Presentation of the funded projects in 2010 for the « Cosinus » Programme

ACRONYM and project title — page

« Cosinus » programme

YEAR 2010

Project title	FOSTER – Spatio-temporal data mining : application to soil erosion monitoring

Abstract

Recently, the environment, sustainable development and natural risk management have become a major challenge. New Caledonia probably represents a first rank environmental observatory. Thus, the bulk of the great barrier reef of New Caledonia has been declared a World Heritage by the UNESCO. Protecting and managing such a fragile environment owing to major mining projects and an increasing population are major challenges. Human activities enhanced erosion of naturally fragile areas (thick tropical regoliths), resulting in dangerous processes for persons, goods and resources. On another hand, an increasing landslide frequency in the southern part of French Alps has been correlated with climate change and geodynamic forcing (earthquakes). Additionally, wrong land use has reinforced the process. These examples underline the need for a global approach of environmental managing, especially as far erosion is concerned. These last years, the increasing amount of geosciences data have lead to new promising applications. For example, the use of very high resolution satellite images now enables the observation of small objects. However, actual data analysis techniques suffer from the huge amount of complex data to process. Indeed, this environmental data is often heterogeneous, multi-scale, incomplete, and composed of complex objects. In this context, this project aims at providing to geologists a semi-automatic and complete process for monitoring soil erosion. This process will be based on multi-temporal very high resolutions satellite images coupled with digital elevation model (DEM), sensor data and/or expert knowledge. The project will focus more precisely on two important aspects of this process: segmentation of satellite images based on collaborative methods, and construction of descriptive (patterns, clustering, …) and predictive (decision trees, …) spatio-temporal models. New methods, algorithms and softwares on environmental data will be proposed to assist experts in their knowledge discovery. This project is composed of academic laboratories and one company. This multidisciplinary consortium is composed of 1) computer scientists of the LIRIS, LSIIT, LISTIC and PPME laboratories, specialists in data mining and image processing, 2) geologists of the PPME and IPGS laboratories, with a strong knowledge in soil erosion mechanisms and 3) the

	Bluecham Company, specialist in decision support systems. Skills of each partner cover most of the KDD (Knowledge Discovery in Databases) process, from selection and preprocessing to geologic knowledge discovery. Moreover, our industrial partner will transfer our works to the society. Thus, thanks to this consortium, we can propose a project gathering different approaches in the same process, which should lead to significant results.
Partners	PPME LIRIS LSIIT LISTIC BlueCham
Coordinator	Nazha SELMAOUI - PPME nazha.selmaoui@univ-nc.nc
ANR funding	921034€
Starting date and duration Cluster label	January 2011 - 36 months

« Cosinus » programme

YEAR 2010

Project title	HAMM – Hybrid Architectures and Multiscale Methods
Abstract	This research project aims at the development, analysis and software implementation of mathematical models for multiscale applications on hybrid architectures. Large scale multiscale applications are indeed within reach with the emerging computing infrastructures, but they require accurate and robust multiscale numerical methods that take into account these new architectures. This is very challenging as current software were not designed for these methods and architectures. As a matter of fact, scientists and engineers need to manage (i) the complexity of the underlying multiscale models, usually expressed in terms of a partial differential equation (PDE) system completed with algebraic closure laws, (ii) the complexity of numerical methods used to solve the PDE systems, and finally (iii) the complexity of the low level computer science services required to have efficient software on modern hybrid hardware (e.g. multicore CPU/GPU). Robust and effective multiscale methods as well as advanced programming techniques need to be combined to fully benefit from massively hybrid parallel architectures. Even though this project revolves around applied mathematics and multiscale numerical methods, it is truly multidisciplinary and critically requires the expertise of specialists of the applications, computer scientists and computer designers. That's why we create a strong consortium : (i) CEA and IFP leading the large scale multiscale application specifications, (ii) UJF leading the multiscale numerical methods development and efficient implementation on hybrid architectures and finally (iii) BULL leading the benchmarking activities and ultimately the sizing of future super computing infrastructure according to nowadays constraints such as energy costs savings.
Partners	UJF IFP Bull CEA

Coordinator	Christophe PRUD'HOMME - UJF christophe.prudhomme@ujf-grenoble.fr
ANR funding	876124€
Starting date and duration	October 2010 - 48 months
Cluster label	

« Cosinus » programme

YEAR 2010

Project title	HORUS – Urban and inter-cities transportation network timetable optimization
Abstract	HORUS project Consortium composed of the PRiSM laboratory (CaRO team), EURODECISION and the CReSTIC laboratory (SysCom team). Building optimized transportation services (urban, peri-urban and interurban) is a major issue for public and private operators. The great complexity of this task usually justifies its decomposition into sequential steps : timetabling of the bus, tramways or trains (vehicle scheduling), assignment of personnel services to the vehicles (crew scheduling) and building individual shifts. To achieve this task, only the major companies have access to decision support systems including optimization modules. In the last decade, EURODECISION integrated its crew scheduling software component, LP-EasyDriver, to the activities of RATP and, more recently, VEOLIA. The optimization methods of this tool, as well as similar ones from competitors, are based on column generation techniques and hybrid approaches. By expanding their functional scope, these methods may also be applied to the railway sector. The main goal of the HORUS project is to prepare the next generation transportation services by overcoming many challenges : - extension to the railway sector (higher combinatorial and business complexities), - design of global optimization of the vehicle and crew scheduling problems, - access to multi-core computing through algorithms parallelization, - access to the optimization algorithms through free software and shared Web services. From an academic point of view, PRiSM and CReSTIC, two university laboratories well known for their expertise in combinatorial optimization (advanced column generation and metaheuristic methods) and parallel computing, are proposing significant contributions in the sequential and parallel resolution methods dedicated to these problems : arcs-states modeling, column generation, innovative metaheuristics (Trust Branching Path) and hybridization techniques including metaheuristics and exact approaches. From an industrial point of view, EURODECISION is a technical center of expertise on previously mentioned methods and transport business

	services. EURODECISION is then able to capitalize on the proposed innovations by diffusing them directly to the transport operators or through business software editors.
Partners	PRiSM EURODECISION CReSTIC
Coordinator	Alain BUI - PRiSM Alain.Bui@prism.uvsq.fr
ANR funding	595521€
Starting date and duration	October 2010 - 36 months
Cluster label	Advancity (ex Ville et Mobilité Durables)

« Cosinus » programme

YEAR 2010

Project title	MAPPI – New sequential and distributed approaches in algorithmics and computational biology for the analysis of data generated by Next Generation Sequencers.
Abstract	These last few years, a breakthrough in DNA sequencing techniques has been realized through new approaches and new devices under the name of Next Generation Sequencing (NGS). These sequencers are able to produce small raw sequences (called reads) at such a rate that the volume of data to process becomes staggering, reaching such a point that managing these data for mapping and assembling becomes the main bottleneck of the new technology, the fastest actual software not being able to scale to process that volume. Moreover, these new sequencers also enable heterogeneous sources to be sequenced together. This opens the way to meta analysis, that is, considering at a glance genomes or transcriptomes of all a population of living organisms. This is the motivation of our 36 months project, which aims to propose new, relevant and efficient sequential, distributed and parallel algorithms to face the challenge of performing intensive computation for mapping, assembling, metassembling the massive volume of reads produced by the Next Generation Sequencers (NGS). This project gathers four partners. LIAFA (University Paris-Direrot), LIFL (Lille) and IRISA (Rennes) are recognized research groups in computer sciences, whose expertise is complementary on the data and techniques to develop, each being a recognized specialist in at least one of the topics of the project: index data structure, string algorithms, parallel algorithms, biological sequence analysis,\dots These groups will design new algorithms, and develop open source software implementing these algorithms. For these latters not to remain theoretical or only prototypes and for the whole project to have a direct application domain, this project is linked to a huge biological project named Tara Oceans through the Genoscope laboratory (CEA) which is the fourth partner of the MAPPI project. Genoscope is specialised in sequencing techniques and disposes of the latest sequencing devices. Tara Oceans is a unique

	multidisciplinary project gathering oceanographers, ecologists, biologists and physicists expert in marine life whose goal is to exhaustively study phytoplankton in several oceans. The role of the Genoscope is to sequence DNA and RNA samples of protits collected at each spot of the expedition from phytoplankton. Tara will provide metagenomic data (DNA from the cells of a whole sample), and metatranscriptomic data (RNA that captures gene expression). The total number of such samples will range between 2000 and 4000 for the whole project, which should induce more than 100TB of data. The software we plan to deliver will be integrated in a bioinformatics pipeline in the Genoscope.
Partners	LIAFA LIFL INRIA Rennes - Bretagne Atlantique Génoscope
Coordinator	Mathieu RAFFINOT - LIAFA raffinot@liafa.jussieu.fr
ANR funding	456830€
Starting date and duration **Cluster label**	October 2010 - 36 months

« Cosinus » programme

YEAR 2010

Project title	NEWCASTLE – Wavelets and order N for very large scale calculations of Electronic structure
Abstract	Ab initio methods are well used in the fields of solid state physics and quantum chemistry because they are free of parameter. They solve the Schrödinger equation in the framework of the density Functional Theory. Traditionally, these methods use gaussian or plane wave basis sets which limit the scalability. Moreover the CPU time varies as a power of 3 of the number of atoms. With the software BigDFT, we have shown that it is feasible to use real space basis sets as wavelets with the same accuracy as traditional basis sets. Due to adaptive properties, the memory consumption can be reduced drastically compared to plane wave basis set. BigDFT has also a better scalability. We have demonstrated an efficiency over 85% up to 1000 cores. Moreover, order N method which scale linearly with the number of atoms can be used with real space basis sets. The first goal of this project is to develop an ab initio method based on wavelet which can exploit petascaling supercomputers by means of order N methods and a better memory consumption using softer pseudopotentials. Hundred of thousands cores could be used to solve multiscale problems as involved in biology or new devices as OLED. Moreover all developed components could be re-used in other softwares (ABINIT, CP2K, Octopus, ...) as it is already done with present versions of BigDFT (Poisson solver and exact exchange calculations). The second goal is to have a robust and reliable Message Passing Interface library based on OpenMPI with fault tolerance. In order to deliver a test suite based on BigDFT for the future petascaling supercomputer (PRACE project), we had tested the supercomputers of national centers (CCRT and CINES) up to 600 cores. We found that the collective communications had not an efficient scalability. Moreover, all MPI libraries had some trouble as the number of cores increases which were not very reproducible. We think that it is very important to have a test suite to optimize algorithms for collective communications versus the number of cores and that fault tolerance is a very crucial concept with the arrival of petascaling computer. With this project, we would

	have a reliable ab initio methods adapted for large scale simulation, a robust message passing interface with fault tolerance for massively parallel computers and a test suite to optimize algorithms of collective communications. All developments will be available and re-usable for other softwares.
Partners	INAC/SP2M ESRF CEA-DIF-DPTA BULL
Coordinator	Thierry DEUTSCH - INAC/SP2M thierry.deutsch@cea.fr
ANR funding	660088€
Starting date and duration	January 2011 - 36 months
Cluster label	

« Cosinus » programme

YEAR 2010

Project title	OPARUS – Optimization and Parallelism for Analysis and Reconstruction of ultrasonic NDT
Abstract	The project OPARUS "Optimization and Parallelization for Analysis and Reconstruction of ultrasonic NDT" focuses on the development of modelling tools and treatment for reconstruction of multi-element ultrasonic non destructive testing data. The objective of the project is to optimize the performance of numerical algorithms and to extend significantly the actual limitations on the use of accurate models for visualization, analysis and diagnostic and provide tools for innovative modelling in agreement with the industrial constraints. The kinds of technology pointed are clearly oriented on multi-core processors and GPGPU, all included in single "workstation". The technological barriers concerns both the capacity to exploit the new parallel architectures and to implement optimized algorithms dedicated to the reconstruction of data for analysis. The developments are focused on applications representative of industrial cases. The reconstruction tools based on these developments will be integrated into three systems and software: the M2M acquisition systems, CIVA platform of simulation and analysis and the analysis production software NDT-kit. OPARUS is a 3 years industrial research project which includes industrial end users of NDT (EADS and EDF), high-tech companies, (M2M, which develops multi-elements systems and CAPS-Enterprise, expert in parallel computing solutions), CEA LIST developing the CIVA software platform, and EFI, university laboratories specialist of parallel architectures, image processing and complex systems. Project results will be directly integrated in systems and software already operating in industrial environments. From a scientific point of view results will be communicated through journals and conferences dedicated to non destructive testing, and in journals for research in computer science and architecture.
Partners	CEA LIST EADS EDF M2M

	IEF CAPS
Coordinator	Stéphane Le Berre - CEA LIST stephane.leberre@cea.fr
ANR funding	769331€
Starting date and duration	October 2010 - 36 months
Cluster label	SYSTEM@TIC Paris région

« Cosinus » programme

YEAR 2010

Project title	OPTIDIS – Optimzation of dislocation dynamics code
Abstract	Plastic deformation is mainly accommodated by dislocations glide in the case of crystalline materials. The behaviour of a single dislocation segment is perfectly understood since 1960 and analytical formulations are available in the literature. However, to understand the behaviour of a large population of dislocations (inducing complex dislocations interactions) and its effect on plastic deformation, massive numerical simulations are necessary. Since 1990, simulation codes have been developed by French researchers. Among these codes, the code TRIDIS developed in the laboratory SIMAP in Grenoble is the very first 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 numerical methods and that should be able 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 introduced when the code NUMODIS is mature enough to implement parallel routines. The objective of the project in to develop and validate the algorithms in order to optimise the numerical efficiency and performance of the code NUMODIS. We are aiming at developing a code able to tackle realistic materials problems such as the case of interaction between dislocations and irradiation defects in a grain plastically deformation after irradiation. These kinds of studies where "local mechanisms" are correlated to macroscopic behaviour is a key issue for nuclear industry in order to understand materials ageing under irradiation, and hence predict power plant service life. To carry out such studies, massive numerical optimisations of NUMODIS that involve complex algorithms, as well competences in applied mathematics and parallel computing methods are critically needed. The project OPTIDIS involves a team collaboration between researchers specialized on one hand in dynamics dislocations and on the other hand in numerical methods.

	This project is divided in 8 tasks over 4 years. Two PhD thesis will be directly financed by the project. One will be dedicated to numerical development, validation of complex algorithms and confrontation 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 development made in OPTIDIS. In both cases, these simulations will be compared with experimental data obtained by experimentalists.
Partners	CEA/DEN/DMN/SRMA SIMaP / Grenoble INP ICMPE / Paris-Est HIEPACS / INRIA
Coordinator	Laurent DUPUY - CEA/DEN/DMN/SRMA laurent.dupuy@cea.fr
ANR funding	365232€
Starting date and duration	October 2010 - 48 months
Cluster label	SYSTEM@TIC Paris région

« Cosinus » programme

YEAR 2010

Project title	PETALH – Preconditioning scientific applications on pETascALe Heterogeneuous machines
Abstract	The modern numerical simulations coupled with ever growing and more powerful computational platforms have been a major driving force behind a progress in numerous areas as different as fundamental science, technical/technological applications, life sciences. These simulations often involve a chain of linear algebra problems as solving a large sparse linear systems of equations or finding a set of eigenvalues and eigenvectors. Solving these problems is very time and memory consuming and often represent an important fraction of the total simulation time. This project focuses on solving large sparse linear systems of equations using iterative methods. More specifically it focuses on parallel preconditioners that are suitable for the emergent hierarchical models of clusters of multi-core processors, as used for example in petascale machines. These preconditioners are important to accelerate the iterative methods, and they represent the main topic of research in this context. This is due to the fact that there is no general preconditioner that works for all applications. Very often the best preconditioner is the one which is adapted to the application to be solved. This project is a continuation of PETAL project that was funded by ANR Cosinus 2008 call. PETAL was submitted as a 3 years project, and was awarded 2 years. Early results we have obtained in the first year are very promising. They have proven the soundness of our approach. This proposal continues this line of research. In particular, PETAL has allowed us to overcome some of the important challenges related to the generalization and suitability for parallel computing of two important preconditioning techniques: filtering decomposition and nested decomposition. The first preconditoner is an incomplete factorization where it is possible to ensure that the factorization will coincide with the original matrix for some specified vector, called a filtering vector. Satisfying this filtering condition is an important factor for accelerating the convergence of the iterative method. The nested based preconditioner has the same property for a specific vector of all ones. However the

construction is different and takes advantage of a nested structure of the input matrix. The previous research on these methods considered only matrices arising from the discretization of PDEs on structured grids, where the matrix has a block tridiagonal structure. To the best of our knowledge, there was no previous result on the parallelization of filtering preconditioners and there was no successful result on the parallelization of nested preconditioners. The nested preconditioner is in particular very used in the oil industry, where it is known to be very efficient. One of the important results of PETAL is the development of a new and general approach to ensure that a filtering condition is satisfied in a matrix decomposition. This approach is based on an innovative way of organizing the computations that allows on one side to satisfy a filtering property and on another side to perform a parallel computation. This approach has been used to develop two preconditioners. The first preconditioner is based on a block approached decomposition, and we refer to as block filtering preconditioner. The second preconditioner is based on a nested approach, that we refer to as nested filtering preconditioner. The goal of PETALh is to continue this research with a specific focus on parallel preconditioners suitable to be used on the emergent hierarchical models of clusters of multi-core processors. Our project also includes a study of their suitability for heterogeneous machines formed by multicore processors and GPUs. This is beyond the goal of the initial PETAL project in which we focused only on clusters of multicore processors.

Partners	INRIA Saclay - GRAND-LARGE UPMC - CNRS UMR7598 INRIA Bordeaux Sud-Ouest IFP CEA
Coordinator	Laura GRIGORI - INRIA Saclay - GRAND-LARGE Laura.Grigori@inria.fr
ANR funding	388769€
Starting date and duration	January 2011 - 24 months
Cluster label	

« Cosinus » programme

YEAR 2010

Project title	REALISTIC – laRge scalE And Long TIme Computations
Abstract	The REALisTIC project objective is to break down the theoretical and technological lock of large scale and long time processes simulations problems. Today, the intensive computing of industrial applications referring to heating and quenching installations is by far out of realism in term of computing time. The average equipments dimensions (tens of meters) and the physical time (tens of hours) considered in such installations and therefore in their numerical simulations counterparts do not allow industrials end-users to access this technology. The numerical simulation is an expected solution to help industrial accessing to optimize processes. These optimizations are as well an economical issue, a key approach for gaining a competitive advantage and also in the mid term a real mean to access energy saving and CO2 production optimization. The worldwide tendency is to provide industry with hardware platforms hosting new generation simulation software, most of the time using the latest generation parallelism algorithms. The REALisTIC project intends to contribute to the development of efficient software delivered by french laboratories and french software editors. REALisTIC project gathers a renowned research laboratory, specialist of numerical simulation, 2 software editors companies and 4 key industrial companies operating in the steel-making business,into a partnership well balanced and demonstrated a real implication. The project intends to analyze and propose alternative numerical algorithms capable to reach the « few days » target for simulation time of a classical heating or quenching installation, coupled with the metallurgical evolution of the parts. The work program of the project is divided into 4 main work-packages. The focus will be set on self adaptive meshing techniques and anisotropic time steps approaches (WP2) . These developments will be incorporated into candidate simulation software(ThosT and Forge) (WP3) to deliver demonstrator releases to industrial end-users. The end-user will use these demonstrator softwares to perform a validation action and quantify the added-value for their business (WP4). The

	project is attached to both the "Intensive Computing" and "Conception & optimisation" thematics axis. The theoretical work will be a contribution to researchers to reach really more efficient computing time, the adaptation to real application will promote the use of simulation as a cost-effective solution to improve processes and gain competitivity.
Partners	Sciences Computers Consultants Aubert Duval Industeel-ArcelorMittal AREVA Creusot Forge Snecma ARMINES Centre CEMEF de l'Ecole des Mines de Paris TRANSVALOR S.A.
Coordinator	Chantal DAVID - Sciences Computers Consultants chdavid@scconsultants.com
ANR funding	749688€
Starting date and duration	December 2010 - 48 months
Cluster label	PEGASE Pôle Nucléaire de Bourgogne SYSTEM@TIC Paris région VIAMECA

« Cosinus » programme

YEAR 2010

Project title	SIM-DREAM – A new paradigm in numerical simulation – Separated variables decomposition for a priori model reduction
Abstract	Many problems in science and engineering remain today intractable, despite the impressive progresses in computer science and the computational resources today available, because their numerical complexity is simply unimaginable. Among the models that remain today intractable, we can distinguish two main families: - The first family of models consists of engineering models usually encountered in computational mechanics, defined in large and complex 3D geometries, involving many multiphysics couplings, many scales (in space and time), strong non-linearities, and whose transient simulation needs extremely small time steps. Despite the maturity of such models, as soon as the loading becomes complex (e.g. cyclic loadings, uncertainties …), the material involves uncertainties, heterogeneities in the microscopic scale, multiphysics couplings, or as soon as one focuses in parametric analysis … the model becomes simply unsolvable. - The other family of challenging models concerns those models defined in highly dimensional spaces. For example, this kind of models appears naturally in the modeling of the structure and properties of materials at the finest scales. These models exhibit the redoubtable curse of dimensionality when usual mesh-based discretization techniques are applied. For alleviating the difficulties related to the first family of models, many authors considered the use of proper orthogonal decomposition (POD) based techniques. However these techniques are far to be optimal, because the reduced basis are only optimal when constructed "a posteriori", being only approximated when constructed "a priori". Obviously, the optimal alternative lies in the simultaneous construction of both the solution and the associated reduced basis for expressing such a solution. The partners of this project (P. Ladeveze, F. Chinesta, A. Ammar and A. Nouy) proposed some years ago a novel efficient technique able to circumvent the challenging issues just described. They called this technique: Proper Generalized Decomposition (PGD). It is based on a separated representation of the

unknown solution. We would like to recall that using these approaches, we have reduced the computing time related to the solutions of problems belonging to the first family of models (defined above) in several orders of magnitude (millions in some cases) and on the other hand we solved successfully highly multidimensional models never until now solved because they were considered suffering of the irremediable curse of dimensionality. Because the novelty and the youngness of the PGD method, many aspects have not been addressed. The aim of this proposal is to push back the limits of this method. Definitively, these developments could lead to a real change of paradigm in computational mechanics. Imagine the possibility of solving efficiently any parameterized (including parameterized geometries), multi-scale and multi-physics model. Inverse identification and optimization would be a direct post-treatment. This proposal groups the "inventors" of the PGD method, all of them being highly recognized within the international community of computational mechanics. The partners are developing actively and independently these methods for many years. This proposal aims at grouping together researchers from the different partners in order to facilitate and accelerate the developments of new ideas within the context of PGD methods. This project could have a significant impact from both fundamental and applicative points of view. The solution of the complex systems encountered in high-tech applications (aeronautic and space among many others) are waiting for new proposals able to solve challenging models without using impressive computer resources. Alternatives other than the increase of computational resources and the speed-up of standard techniques merit to be considered seriously.

Partners	GEM LMT LAMCOS
Coordinator	Francisco CHINESTA - GEM francisco.chinesta@ec-nantes.fr
ANR funding	601827€
Starting date and duration	October 2010 - 36 months
Cluster label	EMC2 (Ensembles métalliques et composites complexes)

« Cosinus » programme

YEAR 2010

Project title	SIMINOLE – Large-scale simulation-based probabilistic inference, optimization, and discriminative learning with applications in experimental physics
Abstract	Simulation lies at the heart of most of today's large scale experiments. Since the appearance of heavy computational machinery, simulation has become the third pillar of scientific discovery beside experimentation and theoretical model building. Its most important role is to connect models at different levels of resolution. Simulation can complement or, in certain cases, replace expensive experimentation; it can be used to validate high-level models using low-level experimental data; it can serve as an engineering aid for designing tools, machines, or detectors. On the other hand, simulation has also become the bottleneck of these applications so a lot of research has been devoted to find how to carry out simulations more efficiently. Most of the time these studies follow one of two approaches: they either delve into the inner workings of the simulator and try to improve it algorithmically, or they attack the problem by implementing simulators on various high-end computing devices. In this project we follow a third approach: we propose to use simulators more efficiently by considering them as a black box, and minimizing the number of calls to the simulator for accomplishing certain tasks. Simulators can be used in different ways for solving particular problems. In this project we identified three common scenarios. In probabilistic inference, the goal is to find values for some input parameters that generate simulations similar to observed data. Our goal is to formalize a data-driven simulation setup, and to replace the sub-optimal naive exhaustive search by an approach based on Monte-Carlo Markov chain (MCMC) techniques. In the second scenario, simulation is used in an optimization loop. When designing complex instruments, tools, or machines, it is a common situation that the simulated instrument is assigned a utility (or cost), and the goal of the procedure is to find regions of the parameter space where the utility is high (or the cost is low). As in the previous scenario, exhaustive search is highly sub-optimal. In this task of the project we

will formalize the problem as utility-driven simulation in a stochastic optimization setup, and apply powerful adaptive techniques developed recently for optimizing expensive black-box functions. In the third scenario, a large set of simulations is used to "discover" interesting features, for example, features that predict well certain generating parameters. These "observables" are then used on real data to estimate or reconstruct generative parameters. The goal of this task is to optimize the use of simulations by replacing the "manual" discovery of observables using machine learning algorithms. The research outlined above is directly motivated by the design and inference problems we are facing in two major astroparticle physics experiments, the Pierre Auger experiment and the JEM-EUSO experiment. Their goal is the same: to study the properties of ultra-high energy cosmic ray particles by observing the particle cascade generated by the collision of the cosmic ray particle and atmospheric particles. The Auger experiment employs two independent terrestrial detectors covering 3000 square kilometers on the Argentinian pampas, whereas the JEM-EUSO telescope will be on orbit on the Japanese Experiment Module of the International Space Station starting in 2015. The techniques outlined in the previous paragraph will be directly usable for the statistical data analysis in both experiments and for the design of the on-board software of the JEM-EUSO experiment. The methodological development is motivated directly by these two concrete applications, but the proposed techniques will be generally usable in other simulation-heavy application domains.

Partners	LAL LTCI INRIA Saclay - Île-de-France - Equipe-Projet TA0
Coordinator	Balázs Kégl - LAL balazs.kegl@gmail.com
ANR funding	1042903€
Starting date and duration	October 2010 - 48 months
Cluster label	Cap Digital Paris-Région

« Cosinus » programme

YEAR 2010

Project title	SIMUDMRI – Simulation of diffusion MRI signals in biological tissue
Abstract	Water diffusion magnetic resonance imaging (DMRI) is a method which uses a combination of applied magnetic fields to measure, statistically, the diffusion of water molecules due to Brownian motion. Its spatial resolution is on the order of millimeters. An apparent diffusion coefficient (ADC) or diffusion tensor (DT) is computed for each voxel based on a model relating these quantities to signal attenuation. In the past three decades, DMRI has been successfully used to track brain white matter fibers and to detect acute brain ischemia. Diffusion functional MRI has also begun to gain momentum as an active area of research. The goal of this project is to go beyond the ADC or the DT to get more detailed information on tissue properties from the DMRI signal. The cellular structure inside the human brain varies on the scale of micrometers, which is much smaller than the size of a voxel. There may be thousands of irregularly-shaped cells within a voxel, and they all contribute to the environment seen by water molecules whose displacement is measured by the MRI scanner. In the typical DMRI experiment, the time interval over which water diffusion is measured is in the range of 50-100 milliseconds. Using the diffusion coefficient of 'free' water at 37 degrees Celsius, D = 3e-9 m2/s, we get an estimated diffusion distance of 15-25 micrometers. Clearly, in a DMRI experiment, water molecules encounter numerous times inhomogeneities in the cellular environment, such as cell membranes, fibers, and macromolecules. We simulate the DMRI signal at the scale of a single voxel, while taking into account realistic cellular structure and the true shape and duration of the diffusion gradients. We will obtain the geometrical description of several samples of brain tissue from 3D electron microscopy. The membranes of the various cells making up each sample will then be extracted using dedicated segmentation tools. The finite duration and ramp time of the gradient coils will be accurately reflected in the simulation. The numerical simulation will approached in two directions. One is based on the numerical solution of the Bloch-Torrey partial differential equation using Green's

functions. The other is Monte Carlo simulation. The end result is a hybrid code incorporating both methods: the Green's function method will be used to speed up the Monte Carlo simulation in regimes where the results of the simulation satisfies the Bloch-Torrey equation and where the geometry can be modelled by simple objects such as smooth surfaces. The two methods will be coupled in a unified numerical code where the coupling will occur across spatial regions and also in time. The end goal of this project is the simulation of DMRI signal taking into account realistic 3D cellular structure, gradient sequence, while producing results which are reliably accurate, scalable, all in reasonable simulation time. This simulation tool will be valuable in understanding the tissue microstructure that gives rise to the DMRI signals and also aid in the design of new diffusion imaging protocols.

Partners	INRIA Saclay - Ile-de-France Equipe-Projet DEFI CEA I2BM NeuroSpin
Coordinator	Jing-Rebecca Li - INRIA Saclay - Ile-de-France Equipe-Projet DEFI jingrebecca.li@inria.fr
ANR funding	400036€
Starting date and duration **Cluster label**	November 2010 - 36 months

« Cosinus » programme

YEAR 2010

Project title	SKIPPI – System for Kansei based Innovation and Product Process Integration (SKIPPI)
Abstract	Recent issues in Design Science tend toward the digitization of the earliest phases of the design process. These changes answer to an industrial need: the reduction of development delays while increasing the variability of the offer. This evolution leads to a concurrent formalization and an early digitization of the design information. This digitization is based on the formalization of the implicit cognitive abilities and skills of the designers, and then on the definition of design rules aiming at support and enhance some routine parts of the activity. These rules are conveyed by adequate algorithms in order to build new computer aided design tools. At this level of the design process, the data are by nature vague and ill defined. The algorithms used are often based on traditional statistical approaches, even if specific tools coming from the field of artificial intelligence proved to be more efficient for this kind of data. Some scientific approaches cope with this ambition already exist, as Kansei engineering. However these methods are still partial and do not cover the whole design process encompassing the product and the process. This causes a rupture in the digital chain, giving birth to two distinct systems: an upstream system for global design, and a downstream system for detailed design and developement. This rupture impacts the coherency of the product. Moreover it is often intensified by a de-contextualization of these approches from the project, from the lauching phase and from the process. Indeed, the brand identity is rarely taken into account, in spite of the high strategical impact of the brand values on the design orientations. There are also no specific relations between product and process specifications in these approaches. The objective of SKIPPI project is the development of a software prototype for supporting decision while designing and processing the brand identity thanks to the use of specific dimensions as semantics, sensoriality, styling in relation with product and process parameters. This system will be based on three technologies: Kansei engineering, artificial intelligence and integrated product process engineering. It will enable to deal with fuzzy information on the base of heterogeneous data, going from brand values to product and process parameters or conversely from product and process parameters to brand values on the base of mainly lexical information. This system will help the designers for interpreting a brief before to generate design solutions. It will enable to simulate and argue the proposal of new design candidates according to the values of the brand. Conversely, engineers,

	designers and marketing managers will use the system to evaluate the adequacy of a new material, or of new processes, or new design solutions according to the values of the brand. The system will be manily based on lexical data. So it will provide an open field to the designers for visual synthesis and creativity.
Partners	Arts et Métiers ParisTech LIP6 - UPMC GRENOBLE INP G-SCOP ILOBJECTS PSYCLE DIEDRE DESIGN SAGEM WIRELESS
Coordinator	Améziane AOUSSAT - Arts et Métiers ParisTech ameziane.aoussat@paris.ensam.fr
ANR funding	1209115€
Starting date and duration	December 2010 – 36 months
Cluster label	Cap Digital Paris-Région

« Cosinus » programme

YEAR 2010

Project title	SOHUSIM – Soft Human Simulation
Abstract	This project deals with the problem of the modeling and the simulation of soft interaction between humans and objects. At the moment there is no software capable of modeling the physical behavior of human's soft tissues (muscles, fat, skin) in mechanical interaction with the environment. The existing software such as LifeMod or OpenSim, models muscles as links of variable length and applying a force to an articulated stiff skeleton. The management of soft tissues is not taken into account and does not constitute the main objective of this software. A first axis of this project aims at the simple modeling and the simulation of a passive human manipulated by a mecatronics device with for objective the study and the systems design of patient's manipulation with very low mobility (clinic bed). The second axis concentrates on the detailed modeling and the simulation of the interaction of an active lower limb with objects like orthesis, exoskeleton, clothes or shoes. The objective being there also to obtain a tool for design of devices in permanent contact with the human who allows determining the adequate ergonomics in terms of forms, location, materials, according to the aimed use. Problems in this project are situated at the level of the modeling of "soft materials» in interaction with diverse objects. The "soft materials» which we intend to model can be artificial materials such as the silicone or natural one such as latex. These materials can constitute homogeneous objects (ex: full rubber ball) or heterogeneous (ex: foam of silicone containing micro air bubble), we can envisage the modeling of inflatable structures. Besides the inert soft materials, we also envisage and especially the modeling of the human at the level of his alive organic tissues (muscle, fat). This level of modeling will require a coupling with data from medical images of MRI type to be the most similar possible with the considered human. The modeling of muscles will have to consider the active character of this organic tissue. These soft organic structures will thus have variable properties for example in terms of shape or stiffness. The modeling of the deformations and the phenomena of friction will be an important aspect of the project. The appropriate function of muscles being the mobility, the interaction with the skeleton and its consequences at the posture level and the movements of the modeled human must be considered by the implementation of biomechanical models. This modeling / simulation of soft tissues will constitute a real step ahead in the field of the simulation of the humans. This will allow, to study the behavior of soft tissues in interaction with the environment on aspects patient's manipulation, to simulate the effects of prolonged pressure forces, to

	predetermine the zones of excessive frictions, formation of bedsores and hurts linked to the use of objects. The dynamic aspects of its simulations will allow studying the energies involvement during phases of movements of the current life. Tools developed in this project will have vocation to be integrated, in the form of module, into software of computer-aided design (CAD).
Partners	INRIA Grenoble - EVASION / LJK FATRONIK FRANCE INRIA / Université Sciences et Techniques du Languedoc Montpellier I HPC Project CHU Montpellier - Lapeyronie
Coordinator	Francois FAURE - INRIA Grenoble - EVASION
ANR funding	569731€
Starting date and duration **Cluster label**	October 2010 – 48 months

« Cosinus » programme

YEAR 2010

Project title	SPUTNIK – Simulating experiments on protein structure and dynamics
Abstract	The goal of this project is to bridge the gap between simulations and experiments in the study of biological macromolecules, in particular proteins. It brings together experts from the fields of molecular simulation, neutron scattering, nuclear magnetic resonance spectroscopy, and scientific computing, which are the most important techniques for exploring the structure and dynamics of biomolecules. Simulation techniques are already an important tool in the interpretation of many biophysical experiments. However, the state of the art in molecular simulations is to simulate a single protein in a small quantity of solvent. Such a minimal system does not represent the reality of a sample used in experiments, which consists of either a protein crystal or a dilute solution of proteins, plus impurities and other imperfections. The typical simulation system does not capture important characteristics of an experimental setup either; unwanted but inevitable effects such as multiple scattering and absorption (in neutron scattering) or the destruction of a crystal by radiation (in crystallography) are not taken into account. This creates discrepancies between simulation and experiment that diminish the utility of simulations to help with the interpretation of experimental data. An important goal of the proposed project is to develop "virtual experiments" in the sense of including the abovementioned unwanted effects in the simulation. In this way experimental setups can be designed avoiding as much as possible the trial and error approach and preliminary studies can be performed "in silico". The approach is also indispensible to interpret "real world" experimental results and to make a connection with theoretical models. To achieve the goals described above the partners of this project can rely on their mixed competences in modeling and simulating biomolecular systems. Three of the four partners have already successfully collaborated in this context within the ANR project THALER (ANR-06-CIS6-012-01). Two other important issues of the project are related to molecular simulation of large biomolecular systems. They

concern the time-consuming calculation of electrostatic interactions in these systems on massively parallel computers and the efficient handling of input/output (I/O) operations. Since the advent of vector computers in the field of scientific computing in the 1980s, I/O has been a concern since it is a major bottleneck for the development of fast computer programs for the simulation of many-particle systems. For both the calculation of electrostatic interactions and the solution of the parallel I/O problem the partner at the Research Centre Jülich (FZJ) has developed solutions which are a good starting point for extensions of the simulation engine developed in the THALER project. The CBM and the FZJ group have long-standing contacts and collaborated in the past essentially on the dynamics of liquids. The present project will be an opportunity to take advantage of the excellent environment of the FZJ group concerning computer equipment and technical expertise.

Partners	CNRS ILL ENS FZJ
Coordinator	Gerald Kneller - CNRS gerald.kneller@cnrs-orleans.fr
ANR funding	703681€
Starting date and duration	January 2011 - 36 months
Cluster label	

« Cosinus » programme

YEAR 2010

Project title	SYNE2ARTI – From Synthetic Gene Networks to Artificial Tissues
Abstract	Synthetic biology, or bioengineering, aims at designing and constructing in vivo biological systems that performs novel, useful tasks. This is achieved by reingeneering existing systems.In contrast to traditional biotechnology, the focus is on developing biological or computational tools that help with an efficient construction of new systems. In this project, we focus in applications of synthetic biology for tissue engineering. Because engineered tissues might supplement or even completely replace defective tissues in patients, their development is therapeutically highly relevant. This requires firstly to reprogram cell growth, differentiation and death, which necessitates extensive modifications of intracellular networks, and secondly, to engineer cell-cell communications so as to coordinate the development of the tissue at the cell population level. This task is delicate since, on one hand, manipulations of intracellular molecular interactions can affect cell-cell and cell-substrate interactions interfering both with tissue architecture and function, and, on the other hand, changes in tissue architecture and function can feed back to the control of gene expression and post-transcriptional modifications. Consequently, the global behavior of engineered tissues emerges from local interactions between extensively modified cells, and this multi-scale aspect needs to be taken into account to successfully engineer such tissues. Unfortunately, because of this multi-scale aspect, the numerical simulation of the (monolithic) models developed to assist the tissue design is not computationally tractable: one has to analyze large differential equation systems (up to 10^6 ODEs) with uncertain parameters defined by probability distributions (random ODEs). Besides large computational resources, such analyses require novel methodological developments for multi-scale numerical simulations. To deal with this problem, we propose to use of abstraction. More precisely, i) we compute an abstract representation of the original random ODE system in the form of a continuous time stochastic automaton, and ii) we use this abstract representation in place of the original ODE

model for the numerical simulation of individual cell based tissue models. The abstract system is defined such that its behaviors are approximately equivalent to the ones of the original system (approximate behavioral equivalence). Three major tasks can be identified. The first one is to develop a theoretical framework to formally define this notion of approximate behavioral equivalence of random processes, and to provide algorithms and tools to compute abstractions of an original random ODE model and assess their quality. The second task is to apply this framework to the computation of an abstract representation of the complex intracellular network of engineered stem cell that are genetically modified to create an insulin-producing tissue. The last task is to develop an individual cell based model of monolayer stem and pancreatic (beta) cell tissues, and to extend it with the abstract representation of the engineered cells computed previously. The resulting model can then be used to predict the tissue development and suggest potential system improvements.

Partners	Contraintes Verimag Bang MIT
Coordinator	Gregory Batt - Contraintes gregory.batt@inria.fr
ANR funding	387370€
Starting date and duration	October 2010 - 36 months
Cluster label	