Activity Report 2011

Project-Team HIEPACS

High-End Parallel Algorithms for Challenging Numerical Simulations

Table of contents

<div align="center">

Project-Team HIEPACS

</div>

Keywords: High Performance Computing, Multiscale Models, Code Coupling, Parallel Algorithms, Parallel Solver

HiePACS is a research initiative of the joint Inria-CERFACS Laboratory on High Performance Computing (https://inria-cerfacs.inria.fr/).

1. Members

Research Scientists

    Luc Giraud [Inria, Team Leader since Oct. 2011, Senior Researcher (DR), HdR]
    Emmanuel Agullo [Inria, Junior Researcher (CR)]
    Olivier Coulaud [Inria, Senior Researcher (DR)]
    Jean Roman [Inria, Team Leader until Sept. 2011, Senior Researcher (DR on secondment), Professor at IPB, 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]
    Yohan Lee-Tin-Yien [Inria, funding from ADT ParScaLi, ended Dec. 2011]
    Bérenger Bramas [Inria, funding from ADT ScalFMM, since Dec. 2010]
    Matthieu Lecouvez [Inria, funding from HiePACS, since Oct. 2011]

PhD Students

    Rached Abdelkhalek [TOTAL, CIFRE, since Jan. 2008]
    Mathieu Chanaud [Inria, funding from Inria and CEA, defended Oct. 2011]
    Yohann Dudouit [CERFACS, funding from TOTAL, since Oct. 2010]
    Arnaud Etcheverry [Inria, funding from ANR-OPTIDIS, since Oct. 2011]
    Andra Hugo [Inria, funding from Inria-Région Aquitaine, since Oct. 2011]
    Stojce Nakov [Inria, funding from TOTAL, since Oct. 2011]
    Alexis Praga [CERFACS, since Oct. 2011]
    Pablo Salas Medina [Inria, funding from Europe FP7/ITN (Marie Curie) MyPlanet, since June 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 Sept. 2010]
    Mawussi Zounon [Inria, funding from ANR-RESCUE, since Oct. 2011]

Post-Doctoral Fellows

    Mikko Byckling [CERFACS, ALGO Project, since Oct. 2010]
    Yan-Fei Jing [Inria, since Oct. 2011]

Administrative Assistants

    Barta Beneddine [Inria, ended June 2011]
    Marie Sanchez [Inria, July-Oct. 2011]
    Chrystel Plumejeau [Inria, since Nov. 2011]

Others

Florent Lopez [Internship, July-Aug. 2011]
Andra Hugo [Internship, March-Sept. 2011]
Stojce Nakov [Internship, March-Sept. 2011]
Lucas Verdonk [Internship, Feb-Aug. 2011]
Mawussi Zounon [Internship, March-Sept. 2011]

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://science.energy.gov/ascr/news-and-resources/program-documents/).

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 emerging 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/), EESI (European Exascale Software Initiative – http://www.eesi-project.eu/pages/menu/homepage.php) 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 sparse 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

- With the Inria GRAND-LARGE Project-Team, we are involevd in the G8 project entitled "Enabling Climate Simulation at Extreme Scale" (https://wiki.engr.illinois.edu/display/G8/G8+ESC++–+Enabling+Climate+Simulations+at+Extreme+Scale) which also involves research groups from Europe, Japan and North America.

- With University of Tennessee (ICL) and University of Colorado at Denver an associated team has been initiated, which name is MORSE (http://www.inria.fr/en/teams/morse). The goal of Matrices Over Runtime Systems at Exascale (MORSE) project is to design dense and sparse linear algebra methods that achieve the fastest possible time to an accurate solution on large-scale multicore systems with GPU accelerators, using all the processing power that future high end systems can make available. To develop software that will perform well on petascale and exascale systems with thousands of nodes and millions of cores, several daunting challenges have to be overcome, both by the numerical linear algebra and the runtime system communities. By designing a research framework for describing linear algebra algorithms at a high level of abstraction, the MORSE team will enable the strong collaboration between research groups in linear algebra and run-time systems needed to develop methods and libraries that fully benefit from the potential of future large-scale machines. The first outcome of this associated team is the release of the MAGMA package (http://icl.cs.utk.edu/magma/).

- The thesis of Mathieu Chanaud (in collaboration with CEA/CESTA) has led to the design and the parallel implementation of an hybrid solver combining a parallel sparse direct solver and full multigrid cycles. A 1.3 billion unknown sparse linear system, arising from the discretization of the 3D Maxwell equations on a fully unstructured mesh, has been solved very efficiently on the CEA/DAM TERA100 supercomputer.

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, Iain Duff, Pierre Fortin, Luc Giraud, Abdou Guermouche, Andra Hugo, Guillaume Latu, Stojce Nakov, Jean Roman, Mawussi Zounon.

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 run-time (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. From this perspective, in the context of the PhD of Andra Hugo, we investigate new approaches for composing parallel applications within a runtime system for heterogeneous platforms. The main goal of this work is to build an improved runtime system which is able to deal with parallel tasks (which may use different parallelization schemes or even different parallelization supports). 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 framework that hosts all these research activities is the associated team MORSE (http://www.inria.fr/en/teams/morse).

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.

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, Luc Giraud, Abdou Guermouche, Andra Hugo, Yan-Fei Jing, Matthieu Lecouvez, Yohan Lee-Tin-Yien, Jean Roman, Pablo Salas Medina, Stojce Nakov, Xavier Vasseur, Mawussi Zounon.

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 that started last 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 have considered a full geometric multigrid solver for the solution of methodology for solving large linear systems arising from Maxwell equations discretized with first-order Nédelec elements on fully unstructued meshes. 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. Mathieu Chanaud defended his PhD in October 2011.

The direct solver can handle linear systems with up to a few tens of 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.

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. In that context, Yan-Fei Jing, who joint HiePACSas post-doc, is investigating how reliable block Arnoldi procedure can be combined with deflated restarted block GMRES technique.

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 of this technique to build a sparse approximation of the Schur complement for algebraic domain decomposition has also been investigated and could be integrated in the `MaPHySpackage` in the future.

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: Bérenger Bramas, Arnaud Etcheverry, 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. Funded by the ANR-OPTIDIS, Arnaud Etcheverry started a PhD in October to study parallel scalable FMM techniques for the dislocation calculations.

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, 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 developed 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 lattice. 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

Participants: Bérenger Bramas, Olivier Coulaud, Aurélien Esnard, Pierre Fortin, Luc Giraud, Jean Roman.

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

Participants: Emmanuel Agullo, Mikko Byckling, Mathieu Chanaud, Luc Giraud, Abdou Guermouche, Yohan Lee-Tin-Yien, Stojce Nakov, Jean Roman, Xavier Vasseur.

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.

For the solution of large systems arsing from PDE discretization, the geometric full multigrid technique based on a few levels in the grid hierarchy and an efficient parallel sparse direct solver on the coarsest level can be considered. Originally developed for 3D Maxwell solution in collaboration with CEA-CESTA, the approach can be extended to other application fields.

Many 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 such as implicit time tepping schemes or inverse iterations. Many of the numerical approaches and possible outcoming software are well suited to tackle these challenging problems.

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).

4.4. Scalable numerical schemes for scientific applications

Participants: Rached Abdelkhalek, Olivier Coulaud, Yohann Dudouit, Luc Giraud, Guillaume Latu, Alexis Praga, Jean Roman, Pablo Salas Medina, Xavier Vasseur.

We are also collaborating with application research group to design or improve numerical schemes in the view of large scale parallel simulations.

Seismic wave propagation in heterogeneous media requires to properly capture the local heterogeneity and consequently requires locally refined meshes. In close collaboration with TOTAL we study new parallelizable schemes for the solution of the elastodynamic system with local spatial refinments based on discontinuous Galerkin techniques. The objective is to design novel parallel scalable implementations for large 3D simulations. A second work is currently carried on with TOTAL for Seismic modeling and Reverse Time Migration (RTM) based on the full wave equation discretization. These tools are of major importance since they give an accurate representation of complex wave propagation areas. Unfortunately, they are highly compute intensive. To address this challenge we have designed a fast parallel simulator that solves the acoustic wave equation on a `GPU` cluster.

Thermoacoustic instabilities are an important concern in the design of gas turbine combustion chambers. Most modern combustion chambers have annular shapes and this leads to the appearance of azimuthal acoustic modes. These modes are often powerful and can lead to structural vibrations being sometimes damaging. Therefore, they must be identified at the design stage in order to be able to eliminate them. However, due to the complexity of industrial combustion chambers with a large number of burners, numerical studies of real configurations are a challenging task. Such a challenging calculations performed in close collaboration with the Computational Fluid Dynamic project at CERFACS.

The chemistry and transport models (CTM) play a central role in global geophyscal models. The solution of the CTM represents up-to 50 % on the computing ressources involved in global geophyscal simulations. Therefore, the availability of efficient scalable parallel numerical schemes on emerging and future supercomputers is crucial. The purpose of this research activity is to study, design and implement novel numerical schemes following the work initiated by D. Cariolle in the framework of the ANR Solstice project. Alexi Praga, PhD hired by CERFACS, is conducing this research action under the joint supervision of `HiePACS` and the Aviation and Environment project at CERFACS in close collaboration with CNRM/Meteo-France.

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 for the 2009-2011 period. 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 first public release of the software is planned eraly 2012. 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 (Environment for Computational Steering) is a software environment for the steering of legacy parallel-distributed simulations with simple GUI or more complex (possibly parallel) visualization programs (see Figure 1). In order to make a legacy simulation steerable, the user annotates the sourcecode 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 remotely accessible. EPSN provides a distributed data model, that handles common scientific objects such as parameters, structured grids, particles/atoms and unstructured meshes. It is then possible to dynamically connect EPSN with a client program, that provides a GUI with some visualization & interaction features, as for instance SIMONE (SImulation MONitoring for Epsn). Once a client is connected, it interacts with the simulation via EPSN API. It is possible : 1) to control the execution flow of the remote simulation; 2) to access/modify its data onthefly; and 3) finally to invoke advanced user-defined routines in the simulation. The current version of EPSN is fully based on CORBA for communication on heterogeneous system and VTK/Paraview for visualization. A new release of EPSN, that will be fully based on MPI to handle efficient communication, is currently under development. A prototype is already working.

EPSN has been supported by the ACI-GRID program (grant number PPL02-03), the ARC RedGRID, the ANR MASSIM (grant number ANR-05-MMSA-0008-03) and the ANR CIS NOSSI (2007). More informations are available on our web site: http://www.labri.fr/projet/epsn. This software is publicly available at Inria Gforge (http://epsn.gforge.inria.fr).

5.4. MPICPL

MPICPL (MPI CouPLing) is a software library dedicated to the coupling of parallel legacy codes, that are based on the well-known MPI standard. It proposes a lightweight and comprehensive programing interface that simplifies the coupling of several MPI codes (2, 3 or more). MPICPL facilitates the deployment of these codes thanks to the *mpicplrun* tool and it interconnects them automatically through standard MPI inter-communicators. Moreover, it generates the universe communicator, that merges the world communicators of all coupled-codes. The coupling infrastructure is described by a simple XML file, that is just loaded by the *mpicplrun* tool. Future releases will incorporate new features for checkpoint/restart and dynamic parallel code connection.
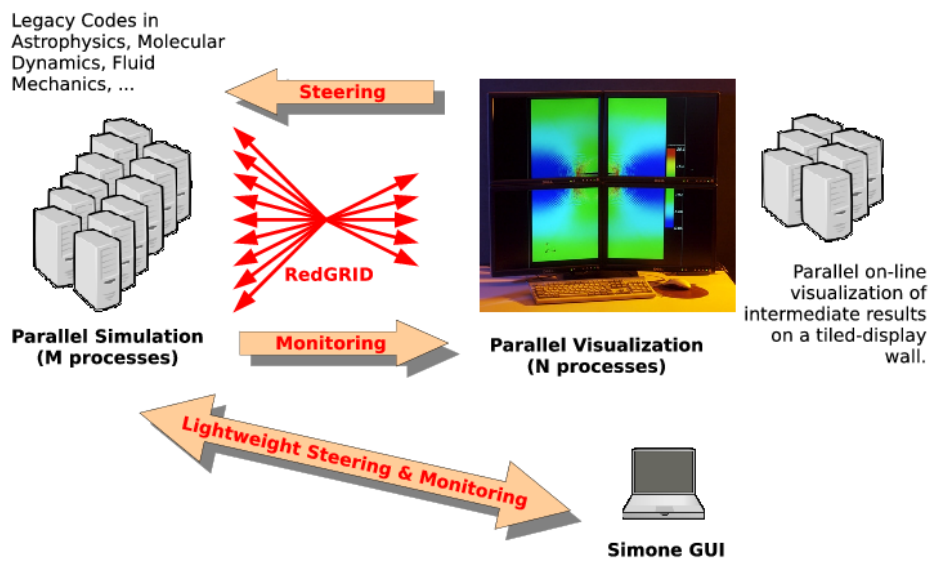
Figure 1. EPSN: software environment for $M \times N$ computational steering.

MPICPL was developed by the Inria HiePACS project-team for the purpose of the ANR CIS NOSSI. It uses advanced features of MPI2 standard. The framework is publicly available at Inria Gforge: http://mpicpl.gforge.inria.fr.

5.5. MONIQA

MONIQA (MONitoring graphic user Interface for Qm/mm Applications) is a GUI specially designed for the monitoring & steering of the QM/MM application in the ANR CIS NOSSI project. It is based on Tulip, a graph visualization software http://tulip.labri.fr), used to display atoms and molecules. It proposes two working modes : offline or online. The offline mode is mainly used to load input files of DL_POLY & Siesta, and to prepare the quantum region for the QM/MM coupling. In online mode, the end-user can monitor & interact with the running QM/MM application thanks to EPSN. It is thus possible to visualize molecular and physical data (distances, angles, charges, energies), and to change simulation parameters on-the-fly, such as the target temperature of the system, thermo or barostat parameters, verbosity of output, ... MONIQA is based on QT4. It was developed specifically for the ANR NOSSI project and is available (restricted access) at Inria Gforge: http://nossi.gforge.inria.fr.

5.6. ScalFMM

ScalFMM (Parallel Fast Multipole Library for Large Scale Simulations) is a software library to simulate N-body interactions using the Fast Multipole Method. ScalFMM is based on the FMB prototype developed by Pierre Fortin during his PhD thesis.

In the framework of the INRIA technologic development actions; 24 man-month engineer (Bérenger Bramas) have been allocated to this software activity started in January 2011.

ScalFMM intends to offer all the functionalities needed to perform large parallel simulations while enabling an easy customization of the simulation components: kernels, particles and cells. It works in parallel in a shared/distributed memory model using OpenMP and MPI. The software architecture has been designed with two major objectives: being easy to maintain and easy to understand. The code is extremely documented and the naming convention fully respected. Driven by its user-oriented philosophy, ScalFMM is using CMAKE as a compiler/installer tool. Even if ScalFMM is written in C++ it will support a C and fortran API soon.

The ScalFMM package is very much a first outcome of the research activity described in Section 3.4.

5.7. 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.

- 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 have developed a new parallel plateform based on a combination of a geometric full 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).

6. New Results

6.1. Algorithms and high-performance solvers

6.1.1. *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. This work has been integrated in the latest release of the MAGMA package (http://icl.cs.utk.edu/magma/).

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

A first release of the MaPHyS package should be made available early in 2012 thanks to the developments conducted in the last year of the ADT. An approximation of the local Schur complement has been studied that is based on approximated inverse technique. This work is a natural extension of part of the PhD research of Mikko Byckling. Furthermore, during his master internship, Stojce Nakov has investigated the design of a Krylov subspace method, namely the conjugate gradient, on a run-time system in order to best exploit the computing capabilites of many-GPU nodes and manycore systems. In the framework of his starting PhD funded by TOTAL, Stojce Nakov will continue his work to design a new implementation of a hybrid linear solver (see Section 3.3) for heterogeneous manycore platforms.

6.1.3. *Resilience in numerical simulations*

In his master internship work, Mawussi Zounon investigated recovery strategies for core faults in the framework of parallel preconditioned Krylov solvers. The underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. He will continue this work in the framework of his PhD funded by the ANR-RESCUE. Notice that theses activities are also part of our contribution to the G8-ECS (Enabling Climate Simulation at extreme scale).

6.1.4. *Full geometric multigrid method for 3D Maxwell equations*

In the context of a collaboration with the CEA/CESTA center, Mathieu Chanaud continued his PhD work on 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 novel parallel solver was able to solve a 1.3 billion system given a 20 million unknown problem at the coarsest level. 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.

6.1.5. *Scalable numerical schemes for scientific applications*

A 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 early 2012.

For the solution of the elastodynamic equation on meshes with local refinments, we are currently collaborating with Total to design a parallel implementation of a local time refinement technique on top of a discontinuous Galerkin space discretization. This latter technique enables to manage non-conforming meshes suited to deal with multiblock approaches that capture the locally refined regions. this work is developed in the framework of Yohann Dudouit PhD thesis. A software prototype is currently developed to address these simulations.

The calculation of acoustic modes in combustion chambers is a challenging calculation for large 3D geometries. It requires the calculation of a few of the smallest eigenpairs of large unsymmetric matrices in a paralell environment. A new block Arnoldi approach is currently developed to best benefit from the continuation scheme used in this application context. This is part of the PhD research activity of Pablo Salas.

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. The key idea is to use a graph/hypergraph model that will incorporate information about the coupling itself. Then, we propose new algorithms to perform a coupling-aware partitioning in order to improve the load-balancing of the whole coupled simulation.

Let us consider two coupled codes, modeled by two graphs (or hypergraphs) A and B, connected by inter-edges $I(A, B)$ that represents the coupling communications between codes. Formally, the problem consists in partitioning A in M and B in N with accounting for $I(A, B)$. This algorithm should optimize both the edge cut for each graph and the coupling communications while maintaining each graph balance. Our general strategy is divided in three main steps:

1. first, we freely partition A in M to obtain the partition A/M;
2. then, we projects this partition to B according to $I(A, B)$, that provides the partition B/M;
3. finally, we compute the partition B/N by repartitioning B from M existing parts into N.

The final repartitioning step is particularly tiedous, because it must handle a variable number of processes. However, as far as we know, the state-of-the-art graph/hypergraph repartitioning tools are limited to a fixed number of processes (i.e. $M = N$). To overcome this issue, we have proposed a new repartitioning algorithm – assuming the load is constant – based on hypergraph partitioning technics with fixed vertices. Our algorithm uses an *optimal* communication pattern, that we have proved to minimize the total number of messages between the former and newer parts. Experimental results validate our work comparing it with other approaches [20]. We currently investigate how to extend our algorithm for the dynamic load-balancing of parallel adaptive codes ($A = B$), whose load evolution is variable and difficult to predict. In this case, it would be convenient to dynamically adapt the number of processes used at runtime ($M \neq N$), while minimizing migration cost during the repartitioning step. This work is currently conducted in the framework of Clément Vuchener PhD thesis.

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

As a different approach of EPSN, we conceived and developed an in-transit visualization framework for interfacing an arbitrary HPC simulation code with an interactive ParaView session using the HDF5 parallel IO library as the API. The library called H5FDdsm is coupled with a ParaView plugin ICARUS (Initialize Compute Analyze Render Update Steer).

Because our interface is based on files, stored in a distributed shared memory (DSM), we sought during this year different redistribution strategies to optimize the bandwidth and the transfers between the simulation and the ParaView servers hosting the DSM. This work showed real benefits, particularly on one of our Cray XE6 testing machines using a block cyclic redistribution. On these large HPC machines that do not support the dynamic MPI process management set of functions, we improved our connection system so that simulation and post-processing can be coupled within an MPMD job. Taking also advantage of one-sided communication models and of the Cray Gemini interconnect communication performance, our framework has been sensibly improved and should be optimal in the coming months.

The interface has also been enhanced with a steering interface that allows us to control the simulation work-flow and send back not only parameters, but also complete meshes in parallel, which can then be read by the simulation using either our steering interface or HDF5 calls. This has been demonstrated with SPH-flow, a CFD code developed by Ecole Centrale de Nantes and HydrOcean, replacing dynamically and in parallel a falling wedge with a deforming sphere.

This work has been realized and is currently carried on at CSCS - Swiss National Supercomputing Centre in the framework of Jérôme Soumagne PhD thesis (under the co-supervision of Mr. John Biddiscombe) and within the NextMuSE European project 7th FWP/ICT-2007.8.0 ([17], [18], [19]).

6.4. Material physics

6.4.1. *Hybrid materials*

The study of hybrid materials based on a coupling between molecular dynamics (MD) and quantum mechanism (QM) simulation has been conducted in collaboration with IPREM (Pau) within the ANR CIS 2007 NOSSI (ended December 2011). 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 was 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 have been investigated in the project and new results obtained as described bellow.

Optical spectra.

Some new improvements in our TD-DFT code have been introduced. Our method is based on the LCAO method for densities and excited states that computes electronic excitation spectra. We have worked in two directions:

- As the method introduces a regularization parameter to obtain regularized spectra we have used it to build better algorithms. In particular, we have developed a new hierarchical algorithm that builds a well adapted frequency distribution to better capture the biggest peaks (strongest oscillator strengths) in the spectrum. Moreover, a nonlinear fit method was added and used to compute the transitions and the oscillator strengths of the spectrum.

- In our algorithm, we used a coarse grain paradigm to parallelize the spectrum computation. This approach leads to a memory bottleneck for large systems. In that respect, we have explored a new parallel approach based on a fine grain paradigm (matrix-vector parallelization) to better exploit the manycore achitecture of the emerging computers.

Finally, we have improved the packaging of the code to prepare a public release of the code. Our TD-DFT code will be soon available on request.

QM/MM algorithm. For structure studies or dynamical properties, we have coupled QM model based on pseudo-potentials (SIESTA code) with dynamic molecular (DL-POLY code). Therefore we have developed a new algorithm to avoid accounting twice for the forces and the quantum electric field in the molecular model. All algorithms involved in the coupling have been introduced both in SIESTA and in DL-POLY codes. The following new developpements needed by the coupling have been introduced in the SIESTA code:

- We have implemented a fast evaluation of the molecular electrostatic field on the quantum grid.

- We have introduced a non periodic Poisson solver based on the parallel linear Hypre solver. This solver allows us to use computation domains as small as possible.

- We have implemented the ElectroStatic Potential (ESP) fit method to obtain more physical point charges than those given by SIESTA with the Mulliken method. These point charges are used by the MM codes to compute electrostatic forces.

Thanks to all our develpements introduced in SIESTA a collaboration with the SIESTA research team has started. This enables us to have acces to their private svn like repository. Preliminary resuls on a water dimer and a water box systems show good agreement with other methods developed in SIESTA and DL-POLY teams.

All these results were presented in the final international NOSSI workshop in Biarritz on december.

6.4.2. *Material failures*

We have started in the context of the OPTIDIS ANR to work on dislocation simulations. The main characteristic of these simulations is that they are highly dynamical. This year, we have started the study of the state of the art on this topic in two directions. The first direction concerns the study of the algorithms used in such simulations and how we can efficiently parallize them on manycore clusters. In the second one for isotropic materials, we are investigating how to adapt our fast multipole method to compute constraints and then forces in this kind of simulations.

7. Contracts and Grants with Industry

7.1. Contracts with Industry

CEA research and development contract:

- 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).

TOTAL research and development contracts:

- Parallel hybrid solver for massivelly heterogeneoux manycore platforms (Stojce Nakov (PhD); Emmanuel Agullo, Luc Giraud, Abdou Guermouche, Jean Roman).

- Parallel elastodynamic solver for 3D models with local mesh refinment (Yohann Dudouit (PhD); Luc Giraud and Sébastien Pernet from ALGO-EMA at CERFACS).

8. Partnerships and Cooperations

8.1. National initiatives

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

Participants: Olivier Coulaud, Aurélien Esnard.

Grant: ANR 2007 – CIS
Dates: 2008 – 2011
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. *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.

8.1.3. *RESCUE: RÉsilience des applications SCientifiqUEs*

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

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.

8.1.4. *BOOST: Building the future Of numerical methOdS for iTer*

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

Grant: ANR-Blanc (applied math theme)
Dates: 2010 – 2014
Partners: Institut de Mathématiques de Toulouse (coordinator); Laboratoire d'Analyse, Topologie, Probabilités in Marseilles; Institut de Recherche sur la Fusion Magnétique, CEAr/IRFM and Inria-HiePaCS
Overview: This project regards the study and the development of a new class of numerical methods to simulate natural or laboratory plasmas and in particular magnetic fusion processes. In this context, we aim in giving a contribution, from the mathematical, physical and algorithmic point of view, to the ITER project.

The core of this project consists in the development, the analysis, the implementation and the testing on real physical problems of the so-called Asymptotic-Preserving methods which allow simulations over a large range of scales with the same model and numerical method. These methods represent a breakthrough with respect to the state-of-the art. They will be developed specifically to handle the various challenges related to the simulation of the ITER plasma. In parallel with this class of methodologies, we intend to design appropriate coupling techniques between macroscopic and microscopic models for all the cases in which a net distinction between different regimes can be done. This will permit to describe different regimes in different regions of the machine with a strong gain in term of computational efficiency, without losing accuracy in the description of the problem. We will develop full 3-D solver for the asymptotic preserving fluid as well as kinetic model. The Asymptotic-Preserving (AP) numerical strategy allows us to perform numerical simulations with very large time and mesh steps and leads to impressive computational saving. These advantages will be combined with the utilization of the last generation preconditioned fast linear solvers to produce a software with very high performance for plasma simulation. For HiePACS this project provides in particular a testbed for our expertise in parallel solution of large linear systems.

8.2. European Initiatives

8.2.1. FP7 Projet

Title: MYPLANET

Type: PEOPLE

Instrument: Initial Training Network (ITN)

Duration: October 2008 - September 2012

Coordinator: CERFACS (France)

Others partners: 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.

See also: http://www.cerfacs.fr/myplanet/

Abstract: 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.

8.3. International initiatives

8.3.1. Inria Associate Teams

8.3.1.1. MORSE

Title: Matrices Over Runtime Systems at Exascale

Inria principal investigator: Emmanuel Agullo

International Partner:

> Institution: University of Tennessee Knoxville (United States)
>
> Laboratory: Innovative Computing Lab
>
> Researcher: George Bosilca

International Partner:

> Institution: University of Colorado Denver (United States)
>
> Laboratory: Department of Mathematics and Statistical Sciences
>
> Researcher: Julien Langou

Duration: 2011 - 2013

See also: http://www.inria.fr/en/teams/morse.

The goal of Matrices Over Runtime Systems at Exascale (MORSE) project is to design dense and sparse linear algebra methods that achieve the fastest possible time to an accurate solution on large-scale multicore systems with GPU accelerators, using all the processing power that future high end systems can make available. To develop software that will perform well on petascale and exascale systems with thousands of nodes and millions of cores, several daunting challenges have to be overcome, both by the numerical linear algebra and the runtime system communities. By designing a research framework for describing linear algebra algorithms at a high level of abstraction, the MORSE team will enable the strong collaboration between research groups in linear algebra and runtime systems needed to develop methods and libraries that fully benefit from the potential of future large-scale machines. Our project will take a pioneering step in the effort to bridge the immense software gap that has opened up in front of the High-Performance Computing (HPC) community.

8.3.2. *Visits of International Scientists*

The following researchers have visited `HiePACS` in 2011

- George Bosilca, University of Tennessee at Knoxville visited from June 15 to August 15.

- Ichitaro Yamazaki, from Lawrence Berkeley National Laboratory visited from August 29 to September 9.

- Hatem Ltaief, from KAUST visited from October 10 to October 14.

8.3.3. *Participation in other International Programs*

8.3.3.1. 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.

8.3.3.2. ECS : Enabling Climate Simulation at extreme scale
Participants: Emmanuel Agullo, Luc Giraud, Abdou Guermouche, Jean Roman, Mawussi Zounon.

Grant: G8
Dates: 2011 – 2014
Partners: Univ. Illinois at Urbanna Champaign, Inria, Univ. Tennessee at Knoxville, German Research School for Simulation Sciences, Univ. Victoria, Titech, Univ. Tsukuba, NCAR, Barcelona Supercomputing Center.

Overview: Exascale systems will allow unprecedented reduction of the uncertainties in climate change predictions via ultra-high resolution models, fewer simplifying assumptions, large climate ensembles and simulation at a scale needed to predict local effects. This is essential given the cost and consequences of inaction or wrong actions about climate change. To achieve this, we need careful co-design of future exascale systems and climate codes, to handle lower reliability, increased heterogeneity, and increased importance of locality. Our effort will initiate an international collaboration of climate and computer scientists that will identify the main roadblocks and analyze and test initial solutions for the execution of climate codes at extreme scale. This work will provide guidance to the future evolution of climate codes. We will pursue research projects to handle known roadblocks on resilience, scalability, and use of accelerators and organize international, interdisciplinary workshops to gather and disseminate information. The global nature of the climate challenge and the magnitude of the task strongly favor an international collaboration. The consortium gathers senior and early career researchers from USA, France, Germany, Spain, Japan and Canada and involves teams working on four major climate codes (CESM1, EC-EARTH, ECSM, NICAM).

9. Dissemination

9.1. Animation of the scientific community

Olivier Coulaud has been member of the scientific committee of the international conference Supercomputing 2011 and local chair of the topic "High Performance and Scientific Applications" at EuroPar 2011. He is member of the Inria COST GTAI committe (in charge of incentive actions), of the C3I GENCI committe and of the scientific board of the regional computing mesocentre. Moreover, he is the leader of the Inria PlaFRIM computing plateform.

Luc Giraud has been co-chair of the Application area at Supercomputing 2011, member of the scientific committee of the international conferences PDSEC-11, PDCN-11 and local chair of the topic "Parallel numerical algorithms" at EuroPar 2011. He was member of the selection committee for the ANR MN programme. He was also involved in the EESI working group 4.3 entitled "Numerical Libraries, Solvers and Algorithms". Luc Giraud was co-organizer of the EDF-CEA-Inria scientific school (Toward petaflop numerical simulation on parallel hybrid architectures) which took place at the begining of june 2011.

Jean Roman was president of the Project Committee of Inria Bordeaux - Sud-Ouest and member of the National Evaluation Committee of Inria until june 2011. He has been member of the scientific committee of the international conference EuroMicro PDP'11 (IEEE) and of the national conference Renpar'11. He was one of the co-chairs of EuroPar 2011. Jean Roman was co-organizer of the EDF-CEA-Inria scientific school (Toward petaflop numerical simulation on parallel hybrid architectures) which took place at the begining of june 2011. 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. He is now in charge at the national level of the Inria scientific activities concerning High Performance Computing.

Finally, the HiePACS members have contributed to the reviewing process of several international journals (BIT, Concurrency and Computation: Practice and Experience (CCPE), Computers and Fluids, Geophysical Prospecting, IEEE Trans. on Parallel and Distributed Systems, International Journal for Numerical Methods in Fluids, Parallel Computing, SIAM J. Scient. Comp., SIAM J. Numer. Analysis), to the reviewing process of international conferences (Supercomputing 2011, EuroPar 2011, IPDPS 2011, ICPP 2011, ...), to the referee of PhD dissertations for various French universities (ENS-Lyon/computer science, Lyon 1/applied maths, Rennes/computer science, Pau/Comp. Chemistry, ...) and have acted as experts for the research agency ANR-MN.

9.2. Teaching

In the following are listed the lectures given by the HiePACS members.

Undergraduate level

1. A. Esnard: Operating system programming, 36h, University Bordeaux I; Using network, 23h, University Bordeaux I.

2. A. Esnard: in charge of the computer science certificate for Internet (C2i) at the University Bordeaux I.

Post graduate level

1. O. Coulaud: Paradigms for parallel computing, 28h, ENSEIRB-MatMeca, Talence; Code coupling, 6h, ENSEIRB-MatMeca, Talence.

2. A. Esnard: Network management, 27h, University Bordeaux I; Programming distributed applications, 60h, ENSEIRB-MatMeca, Talence.

3. L. Giraud: Introduction to intensive computing and related programming tools, 20h, INSA Toulouse; Introduction to high performance computing and applications, 20h, ISAE-ENSICA; On mathematical tools for numerical simulations, 10h, ENSEEIHT Toulouse; Parallel sparse linear algebra, 11h, ENSEIRB-MatMeca, Talence.

4. A. Guermouche: Network management, 92h, University Bordeaux I; Network security, 64h, University Bordeaux I; Operating system, 24h, University Bordeaux I.

5. J. Roman: Parallel sparse linear algebra, 10h, ENSEIRB-MatMeca, Talence; Parallel algorithms, 22h, ENSEIRB-MatMeca, Talence.

6. X. Vasseur: Solution of PDE, 16 h, ENSEEIHT Toulouse; Linear Algebra and Optimization, 25 h, ISEA-ENSICA, Toulouse; Introduction to MPI, 11 h, ENM, Toulouse; Introduction to Fortran 90, 5 h, CERFACS, Toulouse.

Doctorate level

1. E. Agullo: "Dense linear algebra on manycores", CEA-EDF-Inria School, June 6-10, 2 hours, 2011.

2. E. Agullo, A. Guermouche, L. Giraud, J. Roman and X. Vasseur: Formation en algèbre linéaire creuse parallèle, November 28-December 2, 30 hours, Maison de la Simulation.

Defended PhD thesis

1. Mathieu Chanaud, *"Conception d'un solveur haute performance de systèmes linéaires creux couplant des méthodes multigrilles et directes pour la résolution des équations de Maxwell 3D en régime harmonique discrétisées par éléments finis*, Université Bordeaux I, defended on 18 Oct. 2011, advisors: D. Goudin (CEA) and J. Roman.

PhD in progress :

1. Rached Abdelkhalek, *"Modélisation et imagerie sismique sur accélérateurs matériels"* , starting Jan. 2008, advisors: O. Coulaud, G. Latu and J. Roman.

2. Yohann Dudouit, *"Scalable parallel elastodynamic solver with local refinement in geophysics"*, starting Oct. 2010, advisors: L. Giraud and S. Pernet (CERFACS).

3. Arnaud Etcheverry, *"Toward large scale dynamic dislocation simulation on petaflop computers"*, starting Oct. 2011, advisors: O. Coulaud and J. Roman.

4. Andra Hugo *"Composabilité de codes parallèles sur plateformes hétérogènes"*, starting Oct. 2011, advisors: A. Guermouche, R. Namyst and P-A. Wacrenier.

5. Alexis Praga, *"Parallel atmospheric chemistry and transport model solver for massivelly platforms"*, starting Oct. 2011, advisors: D. Cariolle (CERFACS) and L. Giraud.

6. Stojce Nakov, *"Parallel hybrid solver for heterogenous manycores: application to geophysics"*, starting Oct. 2011, advisors: E. Agullo and J. Roman.

7. Pablo Salas Medina, *"Parallel eigensolvers for large scale combustion chamber simulations"*, starting June 2010, advisors: L. Giraud and X. Vasseur.

8. Jérôme Soumagne, *"An In-situ Visualization Approach for Parallel Coupling and Steering of HPC Applications using Files in Distributed Shared Memory"*, starting April 2009, advisors: J. Bidiscombe, A. Esnard and J. Roman.

9. Clément Vuchener, *"Algorithmique de l'équilibrage de charges pour des couplages de codes complexes"*, starting Sept. 2010, advisors: A. Esnard and J. Roman.

10. Mawussi Zounon, *"Numerical resilient algorithms for exascale"*, starting Oct. 2011, advisors: E. Agullo and L. Giraud.

10. Bibliography

Major publications by the team in recent years

[1] R. ABDELKHALEK, H. CALANDRA, O. COULAUD, G. LATU, J. ROMAN. *Fast seismic modeling and reverse time migration on a graphics processing unit cluster*, in "Concurrency and Computation: Practice and Experience", 2011 [*DOI :* 10.1002/CPE.1875], http://hal.inria.fr/hal-00653499/en.

[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, n⁰ 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, n⁰ 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, n⁰ 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, n⁰ 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] M. CHANAUD. *Conception d'un solveur haute performance de systèmes linéaires creux couplant des méthodes multigrilles et directes pour la résolution des équations de Maxwell 3D en régime harmonique discrétisées par éléments finis*, Université Sciences et Technologies - Bordeaux I, October 2011.

Articles in International Peer-Reviewed Journal

[9] R. ABDELKHALEK, H. CALANDRA, O. COULAUD, G. LATU, J. ROMAN. *Fast seismic modeling and reverse time migration on a graphics processing unit cluster*, in "Concurrency and Computation: Practice and Experience", 2011 [*DOI :* 10.1002/CPE.1875], http://hal.inria.fr/hal-00653499/en.

[10] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, J. ROMAN. *Parallel hierarchical hybrid linear solvers for emerging computing platforms*, in "Compte Rendu de l'Académie des Sciences : Serie Mécanique", 2011, vol. 339, n° 2-3, p. 96-103 [*DOI :* 10.1016/J.CRME.2010.11.005], http://hal.inria.fr/inria-00548916/en.

[11] 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° 5, p. 2004-2020 [*DOI :* 10.1016/J.JCP.2010.11.047], http://hal.inria.fr/hal-00547406/en.

[12] C. LUIZ MARIANO, S. GRATTON, R. LAGO, X. VASSEUR. *A Flexible Generalized Conjugate Residual Method with Inner Orthogonalization and Deflated Restarting*, in "SIAM Journal on Matrix Analysis and Applications", 2011, vol. 32, n° 4, p. 1212-1235 [*DOI :* 10.1137/100786253], http://hal.inria.fr/hal-00650239/en.

[13] F. SOURBIER, A. HAIDAR, L. GIRAUD, H. BEN HADJ ALI, S. OPERTO, J. VIRIEUX. *Three-dimensional parallel frequency-domain visco-acoustic wave modelling based on a hybrid direct/iterative solver*, in "Geophysical Prospecting", 2011, vol. 59, n° 5, p. 834–856 [*DOI :* 10.1111/J.1365-2478.2011.00966.X], http://hal.inria.fr/hal-00647179/en.

[14] I. TARRASS, L. GIRAUD, P. THORE. *New curvilinear scheme for elastic wave propagation in presence of curved topography*, in "Geophysical Prospecting", 2011, vol. 59, n° 5, p. 889–906 [*DOI :* 10.1111/J.1365-2478.2011.00972.X], http://hal.inria.fr/hal-00647160/en.

International Conferences with Proceedings

[15] E. AGULLO, C. AUGONNET, J. DONGARRA, M. FAVERGE, J. LANGOU, H. LTAIEF, S. TOMOV. *LU Factorization for Accelerator-based Systems*, in "9th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 11)", Sharm El-Sheikh, Égypte, June 2011, http://hal.inria.fr/hal-00654193/en.

[16] E. AGULLO, C. AUGONNET, J. DONGARRA, M. FAVERGE, H. LTAIEF, S. THIBAULT, S. TOMOV. *QR Factorization on a Multicore Node Enhanced with Multiple GPU Accelerators*, in "25th IEEE International Parallel & Distributed Processing Symposium", Anchorage, États-Unis, May 2011, http://hal.inria.fr/inria-00547614/en.

[17] J. BIDDISCOMBE, J. SOUMAGNE, G. OGER, D. GUIBERT, J.-G. PICCINALI. *Parallel Computational Steering and Analysis for HPC Applications using a ParaView Interface and the HDF5 DSM Virtual File Driver*, in "Eurographics Symposium on Parallel Graphics and Visualization", Llandudno, Wales, Royaume-Uni, T. KUHLEN, R. PAJAROLA, K. ZHOU (editors), Eurographics Association, 2011, p. 91–100, Honourable Mention Award [*DOI :* 10.2312/EGPGV/EGPGV11/091-100], http://hal.inria.fr/hal-00651814/en.

[18] J. SOUMAGNE, J. BIDDISCOMBE, A. ESNARD. *Data Redistribution using One-sided Transfers to In-memory HDF5 Files*, in "Proceedings of the 18th European MPI Users' Group conference on Recent advances in the message passing interface", Santorini, Grèce, Y. COTRONIS, A. DANALIS, D. NIKOLOPOULOS, J.

DONGARRA (editors), EuroMPI'11, Springer Berlin / Heidelberg, 2011, p. 198–207 [*DOI :* 10.1007/978-3-642-24449-0_23], http://hal.inria.fr/hal-00648396/en.

[19] J. SOUMAGNE, J. BIDDISCOMBE. *Computational Steering and Parallel Online Monitoring Using RMA through the HDF5 DSM Virtual File Driver*, in "Proceedings of the International Conference on Computational Science, ICCS 2011", Singapour, 2011, vol. 4, p. 479–488 [*DOI :* 10.1016/J.PROCS.2011.04.050], http://hal.inria.fr/hal-00652816/en.

National Conferences with Proceeding

[20] C. VUCHENER, A. ESNARD. *Repartitionnement d'un graphe de M vers N processeurs : application pour l'équilibrage dynamique de charge*, in "20ème Rencontres francophones du parallélisme (RenPar'20)", Saint-Malo, France, 2011, 8, http://hal.inria.fr/hal-00648404/en.

Conferences without Proceedings

[21] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, Y. LEE-TIN-YIEN, J. ROMAN. *Performance of a sparse hybrid linear solver on a multicore cluster*, in "SIAM Conference on Computational Science and Engineering (CSE 2011)", Reno, États-Unis, February 2011, http://hal.inria.fr/inria-00548911/en.

[22] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *Parallel algebraic domain decomposition solver for the solution of augmented systems*, in "The Second International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering", Ajaccio, France, April 2011, http://hal.inria.fr/hal-00652931/en.

[23] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN, S. LANTERI. *Algebraic Schwarz preconditioning for the Schur complement: application to the time-harmonic Maxwell equations discretized by a discontinuous Galerkin method*, in "The Twentieth International Conference on Domain Decomposition Methods", San Diego La Jolla, États-Unis, February 2011, http://hal.inria.fr/inria-00577639/en.

[24] P. SALAS, L. GIRAUD, J.-D. MÜLLER, G. STAFFELBACH, T. POINSOT. *Stability of azimuthal modes in annular combustion chambers*, in "Third workshop INCA", Toulouse, France, November 2011, http://hal.inria.fr/hal-00651847/en.

Research Reports

[25] E. AGULLO, J. DONGARRA, R. NATH, S. TOMOV. *Fully Empirical Autotuned QR Factorization For Multicore Architectures*, INRIA, February 2011, n^o RR-7526, http://hal.inria.fr/inria-00569514/en.

[26] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, A. HAIDAR, J. ROMAN. *Parallel algebraic domain decomposition solver for the solution of augmented systems*, INRIA, January 2011, n^o RR-7516, http://hal.inria.fr/inria-00559133/en.

Other Publications

[27] X. VASSEUR, S. GRATTON, X. PINEL. *A Multigrid Method for the Solution of Linear Systems with Multiple Right-Hand Sides*, May 2011, Session 8, http://hal.inria.fr/inria-00580794/en.