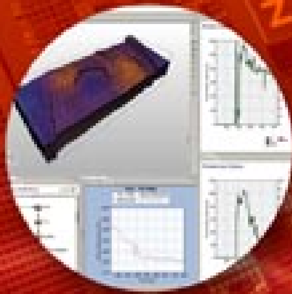
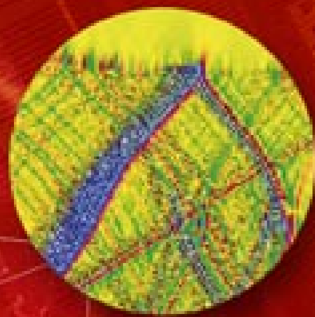
La Recherche

JUNE 2007 - SPECIAL ISSUE IN PARTNERSHIP WITH BULL AND INTEL





Supercomputer Challenges


intel Leap ahead™
MULTIPLY ENERGY EFFICIENCY
AND MAXIMISE COOLING.
THE WORLD'S FIRST QUAD-CORE PROCESSOR FOR MAINSTREAM SERVERS.
The new Quad-Core Intel® Xeon® Processor 5300 series delivers up to 50% more performance* than previous Intel Xeon processors in the same power envelope. Based on the ultra-efficient Intel® Core™ microarchitecture it's the ultimate solution for managing runaway cooling expenses. Learn why great business computing starts with Intel inside. Visit intel.co.uk/xeon
intel Xeon® inside™
*Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Visit intel.com/performance ©2007 Intel Corporation. Intel, the Intel logo, Intel Leap ahead, Intel Leap ahead logo, Intel Xeon and Xeon Inside are trademarks of Intel Corporation in the United States and other countries.

High performancecomputing at the service of **competitivity**

A year ago, *La Recherche* compiled a remarkable survey on high performance computing, a major strategic issue for numerous researchers and engineers employed by public bodies and private companies who could observe on a daily basis the increasingly crucial role played by this type of computing in their areas of activity. Given the unanimous view of our laggard status as compared with other countries, whose ambitions are similar to ours in terms of scientific, social and economic development, the survey invited our country to take action in this area.

What stage have we reached today?

2006 was marked by some major advances. New computers now offer levels of processing power ranging from just a few teraflops to tens of teraflops.(1) These machines can be used for either general purposes in the context of basic research or in various specific scientific disciplines such as meteorology, defence, energy etc. A non-commercial partnership called Genci,(2) has also just been established by the government to coordinate policies for the equipping of large national public scientific computation centres. These initial steps are rapidly giving rise to others advances.

In late 2006, the European Commission launched a consultation process with a view to creating a network of high performance computing infrastructure at European level. Its objective is to provide the European research community with high performance computing resources comparable to those enjoyed by their American and Japanese counterparts by 2010. This proposal is a welcome one. It is the result of projects initially developed by Germany, France and Great Britain and, on a broader basis, since last June on the initiative of the European Commission in the context of HET, the High Performance Computing in Europe Taskforce.(3) The latter recently confirmed, moreover, that together with theory and experimentation, as identified by Michel Serres, numerical simulation is without doubt the "third pillar of research."

The conditions necessary for success

Along with the European Commission and the national bodies, we firmly believe that success in the years to come will depend on the creation of large computation centres, equipped with extremely powerful machines operating on the level of petaflops,(4) and, beyond this, acting as driving forces of scientific excellence facilitating close and sustainable collaboration between "users," computer scientists and mathematicians. In effect, computer science is evolving rapidly. New computer architectures and new methods are constantly emerging and obliging users to make considerable efforts to adapt – if not significantly update – their approaches to problems. Thus, along with these machines, it is also necessary to develop computing and data grids, centralized and local storage facilities based on consistent modalities.

To do this, industry – in particular – European industry must mobilise to develop computer architectures with immediate targets of the level of 100 or 1000 teraflops. Numerous scientific and technological obstacles must be overcome, including the problem electricity consumption (currently on the level of one megawatt for around ten teraflops!). This will involve renewed research in the areas of micro-electronics, mathematics and computing and the revival of a successful and thriving European computer industry. Finally, the lack of clearly designated third-level education in "high performance computing" in France must also be addressed. The resulting careers must be recognized in both the academic world and the private sector. As illustrated by this special edition of *La Recherche*, the consensus with regard to the outlook in this area is largely shared and the effort that needs to be made has been identified. What is important now is to make every possible attempt to translate these ideas into action.

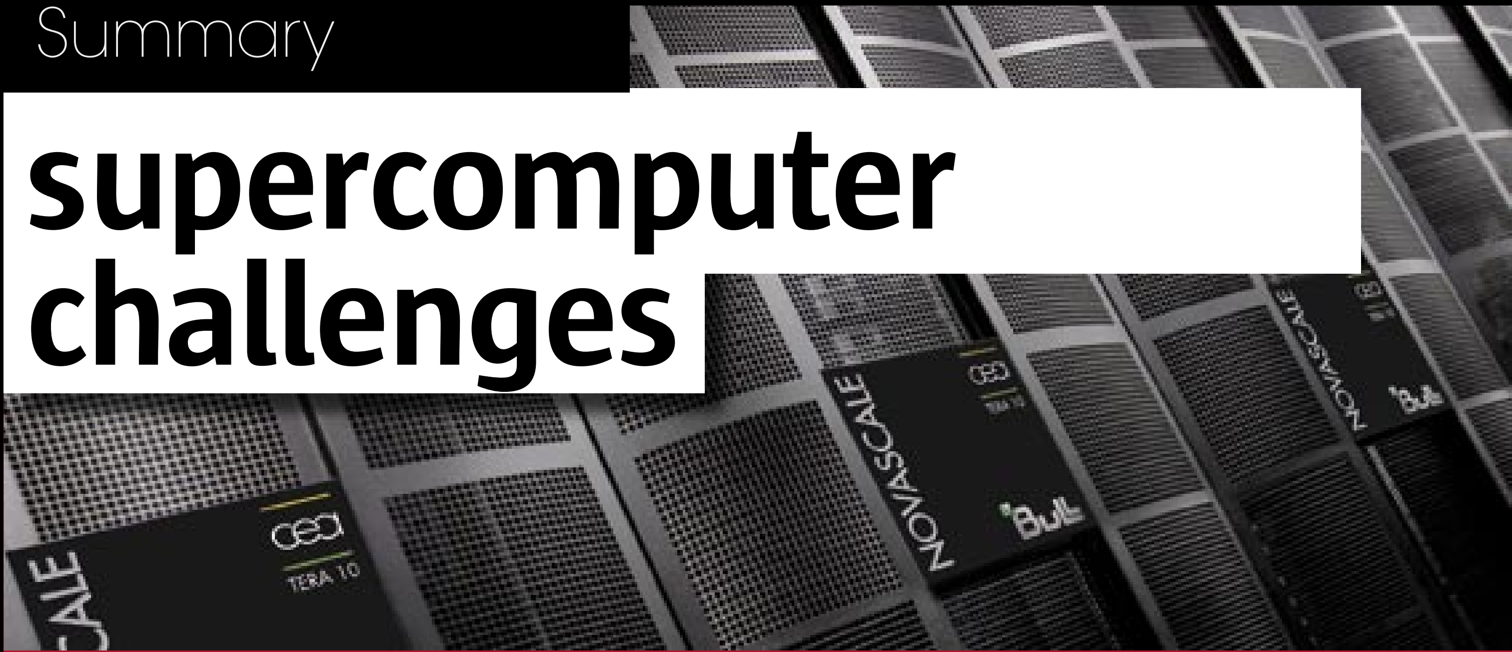
© L. GODART/CEA

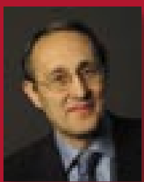
Bernard Bigot
High Commissioner of the French Atomic Energy Commission (CEA)

(1) Teraflop: a trillion operations per second
(2) Genci: Grand équipement national pour le calcul intensif ("Major National Infrastructure for High Performance Computing")
(3) HET: High Performance Computing in Europe Taskforce: http://www.hpcineurope taskforce.eu/
(4) Petaflop: a thousand trillion operations per second

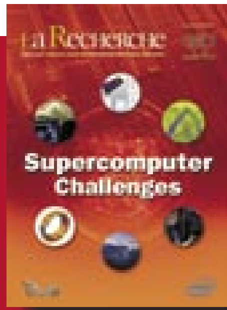
Summary

supercomputer challenges





La Recherche
L'ACTUALITÉ DES SCIENCES



Société d'Éditions Scientifiques
Bernard Wouts, Publishing Director
Stéphane Khémis, Head of the Editorial Committee
Olivier Postel-Vinay, Advisor to the Management

74, avenue du Maine - 75014 Paris
Tél. : 01 44 10 10 10
e-mail editorial: courrier@larecherche.fr

SUBSCRIPTIONS/BACK ISSUES/BINDING
E-mail address: larecherche@presse-info.fr
La Recherche Service Abonnement
B 604, 60732 Sainte–Geneviève Cedex
Tarif France : 1 an 11 nos, 52,60 € ;
1 an 11 nos + 4 hors-série, 72 €
Tarif international : nous contacter.

Suisse : Edigroup, Case postale 393, 1225 Chêne Bourg.
Belgique : Edigroup, Bastion Tower-Etage 20,
place du Champ-de-Mars, 5. 1050 Bruxelles
Canada : Express Mag, 8155, rue Larrey, Anjou Québec H1J 2L5.

Scientific Director: Jean-Michel Ghidaglia
Editor-in-Chief: Aline Richard
Assistant: Juliette Quenin (44 33)
Deputy Editor-in-Chief
Luc Allemand (44 23)

Editor-in-Chief for supplement: Isabelle Bellin
Art direction and graphics: Philippe Lewit
Illustrations for supplement: Julie Épessé
Layout: Carol Müller
Editorial Secretary for supplement: Stéphane Demazure
Design and Production of supplement:
Éditorial et Régie :
22, rue Chauchat 75009 Paris - Tél. 01 53 34 12 22
Collaborating on this issue:
Isabelle Bellin, Didier Gout, Léo Gerat,
Dominique Ritman, Anne Lefèvre-Balleydier,
Pierre Vandeginste, Yves Sciama.
English Translation of supplement: Emendo Bioscience Ltd
- info@emendo.co.uk

Managing Director: Frédéric Texier
Direct Marketing and Subscriptions
Director: Virginie Marliac (44 16)
Chargé du marketing : Arnaud Cobo
Distribution (Distributors/Libraries)
Céline Balthazard (44 11) n° Vert : 0 800 30 76 02
Manager: Isabelle Parez (44 95)

Accounts: Marie-Françoise Chotard (01 44 10 13 43)
Bookshop Distribution: DIF'POP'
Tel.: France (0)1 40 24 21 31 Fax (0)1 40 24 15 88

Head of Project Development: Grégory luneau
Webmaster: Lymédias (01 53 36 06 01)
Production: Christophe Perrusson (44 64)

ADVERTISING: Le Point Communication (44 54)
Head of Advertising (Scientific market)
Laurent Allègre (44 57)
Commercial and Techhnical Assistant
Françoise Hullot (f.hullot@interdeco.fr)

Special issue of La Recherche Jurisdicion
Printed in Italy. legal deposit 2nd quarter 2007

The race for processing power: the

Whether in response to the needs of research or industry, today, simulation is frequently equated with supercomputers. Thirty years after the advent of supercomputing, the landscape has changed in terms of both hardware and software architectures. And this is just the beginning ...

Pierre Leca is Director of the Department of Simulation and Information Sciences at CEA/DAM (Military Applications Division, French Atomic Energy Commission), Île-de-France

A new scientific discipline was born during the 1970s with the arrival of CRAY computers: HPC or High Performance Computing – a discipline at the crossroads between computer science, physical and mathematical modelling. Since then, HPC has assumed an essential role in the majority of sectors, both in research and industry and now represents a crucial and strategic sector for some countries which are adopting long-term policies in this regard.

A worrying state of affairs for Europe

Perusal of the Top 500, the list of the 500 most powerful computers in the world, which has been updated every six months since 1993, says it all about the overwhelming dominance of America in the area of high performance computing: in November 2006, the United States had over 60 percent of supercomputers and the same share of total performance. As opposed to this, Europe only had around 20 percent of total computer power.

This fact is a matter of concern for the competitivity of our research and industry, all the more so given that industry's share of this computer processing power has also been increasing steadily, rising from 20 percent in 1993 to almost 50 percent in 2006. In fact, high performance computing is less and less the preserve of the public sector (research, defence, government bodies etc.) than was the case in the recent past.

However, the most amazing aspect of the developments in relation to the Top 500 is the exponential growth in processing capacity represented by all of these 500 computers, which has never been denied from the outset. In effect, installed processing power has been increasing over tenfold every four years, rising from 1 teraflop[1] (or 10^{12} operations/sec) in 1993 to over 3 petaflops[2] (or 3 x 10^{15} operations/sec) in 2006. This exceeds the increase expected on the basis of Moore's Law (a "law" concerning microprocessors which predicts that the number of transistors on a chip will double every 18 months; we return to this topic below). Thus, based on the increase in processor frequency, it should only have led to an increase of a factor of eight every six years.

So where does this acceleration in the power of computers, which exceeds the power of the microprocessors of which they are made, come from? From the development of parallel architectures, in other words the interconnection and coordination of several processors within the computer. Thus, the average number of processes per Top 500 computer has increased from around 100 in 1993 to well over 1000 in 2006.

in November 2006, the United States had over 60 percent of supercomputers and the same share of total performance

A deluge of data

However, the growth in processing power among the Top 500 computers is not the only factor that needs to be considered to understand all of the developments in the area of high performance computing. It does not explain, in particular, the phenomenal increase in the streams of data to be processed. Two new factors have been at work here for some years now. First, the reduction in the cost of information storage devices: today, a terabyte disk (10^{12} bytes) costs less than EUR 1,000, thus making the installation of tools enabling the storage of several petabytes (10^{15} bytes) economically viable. To put this in context, the digitisation of the National

three revolutions in progress

© C. DUPONT/CEA

Library of France, Bibliothèque François Mitterrand, would only represent a storage volume of between 30 and 40 terabytes. Second, the evolution of high performance computing applications has resulted in the consumption and production of considerable volumes of data. By way of example, the simulations carried out by the CEA/DAM's (Military Applications Division of the French Atomic Energy Commission) Tera-10 supercomputer, which implements three-dimensional models of non-stationary phenomena, have been producing over 10 terabytes of data per day for one year now. In another area of use, the analysis of data produced by large instruments, such as the BarBar detector at the Stanford Linear Accelerator Center (United States), results in the processing of 1 petabyte of data.

Moore's Law and computer architecture

Nonetheless, this increase in computer performance also has consequences that constitute an increasing cause for concern: it has led to an increase in the electrical power consumed and the heat emitted

The Tera-10 supercomputers are developing processing power in excess of 50 teraflops. This machine is constructed using 4352 Intel® Montecito dual-core processors linked by a Quadrics high-performance interconnection network.

by the machines. A micro-processor of the latest generation can consume up to 130 watts which translates into heat release equivalent to that of a large electric bulb over a few square centimetres! This increase in electricity consumption and, hence, heat release per chip cannot continue in the absence of a mechanism for the evacuation of the heat. For some years now this has prompted – with far-reaching consequences – the fundamental re-examination of the architecture of the microprocessors at the heart of all of the systems.

In effect, according to technologists, if Moore's law is not reviewed for the coming decade, then the consequences must be reviewed. Up to now, the momentum of the improvement in performance was created through the increase in clock frequency which is associated with the execution of instructions and implementation of cache memories. Today, this "free" increase in performance is no longer appropriate. Over the past two years, the transistor fabrication process has been reduced through the simultaneous implementation of "multiple cores" within one and the same microprocessor (i.e. several processors on the chip) and the capacity to manage several execution streams simultaneously ("hyperthreading mechanism"). Thus, it should be considered that in the course of time, Moore's Law will be reflected more in terms of the increase in the number of cores than in the increase in frequencies.

© C.DUPONT/CEA

Pierre Leca, Director of the Department of Simulation and Information Sciences at CEA/DAM (Military Applications Division, French Atomic Energy Commission), Île-de-France.

The revolution of parallelism at every stage

The Top 500 list has not included any monoprocessor computers since November 1997 and since then the entire high performance computing community has had to take the issue of parallelism into consideration. Ten years later, with the widespread availability of multi-core microprocessors in all computer devices, from the PC to the supercomputer, it is the entire software community that must now take this new paradigm into account.

Moreover, the non-linearity in the relationship between frequency and electricity consumption could prompt a further increase in this dynamic. The example of IBM's BlueGene architecture demonstrates this trend pushed to the extreme. The American manufacturer opted to interconnect a very large number of processors that are less powerful than is possible using the available technology. In effect, BlueGene has no fewer than 130,000 700 megahertz processors which enables it to attain over 360 teraflops of processing power while only consuming 1.5 megawatts of electricity.

A micro-processor of the latest generation can consume up to 130 watts

The Tera range of computers installed at the CEA/DAM (Military Applications Division, French Atomic Energy Commission) as part of the Simulation Programme constitute a more representative example of generalist supercomputers: the Tera-1 computer installed in 2001 had 2,500 mono-core microprocessors providing 5 teraflops of computing power. To attain a performance level 12 times higher than this, the Tera-10 computer built by Bull and installed four years later uses almost 5,000 bi-core microprocessors or almost 10,000 computing cores.

The software and algorithm revolution

All the evidence points to the fact that it will be possible to build a computer with 1 petaflop of processing power before 2010: this machine will use between 100,000 and 1,000,000 cores, depending on the compromise that will be made between microprocessor frequency and electricity consumption. Moreover, these tens of thousands of cores will be distributed in the computer on a hierarchical basis, first within the microprocessor, then in a shared-memory multiprocessor and, finally, within clusters of multiprocessors.

By way of example, here is the possible configuration of such a supercomputer in 2010: based on a quad-core microprocessor running at 2 gigahertz, capable of executing four instructions per cycle, by incorporating 16 of these microprocessors in one multiprocessor, it would be possible to attain a performance rate of 0.5 teraflops. The interconnection of 2,000 of these multiprocessors would make it possible to construct a 1 petaflop computer using 128,000 cores. Of course, what is involved here is a potential processing power which can only be exploited in conjunction with the

adaptation of algorithms and numerical methods.
As has always been the case, these new hardware architectures will impose an adaptation of the corresponding software. Let us recall the developments triggered over the course of the 1980s by the arrival of the first vectorial computers. For some, the spread of parallelism throughout the IT pyramid represents the most significant change since the object-oriented programming revolution. For the experts in scientific computation who will be in the eye of this particular storm, the consideration of multiple factors such as massive parallelism, hierarchy within the computer and the management of data streams, will necessitate the development of a number of approaches from the perspective of numerical methods, programming models and software architectures.
How much could it cost to own a supercomputer of this kind based on a duration of use of five years? An acquisition cost of EUR 80 million may be assumed. Electricity consumption can be estimated as follows: first and foremost, the 32,000 microprocessors will account for almost 4 megawatts of electricity. To this must be added the electricity consumed by the memory and disks which is estimated at around 3 megawatts if a balanced architecture is required. Thus, a total of 7 megawatt which should be multiplied by a factor of two to account for the electricity consumed by cooling equipment and other ancillary devices.
At current electricity prices, this represents a cost of EUR 7 million per year, or EUR 35 million over the five-year period of use. An estimated maintenance budget of EUR 25 million, based on the acquisition cost, may also be added to this. Finally when other ancillary costs are included, such as the cost of the necessary adaptation of the computer room and other infrastructure, it may realistically be estimated that the simple act of running such a supercomputer over a five year period costs nearly as much as it does to acquire it in the first place.

The revolution in the scientific communities: join the race or get left behind

This “cost of ownership” as well as all the other changes this type of computer is causing in the context of the high-performance computing community, in particular the integration of computing power with the management of data flow and the need to capitalize on the knowledge in enduring software solutions (“the computers come and go, the software stays”), is prompting an examination of the installation of these large computational resources, their use and administration and that of very large computer infrastructure. It is essential that the user communities organize themselves in pluri-disciplinary teams around this infrastructure.
Moreover, the advances promised by high performance computers in the area of simulation can only be achieved if the complexity of this phenomenon is mastered: i.e. modelling complexity, the complexity of mathematical and algorithmic methods, computational complexity, the complexity of verification, validation and software production techniques. What the leaders in the field will be doing tomorrow is to establish this kind of integrated approach and this is the real challenge facing computing communities in Europe today.
This task necessitates collaboration between engineers and researchers specializing in modelling, mathematics, software engineering and computer architecture. In effect, the field of high-performance computing is located at the intersection of several disciplines, an area in which the existence and growth of an industrial and technological base are of fundamental importance for the anticipation of future developments.
The success of the project resulting in the installation of the Tera-10 computer at the CEA/DAM, the increasing power of the Ter@tec association (see article on Ter@tec p. 40), the momentum of the System@tic centre of excellence in competitivity and the inclusion of the topic of high performance computing in the 7th Framework Programme for Research of the European Union have created a context that is favourable to launching a virtual cycle involving research, development, innovation and industrial objectives. Based on the example of the United States, China and Japan, Europe now has all the elements necessary to meet the new challenges posed by supercomputing. **P. L.**

The Tera-10 supercomputer, built by Bull, was delivered to the CEA/DAM in late 2005.

Flops : acronym denoting Floating Point Operations Per Second, unit of processing speed of a processor expressed in the number of operations per second.

(1) One teraflop: a trillion floating point operations per second.

(2) One petaflop: a thousand trillion floating point operations per second.

Websites
Top 500 : www.top500.org
The latest news on HPC (High Productivity Computing): www.hpcwire.com; www.hoise.com

Jean Gonnord : "2007 will be the year of high performance computing in Europe"

Interview with Jean Gonnord, Director of the Numerical Simulation and Information Technology Project at CEA/DAM Île-de-France. Will France and Europe manage to catch up in the area of high performance computing?

La Recherche. First, let's go back to 2006, an eventful year in terms of high performance computing at the CEA....

Jean Gonnord. A year punctuated with success stories: the Tera-10 supercomputer, the most powerful in Europe (60 teraflops), the first to have been designed by a European manufacturer (Bull), was delivered on time at the end of 2005 and rendered operational in advance of the schedule in June 2006. Ring-fenced for the French defence programme, the Tera-10 immediately assumed 5th position in the Top 500.(1) Six months later, in the Top 500 of November 2006, the machine was only 7th in the world ranking: this is what the race for computing power is all about. However, our calling is not to beat records but to compute! We evaluated our needs in terms of computing power on the basis of the progressive improvement expected from our models, i.e. a factor of ten every four years. The development programme for our computers is based on this progression: Tera-1 (5 teraflops) installed late 2001 by HP was replaced by the Tera-10 (60 teraflops). Tera-100 (0.5 petaflop) will be installed in 2010. We plan to have a capacity of 50 petaflops in 2018.

L. R. With the exception of the Japanese Earth Simulator computer, the world's largest computers are clearly reserved for defence purposes. Will this trend continue?

J. G. In effect, following the decisions to stop nuclear tests, the defence programmes were the first to take R&D into the area of high performance computing with large-scale numerical simulations. However, the majority of large projects for 2010 now involve civil applications and the need for high performance computing capacity here is exploding. Supercomputers are already one of the main keys to global competitivity, both for research and industry. Furthermore, the CEA has been implementing a policy of opening access to its resources for several years now. This has resulted, in particular, in the creation of the Bruyères-le-Châtel (Essonne) High Performance Computing Centre which brings the CEA's expert teams together with defence resources (Tera-10), civil resources shared between the CEA and its partners within the framework of the CCRT(2) and the necessary test resources. Similarly, before the Tera-10 became entirely ring-fenced for defence requirements, we made Europe's most powerful computer available to teams working on major scientific and industrial challenges which are breaking new ground on a global scale and are the subject of this special edition of *La Recherche*. Finally, less than one year after the installation of Tera-10, a 43-teraflop computer designed by Bull with similar architecture and comparable processing power is currently being installed at the CCRT. It should be ready for operation in September 2007.

L. R. Ter@tec science park, the emblem of this synergy between defence, industry and research, was established around the scientific computation centre. What projects are currently under way there?

J. G. This is basically the best illustration of this policy of openness. The first success, the precursor to these collaborative endeavours, produced the NovaScale server in 2004. This server is marketed by Bull and a version of it is used in the Tera-10. The science park brings together research laboratories and industries, all actors and sponsors involved in numerical simulation to promote high performance computing applications and to develop the future technologies of powerful computers (see article on Ter@tec p. 40). Today, it has over 40 partners. The CNRS, the Inria, Airbus, Total etc. joined us this year. Depending on their objective, the projects proposed within Ter@tec are supported by the CEA and the partners involved, by the French National Research Agency (*Agence nationale pour la recherche*, ANR(3)), by the System@tic(4) centre of competitive excellence, of which Ter@tec is an essential element, or by the European Union. Three of the five major projects approved by System@tic are projects that are being carried out on the initiative of partners of Ter@tec.

L. R. 2006 was a turning point for the scientific computation community at national level too ...

J. G. That's undeniable. It witnessed the second call for projects from the "high performance computing" sector of the completely new ANR: 14 projects representing a total value of EUR 8 million were approved in 2006. These projects are spread throughout France and three of them originated in Ter@tec. This represents a very important, indeed essential, boost. It is only regrettable that none of the projects involve the development of hardware technologies and architectures. This sector has been abandoned for over 20 years and presents a gap that urgently needs to be filled to regenerate essential skills in this area. Other good news involves the announcement last July by the government of the establishment of Genci (*Grand équipement national pour le calcul intensif*), a non-commercial partnership


© Jean-Michel RILLON

which will double the investments hitherto allocated to high performance computing, coordinate these investments and be France's sole interlocutor at European level. (see article on Genci, p 41).

L. R. What is the current situation in Europe?

J. G. Following its disappearance from the landscape for ten years, we welcome the return of high performance computing to European research objectives with the launch of the EU's Seventh Research Framework Programme (2007-2013). Europe has set itself two major targets: to get back to the level of the major industrial countries in the terms of providing access to processing power and to master the technologies associated with these supercomputers. To respond to this technological challenge, we have established an alliance called TALOS (Technologies for Advanced Large Scale Open Supercomputing) between Bull, Quadrics, Intel's German subsidiary, the CEA and HLRS (High Performance Computing Centre Stuttgart, one of Germany's largest centres of scientific computers and expertise at the university of Stuttgart). In terms of the provision of access to processing power, Europe intends to participate in the financing of petaflop computers by 2009/2010. CEA's Scientific Computation Centre is, of course, an obvious candidate for the accommodation and management of infrastructure of this kind.

By Isabelle Bellin

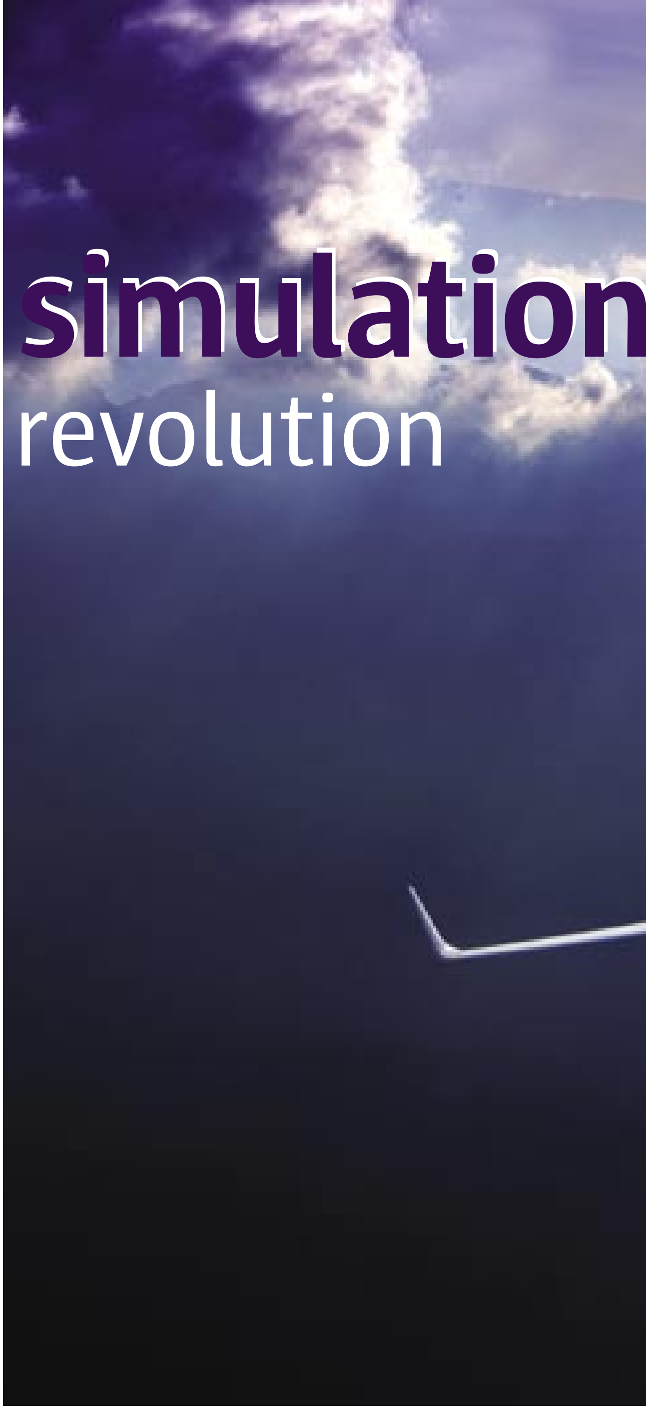
(1) Top 500: ranking of the 500 most powerful computers in the world. ww.top500.org

(2) CCRT: Centre de calcul pour la recherche et la technologie (Computation Centre for Research and Technology). CEA's partners in CCRT are EDF (Electricité de France), the Safran group through its subsidiaries, Snecma, Turbomeca and Techspace Aéro.

(3) ANR: Agence nationale pour la recherche (French National Research Agency) http://www.agence-nationale-recherche.fr/

(4) System@tic Paris Région is a centre of excellence for competitivity established in 2005: its mission is to give the Paris region a global visibility in the area of complex systems on four priority markets: automobile and transport, security, telecommunications and design engineering. http://www.systematic-paris-region.org/

Aerodynamic simulation
a never-ending revolution



© Paul BOWEN / Dassault Aviation

The Falcon 7X, the top of the range model by Dassault Aviation, is the first corporate aircraft to have been fully optimized using numerical computation. Modelling carried out on the Tera-10 supercomputer enabled the French aviation company to test the capacities of future numerical simulation resources.

Didier Gout
is a scientific journalist

"Due to the doubling of computing power every 18 months, aerodynamic design is more accurate and faster with the help of a supercomputer like the Tera-10. We predict that in a few years' time this kind of power will have become a standard tool for us": at Dassault Aviation, there is no holding back with the compliments for the Tera-10, the supercomputer delivered to the CEA in 2006, and the possibilities that a machine with this kind of processing power provides. It was used in one of the "grand challenge" supercomputer operations in the first half of 2006. The objective was to evaluate the possibilities for the improving the aerodynamic design of an entire aircraft using this kind of computer. The aircraft manufacturer's latest baby, the Falcon 7X, a new top-of-the-range corporate aircraft, for which over 125 orders have already been received and whose certification is planned in spring, has already benefited from the use of the most cutting-edge on-board computers (IBM SP and Bull NovaScale): simulations contributed to the remarkable aerodynamic performance achieved by this trijet which is capable of 11,000 kilometres non-stop flight.

The use of high-performance computing is not actually new in the aviation sector, in particular at Dassault, which has always played a leading role in this area thanks to its CATIA software, which was developed by the company's research division. And the performance of the company's corporate aircraft, in relation to fuel consumption in particular, is the direct product of aerodynamic computations and the expertise of the people who carry them out.

Numerical simulation has revolutionized the aerodynamic design of aircraft

High performance computing has been making strides at all levels of aviation design, in particular in the aerodynamics sector, for some years now. The complexity of the flow of air around an aircraft and the levels of performance required necessitate the use of very advanced simulation software. This software resolves the classical fluid mechanics equations (Navier-Stokes equations) through

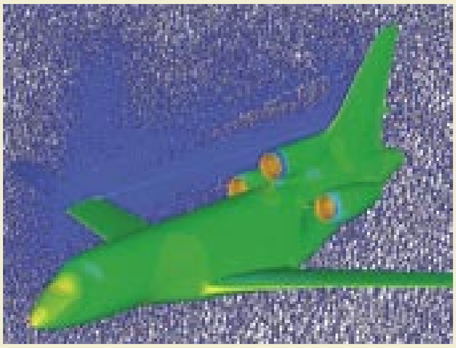
the modelling of turbulence phenomena. The high level of accuracy required necessitates a high level of processing power so that the results of the computations can be obtained within an acceptable period (around a day), compatible with the deadlines in the aircraft design cycle. The combination of the detailed modelling of the physics of airflow and the high levels of processing power have facilitated a revolution in aerodynamic design. The design of the Falcon 7X wing was an iterative process involving the intensive use of automatic optimization and numerical simulation software. These tools are used to generate – at the lowest possible cost – several dozen different wings, the computer analysis of which enables the identification of the best possible compromise between the different performance objectives. The best component shapes are then tested in the wind tunnel so as to validate the results. The eventual gaps between the tests and calculations are then analysed as a back-up to this iterative process. The combined use of the most comprehensive theoretical analysis possible and their immediate comparison with the most modern experimental findings makes it possible to considerably reduce the total number of development tests carried out prior to the ultimate selection of the component shapes to be used in the aircraft.

"These tools are used to generate several dozen different wings"

NUMERICAL COMPUTATION OF THE TECHNICAL CHARACTERISTICS OF A FALCON 7X

Numerical computation makes it possible to optimize the shape of aircraft wings (e.g. here of a Falcon). The area of low pressure characteristic of aircraft lift above the wing is marked in red. Visualization of the wake enables the analysis of the aircraft drag. The evidence of vortices originating from the wing tip can also be noted.

Numerical simulation of a Falcon in landing configuration: the leading edge slats in front and the flaps at the back can be identified. The success achieved in the design of such mechanisms enables the aircraft to land on short runways. The colours indicate the areas of pressure on the wing.



During aerodynamic simulations implemented using the Tera-10, the fluid space that surrounds the aircraft was subject to very detailed discretisation with 115 million elements.

How important is the use of a computer like the Tera-10? According to Dassault Aviation: "*This level of processing power means that the aerodynamic design of an aircraft can be carried out using simulation right up to its final form, thus a process which currently takes over a year can be reduced to a period of a few months and the level of accuracy and precision is also greater.*" This is conclusion is based on the completion of 200,000 hours of computation in just 200 hours thanks to the thousand parallel processors available on the CEA's Tera-10. Given that, traditionally, a single complete design and validation cycle for an aircraft's form requires several dozen of these computations, this represents the equivalent of around a hundred computations.

Operations that are ten times more complex can be carried out using the Tera-10

Using the Tera-10, these computations were carried out based on a more detailed discretisation involving over 100 million elements; current simulations use around 30 million elements as compared with 10 million at the beginning of the design of the Falcon 7X. Based on seven values (e.g. pressure, speed, temperature, kinetic energy, turbulence dissipation rates etc.) to be defined for each node of a discretisation element (which has four), today, 140 million equations need to be resolved for the entire field of aerodynamics around an aircraft, as compared with 14 million a few years ago. Thus, the equation levels are ten times more complex than the Tera-10 computer can cope with.

In reality, the engineers from Dassault Aviation evaluated the drag of the Falcon 7X at its cruising speed (this is referred to as the "cruising point", the most traditional flying point, 0.8 Mach for the Falcon 7X) using the Tera-10. This made it possible to tailor the information requirements corresponding to this problem area. The company now intends to apply this processing power to the improvement of the design of wing-flap devices (flaps and slats) used to land and take off on the shortest possible runways by making use of this more refined discretisation.

Dassault Aviation can also conceive of other applications which may or may not require more discretisation but do require longer processing times. This is the case with the computation of "unsteady" phenomena which vary as a function of type; these are typically the air flows observed at the extremes of the field of flying, in particular for rates of descent or maximum speed. This type of computer architecture will eventually also be used in the automatic optimization of component shapes, a software process enabling the research of shapes. This technology, which is currently being developed, should be applied to the next generation of aircraft. More generally, encouraged by its positive experience with the use of the aerodynamic computation sequence on the vast parallel architecture of the Tera-10, Dassault Aviation can envisage the intensive use of this level of computation in the years to come in both the company's civil and military products, which are being developed by its research department, i.e. the Falcons of the future and manned fighter aircraft (for example, manned or unmanned fighter drones, a prototype of which should fly by around 2011 as part of the European nEUROn project). Thus, Dassault Aviation's future products should benefit from the progress in aerodynamics and the economic gains made possible by the never-ending revolution of the supercomputers. **D. G.**

Xedix,
a system for managing ... masses of data

In the beginning what was involved was the organization of over a billion bytes of data relating to 40 years of nuclear tests. In the end, Xedix is a new type of database management system created to accommodate vast "data warehouses."

Storage Tek data storage robot at the CEA/DAM centre at Bruyères-le-Châtel.

Léo Gerat is a scientific journalist

February 13th 1960, Reggane oasis in the Algerian Sahara: all hell breaks loose on earth. France has just carried out its first nuclear test. January 27th 1996, Fangaaua atoll in the Pacific: a French nuclear warhead exploded for the last time ... in reality. Since this date, France has tested its nuclear weapons using numerical simulation and to this end has acquired powerful computational resources. However, simulation has to remain within the bounds of the possible, it is only valid because it runs "models" which have themselves been validated by and are "rooted" in experimental measures. In other words, the information produced by the 210 empirical tests carried out over a period of almost 40 years is irreplaceable.
Such is the origin of the Xedix project at CEA/DAM (Military Applications Division of the French Atomic Energy Commission), which was born of the desire to preserve 40 years of nuclear experience and to put it at the disposal of weapons designers who will henceforth be working on the basis of simulation. Forty years of documents of every type, experimental measures, photos, films and videos, thought to represent a total volume of 400 gigabytes (Gb)[1]. Didier Courtaud and Pierre Brochard, who carried out this ambitious operation now know

that the volume involved was 1.3 terabytes (Tb).[2] Given that what was involved was a mass of information of a documentary nature, for the most part, it was decided at a very early stage to opt for a system based on SGML. Standard Generalized Markup Language (SGML) is an ISO standard which has been in existence since 1986 and is widely used for the standardization of technical documentation. Formally, SGML is a powerful metalanguage[3] which enables the creation of mark-up languages, i.e. languages that enable the enhancement of files with the help of “tags” for structuring (title, chapters, notes etc.) and formatting (italics, justification etc.) purposes. The best known of these languages is HTML which is used to compile internet web pages.

This mass of information then needed to be made accessible and usable. To do this a “database” needs to be created, for which approved users can formulated “requests” and find whatever they are looking for.

“XML native” databases

At that time (1996), the database paradigm in vogue was the relational model. However, such a large volume of documentary and multimedia information had never been handled in this way before. Although it was not specially adapted to the particular problem in hand, one new concept appeared to be particularly promising, i.e. that of the object-oriented database. Thus, it was initially decided to move in this direction by approaching the O2 project which was born of a collaboration between the Inria (French National Institute for Research in Computer Science and Control) and the Computer Research Laboratory of the University of Orsay.

A few years later, progress was made from SGML to XML (Extended Markup Language), a simplification of the former which is applied like a trail of dust and is seen as promising a brilliant future for all kinds of applications, in particular on the internet. The O2 project ended in 2000. It had become clear that it was not possible to depend on any traditional category of database to resolve the problem and that it was instead necessary to adopt a new approach, i.e. the “native XML database” (NXD), in other words to design a database management system (DBMS) conceived entirely for and with XML. Xedix would, therefore, become an “XML native” DBMS.

From 2001 to 2003, the researchers engaged in large-scale development work, which culminated in the presentation of a complete, autonomous tool capable of managing all kinds of applications, i.e. documentary and multimedia, in particular if they handle large volumes of data. Xedix is capable, among other things, of managing and indexing databases physically distributed on multiple computers.

The world’s largest telescope (Seti@home project in Arecibo, Porto-Rico) detects radio signals from space in real time with the aim of detecting eventual extraterrestrial messages, producing a mass of data for management and thus representing a potential application for Xedix

The moment had arrived to put the baby to the test. This involved the first Xedix test bench, called Xtera 1, implemented in 2003. A fictitious but plausible terabyte database was created and tested on a cluster of 16 PCs fitted with 80 Gb hard disks and coordinated by a master PC. It was then established that the time taken to respond to a request is virtually independent of the size of the database and essentially proportional to the number of XML elements returned in response to the request. In the case of a particularly challenging request (search for a word present in 60 percent of the documents), the average response time was of the order of ten seconds. This initial test demonstrated the capacity of Xedix to manage a terabyte database using hardware architecture of the scope typical found in an SME.

A new test, called Xtera 10, was carried out in December 2005, this time with the bar set at 10 terabytes. The challenge on this occasion was to demonstrate that Xedix can take advantage of high-performance parallel hardware architectures. The TeraNova[4] supercomputer was called on to contribute. Three hundred of its processors were dedicated to a test base occupying 1010 Tb. Again, it was confirmed that the response time to a request is quasi proportional to the number of responses. On this occasion, one response type was obtained in between two and eight seconds. A third test bench is now planned (Xtera 100). This should confirm in spring 2007 that Xedix can manage 100 terabyte databases. To put it in context, this represents three times the volume of data managed by the *Très Grande Bibliothèque* (TGB), the French National Library, in Paris.

(1) Gigabyte (Gb): a billion bytes.

(2) Terabyte (Tb): a thousand billion bytes

(3) Metalanguage: structure enabling the description of languages

(4) TeraNova supercomputer: cluster of Bull NovaScale servers installed at the Ter@tec science park **(see article p 40)**

Processing power: 1.3 teraflops.

(5) Inist: Institut de l’information scientifique et technique ***(Institute for Scientific and Technical Information)*** **attached to the CNRS (French National Center for Scientific Research)**

From post-Mururoa to start-up

Today, Xedix is much more than a project whose objective it is to conserve and provide access to "40 years of nuclear tests." This problem has been resolved. Xedix is now, first and foremost, a product and enterprise that the CEA is preparing to allow to fly under its own steam. This event will be managed by Jean-Claude Sabattier of CEA/DAM, and Pierre Brochard and Didier Courtaud will again be responsible for the technical management.

The DBMS market is growing strongly (8 percent per annum). Gartner predicts a market of USD 15 billion in 2009, of which 10 percent will be accounted for by XML. Xedix aims to take its place in this segment of the market. It is claiming its position as the top industrial implementation of the XML native database concept and can present the results of full-scale testing on a 10 terabyte database, thus firm evidence of something never witnessed before. This kind of tool would be particularly suitable for the management of large collections of documents, but not only this. Collaborative ventures s with the System@tic centre for competitive excellence and the *Institut national des télécommunications* (French National Institute of Telecommunications) have already resulted in a promise to test this potential on an application associated with the "Physiome" international project (see box below) and on access to large video file bases. Beyond clearly documentary or multi-media applications located in the areas of bibliography and media (TGB, Inist[5]), Xedix hopes to attract the interest of sectors such as those working with large scientific instruments (e.g. accelerators, telescopes etc.) that produce considerable volumes of information, the telecommunications sector (which must save the history of every communication) and industries that produce mountains of documentation, such as avionics (those involved in the series of tests carried out on the A380 can only dream of such a tool) and pharmacy (the same situation prevails here with respect to documents describing the history of a medicine up to its official approval). An interesting indicator in relation to this last example is the fact that the FDA (Food and Drug Administration), the American body that authorizes the marketing of medicines, now requires that this documentation be provided in XML. **L. G.**

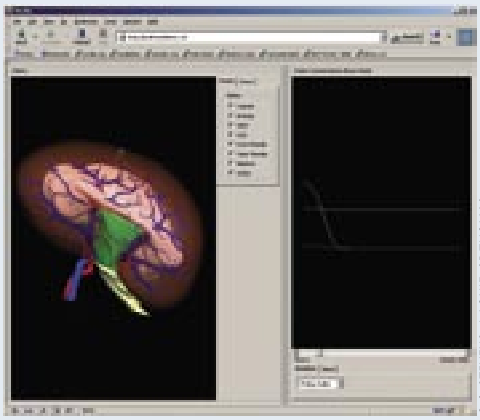
"Xedix is now a product and enterprise that the CEA is preparing to allow to fly under its own steam"

A remarkable application: the physiome

The worldwide project known under the name of "Physiome"[1] is based on an outlandish ambition. Like the "genome" project, which aimed to completely decode the human genetic inheritance, the physiome project aims to provide a computational model of the entire physiology of man. Over time, the researchers involved in this project defined a more reasonable intermediary objective: i.e. the creation of a multimedia global and shared database of the physiology of the human body on all levels. This means that teams dispersed throughout the world can pool their latest discoveries on the functioning of every organ, tissue type, cell type, molecule etc. found in the human body.

In France, Randy Thomas is the global coordinator of the renal Physiome at the Ibisc Laboratory[2] of the CNRS (French National Centre for Scientific Research) and the University of Evry. He manages the QKDB[3] database (Quantitative Kidney Database), the objective of which is to become the "kidney" section of the Physiome project.

As part of the FAME 2 project[4], which is being carried out by the System@tic centre of competitive excellence in the Paris region and in collaboration with the CEA/DAM, Randy Thomas and Fariza Tahi are developing a data warehouse under Xedix for the physiome which will include data of every kind (texts, images, videos etc.) and on all levels, ranging from the genome to the entire organism. **L. G.**



© C. STEVENS, A.LONIE, SR THOMAS

Example of an interface based on a 3D model which should enable interactive access to the information available on the Physiome. Shown here is a virtual rat kidney in zoomable, orientable and removable 3D.

(1) Physiome Project: http://physiome.org

(2) Ibisc: Informatique, biologie intégrative et systèmes complexes (Computer Science, Integrative Biology and Complex Systems Laboratory).

(3) QKBD database: http://physiome.ibisc.fr/qkdb

(4) FAME 2 project: http://www.fame2.org

Managing turbulent combustion phenomena

Parallel computers with the power of the Tera-10 have enabled the simulation of the complete ignition of a helicopter combustion chamber.

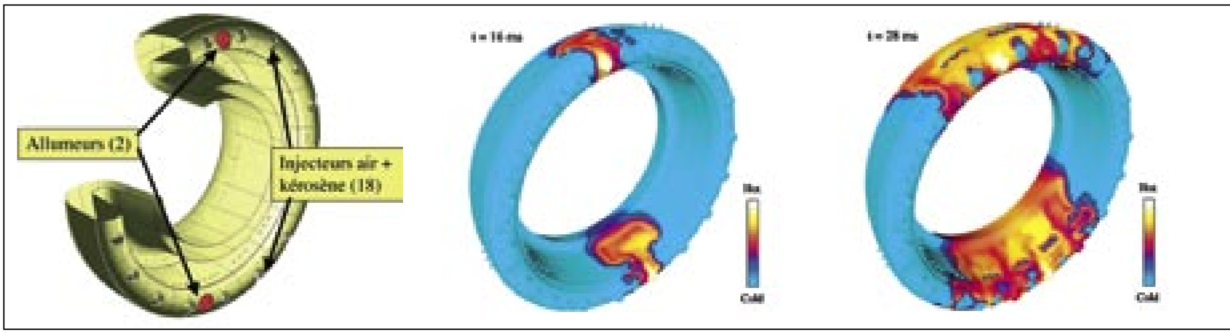
Dominique Ritman
is a scientific journalist

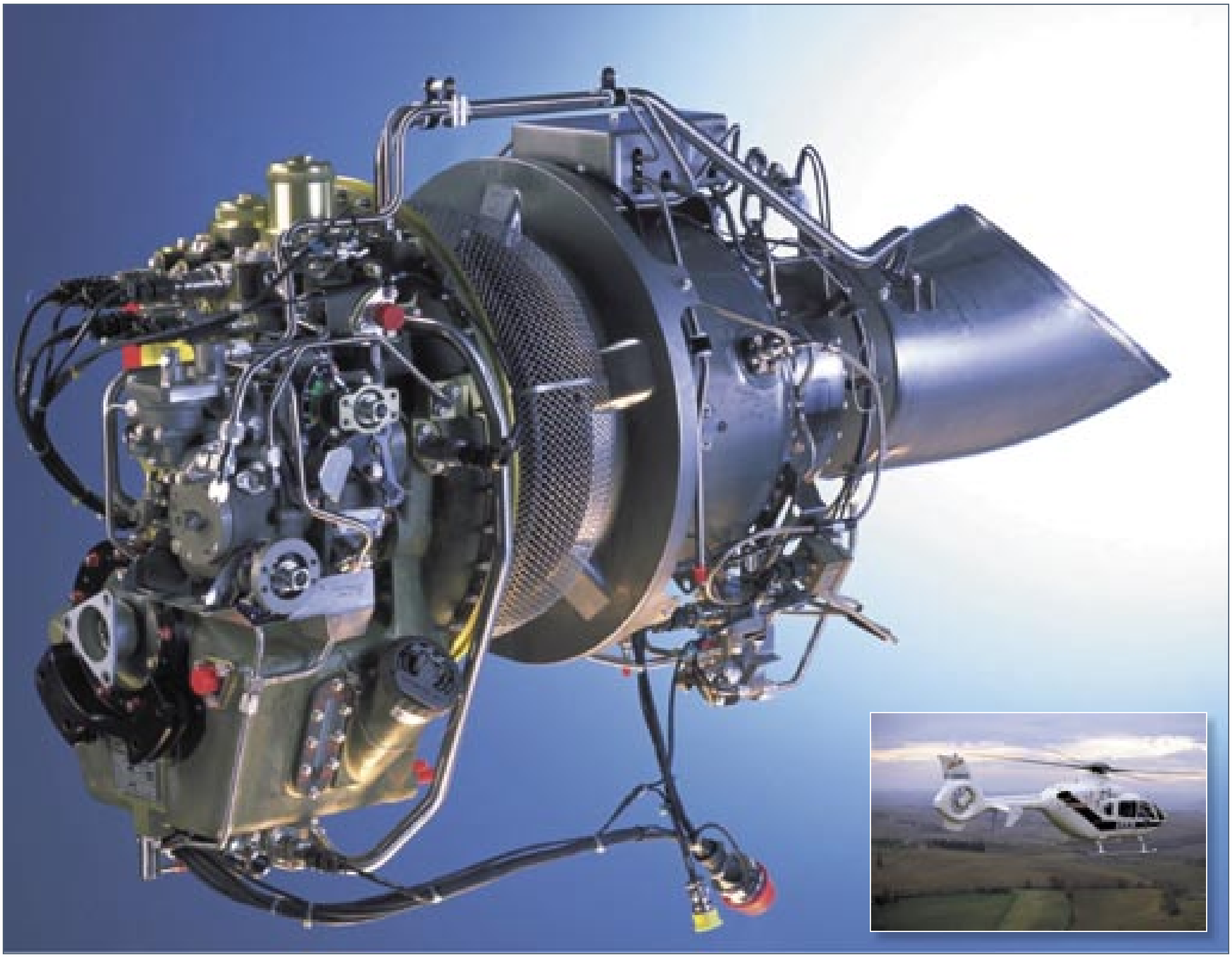
Following in the footsteps of the automobile industry, aviation is one of sectors of industry that has witnessed the most far-reaching changes as a result of the advent of computer simulation and parallel computing. The European Centre for Research and Advanced Training in Scientific Computation (CERFACS) knows a thing or two about this: its shareholders include two global leaders from the aviation and aerospatial sectors, the company EADS and the Safran group (to which Snecma and Turbomeca belong). The design of combustion chambers is one of the priority topics among all constructors of aeronautic engines today. The simulation of the complete ignition of a Turbomeca helicopter turbo-engine, implemented by CERFACS in 2006, represents a global first in this regard.

The increasingly sophisticated technology of combustion chambers should make it possible to respond to the new economic constraints (high price of kerosene) and to anticipate future regulations (reduction of pollutants, such as nitrogen oxide, and greenhouse gas emissions, such as carbon dioxide). Thus, for reasons of cost and design time, numerical simulation has become the mandatory course in this area.

Simulation without over-simplification

The functioning of these chambers is ruled by well-known equations: i.e. fluid mechanics equations (Navier-Stokes equations) coupled with equations relating to the combustion of the different chemical substances present (hydrocarbons, oxygen, nitrogen oxides, carbon dioxide etc.). However, to model such a chamber in its entirety, scientists find themselves having to resolve several million times (corresponding to millions of reputations of computations as a function of time) a system of equations comprising around fifteen variables on 40 million units. This is a particularly complex system which, just a short time ago, could only be resolved at the price of a number of simplifications. The problem, as Thierry Poinsot, CNRS Research Director and CERFACS consultant, explains is that these simplified modellings only provide access to mean values for physical parameters, be it the temperature of combustion gases, pressure etc. Thus, one of the objectives, of industry in particular is to reach a better understanding of the mechanisms that control turbulent combustion, thus making it easier to anticipate every kind of malfunction and, more generally, answer the questions that arise in relation to the





© TURBOMECA

conditions of engine ignition, its re-ignition in the case of unexpected cut out, its cut off and, again, the study of its instabilities.
Scientists used the LES method of simulation[1] (Large Eddy Simulation, i.e. large-scale simulation), developed barely ten years ago while perfecting the AVBP code which enabled the implementation of this model on a vastly parallel computer. AVBP was successfully tested on several computers: i.e. IBM's BlueGene in the United States, a Cray XT3 and, in the first half of 2006, the CEA's Tera-10. While it would have needed between 150,000 and 200,000 hours on a standard computer, the simulation took just a few hours on these machines.
"Instead of calculating a single burner on a chamber sector as is often done for cost-saving reasons, it was possible to simulate the behaviour of the 18 burners," reports Thierry Poinsot. The AVBP code is the product of the collaboration between the CERFACS and a large number of public (in particular CNRS) and privat (IFP, Turbomeca, Snecma, Peugeot, Renault etc.) laboratories connected through one and the same consortium. What remains now is to move on to the industrial phase, as the researcher added: *"This presupposes a significant investment on the part of France in the acquisition on hugely parallel computers."* **D. R.**

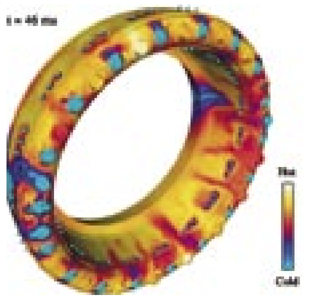


DIAGRAM OF THE STAGES OF COMBUSTION

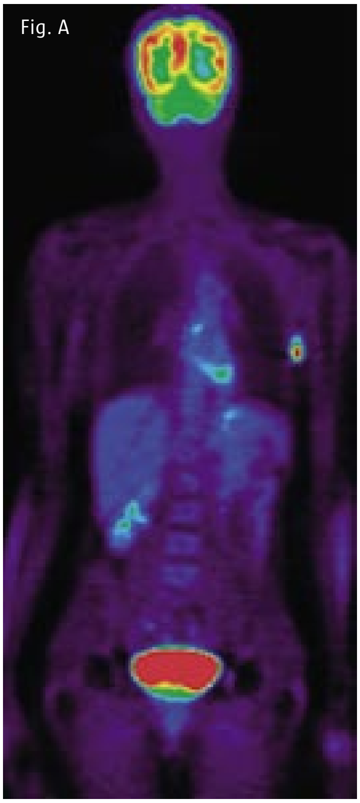
© Infographie CERFACS

(1) Large-scale simulation (Large Eddy Simulation, LES), as opposed to so-called direct simulation (Direct Numerical Simulation, DNS) is based on a separation of the scales: it consists in calculating the large scales by modelling the action of the small scales.

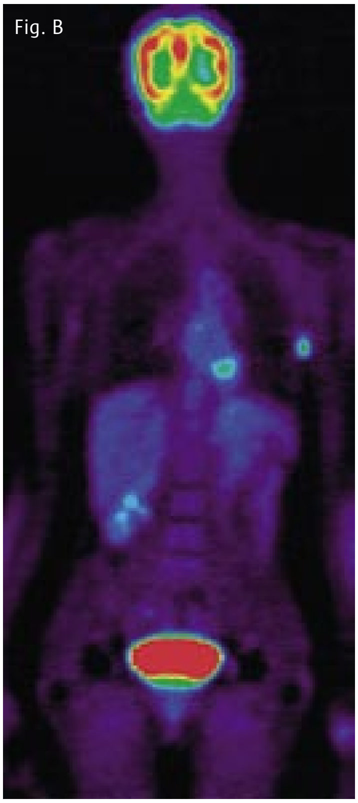
Oncology:
first simulation of a full body scan

Nuclear imaging technologies, such as positron emission tomography, should prove beneficial to the recent numerical simulations of medical scans.

Anne Lefèvre-Balleydier is a scientific journalist



© CEA



© CEA

Fig. A: Result of a real PET scan: the coloured areas are areas with increased metabolism, the tumours are under the patient's left armpit and in the liver.

Fig. B: Result obtained by simulation on the Tera-10 computer.

These two images do not originate from the same source: one comes from a medical scan, which is becoming increasingly common in oncology, i.e. Positron Emission Tomography (PET), and the other was produced by numerical simulation. However, they are very similar: in both cases, it is possible to observe tumours in the liver and under the left armpit. The resemblance is quantitative as well as qualitative: in the image produced by the PET scan the volume of the node under the lung is 23.7 ml (millilitres) and 22 ml in that produced by simulation, thus giving a difference of only 6 percent. For Sébastien Jan, a researcher at the CEA and technical coordinator of an international project on PET simulations since April 2003[1], it is without a shadow of a doubt "beautifully successful" and, within a period of five to ten years, should trigger significant progress in the area of nuclear imagining and, specifically, Positron Emission Tomography (PET). PET, which is used in oncology in particular, provides information on the functional aspect of the disease and is of interest, therefore, in both diagnostics and the monitoring of treatments.

Providing a realistic image of tumours ...

This imaging technology, for which the number of available systems has recently increased thanks to the French Cancer Plan (around one device per one million inhabitants), makes it possible to identify tumoural cells by visualizing glucose consumption. *"In practice, you inject a radiopharmaceutical which, in the context of an oncological test, is a molecule similar to sugar (fluorodeoxyglucose, FDG) marked with fluorine-18, a radioactive tracer with a short half-life (two hours),"* explains Sébastien Jan. *"It spreads through the body and concentrates in*

(1) Established in 2002 by the team at the Swiss Federal Institute of Technology Lausanne, the aim of the Open GATE project, which involves over 20 research laboratories, is to validate a simulator dedicated to PET. In France, the collaboration involves teams from Inserm (French National Institute for Health and Medical Research), the CNRS and the CEA.

areas in which the metabolism is raised, as is the case, in particular, with tumours. This tracer emits positrons which will annihilate themselves with electrons in the neighbouring tissues: for each instance of annihilation, two gamma photons are generated and emitted in two opposing directions." Their simultaneous detection by the detectors on a revolving scanner surrounding the patient (cf. scan image) then enables the construction of an image of the distribution of the tracer in the body. To simulate this full-body PET scan, it is therefore necessary to reproduce all of these processes from the injection of marked sugars to the reconstruction of the image in three-dimensions via the interaction between particles and matter.

© C.Boulze/CEA

Approximately one hour is needed to derive data from a PET scan. During this time,the patient must remain still inside scanner.

... by simulating the path of the particles

This simulation is based on the Monte-Carlo methods: it starts from the physical properties of the different particles (such as the energy, impulsion, load etc.) and of the tissues and substances that they cross (such as their atomic arrangement, their density etc.), and the path of each particle is evaluated on the basis of the probabilities that it has of interacting with them. The objective? To get as close as possible to the real processes so as to optimize the extraction of the relevant signal, in other words the metabolic activity of a tumour. The image produced by a PET scan is generally very noisy, before all kinds of corrections are made to it, first, because the sugar is not fixed by the cancerous cells alone and, second, because the detection of the two photons emitted by the positron-electron annihilation will not always occur under ideal conditions: i.e. the photons can diffuse in the tissues, be attenuated there or originate from two different annihilations causing a fortuitous coincidence of impact. However, given that in simulation all of these events are known, it is possible to take them into account and correct the data accordingly. Even better: it is also possible to implement corrections of the movement to be taken into account, for example, the patient's respiration which tends to smooth the signals originating from the lungs. The problem is that all of this requires vast amounts of computing time.

"But in five or ten years' time, we will see simulations of this type coupled with real hospital scans"

"This is directly related to the dose of tracer injected and the acquisition time, i.e. the number of particles that have to be simulated," stresses Sébastien Jan. For a traditional PET scan, in which the injected radioactive dose is around 300 MBq (millions of becquerels, unit of measurement of radioactivity) and the data acquisition time 40 minutes, around 600 billion positrons are generated and, hence, twice that number of photons. Given that each of the particles is modelled, this process would require a minimum of 20,000 computing hours or around 800 days on a standard computer. By developing a simulation platform (called GATE, Geant4 Application for Tomographic Emission) on the Tera-10 supercomputer, the researchers have succeeded in reducing this time lapse: in the full body scan they opted to simulate, the dose reached 264 MBq and acquisition time 49 minutes, but with 7,000 processors, only three computing hours were required.

"This was a first. But in five or ten years' time, with more distributed computer architectures, I think we will see simulations of this type coupled with real hospital scans, and this would make it possible to customize PET scans while optimising, among other things, the doses of tracer, the acquisition times and, more generally, the acquisition protocols." Sébastien Jan's team is working on this, in particular the optimization of the signal-to-noise ratio in the scan image and enabling better tumoural detection. The aim is also to reduce the processing time to produce simulations using far fewer processors. Thus, what interests the researchers is not so much the "beauty" of their simulation, but the fact that it will be able to become a universally accessible tool.

A.-L. B.

Fig. 1A. Andromeda, photographed in visible light (Mount Palomar, California, 1958), with its two companions, M110 and M32. M32 (on the left) is suspected of having collided with Andromeda 210 million years ago.

Numbers and lights in the sky

For astrophysicians, digital simulation has become a vital addition to the telescope. It allows them to explore space as well as time. Here we present a series of results recently obtained from the COAST project (computational astrophysics), a collaboration of researchers from the CEA-Saclay astrophysics department and its partners (Université Paris 7, CNRS, and Paris Observatory). These spectacular snapshots demonstrate just how much simulation can advance our knowledge of astrophysics on all levels.

Pierre Vandeginste is a scientific journalist

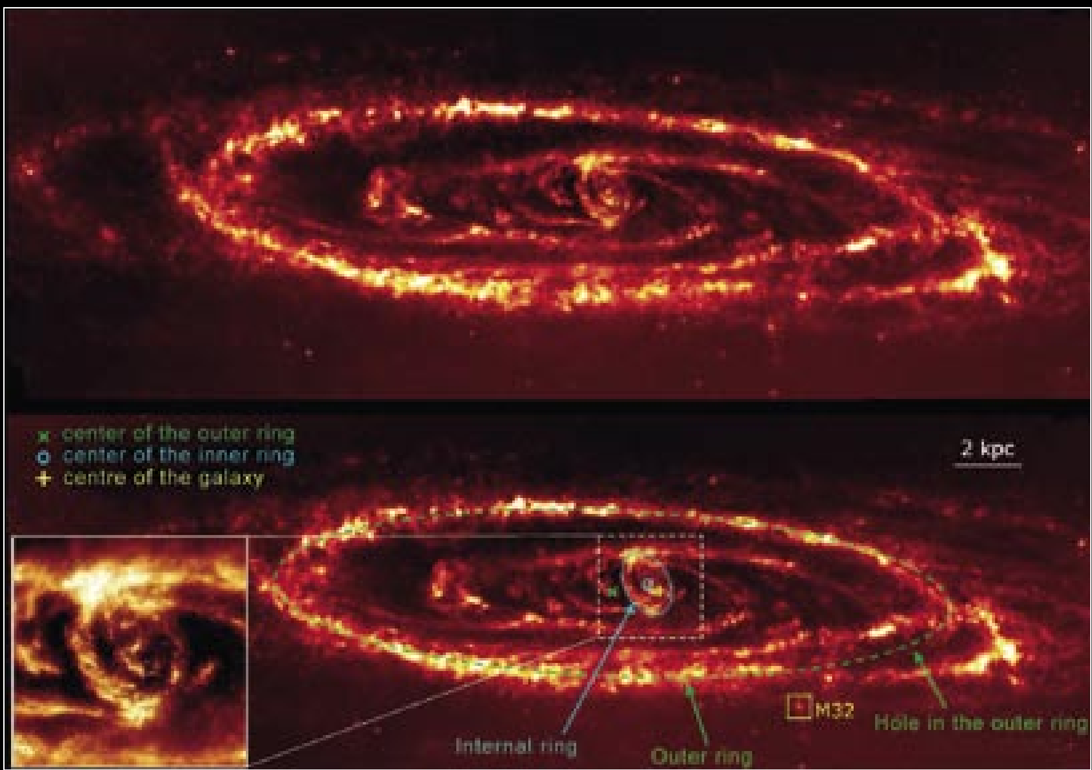


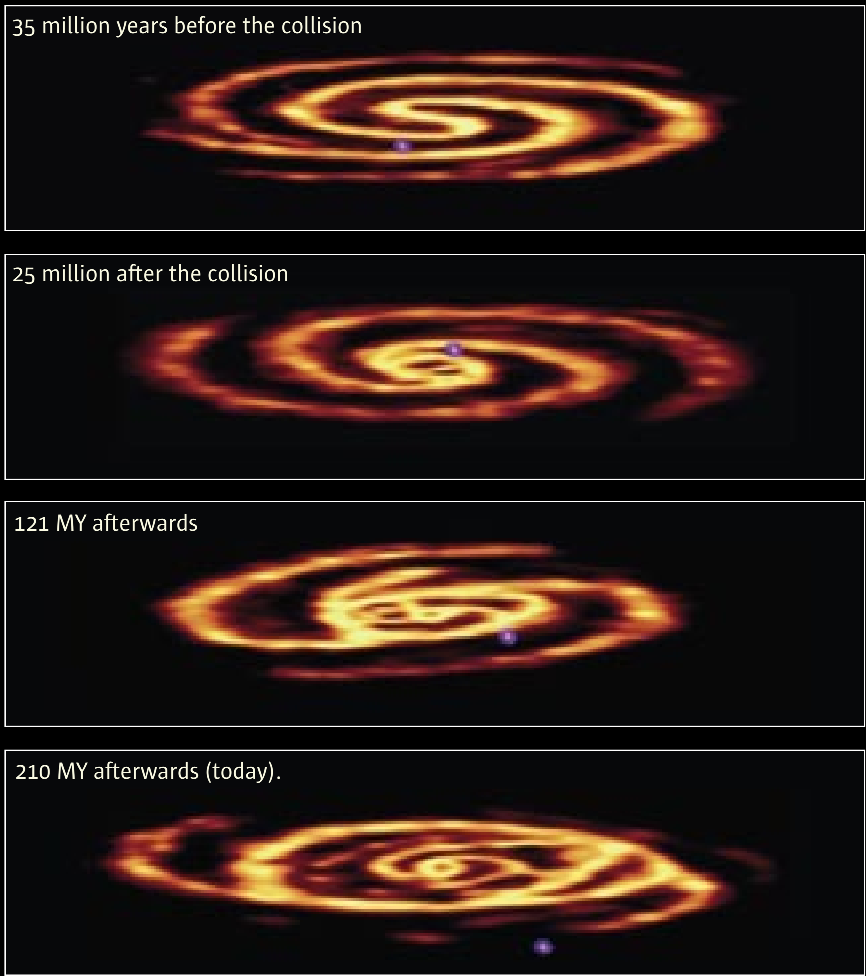
Fig. 1B. Andromeda as seen by the Spitzer satellite in infrared. We see two off-centre rings in the middle of Andromeda as well as a gap in its outer ring.

(1) http://www-dapnia.cea.fr/Projets/COAST/index.htm

Andromeda still shows signs of its collision

Andromeda is the nearest spiral galaxy to our own, the Milky Way. Located two million light-years away, it also the only galaxy in the Northern hemisphere that is visible to the naked eye and the most massive in our local group of galaxies. Although Andromeda appears normal in visible light, an image (1B) taken by the Spitzer space telescope reveals a distorted shape. Launched by NASA in 2003, Spitzer is an infrared telescope, which in particular, allows us to see interstellar dust. This photograph reveals two off-centre rings in the middle of Andromeda and a gap in its outer ring. What event was able to distort Andromeda like this? It was most certainly a collision 210 years ago[2]

(2) www-dapnia.cea.fr/SAP/Actualites/Breves/bournaud061019/hpage.shtml



© F. Bournaud/CEA/CNRS/OP

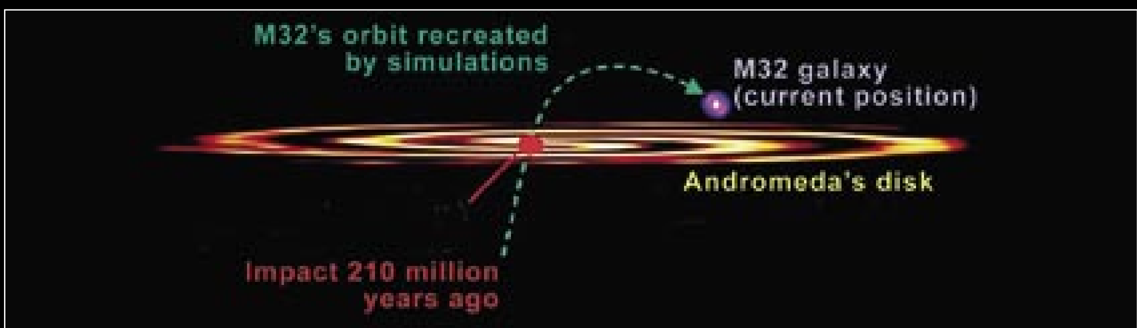
Fig. 1C. Simulation of the collision between M32 and Andromeda. The four images correspond to the following dates: 35 million years (MY) before the collision, 25 MY after the collision, 121 MY afterwards, and 210 MY afterwards (today). The last image is very close to the observation made by the Spitzer satellite.

(3) An almost head-on collision as the origin of two off-centre rings in the Andromeda galaxy, D.L. Block, F. Bournaud, F. Combes, R. Groess, P. Barmby, M. Ashby, G. Fazio, M. Pahre, S. Willner, Nature, 10/19/06

with its much smaller neighbouring galaxy, M32. This finding, explained an international team in last October's Nature[3] magazine, was based on a series of digital simulations performed by Frédéric Bournard (CEA-Saclay, Paris Observatory) on the CCRT's (Centre for Research and Technology Computing) vector supercomputer, NEC-SX6. The model that was used represented each galaxy with a million particles and simulated star movement, the formation of new stars in interstellar gas, and the gravitational influence of the dark-matter halos surrounding the galaxies. Sixty separate simulations were needed to determine the exact trajectory followed by M32 that would be most likely to explain Andromeda's strange shape.

In photograph 1A (see page 22) we see the Andromeda galaxy observed in visible light, as well as two dwarf galaxies, including all that is left of M32 after its collision with Andromeda. Photograph 1B (see page 23), taken in infrared by the Spitzer satellite, reveals its two off-centre rings and a gap in the outer ring. Finally, the sequence of digital simulation photos in 1C (above) shows how the collision unfolded, and figure 1D (below) shows the trajectory of the M32 dwarf galaxy, which caused it to hit and plunge through the Andromeda spiral galaxy.

Fig. 1D. Recreation of the trajectory of M32, the dwarf galaxy that hit and plunged through the Andromeda galaxy 210 million years ago.



© F. Bournaud/CEA/CNRS/OP

Two spiral galaxies meet... and make babies

Small galaxies called "dwarf galaxies" contain less than a billion stars, a hundred times less than massive galaxies such as ours. Where do these tiny galaxies come from? Many of them are found orbiting these massive galaxies, including fourteen discovered around the Milky Way. Until now, we believed that most of these small galaxies seen today were created very early on in the history of the universe and had avoided merging into larger galaxies. However, digital computer simulations[4] performed by a team from the Astrophysics Department of CEA/DAPNIA (research laboratory based on the fundamental laws of the universe) and the Paris Observatory have led to a change in thinking. These simulations demonstrated that dwarf galaxies are produced during collisions between massive galaxies, and would therefore continue to survive.

Frédéric Bournaud and Pierre-Alain Duc simulated approximately one hundred spiral galaxies on CCRT's vector supercomputer, NEG0-SX6. Their digital computer model represented each galaxy by a cloud containing three to six million particles and simulated star movement, interstellar gas, and dark matter. Five hours of calculation were needed on average to carry out each simulation.

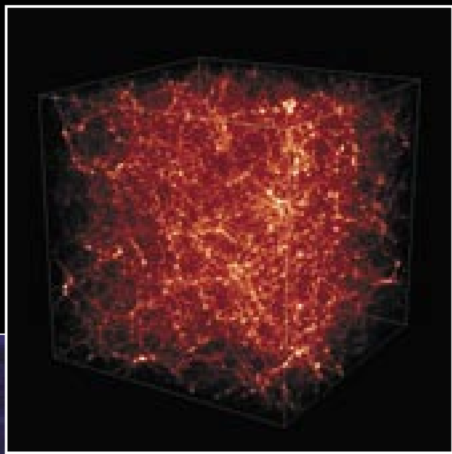
When spiral galaxies collide, they build up long, antenna-like tidal tails. New "tidal dwarf galaxies" can then form in these "antennas". The simulations showed that nearly one-fourth of these dwarf galaxies produced by a collision survived for at least two billion years. Therefore, we can believe that a small fraction of dwarf galaxies originated from a collision, at least in certain environments.

Fig. 2A. Simulation of a collision-merger of two spiral galaxies. The first three photos correspond to three dates: T = 0, T = 300 MY (million years) and T = 1000 MY. The fourth photo shows a detailed look of the third: a tidal dwarf galaxy forming from a collision. The model used has 80 million particles and required 40 hours of vector calculation on the CCRT's NEC-SX6.

(4) http://www-dapnia.cea.fr/Sap/Actualites/Breves/paduco60604/page_fr.shtml html

Fig. 2B. Another simulated collision between two spiral galaxies in slow-motion. This time, there are 30 million years between two images, and the sequence lasts for a total of 330 million years.

Formation of large-scale structures in the **early universe**



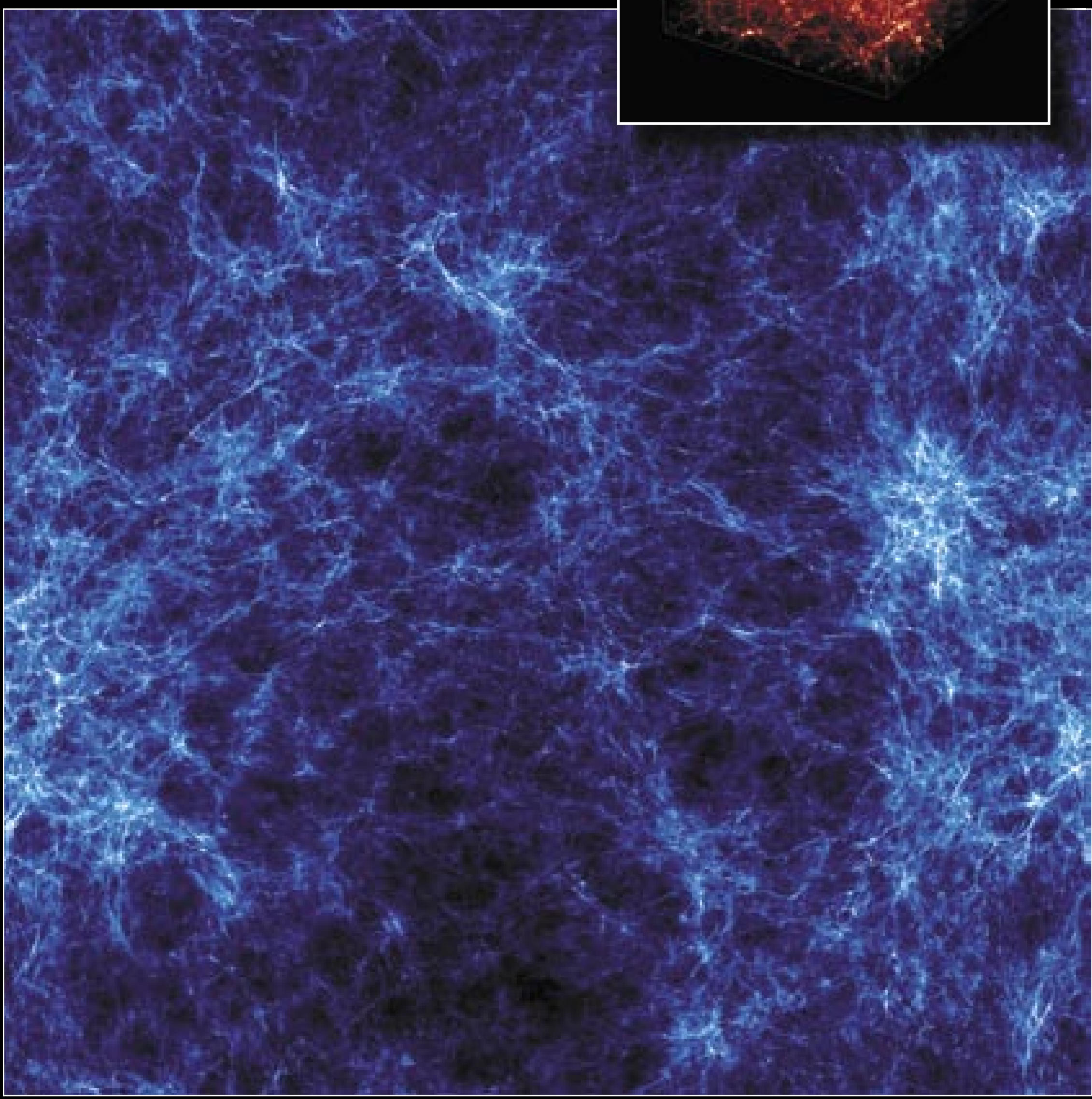
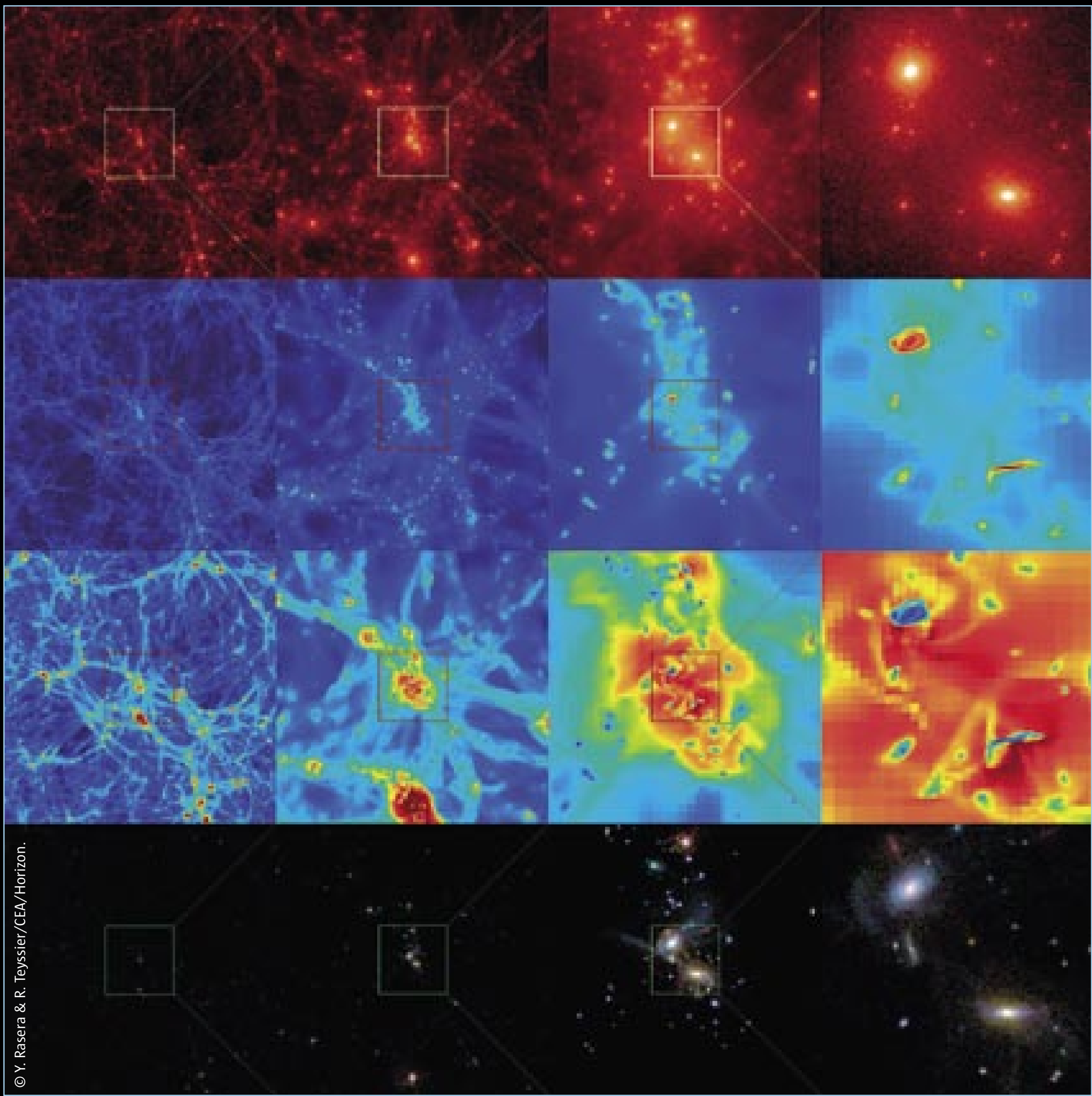


Fig. 3A. Cosmological simulation of the formation of large-scale structures from dark matter in a section of the universe that is one billion years old. The lighter areas correspond to dark matter halos, in which galaxies and their stars form.

Fig. 3B (upper right). Twelve million year later...a section of the current (simulated) universe, this time in volume, with even denser regions correspon-ding to clusters of the current galaxies.



© Y. Rasera & R. Teyssier/CEA/Horizon.

Fig. 3C.
Same type of simulation as 3A, but 2 billion years after the big-bang, in a "universe cube" with 10 million light years on each side. The image shows four zoom levels from left to right, and dark matter, gas, gas temperature and stars are shown (in order) from top to bottom.

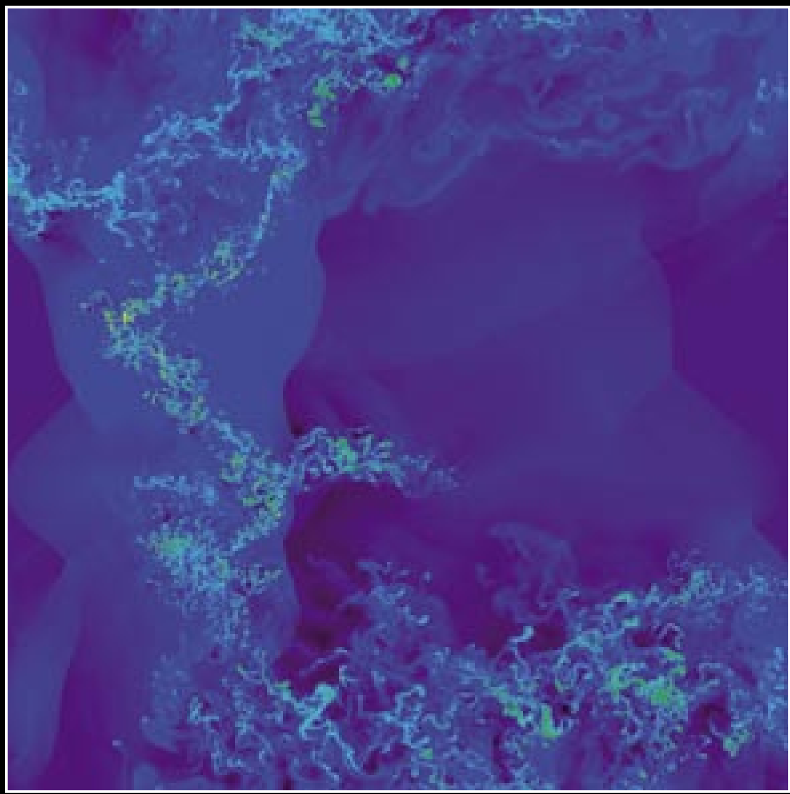
Digital computer simulation goes back in time. One or two billion years is just childhood for a universe that could now blow out thirteen billion candles on its birthday cake. During that period, what we see is the formation of large structures such as filaments and clusters.

Image 3A (see page 26) shows the universe at one billion years old. The light-coloured masses represent dark matter halos, in which stars and galaxies form. In the centre right we see a large cluster of galaxies being formed. At one time, the corner of the universe represented here was 150 million light years on each side; however, the universe has expanded since then and its contents are now spread out over a billion light years. The model[5] being used is made up of ten billion particles. It was run on MareNostrum, the IBM-cluster type supercomputer at the Barcelona Supercomputing Centre. The universe at two billion years (photo above) is slightly more mature and contains galaxies that are already more evolved. Here we see a small cube of ten million light years displayed from several angles (vertically) and "zoomed" three times (from left to right). From top to bottom, the images show (in order) dark matter (1st line), mass density (2nd line), gas temperature (3rd line), and stars (i.e. visible light) (4th line). During this period, gas flows down the filaments to the centre of the dark matter halos and forms "cold" disks, which are the future galaxies where the stars will shine.

(5) http://www-dapnia.cea.fr/Phys/Sap/Actualites/Breves/teyssier070115/page.shtmlfr.shtml

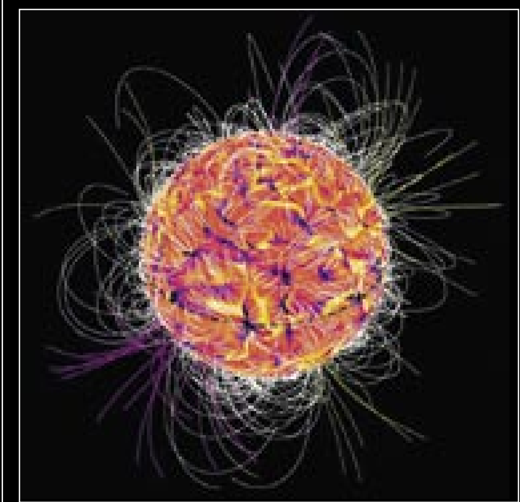
Interstellar Turbulence!

This simulation shows a small piece of the galaxy, 60 million light years on each side. It simulates turbulence phenomena in interstellar gas, in which stars are formed. These simulations demonstrate the behaviour of interstellar gas, notably the turbulence that can develop, which enables a better understanding of the first stages of star formation. Massive parallelism makes it possible to attain a high resolution here.

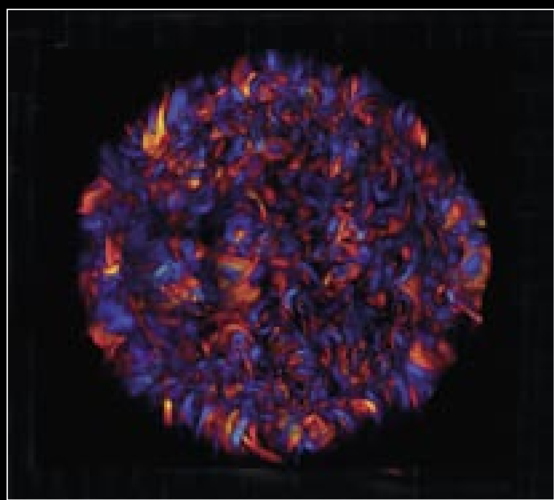


© E. Audit/CEA

The sun's magnetic impulses



© A.S. Brun/CEA

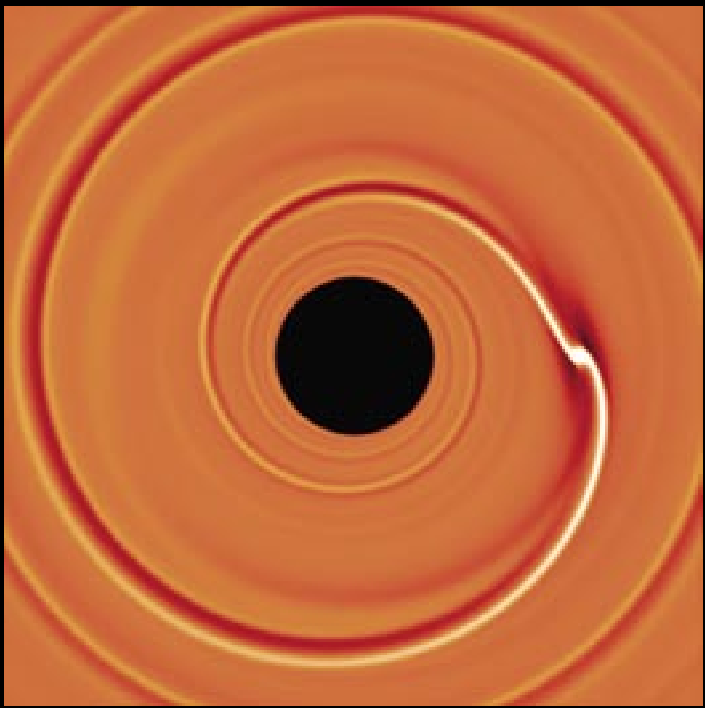


© A.S. Brun/CEA

These images show the re-creation of the solar corona's magnetic field through a high resolution, 3-D simulation of solar magnetism and dynamo. This type of simulation[6] makes it possible to update the physical mechanism behind the sun's rotation, its meridional circulation, and its magnetism. The colours represent the magnetic fields at the surface, and the wavy lines in the photo on the left represent the magnetic fields in the solar corona.

(6) http://www-dapnia.cea.fr/Phys/Sap/Actualites/Breves/sbrun051115/page.shtml

The birth and migration of planets



© F. Masset/CEA

This image could represent our solar system at its birth four billion years ago. The sun (in the centre, not represented) is surrounded by a ring of "protoplanetary" matter. The presence (and mass) of the first planet (in the centre of the light area on the right) distorts the entire disk, which was initially uniform before the development of a spiral structure with over-dense areas (light) and under-dense areas (dark). The disk's reaction in turn affects the planet by exerting gravitational force on it, which will speed up or slow down its rotation and by doing so, increase or decrease its orbital ray - this is "planetary migration". Such simulations can help explain how the planets ended up in their current solar system positions from the positions in which they must have originated.

Filaments and clusters in Saturn's rings

© S. Charnoz/univ. Paris 7 & AIM/CEA. Post-production Frédéric Durillon/Animea

In this image, the physics of Saturn's rings are being simulated on a scale of approximately thirty meters. The rings' matter, "ice cubes" measuring a few centimetres to a meter, is sometimes grouped together in "filaments" and sometimes in "clusters". Clusters are the small satellites that form around Saturn, which can also be destroyed by collisions. Digital simulation calculates the result of the complex balance between gravity, which tends to bring matter together, and tidal waves, which tend to distort or split any object that forms. Cassini (NASA/ESA probe in orbit around Saturn since 2004) has recently shown filaments in the rings that had been predicted by simulations.

It is interesting to note that the mechanism that creates filaments and satellites in Saturn's rings (called Jeans wave) is the same as the one that (on a much bigger scale) forms stars in molecular clouds, as well as spiral arms in the galaxies. Here it is easy to see how simulations deal with the "multiscale" aspect of the universe's physics.

Explosions on a Computer

The enhanced strength of supercomputers opens the door to increasingly realistic simulations of explosive physics.

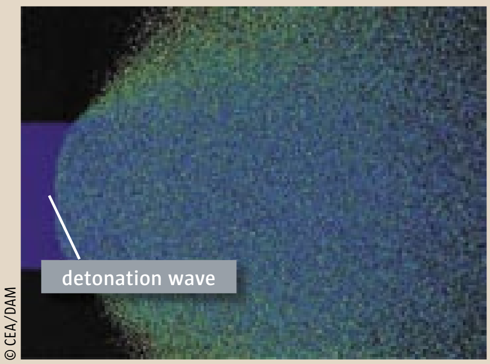
Dominique Ritman is a scientific journalist

There are hardly any experiments more destructive than those aimed at studying a detonation wave propagating in an explosive. Moreover, the mechanisms that they generate take place in a very short time and involve extremely high pressures and temperatures, making their observation particularly delicate. This is why the digital simulations performed at the CEA's Direction des Applications Militaries (DAM- military applications department) in Bruyères-le-Châtel, (Ile-de-France). For example, it took approximately one hundred hours for one thousand Tera-10 processors, the supercalculator installed in 2006 at the CEA-DAM, to model, based on traditional molecular dynamics calculations [1] the collision of two detonation waves propagating in a dense medium comprised of millions of molecules interacting mutually.

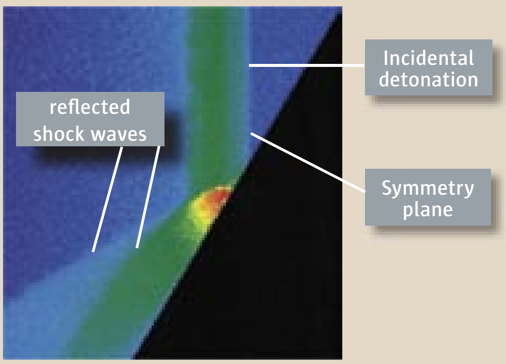
"Therefore, simulating the operation of a weapon requires very precise knowledge of explosive physics"

A nuclear rocket requires an initial source of energy that is very powerful, yet light-weight. The choice falls immediately on condensed explosives. Therefore, simulating the operation of a weapon requires very precise knowledge of explosive physics. Explosive principle: a detonation wave (comparable to a reactive shock wave) propagates in a chemically reactive material and in turn releases a large amount of energy, which continues the propagation of the wave.

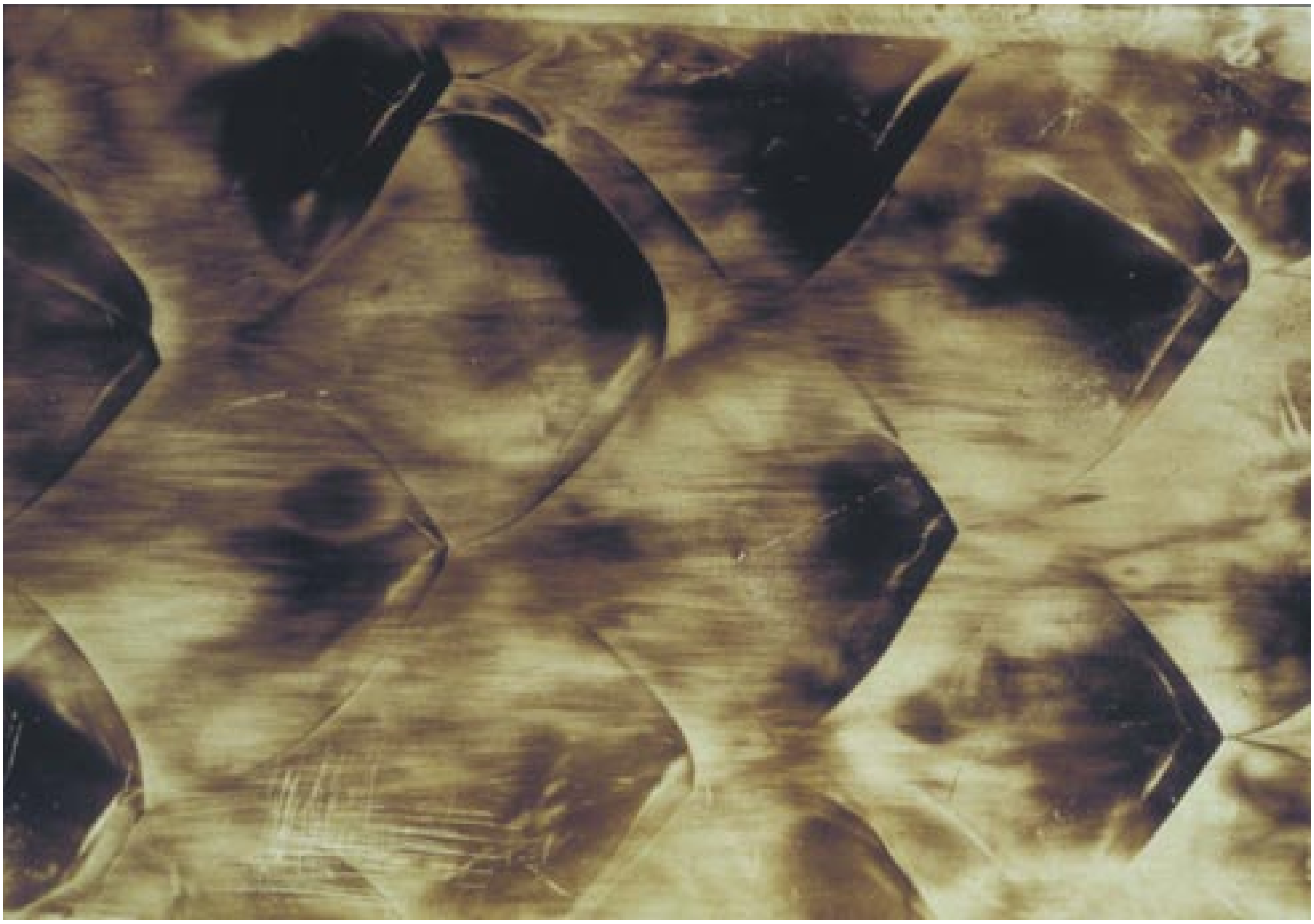
(1) Molecular dynamic equations describe the evolution, based on time, of a system of particles interacting amongst themselves.



Propagation of a detonation wave in a cylinder cartridge without containment: the wave's curve is an important characteristic, which is accurately reproduced by a molecular dynamic simulation.



Symmetrical reflexion of a flat detonation wave. In this complex case, two shock waves are reflected. This 10-million molecule simulation was performed on 1,000 Tera 10 processors.


© CNRS

Detonation wave cells propagating in a gaseous mixture of oxygen and argon: we can observe that in contrast to traditional light wave interference figures (whose propagation follows a linear system described by Maxwell equations), detonation waves, non-linear equation solutions, lead to complex geometric structures.

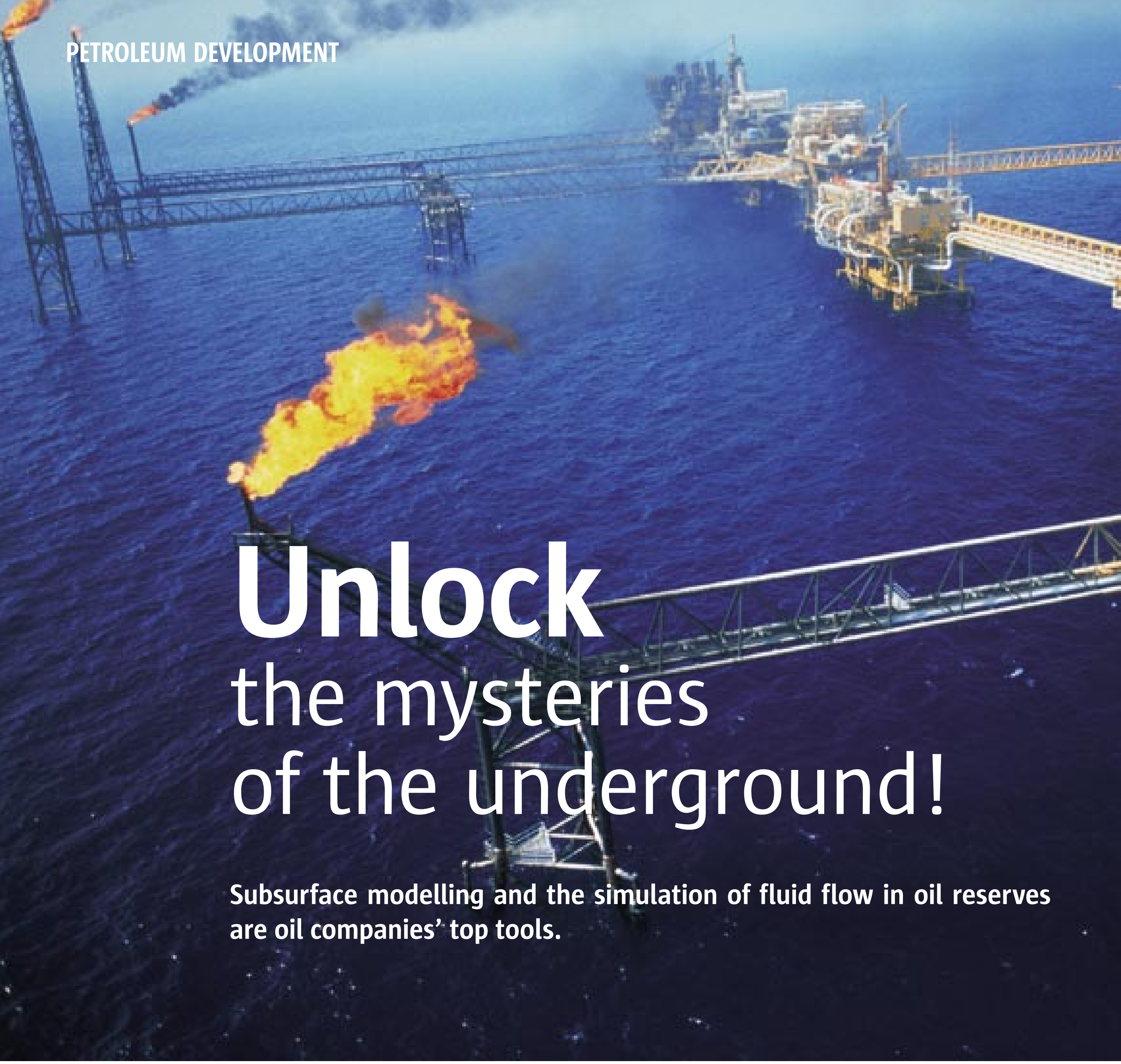
"We are studying two crucial points in particular: first, the chemical processes in the zone where explosive decomposition reactions occur (called reaction zones) and second, the behaviour of the gas produced by the explosive's combustion. Propulsion results from the expansion of this gas," explains Laurent Soulard, an engineer at the DAM. So, while real experiments only allow us to evaluate macroscopic quantities (pressure, temperature, matter speed, etc.), the simulations at Bruyères-le-Châtel provide access to decomposition mechanisms and thermodynamic properties of the reaction zones on an atomic scale.

When two detonation waves collide

Molecular dynamic simulations are based on the resolution of movement equations, which made it necessary to choose a potential energy function to describe the interactions between particles. The length of the simulation depends on the complexity of this function and the number of molecules. Initially, scientists chose a simple function and validated the method based on flat shapes, which only required a few thousand molecules. Next, they performed simulations that resembled real situations more closely by using cylinder-type explosive shapes and a larger amount of molecules.

Latest simulation to date: the collision of two detonation waves performed on Tera-10, which included approximately 10 million molecules. These calculations required a massively parallel molecular dynamic code, which is why the DAM developed the STAMP code. These types of simulations offer two major advantages, points out Laurent Soulard. They provide us with accurate measurements of the temperature, pressure, and so on, in the reaction zone at any given moment. In addition, they ensure that the result obtained is the exact solution to the problem, thereby avoiding the tricky gridding that is inherent in methods for continuous media mechanics.

It remains to be ensured that the simplicity of the potential energy function that is chosen does not create latent bias. *"We are currently working on more realistic functions"*, states Laurent Soulard, *"we may have to wait for the next generation of computers to simulate industrial-type systems. But we are already developing the tools that will be needed."*



Unlock the mysteries of the underground!

Subsurface modelling and the simulation of fluid flow in oil reserves are oil companies' top tools.

Yves Sciama is a scientific journalist

With an average price of forty million dollars reaching all the way to 65 million dollars, the price for deep offshore oil drilling speaks volumes about the importance of current investments in petroleum development. The pioneer days when rudimentary drilling conducted at twenty or thirty meters underground would cause a geyser of oil to shoot up into the air are long gone. "Today, easy oil has given way to technological oil," states Stéphane Requena, specialist in high-performance calculations, in charge of the development of calculation methods at the Institut Français du Pétrole (IFP – French Oil Institute), in Rueil- Malmaison, France. Today, petroleum development involves looking for black gold in deep deposits, at high temperatures and pressures, in sites that are extremely difficult to access, or turning to the development of complex hydrocarbons, such as highly viscous oils.

The enormous investments required for this generates considerable financial risks. The success rate for exploration is approximately 33 percent, and although this rate has not significantly changed, advances in seismics, with a current resolution of approximately ten meters, enable the discovery of reservoirs in increasingly geologically complex areas that are even deeper and farther away. However,

Oil companies now operate in deep offshore to 2,250 meters of water

many uncertainties still remain about the quantity of hydrocarbons that could be extracted and the length of production time.

A few explanations about how oil is extracted help shed some light on the problem. Basically, a deposit is a group of porous rocks full of hydrocarbons – the reservoir – covered with impermeable rocks – the burden (cover). Based on pressure and temperature conditions, gaseous hydrocarbons are found in the upper part of the deposit above the liquid hydrocarbons - in other words, the oil or petroleum – and below that is the water zone where the porous rock contains nothing but water. A deposit often has a complex geometry with an overall inclination of rocks, fault system, rock heterogeneities, etc.

Hydrocarbons are recovered by drilling one or more wells in an oil-saturated layer. In these wells, the natural pressure found in the reservoir is high enough to begin recovery. Next, new wells must be put into place through which water and gas are injected. These fluids "sweep" out the deposit and, in a way, push the hydrocarbons towards the production wells. If needed, recovery can be further improved by drilling new wells (at a fairly higher cost) in poorly "swept" zones, or by injecting chemical products, CO_2, or steam heat. Finally, the current extraction rate reaches an average of 33 percent of oil that is initially there. This rate, which is relatively low, is of significant importance: a slight variation in percentage corresponds to massive volumes and equally large amounts of money.

A dynamic view of exploitation

Thanks in particular to advances in seismics, we now have rather precise images of developed reservoirs. Geological models recreate these structures by using grids of only a few dozen meters on each side, and sometimes less than one meter in height. A deposit stretching over several kilometres in length is thus modelled with grids of several million cells. What is vital, however, for successful development is to anticipate the way in which different fluids are going to move inside the reservoir once it is completed. Asking precise questions is of utmost importance: where should the wells (injectors and producers) be positioned and how many should be drilled to boost recovery? Which fluids should be injected, in which layers and at a what flow rate (injecting too fast often lowers the recovery rate)? How is the deposit's production going to change over time? Thus, the difficulty is going from a precise, yet static, geologic view to a dynamic development view.

The first software programs aimed at this objective began to appear at the end of the 1960s. Since then, they have continued to improve thanks to the advanced computing power available and a better understanding of deposits. The principle behind these models is simple: the reservoir is divided into grids or cells (less refined than those comprising geological models) where porosity and permeability values are assumed to be constant, while varying from one grid to another. These grids are then linked together by equations simulating flows in a porous medium in order to model fluid movements. While this fundamental principle has remained almost the same, the current models have become impressive tools through both their performances and the wide range of possibilities they offer.

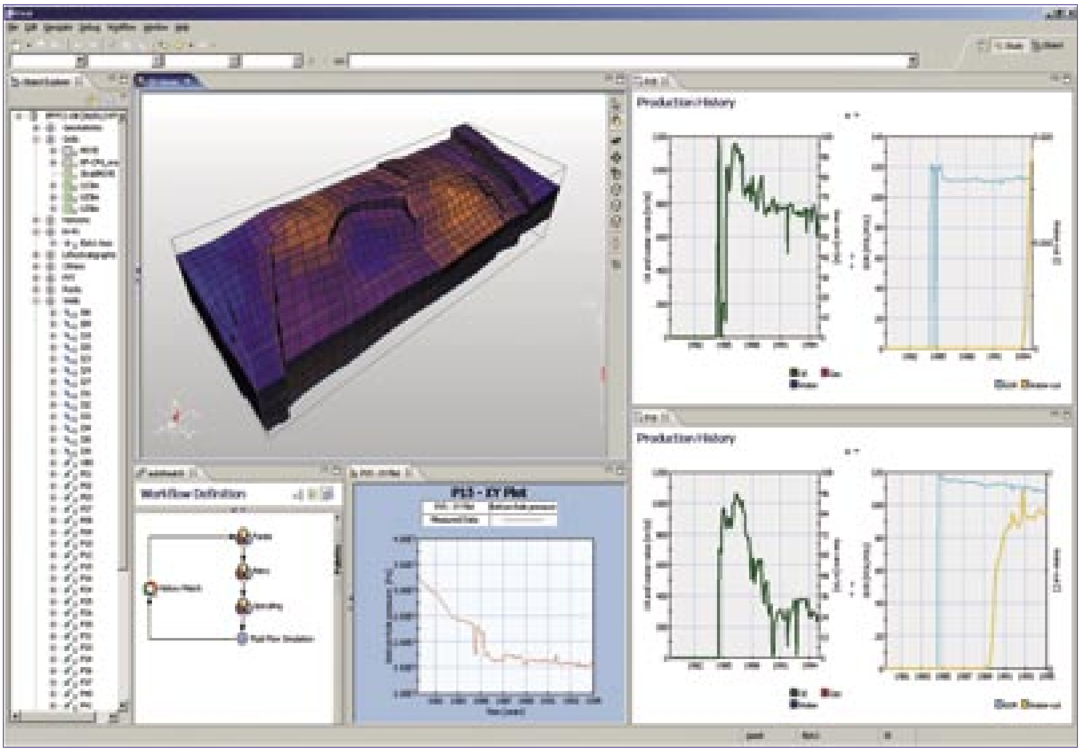


ILLUSTRATION : BEICIP FRANLAB

Result of a dynamic simulation of an oil reservoir, obtained with the First software program: 3D view of grid (left), "history-matching" management tools and well production results (at right, pressure curves, flow rates, ratios).

Modelling in real time

The number of grids and detail of the phenomena represented have greatly increased.

"Today we are capable of performing simulations of several million grids with an average of 3 to 7 unknowns per grid, in just a few hours, using multiprocessor parallel computers," states Patrick Lemonnier, head of the "Flow and Transfer Simulations in Porous Media" Department at IFP. Moreover, these models, which reach a million lines of code, are constantly evolving through the steady addition of new data. Thus, as reservoir development progresses, measurements taken in wells are compared to forecasts made by the software and used to bring simulation closer to reality (which is called history matching).

For several years, we have been seeing the development of 4D seismics: seismic campaigns are now conducted on the same deposits every two years in order to monitor the change in fluids. The current models include data provided gradually by these campaigns as they are produced, data provided in real time by certain sensors (pressure for example) with which the wells are equipped, as well as information obtained during drilling of the new wells.

Of course, this overall switch to real time and the integration of very diverse information requires intensive computational power. At the IFP for example, computing power has been increased by 45 since 2003 and will continue to increase rapidly. "*We are currently developing a simulation platform, called First*[1]*, distributed in April 2007 by our subsidiary Beicip Franlab, which enables us to markedly improve the precision of a reservoir model*", explains Jean-François Magras, specialist of high-performance calculations applied to fluid flows in porous media at IFP.

"Today, easy oil has given way to technological oil"

"And we are participating in the Ter@tec science park's FAME 2 project[2] *(read article on Ter@tec on page 40), which should enable us to mass simulate models of 100 millions of grids, with several hundred processors in just a few years ... a world first."*

Today, with a petroleum recovery rate in constant change and new well architectures (horizontal wells, multibranch, etc), which can increase production tenfold compared to vertical wells, reservoir modelling has earned its stripes. It has become a key element in decision-making, and future enhancements in computer strength allow us to hope for the continued improvement of reservoir modelling results. We want to use hydrocarbons, but their complexity is a challenge that will be difficult to overcome. **Y.S**

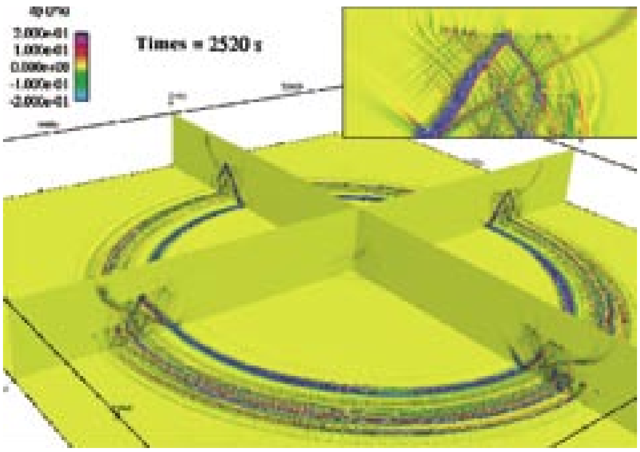
(1) First: Reservoir simulation platform developed by the IFP. Detailed information can be found on the Beicip Franlab site: http:// www.beicip.com

(2) FAME 2 project: www.fame2.org
To find out more: http://www.ifp.fr

Simulation for monitoring nuclear explosions

Infrasound sensors, propagation models, and supercomputers: three conditions for tracking down nuclear tests.

Infrasound can be heard by certain animals (elephants, dolphins, etc.), but are inaudible to humans because their 20-hertz frequency is too low for us. However, their ability to propagate over long distances make them particularly valuable for detecting acoustic events produced several hundred (to even several thousand) kilometres away. This ability was the idea behind using infrasound, acquiring suitable sensors, to monitor nuclear tests in compliance with the Comprehensive Test Ban Treaty (CBTC). But how can a prohibited explosion taking place far away be identified and found through recorded signals? The answer is: three-dimensional sound propagation models, and given the size of these models, very powerful computers. At CEA, the propagation of an infrasound at one-tenth hertz was successfully simulated in ten hours with slightly less than 2,000 Tera-10 processors.



© CEA/DAM

Simulation of infrasound propagation on Tera-10: the pressure wave (infrasounds) created in the middle of the image has propagated in a circle, here over 2,000 km. The cross-section shows that it is reflected at an altitude of approximately 120 km. The large view shows the rich frequency structure and slight skew due to the wind.

The size of the calculations is the result of a number of parameters (infrasound propagation relies heavily on meteorological conditions, including wind speed), as well as the grid quality of the studied region. Researchers in the military applications department (DAM – direction des applications militaires) relied on data collected during an American test, conducted in 1984 in New Mexico. *"The challenge was to find, through digital methods, what the sensors had recorded twenty years ago,"* stated Bruno Després, Director of Research at the DAM's information simulation and science department (DDSI - Département Sciences et Simulation de l'Information). *"The results were beyond expectations, the accuracy proved to be excellent,"* he added.

Modelling must include the wind

The volume on which the simulation was performed was spread out over a wide geographical area, eight times the size of France, to a height of 200 kilometres. The gridding was determined so that the details of roughly 400 meters could be accessed, which resulted in no less than 10 billion grids. In particular, the simulation was able to prove the significance of wind: *"At an altitude of 100 kilometres, the atmosphere behaves somewhat like a wave guide: infrasound emitted on the ground bounce up and then come back down, which is why they can be heard over long distances,"* explains Bruno Després. The signals detected by each sensor are highly dependent on atmospheric winds, which the model accurately displayed.

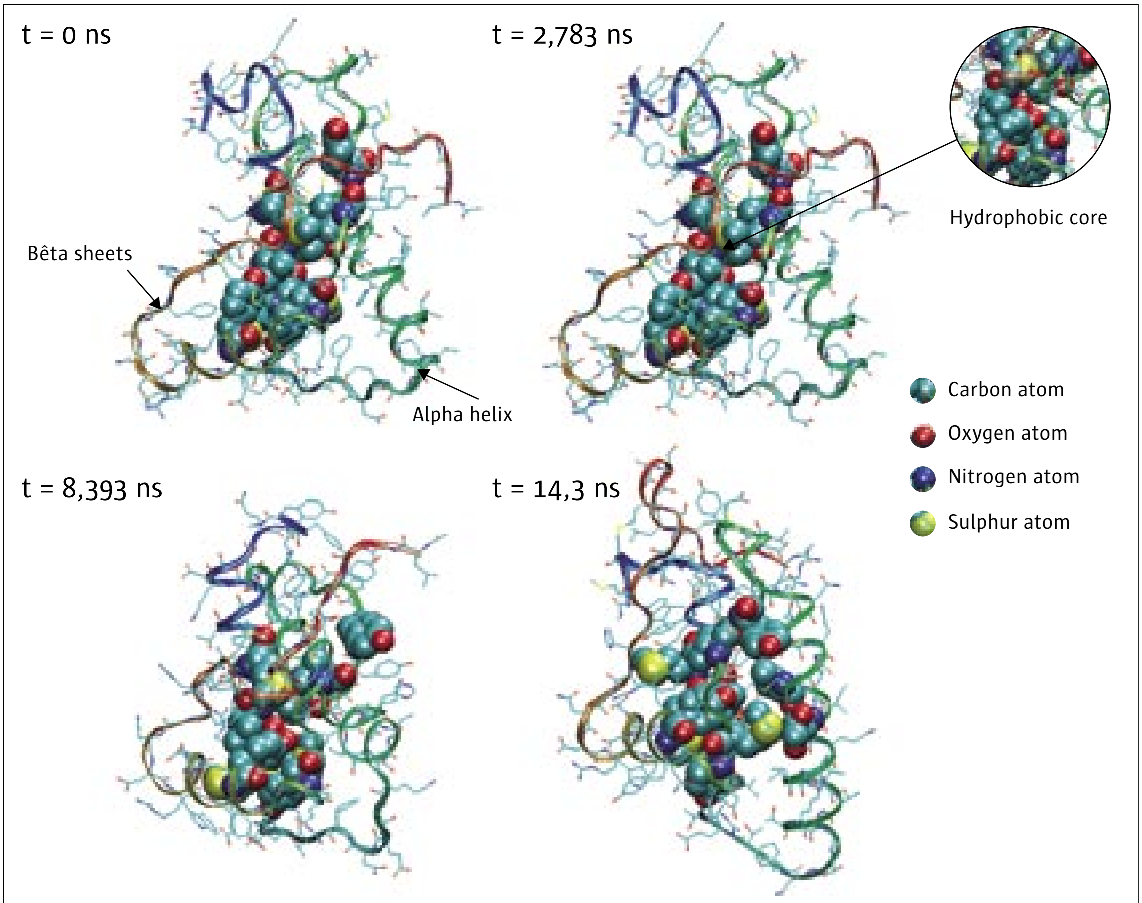
Another remarkable result was the quality of the digital methods and code, the result of the collaborative work between the DSSI and the environment analysis and monitoring department (DASE - Département Analyse et Surveillance de l'Environnement) of DAM, as well as its optimization for the Tera-10 machine. The researchers proved that doubling the number of processors cut calculation time in almost exactly half: this factor characterizes something called scalability, which is nearly 100% here. In addition, notes Bruno Desprès, there is nothing in the principles against using such methods in other applications; for example, to simulate the propagation of resilient acoustic sounds during earthquakes. **D. R.**

Uncovering the "bad" prion

Will digital simulation be able to uncover the structure of a toxic prion in three-dimension someday?

Anne Lefèvre-Balleydier is a scientific journalist

Mad cow disease, scrapie, new strains of Creutzfeld Jacob disease ... there is one single infectious agent involved in all of these syndromes. This agent is an abnormally folded form of a protein called prion (PrP), which forms fibres and plaques that accumulate in the brain tissue, causing neuron degeneration. To fight this toxic prion (PrP^{sc}), we must figure out its three-dimensional structure; in other words, the way in which its long chain of amino acids is folded into loops, helices, and sheets. And that's where the shoe pinches. *"To determine the structure of a protein experimentally, it must be in a crystal or solution form* (then it can be analyzed respectively through X-ray diffraction and nuclear magnetic resonance (NMR). *However, the toxic prion is amorphous and insoluble,"* explains Vincent Croixmarie, a former doctorand of the CEA's condense matter physics department, and currently a researcher for Servier.



DR

Simulation of a normal prion unfolding, from time steps of zero to 14.3 nanoseconds (ns) at 177°C. The protein unfolds normally: its core pulls apart and the directions related to its secondary structures change. But there is no trace of the numerous beta sheet structures that are characteristic of abnormal prions.

Finding the step in which the abnormal prion remains "stuck"

Nonetheless, there is a way to circumvent the problem, because, as Vincent Croixmarie pointed out, *"a protein is like a ball of yarn that can be rolled and unrolled: unfolding it allows us to see how it is folded."* During the thesis[1] that he supported in the spring of 2005, he simulated a normal prion protein unfolding in as many steps as possible, in order to uncover the intermediate step in which the abnormal prion remains "stuck" and thus reveals its structure. His only clue: only the experimental data he had based on the structure of the abnormal prion, which shows a large number of sheets in the protein before the formation of aggregates.

"Unfortunately," states Vincent Croixmarie, "this was a failure. But we have learned a lot about the behaviour of the helices and sheets that make up a protein." The young researcher started from the normal prion protein, which was soluble, and ended by identifying a three-dimensional structure. Turning to molecular dynamic [2] simulation, he modelled this protein as a group of atoms (defined by their mass, electric charge, etc.), connected together by rigid constants describing the forces at work. Next, Vincent Croixmarie chose to impose a certain number of conditions – for example, relatively high temperatures for unfolding the protein and observing what happens, by breaking time down into very short periods (10- 15 seconds) to see the "film" in slow motion. He would never have been able to perform the simulation without the power of a supercomputer such as the Tera-1, the first computer of the Tera generation installed in 2001 at CEA, because the section of the prion that he studied had 102 amino acids, equalling approximately 1,700 atoms. Because each atom was defined according to its three coordinates, the number of variables to change is enormous. Just like cells, protein is surrounded by water; therefore, there must be enough water molecules included in the simulation

Overall, the protein's longest unfolding sequence included 66 million time steps, i.e. 66 nanoseconds of unfolding, and it took one year to complete

Overall, the protein's longest unfolding sequence included 66 million time steps, i.e. 66 nanoseconds of unfolding, and it took one year to complete. This seems like a considerable amount of time, but it is still not enough. "We thought that we would see the alpha helices disappear and then obtain a high rate of beta sheets, that we know prior to the formation of abnormal prion aggregates," stated. However, our proteins demonstrated a high stability: at a very high temperature, the heat destroyed it and its structure became random, and at a lower temperature, the helices remained in place." The protein unfolded too fast, without leaving enough time for the beta sheets to appear.
In other words, the film of this unfolding was too fast. "In order to have a chance of capturing the moment when these famous sheets are formed, you must use a lower simulation temperature," adds Mr. Croixmarie, "but this ends up increasing the calculation quantities again..."

Other avenues to explore

Vincent Croixmarie is also thinking of another angle: "Because there is much more water than protein, simulating molecules of water takes up 80% to 90% of calculation time. But there are ways to solve this problem by simulating the presence of water through its properties: this "default" water, which is still being developed, should make it possible to reduce the calculation time needed to obtain a time step." Even though there have been other hitches to the simulation's success, Vincent is relatively optimistic. According to him, it is possible that the normal prion did not virtually change into an abnormal prion because the process was too simple: maybe the protein will only fold abnormally in a certain environment in the brain.

The fact remains that even if his research did not reveal the toxic protein's structure, they made it possible to find out more about the formation process of the fibres that are characteristic of the abnormal prion found in certain types of diabetes or Alzheimer's disease. Beginning with a fragment of normal prion protein with a sheet structure, Vincent Croixmarie simulated the assembly of several of these proteins into a fibre, which allowed him to observe which atoms and which strength of binding was affected. In the long run, this could open new therapeutic approaches. Until then, the development of simulation methods and calculation power may help uncover the underlying structure of the "bad" prion... **A.-L. B.**

(1) Vincent Croixmarie, Doctorate thesis, Paris 11 Orsay: study of the unfolding of the PrPC protein scaled by molecular dynamic simulation Modelling of amyloid peptide fibres, 2005.

(2) The molecular dynamic simulates the movements of molecular system atoms and the change in their space configuration. Through traditional mechanical equations, it describes the trajectory of atoms according to time.

Philippe Miltin: "Supercomputers, enabling **sovereignty** and **competitiveness**"

The Vice-President of Bull Products and Systems considers supercomputers a prerequisite for innovation in European enterprises and research. They are strategic tools that require Europe's support.

La Recherche. With Tera-10 installed at the CEA in late 2005 and the new machined ordered by the CCRT[1], Bull has entered the arena of major supercomputer manufacturers. How high are the stakes for Bull on the scientific and technical computation market?

Philippe Miltin. It is a vital market for us. Bull wishes to continue building its position as an innovative European manufacturer. High-performance computing is one of the key elements in our strategy. The scientific and technical computing market is the one offering the highest growth rate in IT, with EUR 11.5 billion expected in 2010, as compared to EUR 8 billion in 2006, for an increase of 9%. Our technological innovation comes on the hardware side of our platforms: we develop them by incorporating new technologies that are also used in our infrastructure solutions and management applications.

Moreover, new requirements for high power are emerging, and are not limited to engineering alone. For instance, banks are now also requesting more power, in order to perform purely financial calculations, measure and manage risk, based on mathematical models. Digital simulation also makes it possible to design new products, put them on the market more quickly, with lower development costs and extremely large savings, like in the automobile or oil industry.

Our objective is to win over 10% of the European market. We want to strongly increase our direct sales figure in scientific computation and penetrate new accounts thanks to our "Architect of an Open World" strategy, based on the use of standard components around our NovaScale servers[2]. They are based on Intel processors (Itanium and Xeon). Our inter-connection networks between the servers, used to bring them together into clusters, are also based on market technologies. Our added value comes from that very platform technology, its integration and its software component. To wit, we have developed a software know-how that makes it possible for us to use the platforms to their fullest, optimise, rationalise and simplify the use of our customer applications.

L. R. What did the Tera-10 contract do for the company?

P. M. Tera-10, the leading platform in Europe in terms of power, boosted our credibility. Our objective now is to deploy that know-how across every industry, from the automobile industry to the oil industry, natural sciences, meteorology, finance and others. CCRT's order for a 43-teraflop[3] supercomputer, placed last December, is an important step in that deployment. The machine, which is scheduled to become operational in 2007, will be one of the most powerful in the world. It will be available to the scientific and industrial community. Other clients have already placed their trust in us: the University of Manchester (Great Britain), Dassault Aviation, the Polytechnic University of Valencia, the Universities of Utrecht and Reims, the Southampton Oceanographic Centre and Pininfarina (Italy). Tera-10 provided the structure. We now sell the building blocks to many clients in the industrial and academic communities. Above and beyond national sovereignty issues and the economic and industrial aspects, it is also important to see how research can shape Europe's competitiveness. Only 15% of the supercomputers used by the Top 500[4] are located in Europe, and 49% of them are used by industrial players. The others are in the academic, research and defence communities.

Europe's delay in embracing intensive computing, whether in terms of use or equipment, is thus cause for all the more concern. The United States have realised that scientific computing is vital for the competitiveness of enterprises and for national security: they provide generous

> **"We now want to deploy this know-how across every industry"**

funding to their industrial players, with two projects at DARPA (US Defence Advanced Research Projects Agency) worth USD 250 million each, one for Cray and the other for IBM, to produce petaflop[5] machines by 2010, which will be nearly twenty times more powerful than the Tera-10. This is why it is critical for Bull, the only industrial player in Europe developing technology in the field of scientific supercomputers, to maintain its competitive advantage. Yes, the

© Jean-Michel RILLON

(1) CCRT: Computing Centre for Research and Technology (Bruyères-le-Châtel, Essonne).

(2) NovaScale servers: http://www.bull.com/novascale/index.php

(3) Teraflops: one thousand billion operations per second

(4) Top 500: ranking of the 500 most powerful computers in the world

(5) Petaflops: one million billion operations per second

European Commission also has supercomputer projects. However, China and Japan have projects of the same magnitude, if not greater...
What is at stake in scientific computing goes far beyond the enterprise level, reaching the national and supranational dimension.

L. R. In the international race, where you are facing competitors such as IBM, Cray, SGI or Dell, to name only a few, what are your distinctive features?
P. M. There are three major limits on installing large computer systems: the significant floor space that they require, high electricity consumption and thermal dissipation. Bull offers know-how in the filed of cooling, energy management optimisation and platform density. Our added value can also be found in the integration of standard market components around a server architecture that we developed (our next generation will be designed around Intel's CSI bus) and the development of intellectual property around the software battery. The fact that we use Linux and open-source software makes it possible for our clients to work on the source code of the software suite that we develop. Our competitors have more proprietary technologies, have their own micro-processors, but must deal with the heavy burden of developing their own technology, sometimes with their own interconnection network and their own operating system. Our focus on the European market is also an advantage. Lastly, we offer real added value on components that combine with the basic building-blocks, such as our dedicated accelerators or the way our servers are packaged, in order to cut down on thermal density. The accelerators are specialised in computing specific algorithms, for instance the FFTs (Fast Fourier Transforms), which are very-much used in seismography or matrix transposition or inversion in fluid mechanics. They make it possible to compute 10 to 100 times faster than with the usual microprocessors, but consume 2 to 5 times less electricity. What makes us stand out, now and increasingly as we move into the future, is the quality of the integration system used to bring together all of our supercomputer hardware and software components, and the quality teams that work with our clients to optimise their applications on our servers. **Interviewer : Didier Gout**

Ter@tec

A benchmark player in European intensive computing

Christian Saguez is Chairman of Ter@tec

The strategic importance of digital technologies and, specifically, that of high-performance simulation, is recognised by every player in the scientific, economic and industrial worlds. Since August 2005, Ter@tec Technopole and its 43 partners[1],companies and research organisations, have offered a unique setting for taking up the challenges in this area, by uniting all of France's intensive computing skills around a number of large collaborative R&D projects. Ter@tec provides access to some of the most powerful data computing and storage resources in Europe. The business dynamic it generates creates large numbers of jobs and is expected to help France and Europe return to the fore, in an extremely active global competitive arena – a position they never should have left.

Large-scale Research Projects

As a member of the System@tic competitiveness cluster, Ter@tec is a major player in large-scale projects such as FAME 2 (Flexible Architecture and Multiple Environment) aimed at developing a new generation of servers for intensive computing and information processing; CARRIOCAS (Distributed computation over an ultra-high optical internet network), aimed at developing and building high-speed networks, at 40-gigabits per second; IOLS (simulation infrastructure and software tools). Together, those three projects are expected to provide all of the basic tools required for high-performance computing. Two new projects are also being launched: POPS (petaoperations per second) and EPHOC (High-Performance Environment for Optimisation and Design).

Many projects are supported by ANR[2]. Examples include PARA (Parallelism and Improving Application Output), LN3M (New-Generation Software for Multi-Scale Material Modelling), NUMASIS (Seismology Application Performance Optimisation) and the SCOS platform, which is aimed at uniting French initiatives in open-source software for intensive computing. In addition, joint research laboratories have been instituted, between CEA, the University of Versailles Saint-Quentin (UVSQ), the Paris Ecole Centrale and ENS Cachan.

Unique Infrastructures

Now that the Arpajon Municipality Community (CCA) and CEA have purchased an activity zone in Bruyères-le-Châtel, Ter@tec will be able to develop significant business operations. An extensive real estate complex will be built to host enterprises and laboratories, as well as a computation centre, where petaflop machines will be installed as early as 2010, thereby creating the first-ever European intensive computing technopole. As early as 2007, the available computing power will exceed 50 teraflops, with the CCRT centres (Joint Research and Technology Centres) Bull and HP; the petaflop mark should be reached in 2011, in synergy with European initiatives.

Training and Promotion Action

Ter@tec has established a collaboration agreement with three Master's Degree programmes specialising in high-performance simulation: MN2MC (digital methods for models in continuous environments), M2S (modelling and simulation) and COSY-AHP (from concepts to systems – high-performance architecture). Continuing training activities will also begin this year, targeting in particular SME-SMIs where proficiency with intensive computation tools has become a necessity. Moreover, Ter@tec launched, as early as 2006, a "European High-Performance Computing Forum", designed around a series of conferences and extensive exhibit hall. The 2007 forum will take place on 20 June, at the University of Versailles Saint-Quentin. **C. S.**

(1) Ter@tec is composed of IT companies (Aria Technologies, Bull, Cluster Vision, CS, Data Direct Networks, Distène, ESI Group Eurobios, Fluent, Fujitsu, HP, Intel, Numtech, OpenCascade, Oxalya, Serviware, SGI, Sun, Transtec), industrial players (Airbus, Bertin Technologies, CSTB, Dassault Aviation, EDF, Principia, Snecma, STMicroelectronics, Total), universities and research laboratories (CEA, Cenaero, Cerfacs, CNRS, École centrale Paris, ENS Cachan, IFP, Inria, INT, École des mines Paris, Supelec, UVSQ) and local governments (CCA, Ollainville City Council, Bruyères-le-Châtel City Council).

(2) ANR: National Research Agency

Last July, the Government decided to create a civil-society company dedicated to intensive computing, going by the name of Genci. François Goulard, the Minister of Research, describes its objectives.

Genci Major National Facility for Intensive Computing

Intensive computing is essential if we are to benefit from the advances achieved in modelling and digital simulation. Like theory and experimentation, digital simulation is a must, both in research and in industry. It provides decisive leverage in fields such as the climate, astrophysics, combustion, aeronautics, oil production, nuclear technologies, nanotechnologies, chemistry, pharmaceuticals, and more.

To help France return to computing power compatible with its ranking and power in the scientific and industrial world, and holds its own on the global competitive market, we have set up an organisation responsible for establishing strategic directions and making top-priority investments.

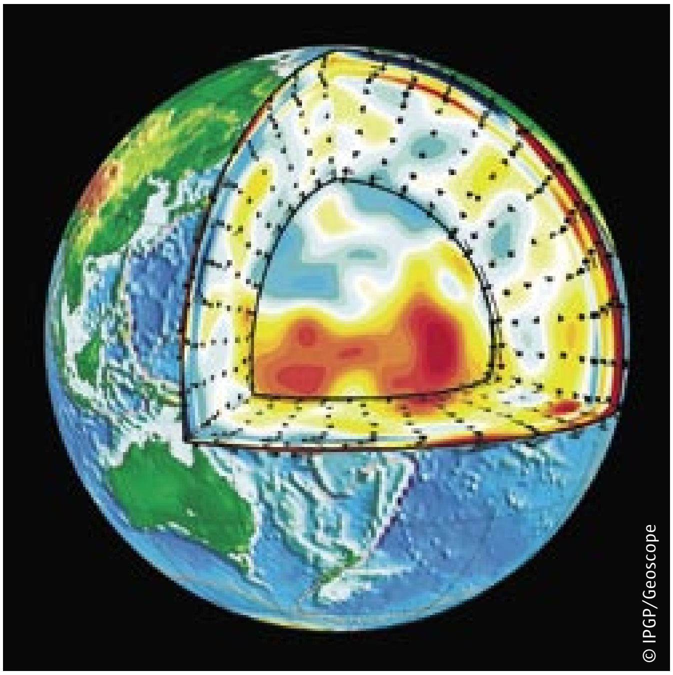
The State is the founder of civil company Genci, alongside CEA (20%), CNRS (20%) and universities (10%) and will provide 50 percent of its annual budget. The most important pillar in the system for rationalising France's investment policy on major computers and significantly increasing France's resources in this sector, Genci is in charge of setting up and coordinating the main facilities at national civil centres. Its facilities will be open to all interested scientific, academic or industrial parties. Cooperation programmes between public and private teams will be given priority. Its annual budget will be EUR 25 million, or twice the resources mobilised by France on intensive computing until now.

By choosing a civil company structure, already adopted for other major scientific infrastructures such as the Soleil synchrotron in Essonne, or Institut Laue-Langevin in Grenoble, France ensured that the company would have a more flexible management system and be able, in particular, to take out stakes in international or European infrastructures of the same kind. This also makes it easier for research partnerships in advanced digital simulation to come about with regional or local authorities and major advanced-technology companies.

At the same time, the second pillar in the system, the strategic committee on intensive computing, will soon be instituted. Its aim is to foster forward-looking consultation on the computing requirements of the entire French scientific community, provide fodder for the elaboration of national policy in this area, unite the relevant stakeholders and coordinate investment.

With an investment plan of EUR 25 million per year over a four-year period, France will be at the same level as the top European countries and make a harmonious fit with the Community policy underway in this area. **F. G.**

François Goulard
François Goulard, Minister of Higher Education and Research



Simulation in Earth Sciences requires increasing computing power. On this cross-section of the globe, focused on the Pacific, the disparities in speed in the Earth's crust were modelled by tomography (slow areas in red, quick areas in blue). Based on this, convection movements can be deduced (velocity field shown with arrows)

WEB

Associations

ORAP
www.irisa.fr/orap
Association aimed at spreading the use of parallel computing through events and newsletters.

TERATEC
www-dam.cea.fr/statique/ouverture/teratec.htm
With 43 members, Teratec is the benchmark association for users and suppliers of high-performance computing equipment.

Information Sites

CLUSTER-HPC.COM
www.cluster-hpc.com
With its daily information and weekly newsletter, ClusterHPC is the French reference in high performance computing. .

CLUSTER MONKEY
www.clustermonkey.net
A comprehensive technical resource site about clustering. A good source of information, in English.

THE TOP 500 SUPERCOMPUTERS
www.top500.org/
A ranking of the 500 most powerful supercomputers in the world.

CNRS COMPUTING GROUP
http://calcul.math.cnrs.fr
A set of resources about high-performance computing in France. Of note: a link to the CIEL project (Online IT codes).

Benchmark Players

BULL
www.bull.com/fr/hpc
Learn all about the French manufacturer's HPC product range. With the NovaScale servers, it delivered the most powerful French computer.

CERFACS
www.cerfacs.fr
Presentation of research findings from European Centre for Research and Advanced Training in Scientific Computing: parallel algorithmics, image and signal processing and processing for associated applications.

COMMISSION ON ATOMIC ENERGY
www.cea.fr
Visit the CEA's main site and find complete reports on nuclear power, controlled fusion, nanotechnologies, the climate and a description of the organisation's main research areas.

CRAY
www.cray.com
A historical player in computing, with massive parallel platforms.

HP
www.hpl.hp.com/techreports
Here, you will find publications and offers regarding HPC from HP, supplier to over 30% of the global market.

IBM
www.research.ibm.com
A goldmine of information about the HPC research programmes conducted by IBM.

INRIA
www.inria.fr
Thematic reports on modelling of living organisms, neurosciences and algorithmics, grid computing and a description of the Institute's programmes on high-performance simulation (Clime, Gamma, etc.).

FRENCH OIL INSTITUTE (IFP)
www. ifp.fr
Overview of IFP activities, highly active on issues in high-performance computing and involved in many collaborative projects: IOLS, FAME2, Scilab, SCOS and others.

MICROSOFT
www.microsoft.com/france/hpc
Learn all about Microsoft's new computing offer, available since 2006: Windows CCS (Compute Cluster Server).

ONERA
www.onera.fr/synindex/calcul-hautes-performances.html
Computing resources and information selected by ONERA.

OXALYA
www.oxalya.fr
HPC technology provider: clusters, administration solutions and remote access portals.

SUN
www.sun.com/servers/hpc
A complete listing of the servers and references offered by SUN.

Major Projects

MARE NOSTRUM
www.bsc.es
The leading computing centre in Europe (and Number 5 worldwide), the Barcelona supercomputing centre offers researchers the opportunity to take up the major challenges in earth, life or computer sciences.

FAME2
www.fame2.org
The FAME2 project focuses on a new generation of servers for intensive computing and data processing. Headed by Bull, the project has 13 partner.

SCILAB
www.scilab.org
Scilab is a digital open source computing software system with a powerful development environment for applications in science and engineering. The consortium, led by INRIA, is composed of 23 members.

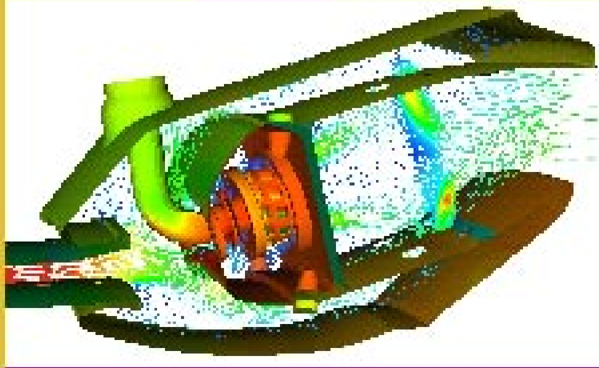
SCOS
www.oscos.org
The SCOS project standardises scientific computing platforms and units computing code development initiatives, including, in particular, open source codes. Led by Oxalya, the consortium is made up of 22 partners.

Ter@tec

EUROPEAN EXPERTISE CENTER IN HIGH PERFORMANCE COMPUTER SIMULATION

A must to improve competitiveness and help companies face scientific and technological challenges

- Bring major industrial companies together with large laboratories
- Facilitates the initiation of common R&D projects among its members
- Offer access to leading-edge supercomputer technology and data treatment facilities, amongst the most powerful in the world
- Provide appropriate trainings and seminars to enhance expertise



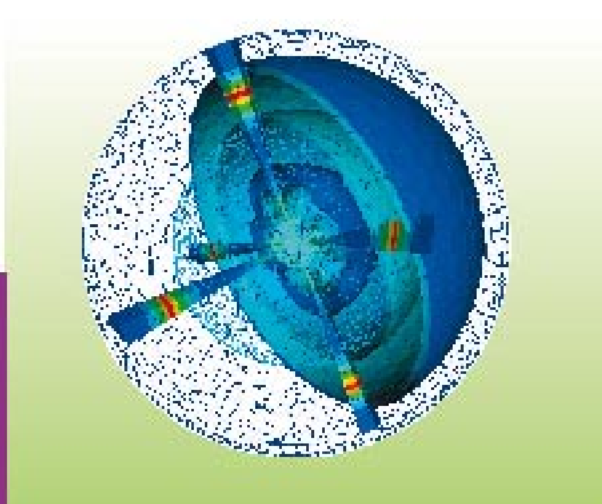
TER@TEC PARTNERS:

Airbus, Aria Technologies, Bertin technologie, Bull, CEA, Cenaero, CSTB, Cerfacs, Cluster Vision, CNRS, CS, Dassault aviation, DDN, Distène, ECP, EMP,ENS-Cachan, EDF, ESI Group, Euroblos, Fluent, Fujitsu, HP France, IFP, INRIA, INT-Evry, Intel, Numtech, Open Cascade, Oxalya, Principia, Safran, Serviware, SGI, ST, Sun, Supelec, Total,Transtec, UVSQ.
Communauté de communes de l'Arpajonnais, Ville de Bruyères-le-Châtel, Ville d'Ollainville.

[redacted]

[redacted]

[redacted]



Ter@tec is a key component of the System@tic Paris-Région competitiveness cluster

Contact: Ter@tec – domaine du grand rué – 91 680 Bruyères-le-Châtel - FRANCE
Christian Saguez – President of the Ter@tec Association
Phone: +33(0)1 41 13 12 86 – christian.saguez@ecp.fr


Pushing back the boundaries
of European research at the rate
of 100,000 billion operations a second.
intel
Itanium 2
inside
Dual-core.
Do more.
Bull, partnering top research establishments
With its Bull NovaScale servers, the scientific computing center at the French Atomic Energy Authority (the CEA) now has over 100 Terabytes of power, and 100,000 billion operations a second at its disposal. Thanks to Bull, the research world can rely on industry standards and open IT environments… and on the expertise of the only truly European hardware manufacturer, trusted by numerous research centers and industrial customers. Putting Europe in the best position to win the High-Performance Computing war. To find out more, click on: www.bull.com
Bull
Architect of an Open World™
TBWA\CORPORATE - © Geoatlas