

# First Brazil-France Workshop

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



This workshop is part 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.

10-13 September, **2012**  
Petrópolis, Brazil

Organizing Committee:

**Pedro Leite S. Dias**  
LNCC/Brazil  
**Antônio Tadeu Gomes**  
LNCC/Brazil  
**Stéphane Lanteri**  
INRIA Sophia-Antipolis/France  
**Fabio Porto**  
LNCC/Brazil  
**Frédéric Valentin**  
LNCC/Brazil

## 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 5 areas: (i) *Resource Prospection*, (ii) *Reservoir Simulation*, (iii) *Ecological Modeling*, (iv) *Astronomy Data Management*, and (v) *Simulation Data Management*. 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 5 aforementioned applications.

## 2. PRACTICAL INFORMATION

The First Brazil-France workshop will take place at the LNCC ([www.lncc.br](http://www.lncc.br)), hosted at the city of Petrópolis (<http://www.visitepetropolis.com/default.asp?id=en>) in the Rio de Janeiro state (<http://www.rioguiaoficial.com.br/en/>). The research center is about 50 km

far away from the Rio de Janeiro International Airport (Tom Jobim Airport, also called Galeão).

The LNCC will set up a free-of-charge transportation between the Airport and the Casablanca Hotel, as well as a daily-transportation from the Hotel and the LNCC. The address of the Casablanca Hotel is: Av. Koeler, No. 60 - Centro, Petrópolis, Phone: +55 24 22432810, e-mail: reservas@casablancahotel.com.br

For any further information or help, please contact our secretary (Tathiana Figueiredo) through the following e-mail address: tathi@lncc.br or by phone: +55 24 22336101

### 3. PARTICIPANTS

#### 3.1. Group 1 and 3.

- Pablo Blanco (LNCC/Brazil)
- Lionel Boillot (INRIA Sophia-Antipolis/France)
- Théophile Chaumont (INRIA Sophia-Antipolis/France)
- Renato Elias (COPPE-UFRJ/Brazil)
- Luc Giraud (INRIA Bordeaux/France)
- Antonio Tadeu Gomes (LNCC/Brazil)
- Stéphane Lanteri (INRIA Sophia-Antipolis/France)
- Alexandre Madureira (LNCC/Brazil)
- Marcio Murad (LNCC/Brazil)
- François Pellegrini (University of Bordeaux/France):
- Pierre Ramet (University of Bordeaux/France)
- Josias José da Silva (COPPE-UFRJ/Brazil)
- Frédéric Valentin (LNCC/Brazil)

#### 3.2. Group 2.

- Reza Akbarinia (INRIA and LIRMM Montpellier/France)
- Luiz Nicolaci da Costa (Observatório Nacional/Brazil)
- Luiz Gadelha (LNCC/Brazil)
- Fabio Porto (LNCC/Brazil)
- Marta Matoso (UFRJ/Brazil)
- Patrick Valduriez (INRIA Sophia-Antipolis and LIRMM Montpellier/France)
- Eduardo Ogasawara (CEFET-UFRJ/Brazil)

## 4. PROGRAM

- **10/09 Welcome, Opening and Group 1 and 3**

10:00-13:00 Morning: Welcome and Opening Section

10:00-10:30 Brazilian Representative from the Brazilian Ministry of Science, Technology and Innovation (MCTI)

10:30-11:00 French Representative from the French Embassy at Brazil

11:00-11:30 Stéphane Lanteri, C2S@Exa

11:30-12:00 Coffee Break

12:00-12:30 Pedro Dias, LNCC

12:30-13:00 Antonio Tadeu, SINAPAD

13:00-14:00 Lunch

14:00-17:30 Afternoon: Group 1 and 3

14:00-14:45 Pablo Blanco. *Anatomically Detailed Models of the Cardiovascular System*

14:45-15:30 François Pellegrini. *Some Software Tools for Parallel Graph Partitioning and Distributed mesh management*

15:30-16:00 Coffee Break

16:00-16:45 Renato Elias. *Parallel Mesh Multiplication: Towards Unstructured Grids Peta Scale Simulations*

16:45-17:30 Pierre Ramet. *Scheduling for Sparse Solver on Manycore Architectures*

- **11/09 Group 1 and 3**

9:30-13:00 Morning: Group 1 and 3

9:30-10:15 Stéphane Lanteri. *Development of Discontinuous Galerkin Methods for Time-Domain Electromagnetics*

10:15-11:00 Luc Giraud. *High Performance Numerical Algorithms and Tools*

11:00-11:30 Coffee Break

11:30-12:15 Lionel Boillot. *Parallel optimization of a Discontinuous Galerkin Method for Seismic Imaging Simulation. Towards the Intel MIC acceleration*

12:15-13:00 Josias José da Silva. *3D Seismic Imaging of Complex Geology*

13:00-14:00 Lunch

14:00-18:00 Afternoon: Group 1 and 3

14:00-14:45 Marcio Murad. *Computational Hydro-Geomechanical Models of Pre-Salt Reservoirs*

14:45-15:30 Frédéric Valentin. *Multiscale Hybrid-Mixed Methods and Numerical Zoom*

15:30-16:00 Coffee Break

16:00-16:45 Alexandre Madureira. *Finite Element Methods for Multiscale Problems*

16:45-17:30 Théophile Chaumont. *Upscaling for the Laplace Problem using a Discontinuous Galerkin Method*

17:30-18:00 First Round Table

20:00- Social Event

• **12/09 Group 2**

9:30-13:00 Morning: Group 2

9:30-10:15 Patrick Valduriez. *Zenith: a Hybrid P2P/cloud for Big Scientific Data*

10:15-11:00 Fabio Porto. *Data Management and Modeling Techniques for Scientific Data*

11:00-11:30 Coffee Break

11:30-12:15 Reza Akbarinia. *DynPart: Dynamic Partitioning for Large-Scale Databases*

12:15-13:00 Luiz Gadelha. *Provenance Management in Many-Task Computing*

13:00-14:00 Lunch

14:00-17:30 Afternoon: Group 2

14:00-14:45 Marta Matoso. *Exploring Provenance in High Performance Scientific Computing*

14:45-15:30 Luiz Nicolaci da Costa

15:30-16:00 Coffee Break

16:00-16:45 Eduardo Ogasawara. *Chiron: An Algebraic-Based Parallel Scientific Workflow Engine* by Eduardo Ogasawara

16:45-17:30 Second Round Table

• **13/09 Group 1 and 3, and General Discussions**

9:30-13:00 Morning: Group 1 and 3

9:30-10:15 Antonio Tadeu Gomes. *SPiNMe: An Environment for the Rapid Prototyping of New Numerical Methods*

10:15-11:00 to be confirmed

11:00-11:30 Coffee Break  
 11:30-12:15 to be confirmed  
 12:15-13:00 Third Round Table  
 13:00-14:00 Lunch  
 14:00-17:40 Afternoon: Discussions

## 5. LIST OF ABSTRACTS

### 5.1. Opening.

- *C2S@Exa: a Large-Scale Initiative on High Performance Computing for Computational Sciences* by Stéphane Lanteri

The C2S@Exa (Computer and Computational Sciences at Exascale) large-scale initiative is a recently launched initiative for the establishment of a continuum of expertise in the computer science and numerical mathematics domains, gathering researchers from INRIA project-teams whose research and development activities are tightly linked to high performance computing issues in these domains. Indeed, 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.

### 5.2. Group 1 and 3.

- *Anatomically Detailed Models of the Cardiovascular System* by Raul Feijóo and Pablo Blanco

Over recent decades we have seen the enormous growth of scientific publications seeking to model certain aspects of the behavior of the human cardiovascular system. Several of these publications tried to represent its behavior using lumped models and distributed models capable to simulate the behavior of blood flow circulation in large deformable arteries. With the increasing accuracy of medical imaging, reconstruction of specific arterial geometry emerged allowing the modeling of local blood circulation in patient-specific arterial districts. Nevertheless, a truly integrative modeling of the cardiovascular system in terms of global/local scales, further coupled with respiratory and neural system -just to mention two- has been missing in part due to lack of data

and in part due to lack of computational resources (modeling the entire cardiovascular system is still unfeasible -?-).

Our aim with this talk is to launch the definite bridging among anatomy, physiology and computational modeling by creating an arterial network following stringent anatomical and physiological considerations. This anatomically-consistent arterial topology incorporates all the arterial segments acknowledged in the medical community going up to 1.200 vessels, and with arteries having lumen radius ranging from 15mm (aorta artery) to 0,15mm (perforator arteries). This model serves as a framework to investigate, in-silico, a large number of cardiovascular conditions, and represents the state-of-the-art achieved within the HeMoLab group at the LNCC (<http://hemolab.lncc.br>). The blood flow in this network is simulated by using 1D models for the arterial segments and 0D models for the terminal elements to account for the remaining resistance of the vascular network present in the peripheral beds, A discussion about the calibration of this model will be raised, involving the setting of vessel and terminal parameters, and some simulations of the hemodynamics in the upper and lower limbs will be presented. Current research is directed towards the modeling of the entire arterial network and the coupling with venous, cardiac, neural and respiratory systems in order to widen the range of physiological and pathophysiological applications.

In this context, we are convinced that this model will contribute to (i) understand the complex systemic interaction taking place at the different scales in the cardiovascular system, (ii) make modeling-based diagnoses, (iii) perform accurate surgical planning, and (iv) aid in the simulation-based medical training and education.

- *Parallel Optimization of a Discontinuous Galerkin Method for Seismic Imaging Simulation. Towards the Intel MIC Acceleration* by Lionel Boillot

Seismic imaging simulation requires high-order discretization methods and deep-parallel algorithms in order to produce accurate images of the subsurface. Our code is based on DG (Discontinuous Galerkin) space discretization method. DG main advantage is the h/p adaptivity: elements with both variable size and polynomial degree. Moreover, it does not involve any mass-matrix inversion. Furthermore, the local nature of DG algorithms means that computations are done element by element. In order to treat big volume of data, a first level of parallelization is obvious with the domain decomposition and the use of the MPI norm. But processors still have lots of small loops to compute in each element. To overcome this, we tried a MPI-OpenMP

hybridization. It helps but acceleration is limited to the number of cores in the actual processors. Further optimizations are needed and especially the use of accelerators. On the one hand, pure GPGPU (General-Purpose computing in Graphics Processing Units) offers hundred of computing units but requires new language and algorithm reformulation. On the other hand, the new Intel MIC (Many Integrated Cores) architecture can run without any modification and contains many cores, more than fifty and multi-threaded. MIC will arrive in the early 2013. I will present on-going work to prepare the code for the MIC: reduction of MPI communication, memory cache optimization, SSE vectorization.

- *Upscaling for the Laplace Problem using a Discontinuous Galerkin Method* by Théophile Chaumont

Operator-based upscaling is a method designed to tackle multi-scale problems, using a finite element discretisation. We seek the discrete solution as the sum of a coarse and fine components. Calculations on the fine component are simplified thanks to artificial boundary conditions.

We present the first attempt to adapt classical FEM operator-based upscaling, to the DGM. We study the Laplace problem as a first step toward the study of the wave equation.

- *Parallel Mesh Multiplication: Towards Unstructured Grids Peta Scale Simulations* by Renato Elias, Jose Camata and Alvaro Coutinho

In this talk we introduce a parallel mesh multiplication scheme that enables very fast unstructured grid generation for high-fidelity edge-based stabilized finite element computations of fluid flow and related multiphysics problems found in fluid-structure interaction and polydisperse mixtures. We present the generation of communication graphs under a non-blocking master-slave MPI paradigm and the potential for generating powerful geometric-based multi-level preconditioners for Krylov-space solvers. For capturing arbitrary immersed surfaces given in standard formats we also discuss the utilisation of parallel octree-based spatial structures to generate the base tetrahedra mesh. Issues of core skew allocation and graph reordering to minimize cache effects are also discussed.

- *High Performance Numerical Algorithms and Tools* by Luc Giraud

In this talk we will describe the various research efforts conducted in the HiePACS project to design and implement efficient numerical scalable software tools to help in the solution of large scientific computational problems. In particular we will review



activities for the solution of large sparse linear systems and n-body calculation based on multipole expansion.

- *SPiNMe: An Environment for the Rapid Prototyping of New Numerical Methods* by Antonio Tadeu Gomes

The development of new numerical methods is of great importance in computational science. Due to their many appealing properties, Finite Element (FEMs), Finite Volume (FVMs) and Finite Difference (DFMs) methods are of particular interest to this talk. Unfortunately, these methods can be time consuming to implement when one would like to apply them to realistic problems or make comparisons with other methods. In fact, before developing a new numerical method, for instance, a researcher needs to be familiar with methods that already exist in order to identify differences and common points between them. The process of working through the existing bibliography can be very time-consuming, particularly when the available numerical results presented for a given method are not sufficient to fully understand its behavior. In such a case, a researcher must implement the method to perform the desired tests, often with significant time-cost. In this talk, we present a new computational environment called SPiNMe (Software Productivity in Numerical Methods) to tackle the aforementioned issues. In the SPiNMe environment, the specification and implementation of such methods can be perpetuated and the experimentation with them can be made reproducible and easily compared with other related methods. The main features of SPiNMe may be outlined as follows: *i*) Flexibility without lack of comparability. The SPiNMe environment allows the various features underlying numerical method definitions to be explored; *ii*) Productivity. The SPiNMe environment provides a way for researchers to rapidly prototype their methods; *iii*) Long term compatibility. The SPiNMe environment only requires from researchers the use of a local web browser. To attain these features, we stick to three key design decisions for the SPiNMe environment. First, a single set of numerical library implementations is provided to the researcher for the implementation of his/her numerical method. Second, the provided numerical libraries can be parametrized either via input data or via "plug-in" code representing the particularities of a specific numerical method. Finally, plug-in code is implemented in a high-level language: therefore, researchers can focus more on prototyping their methods than on coding typical idioms of lower-level languages.

- *Development of Discontinuous Galerkin Methods for Time-Domain Electromagnetics* by Stéphane Lanteri

Nowadays, a variety of solution strategies exist for the computer simulation of electromagnetic wave propagation problems. Despite a lot of advances on numerical methods able to deal accurately and in a flexible way with complex geometries through the use of unstructured (non-uniform) discretization meshes, the FDTD (Finite Difference Time Domain) method is still the prominent modeling approach for realistic time domain computational electromagnetics, in particular due to the possible straightforward implementation of the algorithm and the availability of computational power. In the FDTD method, the whole computational domain is discretized using a structured (Cartesian) grid. This greatly simplifies the discretization process but also represents the main limitation of the method when complicated geometrical objects come into play. Besides, the last 10 years have witnessed an increased interest in so-called DGTD (Discontinuous Galerkin Time Domain) methods. Thanks to the use of discontinuous finite element spaces, DGTD methods can easily handle elements of various types and shapes, irregular non-conforming meshes, and even locally varying polynomial degree, and hence offer great flexibility in the mesh design. They also lead to (block-) diagonal mass matrices and therefore yield fully explicit, inherently parallel methods when coupled with explicit time stepping. Moreover, continuity is weakly enforced across mesh interfaces by adding suitable bilinear forms (often referred as numerical fluxes) to the standard variational formulations. In this talk, we will describe some recent developments aiming at improving the accuracy and the performances of a non-dissipative DGTD for the simulation of time-domain electromagnetic wave propagation problems involving general domains and heterogeneous media. The common objective of the associated studies is to bring the method to a level of computational efficiency and flexibility that allows to tackle realistic applications of practical interest.

- *Finite Element Methods for Multiscale Problems* by Alexandre Madureira

We discuss some finite element techniques that are suitable to deal with problems of multiscale type, and are prone to parallel implementations. In particular, we discuss Residual Free Bubble (RFB) methods, and as an application, we look into some equations that result from neuroscience and dimension reduction techniques (the art of deriving 2D equations from 3D problems in slender domains).

We also propose a RFB type method for a nonlinear, multiscale problem. We derive the method, and present some possible alternatives, depending on linearization choices. We show some error estimates under the assumption that the coefficients are highly oscillatory. The analysis is quite delicate due to the presence of the small scale parameter. The method itself applies in quite general situations, but the error estimates are restricted to the cases of periodic coefficients.

- *Computational Hydro-Geomechanical Models of Pre-Salt Reservoirs* by Marcio Murad
- *Some Software Tools for Parallel Graph Partitioning and Distributed Mesh Management* by François Pellegrini

The advent of massively parallel, NUMA machines, represents a new challenge for software designers. Numerical solvers have to scale up to hundred thousands of processing elements, which requires appropriate software tools for handling efficiently routine tasks such as mesh partitioning and data exchange. The purpose of this talk is to present the current state of development and directions for research of two such tools, Scotch and PaMPA.

Scotch and its offspring PT-Scotch are software tools dedicated to the computation of high quality graph partitions. Scotch is a robust sequential tool that has been developed for almost 20 years. Its parallel offspring PT-Scotch, which aims at providing the same features in parallel, can partition graphs with sizes up to several billion vertices, distributed over thousands of processors.

PaMPA (Parallel Mesh Partitioning and Adaptation) is a middleware library dedicated to the management of unstructured meshes distributed across the processors of a parallel machine. Its purpose is to relieve solver writers from the tedious and error prone task of writing again and again service routines for mesh handling, data communication and exchange, remeshing, and data redistribution.

We will present each of the tools, how they are coupled, and the challenges that we are facing within each project in order to achieve scalability on very large systems.

- *Scheduling for Sparse Solver on Manycore Architectures* by X. Lacoste, M. Faverge, Pierre Ramet

The emergence of many-cores architectures introduces variations in computation costs, which makes precise cost models hard to realize. Static schedulers based on cost models, like the one used in the sparse direct solver PaStiX, are no longer adapted. We describe the dynamic scheduler developed for the super-nodal method of PaStiX

to correct the imperfections of the static model. The solution presented exploit the elimination tree of the problem to keep the data locality during the execution.

- *3D Seismic Imaging of Complex Geology* by Patrícia Perez Ferreira, Josias José da Silva and Luiz Landau

This work present an overview on 3D seismic modeling concepts from the mathematical and physical theory to the computational practice. It is presented key aspects to explain seismic surveys and their importance to the oil industry in the search for petroleum reservoirs. The seismic modeling is performed using a complex model composed of several faults called Overthrust, which was created with the purpose to represent a real geology. Two examples are shown: single and line shots. Additional informations are addressed in the examples.

- *Multiscale Hybrid-Mixed Finite Element Methods and Numerical Zoom* by Frédéric Valentin

This work proposes a new family of finite element methods for porous media flows, named Multiscale Hybrid-Mixed (MHM) methods. The MHM method is a consequence of a hybridization procedure, and emerges as a method that naturally incorporates multiple scales while provides solutions with high-order precision for the primal and dual (or flux) variables. The computation of local problems is embedded in the upscaling procedure, which are completely independent and thus may be naturally obtained using parallel computation facilities. Also interesting is that the flux variable preserves the local conservation property using a simple post-processing of the primal variable.

The general framework is illustrated for the Darcy equation, and further extended to other operators taking part in the modeling of porous media flows (advection-diffusion and elasticity equations, for instance). The analysis results in a priori estimates showing optimal convergence in natural norms and providing a face-based a posteriori estimator. Regarding the latter, we prove that reliability and efficiency hold. Numerical results verify the optimal convergence properties as well as a capacity to accurately incorporate heterogeneity and high-contrast coefficients, showing in particular the great performance of the new a posteriori error estimator in driving mesh adaptativity.

We conclude that the MHM method, along with its associated a posteriori estimator, is naturally shaped to be used in parallel computing environments and appears to

be a highly competitive option to handle realistic multiscale boundary value problems with precision on coarse meshes.

### 5.3. Group 2.

- *DynPart: Dynamic Partitioning for Large-Scale Databases* by Reza Akbarinia

Scientific applications with very large databases, where data items are continuously appended, are becoming more and more common. The development of efficient workload-based data partitioning is one of the main requirements to offer good performance to most of those applications that have complex access patterns. However, the existing workload-based approaches, which are executed in a static way, cannot be applied to very large databases. In this talk, we present DynPart, a dynamic partitioning algorithm for continuously growing databases. DynPart efficiently adapts the data partitioning to the arrival of new data elements by taking into account the affinity of new data with queries and fragments. In contrast to existing static approaches, our approach offers a constant execution time, no matter the size of the database, while obtaining very good partitioning efficiency. We validated our solution through experimentation over real-world data; the results show its effectiveness.

- *Provenance Management in Many-Task Computing* by Luiz Gadelha

Large-scale scientific computations are often organized as a composition of many computational tasks linked through data flow. After the completion of a computational scientific experiment, a scientist has to analyze its outcome, for instance, by checking inputs and outputs of computational tasks that are part of the experiment. This analysis can be automated using provenance management systems that describe, for instance, the production and consumption relationships between data artifacts, such as files, and the computational tasks that compose the scientific application. In this presentation, we explore the relationship between high performance computing and provenance management systems, observing that storing provenance as structured data enriched with information about the runtime behavior of computational tasks in high performance computing environments can enable interesting and useful queries to correlate computational resource usage, scientific parameters, and data set derivation. We briefly describe how provenance of many-task scientific computations specified and coordinated by the Swift parallel scripting system is gathered and queried.

- *Data Management and Modeling Techniques for Scientific Data* by Fabio Porto

The DEXL lab is interested in developing data management techniques in support to scientific applications. In this talk we will present projects currently in progress in the lab. The first project involves the pre-processing of a scientific visualization applications in which the trajectory of thousands of virtual particles through 3D Mesh is computed. Our approach involves the adaptive balancing of the computation of sets of particles through cluster nodes using a greedy balancing algorithm and the modeling of the problem using an algebraic approach. The second project explores the efficient storage of meshes in secondary memory using SciDB, a scientific database. Finally, we will present our work in supporting the modeling and representation of scientific hypotheses in the context of in-silico research.

- *Exploring Provenance in High Performance Scientific Computing* by Marta Mattoso

Large-scale scientific computations are often organized as a composition of many computational tasks linked through data flow. After the completion of a computational scientific experiment, a scientist has to analyze its outcome, for instance, by checking inputs and outputs along computational tasks that are part of the experiment. This analysis can be automated using provenance management systems that describe, for instance, the production and consumption relationships between data artifacts, such as files, and the computational tasks that compose the scientific application. Due to its exploratory nature, large-scale experiments often present iterations that evaluate a large space of parameter combinations. In this case, scientists need to analyze partial results during execution and dynamically interfere on the next steps of the simulation. Features, such as user steering on workflows to track, evaluate and adapt the execution need to be designed to support iterative methods. In this talk we show examples of iterative methods, such as, uncertainty quantification, reduced-order models, CFD simulations and bioinformatics. We briefly describe how provenance of many-task scientific computations are specified and coordinated by current workflow systems on large clusters and clouds. We discuss challenges in gathering, storing and querying provenance as structured data enriched with information about the runtime behavior of computational tasks in high performance computing environments. We also show how provenance can enable runtime and useful queries to correlate computational resource usage, scientific parameters, and data set derivation.

- *Zenith: a Hybrid P2P/Cloud for Big Scientific Data* by Esther Pacitti and Patrick Valduriez

Big data (i.e. data sets that grow so large that they become awkward to deal with) are creating major problems for computational science in terms of data capture, storage, search, sharing, analytics and visualizing. But size is only one dimension of the problems. Other dimensions include velocity, complexity and heterogeneity.

To address these problems in the Zenith team, we adopt a hybrid P2P/cloud architecture. P2P naturally supports the collaborative nature of scientific applications, with autonomy and decentralized control. Peers can be the participants or organizations involved in collaboration and may share data and applications while keeping full control over some of their data (a major requirement for our application partners). But for very-large scale data analysis or very large workflow activities, cloud computing is appropriate as it can provide virtually infinite computing, storage and networking resources. Such hybrid architecture also enables the clean integration of the users own computational resources with different clouds.

- *Chiron: An Algebraic-Based Parallel Scientific Workflow Engine* by Eduardo Ogasawara

Large-scale scientific experiments based on computer simulations are typically modeled as scientific workflows, which facilitate the sequencing of different programs. These scientific workflows are defined, executed and monitored by Scientific Workflow Management Systems (SWfMS). As these experiments manage large amounts of data. It is interesting, if not essential, to execute them in High Performance Computing (HPC) environments, such as clusters, grids and clouds. However, few SWfMS provide parallel support and they usually lack a run-time provenance support mechanism. Also, the existing parallel SWfMS have limited primitives to optimize workflow execution. To address these issues, we developed an algebraic approach to specify the scientific workflow, and it led to the implementation of Chiron: An Algebraic-Based parallel scientific workflow engine. Chiron has a native distributed provenance gathering mechanism and can perform algebraic transformations to obtain better execution strategies. Chiron is efficient in executing scientific workflows, with the benefits of bringing room for declarative specification and run-time provenance support.

## 6. SPONSORS

- INRIA and French Embassy at Brazil
- LNCC and CNPq