, J. ROMAN. *Efficient Parallel Resolution of The Simplified Transport Equations in Mixed-Dual Formulation*, in "Journal of Computational Physics", 2011, vol. 230, n^{o} 5, p. 2004-2020, available online Dec. 2010 [*DOI :* 10.1016/J.JCP.2010.11.047], http://hal.inria.fr/hal-00547406/en.

[15] F. DUPROS, F. DE MARTIN, E. FOERSTER, D. KOMATITSCH, J. ROMAN. *High-performance finite-element simulations of seismic wave propagation in three-dimensional non linear inelastic geological media*, in "Parallel Computing", January 2010, vol. 36, p. 308-325 [*DOI :* 10.1016/J.PARCO.2009.12.011], http://hal.inria.fr/inria-00528479.

[16] L. GIRAUD, S. GRATTON, X. PINEL, X. VASSEUR. *Flexible GMRES with deflated restarting*, in "SIAM Journal on Scientific Computing", October 2010, vol. 32, n^{o} 4, p. 1858–1878 [*DOI :* 10.1137/080741847], http://hal.inria.fr/inria-00542426/en.

[17] L. GIRAUD, A. HAIDAR, S. PRALET. *Using multiple levels of parallelism to enhance the performance of domain decomposition solvers*, in "Parallel Computing", July 2010, vol. 36, n^{o} 5-6, p. 285-296 [*DOI :* 10.1016/J.PARCO.2009.12.006], http://hal.inria.fr/inria-00542444/en.

[18] L. GIRAUD, A. HAIDAR, Y. SAAD. *Sparse approximations of the Schur complement for parallel algebraic hybrid linear solvers in 3D*, in "Numerical Mathematics: Theory, Methods and Applications", August 2010, vol. 3, n^{o} 3, p. 276-294, http://hal.inria.fr/inria-00542450/en.

[19] P. KOVAL, D. FOERSTER, O. COULAUD. *A Parallel Iterative Method for Computing Molecular Absorption Spectra*, in "Journal of Chemical Theory and Computation", 2010, vol. 6, n^{o} 9, p. 2654–2668 [*DOI :* 10.1021/CT100280X], http://hal.inria.fr/inria-00488048/en.

[20] P. KOVAL, D. FOERSTER, O. COULAUD. *Fast construction of the Kohn–Sham response function for molecules*, in "Physica Status Solidi B", February 2010, vol. 247, n^{o} 8, p. 1841–1848 [*DOI :* 10.1002/PSSB.200983811], http://hal.inria.fr/inria-00457652/en.

International Peer-Reviewed Conference/Proceedings

[21] E. AGULLO, C. AUGONNET, J. DONGARRA, H. LTAIEF, R. NAMYST, J. ROMAN, S. THIBAULT, S. TOMOV. *Dynamically scheduled Cholesky factorization on multicore architectures with GPU accelerators.*, in "Symposium on Application Accelerators in High Performance Computing (SAAHPC)", United States Knoxville, July 2010, http://hal.inria.fr/inria-00547616/en.

[22] E. AGULLO, H. BOUWMEESTER, J. DONGARRA, J. KURZAK, J. LANGOU, L. ROSENBERG. *Towards an efficient tile matrix inversion of symmetric positive definite matrices on multicore architectures*, in "9th International Meeting on High Performance Computing for Computational Science (VecPar'10)", United States Berkeley, 2010, http://hal.inria.fr/inria-00548906/en.

[23] E. AGULLO, C. COTI, J. DONGARRA, T. HERAULT, J. LANGOU. *QR Factorization of Tall and Skinny Matrices in a Grid Computing Environment*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)", United States Atlanta, 2010, http://hal.inria.fr/inria-00548900/en.

[24] A. GUERMOUCHE, H. RENARD. *A First Step to the Evaluation of SimGrid in the Context of a Complex Application*, in "Heterogeneity in Computing Workshop", United States, May 2010, p. 1-10, http://hal.inria.fr/hal-00549253/en.

[25] B. HADRI, H. LTAIEF, E. AGULLO, J. DONGARRA. *Tile QR Factorization with Parallel Panel Processing for Multicore Architectures*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2010)", United States Atlanta, 2010, http://hal.inria.fr/inria-00548899/en.

[26] N. RAVEU, L. GIRAUD, H. BAUDRAND. *WCIP acceleration*, in "Asia Pacific Microwave Conference", Japan Yokohama, December 2010, http://hal.inria.fr/inria-00551035/en.

[27] J. SOUMAGNE, J. BIDDISCOMBE, J. CLARKE. *An HDF5 MPI virtual file driver for parallel in-situ post-processing*, in "Recent Advances in the Message Passing Interface - 17th European MPI Users' Group Meeting, EuroMPI 2010", Germany Berlin, Heidelberg, R. KELLER, E. GABRIEL, M. M. RESCH, J. DONGARRA (editors), Lecture Notes in Computer Science, Springer-Verlag, 2010, vol. 6305, p. 62–71 [*DOI :* 10.1007/978-3-642-15646-5_7], http://hal.inria.fr/inria-00551982/en.

[28] J. SOUMAGNE, J. BIDDISCOMBE, J. CLARKE. *In-situ Visualization and Analysis of SPH Data using a ParaView Plugin and a Distributed Shared Memory Interface*, in "5th International SPHERIC Workshop", United Kingdom Manchester, B. ROGERS (editor), 2010, p. 186-193, http://hal.inria.fr/inria-00551987/en.

Workshops without Proceedings

[29] E. AGULLO, J. DONGARRA, B. HADRI, H. LTAIEF. *Tile QR factorization with parallel panel processing for multicore architectures*, in "SIAM Conference on Parallel Processing for Scientific Computing (PP10)", United States Seattle, 2010, http://hal.inria.fr/inria-00548907/en.

[30] E. AGULLO, J. DONGARRA, R. NATH, S. TOMOV. *Autotuning dense linear algebra libraries on multicore architectures*, in "6th International Workshop on Parallel Matrix Algorithms and Applications (PMAA'10)", Switzerland Basel, June 2010, http://hal.inria.fr/inria-00548921/en.

[31] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, Y. LEE-TIN-YIEN, J. ROMAN. *Algebraic parallel domain decomposition solver*, in "Journée GNR MOMAS / GDR Calcul", France Paris, May 2010, http://hal.inria.fr/inria-00549132/en.

[32] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *MaPHyS or the development of a parallel algebraic domain decomposition solver in the course of the solstice project*, in "Sparse Days 2010 Meeting at CERFACS", France Toulouse, June 2010, http://hal.inria.fr/inria-00548915/en.

[33] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *Parallel scalability and complexity analysis of sparse hybrid linear solvers*, in "6th International Workshop on Parallel Matrix Algorithms and Applications (PMAA'10)", Switzerland Basel, June 2010, http://hal.inria.fr/inria-00549126/en.

[34] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *Towards a complexity analysis of sparse hybrid linear solvers*, in "Para 2010 : State of the Art in Scientific and Parallel Computing", Iceland Reykjavík, June 2010, http://hal.inria.fr/inria-00548913/en.

[35] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, J. ROMAN. *Complexity analysis of sparse hybrid linear solvers*, in "3rd "Scheduling in Aussois" Workshop", France Aussois, June 2010, http://hal.inria.fr/inria-00549221/en.

[36] E. AGULLO, A. HAIDAR, L. GIRAUD, A. GUERMOUCHE, Y. LEE-TIN-YIEN, J. ROMAN. *Solveurs parallèles creux hybrides direct/itératif : application à des simulations 3D*, in "Workshop Solveurs de systèmes linéaires de grande taille : les avancées récentes", France Lyon, November 2010, http://hal.inria.fr/inria-00549144/en.

[37] M. CHANAUD, D. GOUDIN, J.-J. PESQUÉ, J. ROMAN. *Hybrid direct-multigrid solver for large 3D electromagnetic problems*, in "11th Copper Mountain Conference on Iterative Methods", United States Copper Mountain, April 2010, http://hal.inria.fr/inria-00555065/en.

[38] M. CHANAUD, D. GOUDIN, J.-J. PESQUÉ, J. ROMAN. *Parallel solver coupling multigrid and direct methods for Maxwell equations*, in "SIAM Conference on Parallel Processing for Scientific Computing (PP10)", United States Seattle, Washington, February 2010, http://hal.inria.fr/inria-00555059/en.

Scientific Books (or Scientific Book chapters)

[39] E. AGULLO, C. AUGONNET, J. DONGARRA, H. LTAIEF, R. NAMYST, S. THIBAULT, S. TOMOV. *Faster, Cheaper, Better – a Hybridization Methodology to Develop Linear Algebra Software for GPUs*, in "GPU Computing Gems", WEN-MEI W. HWU (editor), Morgan Kaufmann, September 2010, vol. 2, http://hal.inria.fr/inria-00547847/en.

Research Reports

[40] L. GIRAUD, A. HAIDAR, Y. SAAD. *Sparse approximations of the Schur complement for parallel algebraic hybrid linear solvers in 3D*, INRIA, March 2010, n^{O} RR-7237, http://hal.inria.fr/inria-00466828/en.