Activity Report 2013

Project-Team HIEPACS

High-End Parallel Algorithms for Challenging Numerical Simulations

Table of contents

Project-Team HIEPACS

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

Creation of the Team: 2009 January 01*, updated into Project-Team:* 2010 January 01.

1. Members

Research Scientists

Luc Giraud [Team leader, Inria, Senior Researcher, HdR]
Emmanuel Agullo [Inria, Researcher]
Olivier Coulaud [Inria, Senior Researcher, HdR]
Jean Roman [Inria, Senior Researcher half-time in HIEPACS, Professor at IPB on secondment, HdR]

Faculty Members

Aurélien Esnard [Univ. Bordeaux I, Associate Professor]
Mathieu Faverge [IPB, Associate Professor]
Abdou Guermouche [Univ. Bordeaux I, Associate Professor]
Pierre Ramet [Univ. Bordeaux I, Associate Professor, joined HIEPACS from Bacchus in Jan 2013]

Engineers

Matías Hastaran [Inria, from Oct 2013]
Julien Pedron [Inria, from Jan 2013]
Florent Pruvost [Inria, SED development engineer 80 % time in HIEPACS from Nov 2013]
François Rue [Inria, SED development engineer 80 % time in HIEPACS from Nov 2013]
Cyrille Piacibello [Inria, from Feb 2013]

PhD Students

Andra Hugo [Inria, Inria-Conseil Régional Aquitaine]
Pierre Blanchard [ENS Cachan, from Sep 2013]
Bérenger Bramas [Inria]
Astrid Casadei [Univ. Bordeaux I, joined HIEPACS from Bacchus in Jan 2013]
Jean-Marie Couteyen [EADS, Astrium Space Transportation, from Feb 2013]
Yohann Dudouit [CERFACS, Total E&P Recherche Developement]
Arnaud Etcheverry [Inria, ANR OPTIDIS project]
Xavier Lacoste [Inria, ANR ANEMOS project, joined HIEPACS from Bacchus in Jan 2013]
Salli Moustafa [EDF, CIFRE]
Stojce Nakov [Inria, Total E&P Recherche Developpement]
Alexis Praga [CERFACS]
Maria Predari [Inria, Inria-Conseil Régional Aquitaine from Oct 2013]
Fabien Rozar [CEA Cadarache, IRFM and Maison de la Simulation, from Feb 2013]
Pablo Salas Medina [Inria, FP7 ITN MyPlanet project, until October 2013]
Mawussi Zounon [Inria, ANR RESCUE project]
Clément Vuchener [Univ. Bordeaux I, until Nov 2013]

Post-Doctoral Fellows

Romain Garnier [Inria, Inria-Conseil Régional Aquitaine from Nov 2013]
Matthias Messner [Inria, until Feb 2013]

Visiting Scientists

George Bosilca [Inria, Total E&P Recherche Developement, until Nov 2013]
Guillaume Latu [CEA Cadarache, IRFM]
Pierre Brenner [Astrium Space Transportation]

Benoît Lize [EADS-IW]
Guillaume Sylvand [EADS-IW]
Xavier Vasseur [CERFACS]
Administrative Assistant
Chrystel Plumejeau [Inria]
Others
Thomas Geffroy [Inria, Intership from Jun 2013 until Oct 2013]
Guillaume Boutin [Inria, Internship from Mar 2013 until Sep 2013]
Pierre Gandanger [Inria, Internship from Oct 2013]
Benjamin Lux [Inria, Internship from Mar 2013 until Sep 2013]
Grégoire Pichon [Inria, Internship from Jun 2013 until Aug 2013]

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 available and exaflop computers are foreseen in early 2020.

For application codes to sustain petaflops 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 modeling, super scalable algorithms that will require paying particular attention to acquisition, management and visualization of huge amounts of scientific data.

In this context, the purpose of the HIEPACS project is to contribute performing efficiently frontier simulations arising from challenging research and industrial that are likely to be *multiscale* and *coupled* 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 hierarchical many-core 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) EESI-2 (European Exascale Software Initiative 2) and likely in the framework of H2020 European collaborative projects.

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 extreme scale platforms. Next, several research prospectives in scalable parallel linear algebra techniques are addressed, ranging from dense direct, sparse direct, iterative and hybrid approaches for large linear systems. Then we consider research plans for N-body interaction computations based on efficient parallel fast multipole methods and finally, we adress research tracks related to the algorithmic challenges for complex code couplings in multiscale simulations.

Currently, we have one major multiscale application that is in *material physics*. We contribute to all steps of the design of the parallel simulation tool. More precisely, our applied mathematics skill will contribute to the modeling 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. Finally, we are also involved in a few collaborative intiatives in various application domains in a co-design like framework. These reserach activities are conducted is a wider multi-disciplinary context with collegues in other academic or industrial groups where our contribution is related to our expertises. Not only these collaborations enable our knowledges to have a stronger impact in various application domains through the promotion of advanced algorithms, methodologies or tools, but they also open new avenues for research in the continuity of our core research.

Thanks to the two Inria collaborative agreements such as with EADS-Astrium/Conseil Régional Aquitaine and with CEA, we have joint research efforts in a co-design framework enabling efficient and effective technological transfer towards industrial R&D.

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 GENCI in France (CCRT, CINES and IDRIS Computational Centers), and next the high-end parallel platforms that will be available via European and US initiatives or projects such that PRACE.

2.2. Highlights of the Year

The `PaStiX` solver is now able to handle efficiently multiple GPU accelerators using runtime systems (`StarPU` or `PaRSEC`). On the Plafrim machine, one GPU card can provide almost the same performance than 12 cores and we get a good scalability while mixing multicores and upto 3 GPUs accelerators.

The first implementation of the Fast Multipole method over a runtime system has been developed in the context of the FASTLA associated team. The main outcome of this work will be published in a paper to appear in the SIAM SISC journal.

3. Research Program

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 extreme scale platforms.

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, which are involved in the IPL C2S@Exa, will be the basis of both academic and industrial collaborations, some are described in Section 4.1, 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 modeling 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.

Currently, we have one major application (see Section 4.1) that is in material physics. We will contribute to all steps of the design of the parallel simulation tool. More precisely, our applied mathematics skill will contribute to the modelling, our advanced numerical schemes will help in the design and efficient software implementation for very large parallel multi-scale simulations. We also participate to a few co-design actions in close collaboration with some applicative groups. The objective of this activity is to instanciate our expertise in fields where they are critical for designing scalable simulation tools. We refer to Section 4.2 for a detailed description of these activities.

3.2. High-performance computing on next generation architectures

Participants: Emmanuel Agullo, Olivier Coulaud, Luc Giraud, Mathieu Faverge, Abdou Guermouche, Matías Hastaran, Andra Hugo, Xavier Lacoste, Guillaume Latu, Stojce Nakov, Florent Pruvost, Pierre Ramet, 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 many-core 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 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 problems. This has to be done in collaboration with our colleagues from the scheduling community like for example O. Beaumont (Inria REALOPT 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 developed 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 tools 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, 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, we investigate new approaches for composing parallel applications within a runtime system for heterogeneous platforms.

The SOLHAR project aims at studying and designing algorithms and parallel programming models for implementing direct methods for the solution of sparse linear systems on emerging computers equipped with accelerators. Several attempts have been made to accomplish the porting of these methods on such architectures; the proposed approaches are mostly based on a simple offloading of some computational tasks (the coarsest grained ones) to the accelerators and rely on fine hand-tuning of the code and accurate performance modeling to achieve efficiency. SOLHAR proposes an innovative approach which relies on the efficiency and portability of runtime systems, such as the `StarPU` tool developed in the RUNTIME team. Although the SOLHAR project will focus on heterogeneous computers equipped with GPUs due to their wide availability and affordable cost, the research accomplished on algorithms, methods and programming models will be readily applicable to other accelerator devices. 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 direct and 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 and collaborating research groups. The objectives will be to develop dense linear algebra algorithms and libraries for multicore architectures in the context the `PLASMA` project and for GPU and hybrid multicore/GPU architectures in the context of the `MAGMA` project. The framework that hosts all these research activities is the associated team 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 ROMA and GRAND-LARGE and the G8 ESC international initiative. 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, Astrid Casadei, Olivier Coulaud, Mathieu Faverge, Romain Garnier, Luc Giraud, Abdou Guermouche, Andra Hugo, Pablo Salas Medina, Stojce Nakov, Julien Pedron, Florent Pruvost, Pierre Ramet, Jean Roman, Xavier Vasseur.

Starting with the developments of basic linear algebra kernels tuned for various classes of computers, a significant knowledge on the basic concepts for implementations on high-performance scientific computers has been accumulated. Further knowledge has been acquired through the design of more sophisticated linear algebra algorithms fully exploiting those basic intensive computational kernels. In that context, we still look at the development of new computing platforms and their associated programming tools. This enables us to identify the possible bottlenecks of new computer architectures (memory path, various level of caches, inter processor or node network) and to propose ways to overcome them in algorithmic design. With the goal of designing efficient scalable linear algebra solvers for large scale applications, various tracks will be followed in order to investigate different complementary approaches. Sparse direct solvers have been for years the methods of choice for solving linear systems of equations, it is nowadays admitted that classical 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. We will continue to work on sparse direct solvers on one hand to make sure they fully benefit from most advanced computing platforms on the other hand because they are a key 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. In this framework, and possibly in relation with the research activity on fast multipole, we intend to study how emerging H-matrix arithmetic can benefit to our solver research efforts.

3.3.1. Parallel sparse direct solver

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

Sparse direct solvers are mandatory when the linear system is very ill-conditioned; such a situation is often encountered in structural mechanics codes, for example. Therefore, to obtain an industrial software tool that must be robust and versatile, high-performance sparse direct solvers are mandatory, and parallelism is then necessary for reasons of memory capability and acceptable solution time. Moreover, in order to solve efficiently 3D problems with more than 50 million unknowns, which is now a reachable challenge with new multicore supercomputers, we must achieve good scalability in time and control memory overhead. Solving a sparse linear system by a direct method is generally a highly irregular problem that induces some challenging algorithmic problems and requires a sophisticated implementation scheme in order to fully exploit the capabilities of modern supercomputers.

New supercomputers incorporate many microprocessors which include themselves one or many computational cores. These new architectures induce strongly hierarchical topologies. These are called NUMA architectures. In the context of distributed NUMA architectures, in collaboration with the Inria RUNTIME team, we study optimization strategies to improve the scheduling of communications, threads and I/O. We have developed

dynamic scheduling designed for NUMA architectures in the `PaStiX` solver. The data structures of the solver, as well as the patterns of communication have been modified to meet the needs of these architectures and dynamic scheduling. We are also interested in the dynamic adaptation of the computation grain to use efficiently multi-core architectures and shared memory. Experiments on several numerical test cases have been performed to prove the efficiency of the approach on different architectures.

In collaboration with the ICL team from the University of Tennessee, and the RUNTIME team from Inria, we are evaluating the way to replace the embedded scheduling driver of the `PaStiX` solver by one of the generic frameworks, `PaRSEC` or `StarPU`, to execute the task graph corresponding to a sparse factorization. The aims is to design algorithms and parallel programming models for implementing direct methods for the solution of sparse linear systems on emerging computer equipped with GPU accelerators. More generally, this work will be performed in the context of the ANR SOLHAR project which aims at designing high performance sparse direct solvers for modern heterogeneous systems. The project involves several groups working either on the sparse linear solver aspects (HIEPACS and ROMA from Inria and APO from IRIT), on runtime systems (RUNTIME from Inria) or scheduling algorithms (REALOPT and ROMA from Inria). The results of these efforts will be validated in the applications provided by the industrial project members, namely CEA-CESTA and EADS-IW.

On the numerical side, we are studying how the data sparsness that might exist in some dense blocks appearing during the factorization can be exploited using different compression techniques based on H-matrix (and variants) arithmetics. This research activity will be conducted in the framework of the FASTLA associate team and will naturally irrigate the hybrid solvers described below as well as closely interact with the other research efforts where similar data sparsness might be exploited.

3.3.2. 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 hierarchically 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 intend to continue our effort on the design of algebraic non-overlapping domain decomposition techniques that rely on the solution of a Schur complement system defined on the interface introduced by the partitioning of the adjacency graph of the sparse matrix associated with the linear system. Although it is better conditioned than the original system the Schur complement needs to be precondition to be amenable to a solution using a Krylov subspace method. Different hierarchical preconditioners will be considered, possibly multilevel, to improve the numerical behaviour of the current approaches implemented in our software libraries `HIPS` and `MaPHyS`. In addition to this numerical studies, advanced parallel implementation will be developed that will involve close collaborations between the hybrid and sparse direct activities. In particular some additional work to complete the initial study conducted with CEA-CESTA on full multigrid method will be undertaken. This activity will be developed either in the framework of the CEA-Inria agreement and/or through joint work with bresilian colleagues within the HOSCAR initiative.

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 side 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. Following the initial work by J. Yan Fei during his post-doc in HiePACS, we will continue the numerical study of the block GMRES variant that combine inexact break-down detection and deflation at restart. In addition a special attention will be paid to situations where a massive number of right-hand sides are given where variants exploiting the possible sparsness (i.e., compression using H-matrix arithmetic) of these right-hand sides will be explored to design efficient numerical algorithms. Beyond new numerical investigations, a software implementation to be included in our linear solver library will be developed.

  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.

- 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. This research activity will benefit from the starting FP7 EXA2CT project led by HiePACS on behalf of the IPL C2S@Exa that involves two other Inria projects namely ALPINES and SAGE.

3.3.4. *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 two 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 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.

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

Participants: Emmanuel Agullo, Bérenger Bramas, Arnaud Etcheverry, Olivier Coulaud, Matthias Messner, Cyrille Piacibello, Guillaume Sylvand.

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 accuracy. 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. 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 achieve high performance.

3.4.4. *Fast multipole method for boundary element methods*

The boundary element method (BEM) is a well known solution of boundary value problems appearing in various fields of physics. With this approach, we only have to solve an integral equation on the boundary. This implies an interaction that decreases in space, but results in the solution of a dense linear system with $O(N^3)$ complexity. The FMM calculation that performs the matrix-vector product enables the use of Krylov

subspace methods. Based on the parallel data distribution of the underlying octree implemented to perform the FMM, parallel preconditioners can be designed that exploit the local interaction matrices computed at the finest level of the octree. This research action will be conduced on close connection with the activity described in Section 3.3. Following our earlier experience, we plan to first consider approximate inverse preconditionners that can efficiently exploit these data structures.

3.5. Efficient algorithmic for load balancing and code coupling in complex simulations

Participants: Astrid Casadei, Olivier Coulaud, Aurélien Esnard, Maria Predari, Pierre Ramet, Jean Roman, 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.

3.5.1. Efficient schemes for multiscale simulations

As mentioned previously, many important physical phenomena, such as material deformation and failure (see Section 4.1), are inherently multiscale processes that cannot always be modeled via continuum model. Fully microscopic 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, 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 ANR NOSSI and the atomic to dislocation coupling in the ANR OPTIDIS.

3.5.2. Dynamic load balancing for massively parallel coupled codes

In this context of code coupling, one crucial issue is undoubtedly the load balancing of the whole coupled simulation that remains an open question. The goal here is to find the best data distribution for the whole coupled simulation and not only for each standalone code, as it is most usually done. Indeed, the naive balancing of each code on its own can lead to an important imbalance and to a communication bottleneck during the coupling phase, that can drastically decrease the overall performance. Therefore, one argues that it is required to model the coupling itself in order to ensure a good scalability, especially when running on massively parallel architectures (tens of thousands of processors/cores). In other words, one must develop new algorithms and software implementation to perform a *coupling-aware* partitioning of the whole application.

Another related problem is the problem of resource allocation. This is particularly important for the global coupling efficiency and scalabilty, 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 each code to avoid that one of them wait for the other(s). And what happens if the load of one code dynamically changes relatively to the other? In such a case, it could be convenient to dynamically adapt the number of resources used at runtime.

For instance, the conjugate heat transfer simulation in complex geometries (as developed by the CFD team of CERFACS) requires to couple a fluid/convection solver (AVBP) with a solid/conduction solver (AVTP). The AVBP code is much more CPU consuming than the AVTP code. As a consequence, there is an important computational imbalance between the two solvers. The use of new algorithms to correctly load balance coupled simulations with enhanced graph partitioning techniques appears as a promising way to reach better performances of coupled application on massively parallel computers.

3.5.3. *Graph partitioning for hybrid solvers*

Graph handling and partitioning play a central role in the activity described here but also in other numerical techniques detailed in Section 3.3.

The Nested Dissection is now a well-known heuristic for sparse matrix ordering to both reduce the fill-in during numerical factorization and to maximize the number of independent computation tasks. By using the block data structure induced by the partition of separators of the original graph, very efficient parallel block solvers have been designed and implemented according to supernodal or multifrontal approaches. Considering hybrid methods mixing both direct and iterative solvers such as HIPS or MaPHyS, obtaining a domain decomposition leading to a good balancing of both the size of domain interiors and the size of interfaces is a key point for load balancing and efficiency in a parallel context. We intend to revisit some well-known graph partitioning techniques in the light of the hybrid solvers and design new algorithms to be tested in the Scotch package.

4. Application Domains

4.1. Material physics

Participants: Pierre Blanchard, Olivier Coulaud, Arnaud Etcheverry, Matthias Messner.

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 algorithms and their parallelization and in the second way, we will consider a multiscale approach.

A 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 modeling 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) modeling. 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 currently under definition OPTIDIS 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.

4.2. Co-design for scalable numerical algorithms in scientific applications

Participants: Pierre Brenner, Jean-Marie Couteyen, Luc Giraud, Xavier Lacoste, Guillaume Latu, Salli Moustapha, Pierre Ramet, Fabien Rozar, Jean Roman, Pablo Salas, Xavier Vasseur.

The research activities concerning the ITER challenge are involved in the Inria Project Lab (IPL) C2S@Exa.

4.2.1. *MHD instabilities Edge Localized Modes*

The numerical simulations tools designed for ITER challenges aim at making a significant progress in understanding of largely unknown at present physics of active control methods of plasma edge MHD instabilities Edge Localized Modes (ELMs) which represent particular danger with respect to heat and particle loads for Plasma Facing Components (PFC) in ITER. Project is focused in particular on the numerical modeling study of such ELM control methods as Resonant Magnetic Perturbations (RMPs) and pellet ELM pacing both foreseen in ITER. The goals of the project are to improve understanding of the related physics and propose possible new strategies to improve effectiveness of ELM control techniques. The tool for the nonlinear MHD modeling (code JOREK) will be largely developed within the present project to include corresponding new physical models in conjunction with new developments in mathematics and computer science strategy in order to progress in urgently needed solutions for ITER.

The fully implicit time evolution scheme in the JOREK code leads to large sparse linear systems that have to be solved at every time step. The MHD model leads to very badly conditioned matrices. In principle the PaStiX library can solve these large sparse problems using a direct method. However, for large 3D problems the CPU time for the direct solver becomes too large. Iterative solution methods require a preconditioner adapted to the problem. Many of the commonly used preconditioners have been tested but no satisfactory solution has been found. The research activities presented in Section 3.3 will contribute to design new solution techniques best suited for this context.

4.2.2. *Turbulence of plasma particules inside a tokamak*

In the context of the ITER challenge, the GYSELA project aims to simulate the turbulence of plasma particules inside a tokamak. Thank to a better comprehension of this phenomenon, it would be possible to design a new kind of source of energy based of nuclear fusion. Currently, GYSELA is parallelized in a MPI/OpenMP way and can exploit the power of the current greatest supercomputers (e.g., Juqueen). To simulate faithfully the plasma physic, GYSELA handles a huge amount of data. In fact, the memory consumption is a bottleneck on large simulations (449 K cores). In the meantime all the reports on the future Exascale machines expect a decrease of the memory per core. In this context, mastering the memory comsumption of the code becomes critical to consolidate its scalability and to enable the implementation of new features to fully benefit from the extreme scale architectures.

In addition to activities for designing advanced generic tools for managing the memory optimisation, further algorithmic research will be conduced to better predict and limit the memory peak in order to reduce the memory footprint of GYSELA.

4.2.3. SN cartesian solver for nuclear core simulation

As part of its activity, EDF R&D is developing a new nuclear core simulation code named `COCAGNE` that relies on a Simplified PN (SPN) method to compute the neutron flux inside the core for eigenvalue calculations. In order to assess the accuracy of SPN results, a 3D Cartesian model of PWR nuclear cores has been designed and a reference neutron flux inside this core has been computed with a Monte Carlo transport code from Oak Ridge National Lab. This kind of 3D whole core probabilistic evaluation of the flux is computationally very demanding. An efficient deterministic approach is therefore required to reduce the computation effort dedicated to reference simulations.

In this collaboration, we work on the parallelization (for shared and distributed memories) of the `DOMINO` code, a parallel 3D Cartesian SN solver specialized for PWR core reactivity computations which is fully integrated in the `COCAGNE` system.

4.2.4. 3D aerodynamics for unsteady problems with moving bodies

ASTRIUM has developed for 20 years the `FLUSEPA` code which focuses on unsteady phenomenon with changing topology like stage separation or rocket launch. The code is based on a finite volume formulation with temporal adaptive time integration and supports bodies in relative motion. The temporal adaptive integration classifies cells in several temporal levels, zero being the level with the slowest cells and each level being twice as fast as the previous one. This repartition can evolve during the computation, leading to load-balancing issues in a parallel computation context. Bodies in relative motion are managed through a CHIMERA-like technique which allows building a composite mesh by merging multiple meshes. The meshes with the highest priorities recover the least ones, and at the boundaries of the covered mesh, an intersection is computed. Unlike classical CHIMERA technique, no interpolation is performed, allowing a conservative flow integration. The main objective of this research is to design a scalable version of `FLUSEPA` in order to run efficiently on modern parallel architectures very large 3D simulations.

4.2.5. Nonlinear eigensolvers for thermoacoutic instability calculation

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 3D configurations are a challenging task. The modelling and the discretization of such phenomena lead to the solution of a nonlinear eigenvalue problem of size a few millions.

Such a challenging calculations performed in close collaboration with the Computational Fluid Dynamic project at CERFACS.

5. Software and Platforms

5.1. Introduction

We describe in this section the software that we are developing. The first list 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 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

Participant: Emmanuel Agullo [corresponding member].

`MaPHyS` (Massivelly Parallel Hybrid Solver) is a software package whose prototype was initially developed in the framework of a PhD thesis and further consolidated first thanks to the ANR-CIS Solstice funding and the Inria Parscali ADT. 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.

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.

`MaPHyS` can be found at http://maphys.gforge.inria.fr.

5.3. PaStiX

Participant: Pierre Ramet [corresponding member].

Complete and incomplete supernodal sparse parallel factorizations.

`PaStiX` (Parallel Sparse matriX package) is a scientific library that provides a high performance parallel solver for very large sparse linear systems based on block direct and block ILU(k) iterative methods. Numerical algorithms are implemented in single or double precision (real or complex): LLt (Cholesky), LDLt (Crout) and LU with static pivoting (for non symmetric matrices having a symmetric pattern).

The `PaStiX` library uses the graph partitioning and sparse matrix block ordering package `Scotch`. `PaStiX` is based on an efficient static scheduling and memory manager, in order to solve 3D problems with more than 50 million of unknowns. The mapping and scheduling algorithm handles a combination of 1D and 2D block distributions. This algorithm computes an efficient static scheduling of the block computations for our supernodal parallel solver which uses a local aggregation of contribution blocks. This can be done by taking into account very precisely the computational costs of the BLAS 3 primitives, the communication costs and the cost of local aggregations. We also improved this static computation and communication scheduling algorithm to anticipate the sending of partially aggregated blocks, in order to free memory dynamically. By doing this, we are able to reduce the aggregated memory overhead, while keeping good performance.

Another important point is that our study is suitable for any heterogeneous parallel/distributed architecture when its performance is predictable, such as clusters of multicore nodes. In particular, we now offer a high performance version with a low memory overhead for multicore node architectures, which fully exploits the advantage of shared memory by using an hybrid MPI-thread implementation.

Direct methods are numerically robust methods, but the very large three dimensional problems may lead to systems that would require a huge amount of memory despite any memory optimization. A studied approach consists in defining an adaptive blockwise incomplete factorization that is much more accurate (and numerically more robust) than the scalar incomplete factorizations commonly used to precondition iterative solvers. Such incomplete factorization can take advantage of the latest breakthroughs in sparse direct methods and particularly should be very competitive in CPU time (effective power used from processors and good scalability) while avoiding the memory limitation encountered by direct methods.

`PaStiX` is publicly available at http://pastix.gforge.inria.fr under the Inria CeCILL licence.

5.4. HIPS

Participant: Pierre Ramet [corresponding member].

Multilevel method, domain decomposition, Schur complement, parallel iterative solver.

`HIPS` (Hierarchical Iterative Parallel Solver) is a scientific library that provides an efficient parallel iterative solver for very large sparse linear systems.

The key point of the methods implemented in `HIPS` is to define an ordering and a partition of the unknowns that relies on a form of nested dissection ordering in which cross points in the separators play a special role (Hierarchical Interface Decomposition ordering). The subgraphs obtained by nested dissection correspond to the unknowns that are eliminated using a direct method and the Schur complement system on the remaining of the unknowns (that correspond to the interface between the sub-graphs viewed as sub-domains) is solved using an iterative method (GMRES or Conjugate Gradient at the time being). This special ordering and partitioning allows for the use of dense block algorithms both in the direct and iterative part of the solver and provides a high degree of parallelism to these algorithms. The code provides a hybrid method which blends direct and iterative solvers. `HIPS` exploits the partitioning and multistage ILU techniques to enable a highly parallel scheme where several subdomains can be assigned to the same process. It also provides a scalar preconditioner based on the multistage ILUT factorization.

`HIPS` can be used as a standalone program that reads a sparse linear system from a file ; it also provides an interface to be called from any C, C++ or Fortran code. It handles symmetric, unsymmetric, real or complex matrices. Thus, `HIPS` is a software library that provides several methods to build an efficient preconditioner in almost all situations.

`HIPS` is publicly available at http://hips.gforge.inria.fr under the Inria CeCILL licence.

5.5. LBC2

Participant: Aurélien Esnard [corresponding member].

`LBC2` (Load-Balancing for Code Coupling) is a library providing different methods for partitioning or repartitioning graphs and hypergraphs. These methods are designed for effective communications during dynamic load balancing with a variable number of processors or coupled code interactions. Repartitioning is achieved with an enriched graph model partitioned using third-party partitioning tools, as `Scotch`, `PaToH`, `METIS`, `Zoltan` or `Mondrian`.

The framework is publicy available at Inria Gforge as a part of the MPICPL framework: http://mpicpl.gforge.inria.fr.

5.6. MPICPL

Participant: Aurélien Esnard [corresponding member].

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.

MPICPL was developed by HIEPACS for the purpose of the ANR NOSSI. It uses advanced features of MPI2 standard. The framework is publicy available at Inria Gforge: http://mpicpl.gforge.inria.fr.

5.7. ScalFMM

Participant: Olivier Coulaud [corresponding member].

`ScalFMM` (Parallel Fast Multipole Library for Large Scale Simulations) is a software library to simulate N-body interactions using the Fast Multipole Method.

`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. There is two main parts: 1) the management of the octree and the parallelization of the method ; 2) the kernels. This new architecture allow us to easily add new FMM algorithm or kernels and new paradigm of parallelization. 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 library offers two methods to compute interactions between bodies when the potential decays like $1/r$. The first method is the classical FMM based on spherical harmonic expansions and the second is the Black-Box method which is an independent kernel formulation (introduced by E. Darve at Stanford). With this method, we can now easily add new non oscillatory kernels in our library. For the classical method, two approaches are used to decrease the complexity of the operators. We consider either matrix formulation that allows us to use BLAS routines or rotation matrix to speed up the M2L operator.

The `ScalFMM` package is available at http://scalfmm.gforge.inria.fr

5.8. ViTE

Participant: Mathieu Faverge [corresponding member].

Visualization, Execution trace

`ViTE` is a trace explorer. It is a tool made to visualize execution traces of large parallel programs. It supports Pajé, a trace format created by Inria Grenoble, and OTF and OTF2 formats, developed by the University of Dresden and allows the programmer a simpler way to anlyse, debug and/or profile large parallel applications. It is an open source software licenced under CeCILL-A.

The `ViTE` software is available at http://vite.gforge.inria.fr and has been developed in collaboration with the Inria Bordeaux - Sud-Ouest SED team, Telecom SudParis and Inria Grenoble.

In the same context we also contribute to the EZtrace and GTG libraries in collaboration with F. Trahay from Telecom SudParis. EZTrace (http://eztrace.gforge.inria.fr) is a tool that aims at generating automatically execution trace from HPC programs. It generates execution trace files thanks to the GTG library (http://gtg.gforge.inria.fr) that can be later interpreted by visualization tools such as `ViTE`.

5.9. Other software

For the materials physics applications, a lot of development will be done in the context of ANR projects (NOSSI and OPTIDIS, see Section 4.1) in collaboration with LaBRI, CPMOH, IPREM, EPFL and with CEA Saclay and Bruyère-le-Châtel.

- **FAST**

  Participant: Olivier Coulaud [corresponding member].

  `FAST` is a linear response time dependent density functional program for computing the electronic absorption spectrum of molecular systems. It uses an O(N3) linear response method based on finite numerical atomic orbitals and deflation of linear dependence in atomic orbital product space. This version is designed to work with data produced by the SIESTA DFT code. The code produces as principal output a numerical absorption spectrum (complex part of the polarisability, loosely called the polarisability below) and a list of transition energies and oscillator strengths deduced from fitting Lorentzians to the numerical spectrum. Considering the absence of hybrid functionals in SIESTA and that concerning calculation of spectra, generalized gradient Hamiltonians are not usually considered to be notably better than the local density approximation, the present release of `FAST` works only with LDA, which despite its limitations, has provided useful results on the systems to which the present authors have applied it. The `FAST` library is available at http://people.bordeaux.inria.fr/coulaud/Softwares/FAST/index.html.

- **OptiDis**

Participant: Olivier Coulaud [corresponding member].

OptiDis is a new code for large scale dislocation dynamics simulations. its aim is to simulate real life dislocation densities (up until 5.10^{22} dislocations/m^{-2}) in order to understand plastic deformation and study strain hardening. The main application is to observe and understand plastic deformation on irradiated zirconium. Zirconium alloys is the first containment barrier against the dissemination of radioactive elements. More precisely, with neutron irradiated zirconium alloys we are talking of channeling mechanism, which mean to stick with the reality,more than tens of thousands of induced loops so 10^8 degrees of freedom in the simulation.

The code is based on Numodis code developed at CEA Saclay and the `ScalFMM` library developed in our Inria project. The code is written in C++ language and using the last features of C++11. One of the main aspects is the hybrid parallelism MPI/OpenMP that gives the software the ability to scale on large cluster while the computation load rise. In order to achieve that, we use different level of parallelism. First of all, the simulation box is spread over MPI process, we then use a thinner level for threads, dividing the domain using an Octree representation. All theses parts are driven by the `ScalFMM` library. On the last level our data are stored in an adaptive structure absorbing dynamic of this kind of simulation and handling well task parallelism.

The two following packages are mainly designed and developed in the context of a US initiative led by ICL and to which we closely collaborate through the associate team MORSE.

- **PLASMA**
  Participant: Mathieu Faverge [corresponding member].

  The `PLASMA` (Parallel Linear Algebra for Scalable Multi-core Architectures) project aims to address the critical and highly disruptive situation that is facing the Linear Algebra and High Performance Computing community due to the introduction of multi-core architectures.

  The `PLASMA` ultimate goal is to create software frameworks that enable programmers to simplify the process of developing applications that can achieve both high performance and portability across a range of new architectures.

  The development of programming models that enforce asynchronous, out of order scheduling of operations is the concept used as the basis for the definition of a scalable yet highly efficient software framework for Computational Linear Algebra applications.

  The `PLASMA` library is available at http://icl.cs.utk.edu/plasma.

- **PaRSEC/DPLASMA**
  Participant: Mathieu Faverge [corresponding member].

  `PaRSEC` Parallel Runtime Scheduling and Execution Controller, is a generic framework for architecture aware scheduling and management of micro-tasks on distributed many-core heterogeneous architectures. Applications we consider can be expressed as a Direct Acyclic Graph of tasks with labeled edges designating data dependencies. DAGs are represented in a compact problem-size independent format that can be queried on-demand to discover data dependencies in a totally distributed fashion. `PaRSEC` assigns computation threads to the cores, overlaps communications and computations and uses a dynamic, fully-distributed scheduler based on architectural features such as NUMA nodes and algorithmic features such as data reuse.

  The framework includes libraries, a runtime system, and development tools to help application developers tackle the difficult task of porting their applications to highly heterogeneous and diverse environment.

  `DPLASMA` (Distributed Parallel Linear Algebra Software for Multicore Architectures) is the leading implementation of a dense linear algebra package for distributed heterogeneous systems. It is designed to deliver sustained performance for distributed systems where each node featuring multiple sockets of multicore processors, and if available, accelerators like GPUs or Intel Xeon Phi. `DPLASMA`

achieves this objective through the state of the art `PaRSEC` runtime, porting the `PLASMA` algorithms to the distributed memory realm.

The `PaRSEC` runtime and the `DPLASMA` library are available at http://icl.cs.utk.edu/parsec.

6. New Results

6.1. High-performance computing on next generation architectures

6.1.1. *Composing multiple StarPU applications over heterogeneous machines: a supervised approach*

Enabling HPC applications to perform efficiently when invoking multiple parallel libraries simultaneously is a great challenge. Even if a uniform runtime system is used underneath, scheduling tasks or threads coming from different libraries over the same set of hardware resources introduces many issues, such as resource oversubscription, undesirable cache flushes or memory bus contention.

This paper presents an extension of `StarPU`, a runtime system specifically designed for heterogeneous architectures, that allows multiple parallel codes to run concurrently with minimal interference. Such parallel codes run within *scheduling contexts* that provide confined execution environments which can be used to partition computing resources. Scheduling contexts can be dynamically resized to optimize the allocation of computing resources among concurrently running libraries. We introduce a *hypervisor* that automatically expands or shrinks contexts using feedback from the runtime system (e.g. resource utilization). We demonstrate the relevance of our approach using benchmarks invoking multiple high performance linear algebra kernels simultaneously on top of heterogeneous multicore machines. We show that our mechanism can dramatically improve the overall application run time (-34%), most notably by reducing the average cache miss ratio (-50%).

This work is developed in the framework of Andra Hugo's PhD. These contributions have been presented at the international workshop on Accelerators and Hybrid Exascale Systems [19] in Boston.

6.1.2. *A task-based H-matrix solver for acoustic and electromagnetic problems on multicore architectures*

$\mathcal{H}$-Matrix is a hierarchical, data-sparse approximate representation of matrices that allows the fast approximate computation of matrix products, LU and LDL^T decompositions, inversion and more. This representation is suitable for the direct solution of large dense linear systems arising from the Boundary Element Method in $O(N \log_2^\alpha (N))$ operations. This kind of formulation is widely used in the industry for the numerical simulation of acoustics and electromagnetism scattering by large objetcs. Applications of this approach include aircraft noise reduction and antenna sitting at Airbus Group. The recursive and irregular nature of these $\mathcal{H}$-Matrix algorithms makes an efficient parallel implementation very challenging, especially when relying on a "Bulk Synchronous Parallel" paradigm. We have considered an alternative parallelization for multicore architectures using a task-based approach on top of a runtime system, namely `StarPU`. We have showed that our method leads to a highly efficient, fully pipelined computation on large real-world industrial test cases provided by Airbus Group.

This research activity has been conduced in the framework of the EADS-ASTRIUM, Inria, Conseil Régional initiative in collaboration with the `RUNTIME` Inria project, and is part of Benoit Lize's PhD.

6.1.3. A task-based 3D geophysics application

Reverse Time Migration (RTM) technique produces underground images using wave propagation. A discretization based on the Discontinuous Galerkin (DG) method unleashes a massively parallel elastodynamics simulation, an interesting feature for current and future architectures. We have designed a task-based version of this scheme in order to enable the use of manycore architectures. At this stage, we have demonstrated the efficiency of the approach on homogeneous and cache coherent Non Uniform Memory Access (ccNUMA) multicore platforms (up to 160 cores) and designed a prototype version of a distributed memory version that can exploit multiple instances of such architectures. This work has been conducted in the context of the DIP Inria-Total strategic action in collaboration with the MAGIQUE3D Project-Team and thanks to the long-term visit of George Bosilca funded by TOTAL. Geroge's expertise ensured an optimum usage of the PaRSEC runtime system onto which our task-based scheme has been ported.

This work was presented during a PRACE workshop [28] as well as during a TOTAL scientific event [29].

6.1.4. Resiliency in numerical simulations

For the solution of systems of linear equations, various recovery-restart strategies have been investigated in the framework of Krylov subspace methods to address the situations of core failures. The basic underlying idea is to recover fault entries of the iterate via interpolation from existing values available on neighbor cores. The resulting results are reported in the research report [41] currently submitted to an international journal. In that resilience framework, we have extended the recovey-restart ideas to the solution of linear eigenvalue problems. Contrary to the linear system case, not only the current iterate can be interpolated but also part of the subspace where candidate eigenpairs are searched.

This work is developed in the framework of Mawussi Zounon's PhD funded by the ANR RESCUE. These contributions have been presented at the international workshop Sparse Days [27] in Toulouse. More details and results can be found in report RR-8324 [41]. Notice that theses activities are also part of our contribution to the G8 ESC (Enabling Climate Simulation at extreme scale).

6.2. High performance solvers for large linear algebra problems

6.2.1. Parallel sparse direct solver on runtime systems

The ongoing hardware evolution exhibits an escalation in the number, as well as in the heterogeneity, of the computing resources. The pressure to maintain reasonable levels of performance and portability, forces the application developers to leave the traditional programming paradigms and explore alternative solutions. Algorithms, especially those in critical domains such as linear algebra, need to undergo invasive structural changes and be adapted to new programming paradigms to be in agreement with the latest hardware advances. PaStiX is a parallel sparse direct solver, based on a dynamic scheduler for modern hierarchical architectures. In this paper, we study the replacement of the highly specialized internal scheduler in PaStiX by two generic runtime frameworks: PaRSEC and StarPU. The tasks graph of the factorization step is made available to the two runtimes, providing them with the opportunity to optimize it in order to maximize the algorithm efficiency for a predefined execution environment. A comparative study of the performance of the PaStiX solver with the three schedulers on different execution contexts is performed. The analysis highlights the similarities from a performance point of view between the different execution supports. These results demonstrate that these generic DAG-based runtimes provide a uniform and portable programming interface across heterogeneous environments, and are, therefore, a sustainable solution for hybrid environments.

This work is developed in the framework of Xavier Lacoste's PhD funder by the ANR ANEMOS. These contributions have been presented at the international workshop Sparse Days [37] in Toulouse. More details and results can be found in report RR-8446 [46].

6.2.2. Hybrid parallel implementation of hybrid solvers

In the framework of the hybrid direct/iterative `MaPHyS` solver, we have designed and implemented an hybrid MPI-thread variant. More precisely, the implementation rely on the multi-threaded MKL library for all the dense linear algebra calculations and the multi-threaded version of `PaStiX`. Among the technical difficulties, one was to make sure that the two multi-threaded libraries do not interfere with each other. The resulting software prototype is currently experimented to study its new capability to get flexibility and trade-off between the parallel and numerical efficiency. Parallel experiments have been conducted on the Plafrim plateform as well as on a large scale machine located at the USA DOE NERSC, which has a large number of CPU cores per socket.

This work is developed in the framework of the PhD thesis of Stojce Nakov funded by TOTAL. These contributions have been presented at the NVIDIA GPU Technology Conference [25] in San Jose.

6.2.3. Designing LU-QR hybrid solvers for performance and stability

New hybrid LU-QR algorithms for solving dense linear systems of the form $Ax = b$ have been introduced. Throughout a matrix factorization, these algorithms dynamically alternate LU with local pivoting and QR elimination steps, based upon some robustness criterion. LU elimination steps can be very efficiently parallelized, and are twice as cheap in terms of flops, as QR steps. However, LU steps are not necessarily stable, while QR steps are always stable. The hybrid algorithms execute a QR step when a robustness criterion detects some risk for instability, and they execute an LU step otherwise. Ideally, the choice between LU and QR steps must have a small computational overhead and must provide a satisfactory level of stability with as few QR steps as possible. In this paper, we introduce several robustness criteria and we establish upper bounds on the growth factor of the norm of the updated matrix incurred by each of these criteria. In addition, we describe the implementation of the hybrid algorithms through an extension of the `PaRSEC` software to allow for dynamic choices during execution. Finally, we analyze both stability and performance results compared to state-of-the-art linear solvers on parallel distributed multicore platforms. These contributions have been presented at the international conference IPDPS [18] in Phoenix.

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

Last year we have worked primarily on developing an efficient fast multipole method for heterogeneous architecture. Some of the accomplishments for this year include:

1. Implementation of the FMM of multicore machines using `StarPU`. A new parallel scheduler was developed for this purpose. We implemented a state-of-the-art OpenMP version of the code for benchmarking purposes. It was found that `StarPU` significantly outperforms OpenMP. Figures show the traces of an execution of the FMM algorithm with our priority scheduler for the cube (volume) and ellipsoid (surface) with 20 million particles on a 4 deca-core Intel Xeon E7-4870 machine.

2. Implementation of the FMM of heterogeneous machines (CPU+GPU) using `StarPU`. The FMM was also used to demonstrate the flexibility of `StarPU` for handling different types of processors. In particular we demonstrated in that application that `StarPU` can automatically select the appropriate version of a computational kernel (CPU or GPU version) and run it on the appropriate processor in order to minimize the overall runtime. Significant speed-up were obtained on heterogeneous platforms compared to multicore only processors.

These contributions have been presented in minnisymposia at the SIAM conference on Comutational Sciences and Engineering [23], [33] in Boston and at NVIDIA GPU Technology Conference [24]. More details and results can be found in report RR-8277 [40], our paper is accepted for publication in the SIAM Journal on Scientific Computing [11].

Concerning dynamics dislocations (DD) kernels, an efficient formulation of the isotropic elastic far-field interactions between dislocations has been developed. This formulation is suitable for any polynomial interpolation based Fast Multipole Method (FMM) and is currently being implemented in OptiDis.

Meanwhile a much lighter and faster interpolation scheme based on a uniform grid (i.e. Lagrange interpolation) and the Fast Fourier Transform (FFT) was implemented into `ScalFMM`. This last feature was introduced in order to overcome the expensive cost of the Chebyshev FMM in the range of low interpolation orders (up to approx. 10). This should significantly improve the performances of the far-field computation in DD simulations where tensorial kernels are involved but only relatively low interpolation orders are required. This work is developed in the framework of Pierre Blanchard's PhD funded by ENS.

6.4. Efficient algorithmic for load balancing and code coupling in complex simulations

6.4.1. *Dynamic load balancing for massively parallel coupled codes*

As a preliminary step related to the dynamic load balancing of coupled codes, we focus on the problem of dynamic load balancing of a single parallel code, with variable number of processors. Indeed, if the workload varies drastically during the simulation, the load must be redistributed regularly among the processors. Dynamic load balancing is a well studied subject but most studies are limited to an initially fixed number of processors. Adjusting the number of processors at runtime allows to preserve the parallel code efficiency or to keep running the simulation when the current memory resources are exceeded. We call this problem, *MxN graph repartitioning*. We propose some methods based on graph repartitioning in order to rebalance the load while changing the number of processors. These methods are split in two main steps. Firstly, we study the migration phase and we build a "good" migration matrix minimizing several metrics like the migration volume or the number of exchanged messages. Secondly, we use graph partitioning heuristics to compute a new distribution optimizing the migration according to the previous step results. Besides, we propose a direct k-way partitioning algorithm that allows us to improve our biased partitioning. Finally, an experimental study validates our algorithms against state-of-the-art partitioning tools. Our algorithms are implemented in the `LBC2` library and have been integrated in the partitioning tools `Scotch` as a prototype.

This work is developed in the framework of Clément Vuchener's PhD, that will be defended on February 2014. These contributions have been presented at the international conference ParCo [22] in Munchen.

Regarding the problem of dynamic balancing of parallel coupled codes, we start to reuse results on *MxN graph repartitioning*. Given two coupled codes A and B, the key idea is to develop an algorithm of *two-graph co-partitioning*, that partitions two *coupled* graphs G_A and G_B in respectively N_A and N_B with classic objectives (*i.e.*, balancing computational load and minimizing communication cost for each code) and that minimizes the number of messages exchanged between codes in the coupling phase.

This work is developed in the framework of Maria Predari's PhD, that just started in october 2013.

6.4.2. *Graph partitioning for hybrid solvers*

Nested Dissection has been introduced by A. George and is a very popular heuristic for sparse matrix ordering before numerical factorization. It allows to maximize the number of parallel tasks, while reducing the fill-in and the operation count. The basic standard idea is to build a "small separator" S of the graph associated with the matrix in order to split the remaining vertices in two parts P_0 and P_1 of "almost equal size". The vertices of the separator S are ordered with the largest indices, and then the same method is applied recursively on the two sub-graphs induced by P_0 and P_1. At the end, if k levels of recursion are done, we get 2^k sets of independents vertices separated from each other by $2^k - 1$ separators.

However, if we examine precisely the complexity analysis for the estimation of asymptotic bounds for fill-in or operation count when using Nested Dissection ordering, we can notice that the size of the halo of the separated sub-graphs (set of external vertices belonging to an old separator and previously ordered) plays a crucial role in the asymptotic behavior achieved. In the perfect case, we need halo vertices to be balanced among parts.

Considering now hybrid methods mixing both direct and iterative solvers such as `HIPS`, `MaPHyS`, obtaining a domain decomposition leading to a good balancing of both the size of domain interiors and the Scalable numerical schemes for scientific applications size of interfaces is a key point for load balancing and efficiency in a parallel context. This leads to the same issue: balancing the halo vertices to get balanced interfaces.

For this purpose, we revisit the algorithm introduced by Lipton, Rose and Tarjan which performed the recursion of nested dissection in a different manner: at each level, we apply recursively the method to the sub-graphs But, for each sub-graph, we keep track of halo vertices. We have implemented that in the Scotch framework, and have studied its main algorithm to build a separator, called greedy graph growing.

This work is developed in the framework of Astrid Casadei's PhD. These contributions have been presented at the international workshop on Nested Dissection [32] in Waterloo.

6.5. Application Domains

6.5.1. *Dislocation dynamics simulations in material physics*

This year we have focused on the hybrid parallelization of the OptiDis code. As dislocations move in their grain, they expand, shrink, collide and annihilate, which means that we are facing a extremely dynamic n-body problem. Also, we have introduced an adaptive cache conscious data structure to manage the dislocation mesh. Moreover, two main kernels, plugged in our `ScalFMM` library, was built to handle the pairwise force interactions and the collisions between dislocations. Finally the code is written using hybrid parallelism based on OpenMP tasks inside on node and MPI to exchange data between nodes. The code can run on both shared and distributed memories. Future works will mainly focus on tuning the code and manage dynamically this tuning to adapt to different kind of simulations and architectures. On the physical side, we have introduced more *split node* cases to simulate irradiated materials. Now we are able to run simulations with tens of thousand of defaults in materials. Typically, our simulation box can hold lot of tiny dislocation loops such as those induced by radiation on materials, so we can observe how Frank-Read sources interact while they cross the field of loop defects.

This work is developed in the framework of Arnaud Etcheverry's PhD funded by the ANR OPTIDIS.

6.5.2. *Co-design for scalable numerical algorithms in scientific applications*

The study of the **thermo-acoustic stability of large combustion chambers** requires the solution of a nonlinear eigenvalue problem. The nonlinear problem is linearized using a fixed point iteration procedure. This leads to a sequence of linear eigenproblems which must be solved iteratively in order to obtain one nonlinear eigenpair. Therefore, efficient and robust parallel eigensolvers for the solution of linear problems have been investigated, and strategies to accelerate the solution of the sequence of linear eigenproblems have also been proposed. Among the numerical techniques that have been considered (Krylov-Schur, Implicitly Restarted Arnoldi, Subspace iteration with Chebyshev acceleration) the Jacobi-Davidson method was the best suited to be combined with techniques to recycle spectral information between the nonlinear iterations. The robustness of the parallel numerical techniques were illustrated on large problems with a few millions unknowns solved on a few tens of cores.

These results are part of the outcome of Pablo Salas PhD thesis that has been defended on November 15th.

The **Time-domain Boundary Element Method** (TD-BEM) has not been widely study but represent an interesting alternative to its frequency counterpart. Usually based on inefficient Sparse Matrix Vector-product (SpMV), we investigate other approaches in order to increase the sequential flop-rate. We have implement extremely efficient operator using intrinsic SIMD or even ASM64 instructions. We are using this novel approaches to parallelize both in shared and distributed memory and target execution on hundreds of clusters. All the implementations should be in high quality in the Software Engineering sense since the resulting library is going to be used by industrial applications.

This work is developed in the framework of Bérenger Bramas's PhD and contributes to the EADS-ASTRIUM, Inria, Conseil Régional initiative.

In a preliminary work, a **3D Cartesian SN solver** DOMINO has been designed and implemented using two nested levels of parallelism (multicore+SIMD) on shared memory computation nodes. DOMINO is written in C++, a multi-paradigm programming language that enables the use of powerful and generic parallel programming tools such as Intel TBB and Eigen. These two libraries allow us to combine multi-thread parallelism with vector operations in an efficient and yet portable way. As a result, DOMINO can exploit the full power of modern multi-core processors and is able to tackle very large simulations, that usually require large HPC clusters, using a single computing node. The very high Flops/Watt ratio of DOMINO makes it a very interesting building block for a future many-nodes nuclear simulation tool.

This work is developed in the framework of Salli Moustafa's PhD in collaboration with EDF. These contributions have been presented at the international conference on Supercomputing on Nuclear Applications [21] in Paris.

Concerning the numerical simulation of **the turbulence of plasma particules inside a tokamak**, two software tools, providing a post-mortem analysis, have been designed to manage the memory optimization of GYSELA [20]. The first one is a visualization tool. It plots the memory consumption of the code along an execution. This tool helps the developer to localize where happens the memory peak and to wonder how he can modify the code to decrease it. On the same graphic, the names of the allocated structures are labelled, which gives a significant hint on the modifications to apply. The second tool concerns the prediction of the peak memory. Given an input set of parameters, we can replay the allocations of the code in an offline mode. With this tool, we can deduce accurately the value of the memory peak and where it happens. Thank to this prediction we know which size of mesh is possible under a given architecture.

This work is carried on in the framework of Fabien Rozar's PhD in collaboration with CEA Cadarache.

In the first part of our research work concerning the parallel **aerodynamic code** FLUSEPA, an intermediate version based on the previous one has been developped. By using an hybrid OpenMP/MPI parallelism based on a domain decomposition, we achieved a faster version of the code and the temporal adaptive method used without bodies in relative motion has been tested successfully for real complex 3D-cases using up to 400 cores. Moreover, an asynchronous strategy for computing bodies in relative motion and mesh intersections has been developed and the test of this feature is currently in progress. The next step will be to design a new fully asynchronous code based on a task graph description to be executed on a modern runtine system like StarPU. This work is carried on in the framework of Jean-Marie Couteyen's PhD in collaboration with Astrium Les Mureaux.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

ASTRIUM Space Transportation research and development contract:

- Design of a parallel version of the FLUSEPA software (Jean-Marie Couteyen (PhD); Jean Roman).

CEA Cadarache (ITER) research and development contract:

- Peta and exaflop algorithms for turbulence simulations of fusion plasmas (Fabien Rozar (PhD); Guillaume Latu, Jean Roman).

EDF R & D - SINETICS research and development contract:

- Design of a massively parallel version of the SN method for neutronic simulations (Moustapha Salli (PhD); Pierre Ramet, Jean Roman).

TOTAL research and development contracts:

- Parallel hybrid solver for massively 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 at ONERA).

- Novel approaches to express fundamental algorithms using constructs that ensure their performance and scalability (G. Bosilca, visiting senior scientist).

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. PlaFRIM: an experimental parallel computing platform

PlaFRIM is an experimental platforme for research in modeling, simulations and high performance computing. This platform has been set up from 2009 under the leadership of Inria Bordeaux Sud-Ouest in collaboration with computer science and mathematics aboratories, respectively Labri and IMB with a strong support in the region Aquitaine.

It aggregates different kinds of computational resources for research and development purposes. The latest technologies in terms of processors, memories and architecture are added when they are available on the market. It is now more than 1,000 cores (excluding GPU and Xeon Phi) that are available for all research teams of Inria Bordeaux, Labri and IMB. This computer is in particular used by all the engineers who work in HiePACS and are advised by F. Rue from the SED.

The PlaFRIM platform initiative is coordinated by O. Coulaud and an application for its upgrade has been accepted.

8.1.2. Innovative simulation methods for large scale numeric prototypes on emerging architectures computers

Participants: Emmanuel Agullo, Olivier Coulaud, Aurélien Esnard, Mathieu Faverge, Luc Giraud, Abdou Guermouche, Pierre Ramet, Jean Roman.

Grant: Regional council
Dates: 2013 – 2015
Partners: EPIs REALOPT, RUNTIME from Inria Bordeaux Sud-Ouest, CEA-CESTA and l'Institut pluridisciplinaire de recherche sur l'environnement et les matériaux (IPREM) .
Overview: Numerical simulation is now integrated into all the design levels and the scientific studies for both academic and industrial contexts. Given the increasing size and sophistication of the simulations carried out, the use of parallel computing is inescapable. The complexity of such achievements requires collaboration of multidisciplinary teams capable of mastering all the necessary scientific skills for each component constituting the chain of expertise. In this project we consider each of these elements as well as efficient methods for parallel codes coupling. All these works is intended to contribute to the design large scale parallel multi-physics simulations. In addition to this research human activities the regional council also support some innovative computing equipment that will be embedded in the PlaFRIM experimental plateform, project led by Olivier Coulaud.

8.2. National Initiatives

8.2.1. *Inria Project Lab*

8.2.1.1. C2S@Exa - Computer and Computational Sciences at Exascale

Since January 2013, the team is participating to the C2S@EXA Inria Project Lab (IPL). This national initiative aims at the development of numerical modeling methodologies that fully exploit the processing capabilities of modern massively parallel architectures in the context of a number of selected applications related to important scientific and technological challenges for the quality and the security of life in our society. At the current state of the art in technologies and methodologies, a multidisciplinary approach is required to overcome the challenges raised by the development of highly scalable numerical simulation software that can exploit computing platforms offering several hundreds of thousands of cores. Hence, the main objective of C2S@EXA is the establishment of a continuum of expertise in the computer science and numerical mathematics domains, by gathering researchers from Inria project-teams whose research and development activities are tightly linked to high performance computing issues in these domains. More precisely, this collaborative effort involves computer scientists that are experts of programming models, environments and tools for harnessing massively parallel systems, algorithmists that propose algorithms and contribute to generic libraries and core solvers in order to take benefit from all the parallelism levels with the main goal of optimal scaling on very large numbers of computing entities and, numerical mathematicians that are studying numerical schemes and scalable solvers for systems of partial differential equations in view of the simulation of very large-scale problems.

8.2.2. *ANR*

8.2.2.1. SOLHAR: SOLvers for Heterogeneous Architectures over Runtime systems

Participants: Emmanuel Agullo, Mathieu Faverge, Andra Hugo, Abdou Guermouche, Xavier Lacoste, Pierre Ramet, Jean Roman, Guillaume Sylvand.

Grant: ANR-MONU
Dates: 2013 – 2017
Partners: Inria (REALOPT, RUNTIME Bordeaux Sud-Ouest et ROMA Rhone-Alpes), IRIT/INPT, CEA-CESTA et EADS-IW.
Overview:

During the last five years, the interest of the scientific computing community towards accelerating devices has been rapidly growing. The reason for this interest lies in the massive computational power delivered by these devices. Several software libraries for dense linear algebra have been produced; the related algorithms are extremely rich in computation and exhibit a very regular pattern of access to data which makes them extremely good candidates for GPU execution. On the contrary, methods for the direct solution of sparse linear systems have irregular, indirect memory access patterns that adversely interact with typical GPU throughput optimizations.

This project aims at studying and designing algorithms and parallel programming models for implementing direct methods for the solution of sparse linear systems on emerging computer equipped with accelerators. The ultimate aim of this project is to achieve the implementation of a software package providing a solver based on direct methods for sparse linear systems of equations. To date, the approaches proposed to achieve this objective are mostly based on a simple offloading of some computational tasks to the accelerators and rely on fine hand-tuning of the code and accurate performance modeling to achieve efficiency. This project proposes an innovative approach which relies on the efficiency and portability of runtime systems. The development of a production-quality, sparse direct solver requires a considerable research effort along three distinct axis:

- linear algebra: algorithms have to be adapted or redesigned in order to exhibit properties that make their implementation and execution on heterogeneous computing platforms efficient and reliable. This may require the development of novel methods for defining data access patterns that are more suitable for the dynamic scheduling of computational tasks on processing units with considerably different capabilities as well as techniques for guaranteeing a reliable and robust behavior and accurate solutions. In addition, it will be necessary to develop novel and efficient accelerator

implementations of the specific dense linear algebra kernels that are used within sparse, direct solvers;

- runtime systems: tools such as the `StarPU` runtime system proved to be extremely efficient and robust for the implementation of dense linear algebra algorithms. Sparse linear algebra algorithms, however, are commonly characterized by complicated data access patterns, computational tasks with extremely variable granularity and complex dependencies. Therefore, a substantial research effort is necessary to design and implement features as well as interfaces to comply with the needs formalized by the research activity on direct methods;

- scheduling: executing a heterogeneous workload with complex dependencies on a heterogeneous architecture is a very challenging problem that demands the development of effective scheduling algorithms. These will be confronted with possibly limited views of dependencies among tasks and multiple, and potentially conflicting objectives, such as minimizing the makespan, maximizing the locality of data or, where it applies, minimizing the memory consumption.

Given the wide availability of computing platforms equipped with accelerators and the numerical robustness of direct solution methods for sparse linear systems, it is reasonable to expect that the outcome of this project will have a considerable impact on both academic and industrial scientific computing. This project will moreover provide a substantial contribution to the computational science and high-performance computing communities, as it will deliver an unprecedented example of a complex numerical code whose parallelization completely relies on runtime scheduling systems and which is, therefore, extremely portable, maintainable and evolvable towards future computing architectures.

8.2.2.2. SONGS: Simulation Of Next Generation Systems
Participant: Abdou Guermouche.

Grant: ANR 11 INFRA 13
Dates: 2011 – 2015
Partners: Inria (Bordeaux Sud-Ouest, Nancy - Grand Est, Rhone-Alpes, Sophia Antipolis - Méditerranée), I3S, LSIIT
Overview:

The last decade has brought tremendous changes to the characteristics of large scale distributed computing platforms. Large grids processing terabytes of information a day and the peer-to-peer technology have become common even though understanding how to efficiently such platforms still raises many challenges. As demonstrated by the USS SimGrid project funded by the ANR in 2008, simulation has proved to be a very effective approach for studying such platforms. Although even more challenging, we think the issues raised by petaflop/exaflop computers and emerging cloud infrastructures can be addressed using similar simulation methodology.

The goal of the SONGS project is to extend the applicability of the SimGrid simulation framework from Grids and Peer-to-Peer systems to Clouds and High Performance Computation systems. Each type of large-scale computing system will be addressed through a set of use cases and lead by researchers recognized as experts in this area.

Any sound study of such systems through simulations relies on the following pillars of simulation methodology: Efficient simulation kernel; Sound and validated models; Simulation analysis tools; Campaign simulation management.

8.2.2.3. ANEMOS: Advanced Numeric for ELMs : Modeling and Optimized Schemes
Participants: Xavier Lacoste, Guillaume Latu, Pierre Ramet.

Grant: ANR-MN
Dates: 2012 – 2016
Partners: Univ. Nice, CEA/IRFM, CNRS/MDS.

Overview: The main goal of the project is to make a significant progress in understanding of largely unknown at present physics of active control methods of plasma edge MHD instabilities Edge Localized Modes (ELMs) wich represent particular danger with respect to heat and particle loads for Plasma Facing Components (PFC) in ITER. Project is focused in particular on the numerical modelling study of such ELM control methods as Resonant Magnetic Perturbations (RMPs) and pellet ELM pacing both foreseen in ITER. The goals of the project are to improve understanding of the related physics and propose possible new strategies to improve effectiveness of ELM control techniques. The tool for the non-linear MHD modeling is the JOREK code which was essentially developed within previous ANR ASTER. JOREK will be largerly developed within the present project to include corresponding new physical models in conjunction with new developments in mathematics and computer science strategy. The present project will put the non-linear MHD modeling of ELMs and ELM control on the solid ground theoretically, computationally, and applications-wise in order to progress in urgently needed solutions for ITER.

Regarding our contributions, the JOREK code is mainly composed of numerical computations on 3D data. The toroidal dimension of the tokamak is treated in Fourier space, while the poloidal plane is decomposed in Bezier patches. The numerical scheme used involves a direct solver on a large sparse matrix as a main computation of one time step. Two main costs are clearly identified: the assembly of the sparse matrix, and the direct factorization and solve of the system that includes communications between all processors. The efficient parallelization of JOREK is one of our main goals, to do so we will reconsider: data distribution, computation distribution or GMRES implementation. The quality of the sparse solver is also crucial, both in term of performance and accuracy. In the current release of JOREK, the memory scaling is not satisfactory to solve problems listed above , since at present as one increases the number of processes for a given problem size, the memory footprint on each process does not reduce as much as one can expect. In order to access finer meshes on available supercomputers, memory savings have to be done in the whole code. Another key point for improving parallelization is to carefully profile the application to understand the regions of the code that do not scale well. Depending on the timings obtained, strategies to diminish communication overheads will be evaluated and schemes that improve load balancing will be initiated. JOREK uses PaStiX sparse matrix library for matrix inversion. However, large number of toroidal harmonics and particular thin structures to resolve for realistic plasma parameters and ITER machine size still require more aggressive optimisation in numeric dealing with numerical stability, adaptive meshes etc. However many possible applications of JOREK code we proposed here which represent urgent ITER relevant issues related to ELM control by RMPs and pellets remain to be solved.

8.2.2.4. OPTIDIS: OPTImisation d'un code de dynamique des DISlocations

Participants: Olivier Coulaud, Aurélien Esnard, Arnaud Etcheverry, Luc Giraud.

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 behavior of a single dislocation segment is perfectly understood since 1960 and analytical formulations are available in the literature. However, to understand the behavior 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 computation. The objective of the project in to develop and validate the algorithms in order to optimize the numerical and performance efficiency 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 plastic deformation after irradiation. These kinds of studies

where "local mechanisms" are correlated with macroscopic behavior is a key issue for nuclear industry in order to understand material aging under irradiation, and hence predict power plant secured service life. To carry out such studies, massive numerical optimizations of NUMODIS are required. They involve complex algorithms lying on advanced computational science methods. The project OPTIDIS will develop through joint collaborative studies involving researchers specialized in dynamics dislocations and in numerical methods. This project is divided in 8 tasks over 4 years. Two PhD thesis will be directly funded by the project. One will be dedicated to numerical development, validation of complex algorithms and comparison with the performance of existing dynamics dislocation codes. The objective of the second is to carry out large scale simulations to validate the performance of the numerical developments made in OPTIDIS. In both cases, these simulations will be compared with experimental data obtained by experimentalists.

8.2.2.5. 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 ROMA (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.

8.2.2.6. BOOST: Building the future Of numerical methOdS for iTer

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

Grant: ANR-Blanc (applied math theme)
Dates: 2010 – 2014
Partners: Institut de Mathématiques de Toulouse (leader); Laboratoire d'Analyse, Topologie, Probabilités in Marseilles; Institut de Recherche sur la Fusion Magnétique, CEAr/IRFM and 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.

8.3. European Initiatives

8.3.1. FP7 Projects

8.3.1.1. EXA2CT

Type: COOPERATION

Defi: Exascale computation

Instrument: Specific Targeted Research Project

Duration: September 2013 - August 2016

See also: https://projects.imec.be/exa2ct/.

Coordinator: Wilfried Verachtert, IMEC (Interuniversitair Micro-Electronica Centrum), Belgium

Partner: Universiteit Antwerpen, Belgium; Università della Svizzera italiana, Switzerland; Inria (ALPINES, HIEPACS, SAGE teams); Université de Versailles Saint-Quentin-en-Yvelines, France; T-Systems, Germany; Fraunhofer-Gesellschaft, Germany; Intel, France; NAG, UK.

Inria contact: Luc Giraud

Abstract: Numerical simulation is a crucial part of science and industry in Europe. The advancement of simulation as a discipline relies on increasingly compute intensive models that require more computational resources to run. This is the driver for the evolution to exascale. Due to limits in the increase in single processor performance, exascale machines will rely on massive parallelism on and off chip, with a complex hierarchy of resources. The large number of components and the machine complexity introduce severe problems for reliability and programmability. The former of these will require novel fault-aware algorithms and support software. In addition, the scale of the numerical models exacerbates the difficulties by making the use of more complex simulation algorithms necessary, for numerical stability reasons. A key example of this is increased relience on linear solvers. Such solvers require global communication, which impacts scalability, and are often used with preconditioners, increasing complexity again. Unless there is a major rethink of the design of solver algorithms, components and software structure, a large class of important numerical simulations will not scale beyond petascale. This in turn will hold back the development of European science and industry that will not reap the benefits from exascale.

The EXA2CT project brings together experts at the cutting edge of the development of solvers, related algorithmic techniques, and HPC software architects for programming models and communication. It will take a revolutionary approach to exascale solvers and programming models, rather than

the incremental approach of other projects. We will produce modular open source proto-applications that demonstrate the algorithms and programming techniques developed in the project, to help boot-strap the creation of genuine exascale codes.

Inria is involved in that project through the IPL C2S@Exa initiative.

8.4. International Initiatives

8.4.1. Inria Associate Teams

8.4.1.1. 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://icl.cs.utk.edu/projectsdev/morse.

The goal of MORSE (Matrices Over Runtime Systems at Exascale) 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.

8.4.1.2. FASTLA

Title: Fast and Scalable Hierarchical Algorithms for Computational Linear Algebra

Inria principal investigator: Olivier Coulaud

International Partners (Institution - Laboratory - Researcher):

Lawrence Berkeley National Laboratory (United States) - Scientific Computing Group - Esmond Ng

Stanford University (United States) - Institute for Computational and Mathematical Engineering - Eric Darve

Duration: 2012 - 2014

See also: http://people.bordeaux.inria.fr/coulaud/projets/FastLA_Website

In this project, we propose to study fast and scalable hierarchical numerical kernels and their implementations on heterogeneous manycore platforms for two major computational kernels in intensive challenging applications. Namely, fast multipole methods (FMM) and sparse hybrid linear solvers, that appear in many intensive numerical simulations in computational sciences. Regarding the FMM we plan to study novel generic formulations based on H-matrices techniques, that will be eventually validated in the field of material physics: the dislocation dynamics. For the hybrid solvers, new parallel preconditioning approaches will be designed and the use of H-matrices techniques will be first investigated in the framework of fast and monitored approximations on central components. Finally, the innovative algorithmic design will be essentially focused on heterogeneous manycore platforms. The partners, Inria HIEPACS, Lawrence Berkeley Nat. Lab and Stanford University, have strong, complementary and recognized experiences and backgrounds in these fields.

8.4.2. Participation In other International Programs

8.4.2.1. HOSCAR

We are involved in the Inria-CNPq HOSCAR project led by Stéphane Lanteri.

The general objective of the project is to setup a multidisciplinary Brazil-France collaborative effort for taking full benefits of future high-performance massively parallel architectures. The targets are the very large-scale datasets and numerical simulations relevant to a selected set of applications in natural sciences: (i) resource prospection, (ii) reservoir simulation, (iii) ecological modeling, (iv) astronomy data management, and (v) simulation data management. The project involves computer scientists and numerical mathematicians divided in 3 fundamental research groups: (i) numerical schemes for PDE models (Group 1), (ii) scientific data management (Group 2), and (iii) high-performance software systems (Group 3).

We organized the 2013 annual meeting in Bordeaux on September 2-6, 2013 and are contributing to the Group 3 activities.

8.4.2.2. G8-ECS

Title: Enabling Climate Simulations at Extreme Scale

Inria principal investigator: Luc Giraud

International Partners (Institution - Researcher):

Univ. Illinois at Urbanna Champaign & Argonne National Lab. - Franck Cappello,

Univ. Tennessee at Knoxville - George Bosilca,

German Research School for Simulation Sciences - Felix Wolf,

Univ. Victoria - Andrew Weaver,

Titech - Satoshi Matsuoka,

Univ. Tsukuba - Mitsuhisa Sato,

NCAR - Rich Loft,

Barcelona Supercomputing Center - Jesus Labarta.

Duration: 2011 - 2014

See also: https://wiki.engr.illinois.edu/display/G8/G8+ESC++–+Enabling+Climate+Simulations+at+Extreme+Scale

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. Scientific Animation

Emmanuel Agullo has been a member of the scientific committee of the international conferences IPDPS'13.

Olivier Coulaud is member of the C3I GENCI committee and of the scientific board of the regional computing mesocentre. Moreover, he is the leader of the Inria PlaFRIM computing plateform. He is member of the scientific committee of CCGRID2014.

Mathieu Faverge has been member of the technical program committee of the international conference HiPC'13 and reviewed for Parallel Computing journal.

Pierre Ramet is member of the GENCI scientific committee (Mathematics and Computer Sciences). He is also in the decision board of the "MCIA" project (*Mésocentre Aquitain : un environnement Mutualisé de Calcul Intensif en Aquitaine*).

Luc Giraud has been member of the scientific committee of the international conferences Preconditioning'13, IPDPS'13, HiPC'13, Pareng'13, PDCN'13 and PDSEC'13. He is also member of the editorial board of the SIAM Journal on Matrix Analysis and Applications (SIMAX), expert for the PRACE program and for the Italian VQR 2004-2010 programme. He is a member of DAS-SCI (Domaine d'Activité technologiques Stratégiques - Systémes complexes, Conception, architecture et Intégration) of Aerospace Valley (pôle de compétitivité).

Abdou Guermouche has been member of the scientific committee of the international conferences HiPC'13, ICPP'13 and IPDPS'13.

Jean Roman is a member of the Research Direction of Inria; he is the Deputy Scientific Director of the Inria research domain entitled *Applied Mathematics, Computation and Simulation* and is in charge at the national level of the Inria activities concerning High Performance Computing. 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. As representative of Inria, he is member of the Steering Committee of EADS Corporate Foundation (for the research chair entitled *Mathematics of Complex Systems* in Bangalore-India), of the board of ETP4HPC (*European Technology Platform for High Performance Computing*), of the French Information Group for PRACE, of the Technical Group of GENCI and of the Scientific Advisory Board of the Maison de la Simulation. Finally, he has been member of the scientific committee of the international conference EuroMicro PDP'13 (IEEE).

Finally, the HIEPACS members have contributed to the reviewing process of several international journals (ACM Trans. on Mathematical Software, IEEE Trans. on Parallel and Distributed Systems, Journal of Engineering Mathematics, Parallel Computing, SIAM J. Matrix Analysis and Applications , SIAM J. Scientific Comp., ...), to the reviewing process of international conferences (CCGRID 2014, IEEE IPDPS 2014, IEEE PDP 2014, ...), and have acted as experts for the research agency ANR-MN.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

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

Undergraduate level/Licence

1. A. Esnard: Operating system programming, 36h, University Bordeaux I; Using network, 23h, University Bordeaux I.
   He is also in charge of the computer science certificate for Internet (C2i) at the University Bordeaux I.

2. P. Ramet: System programming 24h, Databases 32h, Objet programming 48h, cryptography 32h at Bordeaux University.
   He is also in charge of the "licences professionelles" since 2007 at Bordeaux University.

Post graduate level/Master

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

2. E. Agullo: Operating sysems, 24h, University Bordeaux I; Dense linear algebra kernels, 8h, ENSEIRB-MatMeca; Numerical Algorithms, 30h; ENSEIRB-MatMeca, Talence.

3. A. Esnard: Network management, 27h, University Bordeaux I; Network security, 27h, University Bordeaux I; Programming distributed applications, 35h, ENSEIRB-MatMeca, Talence.

4. M. Faverge: System Programming, 74h; Programming Environment, 26h; Load Balancing and Scheduling, 19h; Numerical Algorithmic, 30h; C Projects, 20h; ENSEIRB-MatMeca, Talence.
   He is also in charge of the second year of Embedded Electronic Systems option at ENSEIRB-MatMeca, Talence.

5. P. Ramet: Scheduling, 8h; Numerical Algorithmic, 30h; ENSEIRB-MatMeca, Talence.
   He also give classes on advanced databases, 30h, Ho Chi Minh City, Vietnam.

6. 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.

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

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

9. 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.

9.2.2. Supervision

Defended HDR

1. HdR: Olivier Coulaud, *Contributions algorithmiques pour les simulations complexes en physique des matériaux*, Université Bordeaux 1, defended on 29 Nov. 2013.

Defended PhD thesis

1. Pablo Salas Medina, *Numerical and physical aspects of thermoacoustic instabilities in annular combustion chambers* Université Bordeaux I, defended on 15 Nov. 2013, advisors: L. Giraud and X. Vasseur (CERFACS).

2. Rached Abdelkhalek, *Accélération matérielle pour l'imagerie sismique:modélisation, migration et interprétation*, Université Bordeaux I, defended on 20 Dec. 2013 advisors: O. Coulaud, G. Latu (CEA Cadarache) and J. Roman.

PhD in progress :

1. Pierre Blanchard, *Fast and accurate methods for dislocation dynamics*, starting Oct. 2013,advisors: O. Coulaud and E. Darve.

2. Bérenger Bramas, *Optimization of time domain BEM solvers*, starting Jan 2013, advisors: O. Coulaud and G. Sylvand.

3. Astrid Casadei, *Scalabilité et robustesse numérique des solveurs hybrides pour machines massivement parallèles*, starting Oct. 2011, advisors: F. Pellegrini and P. Ramet.

4. Jean-Marie Couteyen, *Parallélisation et passage à l'échelle du code FLUSEPA*, starting Feb 2013, advisors : P. Brenner (Astrium Space Transportation) and J. Roman.

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

6. Arnaud Etcheverry, *Toward large scale dynamic dislocation simulation on petaflop computers*, starting Oct. 2011, advisor: O. Coulaud.

7. Xavier Lacoste, *Scheduling and memory optimizations for sparse direct solver on multi-core/multigpu cluster systems*, starting Jan. 2012, advisors: F. Pellegrini and P. Ramet.

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

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

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

11. Maria Predari, *Dynamic Load Balancing for Massively Parallel Coupled Codes*, starting Oct. 2013, advisors: A. Esnard and J. Roman.

12. Fabien Rozar, *Peta and exaflop algorithms for turbulence simulations of fusion plasmas*, starting Nov. 2012, advisors: Guillaume Latu (CEA Cadarache) and Jean Roman.

13. Moustapha Salli, *Design of a massively parallel version of the SN method for neutronic simulations*, starting Oct. 2012, advisors: Laurent Plagne (EDF), Pierre Ramet and Jean Roman.

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

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

9.2.3. Juries

PhD thesis defences

1. M. Malandain, *Simulation massivement parallèle des écoulements turbulents à faible nombre de Mach*, Institut National des Sciences Appliquées de Rouen, spécialité mécanique des fluides numériques, January 15, Referee: L. Giraud.

2. S. Fourestier, *Redistribution dynamique parallèle efficace de la charge pour les problèmes numériques de très grande taille*, Université Bordeaux I, June 20, Member: A. Esnard.

3. S. Henry, *Modèles de programmation et supports exécutifs pour architectures hétérogènes*, Université Bordeaux I, spécialité Informatique, Nov. 14, Member: L. Giraud.

HdR defences

1. C. Tadonki, High Performance Computing and Combination of Machines and Methods and Programming, Université Paris-Sud, Orsay, spécialité Informatique, May 16, Referee: L. Giraud.

9.3. Popularization

In the context of HPC-PME initiative, we started a collaboration with ALGO'TECH INFORMATIQUE and we have organised one of the first PhD-consultant action implemented by Xavier Lacoste led by Pierre Ramet. ALGO'TECH is one of the most innovative SMEs (small and medium sized enterprises) in the field of cabling embedded systems, and more broadly, automatic devices. The main target of the project is to validate the possibility to use the sparse linear solvers of our team in the area of electromagnetic simulation tools developed by ALGO'TECH.

The HIEPACS members have organized the PATC training session on Parallel Linear algebra at "Maison de la simulation" in Saclay from March 26th to March 28th. Mathieu Faverge also gave a full day lecture on dense linear algebra libraries at the PRACE Summer School at CINECA.

10. Bibliography

Major publications by the team in recent years

[1] 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

[2] 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, pp. 177-188

[3] J. DONGARRA, M. FAVERGE, T. HÉRAULT, M. JACQUELIN, J. LANGOU, Y. ROBERT. *Hierarchical QR factorization algorithms for multi-core clusters*, in "Parallel Computing", 2013, vol. 39, n⁰ 4-5, pp. 212-232 [*DOI :* 10.1016/J.PARCO.2013.01.003], http://hal.inria.fr/hal-00809770

[4] 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, pp. 1858–1878 [*DOI :* 10.1137/080741847], http://hal.inria.fr/inria-00542426/en

[5] 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, pp. 276-294, http://hal.inria.fr/inria-00542450/en

[6] P. HÉNON, P. RAMET, J. ROMAN. *On finding approximate supernodes for an efficient ILU(k) factorization*, in "Parallel Computing", 2008, vol. 34, pp. 345–362, http://hal.inria.fr/inria-00346018

[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, pp. 2654–2668 [*DOI :* 10.1021/CT100280X], http://hal.inria.fr/inria-00488048/en

[8] C. VUCHENER, A. ESNARD. *Dynamic Load-Balancing with Variable Number of Processors based on Graph Repartitioning*, in "HIPC 2012", Pune, India, 2012, pp. 1-9, http://hal.inria.fr/hal-00722731

Publications of the year

Doctoral Dissertations and Habilitation Theses

[9] R. ABDELKHALEK. , *Acceleration materielle pour l'imagerie sismique : modelisation, migration et interpretation*, Université Sciences et Technologies - Bordeaux I, December 2013, http://hal.inria.fr/tel-00936989

[10] P. SALAS. , *Aspects numeriques et physiques des instabilites thermoacoustiques dans les chambres de combustion annulaires*, Université Sciences et Technologies - Bordeaux I, November 2013, http://hal.inria.fr/tel-00937020

Articles in International Peer-Reviewed Journals

[11] E. AGULLO, B. BRAMAS, O. COULAUD, E. DARVE, M. MESSNER, T. TAKAHASHI. *Task-Based FMM for Multicore Architectures*, in "SIAM Journal on Scientific Computing", 2013, http://hal.inria.fr/hal-00911856

[12] G. BOSILCA, A. BOUTEILLER, A. DANALIS, M. FAVERGE, T. HÉRAULT, J. DONGARRA. *PaRSEC: A programming paradigm exploiting heterogeneity for enhancing scalability*, in "Computing in Science and Engineering", 2013, vol. 99, 1 p. [*DOI :* 10.1109/MCSE.2013.98], http://hal.inria.fr/hal-00930217

[13] M. CHANAUD, L. GIRAUD, D. GOUDIN, J.-J. PESQUÉ, J. ROMAN. *A Parallel Full Geometric Multigrid Solver for Time Harmonic Maxwell Problems*, in "SIAM Journal on Scientific Computing", December 2013, http://hal.inria.fr/hal-00933526

[14] J. DONGARRA, M. FAVERGE, T. HÉRAULT, M. JACQUELIN, J. LANGOU, Y. ROBERT. *Hierarchical QR factorization algorithms for multi-core clusters*, in "Parallel Computing", 2013, vol. 39, n⁰ 4-5, pp. 212-232 [*DOI :* 10.1016/J.PARCO.2013.01.003], http://hal.inria.fr/hal-00809770

[15] J. J. DONGARRA, M. FAVERGE, H. LTAIEF, P. LUSZCZEK. *Achieving Numerical Accuracy and High Performance using Recursive Tile LU Factorization*, in "Concurrency and Computation: Practice and Experience", September 2013 [*DOI :* 10.1002/CPE.3110], http://hal.inria.fr/hal-00865472

Invited Conferences

[16] L. GIRAUD. *Algebraic preconditioners for parallel hybrid solvers*, in "High Performance Computing in Science and Engineering", Ostrava, Czech Republic, May 2013, http://hal.inria.fr/hal-00933517

International Conferences with Proceedings

[17] G. AUPY, M. FAVERGE, Y. ROBERT, J. KURZAK, P. LUSZCZEK, J. DONGARRA. *Implementing a systolic algorithm for QR factorization on multicore clusters with PaRSEC*, in "PROPER 2013 - 6th Workshop on Productivity and Performance", Aachen, Germany, August 2013, http://hal.inria.fr/hal-00844492

[18] M. FAVERGE, J. HERRMANN, J. LANGOU, B. LOWERY, Y. ROBERT, J. DONGARRA. *Designing LU-QR hybrid solvers for performance and stability*, in "IEEE International Parallel & Distributed Processing Symposium", Phoenix, United States, December 2013, http://hal.inria.fr/hal-00930238

[19] A.-E. HUGO, A. GUERMOUCHE, R. NAMYST, P.-A. WACRENIER. *Composing multiple StarPU applications over heterogeneous machines: a supervised approach*, in "Third International Workshop on Accelerators and Hybrid Exascale Systems", Boston, United States, May 2013, http://hal.inria.fr/hal-00824514

[20] G. LATU, J. ROMAN, F. ROZAR. *Achieving Memory Scalability in the Gysela Code to Fit Exascale Constraints*, in "10th International Conference on Parallel Processing and Applied Mathematics", Warsaw, Poland, Lecture Note in Computer Science, Springer, September 2013, To appear, http://hal.inria.fr/hal-00935519

[21] S. MOUSTAFA, I. DUTKA MALEN, L. PLAGNE, A. PONÇOT, P. RAMET. *Shared Memory Parallelism for 3D Cartesian Discrete Ordinates Solver*, in "Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2013", Paris, France, October 2013, http://hal.inria.fr/hal-00924989

[22] C. VUCHENER, A. ESNARD. *Graph Repartitioning with both Dynamic Load and Dynamic Processor Allocation*, in "International Conference on Parallel Computing - ParCo2013", München, Germany, Advances of Parallel Computing, 2013, http://hal.inria.fr/hal-00857881

Conferences without Proceedings

[23] E. AGULLO, B. BRAMAS, O. COULAUD, E. DARVE, M. MESSNER, T. TAKAHASHI. *Pipelining the Fast Multipole Method over a Runtime System*, in "SIAM Conference on Computational Science and Engineering - CSE 2013", Boston, United States, February 2013, http://hal.inria.fr/hal-00797403

[24] E. AGULLO, B. BRAMAS, O. COULAUD, E. DARVE, M. MESSNER, T. TAKAHASHI. *Task-based Parallelization of the Fast Multipole Method on NVIDIA GPUs and Multicore Processors*, in "GPU Technology Conference", San Jose, California, United States, NVIDIA, 2013, http://hal.inria.fr/hal-00879291

[25] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, S. NAKOV, J. ROMAN. *Pipelining the CG Solver Over a Runtime System*, in "GPU Technology Conference", San Jose, United States, NVIIDA, March 2013, http://hal.inria.fr/hal-00934948

[26] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, J. ROMAN. *Towards resilient parallel linear Krylov solvers with recover-restart strategies*, in "European Numerical Mathematics and Advanced Applications", Lausanne, Switzerland, August 2013, http://hal.inria.fr/hal-00933640

[27] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, J. ROMAN, M. ZOUNON. *Towards resilient parallel linear Krylov solvers: recover-restart strategies*, in "Sparse days 2013", Toulouse, France, CERFACS, 2013, http://hal.inria.fr/hal-00935685

[28] L. BOILLOT, E. AGULLO, H. BARUCQ, G. BOSILCA, H. CALANDRA, J. DIAZ. *Optimized propagators for elastic waves with anisotropy Part 1: Applied Mathematics Part 2: High-Performance Computing*, in "XSEDE – International Summer School on HPC Challenges in Computational Sciences (PRACE workshop)", New York, United States, June 2013, http://hal.inria.fr/hal-00868169

[29] L. BOILLOT, E. AGULLO, G. BOSILCA, H. CALANDRA. *3D Geophysics over a runtime system on a ccNUMA machine*, in "MATHIAS - TOTAL Symposium on Mathematics", Paris, France, October 2013, http://hal.inria.fr/hal-00877316

[30] L. BOILLOT, E. AGULLO, G. BOSILCA, H. CALANDRA. *Combining recent HPC techniques for 3D geophysics acceleration*, in "2nd ECCOMAS Young Investigators Conference (YIC 2013)", Bordeaux, France, September 2013, http://hal.inria.fr/hal-00855878

[31] G. BOSILCA, A. BOUTEILLER, M. FAVERGE, T. HÉRAULT. *Linear Algebra Libraries with DAG Runtimes on GPUs*, in "SIAM CSE 2013", Boston, United States, February 2013, http://hal.inria.fr/hal-00934573

[32] A. CASADEI, L. GIRAUD, P. RAMET, J. ROMAN. *Towards Domain Decomposition with Balanced Halo*, in "Workshop Celebrating 40 Years of Nested Dissection", Waterloo, Canada, July 2013, http://hal.inria.fr/hal-00924977

[33] E. DARVE, M. MESSNER, M. SCHANZ, O. COULAUD. *Optimizing the Black-box FMM for Smooth and Oscillatory Kernels*, in "SIAM Conference on Computational Science and Engineering", Boston, United States, SIAM, February 2013, http://hal.inria.fr/hal-00799885

[34] Y. DUDOUIT, L. GIRAUD, F. MILLOT, S. PERNET. *Parallel local time-stepping for elastodynamic equations*, in "Conference on Mathematical and Computational Issues in the Geosciences", Padua, Italy, SIAM, June 2013, http://hal.inria.fr/hal-00933533

[35] L. GIRAUD, F. CAPPELLO. *Resilience at extreme scale : system level, algorithmic level or both ?*, in "SIAM Conference on Computational Science and Engineering - CSE 2013", Boston, United States, SIAM, March 2013, http://hal.inria.fr/hal-00799309

[36] L. GIRAUD, P. SALAS, X. VASSEUR. *Eigensolvers for thermoacoustics instabilities in combustion chambers*, in "Conference on Computational Sciences and Engineering", Boston, United States, SIAM, March 2013, http://hal.inria.fr/hal-00933498

[37] X. LACOSTE. *Work stealing and granularity optimizations for a sparse solver on manycores*, in "Sparse days 2013", Toulouse, France, June 2013, http://hal.inria.fr/hal-00932823

[38] P. RAMET. *From hybrid architectures to hybrid solvers*, in "Workshop Celebrating 40 Years of Nested Dissection", Waterloo, Canada, July 2013, http://hal.inria.fr/hal-00924979

Scientific Books (or Scientific Book chapters)

[39] O. COULAUD, L. GIRAUD, P. RAMET, X. VASSEUR. *Deflation and augmentation techniques in Krylov linear solvers*, in "Developments in Parallel - Distributed - Grid and Cloud Computing for Engineering", B. TOPPING, P. IVANYI (editors), Saxe-Coburg Publications, March 2013, http://hal.inria.fr/hal-00860339

Research Reports

[40] E. AGULLO, B. BRAMAS, O. COULAUD, E. DARVE, M. MESSNER, T. TAKAHASHI. , *Task-based FMM for multicore architectures*, Inria, March 2013, n⁰ RR-8277, 33 p. , Preliminary version of a paper to appear in SIAM SISC, http://hal.inria.fr/hal-00807368

[41] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, J. ROMAN, M. ZOUNON. , *Towards resilient parallel linear Krylov solvers: recover-restart strategies*, Inria, July 2013, n⁰ RR-8324, 36 p. , http://hal.inria.fr/hal-00843992

[42] M. CHANAUD, L. GIRAUD, D. GOUDIN, J.-J. PESQUÉ, J. ROMAN. , *A Parallel Full Geometric Multigrid Solver for Time Harmonic Maxwell Problems*, Inria, July 2013, n⁰ RR-8335, 25 p. , Preliminary version of a paper to appear in SIAM SISC, http://hal.inria.fr/hal-00847966

[43] O. COULAUD, P. BORDAT, P. FAYON, V. LEBRIS, I. BARAILLE, R. BROWN. , *Extensions of the Siesta DFT Code for Simulation of Molecules*, Inria, February 2013, n⁰ RR-8221, 25 p. , http://hal.inria.fr/hal-00787088

[44] O. COULAUD, L. GIRAUD, P. RAMET, X. VASSEUR. , *Deflation and augmentation techniques in Krylov linear solvers*, Inria, February 2013, n⁰ RR-8265, 25 p. , http://hal.inria.fr/hal-00803225

[45] S. DONFACK, J. DONGARRA, M. FAVERGE, M. GATES, J. KURZAK, P. LUSZCZEK, I. YAMAZAKI. , *On Algorithmic Variants of Parallel Gaussian Elimination: Comparison of Implementations in Terms of Performance and Numerical Properties*, 2013, http://hal.inria.fr/hal-00867837

[46] X. LACOSTE, M. FAVERGE, P. RAMET, S. THIBAULT, G. BOSILCA. , *Taking advantage of hybrid systems for sparse direct solvers via task-based runtimes*, Inria, January 2014, n⁰ RR-8446, 25 p. , http://hal.inria.fr/hal-00925017