# EVALUATION REPORT ON THE INRIA "NumB" PROJECTS

The review for INRIA projects organized under the "NumB" heading was conducted March 27-28, 2008, in Paris. The experts forming the review panel were

David Bader, Georgia Tech
Rosa M. Badia, Barcelona Supercomputing Center
Erich Schikuta, University of Vienna
David Keyes, Columbia University
Rami Melhem, University of Pittsburgh
David Nicol (Chairman), University of Illinois at Urbana-Champaign
Alex Pothen, Old Dominion University
Eric Chaput, Airbus
Jean-Pierre Prost, IBM
Stéphane Requena, IFP
Franco Travostino, eBay

*Summary*: Overall, the level of scientific achievement is high. The research appears to align with INRIA's overall scientific strategy; however, the program would be considerably strengthened with more attention paid to coordination among projects, and mechanisms in place to realign research objectives more frequently. Industrial impact and transfer is a somewhat fledgling effort. To achieve greater impact, more attention will have to be paid to promoting and maintaining software product. For this a greater investment in engineering support staff will be needed.

## Table of Contents

## General Remarks

As will be shown in individual project reports, the strength of the program lies in the technical expertise and accomplishments of the teams. In this respect the projects grouped together in "NumB" are all strong. That said, the reviewing experts concur on a number of other points whose consideration could make the program even stronger.

*Cohesion/Coordination*: INRIA has tremendous potential to assert itself as a unit on the world research stage in the same way that a first ranked university does. But while the strengths of a first rate university and INRIA are similarly rooted in the quality of their researchers, a first ranked university is able to project itself as a "whole greater than the sum of the parts", through cooperation and synergy within their research groups. With respect to the review, there is no evident organized coordination among the different teams. Such coordination as exists appears to be accidental by-products of researchers changing projects and/or locations. The review itself would have gone more smoothly if there had been coordination simply in the presentations made. A number of questions the experts asked were predictable and could have been preempted, if only teams had reviewed each other's presentations ahead of time and anticipated questions directed at seeming overlap and duplication of research. For example, several of the projects listed "scheduling theory" as an important activity. After some digging and detailed questioning the review panel was able to see that there were some differences in the research being pursued---but there were also opportunities for synergy that were **not** being pursued. It is interesting to note that essentially the same criticisms were given in the 2005 review team of the Num C projects:

> "*We think anyway that some, if not all projects could learn more from each other; for example, several teams treat problems related to "control", but there is a lack of exchange of expertise between the different groups. It is suggested that a periodic Num C meeting is organized, with all young scientists present, with the goal of exchange of information among the teams.*

Repetitio est mater studiorum; we hope that INRIA takes repetition of the same critique to heart and initiates positive steps towards addressing it.

*Project direction*: The technical issues the various teams approach are scientifically sound, but appear to be largely established at the time the project is proposed, and also perhaps at 4 and 8 year reviews. This long time-scale for adjusting directions creates the potential for research that is not tracking the quickly changing technical landscape as it should. The stated "NumB" focus on cluster computing is an excellent example of this. While cluster computing by no means has disappeared or is in danger of disappearing, the leading edge of research in high performance computing has moved away from that area. We are concerned that whatever is forcing a focus on cluster computing for another four years, that mechanism inhibits teams' ability to address problems that are recognized as "emerging" elsewhere in the computing world. Even if a team recognizes that the cutting edge is moving away from where they are working, it isn't clear whether the organizational and review structure allows them to dynamically change the research direction appropriately. Whatever the cause, the review panel felt that cluster computing may be over-emphasized at the cost of ignoring related areas. New trends of service

orientation and cloud computing (and multi-core architectures that support them) are emerging in the industry; these may not fundamentally rely on different concepts than grids, however, a stronger alignment with these trends may facilitate access to funding sources and technology transfer to industry

Finally in order to increase the footprint of potential academic users of the INRIA R&D and also to give access to INRIA's researchers to large and different (mpp, clusters of thin or larges nodes, vector and hybrid machines) HPC configurations we advise INRIA management to work closer with the newly French agency called GENCI.

*Degree of interdisciplinary work*: A number of the projects aim to produce results and tools for use by experts in other disciplines, and are making some efforts to include experts from other domains. The fact remains though that most team members are mathematically oriented computer scientists. We believe that the inter-disciplinary work would only be strengthened with sufficient participation with domain experts to guarantee that theoretical work is validated through real life applications, especially applications in the strategic fields of bioinformatics, environment, and medicine.

*Technology Transfer*: Solid efforts are being made to effect tech transfer, but some actions by teams and INRIA administration could improve this. The review panel learned that INRIA is working to make engineers more available to the research teams for software development and maintenance. We recommend that this be amplified, if possible. Despite the new efforts, the review panel has general concerns about engineer availability for software development, maintenance, and user support. The past approach seems to have led to a conservative approach towards software advertisement and technology transfer. There is recognized difficulty in finding startups that would acquire and support software developed by the teams (except for rare exceptions); one suggestion is to create a team of permanent engineers across project teams that would develop and maintain software based on theoretical and experimental work from the research teams, and that would provide code maintenance and user support. We found it surprising that teams did not have a good idea of who was using their software products. Better analysis of software use (which is currently limited to tracking the number of downloads) would allow better assessment of the relevance and the reach of software developed. We also recommend fostering the creation of open source development communities around the software the teams develop; this could increase use and exposure, as well as provide a valuable means of providing feedback to the tool developers.

We also discussed the idea of involving industry representatives on the "NumB" evaluation committee from the targeted customers of INRIA work, in order to be aware about industry expectations/needs and to facilitate the technology transfer to end customers.

Also in order to be widely adopted by a large (and not solely French) community we recommend to INRIA management to make an effort for being more active on normalization committees.

*Evaluation Process*: We appreciate the effort and expense INRIA puts into the evaluation process. The staff was helpful, and the venue was conducive to a productive time of presentation and discussion. A number of things might have been done to streamline the process. The review panel felt that before arriving at the review it did not have a good idea of what would be expected of it. Our initial confusions related to the meaning and organization of "NumB" could have been addressed in our early meeting with INRIA administration. In retrospect we would have appreciated a presentation on how the research agenda of "NumB" projects evolve with respect to INRIA's strategic objectives. We also thought it would have been useful to have some time with the INRIA evaluation team during the experts' final meeting. This would have allowed remaining questions about the review process to be answered, and would provide INRIA early (but unstructured) feedback from the review panel.

## Evaluation of Project ALGORILLE


### *Abstract*

The ALGORILLE project, directed by Jens Gustedt (INRIA Nancy), focuses on the transparent and efficient access to distributed resources on the Grid for new types of applications that can harness this power. The novelty of their approach is its emphasis on algorithmic aspects of grid computing, such as the efficient organization of the computation, by both the service provider and the end user. This review covers the first year of the project, 2007, and as such, the ideas and development are still early. We are in favor of continued support of ALGORILLE, as it has the potential to make significant impact in high-performance and grid computing. A research plan often requires many years of sustained investment, and ALGORILLE must be given this opportunity in order to realize its full benefit to society. Its current approaches, as detailed next, are very good in comparison with other international projects at the end of their first year.


### *Scientific Excellence*

ALGORILLE includes three objectives for the evaluation period.

The first objective deals with the study of structuring applications in space and time so that they may be effectively deployed in the grid environment, and thus, scale up to make efficient use of the distributed resources. In this research theme, the researchers propose to impose certain structural restrictions on the algorithms that may be used, reducing stress that an execution imposes on the computing platform. They apply a layered approach to this challenge and perform research in coarse-grained models of computation, relaxed models of synchronization, and efficient distributed algorithms and data structures. As a concrete example, they have taken an important scientific kernel of wave-front computation, and show several innovations. First, they've developed a new, optimized data layout that, with a minor amount of redundancy, performs better than the traditional row-major layout. Second, they've designed better synchronization techniques to chain together waves of the computation when data is written once by a process and read by several others.

The second objective deals with services in a grid and their resource utilization. The aim is to provide transparent resource management through sequential and parallel task scheduling, migration of computations, data exchange, and distribution and redistribution of data. Their approach is to develop resource management algorithms such as one demonstrated called Data Handover (DHO) that can be implemented as grid services and be plugged into middleware frameworks to reconcile the message passing and shared memory paradigms. As an example, they provide a data redistribution service that minimizes the cost of operation using a balanced communications plan. The implementation is experimentally validated on Grid 5000.

The third objective deals with experimentally evaluating the grid infrastructure, keeping in mind important themes such as reproducibility, extendibility and applicability of simulations, emulations and in situ experiments. Here, they focus on the challenge of accurate evaluation and prediction of distributed algorithms using a three-pronged approach with real experiments on Grid 5000, emulations using their Wrekavoc code, and simulation through the co-development of SimGrid. They have developed a scientific instrument that is fast and scalable, and is already used by a number of researchers.

Overall, the research compares favourably to the international level of researchers. The ALGORILLE project team has outlined an aggressive plan of study, and has performed very well in making substantial progress during their first year. They recognize the top peer-reviewed research journals and conferences in this area, and have already published 2 journal and 17 conference papers, 6 of which are in the top tier conferences. This is quite remarkable when one considers that the project has just ended its first year. Clearly, the international community recognizes the innovations produced under this project.

***Adequacy to INRIA's Scientific Strategy***

This project contributes substantially to INRIA scientific strategy by delivering a simulator for large-scale grids. This will enable the realistic study of large-scale applications on the grid, and provide solutions to application teams for selecting algorithms and optimizing applications that can realize the full potential of the large amount of distributed data and computing resources in the national grid. The three objectives form a coherent plan that crystallizes the idea that algorithms, transparent resource management, and experimental evaluations, are foundational basic research and the key enablers to this vision. The project direction of extending the current resource management work to cope with uncertainties at all levels (infrastructure, application, user) is well aligned with INRIA's strategy on ubiquitous computing.

The project has a strong focus on the Grid infrastructure. However, this committed orientation of research goals bears some risks. It is possible that in the future the Grid will loose its dominant position in research of the HPC community and will be replaced by or reincarnate in other arising computing paradigms, such as cloud computing, or in infrastructures such as pure service oriented architectures. In this context, it is obvious that the success of the project also relies to a certain extent on the funding decision of the successor to Grid 5000, the ALADDIN project, by the French government. We expect that the excellent scientific work of the ALGORILLE project will justify their continued support by INRIA. To guarantee impact on the scientific landscape and continued co-operation with and support of industry partners, it is recommended that the above mentioned risk of the "Grid" is acknowledged and some respective actions are foreseen (e.g. by a SWOT analysis), such as the willingness to refocus or extend the research focus of the project to new arising developments in HPC.

***Industrial Transfer & Partnership***

Already in the first year of this project, there has been successful industrial transfer and partnerships. The ALGORILLE project team has worked with domain scientists such as physicists in opto-electronics and fusion, other computer scientists, as well as Supélec for neuromimetic computations. These collaborations resulted in the joint InterCell project that leverages the ALGORILLE parXXL cellular automata framework for implementations. The team has expressed concern that further work is required to successfully transfer usable products to a broader community. In some sense, they have a good solution that is seeking collaborations with problems to solve. Ideally, this would not be the case if they had started with these science problems from day one. The grid scheduling work has had numerous international collaborations with top research groups. Good synergy has been established with other INRIA project-teams working on multi-criteria scheduling, such as GRAAL and MOAIS. Finally, the experimental evaluation software and methodology, such as using SimGrid and Wrekavoc, are already being used in other's research and cited in publications. Early contacts have been established with the industry, e.g. with Bull around SimGrid, EDF around application gridification, and Microsoft around the possible integration of ALGORILLE application layer in Microsoft Cluster Server. Still, additional effort will be necessary to mature the technologies and establish successful industrial technology transfer of this basic research.

***International Level***

The ALGORILLE project team has held the high reputation in the scientific community. Previously, their publication record has been discussed. In addition there is visibility through service on two journal editorial boards, program committee service and leadership with ten international conferences, as well as thesis and habilitation committees. The project team also provides leadership to scientific committees in France, Europe, and international efforts, as evidenced by the Grid 5000 and ALADDIN projects, GDR ASR, RGE, SC2007 exhibit, and so on.

***Manpower and Means***

The team is well qualified to continue this research, and would be well served by additional engineering support to provide software development and technology transfer of the basic research.

## Evaluation of Project DOLPHIN


### *Abstract*

The DOLPHIN project, directed by Professor El Ghazali Talbi (UR Lille), solves discrete optimization problems with multiple objectives using distributed computation. The three major objectives of the project are: (1) choosing suitable models for multi-objective optimization problems (MOP), and analyzing the structure of such problems; (2) the design of hybrid algorithms that combine different meta-heuristics and exact algorithms for MOP and (3) the implementation of the algorithms for MOP on grid computing resources. The team has solved problems from different areas, some of them from practice, and others model problems from theoretical studies: flow-shop scheduling, routing, network design and optimization, genomics and proteomics. The project brings together researchers in operations research, computational intelligence, and parallel and distributed computation to solve these challenging problems.

Multi-objective optimization models have been developed for bin-packing problems (the problems studied here have two objectives), where the authors have obtained new lower bounds; such models have also been developed for identifying proteins from peptide fragments in mass spectrometry; for association rules mining in DNA microarray data; for mobile network design; for vehicle routing problems, etc.


### *Scientific Excellence*

The team members have designed parallel algorithms for multi-objective optimization problems by partitioning the search space of solutions, and then using exact searches (branch and bound methods) in each search space. Meta-heuristics such as genetic algorithms are combined with Pareto-local search algorithms to solve bi-objective flowshop scheduling algorithms. Machine learning methods have been used to identify characteristics shared by good solutions.

The team has used grid computing (thru the INRIA's Grid5000 infrastructure) to bring to bear thousands of processors distributed across France to solve a flow shop scheduling problem exactly, the first time the exact solution to this problem has been computed. Algorithmic innovations developed for this result using the branch and bound algorithm involve an efficient way to represent the irregular computational trees, load balancing, fault tolerance, and communication optimization.

The team members have developed a software framework called ParadisEO, for solving multi-objective optimization problems MOPs using various meta-heuristics on the Grid. The software is designed to permit reuse of code, be flexible, be portable on a number of parallel architectures, and provide transparency with respect to performance. More than a thousand downloads have been reported by the team. However, the Discussion Forum associated with ParadisEO shows only two messages, both dating to a few years ago, and hence it is difficult to gauge how widely other research groups use the software.

The team has produced strong scientific results that have been published in highly ranked journals such as the Journal of Parallel and Distributed Computing (JPDC), European Journal on Operations Research, and first-rate conferences such as IPDPS. However, there are also many publications in less competitive conferences and poster presentations, which the expert panel found difficult to evaluate. As a specific example, it is hard to assess the work related to proteomics, which we expect to be a significant contribution, as the publications are in two conferences without accompanying papers in the proceedings (one is a poster). In such cases, publications in more competitive conferences with published proceedings would disseminate the results more effectively.

***Adequacy to INRIA's Scientific Strategy***

The accomplishments and plan of the team matches very well with INRIA's scientific strategy. For instance, Section 3.1.1 of INRIA's strategic plan refers to the modeling, simulation and implementation of complex dynamical systems, which is the main thrust of DOLPHIN's activities. The current work concerns complex systems, and increased focus on optimizing dynamic systems is planned for the next four year's work. Also, Section 3.1.3 refers to Grid computing, and indicates the necessity of employing non-local, distributed resource allocation techniques.

***Industrial Transfer and Partnership***

The DOLPHIN team's impact is strongest in the domain of industrial transfer and partnership. A large number of international collaborators in Europe (U.K., Spain, Portugal, Switzerland, and Luxembourg), Mexico, Brazil, and Japan are involved with the project team. In networking, the project has worked on practical projects for France Telecom R&D (strong collaborations), Mobinets, T-Mobile, CTEOS, SOGEP (a delivery subsidiary company of REDCATS (PPR group)) etc. In bioinformatics, the project has collaborated with CEA, the Institut Pasteur at Lille, Genopole, and others. The DOLPHIN team is also involved with French national R&D projects (like ANR DOCK and CHOC, CPER projects like MOST2 and LOVAD, and European projects such as GRAAL.

***Manpower and Means***

The team seems to have had adequate manpower and means (thanks to a lot of external funding) to accomplish its stated goals. It seems, however, that the software packages would be more widely and professionally disseminated and professionally maintained if the team was successful in hiring software engineers who can be supported over a longer term (current support seems to run from year to year), and perhaps better compensated.

To encourage adoption of their tools in the academic community we recommend that the team work more closely with national HPC agencies (like the newly founded GENCI) or HPC users groups (groupe calcul ORAP).

## Evaluation of Project GRAAL

### *Abstract*

The GRAAL project, directed by Frédéric Vivien (INRIA Rhône-Alpes), investigates algorithms and scheduling for distributed heterogeneous platforms. It has three distinct activities. The first deals with theoretical studies of scheduling algorithms, the second deals with the implementation of a middleware for high performance computing on the Grid, and the third deals with both theoretical and implementation aspects of parallel sparse direct solvers. Two software packages have resulted from the work of the project team, namely MUMPS for the solution of large sparse systems of linear equations, and DIET for building and executing applications on a Grid of heterogeneous systems.

The three activities complement each other and are well integrated within the project. For instance, the Grid middleware, DIET, incorporates many of the scheduling algorithms that have been designed by the project-team, and the members of the group that design algorithms use DIET as an experimental platform to test and measure the performance of their algorithms. Similarly, DIET has been used to distribute the computations in MUMPS on the grid, which provided an experimental platform to test the scalability and efficiency of the sparse solvers. The synergy between the three activities in the group is probably due to the fact that the members of the project have been working together for an extended period of time.

Recently, the members of the project have started a collaborative fourth activity with another project team (HELIX) for parallelizing existing bioinformatics algorithms. This activity is still in the exploratory phase and a new software package, Mpi-MkDom2, is in the initial phase of implementation for producing a representation of protein domain families. The DIET toolbox is used for the implementation of Mpi-MkDom2.

### *Scientific Excellence*

The team has produced excellent scientific results that are recognized as such by the international community. These results are published in first-rate conferences and journals and are widely referenced by international researchers in the field. For instance, the theoretical and algorithmic studies are thorough and at the forefront of the state of the art in the field of scheduling. The members of the team are pioneers in introducing the concepts of scheduling divisible loads. They have also designed data structures for achieving fault tolerance as well as load balancing and self-stabilization in large-scale distributed systems. These structures are essential components for achieving efficient data management and replication that are crucial for the implementation of scheduling heuristics. Future plans of the group in the algorithm area include extending the scheduling algorithms to multicore systems and taking into consideration energy efficiency and failure prone platforms, major challenges in current data centers around the world.

The work on sparse solvers is one of the most recognized and successful efforts in that area. MUMPS is used by a large number of researchers in academia and industry. The

value of this work stems from both theoretical innovations and efficient implementation, including out-of-core implementations for systems with limited memory and distributed implementations that are efficient on large grids. The team plans to extend this work to include the efficient parallelization of the symbolic pre-processing steps that precede the parallel numeric solution of linear systems.

The DIET implementation is more recent than the MUMPS implementation but is gaining recognition in the grid community. The tools provided by DIET seem to be well documented and easy to install and use. This is evident by the fact that many international research groups have published results using the tool to execute distributed programs by simply downloading the toolbox and building it without even contacting the DIET team or asking them for any help. Moreover, these independent researchers have reported that DIET is more efficient than other available grid middleware. The DIET team also contributed a reservation engine, called GRUDU, to the Grid 5000 community. Future plans in this area includes integrating replication and data management at the scheduler level in DIET and extending DIET to implement work stealing on desktop networks, two promising directions for a broader adoption of grid computing.

It is hard to assess the work related to bioinformatics, since the group has not produced tangible results in this area as of yet. However, the expert panel predicts that the expertise of the research team in parallel scheduling will ultimately lead to successful and efficient parallel implementations of very important algorithms in this emerging area. The panel encourages the team to increase their activities in that domain, but to make sure that they collaborate closely with biologists to make sure that the algorithms being parallelized have significant scientific value.

In summary, the team has produced excellent scientific results and has presented a solid plan for the next four years.

***Adequacy to INRIA's Scientific Strategy***

The accomplishments and plan of the team very well match INRIA's scientific strategy. For instance, Section 3.1.1 of INRIA's strategic plan refers to the modeling, simulation and implementation of complex dynamic systems, which is the main thrust of GRAAL's activities. Specifically, page 44 mentions high performance solvers and scheduling methods and page 47 mentions high-resolution simulation using very efficient solvers of large linear systems. Also, Section 3.1.3 refers to ubiquitous communication, information and computing. Specifically, page 61 mentions the need to deal with heterogeneous systems and client-server systems on non-reliable platforms. The plan indicates the necessity of employing non-local, distributed resource allocation techniques. Finally, the plan mentions that using computer platforms for the simulation of medical models will require parallel numeric solvers.

***Industrial Transfer & Partnerships***

The team was successful in the domain of industrial transfer and partnership. MUMPS has been integrated in many non-commercial and commercial products including in

applications related to the oil industry. There were also two industrial contracts signed with SAMTECH and CERFACS/CNES for its use. The GRAAL team has been also collaborating, as grid experts, with the DECRYPTHON program led by the AFM, CNRS, and IBM. In 2007, DIET was chosen to be the grid middleware for the Decrypthon grid infrastructure.

***Manpower and Means***

As is clear by its past successes, the team seems to have had adequate manpower and means to accomplish its stated goals. It seems, however, that the developed software packages will be more widely and professionally disseminated and maintained if the team is successful in hiring a software engineer. In addition, the panel recommends the establishment of an open source community for the software packages, which will allow the members of the larger scientific community to make direct contributions to the package evolution and support.

Recently, a number of members in the area of sparse solvers have left the team. This may affect negatively the pace of progress in an area of recognized excellence within INRIA. It is, thus, recommended that a new member be recruited in that area. Finally, the team has to acquire expertise in the bioinformatics area to make tangible and scientifically sound contributions to that area. This may be accomplished through external collaborations or through the addition of a new member specialized in this field.

## Evaluation of Project GRAN-LARGE


### *Abstract*

The Grand-Large research project deals with topics related to computing in Large Scale Distributed Systems (LSDS), where users share an infrastructure whose resources are subject to failure, heterogeneous, geographically distributed, and are administratively independent. More specifically, Gran-Large considers Desktop Grids and Grids although some of the research also considers large-scale parallel computers and wireless sensor networks. The research topics considered in the project include: design of middleware for LSDS, large scale data movements on LSDS, fault tolerance MPI for LSDS, fault-tolerant protocol verification, algorithms, programming and evaluation of scientific applications for LSDS, tools and languages for large scale computing on LSDS, and exploration systems and platforms for LSDS.


### *Scientific excellence*

The project deals with topics that are very relevant to the scientific community. The results obtained by the group are very relevant, and the way the group performs the research, by studying experimentally and formally and by developing real software has a lot of merit. Software environments like XtreemWeb, BitDew or MPICH-V are not only a good research exercise but are also useful tools for the scientific community. The project has evaluated their research work with real-world applications.

Most of the objectives defined in late 2003 by the Gran-Large proposal (this is their first evaluation, since in 2004 was not evaluated), has been accomplished with success. The group has now a large number of PhD students (half of their total personnel) and 4 of them have finished the PhD during the period, which is considered a reasonable number for the size of the team. Also, one of the researchers obtained his HDR. The number of publications of the team during the period is quite impressive, with a good proportion between international conferences and journals. Additionally, the group has been developing and distributing eight software packages related to different topics of their research.

The objectives for the next period evolve from the previous ones, with some topics continued, other discontinued and new ones added, this being an expected evolution. A remarkable evolution is the new target on Petascale parallel computers, with hundreds of thousands of computers.

Some of the topics are discontinued due to the departure of some researchers from the team. This seems to be common in INRIA, the fact that research topics migrate from given INRIA centres to other due to departure of individuals. It might be more sensible that topics remain in a center and that when researchers move from one center to another, they will adapt to new research topics enriching their new destination with their previous research.

### *Adequacy to INRIA's scientific strategy*

The project objectives and research topics are adequate with the INRIA's scientific

strategy for the period 2008-2012. The project fits with objective 3.1.2 Programming: security and reliability of the computer systems and with objective 3.1.3. : Community, information and ubiquity computations.

***Industrial transfer and partnerships***

The project has a considerable amount of external funding coming from different sources: National and regional initiatives, European Commission and industrial contracts. This amount has been stable during the period, and even seems to increase during the last year of the period. One possible criticism here can be that the number of projects that the team is involved seams to be quite large (for example, 14 in the last year) and this can be a source of overhead for the team.

***International level***

The team has excellent international visibility, with collaboration with important groups, especially from the U.S., perhaps overly so (specific collaboration with groups in France or Europe were not explicitly mentioned).
Besides the explicit collaboration mentioned by the team, there is a clearly significant level of international recognition, as evidenced by participation of the senior members of the team in program committees of relevant international conferences

***Manpower and means***

The team has been growing during the period from 16 to 30, therefore almost doubling the number of participants. The distribution between senior researcher, researchers and students is a bit incommensurate with a large number of PhD students (9 on the first year, 15 in the last year). The project recognizes that it can run into problems if they cannot hire experienced researchers.

New researchers have been integrated in the team and others had left the group. Some of these researchers moved to international positions in the US and rest of Europe. Again, this demonstrates the international recognition of the project.

***Perspectives***

In scaling out to peta-computing and beyond, the review panel hopes that the team will resist the (entirely human) temptation to model the future on the past. The team and stakeholders will be best served if they afford a clean-sheet comprehensive look at the new landscape, and specifically ask which dominant application patterns will drive the use cases and requirements at the new scale, and how fault-tolerance properties will combine at various levels in the stack?

## Evaluation of Project MESCAL


### *Abstract*

Led by Dr. Bruno Gaujal, project MESCAL seeks to develop mathematical tools to be used in the evaluation and design of "scalable middleware". The mathematical work done by the team is at the front edge of research in this area, and is certainly germane to INRIA's scientific strategy. Technology transfer to industry is basically good, although at this point seems limited to fairly straightforward tools, and the personal expertise of the tool (and model) builders.


### *Scientific excellence*

The MESCAL project themes include

- Performance models for large computing infrastructures
- Exploitation of large architectures
- Self-organization and fairness
- Dimensioning of platforms and scientific applications.

We consider these each in turn.

*Performance models for large computing infrastructures:* The goals set out in this area (documented in the MESCAL 2007 report) include development of guarantees on behavior, to prove certain properties of large Markov chains derived from models of 802.11, apply the notion of perfect simulation to Petri nets, and explore development of performance bounds based on stochastic orderings. The group has largely met these goals. The results reported on perfect simulation and fluid models are quite impressive.

*Exploitation of large architectures*: The key contribution in this area are tools for allocating resources to a distributed computation, called KaDeploy, and for scheduling, called OAR. The main point of interest of these is that they've been deployed in the Grid5000 system and have some use. This provides some contribution towards technology transfer.

*Self-organization and fairness:* The team has developed some nice theory relating to fairness and balance, and is exploring applications in peer-to-peer computing and cellular systems.

*Dimensioning of platforms and scientific applications*: The group has developed some models of communication across different types of networking architectures. The models appear to be at a flow level (hence high level), and support a simulator. It's not clear whether this work leverages quite a bit of similar efforts from earlier days when supercomputers were found only in government labs. Quite a lot of research was published in the 80's and early 90's that address similar sorts of issues. The most

important value of this work might well be in the applications to which it has been applied (e.g., seismic simulations, EM fields).

Overall, the scientific merit of the work performed by this group is very good. The deeper mathematical work is novel and cutting edge, and the group is exploring ways of bringing analytic models to practical use. Looking to the future, the directions being followed in developing mathematical models and results is sound, and can be expected to yield foundational results in the future, as it has in the past. In terms of priorities this group ought best to "play to its strength"---providing such tools to industry as might be expected for public support of research, but with the recognition that there is (and will always be) a significant gap between the sophistication of leading edge analytic techniques, and what is desired (and understood) outside of the academy.

The team shows also good cooperation with other NumB project teams, like MOAIS or Algorille, with tools like SimGrid or Taktuk being developed in cooperation between at least two teams.

***Adequacy to INRIA's scientific strategy***

The project objectives and research topics are adequate with the INRIA's scientific strategy for the period 2008-2012.

***Industrial transfer and partnerships***

The main mechanism for "industrial transfer" and partnerships has been to obtain contracts with companies interested in using modeling and modeling tools, with project team members effectively acting as consultants on model development and tool use. For the most part it does not appear that industry is ready to use shrink-wrapped tools; the type of partnership that has evolved is natural and necessary.

For the Kadeploy (easy deployment of system images thru the network) and OAR (batch scheduler) an industrial transfer to a company like BULL may be encouraged.

Also the team presented the ComputeNode Grid Manager, which allows one to boot an unused workstation remotely (with no intrusion) and transform it into a compute node of a distributed pool of resources. A startup called ICATIS has been created for the technology transfer but it seems that two years after its creation the company has an uncertain financial future. This tool may be re-incorporated onto the INRIA tools portfolio and another partner (not necessary a spinoff for a so little tool) may be contacted for the industrialization of the solution.

***Manpower and means***

The composition of the team (roughly 2/5 professors and 3/5 PhD students, with a couple of engineers) is appropriate for advancing in the core strength areas of the group. If a greater emphasis were to be placed on tool development and distribution then more engineering support would be needed. (However, we do not recommend such a re-targeting.) The group has developed research partnerships with some leading research

teams around the world.  They appear to have the manpower and means to do what they do best.

## Evaluation of Project MOAIS

### *Abstract*

The MOAIS project-team was created in early 2005 and is directed by Jean-Louis Roch (INRIA Rhône-Alpes). It focuses on end-to-end parallel programming solutions for high-performance interactive computing with provable performances along four distinct axes. The first deals with the study of multi-objective scheduling algorithms with performance guarantees, the second deals with the design of adaptive execution algorithms under the control of the scheduling, the third deals with the development of interfaces for coordination of runtime components, and the fourth applies the resulting algorithms and interfaces to interactive computing applications on parallel platforms. Three software packages have resulted from the work of the project team, namely Kaapi for fine grain scheduling of multi-threaded computations, FlowVR for coarse grain scheduling of interactive computing applications, and TakTuk for large scale remote execution deployment.

The four axes complement each other and are well integrated within the project. Scheduling and adaptive execution algorithms with provable performances are designed based on generic theoretical models. They are implemented and evaluated using the software packages developed by the team, and customized for their use with partners in real life interactive computing applications. The project-team thus combines theoretical work with implementation and validation of the theory on real case scenarios, which have applicability to industry.

### *Scientific Excellence*

The project-team, which only started its activities three years ago, has produced outstanding scientific results**,** with impressive publication records: 127 publications, 19 of which in major international journals and conferences. The three software packages developed by the team are widely used by various scientific teams and communities. Kaapi is used by teams working in combinatorial optimization, exact arithmetic, computer-aided process engineering, meta-heuristics and inference engines. It won the first prize during the Grid Plugtest 2007 N-Queens challenge. FlowVR, demonstrated at SIGGRAPH 2007, is used by virtual reality and scientific visualization communities (more than 50 downloads per month on SourceForge). TakTuk is widely used by the Grid'5000 community, for adaptive remote application deployment.

The work on scheduling produced the best approximations, as of today, for 2-D packing of parallel moldable tasks with dependencies and for bi-objective scheduling algorithms with a theoretical guarantee on each objective. It uses either relaxation methods for makespan/minsum and work/depth approximations, or new approximations of Pareto curves to reduce the number of compromised solutions in makespan/reliability and makespan/memory optimizations. The team also studied the impact of uncertainties on the scheduling algorithms and started to investigate the applicability of game theory to the challenging problem of multi-objective optimization with uncertain data. Work was also performed on reducing the expensive control cost of work-stealing scheduling for a

very large number of tasks. The resulting distributed work-stealing scheduling algorithm has been implemented in Kaapi. Most other algorithms have been assessed through simulations and practical implementations on actual platforms remain to be done.

The work on adaptive execution algorithms showed that adaptive coupling of a sequential algorithm with a parallel algorithm allows for minimization of the work while preserving the depth. Excellent work has been done on recursive cascading of sequential algorithms applied to various applications, e.g. 3D-vision, resolution of multiple linear systems, cryptography, quadratic assignment problems, and dynamic application deployment (base of TakTuk). Excellent work has also been done on adaptive work-stealing algorithms applied to interactive 3-D modeling, parallel prefix computations and distributed STL algorithms, and stream computations on MPSoCs (leading to the development of the Adaptive Work Stealing library).

The work on coordination interfaces focuses on the assembly at runtime of application components in order to adapt their coordinated execution to a large set of parallel architectures and environments (ranging from multi-cores/multi-processors to clusters and grids) by transforming a DAG based representation of their dependencies. The Kaapi model based on this representation has been developed; adaptive work-stealing algorithms and novel fault tolerant protocols (e.g. Theft Induced Checkpointing) have been implemented in the Kaapi software package in order to allow for efficient runtime of parallel applications.

The work on interactive computing applications produced FlowVR, a component oriented middleware for high performance interactive applications. FlowVR allows the coupling of parallel applications through its hierarchical component model and to encapsulate data exchange patterns in components. FlowVR was experimented with for real time 3D modeling and real time physical simulation, with a fruitful collaboration with the PERCEPTION and EVASION projects, bringing complementary skills in virtual reality and computer vision. This work led to a demo at SIGGRAPH 2007, which received very good media coverage and was showcased on the Grimage pioneer platform, developed jointly by the three project-teams to enable real-time marker-less 3D interactions.

The MOAIS project-team research directions for the next four years are in multi(more than two)-objective scheduling on heterogeneous platforms, adaptive algorithms for large data sets addressing out-of-core issues, and dependable global computing, leveraging the flexible rollback recovery protocols the team recently developed.

In summary, the MOAIS project-team has produced very good scientific results and proposes a strong plan for the next four years.

***Adequacy to INRIA's scientific strategy***

The achievements and research directions of the project-team very well match INRIA's scientific strategy, in particular with regard to Sections 3.1.3 (Ubiquitous Communication, Information and Computation) and 3.1.4 (Interactions with Real and

Virtual Worlds). For instance, page 61 mentions that the deployment and execution of distributed applications raise many management challenges associated with the heterogeneity and volatility of infrastructure resources, and the expected performance of the applications. These challenges correlate well with MOAIS project-team objectives.

Also, page 68 mentions that the ubiquity of computing resources required to interact with a virtual world with realism necessitates taking into account the limitations and the variability of these resources. This is precisely what the MOAIS project-team does, through the development of adaptive execution algorithms and their customization for interactive computing applications, such as virtual reality and computer vision.

***Industrial Transfer & Partnership***

The MOAIS team was very successful in the domain of industrial transfer and partnership. Through the validation of its theoretical and prototyping work on real life applications, the project-team collaborated with several industrial companies and scientific teams. The real-time 3D modeling code developed by the PERCEPTION and the MOAIS project-teams was transferred to the 4D View Solutions company, created in September 2007. In collaboration with STMicroelectronics, the team developed an adaptive code for stream computation on MPSoCs, with applications to video compressing and filtering. This work led to the development of the AWS specialized library. Kaapi was used by companies and research teams, like ProBayes specialized in probabilistic inference engine, IFP in the context of their CAPE-Open standard activities, Linbox and LJK in exact arithmetic, the Opale team of the PRiSM laboratory and the GSCOP laboratory in combinatorial optimization, and the University of Luxemburg in relation to their meta-heuristics work. The TakTuk adaptive application deployment package was contributed to the Grid'5000 community. The makespan/min-sum scheduling algorithm was implemented in the Grid'5000 OAR batch scheduler by the MOAIS team.

***Manpower and Means***

While the team has had adequate manpower and means to accomplish its prior objectives, we believe that additional engineering manpower is required by the team in order to maintain, document, and adapt its software packages on existing and new architectures. The 2-3 year cycle of engineer hiring is not appropriate to maintain the viability and high quality of these packages. Another key issue for the team is likely to be keeping up with the fast evolution of computing and video technologies. It is essential in order for the team to stay current that it maintains and even reinforces its existing relationships with STMicroelectronics and Bull.

## Evaluation of Project PARIS

*Abstract*

The PARIS project-team, headed by Thierry Priol (INRIA Rennes Bretagne Atlantique), defines its overall goal to be contributing to the programming of parallel and distributed systems for large scale numerical simulations along five research themes: Operating System and Runtime for Clusters and Clusters of Clusters, Middleware for Computational Grids, Large-scale Data Management for Grids, P2P Foundations, and Advanced Programming Models for the Grid.

The PARIS project started in 1999 and is now in its final phase. PARIS developed from its origin with originally 9 project members to a quite large project team with a staff of 34 researchers on all academic levels now. As explained at the review meeting a reorganisation of the project team is expected within a short time frame, which will refocus the research agenda and define two new project teams with two new leaders. By this strategic move of the original research topics to the more general topic "distributed systems" (from Grids to Service Infrastructures) a broadening of the research focus will be achieved.

This development is highly appreciated due to the actual situation that the project has a strong focus on the Grid infrastructure at the moment. It is possible that in the near future Grid research will be cut back to pure HPC application domain and other arising computing paradigms, as cloud computing, or infrastructures (e.g., pure service oriented architectures), as these are expected to be the new dominating areas of research. An extension to new topics will guarantee impact on the scientific landscape and continued co-operation with and support of industry partners.

The specific objectives of the team for the evaluation period comprise three main areas of research:

- operating systems for clusters,
- grid middleware,
- P2P systems.

The first objective "Operating System and Runtime systems for Clusters and Cluster Federations" consists basically of two strongly related sub-objectives, the development of Kerrighed, a full Linux-based single system image operating system, and of Grid-aware operating systems. Kerrighed is an operating system for cluster architectures, which presents to the user a single system image on top of the cluster, in other words the view of a uniform, single SMP system. Kerrighed was developed with ease of use, high performance of applications, high availability of the cluster, efficient resources management, and high customizability of the operating system in mind. The main achievements are the development of a high performance communication system, a customizable scheduler framework, a distributed file system, a reconfiguration, and a checkpoint/restart mechanism. Kerrighed evolved to a well-renowned open source project

with an established user community. It is used by an increasing number of applications and projects and aims for production level quality. The Kerrighed research influenced very strongly the development of the Grid-aware operating systems, Vigne and XtreemOS. Vigne is a Grid operating system for Linux aiming for scalability and transparency. However it misses some important aspects as support for virtual organisations, security and a Grid file system. XtreemOS is a Linux-based operating system to support virtual organizations for next generation Grids for workstations, clusters, and mobile devices. Its development is funded by the EU integrated project XtreemOS headed by Christine Morin, the main (permanent) researcher of Kerrighed.

The second objective "Middleware for Computational Grids" deals with issues for the design of a software component-based model for the development of high performance computing applications for computational Grids. The aims are to provide an environment for the integration of distributed objects within a component architecture, the automatic deployment of application onto available resources on the Grid, and the design of adaptive components adapting their behaviour to the resource situation on the Grid dynamically.

This middleware research is complemented by the third objective "P2P System Foundations" which tackles the problem of managing large-scale data set on the Grid. Based on P2P mechanisms the JuxMem software platform provides a transparent, consistent, persistent, and resilient access to a shared data storage space on the Grid by appropriate models and protocols.

**Scientific Excellence**

PARIS's specific strengths lie in comprehensive coverage of engineering aspects of the support and development of high performance applications, specifically in the numerical simulation domain, by focusing on operating systems, infrastructures and programming models. Hereby the PARIS project team managed to cover in its research activities both basic and applied research.

The results obtained by the PARIS team are very relevant to the scientific community and influenced strongly the scientific landscape far beyond the borders of France. The project achieved internationally highly recognized research results and compares well to the top level of international research. Under the excellent lead of Thierry Priol the project developed favourably in the last years and the size of the team increased dramatically (it was more than tripled). The publication record is high (17 journal papers, 65 conference papers, about 10 PhD defended), which is also due to the large number of project team members. However it is recommended to focus stronger on journal publications additionally to the already large number of published peer-reviewed conference papers.

***Adequacy to INRIA's scientific strategy***

The project is directly in line with the scientific strategy of INRIA by its work on a supporting infrastructure for high performance numerical applications. Specifically the comprehensive coverage of the engineering aspects of the support and development of

high performance applications provides for successful development and deployment of large scale software systems on grids and P2P systems. This will enable other project teams to exploit the available data and computing resources of the national Grid infrastructure for real applications.

**Industrial Transfer and Partnerships**

A considerable amount of funding (600-1000 KEuro per year) is coming from external sources of different origin, national bodies, European Commission, and industry.
The project shows very good integration into the European research landscape and research community. The objectives of PARIS are predominantly established within two large EU funded projects. The operating system objective is the driving force of the XtreemOS project. Also PARIS members have predominant positions in the EU funded Network-of-Excellence CoreGrid; Thierry Priol is the scientific coordinator of CoreGrid. CoreGrid specifically aims for establishing an internationally recognised and sustainable European Research Laboratory of well-established researchers with a joint program of activities in the area of Grid and Peer-to-Peer technologies. This dominant representation of PARIS members within EU funded initiatives guarantees a dissemination of PARIS results beyond the border of France and also provides for certain sustainability by external funding of the project's research activities beyond the duration of PARIS. Further the spin-off company Kerlabs is acknowledged as success story. This company was created in 2006 by members of the PARIS team and develops cluster solutions using the Kerrighed technologies.

**Manpower and Means**

Due to this sustainable integration of the projects research objectives in industry and other research projects the two newly to be installed sub projects are to be expected to be successful as well. A possible threat is that Thierry Priol is expected to be the leader of the new ALADDIN-Grid'5000 initiative so that the project is loosing its experienced project management and leadership. Therefore the choice of the new sub leaders has to be done very carefully to guarantee the continued success of the projects. As mentioned at the review meeting this is a well known issue to the consortium and respective actions and decisions are on the way. So it is to be expected that the PARIS project will be successfully continued by the two new sub projects. Therefore it is recommended that these two newly to be established sub projects will get the continued support by INRIA.

## Evaluation of Project RESO

### *Abstract*

The RESO project is concerned with the communication software, services, and protocols that support Computing/Data Grids operating over short or long-distance networks. Specifically, RESO's scope includes a) optimized data communication architectures for end-systems and programmable access equipment, b) algorithms and protocols for efficient transport and Quality of Service (QoS) for heterogeneous traffic at high speed, and c) network protocols and services. The RESO activities span requirements mining, problem identification, and solutions validated by theory and practice. RESO's outcomes appeal to a diverse audience of researchers and developers of Grid middleware, advanced Grid end-users, service providers, network operators, and equipment manufacturers.
NOTE: Due to restrictions on intellectual property, the reviewers were unable to develop a view of the new "semantic networking" initiative in collaboration with Alcatel-Lucent. This topic appears to be a defining objective for the next four years.

### *Scientific excellence*

The RESO deliverables marry networking themes with Grid themes. This topic area has been traditionally under-represented, given that the Grid paradigm arose from CS circles and the networking problems were typically addressed in an ad-hoc fashion by highly capable system administrators.

The reviewers believe that RESO has delivered on the objectives delineated in 2004. Clearly, the RESO effort has advanced the knowledge on how to improve predictability and reliability of data transfers across local and wide area networks. Their lessons learned have had an impact on other researchers, standardization bodies, and the industry. Within INRIA, the RESO project has had synergisms with a number of projects (e.g., RUN-TIME, ALGORILLE, RAP, MADYNE, MAESTRO) and initiatives (Grid 5000 and Renater).

The objectives proposed for the next four years have merit. They build upon the strength demonstrated during the review period and hold potential to maintain, possibly augment scientific excellence. The reviewers recommend the expansion of liaisons with like-minded initiatives abroad such as FIND, FIRE, Geant.

Upon successful launch of the new "semantic networking" theme, it might be appropriate to entertain a spin-off and make semantic networking a new project in its own rights.

### *Adequacy to INRIA's scientific strategy*

In general, RESO's activities appear to be well aligned with INRIA's strategy for the 2008-2012 period. Specifically, RESO is seen contributing to the strategic objective 3.1.3 (Communication, information et calcul ubiquitaires).
The reviewers singled out the RESO project as the sole Num-B effort concerned with the network. Going forward, the reviewers believe that Com-B (Réseaux et télécommunications) might be a more fitting domain for the RESO project.

***Industrial transfer and partnerships***

Throughout the review period, the RESO project has had several engagements with the industry. ORFA (user-level remote filesystem access protocol) and Tamanoir (software for high speed active networks) represent two successful transitions from the research domain to industry realities (Myricom Inc. and 3DDL respectively).
Seven of the industry engagements resulted in grants for a total funding envelope of 265 KEuros over the past 4 calendar years. The relationships with FT and Alcatel-Lucent started in 2007 are a promising sign of RESO's relevance to industry at the beginning of a new 4-year cycle.

***International level***

The RESO project has international visibility through its leaders such as Pascal Vicat-Blanc and Laurent Lefevre, who enjoy international recognition in a number of research communities. They are well positioned to relate RESO activities to world-class research and to mentor emerging researchers to follow their footsteps.

***Manpower and means***

The RESO manpower has had moderate growth over the review period (from 17 to 20.5). Specifically, two DR/Professor level resources joined RESO. The growth in senior staff hints that it might be possible to increase the number of PhD students during the next funding cycle.

The RESO staff is highly engaged in teaching activities, with about 2000 hrs of lecture over the 4 calendar years. This bodes well for new PhD candidates and talent renewal. RESO's external funding has changed profoundly over the past 4 years. In 2007, RESO signed up new national initiatives (e.g., ANR), European projects (e.g., EC-GIN and AEOLUS), and industry partners. The renewal of funding channels is seen as a positive, even though RESO's senior leaders will have to manage churn in their liaisons.

The nature of RESO's activities calls for continued access to network testbeds of geographic scale. The reviewers are concerned that Grid '5000 does not provide sufficiently many compelling Grid applications and use cases. The RESO project will be best served by international testbeds (e.g., Geant/Dante, Surfnet/DAS3) seeing a healthy proliferation of 3rd party application stacks.

## Evaluation of Project RUNTIME

### *Abstract*

The RUNTIME project focuses on low-level issues related to run-time environments for parallel architectures. Led by Professor Raymond Namyst, the projects works on problems whose solutions provide higher performance communication for clusters, multi-core thread-scheduling, and integration of multi-threading and communication.

### *Scientific Excellence*

The research problems are well chosen, and timely. Some good successes are reported in just-in-time optimization of communication. In the realm of thread scheduling, the team is in line with current thinking that optimizations require some directives from the programmer, but a run-time system can still hide the lowest level details. The team's ideas are finding expression in versions of MPI, and libraries of vendors of communication networks. The insights of considering memory affinity when scheduling are important, and the team's work recognizes this.

The team's projection of problems to address in the coming years are good ones, and relevant. They ***must*** however in the future include consideration of compilers and the information they can provide for real-time optimization. World leaders in optimization of performance on multi-core architectures are quite persuasive in taking this research direction.

The team's publications are mostly in workshops. The relationships that are being developed with other research groups around the world is evidence that the work is valued, but without significant publication in leading parallel processing journals (e.g., *Parallel Computing*, *IEEE Transactions on Parallel and Distributed Computing*), the work risks being unnoticed by the larger world.

### *Adequacy to INRIA's Scientific Strategy*

The sort of low-level run-time optimizations pursued by this team are an important component of making high-performance computing usable and accessible to scientists and engineers in academia, government, and industry; users who develop simulations of dynamic physical systems particularly depend on this work. As such, the work is in line with INRIA's scientific strategy in the area of modeling, simulation, and optimization of dynamic complex systems (3.1.1.)

### *Industrial Transfer and Partnerships*

The RUNTIME project has done a good job in collaborating with others. Partnerships with Argonne National Laboratory and (soon) the University of Illinois at Urbana-Champaign are critical connections with places where cutting-edge research in parallel run-time systems is being conducted; the team collaborates with other universities as

well. A notable success is the use of the team's Open-MX software on the IBM BlueGene/P supercomputer at Argonne. Coordination with Myricom is an important industrial connection, as Myricom networks are exactly the kind of architecture that may benefit from RUNTIME's research. There are connections between RUNTIME and half a dozen other INRIA project teams, and the team has acquired half a dozen contracts as well. The team is also involved on many French ANR R&D projects such as PARA, NUMASIS and LEGO.

The team has started industrial collaborations and technology transfer with CEA (French National Nuclear Research Center), but these efforts need encouragement.

Wider success of RUNTIME's research may be achieved by working more closely with the OpenMP and MPI forums, in order to get some of RUNTIME's advances reflected in in standards, and implemented in compilers.

***Manpower and Means***

System building is a labor-intensive activity. Success of this activity depends very much on highly skilled and knowledgeable personnel at all levels. This is a potential problem area for the team.

The team said that they sometimes suffer from a lack of access to large-scale high performance computing facilities for testing and improving their tools. We advise INRIA to work more closely to the newly French HPC agency called GENCI. Such cooperation will allow INRIA R&D access to larger scale (in term of number of cores and supported HPC architectures), HPC facilities and also enable them to distribute their software to a significant community of academic users.

## Evaluation of Project SAGE

### *Abstract*

SAGE is primarily about algorithms for large-scale nonlinear and linear systems, implemented for distributed memory platforms, and applied to geophysical problems – mostly flows in heterogeneous permeable media described by partial differential equations. SAGE also maintains a software platform called Hydrolab for this class of applications. Off of this vertically integrated core of activity hang naturally related algorithmic projects, corresponding parallel implementations, and other PDE-based applications. While methods for the largest-scale problems are always iterative, sparse direct solvers have key roles within subproblems of fixed size, such as subdomain solves or coarse grids, and a member of SAGE has made notable contributions to direct solvers. SAGE personnel have also worked on Krylov solvers for linear systems and eigensystems. Parallel implementations have been made in both direct and iterative contexts. Geophysical applications include different types of transport, including those with interfaces to be tracked (free boundary problems) and those modeled with continuous fields for passive transport and different aspects of stochasticity in coefficient distribution and in the geometry itself.

### *Scientific Excellence*

SAGE possesses a core of scientific excellence, as reflected in the publications, software, and reputations of its principal investigators, in the area of scalable solvers. Appropriately, it is not researching and inventing into a vacuum, but is driven by a well-chosen application for this set of enabling technologies, namely flow in porous media. Its excellence is not so much in the porous media work, itself, though the project does claim a few superlatives in the end applications, but mainly in the solvers. If SAGE had chosen to study the drift-diffusion equations of semiconductor physics, for example, it could have done as well, but engaged an entirely different set of physical collaborators. Such is the nature of enabling technology research. Were SAGE larger, it would benefit from working simultaneously in multiple applications, but the application chosen is both appropriate and scoped to its size. Surprisingly, much effort is devoted to two-dimensional problems. Whenever the science justifies two dimensions (e.g., severe scaling or anisotropy that causes the third dimension to be treated in special way, as in atmospheric modeling where height is distinguished or tokamak modeling where the toroidal direction is distinguished), this simplification should be pursued, because the dimension drastically changes the "sweet spot" for the switchover from direct to iterative methods. Though much of linear algebra is dimension-independent at the sparse matrix level, other linear algebra explicitly exploits domain information and is sensitive to the physical dimension in which the problem is discretized. Given the difference in physics and computational complexity drivers, three dimensions should always be aggressively pursued. Because of the significant effort devoted to two-dimensional examples in SAGE, scalability to truly massive parallelism of the scale of interest for France's new 100 Tflop/s computational capability was not reported by SAGE.

SAGE has invested effort in multiplicative Schwarz preconditioning. This makes sense

for strongly advective problems, where a proper sequential (causal) ordering of domains can save useless computation; however, in diffusive problems, it fails to fully exploit parallel concurrency and is of limited utility at the very large-scale. Probably it could be de-emphasized if SAGE is to reorient towards large-scale industrial customers, e.g., in reservoir exploitation or waste isolation for real-world environments.

Strong and relevant theoretical work has occurred in condition number estimation and rank-revealing QR. It is nice to see that internationally known abstract work of this nature can be supported by applications.

***Adequacy to INRIA's scientific strategy***

As the major player in large-scale simulation in France, INRIA should definitely support projects with the scope of SAGE. France does not have an abundance of linear algebra projects, and each one is important and well connected to the "small world" of other such projects internationally. The vertically integrated principal portfolio of SAGE aligns well with INRIA's scientific strategy, overall. Within the realm of coupled problems, SAGE appropriately leverages its implicit solver capabilities with an extra Newton loop, rather than the weaker forms of coupling often employed in applications. Whether SAGE really belongs grouped with the large computer science-type projects in NumB, or would be better grouped with more numerical projects, such as those of NumD, is another question.

Outside of direct and iterative linear algebra, SAGE has not defined itself as centrally within INRIA. It aspires to contribute to Uncertainty Quantification beyond the brute force and embarrassingly parallel means of doing statistics on an ensemble of calculations, that is, by directly inferring distributions of outputs from distributions of inputs: this a good area with many mathematical challenges, but as such it is a stretch from the training of current personnel, and this direction would certainly be more mathematical than almost anything else in NumB. A more geometric version of Uncertainty Quantification comes in SAGE's Discrete Fracture Network application, in which the randomness is in the domain, as opposed to in the coefficients. The review team strongly affirms the importance of fundamental new attacks on UQ through new mathematics, rather than brute force distributed computing, so this ambition is aligned with what INRIA's scientific strategy should include, and it is grouped with geophysics applications very well, but SAGE is subcritical in staffing to address it.

***Industrial Transfer and Partnerships***

SAGE is well partnered with other academic and laboratory groups, but apparently not with industrial users. SAGE is both a provider and a consumer in its partnerships. Wisely, it uses Hypre and SUNDIALS (freely available from LLNL) rather than duplicating the effort to invent code of such quality. Within France, SAGE uses the excellent MUMPS sparse direct solver. Through the contributions of Laura Grigori, SAGE has contributed to the development of SuperLU_DIST, another important sparse direct solver, from LBNL. On the geophysical side, SAGE partners with Rennes on the Micas project. This consortium uses Hydrolab. As a project deliverable, SAGE intends to adopt a distribution policy for international release of Hydrolab, which should be encouraged. SAGE has already worked on documentation and a web portal. Ifremer

(engaged in ocean exploration) has been a four-year customer of Hydrolab. Future customers could include ANDRA, France's nuclear waste isolation project, and partners of MOMAS including EDF (nuclear plants), CEA (nuclear waste isolation), and BRGM (water resources). A potential industrial customer could be Veolia, which works on water management and distribution. The review committee strongly encourages pursuing visible collaborators like ANDRA, since they confer credibility, on one hand, and they also drive the research with difficulties from real life, which academic researchers often would not dream of adding to their portfolios.

***International Level***

SAGE personnel have international visibility in computational linear algebra, so much so that one reviewer was surprised to discover how small SAGE is. There is less visibility in the application domain of the geosciences. SAGE personnel have some international leadership positions. For instance, Erhel is influential in ERCIM, to which all major European players in the numerical analysis behind large-scale PDE-based simulations (and other related) topics belong. SAGE's visibility is also strongly influenced by its national and international collaborators. No less than INRIA Director Michel Cosnard, himself an important computational linear algebraist, publishes with SAGE researchers. In the USA, James Demmel, Sherry Li, Esmond Ng, Yousef Saad, and Masha Sosonkina are all collaborators. Former SAGE researcher Laura Grigori was important to bridging the two nations' efforts, and should be replaced, if possible, by someone equally fluent in English and available for collaborative travel.

***Manpower and Means***

In terms of manpower, SAGE has the smallest team of those presented under NumB, and it is also one of few teams that appear to be shrinking. This shrinking is not due to declining resources, but to researchers dropping to lower duty cycles on SAGE due to other commitments or due to transfers or graduations. There is some replenishment by students, and the co-advising of students in effect enlarges and leverages the team of researchers engaged in SAGE projects by their generally unfunded advisors. There is relatively high impact per person from this small team; however, there is a spectre that the software legacy will go subcritical. A multidisciplinary team cannot lose too many performers without losing disciplines. SAGE depends on the collaborations and acknowledges this; for instance its work on UQ methods requires partnership with ENS for expertise in stochastic PDEs. As mentioned above, SAGE is well partnered within the USA at the senior collaborator level. SAGE is also well networked with Francophone Africa and Middle East in terms of attracted talented doctoral students. SAGE looks to hire two doctoral students to replace graduating students and one engineer. The student-senior balance is thus maintained, but SAGE should endeavour to make another senior recruit or two in order to control their destiny more in choosing research subjects.

## Evaluation of Project SCALAPPLIX


***Abstract***

The ScAlApplix project is a joint multidisciplinary research activity of INRIA Futurs with LaBRI laboratory for computer science and IMB Mathematical Institute for applied mathematics. Led by Prof. Jean Roman, since its creation in fall 2002, the project deals with all steps of high performance parallel computational suite in the fields of CFD and multi-scale methods for material simulation. Significant contributions have been made during the evaluation period in the development of high performance mathematical methods, high performance algorithms and tools for very large scientific problems.


***Scientific excellence***

The ScAlApplix scientific originality and strength reside in the complementary of the scientific themes developed within a multidisciplinary team, all contributing to the same goal of designing new high-performance, robust and precise adaptive and multi-resolution scheme, and parallel algorithms and tools deriving from these schemes, for the simulation of multi-scale and multi-physics complex phenomena.

The results achieved during the evaluation period, clearly demonstrate the value of such a multidisciplinary approach, which has to be maintained and further developed in the view of the expert panel, whatever the evolution of the project organization will be.
The major scientific contributions hereafter assessed by the panel, have delivered to the objectives defined in 2002 according to the three branches:

- schemes and numerical strategies,
- high performance algorithms and tools,
- computational steering for parallel and distributed numerical simulations.

*High order Residual Distribution Scheme (RDS)*: The goal was to develop a high order scheme compact enough for efficiency in large-scale parallel application. It was achieved by bringing simplicity in well-mastered matrix algebra in an RDS scheme, including extension of formulation for time dependency and hp refinement. The current and future work to further demonstrate suitability compared to Discontinuous Galerkin methods for 3D applications of industrial relevance are valuable contributions. A recommendation for the future would be to combine adaptation hp and uncertainties management for 3D Navier-Stokes applications.

*High performance Multi-scale simulation for crack propagation*: The contribution in this field was to understand the stability and accuracy of the coupling between continuous/molecular schemes. An efficient implementation for 3D parallel multi-scale simulations, crack propagation for example, has been achieved in the LibMultiscale framework and other ScAlApplix parallel tools available for technology transfer.

*Sparse linear solvers:* The team has developed PaStiX a scientific library that provides a

high performance solver for very large sparse linear systems based on direct and ILU iterative methods, providing first sparse linear solver using MPI/Thread for SMP nodes and multi-core machines. The technology readiness level is demonstrated by a successful integration of the sparse linear solvers in real industrial simulation at CEA.

*Graph partitioning and sparse matrix ordering*: The problem has been addressed by the means of a new parallel implementation of the widely used SCOTCH graph partitioner. An important value of this work was to remove the bottleneck of ordering large sparse matrices and graphs. The development of a dynamic partitioning capability is a good step in the roadmap towards the application to parallel adaptive re-meshing.

The future orientation proposed by the project-team, is to split ScAlApplix in two separated projects, namely:

- high order, massively parallel tools for complex physical problems of essentially hyperbolic type;
- methodological studies and creation of software tools for very large-scale 3D simulations (frontier simulations) in cooperation with CERFACS.

The expert panel strongly recommends that the multidisciplinary competence in applied mathematics and computer science be maintained in the new projects in order to keep the benefit of what has been developed as experience in ScAlApplix.

***Adequacy to INRIA's Scientific Strategy***

The project is in line with INRIA Strategy 2008-2012, to which it contributes in modelization, simulation and optimization of complex dynamic systems, using multi-physique and multi-scale mathematical models.

The cooperation between INRIA and university and CNRS laboratories is well supported by the project.

***Industrial Transfer and Partnerships***

Industrial transfer is taking place via PhD thesis in collaboration with industrial partners like CEA, and EDF. The collaboration with industrial partners consists of delivering INRIA tools under open-source, and providing support for the improvement of performance of industrial codes.

The greatest successes in technology transfer were obtained with CEA using LibMultiscale and PASTIX within ODYSSE code. The increasing use of SCOTCH as partitioner is reinforced by the parallel version that is newly available.

Fluidbox and the RDS scheme are for the moment too confidential to allow a proper industrial transfer. Cooperation with CERFACS and ONERA would help to get quicker access by industry.

The overall amount of external funding has increased during the last year, with fewer industrial contracts, and more funding from European projects and national research

agency (ANR).

***Manpower and Means***

The team has developed during the period 2004-2007 up to 29, with an equivalent number of PhD compared to Professors and Associate. A better focus is expected from the split of ScAlApplix into two separated projects.