Fifth Workshop of the CNPq-Inria HOSCAR Project on

High Performance Computing
and Scientific Data Management
Driven by Highly Demanding Applications

Inria Sophia Antipolis-Méditerranée, France

September 21-24, 2015

http://www.lncc.br/hoscarmeeting2015

Abstract

This workshop is the final workshop of the collaborative project between the CNPq/Brazil - Inria/France which involves Brazilian and French researchers in the field of computational science and scientific computing. The general objective of the workshop is to setup a Brazil-France collaborative effort for taking full benefits of future high-performance massively parallel architectures in the framework of very large-scale datasets and numerical simulations. To this end, the workshop proposes multidisciplinary lectures on the following subjects: exploration of massively parallel architectures with high-performance programming languages, software components and libraries, large scientific data management and processing, and numerical schemes and scalable solvers for systems of differential equations

1 Summary of the Project

The prevalence of modern multicore technologies has made massively parallel computing ubiquitous and offers a huge theoretical potential for solving a multitude of scientific and technological challenges. Nevertheless, most applications and algorithms are not yet ready to utilize available architecture capabilities. Developing large-scale scientific computing tools that efficiently exploit these capabilities will be even more challenging with future exascale systems. To this end, a multidisciplinary approach is required to tackle the obstacles in manycore computing, with contributions from computer science, applied mathematics, and engineering disciplines.

Such is the framework of the collaborative project between the CNPq - Inria which involves Brazilian and French researchers in the field of computational science and scientific computing. The general objective of the project is to setup a Brazil-France collaborative effort for taking full benefits of future high-performance massively parallel architectures in the framework of very large-scale datasets and numerical simulations. To this end, the project has a multidisciplinary team with *computer scientists* who aim at exploiting the massively parallel architectures with high-performance

programming languages, software components, and libraries, and *numerical mathematicians* who aim at devising numerical schemes and scalable solvers for systems of Partial Differential Equations (PDEs). The driving applications are related to important scientific questions for the society in the following 4 areas: (i) *Resource Prospection*, (ii) *Reservoir Simulation*, (iii) *Cardiovascular System*, and (iv) *Astronomy*.

The researchers are divided in 3 fundamental groups in this project: (i) *Numerical schemes for PDE models*; (ii) *Scientific data management*; (iii) *High-performance software systems*.

Aside research goals, the project aims at making overall scientific results produced by the project available to the Brazilian and French scientific communities as well as to graduate students, and also establishing long-term collaborations beyond the current project. To this end, another objective of the project is the integration of the scientific results produced by the project within a common, user-friendly computational platform deployed over the partners' HPC facilities and tailored to the 4 aforementioned applications.

2   Practical Informations

The 5th Brazil-France workshop will take place at Inria Sophia Antipolis - Méditerranée [1]. The city of Sophia Antipolis is about 15 km far away from the Nice International Airport [2].

3   Participants

3.1   Brazilian Participants

- Pedro Dias, LNCC

- Alvaro Coutinho, High Performance Computing Center and Department of Civil Engineering, COPPE/UFRJ

- Rafael Keller Tesser, UFRGS

- Marta Mattoso, Computer Science Department, COPPE/UFRJ

- Philippe Navaux, UFRGS

- Fabio Porto, LNCC

- Antonio Tadeu Gomes, LNCC

- Frédéric Valentin, LNCC

---

[1]http://www.inria.fr/en/centre/sophia
[2]http://en.nice.aeroport.fr

3.2 French Participants

3.2.1 Researchers

- Reza Akbarinia, Inria Sophia Antipolis - Méditerranée, Zenith project-team

- Marie Bonnasse-Gahot, Inria Bordeaux - Sud-Ouest, Magique3D project-team and Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Alexandra Christophe-Argenvillier, Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Stéphane Descombes, Université Côte d'Azur and Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Julien Diaz, Inria Bordeaux - Sud-Ouest, Magique3D project-team

- Luc Giraud, Inria Bordeaux - Sud-Ouest, Hiepacs project-team

- Cédric Lachat, Inria Bordeaux - Sud-Ouest, Tadaam project-team

- Marie-Hélène Lallemand-Tenkes, Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Stéphane Lanteri, Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Raphaël Léger, Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Jean-François Méhaut, Inria Grenoble - Rhône-Alpes, Corse project-team

- Pierre Ramet, Inria Bordeaux - Sud-Ouest, Hiepacs project-team

- Claire Scheid, Université Côte d'Azur and Inria Sophia Antipolis - Méditerranée, Nachos project-team

- Patrick Valduriez, Inria Sophia Antipolis - Méditerranée, Zenith project-team

- Brice Videau, Inria Grenoble - Rhône-Alpes, Corse project-team

- Feng Xing, Inria Sophia Antipolis - Méditerranée, Coffee project-team

3.2.2 Representatives from Inria's management teams

- Tania Castro, European and International Partnership Department

- Thierry Priol, Director of European and International Partnership Department

- Jean Roman, Deputy Scientific Director

3.3 Guests

- Fabrice Dupros, BRGM, France

- Mariano Vazquez, Barcelona Computing Center

- Xavier Vigouroux, Atos, France

4 Program

- **September 21**

    9:30-10:30 Welcome coffee

    10:30-10:45 Workshop introduction
        Stéphane Lanteri (Inria) and Frédéric Valentin (LNCC)

    10:45-11:15 Inria's perspectives on Brazil-French cooperation after HOSCAR
        Thierry Priol, European and International Partnership Department

    11:15-11:45 LNCC's perspectives on Brazil-French cooperation after HOSCAR
        Pedro Dias and Frédéric Valentin, LNCC

    11:45-12:30 Overview talk - Group 1
        Achievements and perspectives on numerical schemes for wave propagation problems
        Stéphane Lanteri (Inria) and Frédéric Valentin (LNCC)

    12:30-14:00 Lunch

    14:00-14:45 Overview talk - Group 1
        A review of NACAD's developments within HOSCAR project
        Alvaro Coutinho, COPPE/UFRJ

    14:45:15:30 MHM for time-domain electromagnetics
        Raphaël Léger and Claire Scheid, Inria SAM

    15:30-16:00 Coffee break

    16:00:16:30 MHM for time-domain elastodynamics
        Marie-Hélène Lallemand-Tenkes, Inria SAM

    16:30-17:00 Distributed mesh and graph computations within PaMPA and PT-Scotch libraries
        Cédric Lachat, Inria BSO

    17:00 End of 1st Day

- **September 22**

  9:30-10:15 Guest talk
  > High Performance Computational Geosciences at BRGM
  > Fabrice Dupros, BRGM

  10:15-11:00 Overview talk - Group 2
  > Supporting efficient data analysis on numerical simulation output
  > Hermano Lustosa and Fabio Porto, LNCC and Patrick Valduriez, Inria SAM

  11:00-11:30 Coffee Break

  11:30-12:00 Integrating Big Data and relational data with a functional SQL-like query language
  > Patrick Valduriez, Inria SAM

  12:00-12:30 Analyzing related raw data files through dataflows
  > Marta Mattoso, COPPE/UFRJ

  12:30-14:00 Lunch

  14:00-14:30 Modelling of elastic Helmholtz equations by a hybridizable discontinuous Galerkin (HDG) method for geophysical applications
  > Marie Bonnasse-Gahot, Inria BSO and SAM

  14:30-15:00 Implicit hybridizable discontinuous Galerkin method for time-domain electromagnetics
  > Alexandra Christophe-Argenvillier, Inria SAM

  15:00-15:30 Hybrid dimensional Darcy flow in a fractured porous medium and parallel implementation in code ComPass
  > Feng Xing, Inria SAM

  15:30-16:00 Coffee break

  16:00-16:30 Simulation of overdecomposition based load balancing for parallel applications
  > Rafael Keller Tesser, UFRGS

  16:30-17:00 Implementation of a flexible, scalable simulator for the family of MHM methods
  > Antonio Tadeu Gomes, LNCC

  17:00 End of 2nd Day

- **September 23**

  9:30-10:15 Guest talk

  HPC-based multi-physics simulations for the energy realm

  Mariano Vazquez, BSC

  10:15-10:45 Julien Diaz, Inria BSO

  10:45-11:30 Coffee Break

  11:30-12:00 Parallel hybrid implementation of an hybrid sparse linear solver

  Luc Giraud, Inria BSO

  12:00-12:30 Blocking strategy optimizations for sparse direct linear solver on heterogeneous
  architectures Pierre Ramet, Inria BSO

  12:30-14:00 Lunch

  14:00-15:00 Visit of Inria's Virtual Reality Lab

  14:00-17:00 Internal meeting - From HOSCAR to HPC4E

  17:00 End of 3rd Day

  20:30 Dinner in Antibes

- **September 24**

  9:30-10:15 Guest talk

  No escape, we have to collaborate

  Xavier Vigouroux, Atos

  10:15-11:00 Overview talk - Group 3

  High-performance software systems

  Jean-François Méhaut, Inria GRA and Philippe Navaux, UFRGS

  11:00-11:30 Coffee Break

  11:30-12:00 FPHadoop: efficient execution of parallel jobs over skewed data

  Reza Akbarinia, Inria SAM

  12:00-12:30 BOAST : performance portability using meta-programming and auto-tuning

  Brice Videau, Inria GRA

  12:30-14:00 Lunch

  14:00-16:30 Summary of the Workshop

  16:30 End of the Workshop

  17:00 End of 4th Day

5   List of Abstracts

September 21

- Raphaël Léger, **Stéphane Lanteri**, Claire Scheid, Diego Paredes and **Frédéric Valentin**

- Inria Sophia Antipolis-Méditerranée, Nachos project-team and LNCC and Universidad Catlica de Valparaiso

- Achievements and perspectives on numerical schemes for wave propagation problems

In this talk we will review our achievements of high order numerical schemes for the solution of PDE models of electromagnetic and elastodynamic wave propagation problems, and also present some perspectives in the contact of the HOMAR associate team which was launched on January 2015.

- **Raphaël Léger**, **Claire Scheid**, Stéphane Lanteri, Diego Paredes and Frédéric Valentin

- Inria Sophia Antipolis-Méditerranée, Nachos project-team and LNCC and Universidad Catlica de Valparaiso

- MHM for time-domain electromagnetics

In this work, we are concerned with the numerical modeling of time dependent electromagnetic waves propagation problems with strong multiscales features (possibly in space and time). The starting point PDE model is sthe system of time-domain Maxwell equations. In this context we would like to contribute in the design of innovative numerical methods particularly well suited to the simulation of such problems. Indeed when a PDE model is approximated via classical finite element type method, it may suffer from a loss of accuracy when the solution presents multiscale features on coarse meshes. To address this issue, we rely on the concept of multiscale basis functions that is one solution to allow for accuracy even on coarse meshes. These basis functions are defined via algebraic relations. Contrary to classical polynomial approximation, they render by themselves a part of the high-contrast features of the problem at hand. Recently, a new family of finite element methods has been introduced, referred as Multiscale Hybrid-Mixed methods (MHM), which is well adapted to the simulation of high-contrast or heterogeneous problems. The underlying approach relies on a two level discretization. Shortly, basis functions computed on a fine (second level) mesh allow for the reconstruction of the solution on a coarse (first level) mesh. Such MHM have been initially designed in the context of stationary problems, such as Darcy flows. In this work, we propose to extend the concept of MHM to time dependent electromagnetic wave propagation problems. The model problem relies on the time dependent Maxwell's equations. The continuity of the electric field is relaxed via the introduction of a Lagrange multiplier. The solutions are expressed on a basis computed at the second level that incorporates the heterogeneity of the problem via the resolution of a PDE. Several schemes are proposed from implicit to explicit time schemes and continuous finite elements to discontinuous ones for the spatial discretization of the local problems at the second level. We will present some results on the validity of the algorithm from both theoretical and numerical point of view and first numerical results in 2D.

- **Marie-Hélène Lallemand-Tenkes** and Frédéric Valentin

- Inria Sophia Antipolis-Méditerranée, Nachos project-team

- MHM for time-domain elastodynamics

In this talk, we propose and define a Mixed Hybrid Multiscale (MHM) method for solving elastic waves propagation in heterogeneous and/or anisotropic media. That method is particularly well suited to media exhibiting multiscale features in the material(s) in presence. The MHM method presented in this work is based on the hybrid mixed velocity-stress formulation of the elastodynamic system. It is particularly attractive as it is structured as fully parallelizable and defined in a general framework. Semi-discretization in space then time integration is defined in the MHM method. The mixed hybrid formulation is then expressed as a collection of global and local problems, where the global problem is related to the Lagrange multiplier, defined at each interface element of some given (coarse) mesh, and is directly related to the traction. Local problems are associated with the splitting of the solution couple velocity/stress in each coarse element. Multiscaling is taken into account via that Lagrange multiplier by introducing a two level discretization strategy. Discretization of the Lagrange multiplier at the coarse level directly defines the discrete coarse solution of both velocity and stress tensor via the fine discretization of each coarse element spaces. The full algorithm is given and we draw the studies that are persued and the perspectives we may consider.

- **Cédirc Lachat** and François Pellegrini

- Inria Bordeaux - Sud-Ouest, Tadaam project-team

- Distributed mesh and graph computations within PaMPA and PT-Scotch libraries

This talk presents some new features of PaMPA, a library dedicated to the management of distributed meshes, including parallel repartitioning and parallel remeshing features. PaMPA performs parallel remeshing by using any sequential remesher (e.g. MMG3D). We present the first scalability results where we generate high quality, isotropic tetrahedral meshes of above a billion elements. PaMPA can also manage concurrently several interlinked unstructured meshes, so as to perform parallel multi-grid computations. A great deal of the algorithms are based on graph partitioning. Many of its algorithms are implemented in PT-Scotch library. New research on static graph mapping on large clusters will also be presented.

September 22

- **Fabrice Dupros**

- BRGM

- High Performance Computational Geosciences at BRGM

Due space and time scales, Geosciences are among the most demanding in computational resources. Considering the complete workflow (from 3D geological modelling to simulations), we will illustrate the impact of high performance computing to tackle large-scale problems. Scientific data management, complex workflow and uncertainties propagation will also be discussed with examples from geophysics and energy area. Past and ongoing collaborations with Brazil will also be described.

- Hermano Lustosa, **Fabio Porto** and **Patrick Valduriez**

- LNCC and Inria Sophia Antipolis-Méditerranée, Zenith project-team

- Supporting efficient data analysis on numerical simulation output

Numerical Simulations generate an ever increasing amount of data that comes from more precise and longer simulations, as much as as a function on the number of trials exercised during a parameter sweep evaluation. As the size of data increases, designing specific programs to efficiently access quantitatively the simulation results becomes a daunting task. In this talk we will present the experiments we have carried on using different database systems to manage numerical simulation data. We compare a traditional relational database, a column-store and a multidimensional array system. Our conclusions indicate that a single solution does not cover all the different needs and some work needs to be applied to gain efficiency in multidimensional array systems whenever a irregular space distribution is involved.

- **Patrick Valduriez**

- Inria Sophia Antipolis-Méditerranée, Zenith project-team

- Integrating Big Data and relational data with a functional SQL-like query language

Multistore systems have been recently proposed to provide integrated access to multiple, heterogeneous data stores through a single query engine. In particular, much attention is being paid on the integration of unstructured big data typically stored in HDFS with relational data. One main solution is to use a relational query engine that allows SQL-like queries to retrieve data from HDFS, which requires the system to provide a relational view of the unstructured data and hence is not always feasible. In this paper, we introduce a functional SQL-like query language that can integrate data retrieved from different data stores and take full advantage of the functionality of the underlying data processing frameworks by allowing the ad-hoc usage of user defined map/filter/reduce operators in combination with traditional SQL statements. Furthermore, the query language allows for optimization by enabling subquery rewriting so that filter conditions can be pushed inside and executed at the data store as early as possible. Our approach is validated with two data stores and a representative query that demonstrates the usability of the query language and evaluates the benefits from query optimization.

- Vtor Silva, Daniel de Oliveira, Patrick Valduriez, Alvaro Coutinho and **Marta Mattoso**

- COPPE/UFRJ and Inria Sophia Antipolis-Méditerranée, Zenith project-team

- Analyzing related raw data files through dataflows

Computer simulations may ingest and generate a large number of raw data files. Most of these files follow a de facto standard format established by the application domain, e.g., SEGY for seismic, HDF5 or NetCDF for computational mechanics. Although these formats are supported by a variety of programming languages, libraries and programs, analyzing thousands or millions of files requires developing specific programs. Database Management Systems (DBMS) are not suited for this, because they require parsing the raw data file, structuring its contents to load it in a database so it can be queried, which gets heavy at large-scale. When computer simulations are managed by a Scientific Workflow Management System (SWfMS), they can take advantage of provenance data to relate and analyze raw data files produced during workflow execution. When the SWfMS is dataflow-aware, it can register provenance data and the relationships among elements of raw data files altogether in a database which is useful to access the contents of a large number of files. In this talk, we present a dataflow approach for analyzing element data from several related raw data files. Our approach is complementary to the existing single raw data file analysis approaches. We use a Reverse Time Migration workflow from Oil and Gas domain as a case study. The cost for raw data extraction and loading is approximately 3.7% of the total application execution time. Its analytical value allows for much more significant improvements on the total simulation execution life cycle. Extras: Systems like NoDB, RAW and FastBit, have been proposed to index and query raw data files without the overhead of using a DBMS. However, these solutions are focused on analyzing one single large file instead of several related files. In this case, when related files are produced and required for analysis, the relationship among elements within file contents must be managed manually, with specific programs to access raw data. Thus, this data management may be time-consuming and error-prone. However, SWfMS register provenance at a coarse grain, with limited analysis on elements from raw data files.

- Henri Calandra, **Marie Bonnasse-Gahot**, Julien Diaz and Stéphane Lanteri

- Inria Bordeaux - Sud-Ouest, Magique3D project-team, and Inria Sophia Antipolis-Méditerranée, Nachos project-team, and TOTAL

- Modelling of elastic Helmholtz equations by a hybridizable discontinuous Galerkin (HDG) method for geophysical applications

The advantage of performing seismic imaging in frequency domain is that it is not necessary to store the solution at each time step of the forward simulation. But the main drawback of the elastic Helmholtz equations, when considering 3D realistic elastic case, lies in solving large linear systems, which represents today a challenging tasks even with the use of high performance computing (HPC). To reduce the size of the global linear system, we develop a hybridizable discontinuous Galerkin method (HDGm). It consists in expressing the unknowns of the initial problem in function of the trace of the numerical solution on each face of the mesh cells. In this way the size of the matrix to

be inverted only depends on the number of degrees of freedom on each face and on the number of the faces of the mesh. The solution to the initial problem is then recovered thanks to independent elementwise calculation.

- **Alexandra Christophe-Argenvillier**, Stéphane Descombes and Stéphane Lanteri

- Inria Sophia Antipolis-Méditerranée, Nachos project-team

- Implicit hybridizable discontinuous Galerkin method for time-domain electromagnetics

Discontinuous Galerkin (DG) methods have been the subject of numerous research activities in the last 15 years and have been successfully developed for various physical contexts modeled by elliptic, mixed hyperbolic-parabolic and hyperbolic systems of PDEs. One major drawback of high order DG methods is their intrinsic cost due to the very large number of globally coupled degrees of freedom as compared to classical high order conforming finite element methods. Different attempts have been made in the recent past to improve this situation and one promising strategy has been recently proposed by Cockburn et al. in the form of so-called hybridizable DG formulations. The distinctive feature of these methods is that the only globally coupled degrees of freedom are those of an approximation of the solution defined only on the boundaries of the elements of the discretization mesh. The present work is concerned with the study of such a hybridizable DG method for the three-dimensional time-domain Maxwell equations time integrated by an implicit scheme.

- Konstantin Brenner, Roland Masson and **Feng Xing**

- Inria Sophia Antipolis-Méditerranée, Coffee project-team

- Hybrid dimensional Darcy flow in a fractured porous medium and parallel implementation in code ComPass

We consider a tracer model in a porous medium, which includes a complex network of planar fractures. The solute is transported by a velocity field calculated from a hybrid dimensional Darcy flow model accounting for the flow within the 2D fracture network, the flow in the surrounding 3D matrix and for the mass exchanges between the matrix and the fracture network. The Darcy fluxes are computed using the Vertex Approximate Gradient finite volume scheme (VAG) on general polyhedral meshes, and the tracer discretization uses a two point upwind flux, which can be combined with a second order MUSCL type reconstruction. We implement this model in the framework of code ComPass (Computing Parallel Architecture to Speed up Simulations), which focuses on parallel high performance simulation (MPI). Good strong scalability is shown by the numerical results.

- Arnaud Legrad, Philippe O.A. Navaux and **Rafael Keller Tesser**

- UFRGS

- Simulation of overdecomposition based load balancing for parallel applications

In a previous work, we used overdecomposition based load balancing to improve the performance of a seismic wave model called Ondes3D. This was time-consuming work, since we had run several executions with number of different load balancer heuristics. Another difficulty is that this kind of experiments is often run in time-shared systems. So, the system may not be available to you when you need it. In this presentation, we will discuss our ongoing effort to solve these problems using simulation. For this purpose, we are developing a strategy to simulate overdecomposition based load balancing using SimGrid. The main idea is to use traces from the application to replay its execution using different load balancers. One advantage of this strategy is that the application would be executed only once. Another one is that the replay is much faster than the actual execution, since it does not actually execute the computation. Besides that, as the simulation results are deterministic, there would be no need to collect more than one sample per configuration. A secondary benefit is that the target machine would only be needed for the initial tracing execution. For all these reasons, this strategy should considerably decrease the time need to test load balancing heuristics. For our purposes, we modified the MPI replay capabilities included in SimGrid, to support the migration of tasks. Besides that we intend to to integrate load balancing heuristics into the simulation. Once that is done, we should be ready to use SimGrid to run load balancing experiments.

- **Antonio Tadeu Gomes**, Diego Paredes and Frederic Valentin
- LNCC and Universidad Catlica de Valparaiso
- Implementation of a flexible, scalable simulator for the family of MHM methods

In the context of the Hoscar project weve been progressing work on the implementation of a flexible, scalable simulator based on the family of Multiscale Hybrid-Mixed (MHM) methods. The MHM method allows solving (global) problems on coarse meshes while providing solutions with high-order precision by exploring the loosely-coupled strategy of embedding independent (local) subproblems in the upscaling procedure. Our original approach to this implementation has been based on the use of different programming languages to tackle different simulation issues. More specifically, we've adopted Erlang as the base implementation for the communicating processes of the new library, and C++ for the numerical computing processes that ultimately solve the global and local problems. In previous talks we've presented the advantages of adopting Erlang not only in terms of high productivity but also with regard to fault tolerance and low impact on performance for stationary, 2D scalar problems. Weve also shown how the Erlang processes are loosely integrated with numerical computing processes, thus indicating the potential of the Erlang implementation in being adopted in other contexts. In this talk we will discuss about some early design decisions on the implementation of the MHM simulator that have been reconsidered so to give better support for transient, 3D and vectorial problems, whilst keeping the original multi-language approach. We will also show some preliminary results obtained from this redesign.

September 23

- **Mariano Vazquez**

- Barcelona Supercomputing Center

- HPC-based multi-physics simulations for the energy realm

Thanks to a strong commitment through collaboration projects, the CASE department of BSC develops research lines with direct impact in engineering-related simulations. The field of energy production represents a great challenge due to the complexity of the multi-physics simulations, being HPC a must. In this talk, we will describe our research, specially which is related to the EU-Brazil HPC4E project, very recently awarded.

- Emmanuel Agullo, Hélène Barucq, Lionel Boillot, George Bosilca, Henri Calandra and **Julien Diaz**

- Inria Bordeaux - Sud-Ouest, Hiepacs and Magique3D project-teams, and TOTAL

- Portable task-based programming for seismic wave propagation simulation in time domain

We propose to apply task-based programming to optimize a 3D anisotropic elastodynamics wave propagation simulation based on a Discontinuous Galerkin space discretization, associated with a Leap-Frog time scheme, which leads to a quasi-explicit matrix linear system involving only local computations (i.e. cell by cell) at every half-timestep. The original, message passing parallel implementation uses a domain decomposition with one domain per process, and lead to a uneven work balance between processes. As a result, the optimization for each architecture is time consuming and the solution is never portable. We successfully overcame this problem on shared memory architectures (ccNUMA node and Intel Xeon Phi accelerator) by changing the programming paradigm with a task-based approach and the use of the PaRSEC runtime system. The two key-features for efficient task scheduling are finer granularity than one subdomain per core and work-stealing depending on data locality. The results showed very good parallel efficiency and were portable on these machines. We now plan to address distributed memory architectures in order to target clusters of hybrid multicore nodes and coprocessors.

- **Luc Giraud**

- Inria Bordeaux - Sud-Ouest, Hiepacs project-team

- Parallel hybrid implementation of an hybrid sparse linear solver

In this talk we describe the parallel hybrid implementation based on MPI and threads of MaPHyS an hybrid direct/iterative solver for large sparse linear system. The scalability of the resulting solver will be discussed and illustrated on a few real life test problems solved on a large computing platform with up to more than 24 kcores (Hopper@LBNL).

- Mathieu Faverge, Grégoire Pichon, **Pierre Ramet** and Jean Roman

- Inria Bordeaux - Sud-Ouest, Hiepacs project-team

- Blocking strategy optimizations for sparse direct linear solver on heterogeneous architectures

In the context of solving sparse linear systems, an ordering process partitions the matrix graph to minimize both fill-in and computational cost. We found that the ordering strategy used within supernodes might be enhanced to reduce the number of off-diagonal blocks, and then increases block sizes and kernel performance. This turns to be into the same complexity as the factorization algorithm, but allows for more efficient BLAS kernels. On the other side, supernodes that are too large need to be split to create more parallelism. The regular splitting strategy when applied locally impacts significantly the number of off-diagonal blocks and might have negative effect on the efficiency. In this talk, we present both a new strategy to improve supernodes ordering and splitting strategy that both enlarge the off-diagonal block sizes without changing the computational cost of the factorization. Performance improvement gains on the supernodal solver PaStiX are shown on multi-cores and heterogeneous architectures.

September 24

- **Xavier Vigouroux**

- Atos

- No escape, we have to collaborate

In this presentation will be present how the current technology trends implies strong collaboration to efficiently rely on HPC supercomputer.

- **Reza Akbarinia**

- Inria Sophia Antipolis-Méditerranée, Zenith project-team

- FPHadoop: efficient execution of parallel jobs over skewed data

Although MapReduce has been praised for its high scalability and fault tolerance, it has been criticized in some points, in particular, its poor performance in the case of data skew. There are important cases where a high percentage of processing in the reduce side is done by a few nodes, or even one node, while the others remain idle. There have been some attempts to address the problem of data skew, but only for specific cases. In particular, there is no proposed solution for the cases where most of the intermediate values correspond to a single key, or when the number of keys is less than the number of reduce workers. In this talk, we present FP-Hadoop, a system that makes the reduce side of MapReduce more parallel, and efficiently deals with the problem of data skew in the reduce side. In FP-Hadoop, there is a new phase, called intermediate reduce (IR), in which blocks of intermediate values, constructed dynamically, are processed by intermediate reduce workers in parallel, by using a scheduling strategy. By using the IR phase, even if all intermediate values belong to only one key, the main part of the reducing work can be done in parallel by using the computing resources of all available workers. We implemented a prototype of FP-Hadoop, and conducted extensive experiments over synthetic and real datasets. We achieved

excellent performance gains compared to native Hadoop, e.g. more than 10 times in reduce time and 5 times in total execution time.

- Thierry Deutsch, Luigi Genovese, Dimitri Komatitsch, Jean-Franois Mhaut, Kevin Pouget and **Brice Videau**

- Inria Grenoble - Rhône-Alpes, Corse project-team and CEA/INAC and LMA

- BOAST : performance portability using meta-programming and auto-tuning

Porting and tuning HPC applications to new platforms is of paramount importance but tedious and costly in terms of human resources. Unfortunately those efforts are often lost when migrating to new architectures as optimization are not generally applicable. That's why we are promoting scientific application auto-tuning. While computing libraries might be auto-tuned, usually HPC applications are hand-tuned. In the fast paced world of HPC nowadays, we believe that HPC applications kernels should be auto-tuned instead. Unfortunately, the investment to setup a dedicated auto-tuning framework is usually too expensive for a single application. Source to source transformations or compiler based solutions exist but sometimes prove too restrictive to cover all use-cases. We thus propose BOAST a meta-programming framework aiming at generating parametrized source code. The aim is for the programmer to be able to orthogonally express optimizations on a computing kernel, enabling a thorough search of the optimization space. This also allows a lot of code factorization and thus code base reduction. We will demonstrate the use of BOAST and show results obtained on state of the art scientific applications.

6 Sponsors

- INRIA
- CNPq