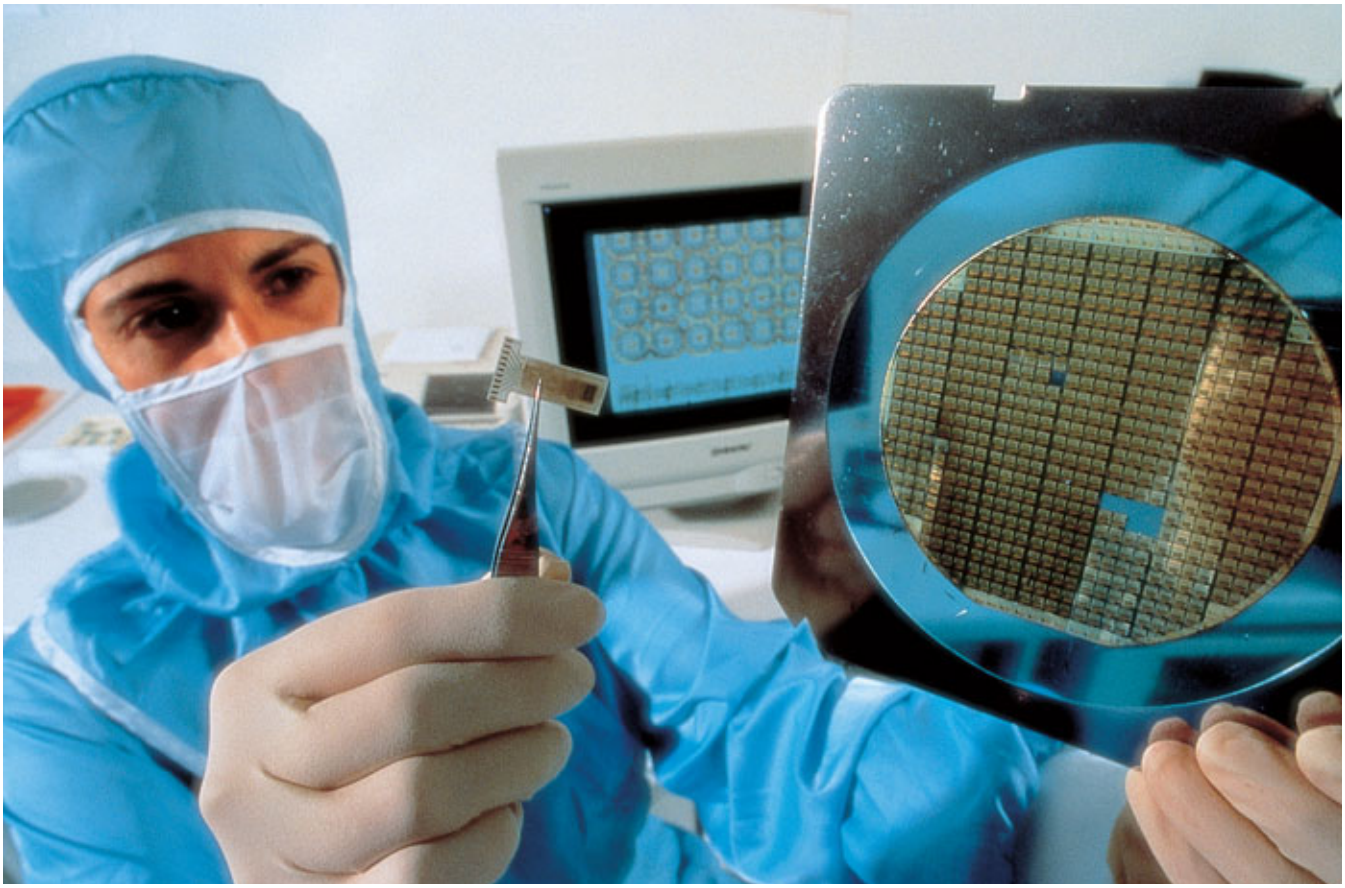
SIMULATION OF BIOLOGICAL SYSTEMS

It is now possible, with DNA arrays, to measure the expression of genes in a living organism. Reconstructing from these data, by way of modeling, the genetic network thus expressed would make it possible to simulate it and, for example, predict the behavior of a cell or the phenotype of an individual, in other words the expression of his genes. Numerical simulation can be seen to be equally crucial in the technological development of the arrays themselves.



CEA/REA/Allard

Micam® multiplexed arrays (128-spot format), on silicon wafer (right foreground) and in individual assembly (left).

Recently, **DNA**(1) arrays (see Box 1) have allowed simultaneous measurement of the **expression** of thousands of **genes** in a cell or a tissue (Figure 1). For instance, in bakers' **yeast**, it is now possible to analyze the expression of the entire **genome** (about 6,300 genes); in humans, simultaneous analysis is achieved for 20,000 genes, i.e. some two thirds of the genome. This mass of information could enable a better understanding of the architecture of biochemical or genetic networks within a cell (Figure 2) and the logic of their regulating mechanisms, which ultimately specify a cell's behavior or the production of a phenotype.(2) It is thus possible to measure the *differential* expression of a major proportion of the genes and to compare, for instance, cancerous cells and healthy cells. If it is possible to deduce the underlying gene network from these data, through a **modeling** stage, this network can be simulated and the cell's state (cancerous or healthy) may be predicted according to the degree of expression of the main regulating gene. While it is still important to study genes and **proteins**(1) individually, it is now becoming possible, and useful, to investigate the structure and dynamics of biological systems by way of global approaches. Since a genetic network is more than an assembly of genes

(1) See Box 1, in *Modeling biological macromolecules*.
(2) The set of characters that are observable in an individual. For a cell, this will be the set of manifest characters resulting from the expression of its genes (cell morphology, state of health of the cell, etc.)

What is a DNA chip?

1

A **DNA** chip, or DNA array, or biochip (or GeneChip) is a device allowing in principle to detect the presence of a strand of DNA – the molecule serving as support for the genetic information in all living beings –, by pairing this strand with its complementary, the so-called probe, fixed on the chip.
Its principle, described at the end of the 1980s, is based on the DNA's property of **hybridization**, in other words the ability of the **bases** in one strand of this DNA (thymine, adenine, cytosine and guanine) to recognize spontaneously complementary bases (thymine and adenine on the one hand, cytosine and guanine on the other) and pair together like the two parts of a zip fastener, thus forming a double helix.
The chip can thus identify a given sequence of **nucleotides**, i.e. the chaining of the bases in a DNA fragment, by bringing it into the presence of other strands whose sequence is known. The technique can be used in many applications, from diagnosis and screening of new medicines and their action sites, to detection of contaminants and pollutants, through **genomics** research (the study of **mutations**, in particular). A DNA chip allows very fast visualization of differences in expression among **genes**, even at the scale of a complete genome.
DNA chips nowadays have the ability to identify in a single operation as many as tens of thousands, or even hundreds of thousands, of DNA samples.
While the principle is simple, practical implementation involves a combination of advanced technologies in the field of microelectronics, **nucleic acid** chemistry, image analysis and bioinformatics. To fabricate the chip, the DNA fragments are amplified by **polymerization** through the polymerase chain reaction (PCR) technique and are then bound (by means of **electrostatic** interactions) on a glass, polymer, silicon or metal substrate to create hybridization sites.
During fabrication of the probe, the sample DNA is **labeled** with a **fluorophore**. Once hybridization is effected, each spot is excited by laser, the emitted fluorescence (the tell-tale sign of hybridization) being detected through a fluorescence microscope. Data analysis actually consists in analysis of an image, allowing not only location, but also quantification of the fluorescent signal emitted by each DNA fragment.
CEA, for its part, has developed, together with the CIS bio International company, the so-called Micam® (Multiplexed Integrated Chip for Advanced Microanalysis) electrochemical addressing technique,(2) in which the chip comprises a silicon substrate overlaid with a gold microelectrode at each spot. This technology, intended for the high-performance chip market, is to be marketed by the Apibio company (set up in 2001 by CEA and bioMérieux), which has exclusive rights to it.

(1) See Box 1, in *Modeling biological macromolecules*.
(2) See note 3 in the main text.

and proteins, merely investigating the architecture of connections will not be sufficient for an understanding of all its properties. Evolution of the system over time is just as important, and the components' interaction dynamics must also be examined.

Two basic principles

Understanding of a genetic network is based on two principles.
The first concerns the *genetic information flow*, which consists in specifying the relations existing from **sequence** space to *function* space. The genome contains information that may serve to construct "objects" as complex as… a human. In terms of information, the complexity of a fully developed organism is held in the complexity of its genome. But what are the codes that translate a sequence into structure and function? These codes must be represented in an understandable form if they are to be implemented for the construction of models. Biologists are therefore seeking methods making it possible to find them, when starting from the gene sequence and activity data gained by means of DNA arrays.

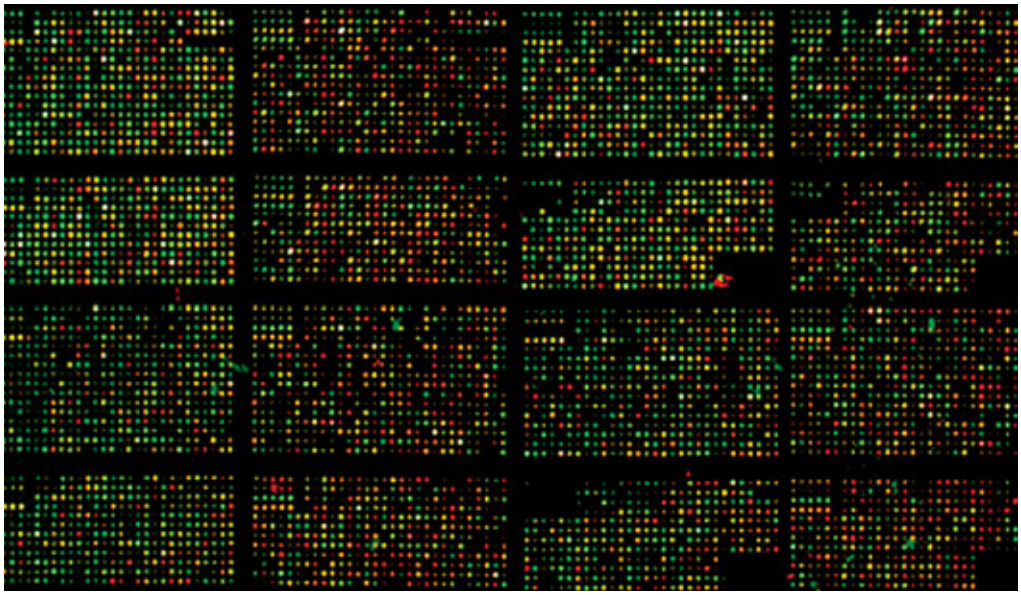


Figure 1. Expression profile for the complete genome of bakers' yeast obtained with a DNA array. Each spot corresponds to one gene. Induced genes (whose expression is activated in a given biological condition) show up in red, repressed genes (whose expression is reduced) in green, and invariant genes in yellow.

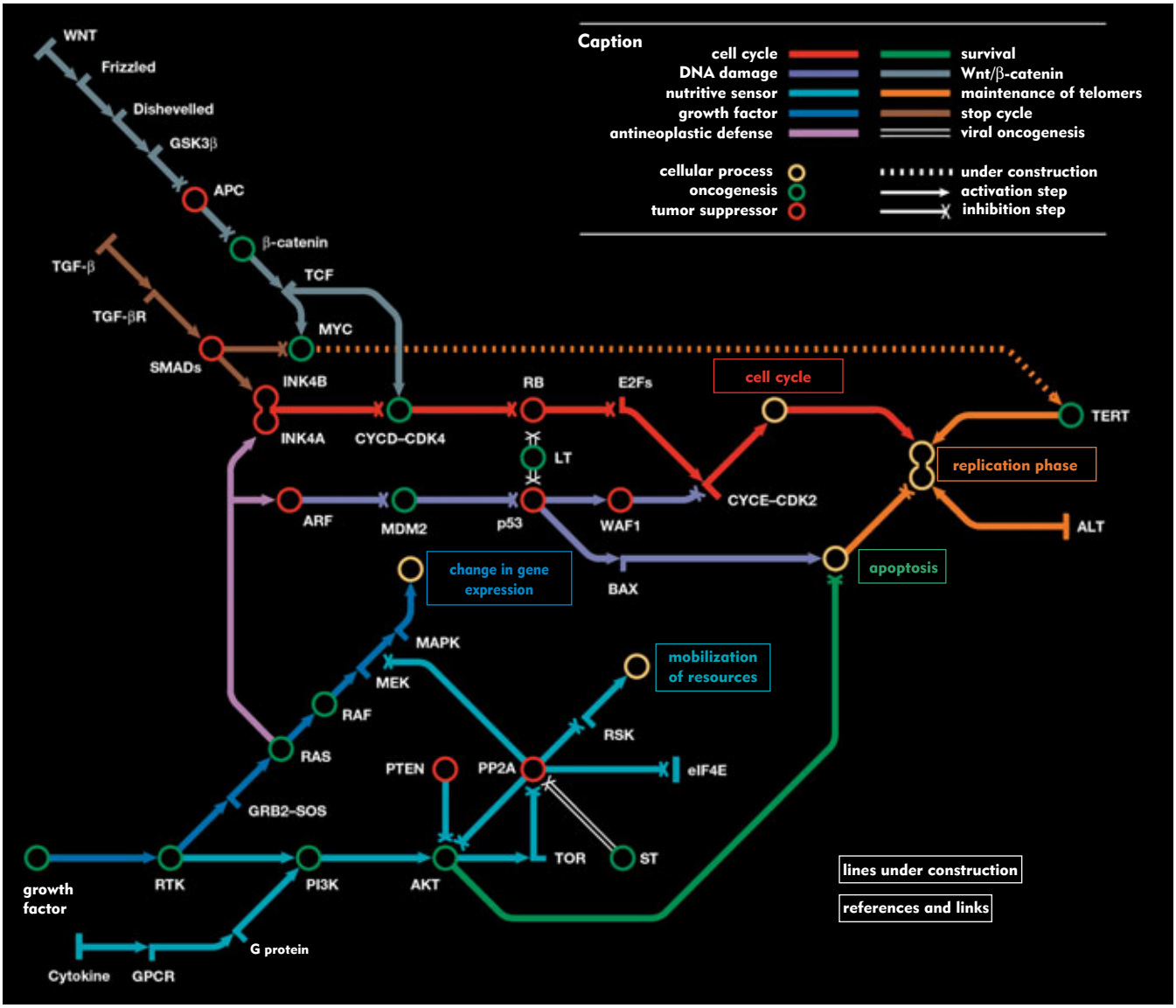


Figure 2. Still highly incomplete, this "subway map"-like diagram rehearses the entire set of genes so far known to be involved in cancerization of a cell, together with their interactions. It is probable that genomics will enable discovery of many others, whose function is as yet unknown. One of the goals, at stake in the simulation of genetic networks, is to place these genes on this network and determine the nature of their interactions with the rest of the network (map established by William C. Hahn, Robert A. Weinberg and Claudia Bentley, and reproduced with the authorization of Nature Reviews Cancer, *Vol. 2, No. 5, May 2002; © 2002 Macmillan Magazines Ltd).*

The second principle is investigation of *complex dynamic systems*. Once the architecture of a genetic network is understood, it becomes possible to investigate its dynamics, in other words the system's behavior over time under a variety of conditions. This problem is approached by measuring the variations over time of the genes' expression. The aim is to have the ability to predict the *attractor* of a genetic network, i.e. the stable phenotype of a cell, so that this network may be directed, possibly some day, towards a chosen attractor: from cancerous cell to benign cell, from aging cell to juvenile cell.

Five parameters to take on board

The dynamic analysis of a genetic network requires models to be created. The choice of model is often determined by what question the experimenter seeks to answer and the degree of abstraction he can accept. Five parameters must be taken into account.

Level of biochemical detail

Models may be highly abstract, such as Boolean networks,(3) or conversely very concrete, such as complete biochemical interaction models using kinetic parameters. The former approach enables the analysis of very large systems. The latter corresponds more closely to biochemical reality; however, being more complex, it remains restricted to the study of small systems. The requirement thus arises for methods having the capacity to handle a very large amount of data globally, while offering an acceptable level of detail, without going as far as the exact reaction.

(3) Genetic network based on Boolean logic, allowing considerable simplification of the interactions between genes. Each gene is considered as a binary variable (0 for an expressed gene, 1 otherwise) regulated by other genes according to Boolean functions.

Boolean or continuous?

Boolean models are based on the principle that gene **activation** or **repression** response is quite abrupt. Now, gene expression tends to be on a continuum rather than binary. Moreover, essential concepts for gene regulation mechanisms cannot be represented by Boolean variables. For example, retroinhibition[4] stabilizes a network by allowing control of a protein's or a gene's homeostasis,[5] and reduces sensitivity to external variations. In a Boolean circuit, retroinhibition will cause oscillation, rather than increased stability.

Deterministic or stochastic?

An implicit assumption of continuous models is that the fluctuations of a single molecule may be ignored. However, there are many examples in genetic networks showing that the presence of a single copy of messenger **RNA** (mRNA)[1] can play an essential role in certain biological processes that cannot then be modeled by means of purely **deterministic** models. Analysis of mRNA, and protein, abundance and degradation could enable identification of the genes for which **stochastic** modeling would be necessary.

(4) Inhibition of a metabolic synthesis pathway by its final product.
(5) Stabilization, in a living organism, of the various physiological constants, in terms of chemical composition, temperature and volume.

The spatial dimension

The spatial dimension can play a major role as regards intracellular compartmentalization (nucleus, cytoplasm, mitochondria, membrane), but equally for interactions between cells. Most biological processes in multicellular organisms (during development, for instance) require interactions between various types of cells. The space concept adds a considerable degree of complexity to models. Some information may be extracted from "non-spatial" models, but it will eventually be necessary to develop models that include this dimension.

Data availability

An exhaustive biological model should take into account RNA concentration, but equally protein concentration, its location, etc., since each molecular variable carries some unique information on cell functioning. The technological limitations of measurement mean obtaining this information is a complicated affair, even if constraints and redundancy in biological networks do suggest a possible way of understanding the operation of a biological system without modeling all of its parameters. Eventually, it will be important to develop innovative high-throughput measurement technology for the molecular parameters mentioned above.
Of course, there is no point in simulating genetic networks unless this allows predictions to be made as to biological processes and the evolution of diseases, and ultimately the development of effective therapies for the benefit of patients. ●

Xavier Gidrol
Évry Génopole (Essonne *département*)

Photonic modeling of DNA chips

In the design of DNA chips (or DNA arrays, as they are also called by biologists), the complexity of the sensors' optical and chemical architecture imposes a major numerical simulation phase, centered on the fluorescent labeling technique.

On **DNA** chips (Box 1), recognition of the **hybridization** (Box 1) of two complementary strands is effected by **fluorescent** labeling[1] (Box 2). This highly sensitive detection method is widely used in biology and analytical chemistry. In the context of DNA chip technological development, the reproducibility and quantity of the signals emitted by the spots is a factor to be taken on board during the design phase, and in chip readout. This concern is of paramount importance, in that analysis of the amount of fluorescence light emitted by each spot on the biochip must be related to the num-

(1) Addition of a chemical group or a radioactive atom to a molecule in order to monitor it and locate it through this tag.

Fluorescence

2

The principle of fluorescence is simple: a molecule (a **fluorophore**) absorbs a photon with a given energy and then re-emits, usually very quickly, a photon of lower energy. This energy loss is accounted for by induced thermal agitation or by alteration to the structure of the molecule.
Consequently, the color of emission light is different from that of the excitation: this enables **labeling**, by separation of these colors by means of an appropriate optical device (the simplest example being a prism).
Fluorescence detection is highly sensitive, since, thanks to advances made in recent years with detectors and image capture devices, it is possible to make out a single molecule.
Unfortunately, fluorescence is extremely sensitive to parameters such as the fluorophores' chemical environment, which can extinguish fluorescence (this phenomenon is generally known as "quenching"), or excessive exposure to excitation light, which may destroy the tags ("photobleaching" phenomenon).

ber of hybridized DNA strands, in order to determine the **expression** rate of the **genes** to be analyzed. Although an analytical **model** allows an understanding of the physics of the problem, the complexity of the optical and chemical architecture of the sensors to be designed means a major **numerical simulation** phase is necessary (see Box A, ***What is a numerical simulation?***). This derives its information from measurement campaigns carried out with specific tools, to ensure availability of reliable, sensitive products that can be made industrially.

An analytical model to understand trends

The aim is to **model**, in the simplest and most relevant fashion, the behavior of fluorescence on plane surfaces composed of a substrate (microscope slide, silicon wafer(2)) that may be overlaid with thin mineral (e.g. silica) or organic coatings. Such coatings have the purpose, in particular, of ensuring good DNA binding onto the surface of the "biochip."

The goal is to identify the parameters involved, in the first order, in the optical behavior of biochips. For that purpose, **fluorophores** (Box 1 and Box 2) are assimilated to small light sources with the property of having a particular orientation: dipoles. Light exciting the fluorescent tag induces oscillation in the electron cloud, and the direction of this oscillation is imposed by the molecule's atomic structure. Coupling between source and surface is processed analytically through a formalism based on an electromagnetic approach to light radiation: this dipole may be assimilated to an antenna emitting close to the ground, as the German physicist Arnold Sommerfeld did, some… 90 years ago!

This preliminary study enabled researchers to decouple light emission phenomena related to the biochip's optical and geometric characteristics from effects related to its chemical and biological environment. Observation of the signals emitted by biochips fabricated on glass or silicon was able to corroborate this.

Predicting signals on complex systems

The technological evolution of biochips is tending more and more markedly towards complex microsystems. Ever-increasing demand for miniaturization, and integration of such functions as spot addressing by CMOS(3) transistors (Micam® technology) is pushing this through. Understanding and predicting fluorescence signals on such components becomes a task significantly more difficult than for conventional chips. Investigation of fluorescence characteristics in the vicinity of many materials on structured surfaces, owing to the complexity of the problem, calls for a numerical approach. To meet this requirement, CEA researchers have produced a simulation tool allowing the photonic balance sheet of a biochip to be drawn up: they resolve, with no approximations, the electromagnetism equations describing the interaction of the fluorophore, still assimilated to a dipole, for any environment.

(2) Thin slice of material (usually silicon) on which integrated circuits and other microelectronic components and devices are fabricated before they are cut out.
(3) On MICAM®-type biochips, to each contact stud is associated a CMOS transistor (high-density integrated circuit using both N- and P-type transistors, allowing them to use practically no energy when not switching). It is then possible to activate them, and, through a phenomenon similar to an electrodeposition reaction, to position specifically on these studs the required DNA strands.

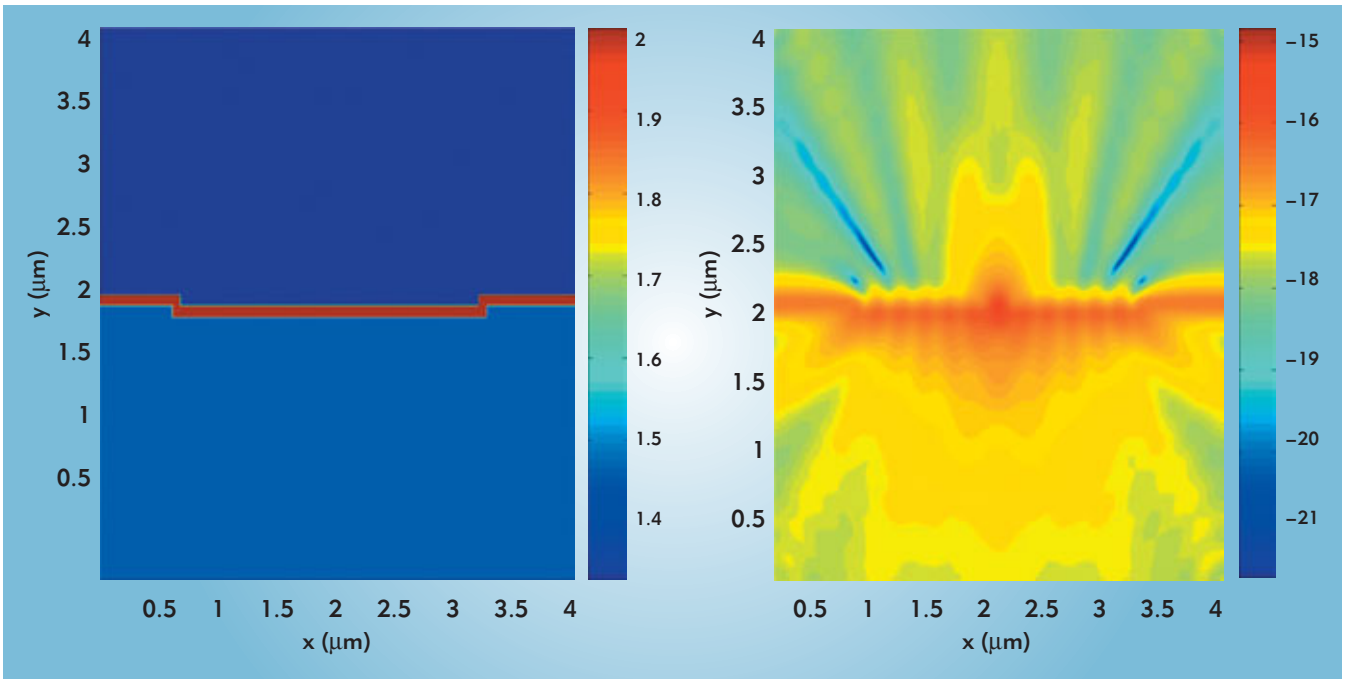


Figure 1. Analysis of fluorescence emitted on a Lightscan® track. The figure at left shows the intensity of fluorescent light (logarithmic scale) in the structure with the geometry represented at right.

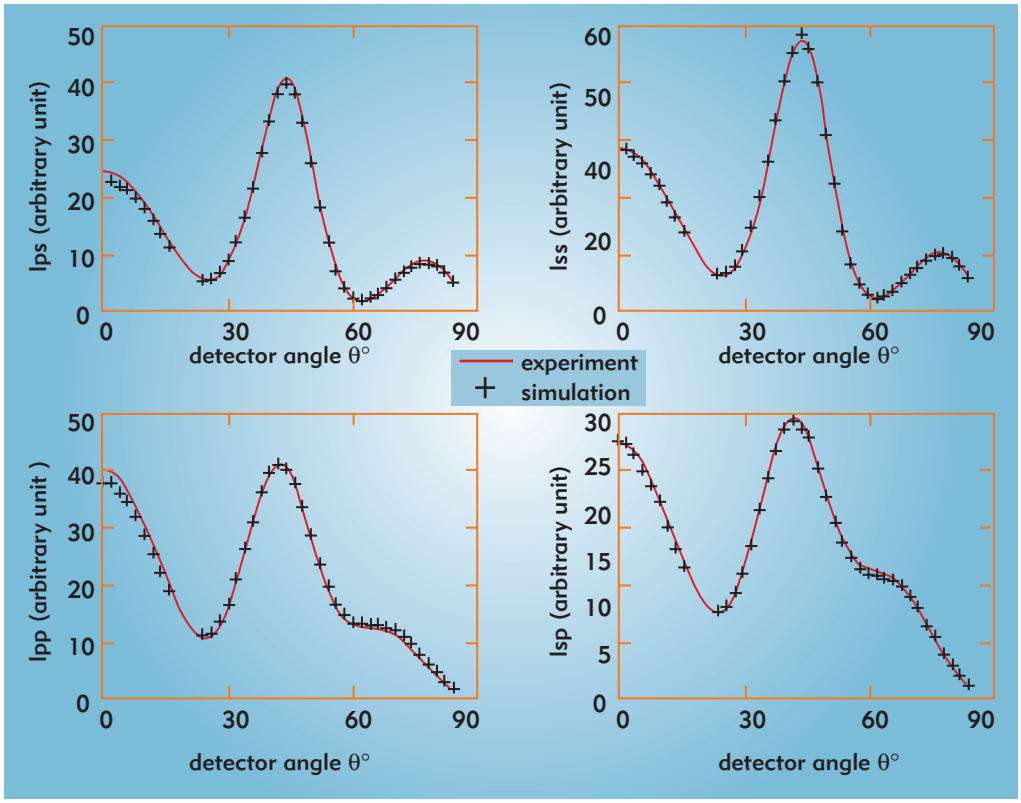


Figure 2. Comparison between theory and experiment for the intensity of light emitted by DNA biochips (fabricated on a silicon substrate coated with a thin layer of silica a few hundred nanometers thick) as a function of measurement angle. The curves correspond to an analysis of fluorescent light polarization along two orthogonal directions ("s" and "p") after illuminating the biochip under light exhibiting rectilinear polarization along these two directions.

Application to a concrete technological case

One major application of this tool was modeling the behavior of fluorescence on Lightscan® biochips jointly developed by the bioMérieux company and CEA's Electronics and Information Technology Laboratory (Leti). This system boils down to producing a biochip scanner inspired… from a compact disk player, a solution which allows cost reductions by a factor of 5–10, compared to a conventional device. Like the compact disk, the biochips must present tracks, so as to allow very exact positioning of the pickup head as it is traversed. Scientists have thus analyzed the behavior of fluorescence on the tracks of these biochips, which thus feature a structuring of the glass (a few tens of nanometers deep over a width of several microns) on which a thin layer, a few tens of nanometers thick, is deposited (Figure 1). Photonic simulation shows that light emitted by the fluorophore is trapped in the thin layer in the form of guided waves (as in optical fibers) and partly released at the level of the structures.

An indispensable experimental approach

Modeling alone cannot guarantee mastery of technological processes. It is indispensable to characterize their outcomes with a dedicated metrological tool. Development of a goniofluorimeter(4) (Figure 2) allowed, in this case, validation of the simulation work. Critical examination of the radiation patterns measured under fluorescence, taking into account the polarization effects of light,(5) led to a reconsideration of the analytical approach and improved modeling. Researchers now have an optical characterization platform enabling them to consider a quantification approach and to suggest fluorescence standards.

There still remains for them to take on board energy-coupling phenomena between the fluorophore and incident radiation, and environmental phenomena such as "photobleaching" or "quenching" (Box 2).

(4) A goniofluorimeter (neologism) serves to determine fluorescent light intensity for a given direction in space. The device developed at CEA–Leti also allows analysis of the polarization state of fluorescent light.

(5) The theory of polarization allows a description of the orientation (in terms of amplitude and variation over time) of the electrical field vector as it describes a light vibration. In the case of fluorescence, a study of light polarization provides a means of analyzing phenomena such as the molecule's change in structure between excitation and re-emission of light.

Stéphane Gétin
Technological Research Division
CEA Grenoble Center

What is a numerical simulation? A

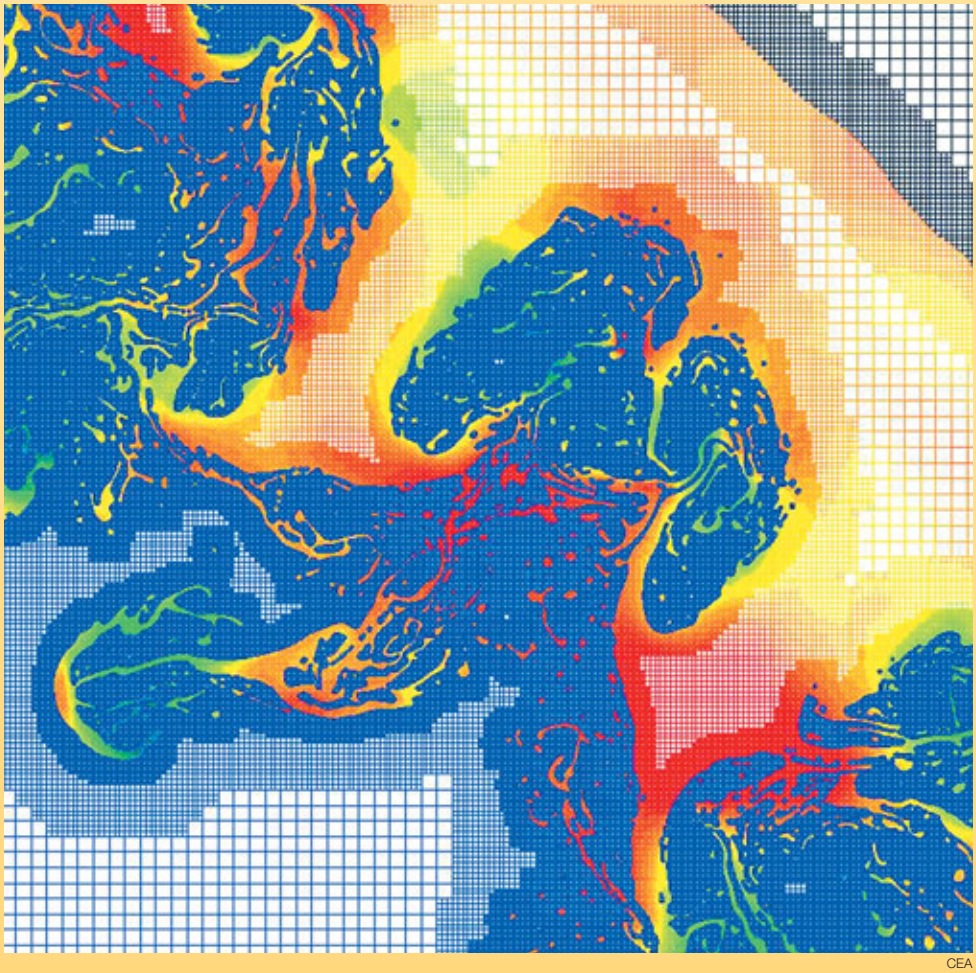
Numerical simulation consists in reproducing, through computation, a system's operation, described at a prior stage by an ensemble of **models**. It relies on specific mathematical and computational methods. The main stages involved in carrying out an investigation by means of numerical simulation are practices common to many sectors of research and industry, in particular nuclear engineering, aerospace or automotive.

At every point of the "object" considered, a number of physical quantities (velocity, temperature…) describe the state and evolution of the system being investigated. These are not independent, being linked and governed by **equations**, generally **partial differential** equations. These equations are the expression in mathematical terms of the physical laws modeling the object's behavior. Simulating the latter's state is to determine – at every point, ideally – the numerical values for its parameters. As there is an infinite number of points, and thus an infinite number of values to be calculated, this goal is unattainable (except in some very special cases, where the initial equations may be solved by analytical formulae). A natural approximation hence consists in considering only a finite number of points. The parameter values to be computed are thus finite in number, and the operations required become manageable, thanks to the computer. The actual number of points processed will depend, of course, on computational power: the greater the number, the better the object's description will ultimately be. The basis of parameter computation, as of numerical simulation, is thus the reduction of the infinite to the finite: **discretization**.

How exactly does one operate, starting from the model's mathematical equations? Two methods are very commonly used, being representative, respectively, of **deterministic computation** methods, resolving the equations governing the processes investigated after discretization of the variables, and methods of **statistical** or **probabilistic calculus**.

The principle of the former, known as the **finite-volume method**, dates from before the time of computer utilization. Each of the object's points is simply assimilated to a small elementary volume (a cube, for instance), hence the *finite-volume* tag. Plasma is thus considered as a set or lattice of contiguous volumes, which, by analogy to the makeup of netting, will be referred to as a **mesh**. The parameters for the object's state are now defined in each mesh cell. For each one of these, by reformulating the model's mathematical equations in terms of volume averages, it will then be possible to build up *algebraic relations* between the parameters for one cell and those of its neighbors. In total, there will be as many relations as there are unknown parameters, and it will be up to the computer to resolve the *system* of relations obtained. For that purpose, it will be necessary to turn to the techniques of **numerical analysis**, and to program specific **algorithms**.

The rising power of computers has allowed an increasing fineness of discretization, making it possible to go from a few tens of cells in the 1960s to several tens of thousands in the 1980s, through to millions in the 1990s, and up to some ten billion cells nowadays (Tera machine at CEA's Military Applications Division), a figure that should increase tenfold by the end of the decade.



CEA

Example of an image from a 2D simulation of instabilities, carried out with CEA's Tera supercomputer. Computation involved adaptive meshing, featuring finer resolution in the areas where processes are at their most complex.

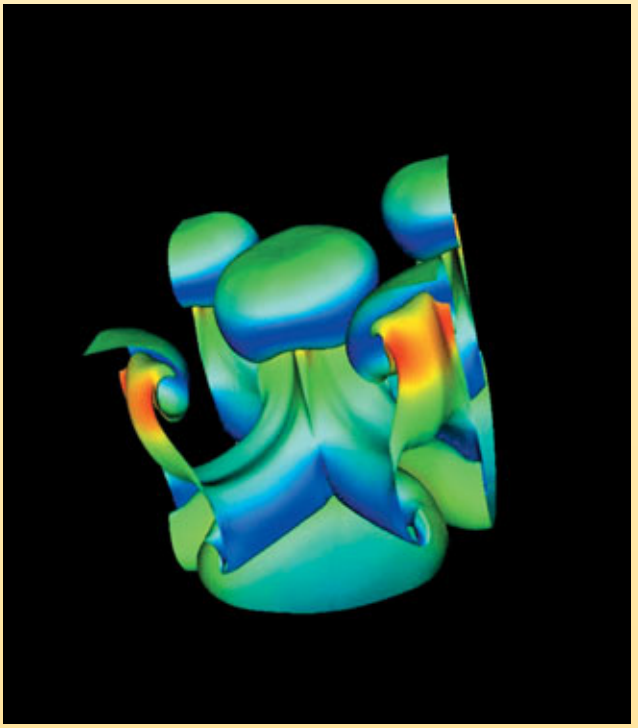
A refinement of meshing, **adaptive remeshing**, consists in adjusting cell size according to conditions, for example by making them smaller and more densely packed at the interfaces between two environments, where physical processes are most complex, or where variations are greatest.

The finite-volume method can be applied to highly diverse physical and mathematical situations. It allows any shape of mesh cell (cube, hexahedron, tetrahedron…), and the mesh may be altered in the course of computation, according to geometric or physical criteria. Finally, it is easy to implement in the context of **parallel computers** (see Box B, ***Computational resources for high-performance numerical computation***), as the mesh may be subjected to partitioning for the purposes of computation on this type of machine (example: Figure B).

Also included in this same group are the **finite-difference method**, a special case of the finite-volume method where cell walls are orthogonal, and the **finite-element method**, where a variety of cell types may be juxtaposed.

The second major method, the so-called **Monte Carlo** method, is particularly suited to the simulation of *particle transport*, for example of neutrons or photons in a **plasma** (see *Simulations in particle physics*). This kind of transport is in fact characterized by a succession of stages, where each particle may be subject to a variety of events (diffusion, absorption, emission…) that are possible *a priori*. Elementary probabilities for each of these events are known individually, for each particle.

It is then a natural move to assimilate a point in the plasma to a particle. A set of particles, finite in number, will form a representative sample of the infinity of particles in the plasma, as for a statistical survey. From one stage to the next, the sample's evolution will be determined by random draws (hence the method's name). The effectiveness of the method, implemented in Los Alamos as early as the 1940s, is of course dependent on the statistical quality of the random draws. There are, for just this purpose, *random-number* methods available, well suited to computer processing.



CEA

3D simulation carried out with the Tera supercomputer, set up at the end of 2001 at CEA's DAM-Île de France Center, at Bruyères-le-Châtel (Essonne* département*).

Finite-volume and Monte Carlo methods have been, and still are, the occasion for many mathematical investigations. These studies are devoted, in particular, to narrowing down these methods' convergence, i.e. the manner in which approximation precision varies with cell or particle number. This issue arises naturally, when confronting results from numerical simulation to experimental findings.

How does a numerical simulation proceed?

Reference is often made to *numerical experiments*, to emphasize the analogy between performing a numerical simulation and carrying out a physical experiment.

In short, the latter makes use of an experimental setup, configured in accordance with initial conditions (for temperature, pressure…) and control parameters (duration of the experiment, of measurements…). In the course of the experiment, the setup yields measurement points, which are recorded. These records are then analyzed and interpreted.

In a numerical simulation, the experimental setup consists in an ensemble of computational programs, run on computers. The **computation codes**, or **software** programs, are the expression, via numerical algorithms, of the mathematical formulations of the physical models being investigated. Prior to computation, and subsequent to it, *environment software* programs manage a number of complex operations for the preparation of computations and analysis of the results.

The initial data for the simulation will comprise, first of all, the delineation of the computation domain – on the basis of an approximate representation of the geometric shapes (produced by means of drafting and CAD [computer-assisted design] software) –, followed by discretization of this computation domain over a mesh, as well as the values for the physical parameters over that mesh, and the control parameters to ensure proper running of the programs… All these data (produced and managed by the environment software programs) will be taken up and verified by the codes. The actual results from the computations, i.e. the numerical values for the physical parameters, will be saved on the fly. In fact, a specific protocol will structure the computer-generated information, to form it into a numerical database.

A complete protocol organizes the electronic exchange of required information (dimensions, in particular) in accordance with predefined formats: modeler,(1) mesher,(2) mesh partitioner, com-

(1) The modeler is a tool enabling the generation and manipulation of points, curves and surfaces, for the purposes, for example, of mesh generation.
(2) The geometric shapes of a mesh are described by sets of points connected by curves and surfaces (Bézier curves and surfaces, for instance), representing its boundaries.

putation codes, visualization and analysis software programs. *Sensitivity* studies regarding the results (sensitivity to meshes and models) form part of the numerical "experiments."
On completion of computation (numerical resolution of the equations describing the physical processes occurring in each cell), analysis of the results by specialists will rely on use of the numerical database. This will involve a number of stages: selective extraction of data (according to the physical parameter of interest) and visualization, and data extraction and transfer for the purposes of computing and visualizing diagnostics.
This parallel between performing a computation case for a numerical experiment and carrying out a physical experiment does not end there: the numerical results will be compared to the experimental findings. This comparative analysis, carried out on the basis of standardized quantitative criteria, will make demands on both the experience and skill of engineers, physicists, and mathematicians. Its will result in further improvements to physical models and simulation software programs.

Bruno Scheurer
Military Applications Division
CEA DAM-Ile de France Center

Frederic Ducros and **Ulrich Bieder**
Nuclear Energy Division
CEA Grenoble Center

The example of a thermalhydraulics computation

Implementation of a numerical simulation protocol may be illustrated by the work carried out by the team developing the **thermalhydraulics** computation software Trio U. This work was carried out in the context of a study conducted in collaboration with the French Radiological Protection and Nuclear Safety Institute (IRSN: Institut de radioprotection et de sûreté nucléaire). The aim was to obtain very accurate data to provide engineers with wall heat-stress values for the components of a pressurized-water reactor in case of a major accident involving turbulent natural circulation of hot gases. This investigation requires simultaneous modeling of large-scale "system" effects and of small-scale **turbulent** processes (see Box F, ***Modeling and simulation of turbulent flows***).
This begins with specification of the overall computation model (Figure A), followed by production of the CAD model and corresponding mesh with commercial software programs (Figure B). Meshes of over five million cells require use of powerful graphics stations. In this example, the mesh for a steam generator (Figures C and D) has been partitioned to parcel out computation over eight processors on one of CEA's parallel computers: each color stands for a zone assigned to a specific processor. The computations, whose boundary conditions are provided by way of a "system" computation (Icare–Cathare), yield results which it is up to the specialists to interpret. In this case, visualization on graphics stations of the instantaneous values of the velocity field show the impact of a hot plume on the steam generator's tube-plate (section of the velocity field, at left on Figure E), and instantaneous temperature in the water box (at right).

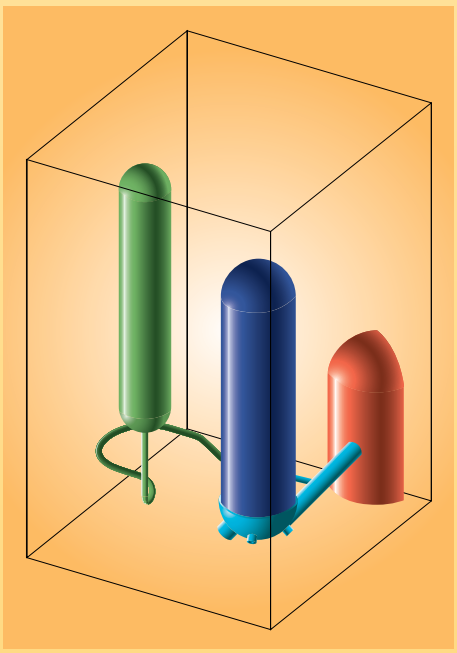


Figure A. Overall computation domain, including part of the reactor vessel (shown in red), the outlet pipe (hot leg, in light blue), steam generator (dark blue), and pressurizer (green).

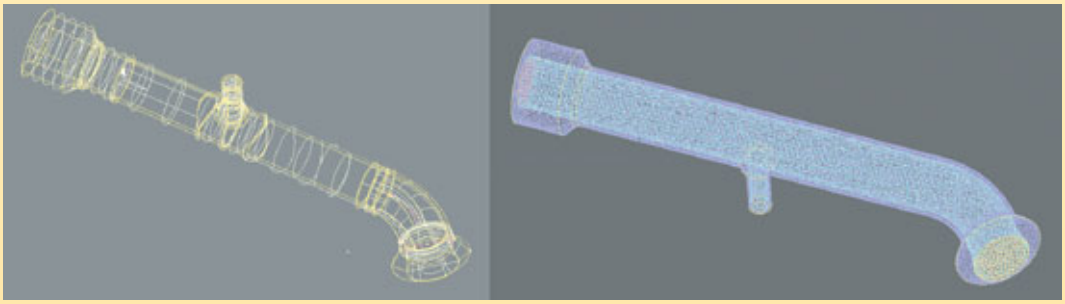
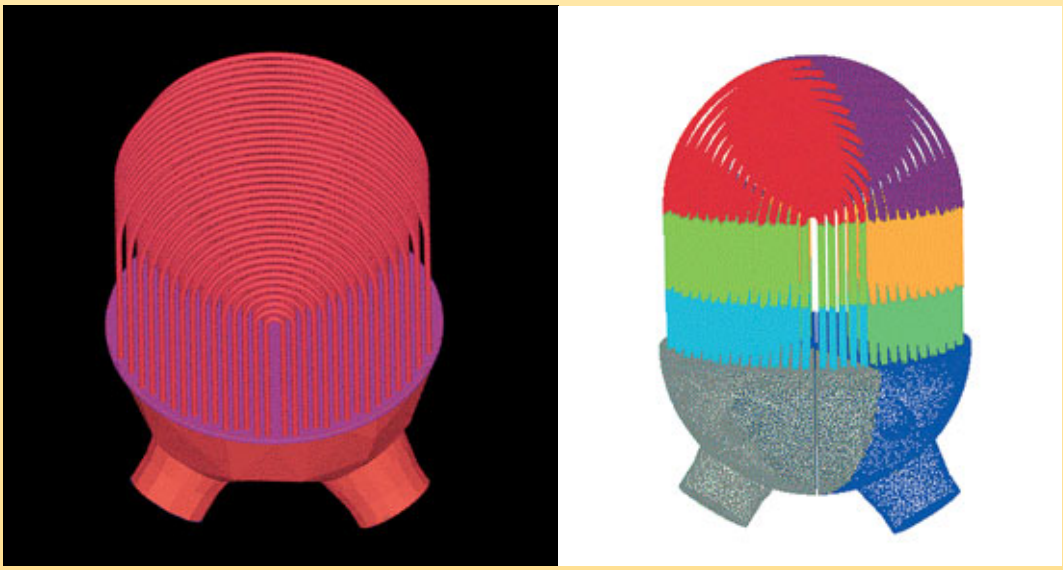


Figure B. CAD model of the hot leg of the reactor vessel outlet (left) and unstructured mesh for it (right).



Figures C and D.

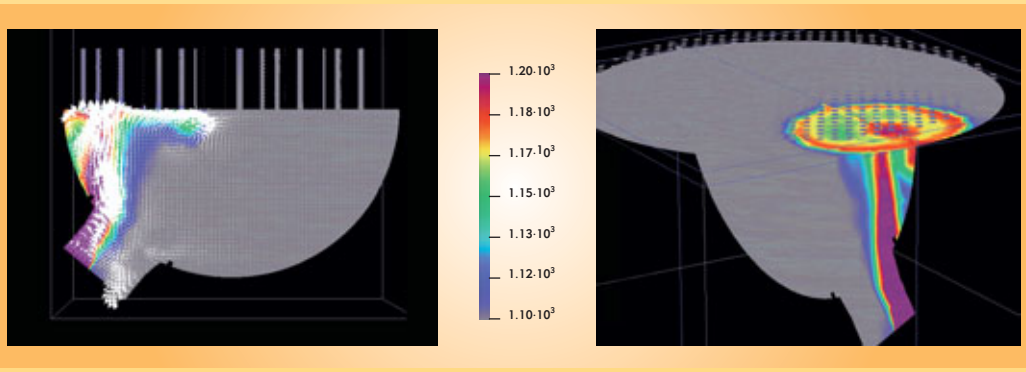


Figure E.

Computational resources for high-performance numerical simulation

B

Carrying out more accurate **numerical simulations** requires the use of more complex physical and numerical **models** applied to more detailed descriptions of the simulated objects (see Box A, ***What is a numerical simulation?***). All this requires advances in the area of simulation software but also a considerable increase in the capacity of the computer systems on which the software runs.

Scalar and vector processors

The key element of the computer is the processor, which is the basic unit that executes a program to carry out a computation. There are two main types of processors, **scalar processors** and **vector processors**. The former type carries out operations on elementary (scalar) numbers, for instance the addition of two numbers. The second type carries out operations on arrays of numbers (vectors), for example adding elementwise the numbers belonging to two sets of 500 elements. For this reason, they are particularly well suited to numerical simulation: when executing an operation of this type, a vector processor can operate at a rate close to its maximum (peak) performance. The same operation with a scalar processor requires many independent operations (operating one vector element at a time) executed at a rate well below its peak rate. The main advantage of scalar processors is their price: these are general-purpose microprocessors whose design and production costs can be written-down across broad markets.

Strengths and constraints of parallelism

Recent computers allow high performances partly by using a higher operating frequency, partly by trying to carry out several operations simultaneously: this is a first level of **parallelism**. The speeding up in frequency is bounded by developments in microelectronics technology, whereas interdependency between the instructions to be carried out by the processor limits the amount of parallelism that is possible. Simultaneous use of several processors is a second level of parallelism allowing better performance, provided programs able to take advantage of this are available. Whereas parallelism at processor level is automatic, parallelism *between processors* in a parallel computer must be taken into account by the programmer, who has to split his program into independent parts and make provisions for the necessary communication between them. Often, this is done by partitioning the domain on which the computation is done. Each processor simulates the behavior of one domain and regular communications between processors ensure consistency for the overall computation. To achieve an efficient parallel program, a balanced share of the workload must be ensured among the individual processors and efforts must be made to limit communications costs.

The various architectures

A variety of equipment types are used for numerical simulation. From their desktop computer where they prepare computations and analyze the results, users access shared computation, storage and visualization resources far more powerful than their own. All of these machines are connected by networks, enabling information to circulate between them at rates compatible with the volume of data produced, which can be as much as 1 **terabyte** (1 TB = 10^{12} bytes) of data for one single simulation.

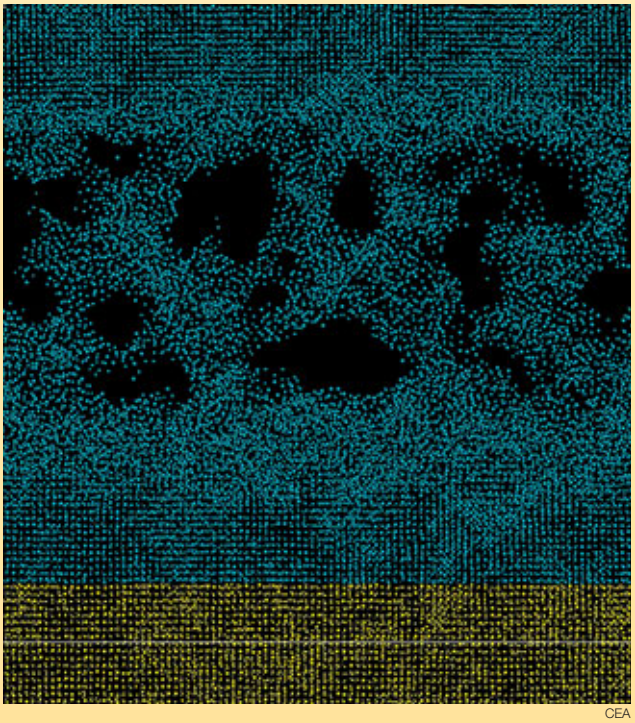
The most powerful computers are generally referred to as **supercomputers**. They currently attain capabilities counted in **teraflops** (1 Tflops = 10^{12} floating-point operations per second).

Currently, there are three main types of supercomputers: vector supercomputers, clusters of mini-computers with shared memory, and clusters of PCs (standard home computers). The choice between these architectures largely depends on the intended applications and uses. Vector supercomputers have very-high-performance processors but it is difficult to increase their computing performance by adding processors. PC clusters are inexpensive but poorly suited to environments where many users perform numerous large-scale computations (in terms of memory and input/output).

It is mainly for these reasons that CEA's Military Applications Division (DAM) has choosen for its Simulation Program (see *The Simulation Program: weapons assurance without nuclear testing*) architectures of the shared-memory mini-computer cluster type, also known as **clusters of SMPs** (symmetric multiprocessing). Such a system uses as a basic building block a mini-computer featuring several microprocessors sharing a common memory (see Figure). As these mini-computers are in widespread use in a variety of fields, ranging from banks to web servers through design offices, they offer an excellent performance/price ratio. These basic "blocks" (also known as *nodes*) are connected by a high-per-

Installed at CEA (DAM-Ile de France Center) in December 2001, the TERA machine designed by Compaq (now HP) has for its basic element a mini-computer with 4 x 1-GHz processors sharing 4 GB of memory and giving a total performance of 8 Gflops. These basic elements are interconnected through a fast network designed by Quadrics Ltd. A synchronization operation across all 2,560 processors is completed in under 25 microseconds. The overall file system offers 50 terabytes of storage space for input/output with an aggregate bandwidth of 7.5 GB/s.

Computational resources for high-performance numerical simulation (cont'd)



CEA

Parallel computers are well suited to numerical methods based on meshing (see Box A, What is a numerical simulation?) but equally to processing ab-initio calculations such as this molecular-dynamics simulation of impact damage to two copper plates moving at 1 km/s (see Simulation of materials). The system under consideration includes 100,000 atoms of copper representing a square-section (0.02 µm square) parallelogram of normal density. The atoms interact in accordance with an embedded atom potential over approximately 4–6 picoseconds. The calculation, performed on 18 processors of the Tera supercomputer at Bruyères-le-Châtel using the CEA-developed Stamp software, accounted for some ten minutes of "user" time (calculation carried out by B. Magne). Tests involving up to 64 million atoms have been carried out, requiring 256 processors over some one hundred hours.

formance network: the cumulated power of several hundreds of these "blocks" can reach several Tflops. One then speaks of a **massively parallel computer**.

Such power can be made available for one single parallel application using all the supercomputer's resources, but also for many independent applications, whether parallel or not, each using part of the resources.

While the characteristic emphasized to describe a supercomputer is usually its computational power, the input/output aspect should not be ignored. These machines, capable of running large-scale simulations, must have storage systems with suitable capacities and performance. In clusters of SMPs, each mini-computer has a local disk space. However, it is not advisable to use this space for the user files because it would require the user to move explicitly his data between each distinct stage of his calculation. For this reason, it is important to have disk space accessible by all of the mini-computers making up the supercomputer. This space generally consists in sets of disk drives connected to nodes whose main function is to manage them. Just as for computation, parallelism of input/output allows high performance to be obtained. For such purposes, parallel overall file systems must be implemented, enabling rapid and unrestricted access to the shared disk space.

While they offer considerable computational power, clusters of SMPs nevertheless pose a number of challenges. Among the most important, in addition to programming simulation software capable of using efficiently a large number of processors, is the development of operating systems and associated software tools compatible with such configurations, and fault-tolerant.

François Robin
Military Applications Division
CEA, DAM–Ile de France Center

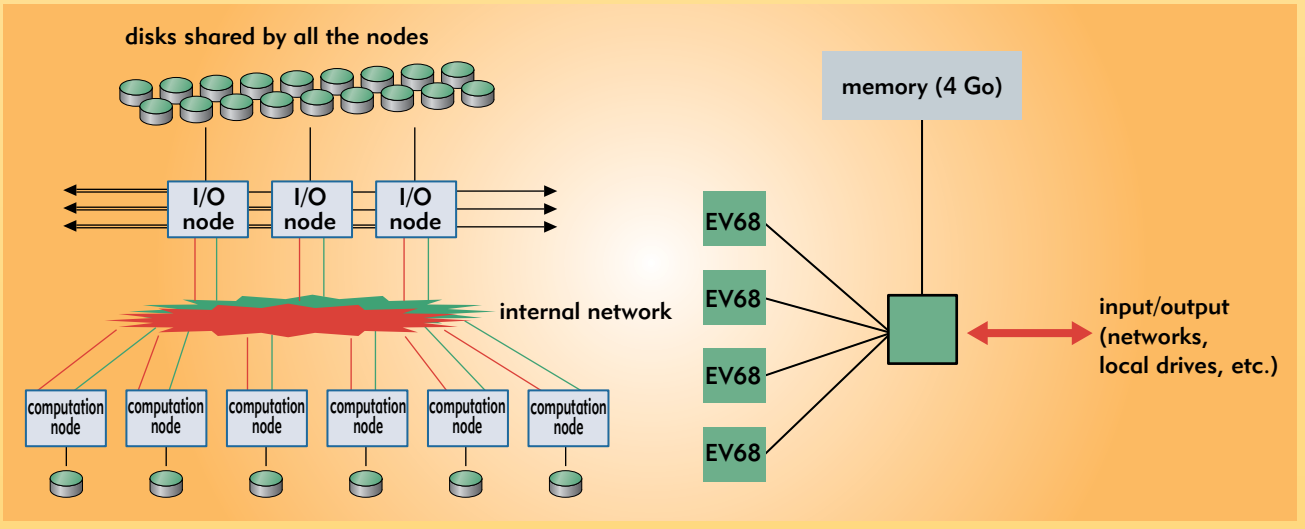


Figure. Architecture of an "SMP-cluster" type machine. At left, the general architecture (I/O = input/output), on the right, that of a node with four Alpha EV68 processors, clocked at 1 GHz.

Modeling and simulation of turbulent flows F

Turbulence, or disturbance in so-called turbulent flow, develops in most of the flows that condition our immediate environment (rivers, ocean, atmosphere). It also turns out to be one, if not the, dimensioning parameter in a large number of industrial flows (related to energy generation or conversion, aerodynamics, etc.). Thus, it is not surprising that a drive is being launched to achieve prediction for the process – albeit in approximate fashion as yet – especially when it combines with complicating processes (stratification, combustion, presence of several phases, etc.). This is because, paraxodically, even though it is possible to predict the turbulent nature of a flow and even, from a theoretical standpoint, to highlight certain common – and apparently universal – characteristics of turbulent flows,(1) their prediction, in specific cases, remains tricky. Indeed, it must take into account the considerable range of space and time scales(2) involved in any flow of this type.

Researchers, however, are not without resources, nowadays, when approaching this problem. First, the equations governing the evolution of turbulent flows over space and time (Navier–Stokes equations(3)) are known. Their complete solution, in highly favorable cases, has led to predictive descriptions. However, systematic use of this method of resolution comes up against two major difficulties: on the one hand, it would require complete, simultaneous knowledge of all variables attached to the flow, and of the forced-flow conditions imposed on it,(4) and, on the other hand, it would mobilize computational resources that will remain unrealistic for decades yet.

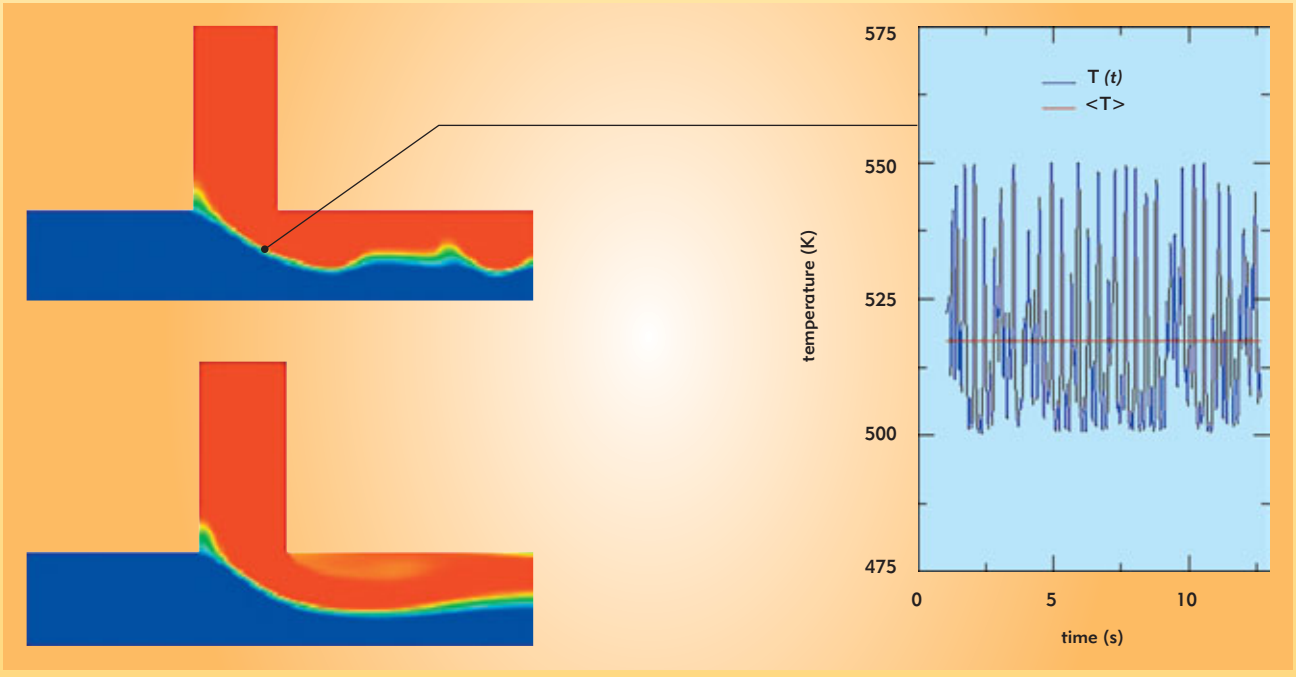


Figure. Instantaneous (top) and averaged (bottom) temperature field in a mixing situation. The curve shows the history of temperature at one point: fluctuating instantaneous value in blue and mean in red (according to Alexandre Chatelain, doctoral dissertation) (DEN/DTP/SMTH/LDTA).

The sole option, based on the fluctuating character of the flow due to turbulent agitation, must thus be to define and use average values. One of the most widely adopted approaches consists in looking at the problem from a statistical angle. The mean overall values for velocity, pressure, temperature… whose distribution characterizes the turbulent flow, are defined as the principal variables of the flow one then seeks to qualify relative to those mean values. This leads to a decomposition of the motion (the so-called Reynolds decomposition) into mean and fluctuating fields, the latter being the measure of the instantaneous local difference between each actual quantity and its mean (Figure). These fluctuations represent the turbulence and cover a major part of the Kolmogorov spectrum.(1)

This operation considerably lowers the number of degrees of liberty of the problem, making it amenable to computational treatment. It does also involve many difficulties: first, it should be noted that, precisely due to the non-linearity of the equations of motion, any average process leads to new, unknown terms that must be estimated. By closing the door on complete, deterministic description of the phenomenon, we open one to modeling, i.e. to the representation of the effects of turbulence on mean variables.

Many advances have been made since the early models (Prandtl, 1925). Modeling schemas have moved unabated towards greater complexity, grounded on the generally verified fact that any new extension allows the previously gained properties to be preserved. It should also be noted that, even if many new developments are emphasizing anew the need to treat flows by respecting their *non-stationary* character, the most popular modeling techniques were developed in the context of *stationary* flows, for which, consequently, only a representation of the flow's temporal mean can be achieved: in the final mathematical model, the effects of turbulence thus stem wholly from the modeling process.

It is equally remarkable that, despite extensive work, no modeling has yet been capable of accounting for all of the processes influencing turbulence or influenced by it (transition, non-stationarity, stratification, compression, etc.). Which, for the time being, would seem to preclude statistical modeling from entertaining any ambitions of universality.

Despite these limitations, most of the common statistical modeling techniques are now available in commercial codes and industrial tools. One cannot claim that they enable predictive computations in every situation. They are of varying accuracy, yielding useful results for the engineer in controlled, favorable situations (prediction of drag to an accuracy of 5–10%, sometimes better, for some profiles), but sometimes inaccurate in situations that subsequently turn out to lie outside the model's domain of validity. Any controlled use of modeling is based, therefore, on a qualification specific to the type of flow to be processed. Alternative modeling techniques, meeting the requirement for greater accuracy across broader ranges of space and time scales, and therefore based on a "mean" operator of a different nature, are currently being developed and represent new ways forward.

The landscape of turbulence modeling today is highly complex, and the unification of viewpoints and of the various modeling concepts remains a challenge. The tempting goal of modeling with universal validity thus remains out of order. Actual implementation proceeds, in most cases, from compromises, guided as a rule by the engineer's know-how.

(1) One may mention the spectral distribution of turbulent kinetic energy known as the "Kolmogorov spectrum," which illustrates very simply the hierarchy of scales, from large, energy-carrying scales to ever smaller, less energetic scales.
(2) This range results from the non-linearities of the equations of motion, giving rise to a broad range of spatial and temporal scales. This range is an increasing function of the Reynolds number, Re, which is a measure of the inertial force to viscous force ratio.
(3) The hypothesis that complete resolution of the Navier–Stokes equations allows simulation of turbulence is generally accepted to be true, at any rate for the range of shock-free flows.
(4) This is a problem governed by initial and boundary conditions.

Frédéric Ducros
Nuclear Energy Division
CEA Grenoble Center

Biological macromolecules

Cells are the fundamental entities in all living organisms. Apart from water, biological **macromolecules** are the major constituents of the cell, where they perform a multiplicity of functions. A biological macromolecule comprises sub-units of low molecular weight, added one to the other to form a long, chain-shaped **polymer**. Usually, each chain is formed of only one family of sub-units and the precise sequence of sub-units is essential to the function of the macromolecule. There are four major categories of macromolecules.

Proteins are probably the most important macromolecules, since they play a predominant role in most biological processes. For instance, **enzymes** are proteins that **catalyze** the majority of chemical reactions in the cell. Other classes of protein have more of a structural role or are involved in signaling,(1) regulation of **metabolism** or the immune system. Proteins are **amino acid** polymers – about twenty different types of amino acids are commonly found – and a protein may comprise several chains, each containing a few hundred amino acids. Proteins are often associated with other molecules, which assist in their biological tasks. The three-dimensional structure of proteins is very complex but critical to their function.

Nucleic acids – deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) – are **nucleotide** polymers. **DNA**, in its double-strand form (two nucleotide chains arranged in a double helix), is the genetic material which, amongst other things, codes instructions for the amino-acid sequences of all the proteins synthesized by the cell. **RNA**, usually in a single-strand form (one nucleotide chain), is essential for protein synthesis.

Lipids, fundamental constituents of cell membranes, also play an important part in metabolism and as energy reserves. The lipids include a number of classes, such as phospholipids, triglycerides and steroids.

Polysaccharides are polymers of simple sugars, such as fructose and glucose. They play a structural role, particularly in plants (cellulose is a polysaccharide), are involved in molecular recognition and can serve as energy reserves.

(1) Transmission of signals allowing cells to communicate amongst themselves.