La Recherche

cea

High Performance Computing

Tera 10: beyond 50 thousand billion operations per second

intel®
Leap ahead™

Bull
Architect of an Open World™

The **lightning** growth of intensive **computing**

Today, researchers, engineers and entrepreneurs in virtually every field are finding that high-performance computing (HPC) is starting to play a central role in maintaining their competitive edge: in research and business activities alike.

Advances in scientific research have always depended on the link between experimental and theoretical work. The emergence of the supercomputer opened up a whole new avenue: high-powered computer simulation. By approaching many traditional research and engineering questions from a new angle, simulation lets us explore even the most complex phenomena.

We can estimate the cost of a typical basic arithmetic operation by comparing mechanical machines from the 17th and 19th centuries (developed by Pascal and Babbage respectively), with modern computers that are based on von Neumann's architecture (1945) and exploit the capabilities of semi-conductors (1956). This cost has fallen by a factor of 10^{15} between the earliest machines and the latest generation of computers, with their intrinsic power expressed in teraflops (or 10^{12} operations a second).

In the mid 1970s, Cray developed an architecture promising a million operations a second for a cost of just a million dollars. Thirty years on – with the advances in electronic circuit technology and the science of computer simulation – a million times that processing power is now available for an equivalent cost. Today's most powerful machine delivers 280 teraflops, and there are plans in the near future for a 1 petaflop machine in China and a 5 petaflop machine in Japan.

Where will this progress lead us? The petaflop (10^{15}) machines will soon be with us, which suggests that high-performance computing will continue to progress at the same kind of rapid pace for years or even decades to come. The performance of algorithms is also increasing at a comparable rate. For some linear algorithms, the number of processing operations required has fallen by a factor of 10^8 over the past 35 years. No matter what kind of hardware is involved, high-performance computing is a major strategic challenge on all three key fronts: industrial, technological and scientific.

However, a recent conference held at CERFACS***** concluded that a number of looming questions remain unanswered. If these are to be tackled effectively, depending on the subject they will require a 100 to 10,000-fold increase in computing capacity within the next 10-20 years.

In this frenetic race, there is general agreement that France has not progressed as fast as we would have wanted over the past few years. In order to be effective and consistent, it is not enough just to give the scientific community the most powerful computers. Active support for research and higher education into algorithms and the associated computer sciences is also essential. In 2005, we saw the first signs of this starting to materialize, with a broad, inter-ministerial consultation and the consolidation of a number of computing centers at the national, and even European, level. This issue surveys these latest developments and various proposals for the future, while also casting a backwards glance at the main steps leading up to the situation we know today. We hope this contributes to mobilizing the forces needed to achieve the progress we're all hoping for in scientific computing!

So here we are, in 2006, in the middle of teraflop era. By 2015, will we already have entered the exaflop (10^{18}) era? And where will France and Europe stand? There is everything to play for in this debate.

Bernard Bigot

Bernard Bigot
High Commissioner for Atomic Energy.

*CERFACS :
Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, The European Centre for Research and Advanced Training in Scientific Computation.

High performance computing
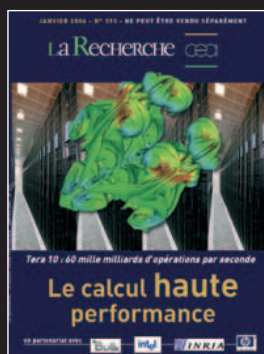
This special issue is adapted from "Le calcul haute performance", which has been distributed in January 2006 with *La Recherche* n° 393 and realised in partnership with Bull, Intel, Inria and Hewlett-Packard.

p. 8

p. 16

Summary

p. 20

p. 26

p. 34

p. 50

Nine questions
about numerical
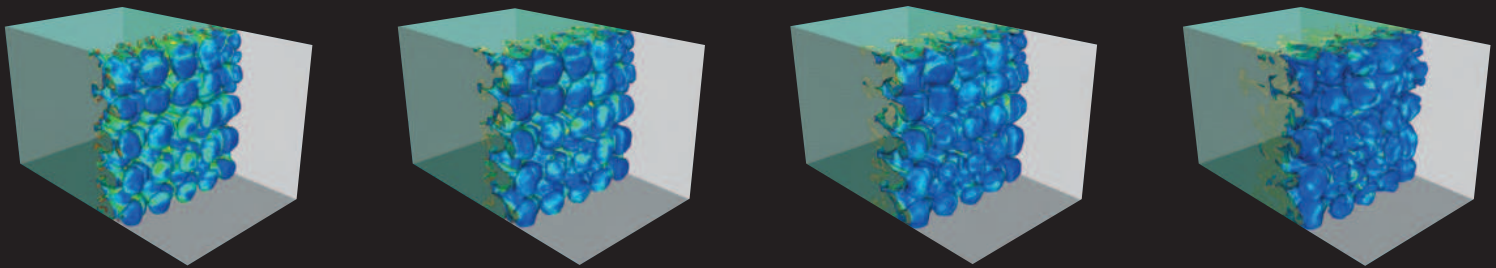
Jean-Michel Ghidaglia

is a professor at École Normale Supérieure de Cachan and scientific director of *La Recherche*.

Benoît Rittaud

is a lecturer at Université Paris-XIII.

It saves time and a lot of money especially for automobile and aeroplane manufacturers. It allows meteorologists to predict what the weather will be like in five days' time. The power of numerical simulation is increasing in step with the power of computers to become an indispensable tool in research.

Is it just calculation?

No! Because we can't use numbers in simulations without first having a mathematical model. So modelling and simulation always go hand in hand. Imagine, for example, that we want to simulate the movement of the Planets around the Sun. First, we need to work out a model of the phenomenon, that is, determine the basic parameters that describe it in both a simple and realistic way. The constraint of simplicity means that the Sun and each Planet are treated as finite points. The result is that, knowing the position and speed of eac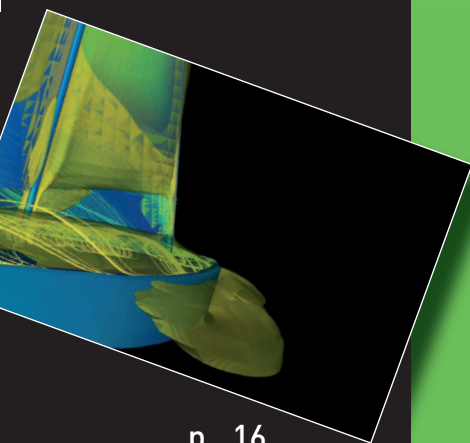h star, the entire movement can be characterised. To complete the model, equations are written that quantitatively describe the reciprocal attraction between the Planets and the Sun. Depending on the degree of refinement desired for describing this phenomenon, solving it will be more or less complicated. If we hypothesise that the Planets' masses are sufficiently low that we can ignore their reciprocal interactions (keeping only those which
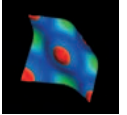
simulation
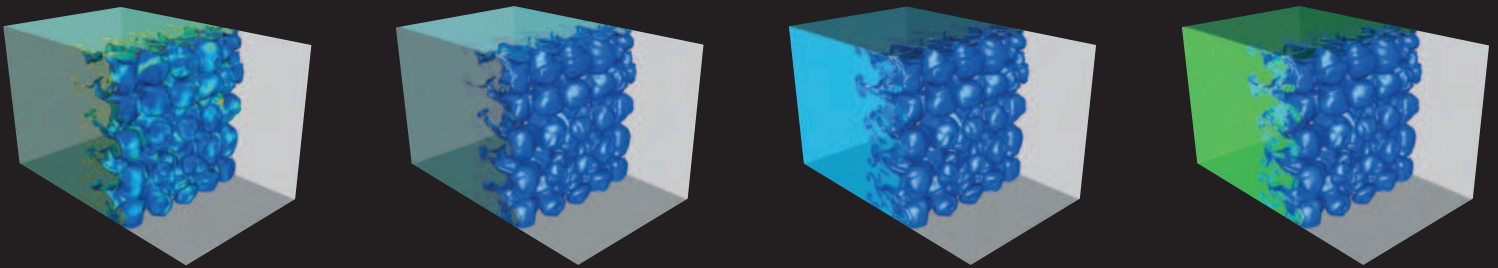
WHAT HAPPENS AT THE INTERFACE BETWEEN TWO DIFFERENT MEDIA WHEN A SHOCK WAVE SPREADS? A zone of mixing forms, the turbulence that appears and changes over time can be reproduced by computation using the Triclade code. ©MIREILLE BOULET/CEA

relate to the Sun), then solving the classic mechanics equations can be done explicitly.

More often however, as in the case where we take into account the gravitational attraction between all the Planets, theoretical constraints prevent the model's equations from being completely solved, i.e., we can't obtain an explicit answer to the problem. So we need to substitute this inexpressible answer with an approximation. It is only when, after starting from these calculations, the trajectories of different planets are shown on the computer screen that we can say we have created a true digital simulation.

Is a computer really needed?

In the 18th Century, the Frenchman Pierre Simon de Laplace was already making simulations even though this term was not used at the time. He used a simple version of the planetary model that we just described and did his calculations by hand referring to numerical

[1] Jean-Michel Ghidaglia, 'Grenouilles in silico', *La Recherche,* October 2004, p. 96.

tables like the old logarithm tables. But the more and more pressing need for simulations of models in physics led to the invention of programmable calculating machines, future computers. So while today simulations are reliant on the tools of information technology, the fact remains that from a historical point of view, computers were born from the idea of simulation! It was in the context of the 'Manhattan project', the military nuclear programme that gave the United States the first atom bomb, that the American mathematician of Hungarian origin John von Neumann, wanting to calculate the variations in pressure and temperature in the immediate vicinity of the bomb, set about building the first computer which saw the light of day in 1946.

The ever increasing demand for simulation was also the reason for the great advances in computer design. It was the enormous need for numerical simulation for weather forecasting that led the American computer
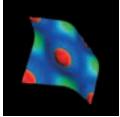
scientist Seymour Cray to design vector computers* in the 1980s.

Weather forecasting, in fact, relies on the changes in countless parameters measured in meteorological stations such as wind speed, temperature or atmospheric pressure [1]. Not only does this create a colossal amount of data to process, but also all the calculations need to be done in a very short time: finding out tomorrow's weather should be done…. before tomorrow! Besides these conceptual breakthroughs which affect the architecture of computers or programs, improvements in the equipment itself, the hardware, has increased the capacity for numerical simulation exponentially. Simulations that were totally unimaginable with the most powerful machines only 30 years ago can be carried out using the capacity of a basic personal computer available in the shops today.

Can it replace experiments?

In many fields, the power of current supercomputers means that simulations may rival certain experiments in terms of quality and cost. For example there are an infinite number of possible shapes for an aeroplane wing and knowing which one to choose for giving maximal lift* is very tricky. The body of wind tunnel experiments needed to determine this would take several days. It is less costly to spend the time developing a good model and making it function on a computer. Starting from the shape designed by the engineer, the optimal form can be determined in several hours by numerical simulation. This is what is referred to when we talk of using 'digital wind tunnels' for designing a suite of computer programs that recreate a physical situation in aerodynamics and in other contexts.

However accurate the model of our aeroplane wing, the numerical data necessary for the calibrating the simulation can only be obtained from experiments. For example, certain coefficients will occur in the model's equations. Experiments, therefore, indirectly feed into simulation. Exploring the numerous situations that simulation make possible allows unexpected behaviour to be observed or predicted. This sometimes suggests new experiments and so our knowledge of physics advances. So far from usurping experimentation, the role of simulation is rather to give a new take on reality, moving on from the traditional partnership between theory and experiment that has been steadily developing over the centuries. Numerical simulation is a third way of studying phenomena, complementing the other two, and often called *in silico* study because the basic material of computers is silicon.

How are models developed?

By translating and simplifying a phenomenon in the

***A vector computer** carries out basic operations on a set of numbers (8 for example) simultaneously, unlike a sequential computer which carries them out one by one.

***Lift** is the vertical force which acts on an aerodynamic body placed in a current of air. This force keeps a plane up in the air.
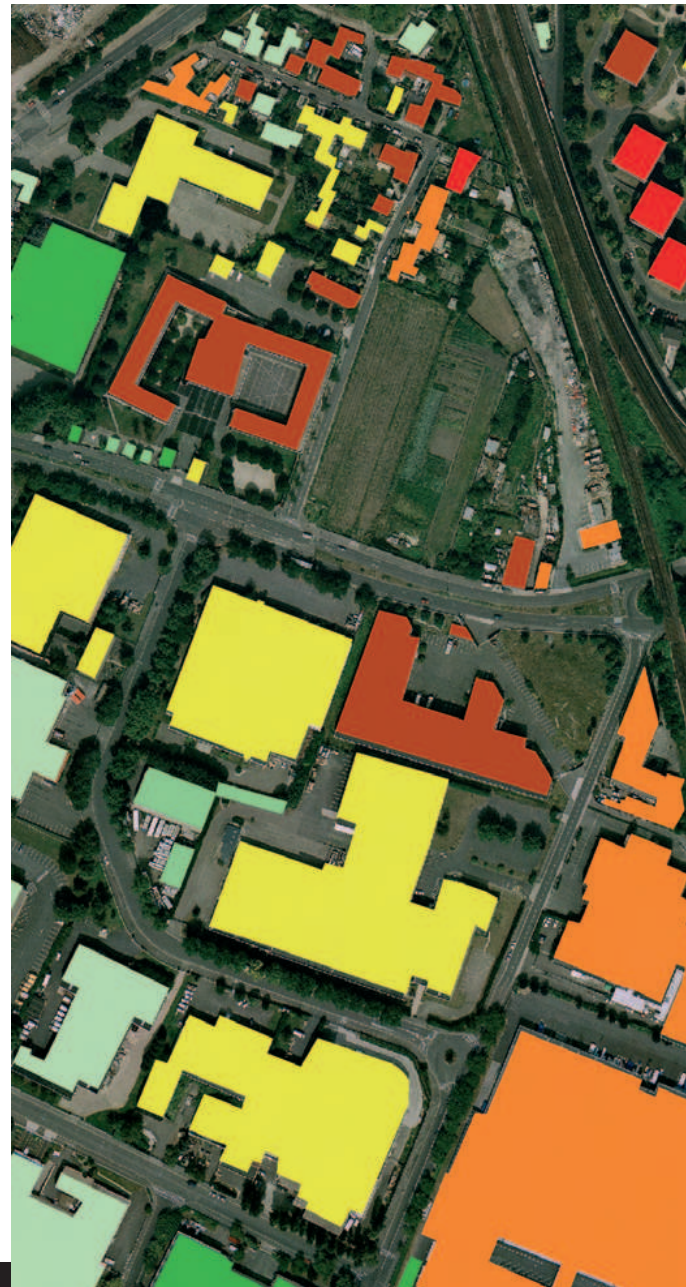
THE MITHRASIG PROGRAM, launched in 2005 by Institut Géographique National, is used to draw maps of the noise level on the scale of an entire town in 2 or 3-dimensions. Cross-referenced to demographic data, the populations exposed to the most noise can be localised on these maps.
©IGN

language of mathematical equations (the only language a computer understands), a model may come from careful observations. For example, the great mathematical treatise of the Greek astronomer Claudius Ptolemy in the 2nd Century (a book better known under its Arabic name *Almagest*) described a model designed to 'save the appearances', that is to note the position of planets in the sky with some precision without really using a theory (the idea of scientific theory was however basically born in ancient Greece).

From the 17th century, models were to become more and more inspired by experiments, the idea of the scientific experiment only being fully accepted after the era of Galileo. Eventually, models were to be deduced from theories, themselves conceived without
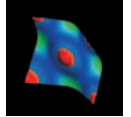
[2] P. Jensen
and X. Blasé,
'Les matériaux
virtuels',
La Recherche,
April 2002,
p. 40.
[3] http://
setiathome.
berkeley.edu/
sah_about.php

in the physical behaviour of water and concrete allow us to make the model's equations more specific. What fundamentally differentiates one medium from another are the internal movements of molecules in the medium under stress or strain as from an external force. It is clear that from this point of view, a solid and a liquid behave in very different ways. The third step often requires experimentation, particularly so that numerical values for the model's parameters can be determined.

Can we make predictions using simulation?

Like a crystal ball, a predictive model (and the simulation that goes with it) can be used to anticipate the future of a system or the behaviour of the system in another configuration in which it has never been observed. Predicting novel situations is an essential asset of digital simulation, for example, in the context of nuclear safety. It goes without saying that predicting incidents, ways of dealing with them, and even the risks at stake following a dangerous manoeuvre, are so much part and parcel of our safety that we can't imagine they would be better understood using real experiments. The 'serious accident codes' are simulation program platforms that enable the constructor or the nuclear safety authorities to test the behaviour of an installation in extreme case scenarios. Evidently, no experimental protocol can ever totally validate these programs.

In this context, we expect simulation to give qualitative results to confirm that we have envisaged certain scenarios correctly and that, for example, the conduct for operators to adopt is in accordance with the gravity of an incident that might be minor in itself, but that could have more serious repercussions.

For a predictive model to be a good starting point for a simulation, it is best to be sure that there is a very good understanding of the fundamental phenomena at work. Otherwise extrapolation is used but its reliability is more or less random. When it is not possible to do otherwise, i.e., it is too difficult (or too expensive) to start from a complete knowledge of the phenomenon, we turn to 'black boxes'. A 'black box' is an *ad hoc* formula, i.e., it is not specific to the situation, but when we key in the input parameters, it gives the expected output numbers in return. However, it is more 'reassuring' to be able to justify a formula by theoretical arguments rather than by a simple calculation of empirical data.

Who uses numerical simulation most?

We have already mentioned several areas: celestial mechanics, nuclear physics, weather forecasting, aeronautics and theoretical physics. Broadly speaking, it's physics that monopolises the large majority of si-

recourse to experiments. Albert Einstein, for example, came up with the theory of general relativity in 1915 by a purely intellectual exercise, a 'thought experiment' (a process also used by Galileo). The equations of the theory produce a model of the curvature of light rays when they pass in the vicinity of a body of very high mass. This could only be compared to experimental results in 1919 during a solar eclipse. Today several physics theories, like string theory, provide models that can not yet by tested in experiments. So we are left with the possibility of using digital simulation to test these models or to imagine the ideal experiment that would allow us to have a more clear-cut idea about the theory **[2]**.

To obtain a model in physics, for example, to describe the flow of water in a pipe, we take three main steps. The first is to take into account the fundamental principles like the conservation of mass, energy, etc. The second is to adapt this set of equations to the situation in question. The basic equations describing the flow of water in a pipe or the behaviour of a concrete beam under strain are the same; but differences

© NASA
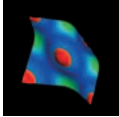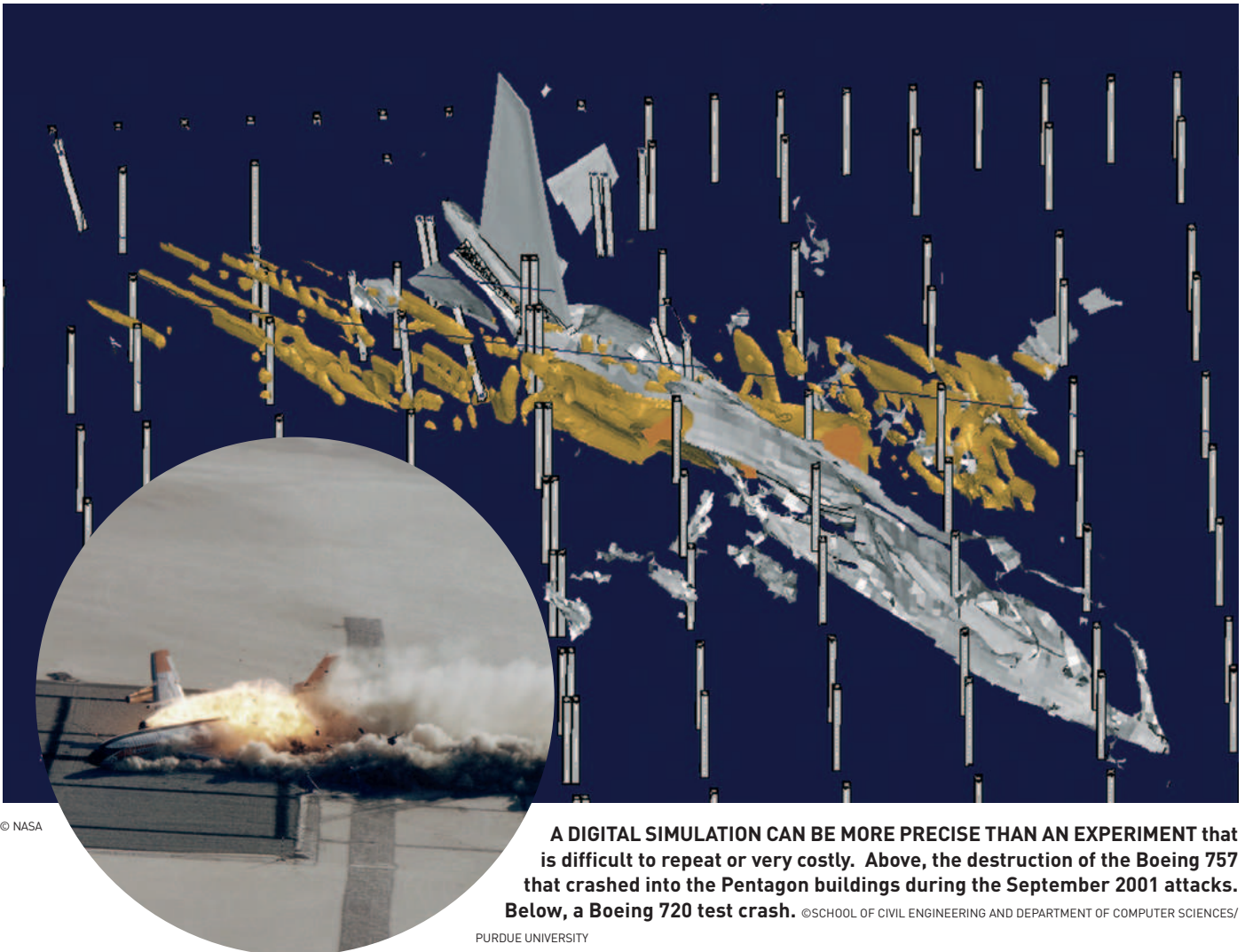
A DIGITAL SIMULATION CAN BE MORE PRECISE THAN AN EXPERIMENT that is difficult to repeat or very costly. Above, the destruction of the Boeing 757 that crashed into the Pentagon buildings during the September 2001 attacks. Below, a Boeing 720 test crash. ©SCHOOL OF CIVIL ENGINEERING AND DEPARTMENT OF COMPUTER SCIENCES/ PURDUE UNIVERSITY

mulations. Still in this area, so-called *ab initio* calculation, which involves finding the properties of everyday materials starting from the basis of atomic physics, is flourishing.

Quantum mechanics governs the behaviour of matter at the atomic scale and the *ab initio* calculation programs can, for example, predict the behaviour of a new material at our scale *in silico* so as to only move into the very costly synthesis or production phases once we are satisfied with the results. Other areas are using simulation more and more. For around 15 years, the financial world has been a major consumer. More recently, biology has become a very large user of computing time. For example, digital simulation is a compulsory step in both visualising the structure of proteins in space starting from experimental measurements and then deducing their biological properties from the structure. Note though that genomics relies on computing methods, but in this case it is data mining and not digital simulation.

Is virtual reality digital simulation?

Virtual reality does in fact implement simulation methods. In this case, the aim is to produce an image which tells our brain to analyse it in an identical way as the real scene would be analysed. But reality is generally complicated… The simple simulation of the movement of a man walking is not obvious with the swinging of the arms, the reaction of the neck which keeps the head horizontal, etc. So it is necessary to find a compromise between the necessary simplicity of the models (to allow the computer to process the data in a reasonably fast time) and the unforgiving judge, our eye, which is so used to seeing real people walking down the street that it becomes capable of relating any differences, even slight ones, to the plausibility of the simulation. Being the exacting viewers of video games and cinematic special effects that we are, forces models to become more and more precise. This demands not only having increasingly powerful computers but also continuously honing the simulation methods.
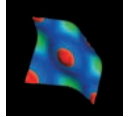
What exactly does grid computing mean?

A computational grid is a set of machines, sometimes geographically very far apart, which cooperate via a communication network to produce a numerical result. To understand the point of this kind of architecture, remember that a normal computer's driving force is a processor which executes the program's instructions one after another.

We call them 'sequential computers'. In the 1980s, a new architecture came out of computing research laboratories which led to computers having several hundreds or even several thousands of processors – parallel computers. Multiple processors in the same setting are linked by a physical network (copper wires or optical fibres). As each processor receives work to do at a given moment, several instructions can be carried out at the same time.

When an instruction requires a result already obtained by one of the processors, it executes it. But when the result is yet to arrive it must wait to be able to execute it. Another way of saying this is if a programme requires 24 hours on a monoprocessor computer, there is no way of knowing that it would need only 6 minutes on a 240 processor machine. Grid computing involves creating a similar architecture to that of parallel machines by linking machines (mono or multiprocessors), which could be at the four corners of the world, by a network like the internet. The efficiency of parallel machines and grid computing depends therefore on the program which is being executed. Their yield is highest when the calculations are independent from one another.
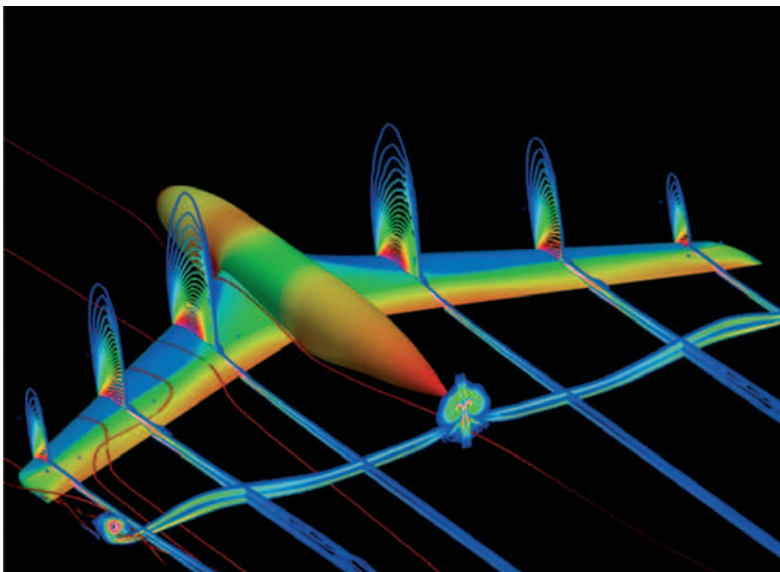
IMPROVEMENTS IN TURBULENCE MODELS and the increases in computation power means that today the airflow around an aeroplane wing can be simulated accurately, effects which influence the stability. © ONERA

One of the first applications of grid computing is the Search for Extra Terrestrial Intelligence (SETI) project. Launched in 1999, the aim is to analyse the signals captured by radiotelescopes so as to detect a possible sign of intelligent extra terrestrial life [3]. Each signal must be independently sounded out in the very many frequency ranges, fully justifying the use of grid computing. For SETI, this grid involves several millions of PC owners volunteering their machines to be linked up to SETI's machines. Other projects are in progress. For example, Decrypton has already compared more than 550,000 proteins in two months thanks to the connection of 100,000 PCs, each of them contributing 10 hours of calculations. It would have been necessary to wait 1,140 years to carry out the same operations with a single computer!

Can anything be simulated numerically?

There are three kinds of limitations to numerical simulation. First, certain models require computing power that is currently not available. This is the case for simulating how a nuclear power plant works starting from physics equations; it's totally out of our grasp at the moment. It is possible to simulate the functioning of the core (neutronics), a gas exchanger (thermohydraulics) or other components, but the totality of the installation is unattainably complex for today's computers. So we have recourse to 'system codes', which are black boxes feeding in codes that simulate each component (the core, the exchanger, etc.) starting from models obtained using equations.

The second limitation is more basic. It results from the fact that some phenomena are poorly understood and are difficult to translate into equations, which, remember, are the only way to dialogue with a computer. Pharmacists do not know how to model the effect of a molecule on the human body, for example, as we do not have a satisfactory mathematical model of the chemistry of a living being.

The third limitation is a theoretical one. There are some mathematical models of physical situations for which no effective method of solving them is known. By 'effective', we mean something that can be solved by a computer in a reasonable time, several days at worst. This does not relate to the computing capacity at any given instant, as for the first limitation mentioned, but to the fact that the number of operations required to resolve the model increases exponentially as the degree of precision increases. *Ab initio* calculation is part of this and is just setting out: only very simple molecules can be studied with this technique. ∎

J.-M. G. et B.R.

www.intel.com/technology/magazine/computing/tera-scale-0606.htm

The Intel® Tera-scale Computing Research Program looks to the future of computing – about a decade from now – when tens or hundreds of processor cores will work together in one system. Intel's researchers are addressing the hardware and software challenges of building and programming systems with dozens of energy-efficient cores.

QUOTES

"Our tera-scale research is directed toward maximizing the advantage of parallel computing for the personal computer."
Joe Schutz, vice president and co-director of the Microprocessor Technology Lab
www.intel.com/pressroom/kits/bios/jschutz.htm

"Looking forward, Intel will really focus on end-user usage models and business value to take advantage of this energy-efficient tera-scale computing."
Steve Pawlowski, Intel senior fellow and Digital Enterprise Group CTO
www.intel.com/pressroom/kits/bios/pawlowski.htm

Intel's Tera-scale Research Prepares for Tens, Hundreds of Cores

Overview:
Bringing Users More Powerful, Energy-Efficient Computers

Intel processors with two cores are here now. Quad-core processors are right around the corner. And the number of cores on a chip will continue to multiply in the coming years, launching an era of vastly more powerful computers.

With these important trends in mind, Intel researchers are today working to shape future Intel® microarchitec-tures. This area of research has been dubbed "tera-scale" computing. Intel has over 80 R&D projects worldwide in the tera-scale area, where researchers are addressing the hardware and software challenges of building and programming systems with dozens of energy-efficient cores.

" 'Multi-everywhere' and valued performance – this is where tera-scale computing comes in: 'multi' to the fullest extent, beyond multithreading and multi-core, to a large number of energy-efficient cores accompanied by fixed function hardware, to pro-vide unprecedented performance in the power and cost envelope for the future applications," explains Shekhar Borkar, Intel fellow and co-director of the Microprocessor Technology Lab (MTL).
www.intel.com/pressroom/kits/ bios/sborkar.htm

Tera-scale computing will be the ful-fillment of Platform 2015, Intel's long-range vision for the collective evolution of computational technologies, interfaces and infrastructures, and the architectural innovation and core competencies enabling that evolution.

Terabytes of Data Require Teraflops of Cycles

"Tera-scale" refers to the terabytes of data to be handled by platforms with teraflops of computing power — a thousand times greater than the today's "giga-scale." For example, a full-body medical scan contains terabytes of information. In the future, an advanced digital health application might assess a patient's health by interpreting huge volumes of data in a scan.

No less complex, you might apply these teraflops of computing power to sifting through your digital photos to identify only snapshots taken at a birthday party. Advanced searching and machine vision applications might be able to find a specific person in a variety of settings and lighting conditions across stills and video.

Tera-scale research brings new processing capabilities to the tasks of crunching and interpreting the world's growing mountain of data. Business and government data stores are becoming larger and more complex. Even at home, consumers are becoming digital pack rats, easily squirreling away hundreds of hours of video, thousands of documents, and tens of thousands of digital photos.

With tera-scale computing, you could create studio quality, photo-realistic 3D graphics on your laptop in real time. You could manage your personal media by automatically tagging and sorting snapshots and home videos. You could use advanced algorithms to fundamentally improve the quality of photos captured on older, low-resolution cameras.

To manage these tasks, tera-scale ar-

chitectures must be optimized for processing massive amounts of data in parallel (see graphic).

Intel is bringing teraflop performance to servers, desktops, and other future platforms. Today, tera-scale computing is limited mostly to supercomputers. However, these supercomputers are composed of separate processors interconnected with cables, which have limited bandwidth and relatively high latency— that is, it takes a relatively long time for a signal to travel from one point to another. Intel is working on connecting many cores on a single chip
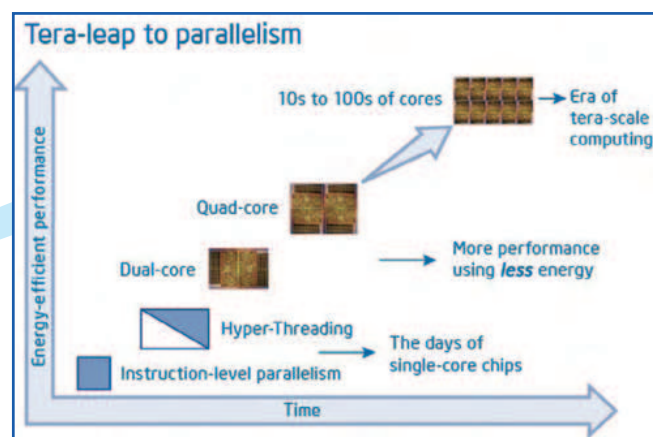
using much shorter, low-power, and low-latency interconnects.

Changes Range from Chip to Platform

Intel's Tera-scale Computing Research Program is laying track for the future, looking out about a decade, when tens or hundreds of cores will work together in one system. To be ready, Intel must make advances in:

• Microprocessor research, by developing scalable multi–core architectures and new types of individual cores and interconnects – continually pushing process technology advancements and energy efficiency.

• Platform research, by optimizing

I/O, communications, memory, and caching for parallel architectures, and increasing virtualization and platform energy efficiency

• Software research, by studying future workloads to direct architectural designs and develop programming tools and techniques to make highly threaded software

Managing Multitudes of Compute Cores

"Management of these core resources will be complex, but entirely transparent to users. They will simply see a much more responsive system with improved power efficiency," says Jerry Bautista, director of technology management for the Microprocessor Technology Labs.

Management techniques for tomorrow's multi-core processors might involve moving processor hot spots around for improved thermal management and shutting down an individual core when its computing power is not required in order to save energy. Cores might also be grouped into separate virtual computers, depending on computing needs.

"Our tera-scale research is directed toward maximizing the advantage of parallel computing for the personal computer. This major inflection point in the history of computing presents challenges across many technical areas. We have research in these areas, from circuit design up through the application software. In MTL, we are strongly focused on applications/us-

age models, interconnect topologies, and microarchitectural innovations. We are making solid progress and are working to establish Intel as a leader in this area," says Joe Schutz, vice president and director of the Intel Microprocessor Technology Lab (MTL).

www.intel.com/pressroom/kits/ bios/jschutz.htm

To be ready, the Tera-scale Computing Program team members identify new and emerging applications that can use the performance. They look for ways to improve existing applications and move them to high levels of execution parallelism.

Optimizing Platforms to Feed Chips the Data

Memory capacity and bandwidth have to match or the cores will "starve" and users will not see the performance benefits. "Tera-scale microprocessors are an important design point for increasing job throughput in future server workloads," says Don Newell, a senior principal engineer from Intel's Systems Technology Lab. "One of the critical issues we have identified is the efficiency and performance of the cache hierarchy. Cache technologies are a key ingredient in alleviating the memory bandwidth pressure that can be caused due to so many cores on a single die."

www.intel.com/technology/ techresearch/people/bios/newell_ d.htm

Creating this balance requires new memory interfaces and packaging techniques. I/O for network and disk storage systems must scale, too.

"Balancing tera-scale levels of I/O

capacity with computation will require new mechanisms, policies, and programming techniques using multiple threads," explains Kevin Kahn, Intel fellow and director of the Communications Technology Lab. Tera-scale platforms will need to support applications of all types, from those with a few high-bandwidth connections to those with tens of thousands of small, low-bandwidth connections. On the platform, network and storage interfaces must move closer to the processor while parallel programming techniques distribute I/O workloads to increase data throughput performance."

www.intel.com/pressroom/kits/ bios/kkahn.htm

Software Design and Education don't Target Parallelism (should this be something like "Changing the way SW is designed")

Today, serial processing is the basis for the majority of programming models. Niche applications, like graphics, parallelize well, but most computer tasks don't. Tera-scale computing researchers have found various attempts to "automatically" parallelize code are not nearly as effective as having the original programmer consider parallelism right from the start.

"Changes in frequency and other advances in processors in the past have transparently increased performance in a typical application without programmers having to do much—they would gain a lot just because they sped up. As we go to multiple cores, software has to make use of the cores. Parallelism just doesn't exist in legacy applications," explains Jim Held, Intel fellow and co-director of

the Tera-scale Computing Program.

www.intel.com/pressroom/kits/ bios/jheld.htm

Programmers for tera-scale computers have to learn methods for creating massively parallel code. We are funding specific research programs as well as symposia to identify critical academic research gaps in the larger community," says Bautista.

Intensive research and collaborative work under way in CTG will help vendors move to tera-scale platforms by:
• Developing hardware and software tools and techniques to help parallelize existing applications to take advantage of multiple cores, in some cases without needing to be rewritten
• Creating hardware or software techniques that help programmers easily take advantage of multiple cores as they write new applications

Transactional memory (TM) serves as an example. In general, the hardware- or software-based TM technology takes care of memory management problems, and programmers have to worry only about the logic of their program.

External Collaboration Drives Thinking into the Future

Intel researchers engage the thought leaders among providers of applications, systems, and services. Industry representation is diverse, including financial services, oil and gas, entertainment, healthcare and information technology. The common goal is to identify the emerging workloads and key requirements for tera-scale computing.

Workload analysis

Discover usages → Prototype a "workload" → Parallelize workload → Test and analyze workload → Architectural enhancements and benchmarking
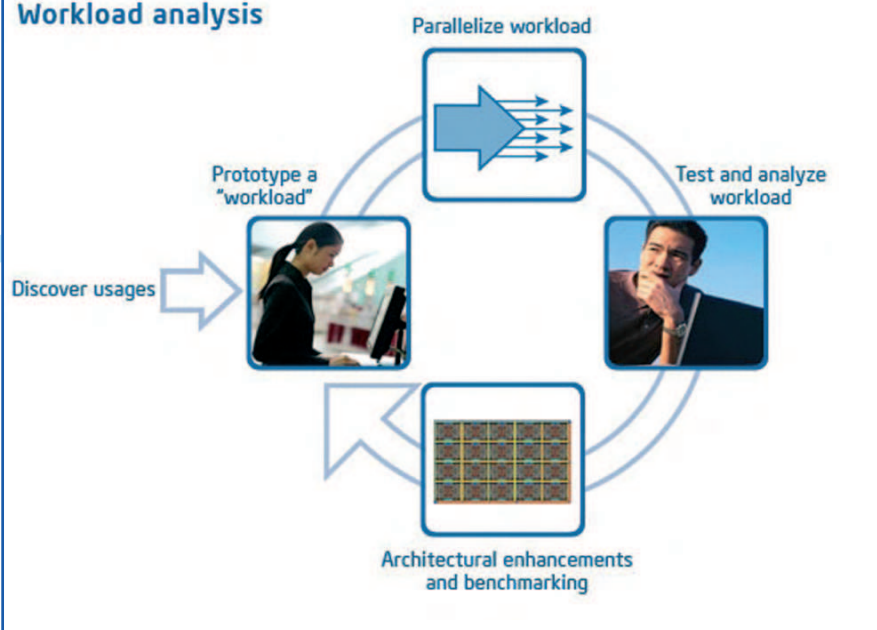
Intel continually refines its view of critical tera-scale architectural features based upon these discussions. Understanding future workloads enables Intel to deliver optimal tera-scale silicon and platforms. Intel also motivates the software industry to create future applications that can take full advantage of advanced technologies.

"Looking forward, Intel will really focus on end-user usage models and business value to take advantage of this energy-efficient tera-scale computing. Multiple cores will be interpreted into platforms that will add value for our customers," says Steve Pawlowski, Intel senior fellow and general manager of Digital Enter-

prise Group (DEG) Architecture and Planning, and DEG chief technology officer.
www.intel.com/pressroom/kits/bios/pawlowski.htm

Summary

The Intel® Tera-scale Computing Research Program looks to the future of computing—about a decade from now—when tens or hundreds of processor cores will work together in one system. Intel has over 80 R&D projects worldwide in the tera-scale area, where researchers are addressing the hardware and software challenges of building and programming systems with dozens of energy-efficient cores.

The program's name, Tera-scale Computing, refers to the terabytes of data to be handled by platforms with teraflops of computing power. Intel's discoveries will help shape the future of the Intel microarchitectures underlying platforms and the applications that run on them.

More Info

You can learn much more about Intel's research in tera-scale computing at the Intel Web site:

Tera-scale Computing: www.intel.com/technology/techresearch/terascale/
Energy-Efficient Performance: www.intel.com/technology/eep/
Platform 2015: www.intel.com/technology/architecture/platform2015/index.htm
Platform 2015 Update: Technologies that are Aware: www.intel.com/technology/techresearch/idf/fall-2005-keynote.htm **Intel® Platforms:** www.intel.com/platforms/
Intel Multi-Core: www.intel.com/multi-core/
The Evolution of Parallel Computing: www.intel.com/platforms/parallel.htm

To learn more about Intel's labs and research projects, visit the Technology & Research area of the Intel Web site. **www.intel.com/technology/**
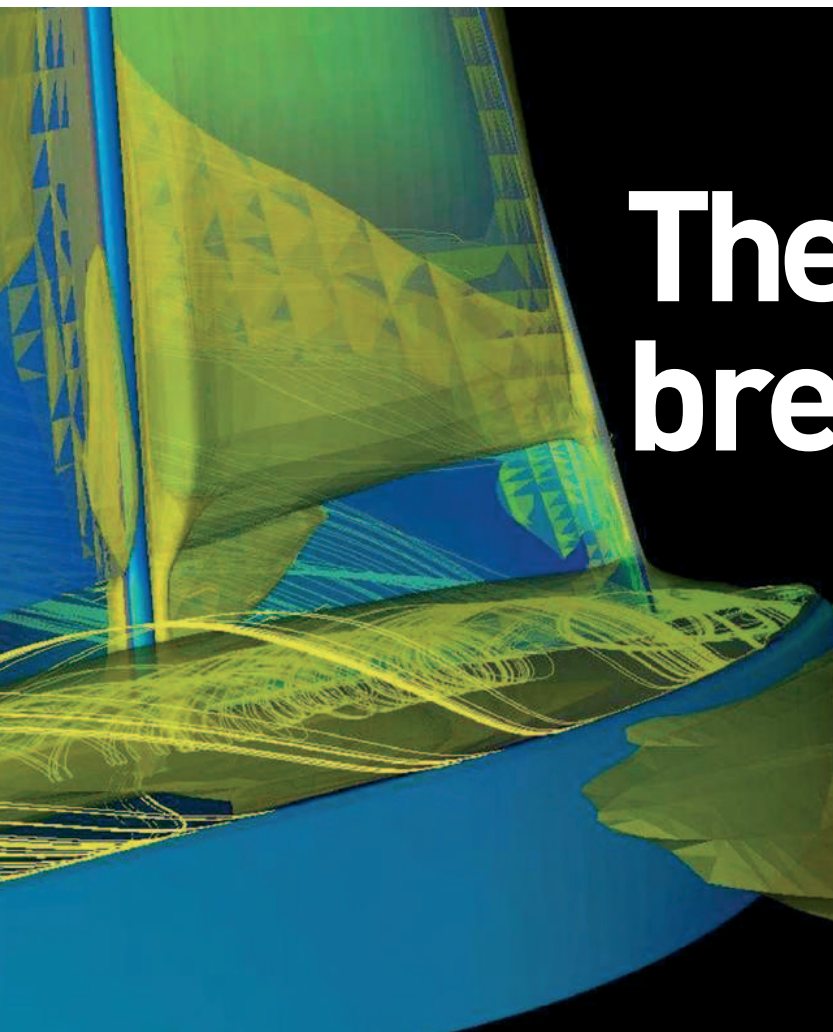
The breakthrough

High performance simulation has become a strategic challenge for both scientists and industrialists. But as Paul Caseau and Christian Saguez underlined in a recent report for the Academy of Technology, France could have a lot of ground to make up in this domain.

NUMERICAL SIMULATION has invaded our everyday life. This picture of Luna Rossa, a racing yacht in the 2003 America's Cup, was done with the Ensight software. © COURTESY PRADA CHALLENGE 2003

Paul Caseau,
former Director
of Research at EDF.

Christian Saguez,
Professor at Ecole
Centrale de Paris
and President
of Ter@tec.
Both are Académie
des Technologies
members.

[1] Académie
des technologies
(2005) : Enquête
sur les frontières
de la simulation
numérique,
Diagnostics
et propositions.
www.academie
-technologies
.fr/pdf/
rapport090505
Simulation.pdf

Mechanics and physics, chemistry and biology – all the major sectors of industry and research use simulation tools. This widespread presence has been consolidated over the last 10 years and is today recognised unanimously. France has reached a more than respectable position in this field especially because of its great schooling in applied mathematics. However, the situation is cause for concern in terms of high performance computing as it is essential to ensure the competitiveness of our industry and to rise to the major scientific challenges. Last May, the Académie des Technologies published a report which analysed France's position with respect to these challenges and gave several recommendations [1].

In industry the strategic role of numerical simulation is clear. It is no longer only used in the analysis and design phases, but all throughout the life cycle of products or systems. The stakes are high for business. High performance numerical simulation increases the capacity for innovation, because it means that new breakthrough technologies can be introduced. Cases can be studied which would be too difficult, too costly or even impossible by experiment. It makes calibration studies easier and considerably reduces the time and the cost of product design. To be in step with the market is also vital for a business. In the automobile industry, the aim is to reduce the design time by more than one half. But downstream simulation also ensures an increased guarantee of quality and monitoring of products. To such an extent that in many industrial sectors computing projects are now a prerequisite for a buyer to accept a product or a system. These tools have become essential to guarantee the development, the competitiveness and the sustainability of business. The way simulation has penetrated into business is in certain cases solely down to the Technical Directors when there is a critical problem to solve or for the smooth day-to-day running of the design/production chain.

It is another story when it is used to design a new product integrating a major technological leap. Then it is crucial that high performance simulation is considered at the level of the Chief Executive or the General Management. In successful cases the business will be able to gain a decisive advantage against its competitors. Having simulation tools for design (test, validation, and production) is a major asset. Simulation can contribute to the overall management of links between suppliers and subcontractors creating a complete virtual system of the business and its supply chain. This too is at the strategic level of business management.

Simulation will spread widely to the service industry in the years to

# of high performance computing

come. It is bound to play a pivotal forecasting role in sectors such as banking, insurance (risk evaluation), telecommunications and in the study and forecast of natural disasters, for example.

High performance simulation is not least a key element in research projects related to the major scientific challenges such as meteorology, climate change, new materials and nanotechnology.

Such high stakes merit a strong commitment both from industry bosses and from the government. It is an absolute necessity if France wants to stay in the leading pack, faced with the more rapid progress of other developed nations today. The French situation is worrying in terms of processing power, software and large applied projects. This can be learned from the analysis of statistics, although not individually revealing, all lead to the same conclusion.

The first criterion of analysis has already been commented on by many observers of France's scientific activity: it is France's position in the Supercomputer Top 500.

France has a significant setback in terms of available computing power; in November 2005, it was alongside Italy and Spain with less than 2% of the total of the Top 500, far behind Germany and the United Kingdom. The number of systems fell from 11 to 8 in 2005 alone representing 1.6% of the total. The only French machine amongst the first 5

is the 60 teraflop being set up at the CEA. Therefore France's ranking is worsening in a Europe that is far behind the United States. More than 95% of the today's machines are built in the USA, less than 0.2% in Europe.

A similar situation is found in two other key areas: the problems tackled and large scale projects. In the USA, the Department of Energy has decided through its ASCI (Accelerated Strategic Computing Initiative) programme to promote its modelling and simulation power by setting up five large centres of excellence. They are: rapid dynamics of materials at the California Institute of Technology; thermonuclear flashes in astrophysics at the University of Chicago; simulation of advanced rocket engines at the University of Urbana Champaign; turbulence at the University of Stanford; accidental fire and explosions at the University of Utah. This ten-year programme will without doubt have strong repercussions for industry considering its quality and the level of funding.

Similarly, Japan has set up several programmes in extreme simulation. The best known, Earth Simulator, has remained at Number 1 of the Top 500 for several years. There is also the Numerical Simulator III Project from the Institute of Space Technology and Aeronautics in Mitaka near Tokyo and the Frontier Simulation Software for Industrial Science Project at the University of Tokyo.

So where is France? We are successful in aerodynamics, in meteorology and climate and, of course, the simulation programme of the CEA-DAM (see page 56). But we can agree that it is rare to see a strategic vision, based on ambitious programmes of extreme simulation, which should be essential.

Practically what problems need solving and what techniques need implementing? The first focus is to consider the geometric complexity of systems, whether moving up to 3D, simulating entire objects, or ➔

**MUSHROOMS?
No, it's turbulence
developing at the
interface of two
media with different properties.**
©DAM/CEA

directly including – without manipulation or simplification – the real life data coming from the major computer-assisted design systems. Another challenge is the processing of multiphysics and multiscale problems. The aim is first to integrate all of the problems coming from different disciplines in the same environment for a given system and secondly to link the analysis of materials at the nano or microscopic scale to the overall properties of the objects (assembly or crash tests). A third focus for progress is improving the probability and stochastic models to be able to make statistical evaluations, a particularly important point in risk evaluation, for example.

Finally, within businesses we must also watch over the overall optimisation of the design/production process. The ultimate purpose of simulation tools is to design an innovative product with the best features and produce it for the best cost. These objectives can only be achieved if the operational system is supported by the technical progress

made possible with simulation. These different points require considerable scientific and technological work in large scale projects including improving hardware and software components. For machines, it is clear that we keep our expertise in the design of architectures for high performance computing so that around 2010 we will have petaflop machines *****. This work involves the study of parallel architectures of large cluster systems, high throughput networks,

Models which will be in common use in 2010 are those being tested today

interconnection networks, grid architectures and all the associated software (exploitation and administration systems, tools and development languages, etc.).

For software, three points merit close attention. Algorithmic studies

must first evolve to get the best out of the new architectures without being constrained by them. Next, the methods and software development tools: quality software needs to be designed in a collaborative manner by many people from different disciplines. We must finally take into account the impact of freeware. Publicising results mainly comes in the form of software and this should really be thought of from the beginning, so the availability of freeware is clearly significant. It offers new opportunities to help France find its rightful place, as proved by the success of SCILAB, the free software for scientific calculation developed by Institut National de Recherche en Informatique (Inria). A major initiative in France in this domain would be desirable.

Models which will be in common use in 2010 are those being tested today. It is important to immediately support a global research effort, led in a coherent way by the public and industrial sectors, to ensure our country is in a reasonable position in ten years. To maintain a good position requires state-of-the-art work at the frontier of knowledge both in model complexity and processing capacity. The analysis done by the Académie des techniques led us to formulate several recommendations in terms of the organisation and general policy. These recommendations are made in the ambition of having machines near to petaflop speeds by 2010.

The first measure is to organise, at the national, European or world level, communities built around the major technological and scientific challenges like the one emerging now on climate. The communities should whenever possible create willing collaborations between research on one hand and industry and services on the other. They should join in the growth of large scale projects with machines and operational software as the output.

*A petaflop: one million billion operations per second.

IN THIS VIRTUAL BEATING HEART, we see the turbulent ebb and flow of blood.

©ENSIGHT IMAGE COURTESY OF LINKÖPING UNIVERSITY
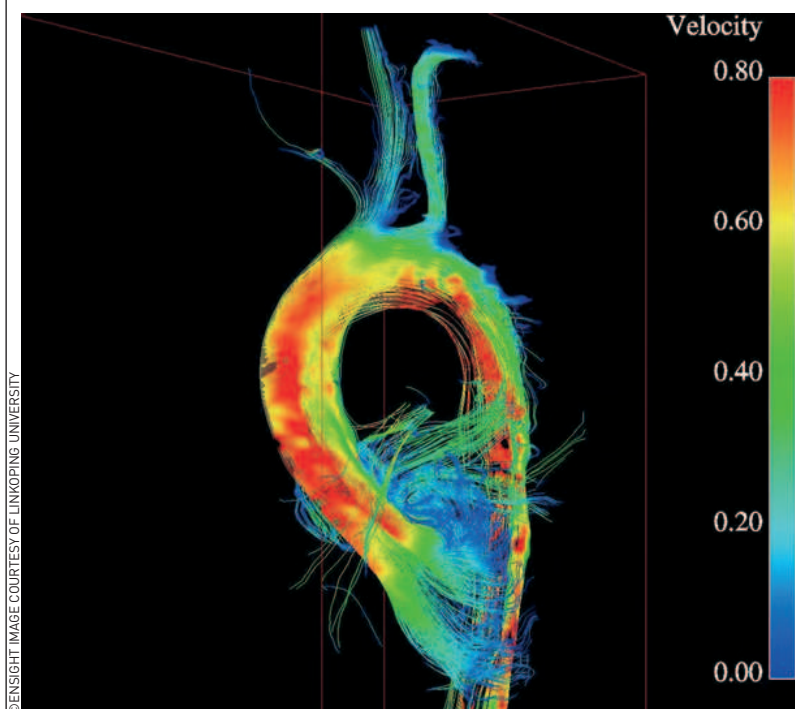
We should keep in mind that high performance simulation is by its very nature multi-disciplinary and problem-oriented. This is particularly true in disciplines linking simulated studies (physics, chemistry and biology) to computing (hardware and software) and to algorithmic techniques. It is essential to set up specific simulation projects so very diverse teams can work together towards a common goal. The national and European political powers should promote these kinds of projects and not only be interested in simulation through applications in computing projects.

Apart from a focus on multidisciplinary applications, the project should consider from the outset all the aspects of distribution, maintenance and development of software. This is essential if we want to make numerical simulation more attractive and do away with the bad image it suffers from, especially among young people. It would also be useful if the academic sector recognised the scientific contribution made by people developing software and if industry recognised the expertise and technical knowl-
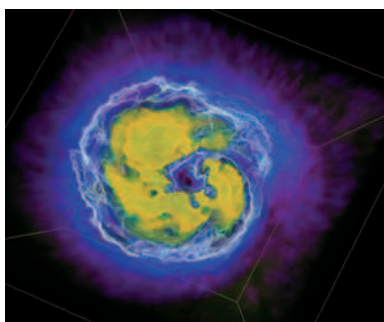
THE DIFFERENT STATES OF WATER IN A CYCLONE. Visualised with the Ensight program: in the form of ice crystals (magenta), snow (blue) and rain (yellow). ©ENSIGHT IMAGE/COURTESY OF LANL

edge of engineers who do this work. Attracting young graduates is essential and can be done by initiatives to communicate successes in the field and possible careers. We also need to develop training specific to the

SIMULATED EXPLOSION OF A SUPERNOVA. Matter is first attracted to the core of the star which collapses. Then at the moment of explosion it is expelled in all directions.
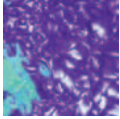
©ENSIGHT IMAGE/COURTESY OF NC STATE UNIVERSITY
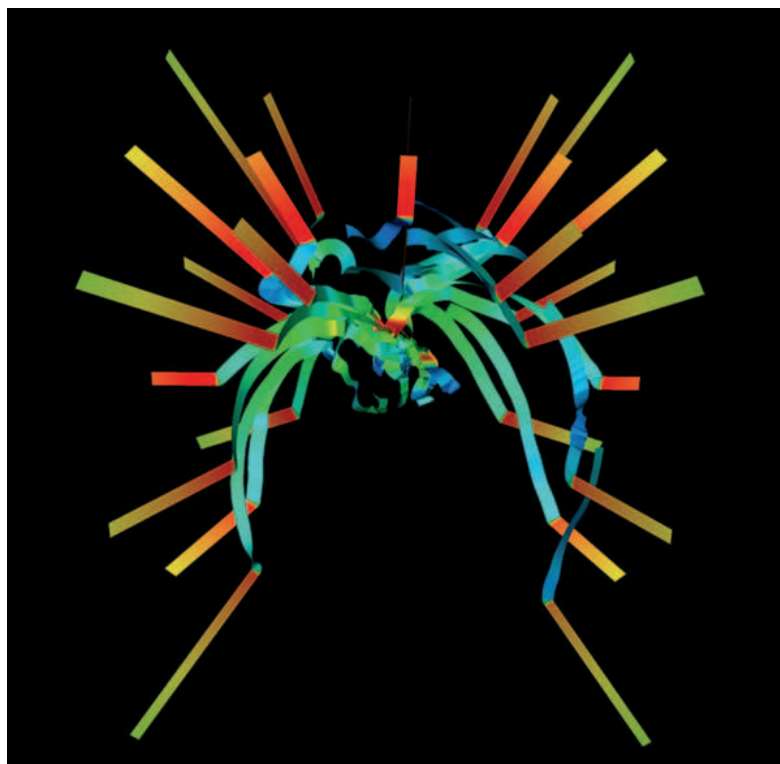
overall simulation approach and encourage the associated careers. Resulting from a voluntary collaboration between industry and research the Ter@tec initiative fulfils these goals perfectly (read page 32). A technopole dedicated to high performance computing, it unites the nation's principal players in research, industry and services around large collaborative projects. Setting up centres like this is a priority. They must be sure of three essential elements: access to very high computing power; creation of national and European projects; and training in numerical simulation techniques, publicising and distributing them. Despite its internationally recognised expertise, the situation is cause for concern, so it is essential that France takes all the actions possible to achieve these goals. Only then will France be able to remain among the top players and guarantee that simulation keeps the position it never never have left. ∎

P.C. and C.S.

TO FIND OUT MORE

▷ Robert Dautray (dir.), *"Le calcul scientifique"* – Rapport du Comité des Applications de l'Académie des Sciences, in *La Vie des Sciences, Comptes-Rendus, série générale, 9,* pp. 63-83, 1992.

▷ Susan L. Graham et Marc Snir (2004) ; "Getting up to speed : the future of supercomputing", National Research Council.

▷ Orap (2006) 1994-2004 : *Promouvoir le calcul haute performance,* www.irisa.fr/orap/Publications/ Bi-orap/livre.pdf

Jean Gonnord

"**Europe is back** in supercomputing"

To make up lost ground Europe should have a more proactive policy in supercomputing centred on a synergy between defence, industry and research.

LA RECHERCHE. A glance at the Top 500 is evidence enough that France and Europe are lagging far behind the United States and Japan in supercomputers. How do you explain this?

JEAN GONNORD. Lagging behind like this is very alarming and is a direct consequence of setbacks in large 'computational projects' at the beginning of the 1990s. The European intensive computing industry collapsed and only a few businesses survived. This was the case for Meiko in Great Britain for example, which after its financial collapse was bought by the Italian firm Finmeccanica. Renamed Quadrics, this company is today producing the 'Rolls Royce' of networks. In France, after a long period in the desert, Bull is coming back to the forefront with the Tera-10 machine.

With the almost non-existent industrial framework and the lack of any real strategy, European countries are using a 'cost base' policy in intensive computing[1]. High power computing (HPC) is considered as a tool used in a few disciplines. Laboratories are investing in HPC using their own research fundings with naturally the aim to get the cheapest machines. This has some odd effects: users practise self-censorship and depend on the American and Japanese makers to define what tomorrow's computing will be like. And this makes Europe fall even farther behind.

By contrast, the computing policies of the United States and Japan, which can be defined as 'strategic opportunity'[1], imply a massive support to the sector's industrial groups.....

J.G. The USA aspire to one thing – world supremacy in this field which they think of as strategic. And they managed to do so: naturally by investing very large budget in the field, but also getting the most out of synergies between defence, industry and research. In real terms, the HPC policy is decided at the level of the President himself who relies on the conclusions of the annual report from the President's Information Technology Advisory Committee (PITAC).

This is then implemented by the Department of Energy (DoE*****), the Department of Defence (DoD) and the major research agencies – the National Science Foundation (NSF) and the Defence Advanced Research Project Agency (DARPA). These agencies fund both civilian and military laboratories, universities and the main computing centres to equip them with really big machines. But, and this is an important point, calls for project proposals are only open to American industry! The Japanese have an almost identical policy, but the main applications are in civilian security.

Can you give us an idea of the American budgets?

J.G. They are considerable. For just the Advanced Scientific Computing Initiative (ASCI) program, since 1995 the DoE has been investing some 100 million dollars per year in its three military laboratories (Lawrence Livermore, Los Alamos and Sandia) just for the sheer power of one machine and 120 million dollars every three years to develop another machine at the technological limits! And that's not all. ASCI is also financing a research and development (R&D) program (Path Forward) aimed at American makers for them to focus on high performance computing (50 million dollars per year) and another, Alliance, to sup-
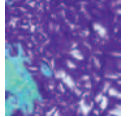
Jean Gonnord

is Head of the Numerical Simulation Project and Computing at CEA/DAM. CEA/DAM

port upstream university research (8 million dollars per year). And this example is just the top of the iceberg. Historically, the major provider of R&D funds in the American computing industry has always been the National Security Agency (NSA) and of course this has not changed especially since September 11th…

Two years ago, China surprised the world by announcing a machine which made 13th place in the Top 500 supercomputers...

J.G. The emergence of this country in supercomputing field is really amazing. The chosen policy is similar to that of the United States, but the stated objectives of the Chinese government are more modest, at least for the moment : to become independent and therefore to control the entire technological chain from manufacturing processors to the final integration of systems. With this in mind, the Minister of Science

THE TERA PROJECT aim is to have 10 petaflops of sustained computing speed by 2017.

and Technology launched an ambitious R&D program planned in five year periods from 1986 onwards having both civilian and military objectives. Nine large computing centres were created. For two years now the installed computing capacity has overtaken that of France! And the rate of progress is impressive.

Even if the first large Chinese supercomputers were bought from the United States, the second generation has been developed and assembled in China using American processors. The next generation will in all likelihood be 100% Chinese. Two projects have been launched to manufacture microprocessors: Godson for scientific computing and ArkII for general public use. Recently, the Chinese have announced that they are going to compete in the petaflop* race… Like the United States, the model for development is based on a defence-industry-research synergy. Europe and France would do well to be inspired by this. Only
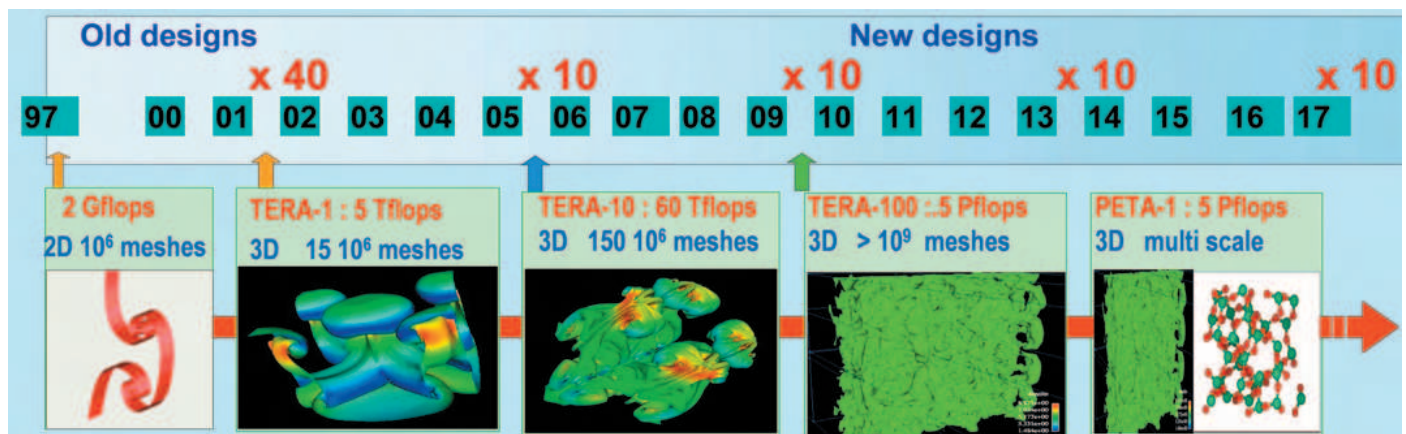
TERA-10, built by Bull, symbolises France's return to the world of high p e r f o r m a n c e computing.
©CEA/DAM

a strategic opportunity policy and putting in place a major European R&D programme would allow us to make up the lost ground.

That's precisely what you've done with the Tera project. When and how did this project start?

J.G. In 1996 after the President of France signed the treaty banning all nuclear testing, the CEA set up the Simulation programme within its military application direction (see the interview with Daniel Verwaerde, page 56). The aim was to guarantee the safety and reliability of weapons for deterrent. The program has two parts : one based on experimentation (with the AIRIX flash radiography machine and the Megajoule Laser being built in Bordeaux) and the other on numerical simulation. Computing is used to reconstruct the different stages in the functioning of a weapon. Around 100 computing engineers and mathematicians have been working on this simulator for almost ten years. They write software, that is millions of lines of code, developed from 'models' established by an equal number of physicists and validated in detail by referring back to past experiments. This colossal task is still ongoing and increasingly sophisticated models are being included in the simulator.

To 'run' the simulator in a reasonable time (several weeks maximum) we would need a much more powerful computer than was available at the time. The required capacity in 2010, when the simulator will be complete, has been estimated as being 100 teraflops of sustained speed*, that's one hundred thousand billion useful operations per second! Our Cray T90 only provided 20 gigaflops* at the time (twenty billion operations per second)! Let me tell you at that time that the prefix 'tera' (for 10^{12}), which stands for 'monster' in Greek, gave the project its name.

Did this pose a particular problem for the vendors?

J.G. The sustained power of 100 teraflops in 2010 was in 1996 well above what they could offer according to Moore's law. Roughly, this law predicts that the power of computers doubles every eighteen months for a fixed cost. This would give us at most 2 to 5 sustained teraflops in 2010 by extrapolating from the power of the very powerful Cray computers we had. Needless to say that such a gain in power which implies a fundamental change of the machine architecture demands considerable scientific and technological jumps. Only the parallelisation of many processors would resolve this problem. But for reasons of costs these processors should be as cheap as possible that is those available on the mass market.

We very soon realised that we would need to push the vendors beyond their limits. To influence their choices we would need to be able to discuss things on an equal footing . In 1997 we brought together a team of top experts on the CEA/DAM-Île-de-France site in Bruyères-le-Châtel. Around fifty engineers were able to interact with the vendors and help in defining an architecture that would fulfill our requirements. A timescale was established to achieve 1 teraflop of sustained speed in 2001 (operation Tera-1), 10 sustained teraflops in 2005 (Tera-10), and 100 sustained teraflops in 2009, all within a strict budget. Now we forecast to bring this capacity up to 10 sustained petaflops in 2017.

In real terms, you launched a call for proposals in 1999 for a machine with a 1 teraflop sustained speed. The specifications were extremely complex with more than 250 criteria and their related penalties! What was the response from the vendors?

J.G. Most of them didn't consider it feasible. Two answered with the best they could offer: IBM and Compaq (in fact Digital which had just been bought by Compaq). The latter won the bid. But with the very fast progress in technology, the machine that was delivered to us at the end of 2001 wasn't exactly what we ordered! However it did allow us to meet our goals and achieve 1.37 sustained teraflops. A really great success…

So what conclusions did you draw from this first experience?

J.G. First of all that it was possible to overtake Moore's law, which vendors normally swear by, to the benefit of all partners. The scientific community also benefited. This machine would never have existed without us or, at least, not so soon. On our side we got the computing resources we needed for nearly five years, tested the simulator during its development, but also learn several lessons for the next machine : Tera-10.
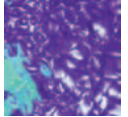
What, for example?

J.G. When we commissioned Tera-1 our main obsession was the power. But once that goal was reached we realised that data management was just as important. I'll just give a few figures: everyday Tera-1 produces more than 3 terabytes of data that's in the range of 1 petabyte per year. Now no machine is safe from breakdown. As we cannot allow the results of a calculation, which might last several weeks on thousands of processors, to be lost we need to save the data very regularly. Unfortunately these operations are very greedy in terms of computing time. We estimate that in one hour the machine should not spend more than five minutes saving data and emptying its memory, which determines the size of the Input/output (I/O)+ system. But that has turned out to be much more complex than expected. We underestimated the I/O capabilities of the machine. Also because of the architecture, the data must be written in parallel with keeping the possibility of reloading the data, not necessarily on the same processors. This poses problems for synchronisation when the machine is functioning at full capacity. Our teams and the vendor needed several months to get around this type of problem.

Wouldn't things have been simpler if you hadn't ordered a blueprint machine?

J.G. Obviously, yes. In computing two years is like an eternity. In 1999 the vendors answered our call for proposals with technologies that only existed on paper. So it took them some time to develop and implement them. The lesson is clear – the time lapse between commissioning and delivery should be as short as possible. Above all, before signing contracts we should insist on technological dem-

THE TERA PROJECT takes up some 2000m². On this plan the rows of processors are shown in the foreground and the data management room in the background.
©CEA/DAM

*DoE
(Department of Energy) leads the United States' nuclear dissuasion programme.

*A **petaflop** is a million billion operations per second (10^{15} operations/s).

*The real power of the computer is expressed in **sustained teraflops**. It is the product of the theoretical power and the yield – the number of operations of a computing code it is able to use. On a parallel machine these yields are of the order of 20% to 25%.

*A **gigaflop:** a billion operations per second (10^{9} operations/s).

*RTD:
Research, Technological Development and Demonstration Activities in the European Union.

"Tera-10 would never have existed without us, at least not so soon"

onstrations that prove the essential elements of the machine work.

From the beginning of the Tera-1 operation you've offered computing time and your expertise to researchers and industry. What were your reasons?

J.G. To give them access to resources which they didn't have and by doing this to make our own project more credible. Numerical simulation is generally validated by one or more experiments. But with the end of nuclear testing we found ourselves in a new situation. Comprehensive experiments were no longer possible so how can we possibly assure the outside world that our project is credible without divulging the details of our methods for obvious security reasons? To demonstrate that we are totally proficient in the technology, that we have the best teams and the most powerful computing resources we began to look further afield. The idea was simple – any major challenges, whatever the subject, solved with our help would consolidate the credibility of our teams and our methods. So we not only offer our computing power, but got our experts to join in projects like genome sequencing or the modelling of the prion movement [2] [fig. 1].

This policy of openness translated into the creation of the CEA Scientific Computation Complex. With its 60 teraflop machine, it is Europe's largest computing centre. How does it work?

J.G. By creating this complex, CEA wanted to get the most out of the synergy in its defence-industry-research programs and the outcomes of the numerical simulation program. Nearly one hundred and fifty CEA/DAM engineers and researchers are now working there. The complex is made up of the Defence Computation Centre with the Tera machine, the CCRT (Centre de Calcul Recherche et Technologie) which is open to all, and finally, a centre for experimentation where our experts work with people from university and industry. The complex is managed by the Ministry of Defence for TERA and by a committee on which each partner is represented in proportion to their investment for the CCRT. Today the CEA has a little over a half of the shares of CCRT. The remainder belong to large corporations (EDF, Snecma, etc) or laboratories like ONERA. With the arrival of Tera-10 the overall capacity of the complex has reached 70 teraflops (60 for defence, 8
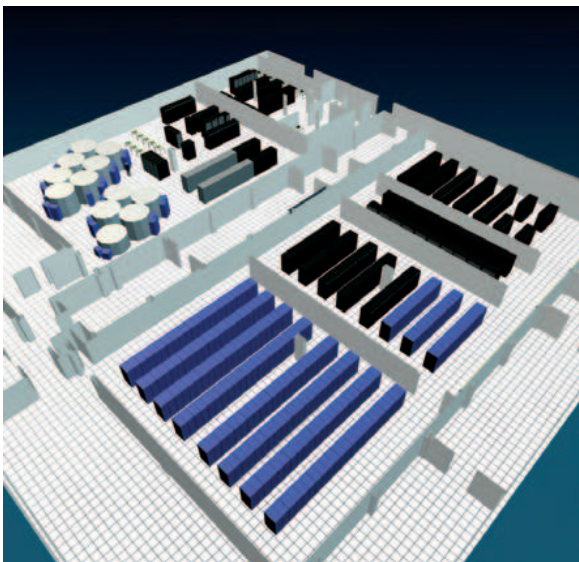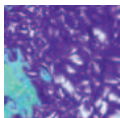
for CCRT and 2 for experimentation) end of 2005. It will pass the 100 teraflop mark beginning of 2007 when the new 40 teraflop CCRT machine will be delivered.

Almost two years ago, a technopole, Ter@tec, was also inaugurated on the DAM-Île-de-France site in Bruyères-le-Châtel...

J.G. The CEA scientific computing complex is in fact at the hub of a much wider operation – Ter@tec (see "Ter@tec : 2 years old and a brilliant future" page 32). The technopole's aim is to unify all parties interested in numerical simulation around the scientific computing complex: researchers, industrials, and technology users and suppliers. And also to share the Defence programme's spin off with the scientific community and industry — and from here to bring Europe back up to the top level in high-performance computing.

Has this collaboration already borne fruit?

J.G. Two associated laboratories have already been created with the University of Versailles and l'École Centrale de Paris, and large industrial groups (Bull, Dassault, EDF, HP, Snecma) are collaborating with us in promoting simulation or in defining the next generation of machines. FAME is one of the first projects to come out of this synergy (read page 30). Uniting Bull, CEA and the University of Versailles, this project supported by the Ministry of Industry has led to the development of a high quality server dedicated to scientific computing. It has been commercialised by Bull under the name of NovaScale since 2002. Fresh from this success, a second project (TeraNova) was undertaken in 2003-2004, this time without state aid, with the University of Versailles and the companies Bull, Dassault and Quadrics. The goal was to create a teraflop machine. The industrial outcomes of these operations are clear. Thanks to them, Bull was able to develop a very general commercial product which can be used in both the management and the scientific markets. They also developed the expertise which places them at the level of the largest corporations. This meant they could answer the Tera-10 call for proposals.

Moving onto the jewel in the crown, the Tera-10 machine. What have been its constraints?

J.G. There again our main goal – a 10 teraflop sustained speed – was far beyond predictions from Moore's law. Like Tera-1, the general machine architecture had to be a SMP cluster type (shared memory multiprocessors). But we had two additional demands. First we wanted a very high sustained speed for a

[1] *Investigation into the frontiers of numerical simulation*, Académie des technologies report by the Simulation working group, May 2005 www.irisa.fe/orap/Publications/AcaTec-rapport_Simulation.pdf
[2] V. Croixmarie *et al., J. of Structural Biology, 150*, 284, 2005.
[3] E. Sartorius and M. Héon *La Politique française dans le domaine du calcul scientifique*, March 2005. www.recherche.gouv.fr/rapport/calcul/2005-017.pdf
[4] http://europa.eu.int/comm./research/future/index_en.cfm

FIG. 1 THE DISPLACEMENT OF AMINO ACIDS IN PRIONS was determined thanks to Tera-1's computational power. Such changes are the cause of neuronal tissue degeneration in spongiform encephalopathy. ©CEA

minimal overall cost including the dissipated power and the floor space. This involved using the first dualcore processors on the market placing us again at the limit of technology. Then we wanted to have large SMP servers for technical reasons (existing codes with a low degree of parallelism and the development of new multiscale models). A tough challenge for the vendors! Finally we wanted fifty to thirty time higher I/O capacities with, of course, the software capable of processing such volumes with maximum reliability. Based on this, our architects wrote a very complete portfolio of specifications with 278 criteria including 53 corresponding to benchmarks defined by our experts. The call for bids was launched in January 2004. Eight makers showed an interest. The call for proposals followed in March.

CEA/DAM

Tera-10 is the most powerful European machine. What's more, for the first time in the history of high-performance computing it was made in Europe. Is that the reason you chose Bull?

J.G. Of course not! Let me just remind you that this machine is one of the crucial elements in a program that must guarantee French weapons for deterrent. Is it imaginable that the CEA/DAM who bears this responsibility could make a choice which might compromise the program for economical reasons or prestige?
Five major makers answered the call for proposals: Bull, Dell, IBM, HP and Linux Networks. Bull made the best proposal. It was able to offer us a homogeneous machine having nodes with 16 dualcore processors and a sustained performance at our Tera benchmark of at least 12.5 teraflops. The Bull machine also had by far the best I/O system and the most reasonable electricity consumption. Finally, BULL proposed an essentially open source solution for system software safeguarding the CEA's freedom of choice in the future. We are obviously very proud that a French business won this challenge. It emphasises the quality of our openness initiative *via*

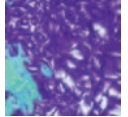Ter@tec and the benefits that the French economy can gain from a defence-industry-research synergy. Finally, the victory of Bull marks Europe's return to the field of high performance computing which is certainly gratifying.

Does this success story show the way for France to get back in the race?

J.G. The conclusions of the report by Emmanuel Sartorius and Michel Héon submitted to the Ministry of Research[3] are very clear. The implementation of a real policy in high powered computing is essential and our methods – grouping resources and the defence industry research synergy – seem to them to be the most appropriate. Times change – and mentalities too! Since the beginning of 2005 we have seen several changes. For example, the National Research Agency (ANR) has included an 'intensive computing' aspect in its program and launched a call for projects last July. Nearly fifty projects were submitted last September and have been evaluated. Another sign is that the System@tic competitiveness initiative, of which Ter@tec is one of the key elements, has just launched a project to develop the new generation of computers (see page 30) with the Ministry of Industry's support. Of course, these efforts do not compare with those undertaken in the United States. But it's a good start.

Will we see a similar initiative at the European level?

J.G. Yes. After a year of effort and persuasion, supercomputing is going to reappear in the budget of the 7th European RTD Framework Programme***** (2007- 2013) [4] which should include an industrial aspect. The beacon project in this initiative will be, if it is accepted, to set up three or four large computing centres in Europe with the mission not of providing computing for a given scientific theme, but to stay permanently in the top three of the Top 500. Undoubtedly, this will mean that major numerical challenges could be solved in the majority of scientific disciplines leading to major technological jumps. The CEA/DAM-Île-de-France scientific computing complex is a natural candidate to host and organise such a structure. But one thing is sure – all of these projects will only make sense if they are based, like in the United States, Japan and now in China, on a solid local industrial network and a proactive policy of States and European Union. ■

"We created Ter@tec to share with the scientific community the outcomes of the defense simulation programme"

Interview by Fabienne Lemarchand

Beyond
50 thousand billion
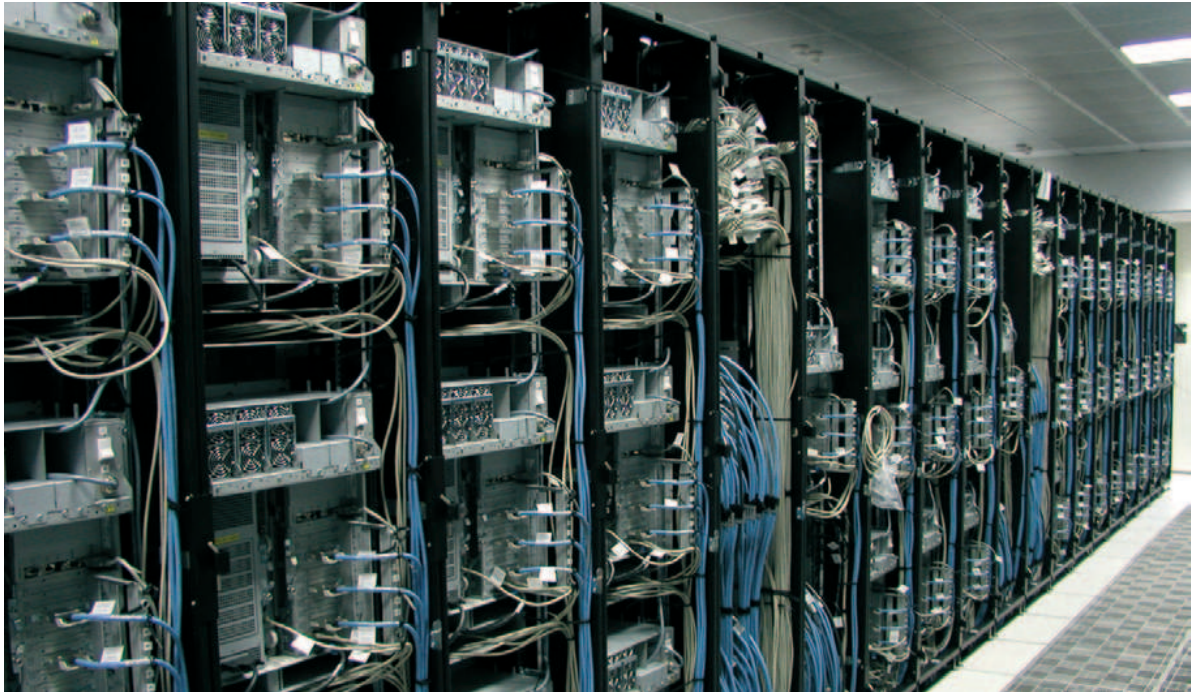
© CEA

Jean Gonnord

Head of the digital and IT simulation project, CEA/DAM.

Pierre Leca,

Head department of Simulation and Information Sciences, CEA-DAM.

François Robin

Deputy head of Simulation and Information Sciences department, CEA-DAM, and operational manager responsible for the TERA-10 project.

Performing more than 50 thousand billion operations a second, TERA-10 will be the most powerful computer in Europe. Profile of this very new supercomputer, designed and build by Bull for the Military Application Directorate (DAM) of the French Atomic Agency (CEA)

Since the French parliament ratified the international Comprehensive Nuclear Test Ban Treaty (CTBT) in 1998, the country's nuclear weapons are no longer tested. Nevertheless, these weapons continue to evolve… To ensure that they remain both effective and secure, computer simulations replaced the physical tests. Predictive models developed by researchers in the Military Applications Directorate (DAM) at the French Atomic Energy Agency (the CEA) are based on non-linear equations, which are solved by iterative calculations, with the correctness of the resulting predictions relying heavily on the size of the mesh. In other words, this kind of modeling requires very powerful computers. The need for processing power, also driven by improvements to the models being used, is growing at breathtaking speed: in essence, by a factor of ten every four years.

This sustained pace has resulted in a new supercomputer being installed every four years. Tera-1, provided by Hewlett-Packard at the end of 2001 gave way at the end of 2005 to Tera-10. This machine, built by Bull, should rank amongst the most powerful supercomputers in the world's Top 500.

Like its predecessor, Tera-10 features an SMP* cluster architecture. Processors (the basic elements that carry our computation) may belong to one of two categories: scalar and vector. The scalar units carry out simple operations, generally on numbers which can be simply defined by their measured value: adding two numbers together, for example.

operations/second !

The vector processors carry out operations on objects that can be defined using a combination of several mathematical entities, such as the addition of two vectors each comprising 500 different elements.

Vector processors are especially well suited to regular calculations, found frequently in computer simulation tasks: during the execution of an operation of this type, a vector processor can function at a speed close to its maximum performance (or 'peak*****'). By contrast, the same operation executed on a scalar processor requires many independent operations (vector element by vector element), carried out at well below peak performance.

Three types of architectures enable processors to be used in parallel: vector supercomputers, 'clusters' of shared memory scalar processors, and PC clusters (the computers we all have in our homes).

Although it is inexpensive, this third kind of architecture is not suitable for environments where many different users carry out calculations that are very 'greedy' in terms of the processor power and memory they require. Vector supercomputers are very expensive but prove more powerful (in scientific computation, they usually achieve more than 50% of peak performance). Clusters of scalar processors, which are at the heart of most mass-market computers, are certainly less powerful (delivering between 5% and 20% of peak performance), but are also significantly less costly. And this means that they offer by far the best price/performance ratio for the CEA/DAM's applications.
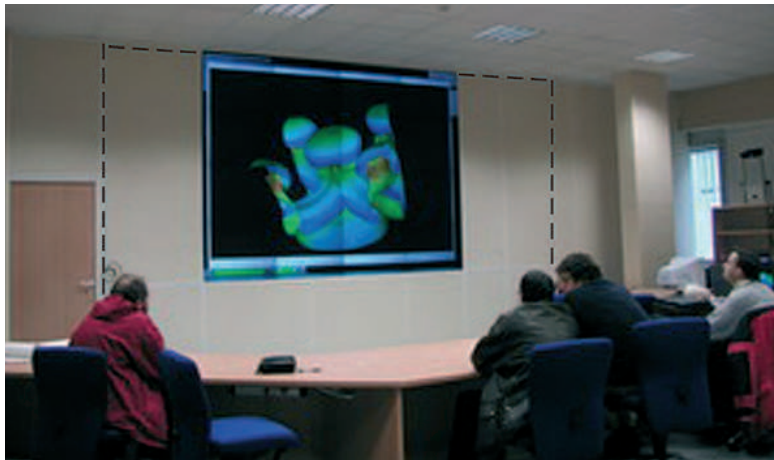
*SMP,
shared memory multiprocessor.
*Peak performance,
optimum performance that a computer can achieve, albeit very briefly.

➔

WITH TERA-10, a 5.5 x 3 meter wall display will replace the current screen (dotted line). © CEA

That is why, beginning with Tera-1, we chose to move away from the traditional scientific computing hardware we had been using up to that point (Cray vector computers) and opted for an SMP cluster architecture.

The basic computing unit consists of a group of sixteen processors sharing the same memory. These units (known as 'nodes') are assembled in parallel, each one carrying out part of the calculation. The nodes are interconnected and share common disk space. Some of them (I/O controllers) are dedicated to controlling the disk space, and this enables the computing nodes to access the disk space to record the results of their calculations or read the initial conditions needed to perform subsequent calculations.

The CEA's experiences gained while using Tera-1 enable the strengths and limitations of these types of supercomputers to be identified.

The first question was to discover whether this type of architecture could actually deliver the necessary computing power. This proved to be the case, on condition that it had access to applications that had been properly configured for parallel processing.

When an application runs on several hundreds or even thousands of processors, it has to effectively be broken down into as many smaller computational units as there are processors. These units are in regular communication with each other, so that the boarder results are cross-checked, which does consume processing time. This means this exchange has to be optimized as far as possible. In addition, the application is fragmented in such a way that every calculation on every section of the program runs for an almost identical length of time. No one part should be effectively putting the brakes on the entire calculation. When applications are written to meet these parallel processing criteria, the performance of SMP clusters really comes into its own. Its overall performance also depends on the supercomputer's loading capabilities.

So the job schedules must be extremely carefully managed to ensure that the maximum number of processors are continually in use. This involves both small and major jobs being scheduled to start at the same time.

The use of an optimized queuing system, taking into account the relative priorities of different calculations, enabled Tera-1 to achieve productivity records beyond 80%. This is relatively unusual for these kinds of systems, where the norm is usually less than or equal to around 70%.

The chosen architecture for the Tera computers nevertheless has some limitations. The biggest drawback is undoubtedly the I/O load due to the shared disk storage of the SMP cluster. The second limitation of these very large-scale systems is the high number of electrical components they contain. Even though, individually, they may be very reliable, none is immune from a failure. With Tera-1, we had around 20 of these a month. Most of these only resulted in a delay of a few hours in completing particular calculations (which often take many hundreds of hours to run). Bigger breakdowns – resulting in actual machine downtime – are thankfully fairly rare: typically less than once a quarter.

Under these conditions, it was essential to establish fault tolerance strategies at every level in the system. So, on Tera-1, critical hardware elements essential to the operation of the whole system (the machine interconnectivity network, part of the shared disk space…) were mirrored, while others were not, for reasons of cost. In case of a breakdown, the software can be called on. For example, if a particular processing node goes down, the calculation it was performing is lost. However, it is then relaunched from a 'restarting point', recorded at regular intervals by the applications running on the computer. Obviously, this risk of breakdown means that major maintenance and surveillance systems have to be in place. Some 15 people are employed in system surveillance at the CEA, with maintenance services being provided by the hardware manufacturer.

For Tera-10, the SMP cluster architecture tested on Tera-1 has been retained. This makes it much easier to transfer software from one platform to the other, however, it still has to be optimized for the new system, which is a heavy undertaking. To respond to current and future developments in production code

*Sustained performance: average performance that a computer achieves on an actual application..

and limit the number of nodes in the cluster, a machine consisting of extremely powerful nodes was chosen (100 gigaflops/node), offering ten times greater levels of sustained performance* than Tera-1. So, although the basic architecture remains the same, the number and nature of the components changes.

The 544 'computing nodes' that make up Tera-10 are each equipped with 16 processors, making a total of some 9,000 processors when the dedicated I/O nodes are taken into account. And if Tera-10's processing performance is ten times better than Tera-1, the disk space for its part has been multiplied by 20, to meet the extreme demands in terms of the quantity of results produced by the simulation software.

When it comes to the 'system' software, which manages the operation of the actual machine, open software was chosen, including Linux for the server operating system and Lustre for the global parallel file system. Bull is providing additional specific elements for both these software components in order to optimize the calculations. Moving from Tera-1 to Tera-10 also involves many adaptations to the IT infrastructure. To ensure that the users get the most from the new power on offer, the supercomputer has to be networked and linked to data storage systems, interfacing and data visualization software, etc. This will enable Tera-10 to be connected to a 5.5 x 3 m image display wall, allowing groups of users to view the results of simulations together.

Currently, Tera-1 is processing between three and five terabytes of data a day. With Tera-10 the volumes will increase in almost direct proportion to the increase in processing power, necessitating the implementation – during 2005 and 2006 – of much more powerful network and data storage equipment. Internal networks have been consolidated and data storage systems increased. There are two levels of storage: data that is more than a year old, which is less likely to be required again in the future, is kept on a magnetic tape system (the most reliable and cost-effective way of storing large volumes of data), stored in tape silos.

The most important data is duplicated and placed into secure zones. For short-term data storage, the magnetic tapes will be replaced by magnetic hard disks, which have come down enormously in price, driven by the home computing market. In this context, four petabytes (four million gigabytes) of disk space will be installed during 2006, for first level data storage.

A large new building (featuring 2,000m² of machine rooms) was specially built for Tera-1. It was designed from the outset to accommodate several generations of simulation hardware (computers and storage), including Tera-100, which is planned for 2009.

Nevertheless, each new version requires some adaptations. So, to move from Tera-1 which consumes around 0.6 MW of power to Tera-10 which requires around three times more than this, it has been necessary to install a new 2MW air cooler, three inverters (which provide battery-powered AC current in the case of electrical power cuts), not forgetting an extension of the fire detection/extinguisher network and the machine room cabling! Once the simulation code used for the design and verification of Tera-10 has been validated, Tera-1 will be switched off and dismantled. This is due to take place in mid 2006. The space freed up as a result will accommodate Tera-100 in 2009.

With Tera-10, the physics experts and designers at CEA/DAM will have access to ten times the simulation capacity provided by Tera-1. But the installation, and then the six-month optimization period required for a system of this scale, remains a challenge. Because here, as with any exceptional piece of equipment, both the technologies and human experience are being pushed to the outer limits. ■

J. G., P. L. and F. R.

TERA-10 consuming three times the power of its predecessor, a new 2MW air cooler has been installed. © CEA

Codename:

Launched last year, in collaboration with Bull, FAME2 is one of the first projects resulting directly from the French Atomic Energy Authority's new open systems policy. Its aim? To develop new servers designed for high-performance computing and multimedia data processing.

Claude Camozzi

is Vice-president Platform Strategy at Bull.

Pierre Leca

is the head of the simulation and information sciences department at the military applications division of the CEA (Atomic Energy Authority), based near Paris.

■System@tic cluster Paris Region: www. polelsc.orgpole_ logiciel_et_ systemes_ complexes.php3

■NovaScale Servers: www. bull.com/ novascale/hpc. html

■Bull White Paper: *The new challenges of High Performance Computing:* http://www.bull. com/download/ whitepapers/ hpc.pdf

The objective of the FAME2 program – launched in 2005 under the sponsorship of Bull and the CEA (the French Atomic Energy Authority) – is to develop a new generation of servers for 2008, specifically designed for high-performance computing (HPC) and multimedia data processing. The acronym FAME (Flexible Architecture for Multiple Environments) denotes an architecture that, as part of its very design, takes into account the key criteria of hardware extensibility (variable numbers of processors, for example) and versatility when it comes to operating system software. The project forms part of the Ter@tec series of initiatives run at the System@tic* competitiveness cluster in the Paris region. It brings together many players from the worlds of research (IRISA, the Universities of Versailles and Évry, the French national telecommunications institute, the École centrale de Paris) and industry: software publishers (ILOG), innovative new companies (RESONATE, CAPS-Entreprise and NewPhenix) and experienced user organizations (Dassault-Aviation and the French petroleum institute).

Why so many partners? Technological evolution developments in HPC are enabling new players to enter the marketplace, but only on the condition that they offer products based on highly competitive technologies and can build up a network of complementary skills. Three characteristics tend to mark these evolutions. The first of these is standardization. Firstly, in software the major code components of processing software

The major code components of processing software have a longer lifespan than information systems

are now tending to have a much longer lifespan than the information systems being used.

Hence, the desire to work with a long-lasting environment when it comes to develop processing code. In practice, this has resulted in increasing use of Open Source software, or software products from ISVs which have become established as the *de facto* industry standard and which can be shared by widespread user and developer communities. This trend towards standardization also applies to hardware. Using COTS (Components On-The-Shelf) is becoming more and more widespread, and is forcing hardware manufacturers to carry out only developments that are strictly essential for them to bring new products rapidly to market.

A second major technological turning point relates to the high volumes of data being crunched in HPC applications. There has been an understanding that the overall productivity of an IT architecture is directly linked to its capacity to manage the information flows and data volumes resulting from simulations or experiments. Alongside these requirements, there is the emergence of active communities focused in the development of specific middleware for this area. And finally, the third major trend is the integration of multimedia characteristics within high-performance simulation applications.

The FAME2 project will take advantage of this evolution and targets the development of a partitioned memory multiprocessor, which by 2008 will be using the new generation of Intel processors. Given that there is a need to keep a limit on processor speed, in order to conserve energy, one of the challenges for the project is to efficiently achieve extremely high levels of

FAME2

parallel processing within the servers. The use of multi-core processors, incorporating several separate processors in one unit, enables several hundreds of processing units to be housed in a single server.

Moving ahead, step by step

This approach inevitably involves some compromises and optimization, both from the hardware architecture point of view (in the hierarchy and consistency of information held in memory) and in terms of the software architecture (managing parallel processing, operating and development environment).

To meet these challenges, the FAME2 project has been organized into four main workstreams. The first is studying the software architecture and will be focusing more specifically on the development of the Linux kernel, the management of independent processing 'threads' and generating code for multi-core processors. The second concerns the hardware architecture, and will focus particularly on NUMA (Non-Uniform Memory Access) architectures which regulate access to the various areas of memory, as well as I/O operations. The third workstream will concentrate on data management: in parallel with the well-known HPC from the energy and aeronautical industries, the management of large-scale XML-format databases, and multimedia dataflows have been chosen as being especially representative of emerging applications. Finally, a fourth workstream will involve taking into account the issues involved in inte-

grating the server with the data center and IT security measures.

The project is scheduled to last for 18 months and a number of milestones have been set including: the key points and basic elements of the architecture, the subsequent demonstration of the feasibility and efficiency of a homogenous internal interconnectivity architecture, and the provision of a demonstrator to emulate a server running the most up-to-date microprocessors with distinctive levels of parallel processing. After this, the next stage will be to create a development and optimization environment that takes account of the new hardware architecture, in order to make typical HPC applications available and to prove the server's effectiveness for these applications. Finally, the last stage will be to offer comprehensive solutions for accessing very large databases. ■

C.C. and P.L.

FROM A PARTNERSHIP BETWEEN BULL, DASSAULT AND THE CEA-DAM, the Teranova platform, a 20 Novascale server cluster, has strenghtened the collaborations that were behind the original project. The Novascale server is the industrial version of the prototypes assessed in the FAME project. ©BULL

*Established in 2005, System@tic Paris Region is a ´competitiveness cluster´ which aims to give the Ile-de-France region just outside Paris a worldwide class capability in complex systems development. It focuses on four main industry sectors: automotive and transport, security/defense, telecommunications and systems design/development

Ter@tec :
2 years old and a brilliant future

The young technopole set up near Arpajon brings together researchers, engineers and industrial partners in one place. All specialise in high performance numerical simulation. It's a unique place in France...
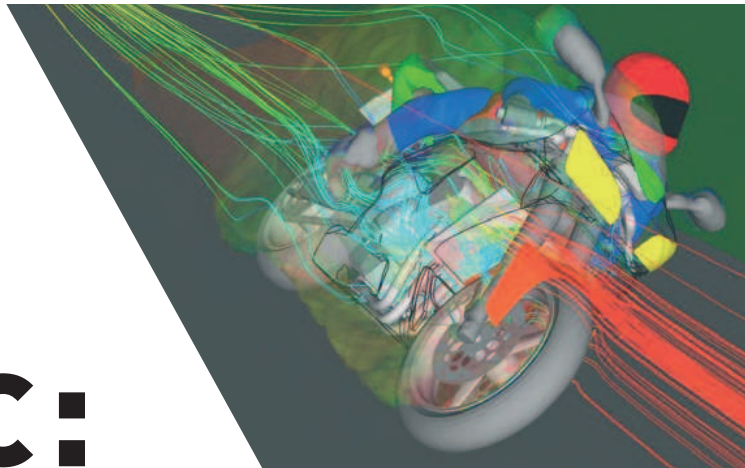
ENSIGHT IMAGE - COURTESY OF ADVANTAGE CFD

Bruyères-le-Châtel in Essonne. Outside the high security centre of the Department of Military Applications of the Commissariat à l'Energie Atomique (DAM/CEA), two two-storeys buildings accommodate sixty or so people coming from the world of research, computing and industry. They all have one thing in common: they are working in the field of intensive computing in the young technopole Ter@tec. *"It's a unique place in France,"* enthuses Christian Saguez, President of Ter@tec.

The idea of Ter@tec originated over the years 2003 and 2004 on the initiative of the Directors of CEA/DAM. Their objective? To open up more to outside influences and to let the scientific and industrial community benefit from their experience and their resources in computing. *"We first created the Centre of Computing, Research and Technology (CCRT). Very soon after, we formed direct relationships with Universities and major industrial groups (EDF, Snecma, Onera, etc.). We put at their disposal a 2 teraflop machine. And little by little, thanks to these collaborations the computing centre has grown,"* recounts Jean Gonnord, Head of the Numerical Simulation and Computing Project at the CEA/DAM. *"Then it*

seemed essential for us to move on from the simple provision of large computational resources," adds Christophe Behar, Director of the CEA/DAM Ile-de-France Centre. *"So we thought of grouping together around the same objective – high performance computing – with industry, whether they are users or from the computing world itself, and with laboratories. This is how the Ter@tec Project was born."*

Since then start-ups (Distène and Numtech) and the constructors (HP and Bull) have joined the technopole. *"Today we have twenty partners*****,"** says Christian Saguez, *"and the demands are increasing from both researchers and industry."*

Because intensive computing is a strategic challenge. It means that research can progress in scientific and technical fields as varied as electronics, aeronautics, the safety of nuclear reactors, climate change, genomics and the properties of material. On the industrial side, intensive computing is synonymous with competitiveness. *"Industry has no time to lose. So it's good if they can turn to simulation for the design and development of competitive new products. As in other domains it has*

replaced experimentation which has become too costly in terms of time and money," explains Christophe Béhar. So where do the constructors stand? *"In the world of computing scientific computation leads the way, it is a niche where the hardware is very sophisticated. And it's this hardware that we will later find in general use in industrial computing,"* explains Christian Saguez.

For the moment Ter@tec offers a total computational power of 7 teraflops: 4 teraflops from CCRT is added to the 2 teraflop Teranova machine from Bull and 1 teraflop from HP.

A technopole looking to the future. *"With the arrival of Tera-10 the DAM decided to increase its computational power up to 50 Teraflops. In addition, with the Arpajon community we are going to buy surrounding land in order to build some 8000 m² of new buildings between now and 2009 destined to house other businesses and to support the increase in CCRT's computing power,"* adds Christophe Béhar. This expansion is even more pertinent as Ter@tec is now one of the key players in System@tic, one of the six world class competitiveness centres. ■

Fabienne Lemarchand

Simulation has replaced experimentation which has become too costly

*****Among Ter@tec partners are:** Communication et Systèmes, Dassault Aviation, Distène, École Centrale de Paris, EDF, ENS Cachan, Institut Français du Pétrole, Institut National des Télécom d'Evry, Snecma Moteurs, Turboméca, Université de Versailles-St-Quentin, and local council organisations of Arpajon, Bruyères-le-Châtel and Ollainville.

High performance computing
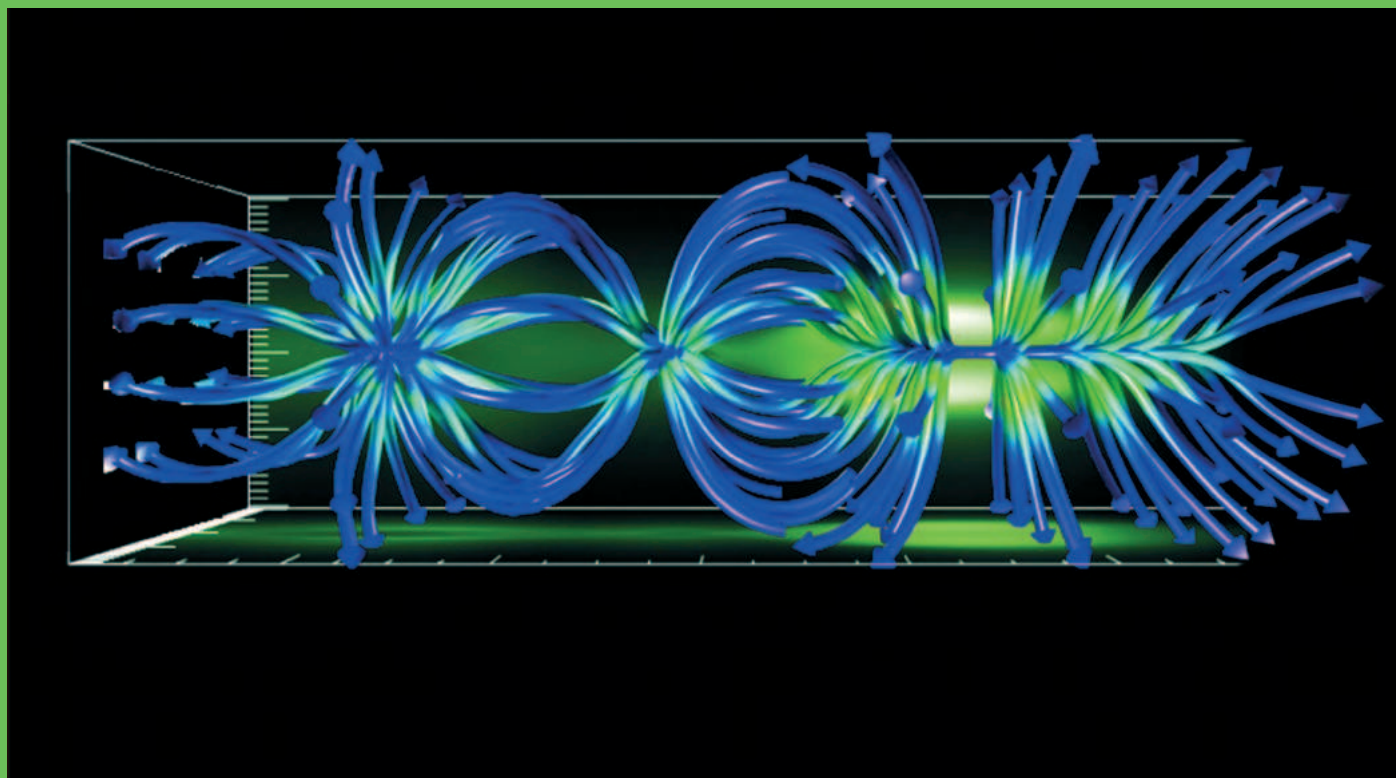
©CHAN JOSHI

Applications
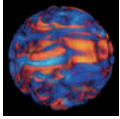
To the origin of the

How stars are born

How stars form is still poorly understood. However modelling it should take a giant step soon with the launch of the European space telescope Herschel planned for 2007. It is generally accepted that an immense and turbulent cloud of gas, several tens of parsec in size, breaks up to form more dense clumps where gravity is sufficiently strong to assemble the matter. Under the effect of gravity the pressure increases, the gas heats up and begins to radiate some of its energy. Cooling then allows the contraction to continue and the gas to carry on heating up until the ignition of the first nuclear reactions of the young star. But many questions still divide the specialists especially this one: where does the turbulence which fragments and regulates the formation of stars in the clouds come from? This turbulence, in fact, determines the mass of the clumps and hence the mass of the star. As it prevents a large proportion of the gas from collapsing it regulates the rate of formation of stars. This turbulence could come from the meeting between cold and condensed zones and other hot and diffuse zones within the clouds. But how can this hypothesis be verified in the absence of observations? Wrapped up in cocoons of gas and dust the forming stars are only visible by sub-millimetric imaging. For the first time, the Herschel space telescope – the largest ever built – will be used to observe molecular clouds at these wavelengths in detail. The assembly of its mirror (3.5 metres in diameter, the largest ever made) has just been completed in September 2005.

⮑ Spatial Telescope Herschel:
http://sci.esa.int/science-e/www
/area/index.cfm?fareaid=16

COSMOLOGY

Galaxies come in all shapes and sizes. Where does this diversity come from? Opinions differ…

Behind the cosmological model of the Big Bang is an extremely simple hypothesis: the universe was at its beginning completely homogeneous. This hypothesis was verified in 1992 with the observations of the Cobe satellite, then again in 2003 with the WMap probe. These instruments measured the diffuse cosmological background, fossil light dating from 380,000 years after the Big Bang, so a portrait of the primordial universe could be painted. It was like a soup of particles sprinkled with lumps whose density was only 0.01% more than the average density. So how did these infinitely small primordial perturbations in density lead to the birth of the galaxies? There are two conflicting models. The first one, hierarchical growth, assumes that the first galaxies were small and grew as and when collisions led them to fuse with their neighbours. In such a process, several billion years would be necessary to form the massive galaxies that surround us. The second model, the dissipating collapse hypothesis, starts from the idea that the growth of the lumps was followed by the collapse of clouds of dust and very heavy gases. This itself was the start of a very bright galaxy full of stars and dust which radiate (this is the dissipation). The process is very rapid; less than a billion years are necessary for the largest galaxies to become materialise. The eventually collisions would only have played an accessory role in their evolution. The model of hierarchical growth was up until recently the favoured model, but it has just been seriously challenged by the discovery that some galaxies are both huge and yet very young.

The two approaches give rise to different models. If we suppose that the interactions between galaxies – fusions and collisions – played a prevalent role in their evolution it becomes necessary to consider the gigantic crashes of the nebula-rich universe into account. This is what Romain Teyssier, an astrophysicist at the CEA is doing. *"It's an important feature of galaxies. Some are subjected to many collisions and others, like the Milky Way, have a much calmer life."* This lifestyle is reflected in their shape: the most unruly ones take the form of elliptical balls, while others have spiralling arms. *"To simulate this change, we take as initial conditions the spread of matter given by the diffuse cosmological background. Then we increment the steps in time to arrive at a current universe. We try to compare the*

At the moment, our models are incapable of reproducing the proportion of spiral and elliptical galaxies observed

When lasers test plasmas

Modelling the cores of giant planets or violent phenomena like the explosion of supernovas or the accretion of matter at the edge of black holes, these are some of the possibilities offered by very high power lasers which test the very unusual properties of plasmas. These ionised fluids are conductors and can only be produced in extreme conditions of temperature and pressure. When a laser beam is focused on a material placed in a vacuum, some of the matter evaporates as hot plasma of several tens of thousands Kelvin. This impulse of gas violently and strongly compresses the non vaporised matter producing a shock wave which causes a rise in pressure and temperature in the material, basically reproducing small pieces of planet in the laboratory. The laser integration line of the prototype of the future megajoule laser set up in the Bordeaux region and essentially destined for military use is widening its remit to include astrophysics experiments at the end of 2005. These experiments may allow the equation of the state of ionised gases to be determined. Although known for perfect gases, this equation, which links the density, the pressure and the internal energy of the gas, is not known for plasmas.

⮑ www.lmj.cea.fr/index.html
⮑ www.luli.polytechnique.fr/
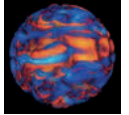
galaxies, *in silico*
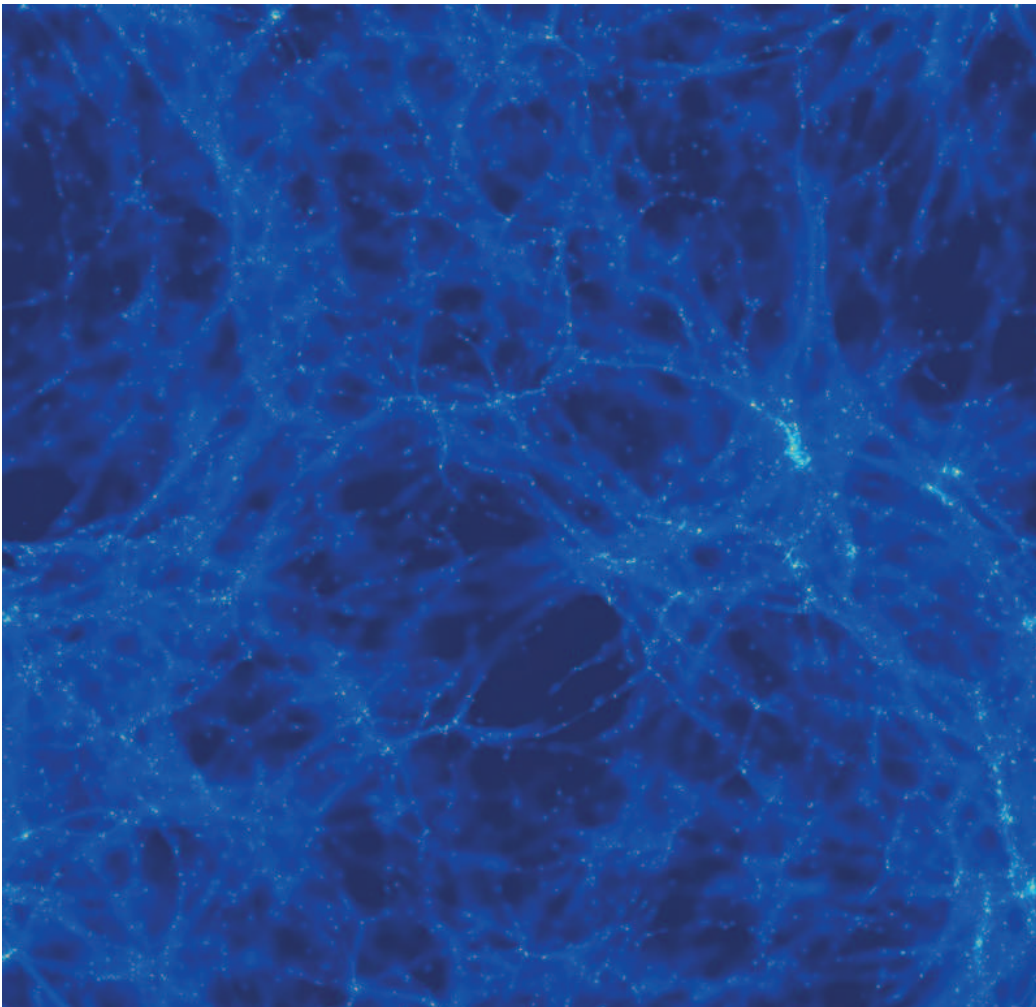
ON THIS HYDRODYNAMIC SIMULATION OF THE DENSITY OF GAS, galaxies are shown in white and space in blue.
© YANN RASERA - ROMAIN TESSIER /CEA

number, the shape and the arrangement of galaxies obtained by computing with those that we observe today." The equations take into account the expansion of the universe, the force of gravity, the mechanics of self-gravitational fluids and the phenomena of atomic physics. But the amount of space to cover poses problems. *"To see galaxies born we need a mesh smaller than the smallest one of them. We will be really satisfied with a million billion points. For the moment, because of insufficient calculating capacity, we only have several billion."* So the only way for these models to see the galaxies is by zooming in on a more limited part of the universe. But many points remain to be revealed, states Romain Teyssier.
"At the moment, our models are incapable of reproducing the proportion of spiral and elliptical galaxies observed. Moreover they give more satellites for our Milky Way than we observe."

Another worry is that simulations based on the hierarchical model predict that objects in the universe all have a lot of mass at the centre. A rule which the dwarf galaxies seem to have flouted. *"Either the model is false or it needs much finer tuning,"* analyses Romain Teyssier. And phenomena such as the explosion of supernovas, jets of matter spewed out from black holes, or dark matter may if taken into account reconcile models and observations.

The model of dissipating collapse, even if less ambitious, agrees better with certain observations. And for good reason: the simulations which follow the collapse of a cloud of gas and dust only use data from telescopes. As Brigitte Rocca, Professor at the University of Paris XI explains, *"Today we have data from deep in the galaxies which goes back to the tender years of the universe, that's looking back more than 12 billion years."* With these observations, calculations based on the hypotheses of simulations describing the galaxies at different epochs can be confirmed of negated. Here there is no limitation due to the computing capacity. The two complementary approaches advance their pawns one after the other. And in 2004, several groups, including Brigitte Rocca's, demonstrated the existence of galaxies that are both massive and very young. Even if they are unusual, these objects challenge the model of hierarchical growth as they don't fit in well with the theory. It will be necessary for simulations to gain in complexity and in speed to succeed in explaining the extraordinary variety of the galactic menagerie and in the end could well depend on two evolutionary processes… ∎

Anne Debroise
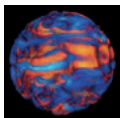
[1] Rocca-Volmerange *et al., Astronomy &Astrophysics, 415,* 931, 2004.
[2] Cimatti *et al., Nature, 430,* 184, 2004.
[3] Glazebrook *et al., Nature,* 430, 181, 2004.

➲ http://aramis.obspm.fr/HORIZON /docs/PROJET-HORIZON.doc
➲ http://dphs10.saclay.cea.fr/Sap /Activites/Science/Cosmologie /Simulation/index.html

**ALLAN SACHA BRUN
is an astrophysicist in the
Department of Astrophysics
at CEA – Saclay.**
©DR

Allan Sacha Brun:
"Modelling turbulence on the surface of the Sun"

The turbulent movements which animate the visible layer of the Sun remain a mystery. For the first time, astrophysicists are attempting to simulate them in three dimensions. Here we interview one of them.
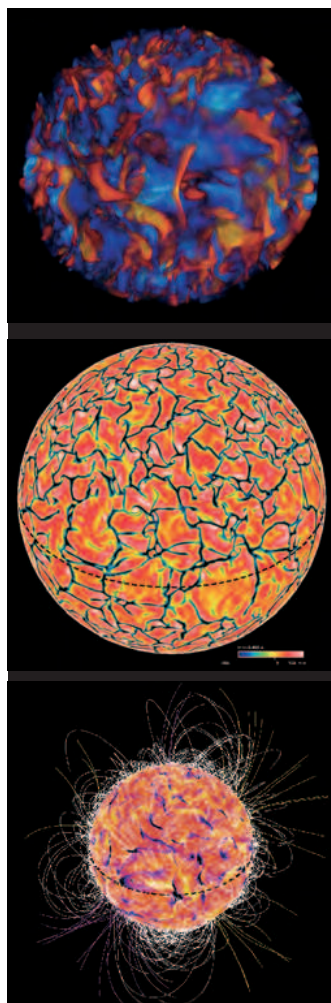

CONVECTION

Why model the Sun?

ALLAN SACHA BRUN: I'm interested in modelling the Sun in short timespans of several minutes or seconds to several years. The goal? To understand all the movements which animate the surface of the Sun – convective movements, large-scale circulations, flares and ejections of solar plasma, etc. As this turbulence and the related magnetic activity can have profound repercussions on the Earth. Here's one example. In 1989, the majority of households in Quebec were deprived of electricity because of a solar eruption; a gust of charged particles disturbed the Earth's magnetic field creating geomagnetic currents which short-circuited the electricity network. Also the Sun is a real-life physics laboratory where complex and extreme dynamic phenomena are produced which are still mysterious…

How would you describe the Sun?

It's a ball of 710,000 kilometres in radius. Its temperature ranges between 15.5 million degrees Kelvin (°K) at the centre to 5,800°K at the surface. It's composed of ionised gas, a plasma. We know its overall structure. A nuclear core which takes up 25% of the radius is where nuclear energy is produced and transferred by radiation in the form of light. Surrounding this is the radiative zone, quite homogeneous and very dense covering 70% of the radius (including the core). Coming

DIGITAL SIMULATIONS show that the turbulent convection (top, rising (red) and falling (blue) movements) generate a magnetic field (middle and bottom).
©A.S. BRUN, CEA / DSM / DAPNIA / SERVICE D'ASTROPHYSIQUE

next is a convective zone which transports the energy in macroscopic movements over the remaining 30% of the radius. It's this very turbulent convection that gives the Sun its granular appearance. But the plasma movements are even more difficult to describe as they are under the influence of the magnetic field and rotation and they spread out over spatial and temporal scales which differ by six orders of magnitude. It is still impossible to model the solar magnetic turbulence in three dimensions completely. But we can get close especially using the new capacity of super computers.

How are these models exactly?

They are detailed descriptions of the non-linear dynamics and the time dependency of turbulent plasmas (conducting liquids) subjected to the coupled effect of the rotation and magnetic field of a sphere. This is the field of magnetohydrodynamics (MHD).

Which data are these simulations compared to?

To observations from Trace satellites, SoHO, for example, or to measurements of the surface and atmospheric magnetism taken by ground instruments like Themis. But we can also listen to the song of the Sun. Acoustic waves are propagated in the interior of the Sun. And the harmonics are characteristic of the composition and internal structure of the star.

What mysteries are these models trying to make clearer?

There are several. We are trying to understand why the tachocline – the transition between the radiative and the convective zones – is so thin; it accounts for only 2% of the Sun's radius. Another question:

why does solar activity (the spots and the global magnetic field) follow cycles of around eleven years but its polarity, cycles of twenty-two years? With current three dimensional models describing the dynamo effect thought to be the source of this magnetic activity, we found cycles of one and a half years… We are also trying to describe the differential rotation of the convective and radiative zones of the Sun. We know that from the surface to the bottom of the convective zone the poles are turning 30% more slowly than at the equator. But this phenomenon is difficult to describe in detail. The three-dimensional models can be used to study the rise and fall of magnetic fields at the surface the cause of groups of solar spots. We think that the tachocline, which is also a zone of shearing, could be the site of organisation of large-scale solar magnetic fields. Finally, we want to describe the interactions between convection, turbulence, rotation and magnetism. How do they trigger flares and generate the fast solar wind? As we see it, there are several questions to answer before we have a dynamic and comprehensive model of the Sun from its centre to its corona. ■

Anne Debroise
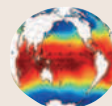
Jean Pailleux :
"The better use of **real data**"

METEOROLOGY

Meteorology was one of the first disciplines to use computers. Today, improving weather forecasting models requires the inclusion of a growing number of statistics. And therefore more powerful supercomputers too.

When were numerical models first used in meteorology?

JEAN PAILLEUX : from 1950. These models predicted the weather from one day to the next. They had just one variable: the wind at 5 kilometres altitude. Plugging this, the parameter at Day Zero, into the only equation (called the Conservation of Absolute Vorticity) gave the value at Day One.
These models covered a large part of the northern hemisphere with a mesh of 400 to 500 kilometres width. The forecast was then refined in step with the improvement in computers. In the middle of the 1960s, a comparable model was used for the first time in the United States in order to determine the daily weather forecast.

Do climatologists and meteorologists use the same models?

Yes. From the 1970s onwards, the models integrated a growing number of classic equations from physics and fluid mechanics to make forecasts several days ahead. The longer term climate forecasts were developed very soon after especially with the growing awareness of global warming in the 1980s. The same models serve to carry out climate predictions for tens or hundreds of years, seasonal

JEAN PAILLEUX is Scientific Assistant to the Research Director at Météo-France (Toulouse). © DR

predictions or predictions of the chemistry of the atmosphere (to follow the outcome of atmospheric pollution) and even the forecast on the scale of a town or a street. But they are models with many options: we can choose the size of the mesh, the place on the globe where it will be the most focused or replace certain parameters with constants.

Calculation time must therefore be a major constraint…

In weather we need a provisional set of data several times per day. We are obliged to make compromises so as not to slow the model down. A numerical forecast at 24 hours should take a maximum of 20 minutes. That's almost one minute for one hour of simulation.
Another constraint: before running the model it is necessary to have collected the initial data from all over the planet and to have processed the information

so it can be used by the model. Building the initial state is what we call 'data assimilation' and is the most costly operation in terms of time. The current models require inputting tens of millions of parameters. Thirty years ago this was a marginal activity but today assimilation predominates.

What progress is expected in weather forecasting?

The most significant progress will come from data assimilation. We try to use more and better satellite observations. The American satellite NOAA18 launched last May will provide new measurements, particularly of telluric radiation. This parameter will indirectly give us information about the temperature and humidity of the air.
The movement from hydrostatic models to other non-hydrostatic models between now and 2008-2010 will also allow us to simulate the atmosphere more precisely. And taking vertical movements into account on the small scale is essential for better predictions of extreme events. For this we need a much bigger computing capacity. In summer 2007, Météo France should have a computer 4 to 8 times more powerful than their current one capable of carrying out 5,000 to 10,000 billion operations per second. ■

Interview by Anne Debroise

➲ The history of the numerical weather prediction by Météo-France: www.meteo.fr/meteonet/decouvr/dossier/previsionmeteo/pre3.htm
➲ The history of computers and weather forecasting: www.metoffice.gov.uk/research/nwp/numerical/computers/history.html
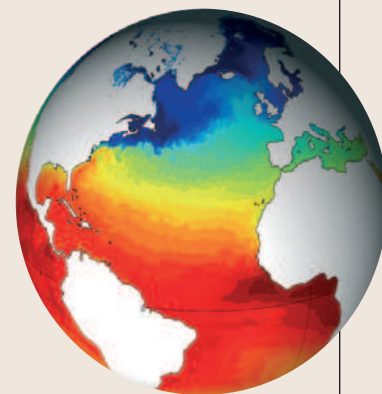
AFTER TEN YEARS IN EXISTENCE, THE MERCATOR PROJECT REACHES MATURITY. Since last autumn, it has been possible to know the state of the oceans of the entire world, at surface and at depth, in real time. And to even have oceanic forecasts fourteen days ahead.
Data are continuously provided by the altimetric satellites Topex/Poseidon and Envisat and by an armada of 1,900 floats which measure the temperature and the salinity between 0 and 2,000 metres depth. They are then fed into a model that has a resolution of one quarter of a degree (that's 28 kilometres at the equator).
This model was developed by the experts from 'Mercator Ocean', a group made up of CNES, CNRS, Ifremer, Institut de Recherche pour le Développement, Météo-France and Service Hydrographique et Océanique de la Marine. Fifty or so teams currently use it whether for research, the Navy, seasonal forecasts, fishing, oil exploration, or monitoring pollution or sea routes. During 2006, the oceanic forecasting model should also take into account the frozen seas at the two poles. It will then be the first to do so.
Another objective is to integrate biology into the model so as to understand the dynamics of marine ecosystems and the related biogeochemical cycles (carbon, oxygen, nitrogen, iron, etc.).

➲ www.mercator-ocean.fr
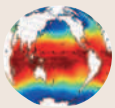
Magnetic jerks

The behaviour of the earth's magnetic field is very strange. It never stops changing over time even to the point of reversing. There are also small disturbances (called magnetic jerks) of unknown origin on the yearly scale. The magnetic field originates 2,900 kilometres under our feet in the outer nucleus. There, liquid iron pulled by the Earth's rotation generates electric currents which in turn create a magnetic field by the dynamo effect. Geophysicists ask: Are the Alfven waves at the Earth's nuclear core the source of the magnetic jerks observed at the surface?

To try to find out, the 'Geodynamo' team in the Laboratoire de Géophysique Interne et Tectonophysique (University of Grenoble), led by Dominique Jault, have undertaken the modelling the nucleus's dynamics. Their approach is novel. It is one of the few experimental initiatives in the world that recreates the working of the terrestrial dynamo. In the experiment, which started in 2005, a copper magnetic sphere is made to turn in another sphere filled with liquid sodium and itself in rotation. Similar flows to those which animate the liquid iron nucleus of the Earth are created. The flow can be described by comparing the first experimental results with the numerical ones obtained from a theoretical model. They are similar to those which would be generated by the rotation of an entire set of parallel cylinders around the rotational axis of the Earth. And the Alfven waves could come from the differential rotation between these cylinders.
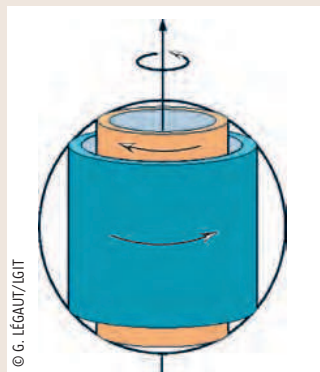

© G. LÉGAUT/LGIT

↪ www.ujf-grenoble.fr/ujf/ fr /recherche/labujf/lgitfrw6.phtml
↪ www.lgit.obs.ujf-grenoble.fr
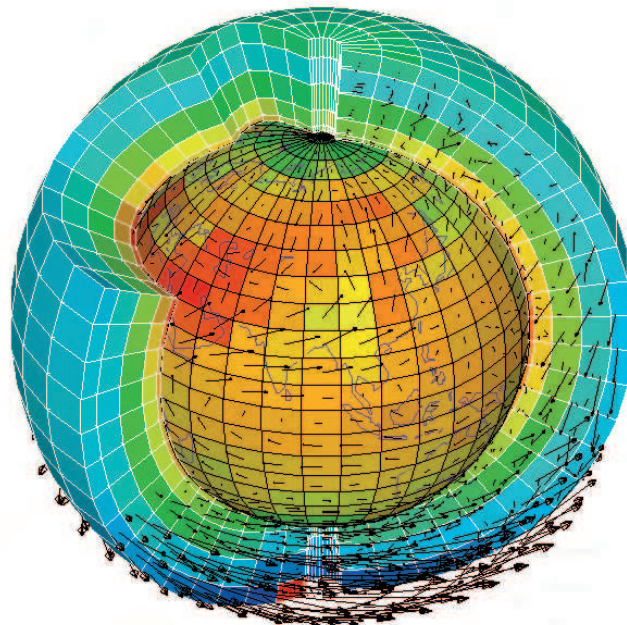
Tomorrow's climate today

CLIMATOLOGY

With supercomputers climatologists can look ahead to future climates. But the current models are not convincing the decision-makers of the reality of global warming as they lack precision. A precision which is largely synonymous with computing power.

"Predicting the climate requires 10,000 to 100,000 times more computational power than is currently available in France," says Jean-Claude André. The meteorologist, who today is leading Cerfacs (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique), is not putting this figure forward blindly. In fact, it comes from an analysis carried out about one year ago by around twenty French specialists in climatology from the Institut Pierre-Simon-Laplace, Météo-France and Cerfacs. The scientists got together to estimate the computing needs of the discipline. The goal? To do away with the approximations and unfocused ideas of long term climate forecasts which today are hampering the struggle against global warming. So the stakes are high. *"Europe must be able to have its own simulations and figures to be able to negotiate international agreements on an equal footing with the United States,"* continues Jean-Claude André.

As global warming is still for many, and particularly for the decision makers, a virtual reality. Symbolic figures brandished by the IPCC (Intergovernmental Panel on Climate Change) are hardly striking, it's true: an increase in global temperatures of 1.4°C to 5.8°C between now and the end of the 21st Century… Why such uncertainty? And what do these values mean at the regional level? Will the Camargue disappear under water? Is Florida's agricultural

Will the Camargue disappear under water? Is Florida's agricultural land going to dry out?

land going to dry out? Are the icecaps melting faster and faster? So many questions in need of clear answers. Part of the uncertainty about the extent of warming is difficult to assess: it is due to our lack of understanding of the way in which different regions of the world will develop and the steps more or less environmentally friendly they will take. But the second part of this uncertainty comes for sure from the models' lack of detail. The climatologists brought together by Jean-Claude André identified ten scientific obstacles to overcome (see opposite). But for this there will need to be at least a 10,000 times increase in the power of European supercomputers! In total, France today has an effective computing power of 0.2 teraflop to run climatic models and Europe has five times more (1 teraflop). This is little in comparison to the United States and Japan (with the Earth Simulator) each having 4 to 5 teraflops. If they continue on their trajectory these two countries will overcome the obstacles around 2020 ten to twenty years ahead of the Europeans. The estimated cost of investment for Europe to stay in the game: 200 to 250 million euros per year over three years. ■

Anne Debroise

Researchers cut the Earth into small cubes (a mesh) as it is impossible to do calculations everywhere. For each of them, they define the temperature, the humidity, the pressure, the wind, etc. By making these cubes communicate with each other, the models build up simplified climates. ©FAIRHEAD/LMD/CNRS
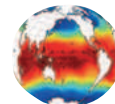
Ten obstacles to overcome

TO DETERMINE THE CLIMATE'S TRUE SENSITIVITY TO THE INCREASE IN CONCENTRATION OF GREENHOUSE GASES: this essentially involves improved modelling of the water cycle in the atmosphere and defining the role of clouds. Clouds contribute to warming as water vapour is a powerful greenhouse gas, but they also block the Sun's rays. Modelling clouds requires a small mesh (of the order of 1 kilometre as against several tens of kilometres used today) and an atmosphere in three dimensions. The current models 'see' the atmosphere as a superposition of two-dimensional layers that do not interact very much.

TO BETTER PREDICT TROPICAL PHENOMENA (El Niño, monsoons, etc.). This involves modelling the heat exchanges between the surface of the ocean and the atmosphere with a vertical resolution of the order of ten centimetres, a hundred times finer than is achieved today.

TO QUANTIFY THE RISK OF CLIMATIC SURPRISES, for example, the risk of the Gulf Stream slowing down. This current runs along the northwest face of Europe and via ocean-atmosphere exchanges makes the climate of this zone considerably milder.

TO SIMULATE THE DEVELOPMENT OF DANGEROUS PRECIPITATION, tempests and cyclones. Are these irregular phenomena becoming more common and stronger? A question that current models struggle to answer. These phenomena, in fact, depend on interactions between the ocean and the atmosphere. And to distinguish the ocean's changes, by their nature slow, we must be able to do simulations over long periods. But to describe local meteorological events also requires a fine resolution, so there is a large gap which today is impossible to narrow.

TO MODEL THE CHANGES IN CARBON SINKS: today we assume that the capacity of the ocean and the terrestrial biosphere to absorb carbon dioxide will remain constant. A hypothesis which is highly speculative...

TO UNDERSTAND THE LINK BETWEEN ATMOSPHERIC OZONE AND CLIMATE CHANGE. This requires better modelling of the vertical exchanges between the different layers of the upper and middle atmosphere.

TO ADD THE BIOGEOCHEMICAL CYCLES INTO THE MODELS, for example, the carbon dioxide and methane cycles in various compartments (soil, vegetation, ocean, etc.). It will be particularly necessary to include the life cycles of phytoplankton, essential species in the assimilation of carbon by the ocean.

TO MODEL THE INFLUENCE OF URBANISATION AND CHANGES IN LAND USE ON THE CLIMATE. Here again the models assume that land use will remain the same. A hypothesis incompatible with demographic projections which predict a one third increase in the world's population between now and 2050.

TO ASSESS THE QUALITY OF THE MODELS by comparing them to the body of available data, in particular, the large amount of data from space observation programmes.

TO CARRY OUT ENOUGH SIMULATIONS to be able to determine the likelihood of imagined scenarios.

A.D.
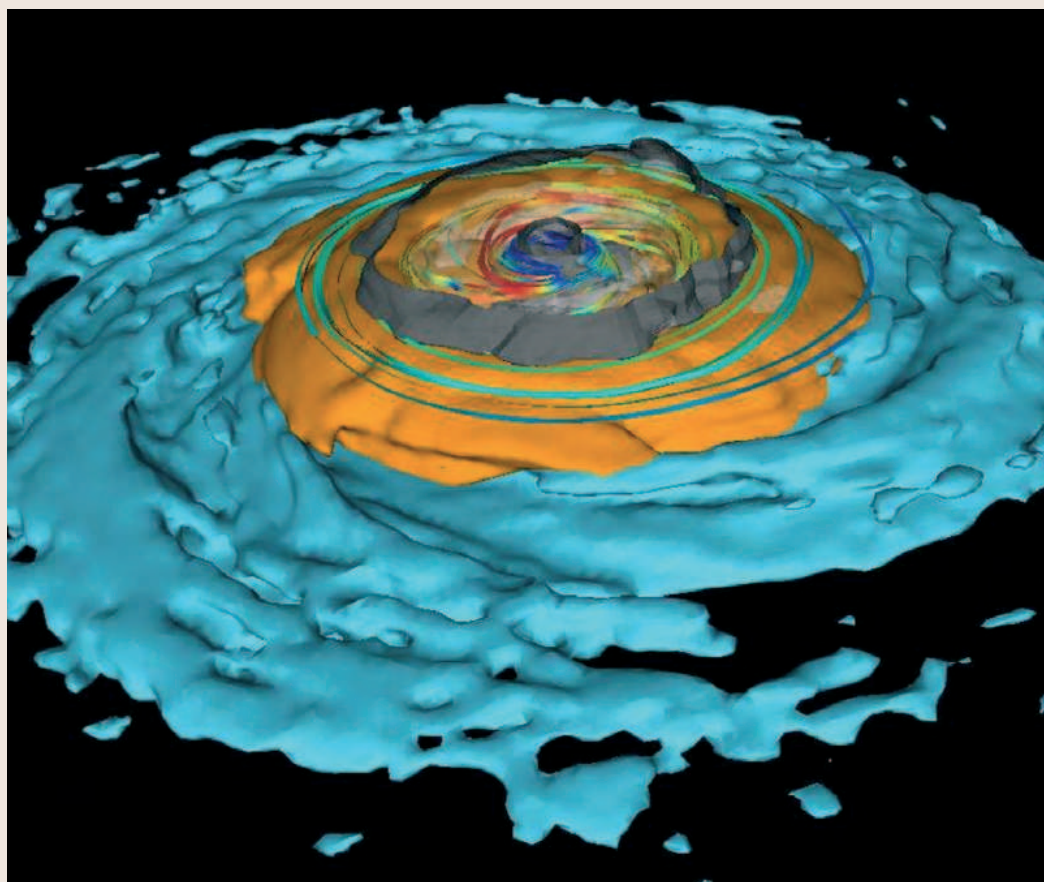
ARE CYCLONES BECOMING MORE COMMON and more intense with global warming? The resolution of current climatic models is insufficient to answer this. © COURTESY LOS ALAMOS NATIONAL LABORATORY. IMAGE REALISED WITH THE ENSIGHT SOFTWARE.

⮩ Intergovernmental Panel on Climate Change: www.ipcc.ch
⮩ Description of climatic models at the Hadley Centre: www.metoffice.com/research/hadleycentre/models/modeltypes.htm
⮩ High resolution simulation with the Earth Simulator: www.es.jamstec.go.jp/esc/eng/

Proteins span the web

BIOINFORMATICS

Proteins get their activity from their three-dimensional structure. These very fragile conformations are tricky to handle. Except by computer…

There was a time when the structural analysis of a protein could only be done at the bench. That was fifteen years ago before structural bioinformatics. Biocrystallography and nuclear magnetic resonance (NMR) were then the only ways of revealing the secret of these molecules: their three dimensional structure on which all their activity depends. By structure we mean the way in which the chain of amino acids making up each protein acquires a particular architecture: it folds itself into loops, helices or other motifs, each fitting precisely together. Biocrystallography involves searching for a solvent in which the protein will crystallise. Then deducing its structure by studying how the crystal diffracts high energy X-rays in a synchrotron. *"To determine the three-dimensional structure of a protein takes on average one year and costs around 150,000 euros,"* explains Gilbert Deléage, Assistant Director of the Institut de Biologie et de Chimie de Protéines (IBCP) and a Professor at the University of Lyon. And the method doesn't work all the time; not all proteins crystallise (particularly membrane proteins which account for 25% of the proteins in the human genome) and crystals do not necessarily diffract X-rays with a sufficient resolution.

Structural bioinformatics and, in particular, molecular modelling have come to speed up this meticulous procedure. The principle is simple. We take the known protein structures in a database and using a program we compare the sequences in such a way that we can imagine the structure of the unknown protein. IBCP is one of the French laboratories to have developed such a programme. *"When we have the linear structure of a protein, we can imagine an infinite number of possible folds,"* explains Gilbert Deléage. *"But we start from the hypothesis that in Nature the number of these folds is not infinite. There is a strong probability that a protein adopts known folds. Experience shows that sequences that are only 30% similar have comparable architectures."*

So IBCP constructed some 13,000 models of the 25,000 proteins in the genome of *Arabidopsis thaliana*, thale cress, geneticists' model plant. This year they have been made available on the web to researchers who can manipulate them in three virtual dimensions. *"The problem with the models is that we need to find a protein's distant cousin whose structure is known so as to compare the sequences,"* admits the researcher.

The databases are sure to be sustained by crystallographers who are adding 10-15 structures every day. But sometimes the resemblance is still too distant. So a threshold is set above which sequences can be considered to be comparable. *"We are working on the extreme cases, what we call molecular modelling at a low level of identity."* This involves identifying secondary structures (helices and sheets) and forcing the model with them. The result – sequences with 10-35% of similarity – can be compared and unknown proteins modelled with 90% reliability.

The other advantage of molecular models is that they allow the geometry of a protein (particularly the spacing of electrical charges) to be compared with those of known molecules. This may elucidate its function. IBCP has thus developed a novel program for identifying functions – active sites or sites of ligand binding shared by sev-

A model brain

Two projects for modelling the activity of the brain were launched in Europe in the summer of 2005. The first, named Bio-i3 (for Bio-Inspired Intelligent Information Systems) is part of the European programme for emerging technologies. Planned to last for four years, it aims to model

© IBM
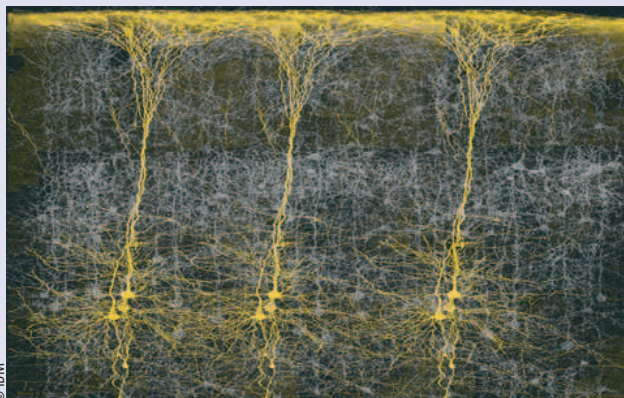
the electrical activity of neurons and to study the properties emerging from the major groups of interconnected neurons, the mechanisms of plasticity, and the changes in these systems. Finally, Bio-i3 will attempt to generate and exploit consciousness of the environment or even consciousness of self. The second programme, Blue Brain Project, is the result of collaboration between IBM and Ecole Polytechnique Fédérale de Lausanne. The goal is to model the exchange of information in circuits of the neocortex, the most complex part of the human brain. This involves describing the electrochemical exchanges between neurons in three dimensions (see the picture opposite). The project heads hope to learn more about the emergence of intelligence so this can by applied to understanding psychiatric and neurological disorders like autism, schizophrenia and depression. The chosen tool is certainly up to the task of this ambitious project: it is Blue Gene, the IBM supercomputer with a top speed 22.8 teraflops. The collaboration is expected to last for two years.

⮑ Bio-i3 : http://domino.research.ibm.com/comm/pr.nsf/pages/rsc. bluegene _cognitive.html ; www.cordis.lu/ist
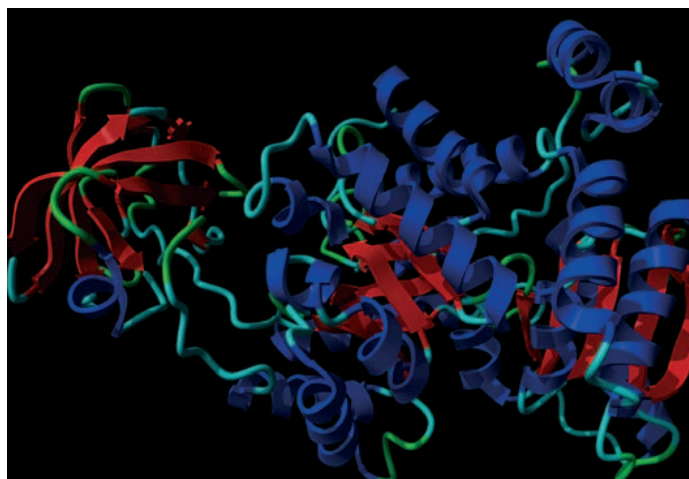
3D model of the pyruvate kinase protein of *Arabidopsis thaliana*. Repeating structures are in red and blue. The image was made with the Lyon bioinformatics server. © INSTITUT DE BIOLOGIE ET CHIMIE DES PROTÉINES

➲ Geno3D, molecular modelling program: http://geno3d-pbil.ibcp.fr
➲ IBCP softwares:
http://sumo-pbil.ibcp.fr
http://modeome3d-pbil.ibcp.fr
www.arianapharma.com

eral proteins. With this, a database of ligands has been set up that can serve as a base for a virtual screen. A screen which is considerably faster and more affordable than the automated ones used today by all the major pharmaceutical laboratories in the search for new drugs…
Virtual pharmacology can already claim some successes. As Mohammad Afshar, one of the founders of the British company Ribotargets in 1997, testifies *"thanks to virtual screening methods, we identified a synthetic molecule having large spectrum antibacterial properties."* Sold to Johnson & Johnson this molecule is the first to have been discovered by digital simulation. Today the major pharmaceutical laboratories systematically use virtual screening and high throughput screening to test molecules with three-dimensional structures.
Fresh from this success, Mohammad Afshar set up a second company in 2003 this time in France. 'Ariana' is working in the step following the detection of an active molecule.

"It is, in fact, relatively easy to find a first active molecule in a therapeutic project. But it must then be absorbed, be stable and non toxic! Researchers therefore spend many years synthesising and testing molecules that fulfil these criteria while conserving the active principle."
Ariana sells a tool to help in decision making that is also based on using databases. *"Our goal is to extract the maximum information from existing data so that chemists only test new hypotheses and they avoid synthesising molecules that are already known or that have predictable properties. Our tool also suggests structures likely to optimise the many pharmacological characteristics of the molecule."*
Even if bioinformatics is increasingly influencing drug development and reducing the clinical tests on humans, it can never replace this. As it does not take into account the extreme complexity and variability of the human organism's responses to the ingestion of an active molecule… ■
Anne Debroise

The first human migrations digitalised

A number of genetic and archaeological facts favour the theory of a single origin of modern man. Laurent Excoffier and his colleagues in the Laboratoire de Génétique des Populations at the University of Berne are today adding new arguments [1]. They have developed a programme called SPLATCHE which for the first time simulates the human species' colonisation of the world as a function of the environmental conditions prevailing at the time. The data (vegetation, availability of water and food, and landscape) are translated into two parameters: the 'load capacity' (maximum number of individuals per locality taking into account its resources) and the 'friction value' (the relative difficulty of migrating in a certain environment). These two parameters are defined at each point of a mesh covering the globe. Each of these points harbours a virtual sub-population (or 'deme') from which the demographic changes and an overview of migration to neighbouring demes are reconstituted over time. All these demographic data are used to simulate the expected genetic diversity at various points of the globe for which we have genetic data. Several simulations have been done with the bases placed at different points of the Ancient World. Each time the simulated and observed genetic diversity are compared. The conclusions of these initial studies: modern man certainly could have had a single origin and could not have evolved simultaneously in several continents. The cradle of mankind is in East Africa in agreement with archaeological data.

[1] N. Ray *et al., Genome Research, 15,* 1161, 2005.
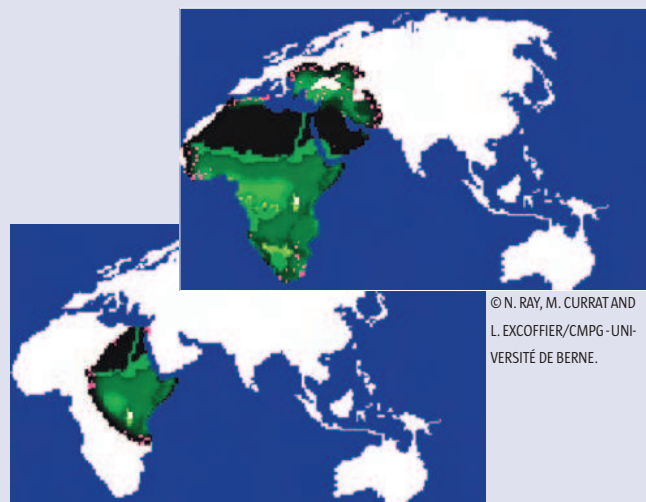
© N. RAY, M. CURRAT AND L. EXCOFFIER/CMPG - UNIVERSITÉ DE BERNE.

Olivier Gascuel:
"At the heart of evolution"

MOLECULAR PHYLOGENY

The joint burgeoning of genomics and computing has revolutionised our way of seeing the evolution of living species.

How has computing helped in the reconstruction of links between living species, what's known as phylogeny?

OLIVIER GASCUEL: Since the end of the 1960s, the spectacular increase in the calculating power of computers and the genome sequencing has given birth to new discipline: molecular phylogeny. We no longer follow the evolution of the morphological characteristics of species, but of their genes and their genomes. Calculations done with programs enable us to reconstitute the history and past sequences starting from sequences (DNA or proteins) known today. We can describe the evolution of plants and bacteria whose fossil records are too rare to be able to use classic paleontological methods. We have made some progress on the history of mammals too.

THIS PHYLOGENETIC TREE includes 21 of the 23 current primate species. It was developed with the PHYML programme starting from DNA sequences of more than 900,000 nucleotides.

© E. DOUZERY ET P. H. FABRE/INSTITUT DES SCIENCES DE L'ÉVOLUTION, MONTPELLIER

OLIVIER GASCUEL is Head of the Equipe Méthodes et Algorithmes pour la Bio-informatique (LIRMM, CNRS/University of Montpellier). © DR

What principles lie behind the reconstruction of ancestral sequences?

We begin by finding homologous genes in different species: similar from the point of view of the sequence, they code for proteins with the same structure and function. This indicates that these genes derive from a common ancestor and we try to establish which changes in an ancestral gene could give rise to the current forms of the gene. So we construct a tree whose leaves correspond to contemporary sequences and the nodes represent ancestral sequences. To develop this tree further, most specialists use the methods of probability that appeared in the 1980s. By following the maximum likelihood principle, they evaluate the probability of having more or fewer changes as a function of passing time between the ancestral and current sequences.

What difficulties do you find?

Calculation time is one of them. To process the sets of data embracing hundreds of species and tens of thousands of sites, it is necessary to find algorithms that avoid exploring 'all' the trees and 'all' the evolutionary scenarios.

A second difficulty relates to the fact that the different sites of the same gene do not always evolve at the same speed. There are moments when the gene evolves very fast because the environment is changing. But not all sites respond in the same way. Those which are essential for the structure or the function of the corresponding protein evolve very little, in general. Those that are less constrained evolve very rapidly. At the moment our group is setting up models that integrate not only the variations in speed and the mode of evolution at different sites, but also the fact that the same site can behave differently over time.

Can we test the validity of the models?

With difficulty, as the few ancestral sequences that are available are degraded. There are, for example, a few genome fragments from Neanderthal man but not much more. Mainly we proceed by cross-checking. A good indicator of reliability is for example to find in an *in silico* reconstruction a rapid period of evolution which fits in well with a change in climate or habitat known to palaeontologists. Very rapidly evolving viruses provide other indications. Virus populations were made to evolve artificially in a laboratory by separating them into sub-populations and doing that several times in succession. At the end of one year we recovered the virus sequences and we applied phylogenetic reconstruction methods. They functioned perfectly well and gave the history of successive divisions. Our models aren't too bad…

What is the point of retracing the history of genomes and studying their evolution?

We enter into the heart of evolution. Nature has not explored every possibility of the physio-chemical world. It combined existing elements, for example, protein sequences. It also transformed and adapted existing proteins to give them different functions. As François Jacob said so well, Nature likes to tinker. ■

interview by A. D.

➲ *Mathematics of Evolution and Phylogeny,* Olivier Gascuel (ed.), Oxford University Press, 2005.
➲ Tree of Life.
http://tolweb.org/tree/phylogeny.html

Antoine Joux:
"Cryptanalysis can only progress"

CRYPTANALYSIS

With the development of supercomputers calculations that were formerly impossible can now be done, weakening security codes. Cryptography relies on the complexity of calculations needed to decode things. Recently a French team has successfully tackled one of the most used protocols.

What is the principle behind hashing?

ANTOINE JOUX: In cryptography, hashing is 'to summarise' a message of a particular size in a shorter message of a predefined size such that two different initial messages correspond to two different summaries. As there are more messages than summaries this goal cannot be achieved: whatever method is used there will always be messages that will be summarized in the same way. This is called collision. In practice, we look for ways to prevent two messages summarized in identical ways from being found easily so as to avoid the possibility that one message can be replaced by the other one.

Is there are great risk?

Suppose that the summary of any message is coded on 160 bits, that's a succession of 160 zeros and ones. The total number of possible summaries is 2^{160}. By taking $2^{160} + 1$ messages we are sure to find a pair of messages summarized in the same

ANTOINE JOUX is the Scientific Adviser for the Direction Générale de l'Armement at Issy-les-Moulineaux. He is also an Associate Professor at the University of Versailles-St-Quentin. © DR

way. But reasoning in this way we don't take into account the paradox called 'the birthday paradox'. This is so called because the probability that in a class of 23 pupils two will have been born on the same day is unexpectedly higher than a one chance in two. Thinking in the same way, the number of messages to test to have one in two chance of finding two with identical summaries is of the order of 2^{80} (the square root of the number of possible summaries). A number which remains out of reach of current computers.

Can we do any better?

In 1998, with Florent Chabaud we found a way of reducing this number for a protocol called SHA-0. By using a technique called differential cryptanalysis, we showed that it is possible to construct two different messages having the same summary in only 2^{61} operations. In 2004,

Eli Biham and Rafi Chen from the Technion Institute of Technology in Haifa, Israel improved these results to get to 2^{56}. Returning to this research I was able to get down to 2^{51} and for the first time achieved a collision. For this I needed to use the Teranova supercomputer set up by Bull on the Ter@tec site. Such a facility is not available to any hacker, but the connection between theory and practice has nevertheless been made.

What are the real consequences of these developments?

For the moment collision only concerns the Sha-0 protocol and this has been overtaken by the SHA-1 protocol for ten years. But cryptanalysis can only progress. A team of Chinese researchers, including Xiaoyun Wang, at Shandong University came up with a method of attacking SHA-1 in 2005. It has been improved even more very recently. The method remains theoretical but current progress should be taken seriously. The National Institute of Standards and Technology the American organization which developed the SHA standard is not going to make any mistakes: it organized a meeting on this subject at the end of October 2005. Even if it is not critical now, the current situation requires some thought on how to protect the future. ■

Interview by Benoît Rittaud

[1] E. Biham *et al.,* Collisons of SHA-0 and reduced SHA-1, Eurocrypt 2005.

Prime or not

Certain methods of cryptography – including RSA – exploit the difficulty of factorizing large numbers. Protocol security decreases as long as mathematics advances. Maninidra Agrawal, Neeraj Kayal, Nitin Saxena from the Indian Institute of Technology in Kampur have found a theoretically rapid and infallible way of knowing if a given number has factors other than itself and one. It's clear that the problem of factorization is not resolved, but it is a breakthrough in a related question. Bad news for the many techniques of encryption? At the moment, no. Resorting to techniques that exploit the arithmetic of prime numbers to cryptographic ends is really more complex than the problem solved by the Indians.

⮑ www.cse.iitk.ac.in/users /manindra/primality_v6.pdf

$2^{25\,964\,951} - 1$

Prime numbers are a link between theoretical mathematics, high precision calculation and cryptography. All progress in prime numbers is therefore likely to have repercussions in code-breaking. The Mersenne numbers, which are in the form 2^n-1, feature among the prime numbers that interest theorists the most. Not being well understood, they must be satisfied by looking for the largest ones. The latest record is by Martin Nowak, a German ophthalmologist. He demonstrated that $2^{25964951}-1$ is a prime number. This discovery was made thanks to shared computing software from the large Mersenne prime numbers Internet research group. By lending the group part of the calculating power of his personal computer, he contributed to the mass of calculations required to find this rare jewel. Imagine being the proud owner of the computer which reveals the next Mersenne number…

⮑ www.mersenne.org
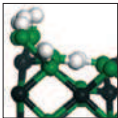
Simulation is **an essential tool!**

Simulations have changed our way of studying the physics of materials. What's the evidence? With computing today we can predict to within a few percent some of the mechanical, optical, electronic or chemical properties of materials. The models used rely mostly on ab initio calculations. However simulation on the macroscopic scale is still out of the reach of even the most powerful computers: a cubic micron of matter (the tip of a hair) already has a hundred thousand atoms to be modelled over seconds while they are oscillating at speed (a fraction of a picosecond, 10^{-12} second)! Fortunately, mostly such precision is not essential. We confine the ab initio calculations to small cells of 100 to 200 atoms characteristic of homogeneous zones. And we couple them to models that are less precise but on a larger scale. These multi-scale simulations in both space and time can represent most phenomena.

The microelectronics industry changes scale

ATOMISTICS

Even smaller! This could be the motto of designers of microelectronic components. But moving into nanoelectronics requires a one thousand times reduction in the size of components. A difficult threshold to cross…

Often applied research and pure research are pitted against each other. *"The ways of doing things however are complementary. The microelectronics industry in quest for miniaturisation is asking us to develop modelling methods for their future products,"* states Frédéric Lançon, Head of the Atomistic Simulation Laboratory (L_SIM) at the CEA in Grenoble. *"And to do this they even take on physicists!"* So what unites the two communities? A common interest in… the nanometre (10^{-9} metres): the size range of components to be produced for industry.
'Atomistic' simulation, the realm of computing electronic structures and quantum mechanics, was included in the International Technology Roadmap for Semiconductors in 2001. It is a reference standard for the microelectronics industry which sets the indus-

trial objectives and the technological barriers to cross. The numerical models used to scale down the transistors and regulate their output are becoming poorly suited to the nanometric scale. *"We are working on the insertion of germanium atoms into silicon, for example,"* explains Frédéric Lançon. Industry will need this to increase the mobility of electrons in the transistor channels. *"The problem is to control the insertion of atoms so as not to create structural defects and keeping experiments to a minimum,"* he says further. *"We can predict the effect of defects by simulating the basic mechanisms of quantum mechanics (ab initio calculations and molecular dynamics) coupled to statistical methods (like the Monte Carlo Method)."* Some

Microelectronics industry needs modelling methods for their future products

structures are calculated for hundreds of cells of one hundred atoms each, with their kinetic changes spread to a million atoms using the CEA com-
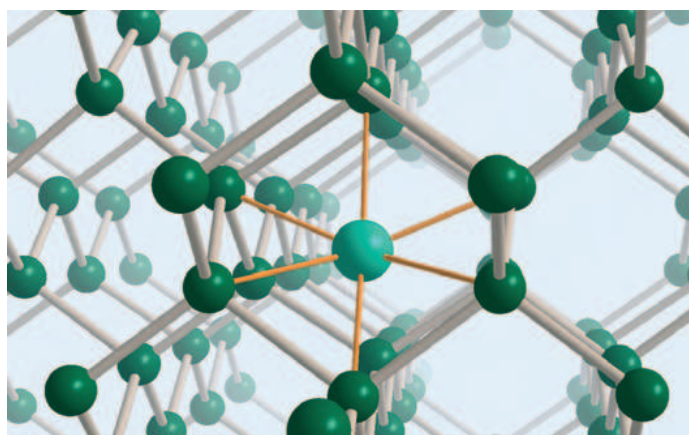
INDUSTRY IS TURNING TO SIMULATION to increase the mobility of electrons in transistor channels. The difficulty is to insert germanium atoms in the silicon without creating defects in the structure. © CEA

puting facilities. The laboratory launched a project (OsiGe_SIM) with the support of the Agence Nationale de la Recherche to link this subject to the study of silicon oxidation with LAAS (CNRS, Toulouse), CRMCN (CNRS, Marseilles), Leti (CEA, Grenoble) and STMicroelectronics.
Most of the research is more forward looking. The European project BIGDFT, coordinated by L_SIM and which started in 2005, includes mathematicians and physicists. The goal? To simulate functionalised nanoobjects by molecule grafts. *"Here too we are going to take*

different degrees of precision in the description of the system to model it," explains Thierry Deutsch, Project Coordinator. *"We only precisely describe the electronic structure of the nanoobject where it varies the most by using wavelet theory. This process is integrated in Abinit, the free international software run by a Belgian lab, and should allow us to simulate thousands of atoms ab initio instead of only the hundreds today."* ■ **Isabelle Bellin**

⮥ L_Sim laboratory: www-drfmc.cea.fr/sp2m/L_Sim

⮥ BigDFT: www-drfmc.cea.fr/sp2m/L_Sim/BigDFT/
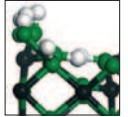
Making **virtual** materials

The mechanical properties of a solid, its plasticity or resistance to breaking, for example, are intimately linked to the dynamics of countless structural defects. By modelling these, physicists can predict the properties of materials.

To design and scale up objects that they produce, industry uses macroscopic codes based on empirical data that they adjust in line with experimental measurements. *"Eventually, our aim is to replace some of this experimental data with results coming from atomistic simulations,"* states Gilles Zérah from the Department of Theoretical and Applied Physics at the CEA (Bruyères-le-Châtel). *"This software of the future will combine the tools which are being developed in different laboratories: software for simulating the atomic scale, dynamic simulations of defects, the industry's structural codes…"*

How do these models describe the mechanical behaviour of a metal, for example, its resistance to bending? Although simulations based on molecular dynamics (knowing the direct interactions between the atoms) are capable of following ten million atoms for several nanoseconds they become inoperable when a deformation occurs involving tens of thousands of atoms over one second. The phenomena to be modelled here are structural defects – missing atoms (vacancies) or additional atoms (interstitial) – and dislocations, long lines of defects. So many weak links which determine *via* their movements and their
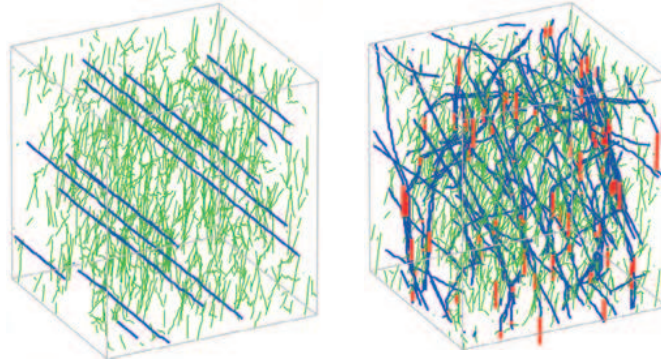
TO PREDICT THE PROPERTIES OF A MATERIAL, physicists try to follow by computation the movement of defects. As seen in these images, two types of dislocations (blue and green lines) in a material under strain can interact and locally cancel each other out at intersections (in red). The shorter strands are no longer mobile. © CEA

interactions the macroscopic properties.

"To simplify this we can consider these defects as separate objects and develop models of their own changes in behaviour on scales of time and space much larger than for atomistic simulation," explains Gilles Zérah. For linear defects researchers are developing models of the dislocation

> **The software of the future will combine tools for simulating the atomic scale, dynamic simulations of defects and the industry's structural codes…**

dynamics. They have thus demonstrated that the mutual cancelling out of dislocations influences hardness, an effect unknown up until then. And therein lies the interest of simulations: to understand the behaviour of materials and above to predict their properties while industry's empirical models will always be confined to the experiments on which

they were based. So together with experiments, numerical simulations widen the scope of a model's validity.

"Our team is developing several software tools suited to this multi scale approach," explains Gille Zérah one of the them is Abinit in development since 1995 with the Catholic University of Louvain (Belgium). Today it is an international standard code for the *ab initio* simulation of materials. We have also developed a model of molecular dynamics for Leti (Laboratoire d'Électronique et de Techniques d'Instrumentation), CEA and more recently in collaboration with LEM (Laboratoire d'Études des Microstructures, CNRS-ONERA), a dynamic model of dislocations (Micromegas) which has been distributed internationally for two years. These software are good candidates for inclusion in the generic computing platform of the IOLS project for System@tic, the new competitiveness centre. ■

I.B.

➲ Abinit program: www.abinit.org
➲ MicroMegas dislocation code: www.zig.onera.fr/mm_home_page

Semiconducting nanofibres

Semiconducting nanofibres will perhaps be the basic electronics components of the future. *"But these miniscule structures of several nanometres in diameter are still very difficult to characterise precisely,"* explains Yann-Michel Niquet from the Atomistic Simulation Laboratory at the CEA in Grenoble. *"Simulations are an excellent way of understanding and optimising their properties – particularly, the transport of electrons and the optical properties. They also make it possible to focus experiments more choosing ranges of diameters to test. Structures of a million atoms (a piece of nanofibre 10 nm in diameter and 250 nm long) can be simulated in one night. "Our methods are semi-empirical,"* specifies the researcher. *"Such ab initio calculations are out of the range of computers. So we accurately simulate the interactions between the atoms in a perfect crystal of silicon and we transpose these physical parameters to the nanofibre."* This shortcut agrees with experimental results. The simulations are done in the framework of a national theoretical project (TransNanoFils) and a huge European project NODE whose purpose is both experimental and technological. Beginning in October, it brings 13 partners together (including Philips, Infineon and IBM) coordinated by the Swedish University of Lund.

➲ http://cmliris.harvard.edu/research/overview/index.php
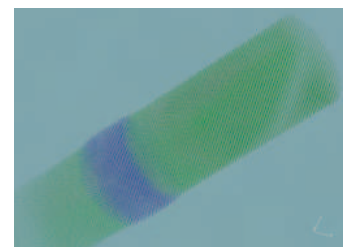
SIMULATIONS are used to optimise the properties of nanofibres and to target experiments. Here a quantum box of indium arsenide (InAs) is included in a gallium arsenide (GaAs) nanofibre. The gallium atoms are in green, indium in blue and arsenic in grey. The diameter of the GaAs fibre is 10 nm and the INAS fibre 6 nm.
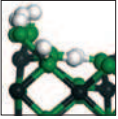
Irradiated steel aging *in silico*

How can the mechanical properties of a nuclear reactor containment vessel be modelled from its microstructure? By joining up physical and algorithmic modelling…

The walls of a nuclear reactor containment vessel and its components are subjected to rough treatment, 'bombarded' with neutron fluxes and brought to temperatures sometimes higher than 340°C! The safety of the installation and its lifespan depends on its mechanical integrity. It's not surprising that the aging of steel, the main material, is closely monitored. How can changes in the mechanical properties of such imposing structures (a reactor measures more than 13 metres high and is 20 centimetres thick, that's more than 300 tonnes of steel) be modelled in the long term (over 50 or 60 years)? At the moment we are using experimental reactors or empirical models where samples are plunged into the reactor. *"However for 5 years, thanks to new computing methods, with the CEA we have been able to begin modelling the irradiation of steel at the atomic scale,"* say Pascal Mialon, Head of the Department of Materials and Mechanics of Components at EDF. *"New experimental tools like the atomic tomography probe allow us to compare simulation and experiments on the atomic scale. We can see the impurities and follow their rearrangements in the structure."*
The aim is now to simulate the

changes in the microstructure at the atomic scale, in particular, defects due to irradiation. A huge European project was started in January 2004 for a period of 4 years. 'Perfect' includes 28 partners coordinated by EDF. These efforts should lead to the creation of a platform combining different multi scale numerical simulation models. Two simulators (one of the containment vessels, the other of the internal structures) should then be built.
"The scientific effort required is enormous," explains Pascal Mialon *"The phenomena at work are complex."* To simulate neutronic bombardment boils down to modelling a game of atomic billiards: under the influence of neutrons, atoms move and defects are created (missing atoms or additional ones). The complete structure

"We can see the impurities and follow their rearrangements in the structure"

gradually rearranges and weakens. *"We are using quantum physics models based on* ab initio *calculations at the atomic scale,"* explains Pascal Mialon. *"The first difficulty is to adapt these calculations to the steel's microstructure in a simplified but representative way. Then we develop molecular dynamics codes to simulate the interactions between the defects and their movement. The long term changes in the microstructure are modelled using statistical methods (Monte-Carlo kinet-*
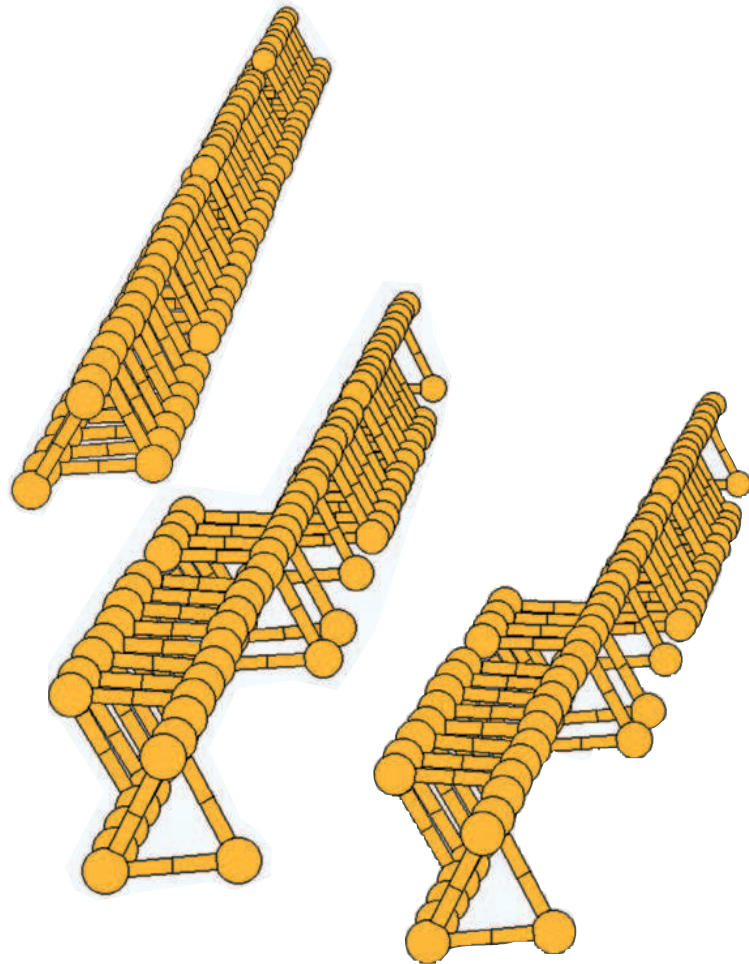
RESEARCHERS HAVE CALCULATED here the relative movement of atoms (dislocation) when the material is deformed. © DOMAIN ET MONNET/EDF

ics). The codes for the dislocation dynamics (movement of lines of slippage) allow us to go up to the macroscopic scale, then to the total structure with our own codes by finite element." This kind of mechanical property simulator should see the light of day between now and the end of 2007.
"In parallel, we are improving the algorithmic methods and programming in order to integrate the enormous changes in spatial and temporal scales which are at work", adds Jean-Louis Vaudescal, Head of the SINETICS Department at EDF. *"For example, for metallic materials we are developing numerical methods for* ab initio *calculations where the calculation time varies linearly with the number of electrons* N *(calculations of the order of* N*). With current methods the calculation time varies with its cube (*N^3*)."*

Metals are particularly difficult to deal with as they are conductors and their atoms interact over long distances. With current machines we know how to do calculations of the order of N on thousands of atoms with insulating or semi conducting materials (by ignoring this interaction). But for metals these methods are still confined to a hundred atoms or so. So the goal with *ab initio* calculations is to simulate sets of around ten thousand atoms. A goal which one day will be within reach with improvements in numerical methods, a continued increase in computing power and parallelisation methods, all of which are being researched. ∎

Isabelle Bellin

↪ Perfect European Project:
http://coscient.edf-labs.net/site/index.htm
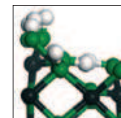
Michel Guttmann:
"Models that guarantee the safety of reactors"

Since the 1960s, the Department of Research in Physical Metallurgy at the CEA has been trying to understand and model the long term effects of radiation on metal and insulators. The current director sums up the state of play.

Michel Guttmann is Research Director and Head of SRMP (Service de Recherche de Métallurgie Physique) at the CEA. © D. R.

How is your work in pure research connected to the challenges of the electronuclear industry?

MICHEL GUTTMANN: They are all basically essential: the predictive modelling of the long term behaviour of materials is essential to ensure the safety of reactors, to prolong their useful lifetime, to store radioactive waste, to develop new reactors... All of this requires the capacity to extrapolate as it cannot be founded only on experiments. Only modelling based on the physical description of the basic mechanisms starting from the atomic scale, paradoxically but undoubtedly the most certain, can guarantee it. Numerical simulation is a central tool in this new physics. It allows us to follow the changes in the intimate structure of a material disturbed by cascades of atomic displacements induced by radiation up to the macroscopic scale. Caused by the mechanisms of the atoms' movements – thermal diffusion and nuclear collisions – occasional defects in the crystal move around and interact leading to the formation of lots of defects, cavities, gas bubbles, piles of atoms of elements from alloys and impurities, and precipitates, etc.

Where is this multi scale modelling at?

Complete modelling up to the induced alteration in the macroscopic behaviour is still a dream. There are twelve orders of magnitude to cover for space and twenty-four for time from the electronic structure of atoms up to the decades of aging for reactor vessels or thousands of years for waste! We are still only paving the way of space and time with small tiles being developed; but it's advancing very quickly thanks to methods like *ab initio* quantum calculations of the basic properties of atoms; molecular dynamics for the initial ballistic damage; various methods of simulation based on algorithms of the Monte Carlo type for the slower changes; field methods for the kinetics of a homogenous chemical, etc. As an example, we recently succeeded in linking *ab initio* computing with a novel Monte Carlo code developed at the CEA to precisely reproduce the changes in defects in irradiated iron over about one hour.
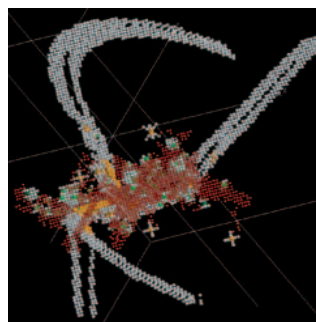
Will modelling be able to replace experimentation?

The opposite's true! They are inseparable. Targeted experiments are the only way to discover and analyse the physical phenomena to be modelled, and then to perfect and prove the models. Current tools allow us to produce and study the damage caused by the irradiation of charged particles on the same scale as our models. This is why we launched the JANNUS Project (Jumelage d'Accélérateurs pour les

THIS NUMERICAL MOLECULAR DYNAMICS EXPERIMENT simulates on the atomic scale the damage induced by a cascade of displacements induced by irradiation and its effect on dislocations. Displaced atoms ejected from their site (red), interstitial atoms (orange); vacancies – the crystalline sites left empty (green); atoms with disturbed environments, the dislocation centre, neighbouring defects (in grey).

©N.V. DOAN/CEA

Nanosciences le Nucléaire et la Simulation), in collaboration with IN2P3 (Institut National de Physique Nucléaire et de Physique des Particules). Two accelerator groups, one at Orsay and the other at Saclay, will allow us to subject samples to several forms of damage simultaneously as in real life while observing them *in situ* by transmission electron microscopy.

Who are your partners?

We are applying our atomic models to the behaviour of individual dislocations, long linear defects which govern the plasticity and the breakage of material. After this, our colleagues in the Department of Applied Metallurgy Research in the CEA take over. From the dislocation dynamics models which describe the collective behaviour on the scale of a basic crystal, they model the polycrystal using codes for crystalline plasticity on finite parts and homogenisation techniques to integrate all the scales up to the macroscopic mechanical behaviour. The need for computing resources is always growing. All these studies are done in cooperation with the CNRS and our industrial partners EDF and Framatome in European programmes – Perfect for fission reactors and the European Fusion Development Agreement for Future Fusion Reactors – and three CNRS-industry research contracts. ∎

Interview by I.B.

Virtual humans

INRETS (Institut National de Recherche sur les Transports et leur Securité) is working at developing more and more realistic and versatile models of the human body. With bones, muscles and much more. This is for performing reliable simulations of all sorts of traffic accidents. *"INRETS is leading the European HUMOS2 project, funded by the European Commission,"* explains Philippe Vezin, one of the researchers involved. *"Between 1997 and 2000, HUMOS (HUman MOdel for Safety) created a digital model of an 'average' man in driving posture. HUMOS2 (September 2002 to February 2006) added a small woman and a tall man and to provided standing-up versions of all three."* Some tools allowed intermediate sizes and postures to be created. The researchers are now tackling intricate problems such as muscle tone which can influence the way a body reacts to impact.

⮑ http://humos2.inrets.fr

A smoke-free Grand Palais

"We have been developing digital tools to study fire for twenty years," explains François Demouge, research engineer in fire safety at the CSTB (Centre Scientifique et Technique du Bâtiment). *"The new thing is that simulation is nowadays part of the regulatory process."* While the main building of the Grand Palais in Paris was being reopened after renovation, a CSTB team was digitally testing the smoke-clearing equipment planned by the architects. *"We were able to confirm its efficiency. The existing openings had been widened enough to meet the safety regulation standards."* This type of simulation is common practice at the CSTB. Many public buildings or tunnels are now tested this way. For example, the future Zenith concert hall in Strasbourg and the exhibition park in Angoulême have both passed this digital trial.
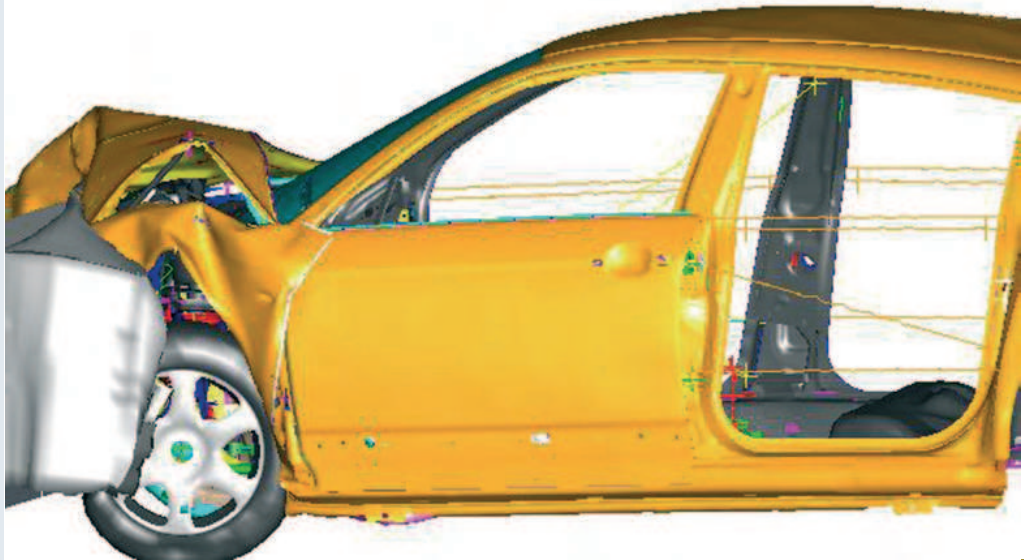
Frontal impact simulation of a Citroën C6. © PSA PEUGEOT/CITROËN

Well-calculated collisions

CAR CRASHES

Computerised car crashes have become an everyday event at PSA-Peugeot-Citroën. The aim is to improve the passive safety of the whole vehicle.

"Counting all car models, we perform about 20,000 crash simulations per year," says Laurent Di Valentin, Head of Modelling, Graphics and Passive Safety at PSA-Peugeot-Citroën. *"It allows us to optimise the passive safety of new vehicles by acting very early during the design process. Of course, we still carry out a real crash from time to time (about one for every ten computer-generated ones). But nowadays, we know in advance what it will show; the real-life trial doesn't bring any surprises."* It's not a trivial exercise as an automobile is a very complex object. It is made of numerous components that obey various 'Materials Science Laws' and the crash causes a wide variety of physical phenomena that are hard to simulate. There are major body deformations, breaks, and more generally, non-linear phenomena, contacts, etc. Beyond the parts the properties of many assembly techniques – bolts,

welding, glue, etc. – need to be simulated.

After the usual impacts (frontal, lateral against barriers or posts, rear), we're interested in 'pedestrian impact', which has been among the numerous tests in the homologation process since October 2005. The software models the impact of legs on the bumper and of heads on the bonnet. When it involves a sophisticated vehicle like the brand new Citroën C6 equipped with accelerometric detection on its bumper and a pyrotech-

> **"We still carry out a real crash from time to time. But nowadays, we know in advance what it will show: it doesn't bring any surprises"**

nic active bonnet, the simulation becomes very subtle.

"The digital models that we're using," says Laurent Di Valentin, *"often involve 600,000 to 700,000 components. An impact typically lasts 140 milliseconds, but most of it happens during the first 110 milliseconds."* The time increments used depend on the speed of sound propagation in the mate-

rials and of the smallest mesh size (10mm). *"At the moment we're using a time increment of 1 microsecond; that figure should decrease when the meshing becomes narrower."* Usually four processors work for 48 hours for a single simulation. But routinely five or six computations are done at the same time in order to test various hypotheses. Engineers have at their disposal a cluster of Fujitsu servers totalling 700 processors.

Nearly 20 years after the first car crash simulation trials, the impact of this technology on the evolution of passive car safety is considerable. *"We start to digitally test a new model,"* Laurent Di Valentin says, *"about one year before the completion of the prototype. So many things can be tested. Because we simulate the entire vehicle very early on we can achieve a broad compromise by adjusting many components. At this stage we can even reconsider the main body parts as it's far from being set in stone."* And all of this is done by breaking digital cars for almost nothing, considering the 'true' cost of one million euros for a real prototype crash. ∎

Pierre Vandeginste

⮑ www.psa-peugeot-citroen.fr
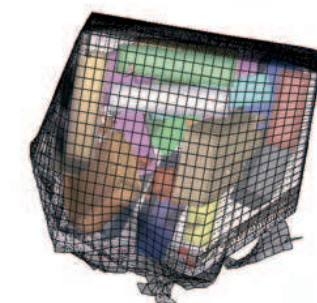
Digital **disasters…**
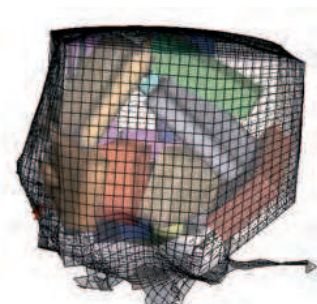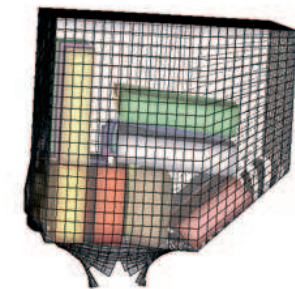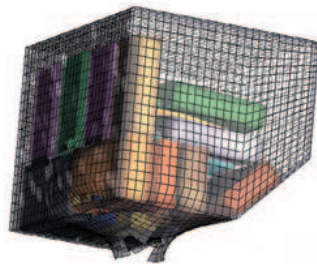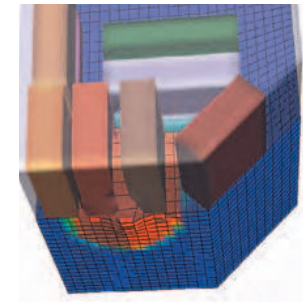# to avoid the real thing

AERONAUTICS

Aviation nowadays uses digital simulation to study all kinds of dramatic situations. This is one of Lille Onera's specialities.

Aeronautics was one of the first industries to exploit the potential of digital simulation. The expression '*digital wind tunnel*' sums up this revolution: a new plane is flown by computing even before its model is tested in a wind tunnel. Since then, the use of digital simulation never ceases to spread to new aspects of aviation. For example, Eric Deletombe, from Onera in Lille, makes civil aircraft land on the sea. The sea may be calm or rough, it depends. This is particularly in the context of the European programme CRAHVI (Crashworthiness of aircraft for high velocity impact).

"A *sea-landing lasts one second; that's the crucial initial phase. A quarter of a second when the tail of the aircraft impacts with the water. Another quarter of a second when the nose plunges. After that it all depends…*" Each quarter of a second is simulated in 15 days of calculation taking up four processors in the parallel computer belonging to Lille Onera.

Why these mock sea-landings? To verify the aircraft's integrity, but also to check the reaction of the luggage racks or the cabin seats: future aircraft models could benefit from these observations. Regarding the luggage, Eric Deletombe and his team are also working on explosions in the baggage hold. For this, the luggage is simulated in detail: a suitcase doesn't behave like a backpack or a metal case. All of them can absorb some energy but can also become missiles. The luggage hold is also simulated: it can of course play a protective role. Finally the plane itself must be modelled in every detail. It is not done here to study just the aerodynamics but the resistance of each rivet to ripping. Not forgetting the last difficult beast to simulate; the explosion itself requires a model of more than one million elements.

With these calculations the 'worst case scenario' can be studied, visualising, for example, where and how the outer shell will 'unbutton' itself along the dotted lines that are the lines of riveting, and more generally, evaluating the devastation in any given situation and the consequences on the aircraft's survival.

What lessons can be learned from this kind of simulation? Ideas on how best to spread the luggage in different types of hold. Information on container design. And finally, they allow in an extreme situation to be studied to potentially lead to novel assembly techniques.

A full list of risk scenarios studied by Eric Deletombe's team doesn't stop here. It includes emergency landing and bird impact. But rest assured, only virtual birds are harmed in the making of these simulations… ∎

P.V.

⮑ www.onera.fr
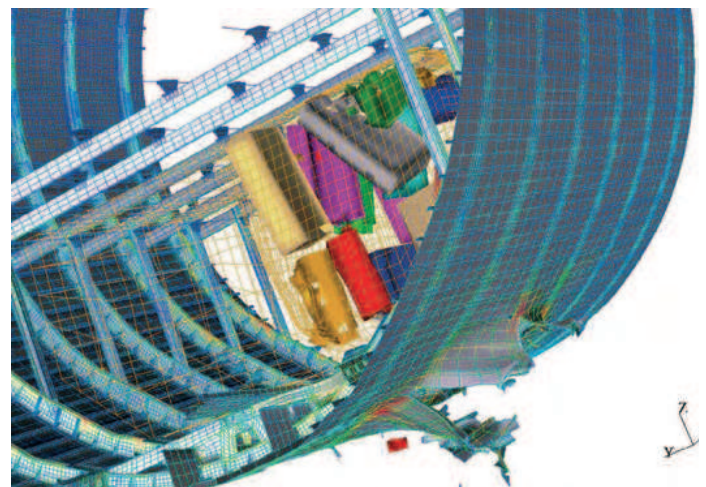
DIGITAL EXPLOSION in the baggage hold of a commercial aircraft.
© ONERA

Jean-Jacques Thibert:
"The difficulty of **modelling turbulence**"
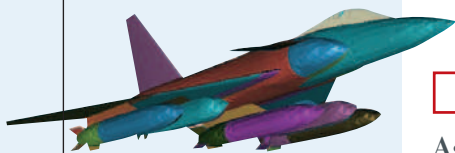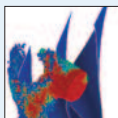
Mesh generation

It's impossible to talk about digital simulation without talking of mesh generation. The methods of using finite elements, at the core of many models, implies calculations… on a finite number of elements, the polyhedrons of the mesh. *"To model the airflow around a plane we 'mesh' its outer volume with a few million tetrahedrons in only a few minutes. It required four hours in 1990,"* says Paul-Louis George, head of the GAMMA project at Inria (Institut National de Recherche en Informatique et Automatique). This world class team is developing algorithms that automatically generate meshes. One of their software, GHS3D, is part of numerous industrial codes used in the aeronautics and space industries. *"Our mesh generator is generic: the size of the meshes is automatically defined according to the metric specifications of the problem, and then adapted with mathematical models (error estimators)."* Each calculation lasts several hours. In general, it is iterated several times with new instructions for adapting the mesh. *"We are researching how to improve these adaptations to better define the metric specifications, to make anisotropic (irregular tetrahedrons) or very large (100 millions tetrahedrons) meshes or even hexahedric meshes…"*

➲ www.inria.fr/recherche/equipes/gamma.fr.html
➲ *Le Maillage facile*, P. J. Frey and P. L. George, Hermes Lavoisier, 2003.

AERODYNAMICS

As surprising as it might seem, turbulence is still difficult to simulate digitally even with the biggest computers…

What is the current state of digital simulation in plane aerodynamics?
JEAN-JACQUES THIBERT: Better forecasting of aerodynamic performance is, of course, a main concern of the industry. Optimising shape allows improvements in flight performance and reductions in fuel consumption, and hence pollution. With digital simulation we reduce the numerous steps necessary for the definition of a new form and more importantly the number of windtunnel trials. Major progress has been made in the last few years, which is especially down to the progress in computers. Proof that aerodynamic simulation is nowadays common practice, our standard platform, elsA, (ensemble de logiciels pour la simulation en Aérodynamique), transferred to the industry in 1998, is used increasingly by Airbus in France and abroad, Snecma, Turbomeca, Eurocopter, CNES, etc. elsA remains nevertheless a real research tool that can be developed further.

What is the scientific basis behind this simulation?
Prediction of airflow around an aircraft relies on solving Navier-Stokes fluid mechanics equations. They allow the calculation of some fundamental physical data to validate the plane's shape (pressure on the surfaces, drag, lift, etc.). Our big problem in aeronautics is

that in the vicinity of the surfaces, the flow is neither laminar nor regular but is more like chaotic eddies. And these are difficult to simulate numerically. We can only tackle them from a statistical angle. This turbulence however is crucial; by directly influencing the drag, for example, it determines the plane's fuel consumption.

Can we expect some improvement in modelling turbulence?
Of course, it is one of the main focuses of development for the elsA code. In parallel, other research aims at improving the speed of computation using numerical techniques and parallelisation methods. At the moment, the turbulence on the scale of the entire plane is estimated. Our objective is to gradually free the models from this constraint. Solving the entire set of Navier-Stokes equations (also called DNS or direct simulation) remains out of the

> With digital simulation we reduce the numerous steps necessary for the definition of a new form and more importantly the number of windtunnel trials

range of current computers; the meshing and time increments are much too small. It is however possible for simple shapes. We are also starting to accurately compute turbulence on the large-scale (the Large Eddy Simulation (LES) approach). But we are a long way from doing so for an entire plane. That would cost 20 million dollars compare with the 30 dollars with current models….

JEAN-JACQUES THIBERT is Head of the Applied Aerodynamics department of ONERA in Châtillon. © DR

Is there really any point in trying to improve the precision of these calculations?
Yes. It's the only way to improve aerodynamics while keeping the numbers of windtunnel trials to a minimum. At the moment, the certification of several characteristics and the limitations of the flight profile (landing and takeoff) are not simulated accurately enough. Only windtunnel trials can reproduce these conditions. It is the same when testing the efficiency of the control surface or some naturally occurring movements like shaking. ■

Interview by Isabelle Bellin

➲ The ONERA Applied Aerodynamics department site: www.onera.fr/daap
➲ Some great images of digital simulation: www.onera.fr/photos/simulations.html
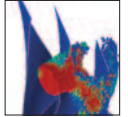
Deep in the heart of the **jet engine**

How can the energetic phenomena that occur at the heart of the jet engine be modelled?
Scientists are working actively on the development of new codes that link aerodynamics, thermal science, chemistry and thermodynamics.

Things are hotting up…! Inside the combustion chamber of an aeroplane engine, the atmosphere is scorching hot and agitated to say the least: air and kerosene mix and burn at nearly 2000°C. At the exit the hot air drives the blades of the turbines… The blades can't resist temperatures higher than 1500°C, the melting temperature of their nickel-titanium alloy… Their resistance is vital to the lifespan of the engine. The answer? Drill some holes that will cool the core of the blade and create a buffer of cold air at its surface. *"In such complex cases digital simulation allows us to deduce temperature patterns and to understand the appearance of hotspots and finally define the optimal geometry for the separation and shape of the channels…,"* explains François Vuillot, Assistant Director of the Department of Digital Simulation of Airflow and Aeroacoustics in Onera. *"Studying thermal interactions between the hot gases coming from the combustion chambers and the blade, we have been able to predict their efficiency."*
Such simulations have only been possible for a few years. And for just cause: they imply

knowing how to model the thermal constraints on the blades as a function of the variable aerodynamic flow of gases sometimes mixed with droplets. *"Previously, each speciality had its own model. Some involved the thermal science of solids, others flow physics,"* says François Vuillot. *"The different calculations provided the conditions at the limits of our energetics models. But our results relied mostly on codes which were independently acquired."* Taking the simple case of the flat panel placed in a cold airflow. The calculated thermal flux can vary by a factor of 5 according to the coupling mode selected and the thermal conductivity of the material! The idea of coupling models in a single platform has gained ground. In the USA, even complete jet engines have been simulated. This work, done at the University of Stanford in the framework of a vast programme on digital simulation ASCI (Accelerated Strategic Computing Initiative),

mainly uses multiphysics computation.
There's nothing comparable in Europe at the moment. On the French side, in April 2004, Onera delivered a second version of its code dedicated to energetics, CEDRE® (Calcul d'écoulements diphasiques réactifs pour l'énergétique). The beneficiaries? Some indus-
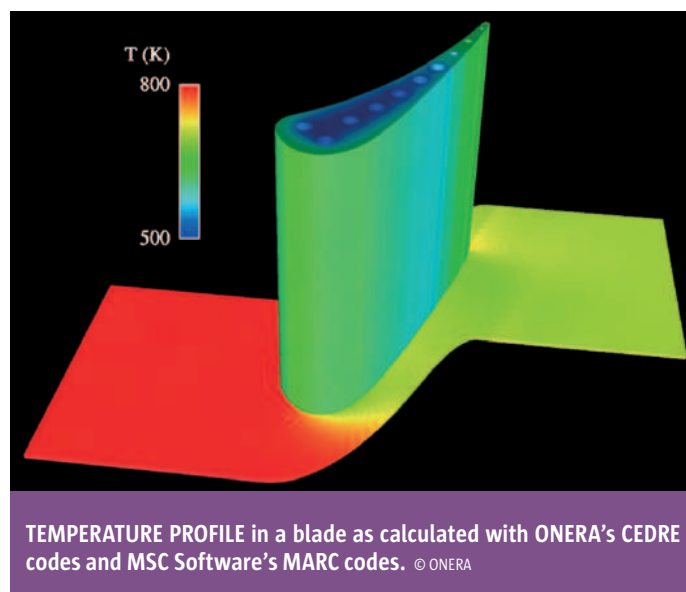
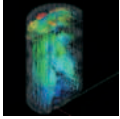**The idea of coupling models in a single platform has gained ground.**
In the USA, even complete jet engines have been simulated

tries and research laboratories (Snecma, SPS, EADS, MBDA, Cnes, ENS-Cachan, etc.). CEDRE is common to the different departments of ONERA interested in flow physics with combustion (flow of gases, liquids or solids, thermal phenom-

ena, radiation, conduction, chemical reactions, turbulence, etc.). *"At the moment, our models of diphasic flows – gas/liquid or gas/solid mixtures – are functional,"* specifies François Vuillot. *We are currently coupling the calculations on 'aerothermy' (aerodynamics and thermal sciences)."* In concrete terms, a calculation is done simultaneously linking these different aspects and exchanging information on the physical quantities calculated, for example, the heat flux transmitted by the liquid to solid up to a stationary point.
"We are also researching how to make CEDRE interact with external codes, like the thermomechanics codes used by industry (ONERA's ZeBuLoN, Dassault Systèmes' Abaqus or MSC Software's Marc). This time, it involves dividing the data while maintaining each code in its environment. We imagine the types of 'sockets' to which codes would be connected. Sockets that would enable the construction of data to exchange and be exchanged," continues the researcher. It is the subject of several national programmes like ARCAE (Onera, Snecma, Turbomeca, CNRS, IVK, DGA, DPAC) for the turbine blade aerothermy. It also involves simulating the transition phases – the changes of engine regimes – more complex and more demanding from the numerical stance. This problem is at the heart of the national Atran programme, grouping Airbus, Snecma and Onera. The latter is also investing in more complex coupling in order to take structural deformations into account. ■ **I.B.**


TEMPERATURE PROFILE in a blade as calculated with ONERA's CEDRE codes and MSC Software's MARC codes. © ONERA

Controlled fusion:
towards models in 5 dimensions

FUSION REACTOR

Producing energy from controlled fusion. Much research is still required before this dream becomes reality. We need to understand, to simulate and to visualise the physical phenomena at work in plasmas.

The decision came several months ago: the International Controlled Fusion Research Reactor (ITER) will be built at Cadarache in Bouches-du-Rhône. But many decades will pass before an industrial reactor is built as the physical phenomena in question are complex. It is necessary to calculate the ideal magnetic topology of the reactor to ensure the best plasma stability from a magnetohydrodynamic (MHD) viewpoint. But also to understand the interactions between the plasma and the walls of the reactor, between the plasma and the electromagnetic waves…. without mentioning the ungraspable physics of the turbulence which regulates the energy confinement in the reactor! *"We are developing two complementary strategies,"* explains Xavier Garbet in charge of the Transport, Turbulence and MHD Group at the CEA (Cadarache), *"an integrated model and a model from first principles said to be ab initio."*
The first one will be used to build a Tokamak simulator like flight simulators for training pilots. With this integrated modelling, we can both prepare
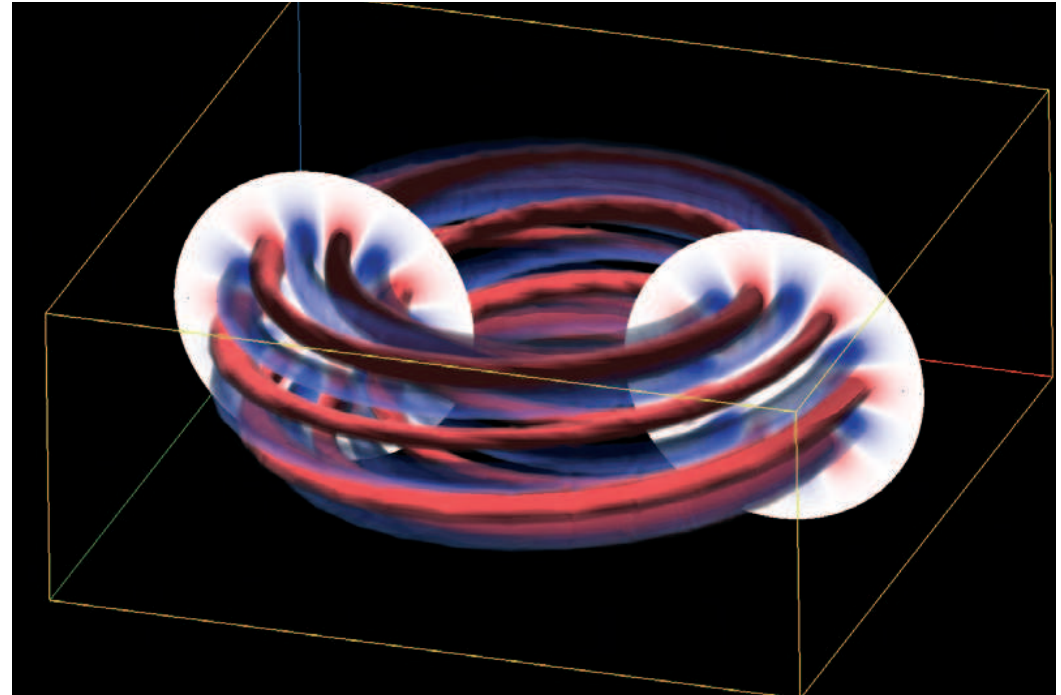
SIMULATED PLASMA. The high values of electrostatic potential appear in red. © CEA/INRIA

experiments (predict the plasma profiles) and interpret them using experimental measurements, and all in a reasonable computing time (several hours for one simulation).
It is also a design tool for future tokamaks. Only the essential physical phenomena are described in a simplified way. *"An integrated modelling initiative was launched in 2003 with the countries of the European Union so as to coordinate efforts and develop common software,"* describes Xavier Garbet. *"Nevertheless to model ITER plasmas and validate this integrated simulation also requires that the phenomena are reproduced by calculation on a small scale."* In particular, we must take into account the plasma's kinetics: gyrokinetic models in 5 dimensions are used to describe the local distribution of particles according to their speed, transverse and parallel to the magnetic fields. *"For example, we must model turbulence on the millimetre scale where the characteristic time is around twenty milliseconds. And this in a plasma several metres in size which may remain confined for 500 seconds,"* continues the researcher. *"The power of calculation necessary is obviously phenomenal and of the order of that used for studies in climatology or theoretical physics."*
Unlike the United States where fusion benefits from earmarked financial support for the development of codes and a dedicated computing centre, Europe has not yet put in place such a research structure. The Cadarache researchers are using the most powerful machines (around 1 teraflop) at the Centre de Calcul, Recherche et Technologie (CCRT) of the CEA at Bruyères-le-Châtel. In spring 2006, they should be able to take advantage of Tera-10's 60 teraflops. *"A single partial simulation of the ITER plasma will draw on the complete capacity of this machine for around eight days,"* underlines Xavier Garbet, *"and much more still to simulate a complete discharge. So we will need a capacity of the order of 100 teraflops or even more."* ■
Isabelle Bellin
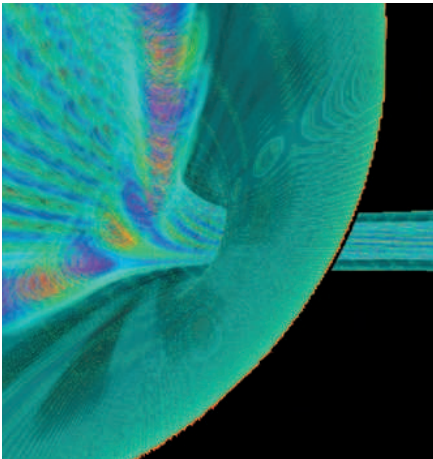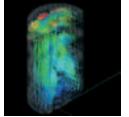
Show me the plasma !

"We are probably the first team in the world to develop specific tools for visualising plasmas," **claims Eric Sonnerdrücker, a lecturer/researcher at the University of Strasbourg in charge of the Calvi Project on the computing and visualisation of plasmas which includes INRIA, CNRS and the Universities of Nancy and Strasbourg.** *"The novelty? It's that we need to process a continuum of an enormous mass of points (of the order of several terabits) in 5 or 6 dimensions."* **The dynamic distribution of particles is described by the Vlasov equations or gyrokinetic models. The researchers are developing algorithms to simplify them, render them usable by computers, and select the points to visualise. They work with the CEA on magnetic confinement and with the Lawrence Berkeley National Laboratory in the United States on inertial confinement.** *"Our visualisation software must be operational between now and three years' time,"* **he hopes. INRIA is also working on the virtual reality tools for visualisation in 5 or 6 dimensions.**

Visualisation of electrostatic potential of a plasma in tokamak © CEA/INRIA

Jean-Pierre Chièze :
"Fantastic methods in basic research"

LASERS

Created in March 2003, the Institut Lasers et Plasmas (ILP), joining together the CEA, CNRS, the University of Bordeaux and Ecole Polytechnique, is focusing on the civilian applications of inertial confinement.

What is the mission of this institute?

JEAN-PIERRE CHIÈZE: The ILP unites French research carried out in more than twenty laboratories on the physics of high energy lasers and the dense hot plasmas which they can create. With them the scientific community can also benefit from the fabulous facilities developed for inertial fusion: the LIL (the Laser Integration Line), the ultra intense multi petawatt laser coupled to it or even the future megajoule laser. A bounty for astrophysics and more generally for the physics of extreme conditions…

What is studied in the field of controlled fusion?

The common point is implosion (by the rocket effect) of a capsule of around two millimetres in diameter, containing a deuterium/tritium mixture. The variations are in the way of bringing energy to the target – by irradiating it directly or placing it in an oven brought to several millions of degrees by laser beams (indirect attack) – and on the way of igniting the thermonuclear reactions: either with the high temperature reached eventually by compression or by 'helping' the ignition with an ultra intense laser beam.

How does modelling fit in at IPL?

It is obviously essential! It is based on fundamental and experimental studies and creates strong links between several communities. It is also a platform for interaction which prioritises the topics of fusion for energy and research in pure physics carried out in the large laser installations. ∎

Interview by I.B.

JEAN-PIERRE CHIÈZE is the Director of the Fédération de Recherche 'ILP-Recherche' and the Assistant Director of Institut Lasers et Plasmas. D.R.

A question of confinement

IN THE SUN, the ionised gases in plasma state are confined by natural gravity. On Earth there are two ways of doing this: magnetic or inertial confinement. In the first, a very hot and very dilute plasma (10^{14} ions per cm^3) is trapped for around one second in a large magnetic structure in the form of a ring (a tokamak like at the future ITER research reactor).

THE INERTIAL CONFINEMENT results from a very rapid compression of the plasma, induced by the energy from powerful lasers such as the future megajoule lasers in France or NIF (National Ignition Facility) in the United States. The very dense (10^{26} ions per cm^3) and very hot plasma only persists for some 10-11 seconds.

FOR MAGNETIC CONFINEMENT the modelling of the turbulent charged particles is based on the Vlasov equations in 'the phase space' in 6 dimensions (3 coordinates of space and 3 of speed). The ring shape of tokamak means that models in 5 dimensions are possible, called 'gyrokinetics'. However hydrodynamics in three dimensions plays an important part in the modelling of inertial confinement. ∎
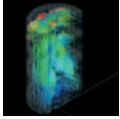
A software platform for nuclear power plants

INDUSTRY

French nuclear research and industry are currently developing a platform of software of a special kind. The goal? An integrated simulation of current or future nuclear reactors and storage or waste depots.

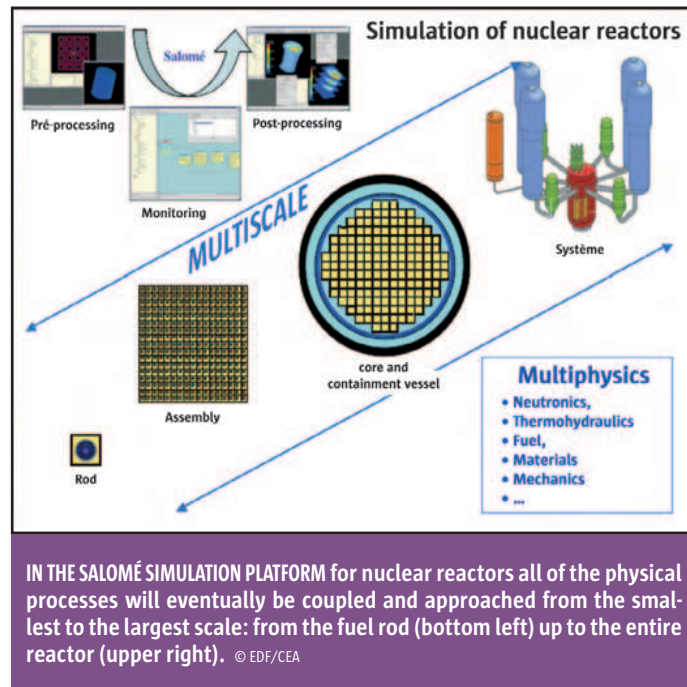It's a question of performance and safety: to design and manage a nuclear reactor is above all a business of simulation… Simulation of neutron physics, of thermohydraulics, of structural mechanics, of materials science, etc. The software should be precise, rapid, user friendly and robust. *"However,"* recalls Christian Chauliac, Head of the Simulation Project in Nuclear Energy Department at the CEA, *"up until 2000 these skills were developed without any overall coherence with programs being designed by the specialists in each domain."* But with the launch of the Simulation Project and the Salomé Platform everything is on the move and the construction of a truly integrated system is starting. A real change in mentality! As Christian Chauliac underlines, *"Before we would develop competing programmes. From now on the fundamental and applied research of the French nuclear industry (researchers, companies, safety specialists) will be shared. This integrated approach is difficult to undertake but it is unique in the world. And it will be a good example to follow."* No less than 30 programs are in the process of being integrated by Salomé! This open source software platform is being developed in the framework of an RNTL project (Réseau National des Technologies Logicielles) which unites about 20 partners: research laboratories (Irisa, Ensam), the nuclear organisations (CEA, EDF), the aeronautics (EADS) and automobile (Renault) industry and the OpenCascade company (the programme editor). Whatever the industrial application the computing needs are the same: Salomé is generic and offers services common to pre and post data processing, enables specialised physics software to be linked and offers the possibility of managing them in a unique environment with a common data exchange format. This is one of the ideas put forward as a computing platform in the context of the new System@tic competitiveness centre.

IN THE SALOMÉ SIMULATION PLATFORM for nuclear reactors all of the physical processes will eventually be coupled and approached from the smallest to the largest scale: from the fuel rod (bottom left) up to the entire reactor (upper right). © EDF/CEA

## No less than 30 programs are in the process of being integrated by Salomé!

"Let's take the example of one of the building blocks connected to Salomé: thermohydraulics modelled with the Neptune Platform," suggests Olivier Marchand, Head of the R&D Department at EDF. *"Since 2001, we have been developing this platform with the CEA, Framatome-ANP and IRSN (Institut de Radioprotection et de Sûreté Nucléaire). The aim is to model more precisely than the existing software, both on the physical and geometrical levels, the disphasic flows – making water act in the form of liquid and vapour – in components like the steam generator, the reactor core or even the entire plant."* Of the major advances, an example is the three-dimensional modelling at the smallest scale, a cubic millimetre. The aim is to zoom in on certain zones of interest within the complex structure (a fuel assembly, for example) with the rest of the flow being simulated on a rougher scale. *"This approach should be operational between now and four years' time. But we will need about ten years to have the entire Neptune Platform and its new generation of software at the scale of a component or of the complete plant,"* states Olivier Marchand. *"As in thermohydraulics,"* adds Christian Chauliac, *"each discipline (materials, neutronicss, waste storage, etc.) is tackled with a multi scale perspective. This alone allows one macroscopic approach to be coupled to another microscopic one (going up to the point of ab initio calculations) while giving the best compromise in terms of accuracy and calculation time."* It also involves linking these different codes to each other to optimise the interactions between disciplines. Applied to thermohydraulics and neutronics this multiphysical approach, for example, allows the effect of the temperature of the surrounding fluid to be combined with the computing of neutron reactions. What remains is to make these links automatic… ∎

Isabelle Bellin
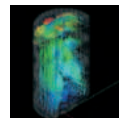
A pressing need for **computing tools**

As Jean-Pierre Chièze, explains many obstacles need to be removed before fusion phenomena by inertial confinement can be reproduced by computing.

Numerical modelling is a determinant in research on fusion plasmas by inertial confinement. We expect that it will be used for the analysis of experiments and the interpretation of measurements, of course. But also that it will be predictive. However there are many difficulties. First, whatever the chosen approach (direct or indirect attack) very different spatio-temporal scales need to be dealt with at the same time: from microscopic processes (the interaction of

laser beams with the plasma, for example) of the order of a femtosecond and of tenths of microns up to the hydrodynamic phenomena of the implosion of the capsule containing the deuterium-tritium mixture, a million times larger (several nanoseconds and several millimetres). The implosion codes therefore contain approximation models to describe the small scale phenomena.

The second difficulty is the extreme variety of physical mechanisms that come into play. In fact, we must model them then integrate the various phenomena regulating the laser-plasma interaction, the implosion of the capsule and the thermonuclear combustion. Some phenomena overlap directly anyway. The challenge is to correctly recreate

their coupling in an integrated simulation in a three dimensional space.

Certain phenomena, like energy transfer, must be described in detail which often demands the calculation of the distribution of particle speeds. The simulation of the capsule implosion requires the simultaneous resolutions of these energy transfer equations, the hydrodynamic equations and equations governing the atomic populations of the plasma; these hydrodynamic radiation codes are a hybrid physical model linking a description of fluids to a description of the kinetics of matter.

The hydrodynamic steps consist of representing, by calculations, the multiple shockwaves and the release due to the implosion of the target. It is crucial but particularly com-

plex. Especially as many types of instabilities appear which amplify the faults in the target. This can lead to it breaking or altering the ignition conditions for the thermonuclear reactions. The quantitative evaluation of these effects relies on the three dimensional hydrodynamic codes which take up the greatest calculating resources, in the order of 100 teraflops.

Finally, simulation of the ignition of reactions by an ultra intense laser source raises specific problems too related to the generation of particles and to the physics of relativist plasmas. ■

Jean-Pierre Chièze

➲ www-fusion-magnetique.cea.fr/
➲ About integrated modelisation:
www.efda-taskforce-itm.org
➲ About JET : www.jet.efda.org/
➲ About ITER : www.iter.org

Digital nuclear reactors

Numerical simulation is today used for the design and management of nuclear reactors. But the current computing power is insufficient to describe their development in the long term.

Piloting a nuclear reactor requires dexterity. The radioactive compounds enclosed in the core are continually transforming under the effect of neutron bombardment. Under the impact, the nuclei of uranium or plutonium break down to release other neutrons

– which maintain the chain reaction – and fission productions accumulate as the fuel is used up. Considering the difficulty of taking direct measurements, regulating this flow of neutrons and adjusting the power of the reactor according to energetic demands requires simulation. Simulation also allows us to spread the production of a new fuel rod over one or two years. For its 50 water pressure reactors, EDF therefore uses some of the neutronics software developed by the CEA such as those included in the SAPHYR (Système Avancé pour la Physique des Réacteurs) system of calculation codes. "*We already have good tools but the industry is*

asking us for more precise models to reduce the margins and uncertainties," explains Anne Nicolas engineer in the Department of Nuclear Studies at CEA Saclay. The aim is to reduce costs and increase the safety of plants still more. "*We are particularly researching how to reduce the uncertainty over the locally produced power in the fuel even though this is already calculated with a good accuracy of the order of several percent. The economic stakes are important: a gain of 1% in the maximum power translates into an increase of 1% of the power capacity of the reactor*," says Robert Jacqmin, Research Director in Neutronics at the Cadarache centre. Other

demands from industry: to be able to increase the useful life of plants and increase the rate of combustion within reactors. "*Today the software has assessed combustion rates of up to 45 Gigawatt days per tonne the while industry wishes to go up to 60 Gigawatt days per tonne and even higher*," explains the researcher.

In practice, the simulations rely principally on the Boltzmann equation, which describes the behaviour of neutrons, and on the measurements taken in test reactors functioning at a virtually zero power like Eole, Minerve and Masurca on the Cadarache site. With the current calculation capacity it is not reason-

able to directly describe the detailed changes in the entire reactor over several years, so models are built in two stages. First we work on the elementary components (an assembly made of several hundreds of fuel rods, for example, describing them with all the nuclear reactions in a very detailed geometric and energetic representation. By reduction of this data we then extract an average result and apply it to a 3D model of the entire reactor. *"The idea is to no longer use these simplifications, but to head towards a direct calculation of the whole of the reactor with a very fine mesh,"* explains Anne Nicolas further.

As a tool for piloting and designing plants, neutronics modelling must also adapt to the new forms of reactors currently being studied. The fuel of current French water pressure reactors is made up of a regular network of tens of thousands of rods containing piles of uranium dioxide blocks arrayed vertically. They can be represented simply by Cartesian geometry. But France is also interested in the so-called Fourth Generation Reactors: reactors at very temperature cooled by gas and with fuel particles arrayed in an random way within a matrix. *"This poses a challenge to stochastic geometry,"* explains Anne Nicolas, *"these new ideas mean we must rethink the modelling."*

The development of future fuel reactors is a major challenge for simulation in neutronics, thermohydraulics and thermomechanics. A challenge which the future Jules Horowitz experimental reactor, now planned at Cadarache, could help to meet and is expected to supply the need for knowledge on new materials and fuels for the fifty years to come. ∎

Laure Schalchli

➲ Cadarache centre website: www-cad.cea.fr
➲ « Physique nucléaire et sûreté », *Clés CEA, n° 45,* 2001 : www.cea.fr/fr/Publications /clefs45/sommaire.html

IOLS :
a generic tool box

Simulating a nuclear installation, air traffic, monetary transactions, climate risks, the aerodynamics of a plane or even a car crash… Each time the complex scientific calculations reproduce phenomena from the smallest to the largest scale arising from various physical and mathematical models. *"In fact, we can simulate the design of products or the optimisation of complex systems with generic multiphysical and multiscale modelling tools,"* **states Jean-Yves Berthou, Head of the Applied Information Science Group at EDF, one of the partners in the IOLS project (Infrastructures et Outils Logiciels pour la Simulation), the uniting theme of the world class competitiveness centre System@tic approved in July 2005.** *"It involves developing transverse software tools and testing them on applications, in the first instance relating to materials and problems in fluid/structure. Two ranges of complementary tools will be developed. First an open source along the lines of the Salomé Software Platform developed around the CEA and EDF. Second with proprietary tools like in the PLM (Product Lifecycle Management) procedure for the development of products by Dassault Systèmes".* **The IOLS Project has 24 partners (12 industrial groups, 9 laboratories and 3 SMEs).**

➲ www.systematic-paris-region.org/

Daniel Verwaerde:
"Perpetuating France's

NUCLEAR WEAPONS

France's nuclear dissuasion capability today relies on the simulation programme of the Department of Military Applications of the Commissariat à l'Énergie Atomique. An update from its first Director.

When did the simulation programme begin?

DANIEL VERWAERDE: In 1996 when France's President decided on a complete and definitive end to nuclear testing. The CEA Department of Military Applications was then given the responsibility of setting up a programme destined to perpetuate France's dissuasion capabilities in the absence of new tests. Financed by the Ministry of Defence, this programme covers 15 years running up to 2010. Remember that in 1996 France decided to significantly and unilaterally reduce the number of dissuasion methods, particularly by closing the 'Plateau d'Albion' and then its nuclear experimen-

DANIEL VERWAERDE is the Director of Nuclear Weapons within the CEA (DAM). An engineer and mathematician, he directed the Department of Applied Mathematics there in charge of Numerical Simulation and Computing for Dissuasion. He is also Professor of Numerical Analysis at Ecole Centrale de Paris. © DR

tation centre. The 'simulation' programme was thus designed to replace the system of current armaments as and when they arrive at the end of their life. And also to remain at the highest level scientifically and to be capable of guaranteeing the reliability and the safety of current and future systems.

What is it exactly?

It revolves around three areas: modelling of physical phenomena, numerical simulation, and the validation by laboratory experiments and reconstruction of past tests. First, we needed to list all of the physical phenomena involved in the functioning of a nuclear weapon to study the way in which they succeed each other and link up. Then to make an inventory of the mathematical equations likely to represent them like those which govern fluid mechanics (Navier-Stokes equations), the transport of neutrons (Boltzmann equation) or the changes in the populations of photons or ions (diffusion and transport equations). But that's

AIRIX the radiograph machine for validating calculations relating to the first step in the functioning of a weapon.
© CEA

nuclear dissuasion without tests"

not all. Once the system of equations is set up we still need to determine the right properties of matter (equations of state, effective neutronic sections, transport coefficients, etc.) in the domains relevant to the functioning of weapons. Then we must solve the system of equations in the real and changing world. And finally validate all the solutions with respect to past information and to the results of experiments.

So how do you proceed?

Modelling physicochemical phenomena requires the resolution of very large systems of equations that are coupled and non-linear. No one knows how to solve them so luckily we can turn to numerical analysis in order to transform them (by derivation) into a suite of linear systems which can be resolved with supercomputers. Overall we cut the calculation field into millions or even billions of small zones (meshes) and for each one of them the arithmeticians replace the equations that we didn't know how to solve by other approximations which are 'digestible' for the computer.

All the power of numerical analysis resides in the capacity to prove that the approximated solution from the computer is as close as possible to the exact solution of the real problem. But this will never be known. Of course we get much closer if there are more meshes. For example, to simulate the state of a system involves tens or even hundreds of millions of meshes. The associated systems of equations therefore have billions of unknowns!

This clearly requires huge computing capacities…

Yes. From 1996 when the pro-gramme started we knew that to run more predictive models, which are therefore more complex with finer meshes, we would eventually need to multiply the power of our computers and their memories by ten thousand! Tera-1 allowed us to gain a factor of 100 on the size of meshes. With the arrival of Tera-10 our computing power is going to be multiplied by 10 again (60 thousand billion operations per second). Then by 10 again by the year 2010. For the CEA each of these steps is associated with a well defined dissuasion capability. So the weapons simulator (like the simulators used for training aircraft pilots or managers of power plants) should be in use in 2010. Of course, this date does not signify the end of simulation: we must adapt our dissuasion capabilities to the changes that are bound to occur. This will certainly lead us to increase the computing power even more.

How can you be sure that the simulation results are good? In other words how do you validate your software?

First of all I would like to emphasise that the validation does not only consist of assessing the software but more fundamentally reassuring ourselves that the users are perfectly proficient and are aware of the limitations. So simulation integrates the training and homologation of users who at the CEA are known as the 'designers'. To return to the initial question we need to distinguish between validation of the parts and overall validation. The first involves a physical phenomenon or several phenomena coupled together. The second evaluates the capacity of software to reconstitute all of the measurements collected during past nuclear tests.

The validation by parts is based on experimentation. The most important experiments call on the use of two of the simulation programme's beacon facilities: the Airix radiograph machine and the megajoule laser. Installed since 1999 in Moronvilliers in Champagne, Airix is used in the validation of the calculations corresponding to the first step in the functioning of a weapon. This machine radiographs matter as it is compressed. Remember that compression is the initial condition for the functioning of a nuclear weapon. In these experiments the fission material is of course replaced by inert material with mechanical and thermal properties that are very similar (said to be 'cold experiments').

The second key instrument of the simulation programme is the megajoule laser in construction near Bordeaux (photo above). With this laser the conditions of pressure and surrounding temperature as would occur in the functioning of a weapon or even in the stars can be reconstituted in the laboratory. It will be possible using the theory of comparisons to study the physical phenomena occurring during the function of a nuclear weapon. The megajoule laser is to the weapons designer what a windtunnel is to an aircraft manufacturer. A plane never flies in a windtunnel just as an weapon will never be tested with a laser. By using 240 identical beams and concentrating their energy (7.5 kJ) on several cubic millimetres of matter for several thousandths of a second, the megajoule laser will recreate on a reduced scale the extreme conditions which prevail in the core of the stars and at the centre of the earth.

Since 2002, a prototype with four beams has been operational. Its energy is insufficient to recreate the conditions that we're interested in for the validation by parts of the software. The megajoule laser should be ready at the beginning of the next decade which will end the construction phase of the Simulation project. The first fusion experiments forecast for 2012-2013 will kick off the new era of its use. ■

Interview by Fabienne Lemarchand

© CEA
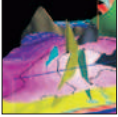
From computer viruses to flu viruses

EPIDEMICS

Is it possible to predict the way an epidemic spreads? Several teams are working on this question using quite different concepts. One of the most innovative is inspired by the behaviour of computer viruses.

It's best to be prepared for the worse. From one day to the next a new strain of the bird flu virus capable of being transmitted from one person to another could appear. And trigger a pandemic causing millions of deaths. How would the disease spread? Would the control measures – quarantine, antiviral treatments or vaccinations – be effective? To know the answers, simulations are making good headway. *"This is a very dynamic field at the moment, even though the entry of the computer on the epidemiology scene is relatively recent,"* observes Marc Barthélemy from the CEA Department of Theoretical and Applied Physics.

In the largest study done to date, Neil Ferguson's group at Imperial College, London has just modelled nothing less than the population of Thailand, that's 85 million individuals [1]. The composition of households, the average size of schools, distances between work places, and the spread of the population over the country were all taken into account. The objective was to test the effects of local control measures once a first case is detected. And the result? Everything would depend on the capacity of the virus to be transmitted from one person to another. If it behaved like the current flu virus three million antiviral treatments would be sufficient to avoid a pandemic. Since then, the same group has thrown itself into an even more ambitious project with the financial support of the American National Institute of Health – modelling the whole of the United States. Clearly this avalanche of data requires considerable computing resources. For Thailand the simulation took up the resources of 10 supercomputers each with two processors for one month; for the United States, machines with 8 or more processors will be necessary.

Such calculating power does not dissuade the critics. *"These models are so specific. How can they be extrapolated to another country?"* questions Marc Barthélemy. With others the researcher has taken a competing approach originating from the recent explosion in network physics [2] and from the observation… of computer viruses. By astonishing good luck their study will help the study of real diseases. *"We understood how these viruses spread so quickly on the web and survive there such a long time, something that was a total mystery in classical epidemiology* [3]*,"* explains the researcher. The reason? The internet network contains nodes with very high connectivity or 'hubs', which play a key role in spreading viruses. It's the same thing in real epidemics where some individuals act like 'super propagators' [4,5]. A totally opposite view to the classical simu-
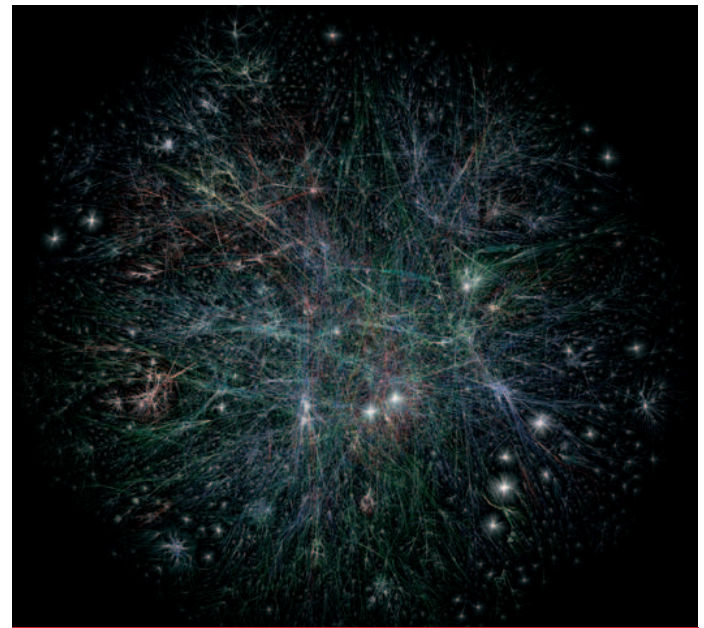
Just as high connectivity nodes play a key role in the spread of viruses on the Internet, certain individuals act like super propagators.
© THE OPTE PROJECT

lations, such as Ferguson's, in which the social network is homogeneous and individuals equivalent….

Based on these new concepts the CEA team is putting the finishing touches to a model of the spread of epidemics by airline networks. Built in collaboration with the Laboratory of Theoretical Physics at the University of Orsay and the University of Indiana in the United States, it uses data from the International Air Transport Association on the connections between the 4000 principal airports on the planet, the frequency of flights, and the number of passengers. Still at the calibration and validation stage, the model has already been used to simulate the SARS epidemic and is now being tested on flu.

The idea is to eventually enlarge it by including the urban transport networks of cities served by these airports. *"With this in mind, moving up to Tera-10 will be really important,"* judges Marc Barthélemy, *"a computation which today takes one week won't take more than an hour or two."* ■

Laure Schalchli

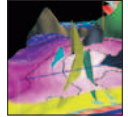[1] N. Ferguson *et al.*, *Nature, 437*, 209, 2005
[2] A. Barrat *et al.*, *Proc. Natl. Acad. Sci. USA, 101*, 3747, 2004
[3] R. Pastor-Satorras et A. Vespignani, *Internet : structure et évolution*, Belin, 2004.
[4] M. Barthelemy *et al.*, *Phys. Rev. Lett., 92*, 178701, 2004
[5] M. Barthelemy *et al.*, *Journal of Theoretical Biology, 235*, 275, 2005.

➲ www2.cnrs.fr/presse/thema/373.htm

Jean-Yves Blanc:
"Going **underground** & knowing where to look"

Petrol companies don't drill at random. They rely on increasingly precise maps of the substratum drawn up from very complex digital models.

The Compagnie Générale de Géophysique acts at the request of petrol companies. Why?

JEAN-YVES BLANC: In the first instance, it's for collecting geological and geophysical data (in particular, seismic measurements) in the regions likely to harbour oil-bearing deposits. Then, to process the data so as to create three-dimensional maps of the substratum. Obviously, the objective is to have as precise an idea as possible of the geometry and volume of subterranean or submarine reservoirs. On the basis of these maps petrol companies decide whether or not to exploit them.

How are these maps drawn up?
Roughly speaking, two lorry-loads of data arrive which we narrow down to two double sided DVDs… This reduction can be done thanks to simulation and the use of a large number of processing sequences. Some of them allow us to 'clean up' the raw data: to remove the noise (linked, for example, to industrial activities); to increase the signal-noise ratio; to take surface effects (superficial rocks don't have the acoustic elasticity of deep rocks) and the shape (like the presence of faults) into account; to eliminate the influence of marine currents, etc. Then we build a mathematical model of the substratum. Close to what we aim to represent, the

results are compared to results acquired with our algorithms. The model is progressively refined by successive iterations. We consider that it approximates reality when the results of the model and of calculations done on real data are very close.

How precise are the maps you deliver to petrol companies?
On average, it's of several metres. The goal is to improve this still more. The stakes are high for petrol companies especially when the oil-bearing layers are thin and uneven! The worse their position, the higher the risk of drilling beside them. And the higher the bill! Remember that an exploration well costs several millions of dollars and a deep offshore well even more. Errors are, however, rarer and rarer. First, because our algorithms are more and more precise. Then because the data collection drives are more concentrated than in the past.

Could your models be perfect?
Of course. The problem is firstly financial though. The data we collect are sufficient to implement more sophisticated algorithms, but the results must be available in a time span compatible with the demands of the petrol companies, several months only. To run more complex models in this limited time would require computer resources 10 to 100 times larger than today's and therefore much higher budgets…

Do you act at the production level too?
More and more so. With the advent of '4D mapping' (3-dimensional seismic maps repeated in time), the oil companies have permanent or semi permanent initiatives on site for the regularly collection of data (every week, for example). The goal is to have as precise a vision as possible of the oilfield so that production can progress. It is an essential tool in decid-

JEAN-YVES BLANC is a computing architect in the Processing and Reservoirs group of the Compagnie Générale de Géophysique. © DR

ing whether to install additional wells, to inject mud or to increase the rate of recovery by several percent…. The data processing is just like I described before. And the precision of the maps too. ■
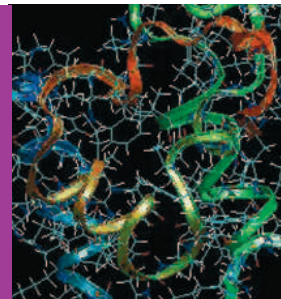
**Interview by
Fabienne Lemarchand**

➲ www.cgg.com
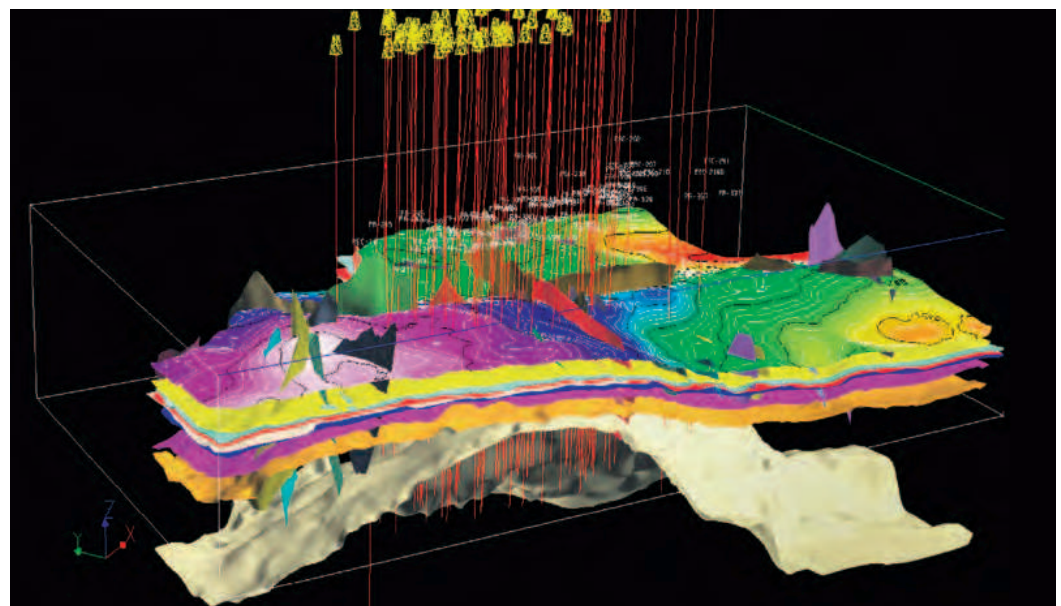
3D maps of the substratum help petrol companies drill in the right place. © CGG

Teratec

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 and IT manufacturers

- Facilitates the initiation of common R&D projects among its members

- Offer access to leading-edge supercomputer technology and data processing facilities, amongst the most powerful in the world

- Provide appropriate trainings and seminars to enhance expertise

TERATEC PARTNERS:
Bull, CEA, Communication et Systèmes, Dassault-Aviation, Distène, Ecole Centrale Paris, EDF, Ecole Normale Supérieure de Cachan, Hewlett-Packard France, Institut Français du Pétrole, Institut national des télécom d'Evry, Safran, Université de Versailles Saint-Quentin. Communauté de Communes de l'Arpajonnais, communes de Bruyères-leChâtel et dOllainville.

RESEARCH
Increase knowledge

collaboration

INDUSTRIAL COMPANIES
Increase competitiveness

IT ENTERPRISES
Design, develop, optimize
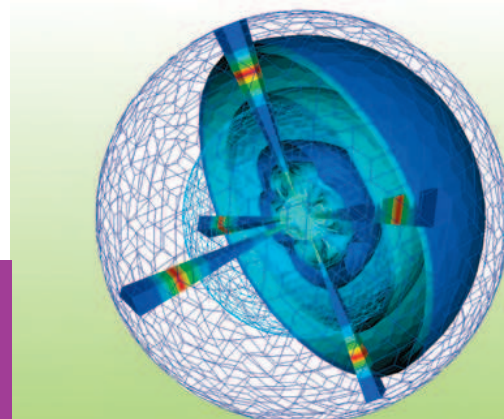
Teratec is a key component of the System@tic competitiveness cluster

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