HET

European High Performance Computing Initiative

The Scientific Case

for a

European Super Computing Infrastructure

Petascale Computing in Europe

Table of Contents

1 EXECUTIVE SUMMARY

This scientific case for high-end scientific computing in Europe has been produced between August 2005 and April 2006 by an international scientific panel, put together by cooperation between Finland, France, Germany, Italy, the Netherlands, Spain and the United Kingdom. The origin of this endeavour was the recognition of the strategic role of High Performance Computing (HPC) at the leading edge, so-called "Leadership-Class Supercomputing", for European Science and Economy. The conviction that isolated European countries would not be able to provide their researchers and engineers with resources competitive on the world stage lies at the heart of the HPCEUR initiative.

In focusing on the requirements for the successful exploitation of Leadership-class Supercomputing – the provision of the order of 1 PetaFlop/s peak performance around 2009 – the panel emphasises that the exploitation of such a resource demands an associated computational infrastructure. We envisage *one or more general purpose, Pflop-class machine operating at the apex of a wide pyramid of computational resources that embrace national, regional and thematic centres, accessible through a data network or grid, with appropriate user support organizations dedicated to ensuring maximum exploitation throughout the entire European community*. Europe has been very successful in establishing various elements of this resource and performance pyramid, notably through the GEANT2 network and its high throughput optical interconnections, and with the DEISA supercomputer grid. A compelling scientific leadership position for Europe's Computational Science (CS) and Scientific-Based Engineering Science (SBES) demands such a strong and sustained commitment to high-end computing.

Computational Sciences and High-End Computation: Alongside their own development as fully fledged disciplines, computational science (CS) and scientific-based engineering science (SBES) certainly have evolved into bringing new methodologies and ideas into mainstream science and engineering ([PITAC-CS] and [SBES]). The thematic panels of the HPCEUR initiative have considered the relevance and importance of CS and SBES in explicit terms. Computational science, fuelled by top-level supercomputers, has established itself as the third pillar of scientific methodology alongside theory and experiment. This approach has the potential to influence profoundly both science and engineering through its ability to deal with the complexity of phenomena. A large part of the associated model hierarchy is based on first principles rather than on effective theories or empirical approaches. It often is the only known method to bridge gaps between theory and experiment.

Capitalising on the potential of CS and SBES brings several urgent challenges to the European scientific and engineering communities. Making adequate expertise available to the scientific teams is a requirement for which several practical suggestions are made by our panel. The procurement of Leadership-class supercomputers is clearly essential for the European scientific community to maintain and enhance its world leadership in several domains.

The International Competition: The importance of Leadership-Class Supercomputing for scientific and economic leadership has been stressed by several reports in the USA and Germany. More importantly, a strong commitment to maintaining a competitive advantage has been shown by the explicit plans put in place in the USA to regain world leadership after the Earth Simulator was installed in Japan. Current plans are extremely ambitious, and will necessitate advanced technological developments to deliver machines in the range of multiple Petaflop/s by 2011. Our panel is convinced that the competitiveness of European science and industry will be jeopardised if sufficiently capable computers are not made available, together with the associated infrastructure necessary to maximise their exploitation; the panel has listed multiple areas at risk. Examples include the capabilities to exploit climatic data in the GIEC international negotiations, to construct ITER and fully exploit the results of the project, to expedite the design of competitive aircrafts and helicopters, to advance biophysics and systems biology to the understanding of complete cells, and to understand the results of the experiments at LHC and FAIR, etc.

Capability Computing: The requirements for scientific computing can be broadly divided into the categories of capacity and capability. Capacity computing implies high throughput of a large number of program runs of small or medium size, on different data sets. This style of computing can be dealt

with in a cost effective manner, by providing adequate computing capacity at the regional or national level, accessed by upcoming grid-based computing methods. **Capability computing** is technologically far more challenging, and is characterised by simulations using a large number of processors in a cooperative mode. Often requiring large amounts of memory and the ability to process extensive data sets, such simulations are dependent on a high-bandwidth, low-latency communication fabric that connects the individual processors. These capabilities imply a single parallel computer that no loosely coupled distributed computing infrastructures can provide. This panel is convinced that the establishment of Europe-wide supercomputer capability is critical for this class of problem.

The Need for a European commitment in Capability Computing: Providing scientists and engineers with access to capability computers of leadership-class must be recognised as an essential strategic priority in Europe. Such resources are likely to remain extremely expensive and require significant expertise to procure, deploy and utilise efficiently; some fields even require research for specialized and optimized hardware. The panel stresses that these resources should be reserved for the most exigent computational tasks of high potential value. This would require putting in place an appropriate process to screen proposals, and to run the resources as permanent research infrastructure. It is clear that the computational resource pyramid must remain persistent and compelling at all levels, including national centres, access and data grids. The active involvement of the European Community along with appropriate member states is critical in establishing a world-leading supercomputer infrastructure in the European e-Science ecosystem. Europe must foster excellence and cooperation in order to compete effectively with similar projects in the USA and Japan and to gain the full benefits of capability computing for science, engineering and industry in the European Research Area. Such resources should act as enablers for research programs, including the FP7 evolution of several FP6 projects.

1.1 The Main Scientific Objectives

This report has been prepared by five thematic panels. Capturing all of the key scientific statements from these panels is clearly beyond the scope of this Executive Summary. We choose instead to present the key scientific objectives and challenges in tabular form to impress on the reader the sheer scope of accomplishments promised through provision of a sustainable top-level infrastructure. Table 1 is organised to identify the key challenges from five distinct areas. In Weather, Climatology and Earth Sciences, we focus on Climate Change, Oceanography and Marine Forecasting, Meteorology, Hydrology and Air Quality and Earth Sciences. In Astrophysics, High-Energy Physics, and Plasma Physics a compelling and recurring theme is that of theory and modelling providing fresh insight and hence contributing to the success of major experimental facilities – from future space experiments such as the European Planck Surveyor satellite to those at large European centers like CERN, DESY and ITER. The key challenges from Materials Science, Chemistry and Nanoscience highlight the crucial role that computer simulations now play in almost all aspects of the study of materials: not only traditional materials science, physics and chemistry, but also in nanoscience, surface science, electrical engineering, earth sciences, biology and drug design. The demands of environmental constraints – cleaner catalysis-based chemical processes, materials able to withstand extreme stress and environments, nano technologies etc. drive many of these developments. Our consideration of research in the Life Sciences identifies the challenges and scientific objectives in systems biology – our top priority – in chromotine and large-scale protein dynamics, protein association and aggregration, supramolecular systems and in Medicine. Finally, the key objectives and challenges in Engineering present compelling exemplars of the sheer breath and impact of simulation. These range from innovation in technology and design (complete helicopter simulation for next generation rotorcraft, “green” aircraft & the virtual power plant), to enhanced understanding and modelling of physical phenomena in e.g. optimising gas turbines (aero-engines or power generation) or internal combustion engines (in terms of costs, stability, higher combustion efficiency, reduced fuel consumption, near zero pollutant emissions and low noise). The scientific and societal impacts are widespread, from disaster management e.g., forest fires, to improvements in medical care (biomedical flows).

Table 1. The Challenges and Outcomes in Science and Engineering to be Addressed through Petascale HPC Provision.

Area	Application	Science Challenges & Potential Outcomes
WEATHER, CLIMATOLOGY AND EARTH SCIENCES	Climate change	Quantify uncertainties on the degree of warming and the likely impacts by increasing the capability and complexity of 'whole earth system' models that represent in ever-increasing realism and detail the scenarios for our future climate.
	Oceanography and Marine Forecasting	Build the most efficient modelling and prediction systems to study, understand and predict ocean properties and variations at all scales, and develop economically relevant applications to inform policy and develop services for government and industry.
	Meteorology, Hydrology and Air Quality	Predict weather and flood events with high socio-economic and environmental impact within a few days. Understand and predict the quality of air at the earth's surface; development of advanced real-time forecasting systems for allowing early enough warning and practical mitigation in the case of pollution crisis.
	Earth Sciences	Challenges span a wide range of disciplines and have significant scientific and social implications, such as the mitigation of seismic hazards, treaty verification for nuclear weapons, and increased discovery of economically recoverable petroleum resources and monitoring of waste disposal. Increased computing capability will make it increasingly possible to address the issues of resolution, complexity, duration, confidence and certainty, and to resolve explicitly phenomena that were previously parameterized, and will lead to operational applications in other European centres, national centres and in industry.
ASTROPHYSICS, HEP AND PLASMA PHYSICS	Astrophysics	Deal with systems and structures which span a large range of different length and time scales; almost always non-linear coupled systems of ordinary and partial differential equations have to be integrated, in 3 spatial dimensions and explicitly in time, with rather complex material functions as input. Grand challenges range from the formation of stars and planets to questions concerning the origin and the evolution of the Universe as a whole. Evaluate the huge amount of data expected from future space experiments such as the European Planck Surveyor satellite.
	Elementary Particle Physics	Quantum field theories like QCD (quantum chromodynamics) are the topic of intense theoretical and experimental research by a large and truly international community involving large European centers like CERN and DESY. This research not only promise to yield a much deeper understanding of the standard model of elementary particles and the forces between them, as well as nuclear forces, but is also expected to discover hints for a yet unknown physics beyond the standard model.
	Plasma physics	The science and technology challenge raised by the construction of the magnetic confinement thermonuclear fusion reactor ITER calls for a major theory and modelling activity. Both the success of the experiment and its safety rely on such simulators. The quest to realize thermonuclear fusion by magnetically confining a high temperature plasma poses some of the computationally most challenging problems of nonlinear physics.
MATERIALS SCIENCE, CHEMISTRY AND NANOSCIENCE	Understanding Complex Materials	The determination of electronic and transport properties central to many devices in the electronic industry and hence progress the understanding of technologically relevant materials. Simulations of nucleation, growth, self-assembly and polymerization central to the design and performance of many diverse materials e.g., rubbers, paints, fuels, detergents, functional organic materials, cosmetics and food. Multiscale descriptions of the mechanical properties of materials to determine the relation between process, conditions of use and composition e.g., in nuclear energy production. Such simulations are central to the prediction of the lifetime of high performance materials in energy technology, such as high-efficiency gas-turbines
	Understanding Complex Chemistry	Catalysis is a major challenge in the chemistry of complex materials, with many applications in industrial chemistry. The knowledge of atmospheric chemistry is crucial for environmental prediction and protection (clean air). Improving the knowledge of chemical processing (from soft chemistry including polymers to the atomistic description of combustion) would improve the durability of chemicals. Supra molecular assemblies open new possibilities for the extraction of heavy elements from spent nuclear fuels. In biochemistry, a vast number of reactions taking place in the human body (for example) are not understood in any detail. A key step in the development of the clean fuels of the future requires the realistic treatment of supported catalytic nanoparticles.
	Nanoscience	The advance of faster information processing or the development of new generations of processors requires the shrinking of devices, which leads inevitably towards nanoelectronics. Moreover, many new devices, such as nanomotors can be envisioned, which will require simulation of mechanical properties at the nanolevel. Composite high performance materials in the fields e.g. adhesion and coatings will require an atomistic based description of nanorheology, nanofluidics and nanotribology. As an example the description of the complex magnetic and mechano-optical properties of nanodevices components is only feasible only on systems in the Petaflop/s range.

LIFE SCIENCES	Systems Biology	The use of increasingly sophisticated models to represent the entire behaviour of cells, tissues, and organs, or to evaluate degradation routes predicting the final excretion product of any drug in any organism. To position Europe in the next 4 years to host the first "in silico" cell.
	Chromatine Dynamics	The organization of DNA in nucleosomes largely modifies the accessibility of transcription factors recognition sites playing then a key role in the regulation of gene function. The understanding of nucleosome dynamics, positioning, phasing, formation and disruption or modifications induced by chemical modifications, or by changes in the environment will be crucial to understand the mechanism of gene regulation mediated by chromatine modelling..
	Large Scale Protein Dynamics	The study of large conformational changes in proteins. Major challenges appear in the simulation of protein missfolding, unfolding and refolding (a key element for the understanding of prion-originated pathologies).
	Protein association and aggregation	One of the greatest challenges is the simulation of crowded "not in the cell" protein environments. To be able to represent "in silico" the formation of the different protein complexes associated with a signalling pathway opens the door to a better understanding of cellular function and to the generation of new drugs able to interfere in protein-protein interactions.
	Supramolecular Systems	The correct representation of protein machines is still out of range of European groups using current simulation protocols and computers. The challenge will be to analyze systematically how several of these machines work e.g., ribosome, topoisomerases, polymerases.
	Medicine	Genome sequencing, massive genotyping studies are providing massive volumes of information e.g. the simulation of the determinants triggering the development of multigenic-based diseases and the prediction of secondary effects related to bad metabolism of drugs in certain segments of population, or to the interaction of drugs with macromolecules others than their original targets.
ENGINEERING	Complete Helicopter Simulation	The European helicopter industry has a strong tradition of innovation in technology and design. Computational Fluid Dynamics (CFD) based simulations of aerodynamics, aeroacoustics and coupling with dynamics of rotorcraft already play a central role and will have to be improved further in the design loop.
	Biomedical Flows	Biomedical fluid mechanics can improve healthcare in many areas, with intensive research efforts in the field of the human circulatory system, the artificial heart or heart valve prostheses, the respiratory system with nose flow and the upper and lower airways, and the human balance system. Although experiments have significantly improved the understanding in the field, numerous questions, the answers of which need a high resolution of the flow field, of the surrounding tissue, or of their interactions, require a detailed numerical analysis of the biomedical problem.
	Gas Turbines & Internal Combustion Engines	Scientific challenges in gas turbines or piston engines are numerous. First, a large range of physical scales should be considered from fast chemical reaction characteristics (reaction zone thicknesses of about tens of millimetres, 10^{-6} s), pressure wave propagation (sound speed) up to burner scales (tens of centimetres, 10^{-2} s resident times) or system scales (metres for gas turbines).
	Forest Fires	The development of reliable numerical tools able to model and predict fire evolution is critically important in terms of safety and protection ("numerical fire simulator"), fire fighting and could help in real time disaster management. The social impact is very important and is concerned with land, buildings, human and animal life, agriculture, tourism and the economy.
	Green Aircraft	ACARE 2020 provides the politically agreed targets for an acceptable maximum impact of air traffic on people and the environment, while at the same time allowing the constantly increasing amount of air travel. The goals deal with a considerable reduction of exhaust gas and noise. Air traffic will increase by a factor of 3, accidents are expected to go down by 80%. Passenger expense should drop (50%) and flights become largely weather independent. The "Green Aircraft" is the answer of the airframe as well as engine manufacturing industry. However, it is only by a far more productive high quality numerical simulation and optimisation capability that such a challenging development will be possible. It will be indispensable to be able to compute the real aircraft in operation, including all relevant multi-disciplinary interaction.
	Virtual Power Plant	Safe production of high quality and cost effective energy is one of the major concerns of Utilities. Several challenges must be faced, amongst which are extending the lifespan of power plants to 60 years, guaranteeing the optimum fuel use and better managing waste. These challenges demand access to Petascale machines to perform advanced simulations along with a new generation of codes and simulation platforms.

Key statements from our thematic panels: The summary statements emphasise the essential role of a sustainable top-level infrastructure.

*"In the fields of **astrophysics, high-energy physics and plasma physics** supercomputers are required to support research where real experiments are not possible, too time-consuming or too expensive and where computing systems of lower performance levels are not adequate. They constitute*

an indispensable tool for top research. In the past few years scientific breakthroughs have been achieved by the use of supercomputers which would not have been possible on computers of lower performance classes"

*"In **Climate research, Air quality and Meteorology**, Europe has been able to establish a scientific community that has played a leadership role and has effectively contributed to international assessments. These have provided the scientific basis for national and international agreements. The European community is already well organised through European projects and collaborative research. The accuracy and details of earth system research and predictions of climate change have been severely limited by the computing resources available to the community."*

*"For the community in **Material science, Chemistry and Nano science**, scientific, technological and societal challenges require the acquisition at the European level of capability computers in the 1-10 Petaflop/s range. This would enable simulations of complex systems over the long space and time scales required for quantitatively reliable results and for improved connectivity between experiment and theory. A clear benefit of these developments is the enhanced industrial exploitation potential in a variety of fields relevant to energy production, to electronic and optical communication, and for ceramics for sensors, plastics and complex fluids in the food industry, or hydrogen storage materials. In addition, there are important societal applications that will lead to improvements to cleaner air, sustainable technology, decontamination, nanotechnology and medicine."*

*"After Genomic projects the emphasis in **Life Sciences** will be the exploitation of the massive collections of genomic data and its translation into biologically relevant information. This will increase dramatically the needs for computational resources. Genomic data must be processed to find complex multigenic relationships and gene regulation networks which trigger the most prevalent pathologies like cancer, diabetes or even aging. The meta-mechanisms of gene regulation from chromatine remodelling to interference-RNAs need to be understood at the atomic level, which would requite massive atomistic simulations. Genetic information on proteins must be transformed into protein structures, which should be then quantitatively analyzed by means of very large (micro to millisecond) molecular dynamics trajectories to decipher their mechanism of action. Special emphasis will be given to key molecular machines, like ribosomes or polymerases, which control crucial cellular processes. The aggregation and association of proteins (and nucleic acids) in labile complexes, impossible to detect experimentally, must be then determined to understand signal transduction pathways. Finally, massive dockings (10^7-10^8 compounds against 10^5 simultaneous flexible targets) will be performed to help in the development of new drugs. The improvement in the understanding of pathologies and aging arising from the combination of genomic data and enhanced simulation and modelling tools will have a tremendous impact in our society, changing our current view of biology, pharmacy and medicine".*

*"The impact of computer simulations in **Engineering** allows the development of highly optimised designs, the investigation of hazards too dangerous to test and reduced development costs. In parallel new scientific investigations are developing understanding in areas such as turbulence and flow control necessary for future engineering concepts. A clear trend is towards complete system modelling, as higher fidelity information is required and automatic optimisation of the whole system, including interference between sub-components, is required. Such simulations encompass the interaction of different aspects of a system, requiring modelling from multiple disciplines and phenomena at radically different scales. The engineering simulations of today and the future lead to requirements for supercomputing resources which are almost unlimited. The demand for technological solutions to some of the great problems of the modern world, such as cheap, safe and available transport at low environmental cost or low cost power generation with low environmental impact, implies an emphasis on large scale computations to research new concepts and the underpinning science".*

2 INTRODUCTION

Over the last 30 years, Computational Science has established itself as the "third pillar" of scientific enquiry alongside theory and experiment. The discipline of Computational Science enhances scientific investigation by enabling practitioners to build and test models of complex phenomena, yielding new information, innovation and fresh insight into the research process that is otherwise unobtainable. This report shows the essential contributions expected from a variety of fields of interest, with major advances in scientific research, plus associated societal impact and ensuing industrial innovation.

For these expectations to be realized, however, it is essential that a pyramid of computational resources topped with world class "Capability HPC Centres" be provided. An appropriate infrastructure, in the form of a high speed network or "grid", has to be added to manage these resources.

The rapid pace of improvement in the areas of electronic and telecommunication hardware makes computers available cheaply for numerous applications. However, the requirements for leading edge scientific applications are rapidly outgrowing the technological progress as characterised by Moore's Law. Such applications now demand very long computation times on extremely large machines built with hundreds or thousands of processors. Some of these need so rapid data exchange that the speed of light becomes the limiting barrier, together with physical constraints of packaging, cooling and power consumption. This has made current leading edge supercomputers very different from other large systems – they present considerable challenges at all levels from design to procurement and operations. Several countries, including Japan, the USA and China are deploying major initiatives to remain competitive across the entire domain of supercomputing.

2.1 The Need for a Strategy

The procurement and deployment of key resources and infrastructures for scientific computation often demands a strategic approach. Over the years, long-term plans have been defined and executed by Japan and the USA. Major strategic revisions have occurred to address a variety of concerns: Should high-end equipment be accessible only within defence related laboratories or more generally? Should high-end computers be derived from market driven components (so called COTS)? Should a coordinated plan be put in place for Computational Sciences and Computational Infrastructure? Several panels have prepared and published key reports addressing these issues. [G-WR], [HECRTF].

As a result, Japan and the USA are pushing the performance and the application of supercomputers very aggressively, recognizing that such resources are key to national success in both science and the economy. These efforts encompass the procurement of very large and effective supercomputers, investments in the computer and software industries, and the development of supercomputing in academia and through independent software vendors. Their determination is also very clear in their announcement of future procurements and collaborations with industries in the pursuit of such ambitious goals.

In this report, we present the scientific case for "Petascale Computing in Europe" over the period 2010-2020. It is the result of collective action, launched to support the concept of a truly international programme across Europe. Our vision is that Capability Computing is not a concept or a line of action which competes with other approaches, such as "Grid Computing". It is the top of the pyramid of the entire computational infrastructure which builds upon the other layers and which brings value to the whole by permitting the infrastructure to tackle and solve major scientific challenges.

A series of meetings of representatives of France, Germany, and the United Kingdom, in 2004, together with Spain in 2005, stressed the crucial need for Europe to remain competitive with respect to the United States and Japan, summarizing this vision while calling for action [SP-ESR]:

> *« While Supercomputers of the highest performance class have become indispensable tools for science and engineering in academia, research and industry, financing of*

the next but one generation of supercomputers will strain the national budgets of the individual European countries. Hence, this paper suggests that the establishment of infrastructures for high performance computing shall be implemented on a European level and discusses first steps to be taken by the three undersigning countries. The goals of the proposed action are the consolidation of Europe's important role in both supercomputer applications and in ensuring the availability of the most advanced supercomputer systems in the EU. »

2.2 High End Computation and Computational Science

Beyond their own development as fully fledged disciplines, computational science (CS) and scientific-based engineering science (SBES) certainly have evolved into bringing new methodologies and ideas into mainstream science and engineering, as expressed recently by two American reports: the PITAC report on "Computational Science" [PITAC-CS] and the NSF report on "Scientific-Based Engineering Science" [SBES].

The advent of Computational Science and advanced supercomputers has made it possible to deal with very complex scientific phenomena, permitting the application of new, more quantitative and predictive theories in areas where experiment is impossible or not practical or where these theories cannot be evaluated by other means In order to understand such phenomena, it is now vital to study the interaction of systems that have been kept conceptually separate in earlier simplified approaches. Examples include the modelling of the earth climate to simulate Global Warming, and the optimization of aircraft designs accounting for aerodynamics, structure, and acoustics. In the case of biology, the explosion of genomic data provides only a partial description of the system under study, from which answers to very specific questions are required to make progress in priority areas such as pharmaceuticals and health. Only simulation and modelling can make this progress possible.

Another aspect of crucial value in science and engineering is the ability to describe phenomena at various length and time scales. In many problems, the relevant scales that must be considered differ by many orders of magnitude. The empirical modelling approach consists of constructing models for the intermediate scales, and validating these through experimentation. This brings into question the validity at the various scales; there is no information in conditions that do not lend themselves to experiments – for example extreme conditions, the scales demanded in astronomy or particle physics, and the spread of time scales in biochemistry. Capability supercomputers now allow scientists to construct a hierarchy of models, where each is founded on the characteristics computed at the lower scale. At the base of this hierarchy the "*ab initio*" approach relies on the theoretical models of matter. The reward for building and validating such a hierarchy is that science and engineering is capable of dealing with important new fields of both practical and theoretical interest.

Supercomputers of the highest performance levels will be a key ingredient in quantifying the uncertainties in numerous models. The area of climatology provides compelling examples where such information is essential for informing political decisions on the optimum approach to "Global Warming". A key requirement in the engineering disciplines is to quantify these uncertainties where they relate to safety, efficiencies of industrial processes, and the effectiveness of complex devices.

In summary, computational science fuelled by top level supercomputers has established itself as the third pillar of scientific methodology alongside theory and experiment, often bridging the gap between both. It has the potential to influence profoundly science and engineering through its ability to treat highly complex phenomena, and to base a larger part of the model hierarchy on first principles rather than on effective theories or empirical approaches, thereby quantifying uncertainties.

2.3 The Scientific Perspective

The main part of the report is devoted to the description of a scientific roadmap, detailing the major challenges, the scientific and societal benefits through making progress towards their resolution, and the prerequisites for being able to tackle these challenges. Preceding it, we have provided two sections that deal with the key components of a Computational Infrastructure, and the user support functions that must be provided to realize the full benefit of that infrastructure.

2.3.1 Weather, Climatology and Earth Sciences

There are few more critical areas of understanding that require research than the societal and economic realities that arise from the effect of climate change and weather and the geology and seismology of our planet. Research in the fields of weather, climatology and earth science is of key importance for Europe, for the following reasons:

- Informing EU policy on the environment and carbon emissions
- Understanding the likely impact of the natural environment on EU infrastructure, economy and society
- Enabling informed EU investment decisions in ensuring sustainability within the EU and globally.
- Developing civil protection capabilities to protect the citizens of the EU from natural disasters

The following sub-paragraphs summarise the key motivations

2.3.1.1 Climate Change

The field of climate change research is developing rapidly. However, whilst there is great confidence in the fact that climate change is happening, there remain uncertainties. In particular there is uncertainty about the levels of greenhouse gas emissions and aerosols and, perhaps even more significant, are uncertainties on the degree of warming and the likely impacts. These latter uncertainties can only be reduced by increasing the capability and complexity of 'whole earth system' models that represent in ever increasing realism and detail the scenarios for our future climate. These uncertainties allow the possibility of political biases in the scientific debate: decision-makers need therefore to access scientific expertise and modelling resources that provide the highest level of certainty and credibility.

2.3.1.2 Oceanography and Marine Forecasting

Our understanding of the oceans is critical to our ability to understanding and predicting climate change. Ocean modelling is an integral and core element of the 'whole earth system' model outlined above and is crucial to reducing uncertainties in climate change and its impacts. Beyond its role in climate, oceanography is also important in other ways and there remain specific challenges. Mankind's activities, including supply of food and energy, transport of goods, etc., are exerting a continuously increasing stress on the open and coastal oceanic environment, a stress that needs to be evaluated and controlled in order to preserve the ocean integrity and available resources. There is a need to be able to build the most efficient modelling and prediction systems to study, understand and predict ocean properties and variations at all scales, and develop economically relevant applications to inform policy and develop services for government and industry.

2.3.1.3 Meteorology, Hydrology and Air Quality

Weather and flood events with high socio-economic and environmental impact may be infrequent, but the consequences of occurrence can be catastrophic to those societies and Earth Systems that are affected. There is of course a linkage to climate prediction and climate change impacts, if severe meteorological and hydrological events are to become more frequent and/or more extreme. Predicting these low frequency, high impact events a few days in advance with enough certainty and early warning to allow practical mitigation decisions to be taken remains difficult. Understanding and predicting the quality of air at the earth's surface is an applied scientific area of increasing relevance. Air quality can cause major environmental and health problems affecting both industrialized and developing countries around the world (e.g., respiratory difficulties, especially for sensitive people, adverse effects on the flora and fauna, …). Here also advanced real-time forecasting systems are necessary tools for allowing early enough warning and practical mitigation in the case of pollution crisis.

2.3.1.4 Earth Sciences

Computational challenges in earth sciences span a wide range of disciplines and have significant scientific and social implications, such as the mitigation of seismic hazards, treaty verification for nuclear weapons, and increased discovery of economically recoverable petroleum resources and monitoring of waste disposal. In the realm of seismic hazard mitigation alone it is well to recall that despite continuous progress in building code, one critical remaining step is the ability to forecast the earthquake ground motion to which a structure will be exposed during its lifetime. Until such forecasting can be done reliably complete success in the design process will not be fulfilled.

All these areas of expertise require increased computing capability in order to provide **breakthrough science**. A programme of provision of Leadership-class computational resources will make it increasingly possible to address the issues of resolution, complexity, duration, confidence and certainty, and to resolve explicitly phenomena that were previously parameterized. Each of the challenges represents an increase at least by a factor of 100 over individual national facilities currently available. A large number of the numerical models and capability-demanding simulations described below will lead to operational applications in other European centres, national centres and in industry.

2.3.1.5 Societal Benefits

Natural disasters claim hundreds of thousands of lives annually and cause vast property losses. To what extent anthropogenic climate change will lead to an increase in occurrence and severity of extreme events and natural disasters is one of today's most important and challenging scientific questions. The countries that will have access to the highest performance in computing will be able to perform experiments which will become the references for future scientific assessments and associated political decisions. Even though Europe has a world-class expertise in climate, oceanography, weather and quality, and Earth sciences modelling issues, European scientists may loose their current prominence if they cannot access the most powerful computing systems.

The economic benefit to society of quantifying the certainty and impact of forecasts, on whatever timescale, is enormous. By providing probabilistic results for impact models to agencies involved in assessing impacts of extreme events or climate change adaptation, mitigation strategies can be developed and the impacts constrained. The societal benefits range from mitigation of high-impact weather by having a more accurate and timely weather nowcasting system to the impact on society from a better air-quality forecast with regard to health, traffic (e.g. fog and rain), agriculture (e.g. ozone influence), etc. Changes in the chemical composition of the atmosphere bear consequences on health and agriculture which both affect the state of the economy.

2.3.1.6 International Back ground

Europe has been able to establish a scientific community that has played a leadership role in climate research, air quality and meteorology and has effectively contributed to international assessments. These have provided the scientific basis for national and international agreements. The European community is already well organised through European projects and collaborative research. The accuracy and details of earth system research and predictions of climate change have been severely limited by the computing resources available to the community. It is to address this problem that Japan has developed a large supercomputing facility, the Earth Simulator in Yokohama, dedicated primarily to Earth system science. Major plans to enhance by orders of magnitudes the computing resources in support of climate research have also been formulated in the United States ("establishing a petascale collaboratory for the geosciences", http://www.isse.ucar.edu/HP_rick/NSA/petascale-science.pdf, 2005).

2.3.2 Astrophysics, High-Energy Physics, and Plasma Physics

2.3.2.1 Astrophysics

Astrophysics is a field of research in which high-end supercomputers traditionally play a dominant role, mainly because very often modelling and simulations must replace planned and controlled

experiments. The most demanding applications deal with systems and structures which span a large range of different length and time scales, and almost always non-linear coupled systems of ordinary and partial differential equations have to be integrated, in three spatial dimensions and explicitly in time, with rather complex material functions as input. The grand challenges that can be foreseen for the coming years range from the formation of stars and planets to questions concerning the origin and the evolution of the Universe as a whole. Moreover, supercomputer resources are needed to evaluate the huge amount of data expected from future space experiments such as the European Planck Surveyor satellite.

2.3.2.2 Elementary Particle Physics

In the context of the standard model of ***Elementary Particle Physics***, quantum chromodynamics (QCD) is considered as the fundamental quantum field theory (QFT) of the strong interaction. It describes the highly non-linear interaction between quarks and gluons which creates bound states like the proton or neutron. Because QCD is a strongly interacting non-abelean (i.e. highly non-linear) theory it leads to a great variety of phenomena, including e.g. topologically non-trivial field configurations, phase transitions, chiral symmetry breaking, confinement and a very complex structure of the physical vacuum. Because of its richness and experimental accessibility QCD is the natural framework to study properties which are potentially shared by theories beyond the standard model. In addition, up to renormalization effects and neglecting the small current quark masses, it is a conformal theory and thus the prime candidate to establish an AdS/CFT duality. The latter postulates that string theories on curved space-times in higher dimensions can be equivalent to conformal quantum field theories in ordinary space-time. For all of these reasons QCD is the topic of intense theoretical and experimental research by a large and truly international community involving large European centers like CERN and DESY and integrating collaboration efforts within such as DEISA and the I3HP hadron physics project. In the future, our understanding of QCD, and possibly other QFTs and string theories, will undergo scrutinizing tests at new experimental facilities like LHC at CERN, FAIR at the GSI or the planned ILC (international linear collider). These experiments not only promise to yield a much deeper understanding of the standard model elementary particles and the forces between them, as well as nuclear forces, but are also expected to discover hints for a yet unknown physics beyond the standard model.

2.3.2.3 Plasma Physics

In ***Plasma Physics*** the science and technology challenge raised by the construction of the magnetic confinement thermonuclear fusion reactor ITER calls for a major theory and modelling activity, in order for the numerous basic physics phenomena to be better understood, but also for all the relevant models to be integrated in a series of codes able to reliably predict the behaviour of both the plasma discharge and the various components of the device, the so-called ‘tokamak simulators’. Both the success of the experiment and its safety rely on such simulators. The quest to realize thermonuclear fusion by magnetically confining a high temperature plasma poses some of the computationally most challenging problems of nonlinear physics. Fusion plasmas are nearly collision-free, and ab-initio modelling of some of the critical phenomena requires a kinetic description of the plasma with a distribution function depending on three space and two velocity coordinates.

2.3.3 Materials Science, Chemistry and Nanoscience

Materials science simulations are playing an increasing role in almost all aspects of the study of materials: not only traditional materials science, physics and chemistry, but also nanoscience, surface science, electrical engineering, earth sciences, biology and drug design. By throwing in the atoms and the laws of physics and chemistry into the computer, one effectively performs 'computer experiments' to observe structures, properties and processes and to conduct measurements that cannot be readily undertaken in the laboratory.

2.3.3.1 Understanding Complex Materials

The progress expected in the understanding of technologically relevant materials comprises the determination of purely electronic properties, such as optical properties (optomechanical and optoelectronic properties and optomagnetic processes) or transport properties (giant and colossal magnetoresistance). These properties are central to many devices in the electronic industry.

The simulation of nucleation, growth, self assembly and polymerization is central to the design and performance of many diverse materials such as rubbers, paints, fuels, detergents, functional organic materials, cosmetics and food.

Finally, the relation between process, conditions of use and composition of a material requires a multiscale description of e.g. mechanical properties of materials (nuclear energy production). As an example, the multiscale simulation of the mechanical properties of irradiated materials would require access to facilities 1000 times larger than those currently available.

These simulations would decisively contribute to the prediction of the lifetime of high performance materials in energy technology, such as high-efficiency gas-turbines ...

2.3.3.2 Understanding Complex Chemistry

The simulation of technologically relevant chemical reactions and processes has huge potential in a variety of fields.

Catalysis is a major challenge in the chemistry of complex materials, with many applications in industrial chemistry. The knowledge of atmospheric chemistry is crucial for environmental prediction and protection (clean air). Improving the knowledge of chemical processing (from soft chemistry including polymers to the atomistic description of combustion) would improve the durability of chemicals. Supra molecular assemblies open new possibilities for the extraction of heavy elements from spent nuclear fuels. In biochemistry, a vast number of reactions taking place in the human body (for example) are not understood in any detail.

Addressing effectively the modelling of these complex chemical systems can only be achieved with a massive upgrade of computing resources. As an example, the realistic treatment of supported catalytic nanoparticles involving several hundred transition metal atoms would require resources of at least a Petaflop/s. This is a key step in the development of the clean fuels of the future.

2.3.3.3 Making Progress in Nanotechnology

The advance of faster information processing or the development of new generations of processors requires the shrinking of devices, which leads unavoidably towards nanoelectronics. Moreover, many new devices, such as nanomotors can be envisioned, which will require simulation of mechanical properties at the nanolevel. Finally, composite high performance materials in the fields e.g. adhesion and coatings will require an atomistic based description of nanorheology, nanofluidics and nanotribology. As an example the description of the complex magnetic and mechano-optical properties of nanodevices components, requires sophisticated computational methods, feasible only on nanosize systems with massively parallel computers in the Petaflop/s range

2.3.4 Life Sciences

2.3.4.1 *Systems Biology:*

These projects are just in the beginnings, but appear as one of the most important issues of research in biology for the next decade. Major European (like EMBL) or National institutions are considering System Biology as the top priority, and are designing entire research programs devoted to this interdisciplinary activity. The main idea of systems biology is to represent using simple models the entire behaviour of cells, tissues, and organs, or to evaluate degradation routes predicting the final excretion product of any drug in any organism. Systems Biology projects start always from the volume of information derived from genomic experiments (protein arrays, DNA chips, yeast two hybrid experiments,…) which is used to define the systems nodes, connexions and parameter, generating a

giant networks defined with different resolutions and focus in different regions, and where due to the low quality of the experimental information, all nodes and connections have a probabilistic nature that increases even more the complexity of the system. Considering even the simplest cell we cannot expect to identify all the important players from experimental techniques. Simulation will then be the only possibility to detect for example hidden players (processes or proteins that are important only under certain conditions), or to understand the behaviour of the cell, tissue or organs under anomalous conditions. The simulations needed imply massive calculations, impossible to perform in a GRID-based platform, and out of the possibilities of any current computer centre. In any case, Europe must be ready in the next 4 years to host the first "in silico" cell, probably a computational implementation of the "subminimal cells" which are now being experimentally developed. Otherwise, Japan and the United States will take a dramatic advantage in the post-genomic race,

2.3.4.2 Chromatine dynamics:

Once the genome has been solved, the next question is to determine how the genome regulates its function. Analysis of databases like TRANSFAC reveals that known sequence rules alone cannot explain the differential expression of genes. Physical properties of DNA must then play a key role in modulating gene expression. A particularly interesting issue of gene regulation is the dynamics of chromatine. We know that the DNA of eukaryotes is organized in nucleosomes which are then wrapped forming long fibbers. The organization of DNA in nucleosomes largely modifies the accessibility of transcription factors recognition sites playing then a key role in the regulation of gene function. The understanding of nucleosome dynamics, positioning, phasing, formation and disruption or the modifications induced by chemical modifications, or by changes in the environment is a great challenge. This would imply extensive molecular dynamics simulations on the scale of the microsecond for systems that contains from 10^5 to 10^6 atoms.

2.3.4.3 Large scale protein dynamics:

To understand the physical reasons behind large conformational changes in protein is one of the main objectives of structural biochemistry. For example, fifty years after the discovery that sequence contain all the information for guiding protein folding we are still unable to reproduce this process and it is now clear that such knowledge will arrive only from physical based simulations, where physical force-field will be combined with atomistic molecular dynamics algorithms. The importance of protein folding simulations, which is now the focus of a large number of computational projects like Blue Gene, should not hidden the relevance of the study of other large conformational changes in protein. Particular, due to its biological and biomedical importance big challenges appear in the simulation of protein missfolding, unfolding and refolding (a key element for the understanding of prion-originated pathologies), allosteric changes, and conformational changes induced by protein-protein interactions. All these calculations imply extremely long molecular dynamics simulations (time scale between micro and millisecond) of systems containing 10^4-10^5 atoms. These simulations are possible only in very fast and parallel computer systems, with fast interconnection between processors.

2.3.4.4 Protein association and aggregation.

It is clear that proteins are not in the cell are in a crowded environment and are continuously making interactions with other macromolecules. The genomic revolution is showing that such interactions are the responsible for the control of cell function. For example, cascades of protein-protein interactions are the responsible for increasing the expression of a given gene as a response to the binding of a ligand to a membrane receptor. One of the greatest challenges for the next decade will be the simulation of these crowded protein environments. Unfortunately, we are facing a formidable problem since protein-protein interactions are controlled by two mechanisms: i) a Brownian diffusive mechanism, and ii) an atomic detailed docking between macromolecules. Combined methodologies, where Brownian and molecular dynamics are combined with docking algorithms incorporating evolutionary information should be used in massive simulations. To be able to represent "in silico" the formation of the different protein complexes associated to a signalling pathway will be a great biological and computational challenge, which will help to bridge cellular biology with chemistry,

opening the door for a better understanding of cellular function and to the generation of new drugs able to interfere in protein-protein interactions.

2.3.4.5 Supramolecular systems:

Current simulation protocols and computers allow us to reproduce reasonably the interaction of small molecules with small or medium-sized proteins, but the correct representation of protein machines is still out of the possibilities of European groups. American groups, supported by massive computer resource have proven that, under ideal conditions, we can reproduce enzymatic catalysis, or the behaviour of large protein machines (like ATPases, ions channels,…) with more than 10^5 atoms. The challenge will be to analyze systematically how several of these machines (ribosome, topoisomerases, polymerases,…) work. The impact of these challenging simulations will be tremendous, not only for fundamental biology and biochemistry but also for nanobiotechnology application, since we know that molecular catalyzers and protein engines are several times more efficient than the best models developed by humans. Again we are facing big computational challenges since MD simulations in the range of the microsecond should be made in systems containing up to 10^7 atoms under different external conditions. Only supercomputers can make this type of simulation possible.

2.3.4.6 Medicine.

The last objective of all research in Life Science should be the improvement in our life quality. That is, all basic research should be addressed as to provide scientific information that can lead to the development of new, more efficient medical treatments. Genome sequencing, massive genotyping studies are providing us with massive information. But connecting this information with human health is not always simple. Particularly interesting is the simulation of the determinants triggering the development of multigenic-based diseases and the prediction of secondary effects related to bad metabolism of drugs in certain segments of population, or to the interaction of drugs with macromolecules others than their original targets. A great area of work appears for bioinformatics in this area, where the challenge is to combine database analysis with very CPU-demanding techniques like maximum likelihood models, systems biology approaches, microscopic and mesoscopic simulations.

2.3.5 Engineering

2.3.5.1 Complete Helicopter Simulation for Next Generation Rotorcraft

The ability for hover, vertical takeoff and landing gives the helicopter a unique place in the transport system, e.g. with a significant role for rescue operations, police missions, traffic control, and remote transport of goods. However, the helicopter is still a costly machine with high maintenance costs, poor fuel efficiency, moderate ride quality, remarkable noise generation and air pollution. These short comings stem from the complexity of the flow around a helicopter. The complex unsteady aerodynamics encountered features shock waves, separation, vortices and interaction. In addition, the dynamics of the blades and the rotor head assembly are complicated. Thus, the ACARE 2020 vision defines targets for an improved efficiency and ride quality, and a reduced environmental impact of rotorcraft. The European helicopter industry has a strong tradition of innovation in technology and design. Computational Fluid Dynamics (CFD) based simulations of aerodynamics, aeroacoustics and coupling with dynamics of rotorcraft already play a central role and will have to be improved further in the design loop. There has been consistent funding from the European Union to develop the codes necessary for rotorcraft analysis, e. g. DACRO, ECARP, HELISHAPE, HELIFUSE and EROS. The GOAHEAD project recently started to generate a database for the validation of complete helicopter simulations and involves the cooperation of the European helicopter industry (Eurocopter, Agusta-Westland), national research centres, and universities.

2.3.5.2 Biomedical Flows

Today biomedical fluid mechanics can improve healthcare in many areas. Among other applications, there is quite an intensive research effort in the field of the human circulatory system, for example

blood flow disturbances caused by arterial malformations, the artificial heart or heart valve prostheses are investigated, the respiratory system with the nose flow and the upper and lower airways, and the human balance system, e.g., the fluid mechanics in the cochlea. Although experiments have significantly improved the understanding in the field of biomedical fluid mechanics, numerous questions, the answers of which need a high resolution of the flow field, of the surrounding tissue, or of their interactions, require a detailed numerical analysis of the biomedical problem. The flow problems encountered are characterized by an extremely complex geometry, moving boundaries, locally and globally unsteady flows, fluid-structure interactions, Newtonian and non-Newtonian fluids, and by single and multiphase flows to mention just a few phenomena. Any computational simulation therefore faces an extreme challenge. Similar features defined by quasi discontinuous flow, transition, and intricate configurations also occur in the numerical analysis of flows over aircraft, helicopters, and turbine blades.

2.3.5.3 Gas turbines and internal combustion engines

Combustion represents more than 80% of energy conversion in the world and has a strong impact on the environment (greenhouse gases, pollutant emissions). The objective of combustion studies is to better understand and model physical phenomena to optimise, for example, gas turbines (aero-engines or power generation) or internal combustion engines, in terms of costs, stability, higher combustion efficiency, reduced fuel consumption, near zero pollutant emissions and low noise. Computational Fluid Dynamics (CFD) offers to design engineers the unique opportunity to develop new technical concepts, reducing development costs by avoiding extensive and very expensive experimental campaigns. From an economic point of view, industrial companies involved in propulsion and energy systems are among the biggest employers in the European Union. To give them more efficient and cost effective system designs is crucial support to promote their competitiveness on the worldwide market.

Scientific challenges in gas turbines or piston engines are numerous. First, a large range of physical scales should be considered from fast chemical reaction characteristics (reaction zone thicknesses of about tens of millimetres, 10^{-6} s), pressure wave propagation (sound speed) up to burner scales (tens of centimetres, 10^{-2} s resident times) or system scales (metres for gas turbines). Turbulent flows are, by nature, strongly unsteady and large eddy simulations (LES), where the largest flow motions are explicitly computed while only the effects of the small ones are modelled,[1] appear as a relevant CFD tool, especially when combustion instabilities or cycle-to-cycle variations in piston engines, occur. Chemistry involves hundreds of chemical species and thousands of chemical reactions, and cannot be handled in numerical simulations without adapted modelling. To design cooling systems requires the knowledge of heat transfer to walls due to convection and radiation.

2.3.5.4 Forest Fires

Regularly forest fires strongly affect south European countries such as Portugal, Spain, France, and Greece and because of climate change, may concern northern regions in the future. The social impact is very important and is concerned with land, buildings, human and animal life, agriculture, tourism and the economy. Accordingly, the development of reliable numerical tools able to model and predict fire evolution becomes of importance in terms of safety and protection ("numerical fire simulator"), fire fighting and could help in real time disaster management.

2.3.5.5 Challenge: Green Aircraft

ACARE 2020 provides the politically agreed targets for an acceptable maximum impact of air traffic on people and the environment, while at the same time allowing the constantly increasing amount of air travel. The goals deal with a considerable reduction of exhaust gas (NO_x 80%, CO_2 50%) and noise

[1] Obviously, the flow motions explicitly determined in large eddy simulations should be resolved on the numerical grid mesh. Accordingly, the thinner this grid mesh is (and the higher the computational costs!), the smaller will be the unresolved scales and thus the smaller the importance of sub-grid scale models.

(10-20 dB). Air traffic will increase by a factor of 3, accidents are expected to go down by 80%. Passenger expense should drop (50%) and flights become largely weather independent.

It is clear that these targets cannot be achieved by enhancement of existing technology, and that they require essential technology jumps. The "Green Aircraft" is the answer of the airframe as well as engine manufacturing industry. It is supposed to minimize environmental impact and at the same time provide more transport capacity through alternative operational scenarios. In conjunction with the European research community, the aircraft industry has started investigations on the large scale deployment of flow control techniques (hybrid laminar flow, dimples, plasma, etc.) as well as on radically different configurations and design solutions (blended wing/body, etc.). However, it is only by a far more productive high quality numerical simulation and optimisation capability that such a challenging development will be possible. It will be indispensable to be able to compute the real aircraft in operation, including all relevant multi-disciplinary interaction.

2.3.5.6 Challenge: Virtual Power Plant

Safe production of high quality and cost effective energy is one of the major concerns of Utilities. This is not only a source of competitiveness for them, but also for their costumers and especially for companies for which the energy is a major contributor to their costs. Several challenges must be faced to achieve this goal, amongst which are extending the lifespan of power plants to 60 years, guaranteeing the optimum fuel use or better managing waste. While numerical simulation has been a strategic tool for decades for utilities that produce, transport, distribute and sell energy, these challenges will not be accessible with old generation simulation tools and computers. We need to address very fine simulations in order to provide reference computations for less detailed approaches, to couple various disciplines (fluid-structure interaction, neutronics and thermalhydraulics interaction, …), to do multiscale calculations (from microstructure to macrostructure), and to compute actual (and not simplified) geometries. Petascale machines are required to perform most of these advanced simulations along with a new generation of codes and simulation platforms. Strong collaboration is also required between physics (thermalhydraulics, material science, structural mechanics, chemistry,…), numerics and computer science (including development of parallelism, and pre- and post-processing)

The Virtual Power Plant challenge is divided into two sub-challenges:

- Safety and optimization of power plants,
- Design of future plants.

2.4 Economical and Societal Benefits

Engineering

Europe has been able to establish leading class enterprises developing advanced complex systems, in areas where technological innovation is a major requirement. The examples developed in the sequel come from aerospace, biomedicine, the power industry, the environment and disaster management, and are by no means a complete list.

Leadership in advanced simulation will bring economical and societal benefits for several reasons:

1. Complete system modelling, leads to higher fidelity information and the opportunity to optimize the entire system. This results in more cost effective, safer and more environment friendly designs. Examples of such are found in the areas of aircraft and helicopter design, automobile design, aviation engines and power plants.
2. Design customization for the individual. Several examples can be found in the evaluation of surgical procedures, in connection with detailed information on a patient anatomy. Since tomographical tools provide three-dimensional information about key features of a patient's anatomy, it will become feasible to better select treatments and surgical procedures after simulating the possible outcome based on the patient's data for modelling tissues, organs and blood flow.

3. Multiscale modelling of materials is key to the design of materials able to withstand very unusual conditions: temperature, pressure, irradiation etc. It will also become possible to model several aging processes. The outcome is essential for building new generation nuclear power plants and ensure their long term efficiency and safety.
4. Design at the atomic level. In nano science, and nano electronics, the scale of the relevant mechanisms has become so small that models based on quantum physics are required. A cascade of modelling tools needs then to be built to be applicable for the design at hand. Leading edge simulations will enable the design of nano systems and will also be required to fully qualify the industrial advanced processes that permit their production.

Climate and management of earth resources

Europe has been able to establish a scientific community that has played a leadership role in climate research, air quality and meteorology and has effectively contributed to international assessments. These have provided the scientific basis for national and international agreements. The European community is already well organised through European projects and collaborative research.

Not withstanding the substantial progress to date, the accuracy and details of earth system research and predictions of climate change have been severely limited by available computing resources. A programme of provision of Leadership-class computational resources will make it increasingly possible to address the issues of resolution, complexity, duration, confidence and certainty, and to explicitly resolve phenomena that were previously parameterized. Each of the identified challenges represents an increase at least by a factor of 100 over individual national facilities currently available. A large number of the numerical models and capability-demanding simulations proposed to provide breakthrough science will lead to operational applications in other European centres, national centres and in industry.

All the challenges identified will provide increased knowledge and improvement of climate prediction, oceanography, weather and air quality, natural disaster forecasts, and have direct implication on energy resources, environmental management, and international negotiations. For climate, the key questions concern more accurate climate change prediction with models representing the synoptic scale processes in the atmosphere and the ocean, the interactions between climate and the biochemical cycles, models that can be run in ensemble mode or sufficiently long to account for uncertainties and the risk of abrupt climate changes. Increased resolution is also at the heart of oceanography. The next generation ocean models needs to represent properly eddies of the order of 20 km to better represent ocean heat transport and different key processes affecting the marine carbon cycle and the marine ecosystems. Very high resolution models that can explicitly represent small scale processes that are in general highly parameterized, such as clouds, are needed to provide better weather forecast and improve air quality forecast at the scale of Europe, which is a vital need for the population.

Key challenges in the management of earth resources include earthquake ground motion simulation and seismic-hazard, high resolution imaging techniques, the improvement of the resolution and the understanding of the earth's interior structure and dynamics, and the understanding of the generation of the Earth's magnetic field. The resolution of these challenges would lead to substantial benefits to society.

Life Sciences

There are numerous examples of economical and societal benefits within the Life Sciences, a number of which are listed below:

1. genome annotation (which is largely computational)is essential for identifying new drug targets;
2. protein fold recognition and structure prediction are essential to allow the engineering of new proteins, as catalysts, biosensors, or nanomaterials (motors, fibers);
3. computational biochemistry is a key part of the drug development pipeline, through docking and structure prediction calculations;
4. protein misfolding is the basis of several major diseases, such as amyloid fiber diseases (Kreuzfeld-Jacob and mad cow disease);

5. emerging pathologies (SARS, avian flu) require very fast computational responses: massive drug design projects were recently performed in a very short time in France by HPC and in the UK by grid computing; several promising lead molecules were identified;
6. simulations if disease evolution have obvious implications for public health;
7. simulations of biological nanomachines (chromatin, membrane transport systems) are essential to understand the molecular mechanisms of cancer and other diseases;
8. computational modelling of the cell will play a role in understanding cell growth and differentiation, and ultimately phenomena such as embryogenesis and development;
9. whole-organ modelling will be increasingly a tool for medicine and surgery in humans and animals.

2.5 Relation with Other National and European programmes

Essential functions of a European supercomputing infrastructure

There are several complementary functions that must be provided by a Computational Infrastructure if it is to prove both effective and persistent:

a) The development and evolution of innovative application programs, models and methods.
b) Preparatory and post processing work, permitting the design and validation of particular models; this may require both data preparation plus the analysis and exploitation of the data generated by the computations,
c) Large scale systematic studies, where each case requires true super computer power. This enables exploration of the parameter space of devices and phenomena, with the ability to deal with multiple combinations of parameter values, thereby enabling the investigation of the statistical behaviour of phenomena.
d) Extremely large, so-called "hero" computations, where the sheer power of the entire computational resource is used to study more detailed models than previously possible. The objective may be scientific insight, where the model would include scientific aspects not previously understood, or an attempt to deal with more detailed data than usually feasible. In industry, it may be necessary to validate models extensively before they are used more routinely in design processes. Extremely large computations may also be required to deal with unexpected situations and incidents, in order to mitigate the consequences or rapidly prepare design changes.
e) Efficient algorithms are an essential ingredient of any HPC project. As larger and larger problems are solved on larger and larger computers, it becomes increasingly important to select optimal, or near optimal, algorithms and solvers. As most problems have a superlinear computational complexity, simply relying on hardware advances to solve these larger problems is ultimately doomed. Moreover, some of the more critical tasks are generic, in the sense that they are not tied to one particular application, or even one particular field, but will occur in most of the challenges listed in this report e.g., the solution of (large and sparse) linear and non-linear systems, computation of the Fast Fourier Transform (FFT) and integration of time dependant differential equations.

Usually, "capacity computing" is deployed against tasks (b) and (c), while "capability computing" provides the only solution to deliver against task (d). In this scientific case we advocate that this essential component of capability computing should be performed through shared European services that complement national facilities. This will add value at all levels, in particular by being more competitive on the innovative aspects permitted by type (d) tasks. We also show that the infrastructure needs to embrace a pyramid of resources to effectively deliver against all of the above, and we consider from a scientific perspective how this infrastructure might be best balanced.

In order to integrate the variety of resource levels, facilitate access for users and simplify the management of the extreme volumes of data required, an appropriate electronic data communication infrastructure is key. Typically referred to as a "Grid", this infrastructure needs to be highly tuned for

HPC usage, and connected to the various tiers of HPC facilities. The DEISA project is currently developing and performing a realistic field test of such an HPC-Grid.

It is quite revealing that the case for a balanced resource pyramid, embracing capacity resources, can be made based on both technological and architectural grounds, as the paper by G. Bell and his co-authors shows [BGS].

3 STRUCTURE

3.1 The Resource and Organisational Pyramid

The resources required to support computational science through a number of large and often diverse computational projects span a hierarchy of levels – desktop, departmental or laboratory level machines, regional centres, and supercomputer centres. These resources need to be organised in a hierarchical multi-tier pyramid and connected by adequate high-speed links and protocols. The reasons for such an organisation are many:

1. Organisational: some functions must be performed at the level of small research teams operating in a truly interactive fashion. This requires desktop and laboratory level equipment. Other aspects may involve a large community or may require a scale of resources such that only a regional or national centre is appropriate,
2. Technical: despite progress in telecommunication technologies, the transfer of data still has a significant cost and "end-to-end" throughput limitations. This has major consequences at all levels. In particular, the network or "Grid" architecture, the massive data repositories and resource centres devoted to both capability and capacity jobs need to be balanced and interconnected. It also means that there are empirical, technology dependent, rules and constraints governing this overall provision.
3. Economical: with current technology, the greater capabilities associated with the largest supercomputers lead to a higher cost performance ratio than that associated with smaller equipment. Therefore they should be applied to large jobs that cannot be performed otherwise; the resource pyramid has an optimal shape that most countries or organizations try to achieve empirically. With continued technological progress, it is envisioned that this trend will increase: smaller equipment benefits through higher levels of "on chip" integration, while larger equipment must integrate massively parallel resources and have fast growing communication requirements.
4. Financial: the requirements are so extreme that only a small number of very large centres may be purchased and operated. Collectively they must have the capability to handle very large jobs while providing the throughput essential to sustaining scientific progress.

While this report targets specifically the infrastructure required to handle capability jobs, it also acknowledges the importance of the remainder of the pyramid. This comprises the grid or network infrastructure, which can be based on state of the art developments within the GEANT-2, DEISA and EGEE projects. The national computational centres are important resources; our view is that the European dimension should consider them as an integral part of the European resource pyramid, whose apex should be an exceptional, very large capability "Leadership" class resource. Such an approach will position European capability resources at a level comparable to the best in the world, resources that to date are only available in Japan and the USA.

3.2 Capability vs. Capacity

Capability class supercomputers are dimensioned and used to perform very large tasks that cannot be handled by other means. This depends on a number of factors:

1. the state of technology: what are the largest systems that may be designed, procured and operated at a given point in time ?
2. the economics: what cost may be justified, based on the expected scientific, industrial and societal impact ?
3. the characteristics of the computational job: why does the job require such a resource ? This may be caused by the sheer size of the computational task in hand, by time to solution constraints, or by the difficulty or impossibility of partitioning the computation. Usually these factors act in combination.

In the forthcoming chapters of this report, it is shown that a wide spectrum of scientific challenges demand capability resources best achieved at the European level. The justification of such an endeavour will be given on scientific grounds.

The effect of a European collaboration to advance the apex of the resource pyramid amounts to positioning it competitively with respect to similar systems in other major countries, notably Japan and the USA. The key driver is to promote scientific competitiveness; these systems should be targeted strategically at scientific challenges with the full support and agreement of the relevant scientific communities. .

3.3 Dedicated versus Multi-disciplinary

Several communities have requirements large enough to demand a dedicated, leading edge computational resource. The question therefore arises of whether an integrated set of dedicated resources offers a better solution than a number of independent, discipline specific topical-based resources.

The panel defends the view that the cost level of European capability resource provision will be such that only a very small number of machines can be procured. To deal with technological obsolescence, an upgrade scheme needs to be put in place across a small set of centres, only one of which will be at the highest level of capability at any time.

The consequence is that these resources should be managed in a multi-disciplinary context, trying to concentrate, when possible, specific areas in specific centres, favouring the synergies that appears in that case. The most relevant computational challenges will then be chosen by the scientific community on a regular basis, say annually. This also means that fair selection and evaluation processes must be put in place and managed though an appropriate European organization. Just such a model may be found around other large shared scientific equipment.

Overall, this discussion emphasizes that the future of supercomputer architectures is largely open, introducing a significant uncertainty and risk when procuring a new machine. This risk is increased by the commitment of several countries like Japan and the USA to introduce newer designs and architectures intended to give them a competitive edge. The visibility of a European capability oriented supercomputing initiative will lead to a far greater ability to influence the computer industry, notably the providers of high-end equipment. The panel does not have sufficient information and expertise to address the potential future of a European-based supercomputer industry.

Another consequence of this discussion is that it is highly unlikely that there will be a single design or architecture that best addresses the capability requirements of all disciplines. While recognising this multi-architecture landscape, nevertheless *the panel recommends an approach that is driven by multi-disciplinary, general-purpose hardware*, at least at the outset. This strongly suggests a managed process, driven by an organisation flexible enough to both decide upon, and subsequently procure, the optimum mix of supercomputer types required to address the major scientific challenges. Such an approach, together with international cooperation at the European level, provides a compelling risk mitigation strategy associated with the procurement of leadership class machines.

3.4 Evolution Strategy at all Levels

The capability and power of supercomputers is expected to continue unabated over the next decade. We have argued that the means and technical directions for this increase is likely to show some surprises, although the overall growth is very likely to conform to the empirical “Moore’s Law” which predicts an order of magnitude increase every 4 to 5 years. New architectures may eventually provide faster growth. More detailed technological data may be found in the International Technology Roadmap for Semiconductors [ITRS] report.

The above scenario means that a capability supercomputer remains competitive for at most 5 or 6 years and that an upgrade strategy must be implemented. The same argument applies to resources at all levels of the pyramid, including national centers. When multiple systems are in use, an “upgrade spiral” model, as implemented in Germany becomes very attractive – indeed, this was the basis for the

joint proposal [SP-ESR]. Of course, such models can be tuned to account for the interests of the partner countries and economic realities which include not only the technology, but the costs of operation and maintenance.

From this perspective a computational infrastructure needs to be financed in the long term. The initial design period is certainly shorter, but an upgrade process needs to be put in place in parallel with the operation of the infrastructure. Such a process needs to rely both on an evaluation of the utility of the infrastructure and a forward looking process to determine scientific goals and perspectives.

3.5 Long Term Perspective & Relation with Scientific Investment

Important considerations in the provision of High Performance Computing include the associated development infrastructure in place around the machines, plus the level of expertise required within the scientific community to ensure effective exploitation of the resources provided. More specific descriptions will be given in the forthcoming thematic chapters, but from a very general perspective this infrastructure should provide for:

1. the development of adequate models, and their evolution according to scientific progress;
2. the development of, or improvements to, the hierarchy of mathematical methods, numerical and statistical methods, and other resolution techniques required to fully exploit the developed models;
3. the development of the associated computer codes, together with associated algorithms and their efficient implementation on the available resources.

The distinctive aspects of these in large part constitute the discipline of Computational Science and Engineering. Its relationship with the scientific progress and scientific organisation are discussed in the PITAC report on Computational Science [PITAC-CS].

The panel wishes to emphasize several issues that are central to this scientific case:

1. The development of the necessary expertise in Computational Science, alongside the investment in the methods and codes that are needed to exploit fully the capability level equipment, present very significant scientific challenges in their own right. Investment in these areas will lead to substantial progress in both scientific research and technological applications.
2. There are strong arguments to emphasize research not only on specific challenges (though of course the challenges are what is driving the need for a dedicated facility), but also for research on generic solvers, that will provide reusable, basic, components to be integrated in many of the project working on specific challenges. A follow-up to this argument is the need for more interdisciplinary groups, composed of applied mathematicians, computer scientists and specialists from the fields underpinned by the challenge. The technology of the more efficient solvers has become sufficiently complex that it is hard for non-specialists to keep abreast of recent developments: solvers need to be tuned for the memory hierarchy found in the processors, they need to be parallel, and algorithmically efficient. The resulting codes need advanced data structure to remain efficient, and this can only be provided by the kind of collaborations alluded to above.
3. Addressing the above requirements requires significant planning and human investment. For example, the development of a large code may involve a collaborative team effort lasting some five years, culminating in a code that may be used for 10 to 20 years.

Therefore, a visible, long-term commitment of the European Community and of the Research organisations is crucial. Such a commitment would convince the scientific community to commit their own expertise and resources; indeed, commitment to a European capability level supercomputing infrastructure would be a clear signal of intent, confirming to leading scientists that Computational science is indeed perceived to be one of the major pillars of scientific progress.

This argument also suggests that a European capability level infrastructure would increase the role and impact of the overall computational resource pyramid: beneficiaries would include national centers, application codes repositories, access and data grids, and so on. We have already shown that computational infrastructure is an enabler for scientific and technological development; a European

leadership-class infrastructure will prove to an enabler for many scientific and engineering programmes.

3.6 Relationship with Grid

Regardless of the architectures deployed in the next five years to enable the European scientific community to maintain its international position in computational science, it will be critical to integrate these new computing platforms with the emerging computational infrastructure – providing integrated access to large-scale data, computing platforms, instruments and users. It is critical that Europe deploys these systems in a manner that creates a pan-European footprint and fully integrates them with the emerging research Grid.

A number of the key scientific disciplines of section 5 require a new scale of networking connectivity to exploit this new frontier of computing. The strength of European science is due in no small part to the infrastructure that supports it, and networking technology is rapidly reaching a level of performance that will change fundamentally our expectations of accessibility of our computational resources and our ability to move tasks and data between them transparently.

Finally, we should point again to the complementary relationship between the Leadership-class computing infrastructure central to this submission, and the computational infrastructure of the Grid. While the Grid should certainly make access to high-end resources more pervasive, with any plans for these services taking into account Grid developments, it does not provide a vehicle for addressing the key scientific goals highlighted in this document. These goals demand a high-end computing roadmap that is concerned with the development and deployment of the hardware and software infrastructure needed to enable large-scale computations for breakthrough science and engineering. Our focus is on simulations that are so memory, CPU, data, or I/O intensive that they require dedicated use of a large resource for extensive periods. Many of the associated attributes – sub-usec latency between co-operating processes, a global address space demanding uniform, sustainable memory bandwidth of O(GB/sec), sustainable real-time I/O transfer rates in excess of 50 GB/sec – are key requirements well beyond any distributed infrastructure currently available. This, coupled with the issue of how best to schedule resources across participating sites, an issue that is unlikely to be solved on the timescale of new HPC provision, reinforces the conviction that a GRID-based solution to our high-end computing roadmap remains a concept, and not a credible solution in the near term. High-end provision should routinely support multi-day runs utilising in excess of thousands of processors, closely coupled through a high-speed interconnect. Such provision is critically dependent on the type of Leadership-class facilities detailed in this document.

3.7 User Support

All across Europe, world-class research teams are using HPC resources to make new discoveries. The breadth of research applications is staggering, encompassing virtually all areas of the sciences, engineering, and medicine, with growing applications in the social sciences and humanities. Although it is the lead researchers that often assume the high-profile roles in the research process, for HPC-related activities there has to be a large supporting team working behind the scenes. An effective HPC facility is much more than hardware; the smooth operation of the facility requires a second tier of highly qualified personnel (HQP) to manage, operate and maintain the facility. It is equally important for these highly trained technical support staff to train and assist researchers in making the best use of this expensive infrastructure.

An investment in people for today and for the future is a critical component of this proposal. In many respects, computing infrastructure can be more readily acquired than human infrastructure. Given adequate funding, the upgrading of the capital equipment is straightforward: one can simply buy whatever is needed. However, human infrastructure is much more challenging to obtain. It can take years to train people with the necessary skill sets, and then they can be easily enticed away from Europe by the lure of better opportunities coupled with higher salaries. If Europe is to invest in people and skills, then it must also invest in creating the environment to attract and retain them.

3.7.1 Support Personnel

A variety of skilled personnel and support roles are therefore essential to the effective operation and maximum exploitation of any HPC facility. The skills and experience needed are extensive, including (i) managing, operating and maintaining the facility, (ii) training and assisting researchers to make the best use of its resources and capabilities, (iii) ensuring maximal productivity of the HPC sites by, for example, checking that software is run on the most suitable computing platform and reworking code to achieve significant performance gains, and (iv) helping to create new applications in support of innovative research initiatives.

System Administration and Operations

Systems administration and operations are primarily concerned with the day-to-day care of the HPC hardware and software infrastructure. The supporting personnel ensure the proper functioning of HPC facilities, providing systems management and operations support. Specific tasks include installing and maintaining operating system(s), performing updates and patches, managing file systems and backups, and ensuring the integrity and security of the user data. These activities are crucial to ensuring that the system is fully functional and available to the community.

Programmer/Analysts

The role of programmer/analysts is to provide specialized technical assistance to researchers, to conduct workshops and training, and to evaluate and implement software tools to make effective use of available resources. HPC hardware typically operates at a sustained rate well below the theoretical peak performance of the system; this is usually due to a lack of parallelism in parts of an application. A creative team of programmer/analysts can double that rate through code optimizations, algorithm re-design, enhanced cache utilization and improved data locality. The added value from such activities can be huge, and can correspond to *twice* the science delivered for the same hardware. These skills can thus dramatically increase the scientific productivity of the research community. By allowing researchers to run their applications faster, analysts support researchers and their students to do better science.

In extreme cases, the support team may be responsible for enabling research that would not have happened without their work. It is therefore extremely important that researchers and their science are matched to the appropriate HPC facilities and personnel resources within a European support model.

Applications Programmers

Frontier science requires world-class software applications. While much of the development of new scientific functionality is traditionally carried out in a researcher's own laboratory, HPC applications programmers often make valuable contributions to this work by virtue of their own scientific, numerical or visualization experience. The additional skills of the support staff, when matched with researcher needs, often play an integral role in enabling ideas, concepts and advice to flow with greater ease in the subject domain of the scientist. This support has the additional benefit of greatly reducing what is normally a challenging startup period for researchers learning to work with HPC. This skill set is imparted to students and postdoctoral fellows as well, giving them both the scientific knowledge and the programming experience necessary to create new computational methods and applications in their various fields, eventually leading to dramatic new insights.

Data Management and Visualization Personnel

The importance of versatile analysis and visualization techniques for simulation work is now self-evident, and both computation and visualization activities are increasingly being driven by "science pull" rather than "technology push."[2] The most challenging aspect of data management and visualization is coping with the massive datasets that are being produced. Simulations in climatology, bioinformatics and astrophysics, for example, regularly produce data sets that are hundreds of terabytes or even petabytes in size. Entirely new techniques and computing resources will be

[2] Science is increasingly driving the rate of technology development, rather than merely tracking it.

necessary to cope with them: in most cases, interactive visualization is the only practical way to glean insights into these datasets. In addition, the effective exploitation of such volumes of data will require a major development effort in distributed computing across high-speed networks (grid computing). This requires the training and the retention of personnel able to manage the data resources and to develop the new tools and techniques required to visualize them.

Training of Researchers

The overall goal of HPC support is to proactively support the needs of a wide variety of researchers by engaging them through their entire HPC lifecycle. In this context, support staff must (i) create awareness of the resources available and the potential for these resources to accelerate research productivity; (ii) provide usage instructions and courses (on topics such as parallelism, programming tools and performance analysis); (iii) help users find the right match between their application and the available technologies; (iv) maintain the HPC facility; and (v) develop new tools or tune existing ones (hardware and software) to optimize the use of HPC resources and the applications running on them. This type of support will enable the researcher to obtain more results in the same unit of time or the same results in less time, thus freeing up facility time for others.

HPC support staff is essential for training Europe's next generation of scientists and engineers in the use and improvement of HPC resources. Interactions of HPC support staff with graduate students and postdoctoral fellows will provide a fertile training ground to develop the next generation of researchers, giving the new scientists a grounding in HPC as part of their disciplinary training. Much like today's researchers use personal computers to support their work, the next generation of researchers will rely on and be able to take effective advantage of HPC resources, thus accelerating their research outputs. To do so effectively, researchers will need appropriate training.

3.8 Building a Community and Software Facilities for Petascale-class computers

In the preceding section we have considered the generic class of HPC support staff required in developing, sustaining and supporting a Petascale-class computing infrastructure. We consider below the broader requirements of community building and the software facilities crucial to the successful exploitation of Leadership Class systems. First, we illustrate these requirements through specific reference to the area of Materials Science, Chemistry and Nanoscience. This analysis is then broadened through a consideration of the human dimension of computational science infrastructure and what type of infrastructure is needed to keep Europe as a leading player on the world stage. Specifically we are concerned with the format of a network of expertise necessary to sustain Europe's leadership role (note that such deliberations are currently underway as part of an ESF funded Forward Look on European Computational Science).

3.8.1 Community building and Software facilities in Materials Science, Chemistry and Nanoscience

We can identify four key requirements in this endeavour:

a) Particular Grand Challenges need to be selected, and consortia (teams) brought together to pursue them, with the relevant software.

b) Special support staff are needed for getting the appropriate software written, and adapted over time to ever evolving computer architectures.

c) It is critical to have high quality software available at the time that the computer facility arrives.

d) Over the longer term (10 years) it is necessary to sustain the wider community in atomic scale simulation, from which new ideas for better algorithms and novel types of application come to feed into the supercomputing. Some support on the European level is needed, as well as keeping contact with national developments.

Motivation (scientific benefits, social benefits)

The development of a supercomputing facility will result in building a well-connected and well-organised "European supercomputing community" within the wider community in atomic scale simulation of materials so as to maintain excellence by world standards. We expect an impact similar to the gathering of experimental elementary particle physics around CERN and materials science around ESRF, ILL etc. This will improve the connection of national activities to form an integrated European activity.

Relation to state of the art

a) European researchers already lead the world in atomic scale simulation of materials.

b) This includes work on national supercomputing facilities, mostly using open source software with few commercial or other constraints.

c) Most of these codes were developed out of research collaborations such as the European electronic structure community PSI-K and molecular dynamics in MOLSIMU. These have grown up over the last 25 years and longer, with support by various programmes of the European Community and particularly the ESF. Indeed the excellence by world standards has come from a strong spirit of European cooperation in this community: one can say that they have developed from the start more on a European basis than the coming together of national efforts.

Roadmap

d) The survey of the three areas of Grand Challenges in Materials Science, Chemistry and Nanoscience suggests that specific projects need to be sharpened up and some selection between them made. We suggest, in parallel with initiating the construction of the supercomputing facility, an open call and workshop of experts for this, probably best under the auspices of CECAM which is already a European organisation. This includes the selection of appropriate software, which in this field tends to be somewhat focused on particular types of application.

e) It is important to start early to adapt or write software for the particular projects chosen, and taking into account the specific architecture of the proposed computer, with access to a smaller but similar machine for development. In some cases a novel way of organising the computations is needed to get the full advantage of the supercomputer, for example depending on the nature of the communications in the computer. It is essential to fund from the Euro HPC project the special support staff for this work, and subsequent adaptation to evolving computer architecture, because 50 years of experience has shown that research funding agencies will hardly ever fund code development: they just want exciting new science without putting in the necessary infrastructural basis.

f) A small number of special workshops could focus on particular hurdles in the current methodology, where it would be really valuable for supercomputing to overcome them and where there are ideas in the wider research community for how one might make progress. Such discussion might lead to specific code development projects, which again should be funded through the EuroHPC initiative for the reason given above. One example is in Molecular Dynamics, where codes that can harness thousands of processors efficiently could extend simulation of ground state properties by several orders of magnitude and the simulation of small enzymes in biological dynamic catalytic processes. In quantum mechanical simulations the calculation of excited states takes about 100 times more disk space and computation time than groundstate properties: thus only rather simple systems have been considered so far, and it is important to extend them.

g) On the longer term (10 years) it is essential that the ideas and progress in algorithms, codes and types of applications, which bubble up in the wider research community, feed into the High Performance Computing community. In the past, this has occurred naturally through the series of workshops organised by the PSI-K and MOLSIMU groupings, and the HPC initiative should help to keep these organisations alive until a proper infrastructural basis can be found for supporting them.

h) Particular attention should be paid to initiatives aiming to provide software tools which can readily be made available on HPC platforms and on national centres. An example is given by the proposed "European Theoretical Spectroscopy Facility" (ETSF), which aims to provide knowledge transfer in theory, code development, distribution and use for computing spectroscopic properties of materials.

3.8.2 Computational Science Infrastructure in Europe

We focus here on the human aspect of *CyberInfrastructure* (defined here as computational science infrastructure) and what type of infrastructure is needed to keep Europe as a leading area in the world. Specifically we are concerned with the format of a network of expertise which is focused on:

1. Code Development;
2. Researcher Training and Support;
3. Access to expertise, code libraries and other information across this interdisciplinary field.

with some minimal organisational structure to implement these three activities. We briefly elaborate below on these points below:

1. Here *code development* is taken to include the following components:

- The whole process from a researcher (often a doctoral student or young postdoc) initiating an algorithm for a new type of simulation (for example) to its incorporation in a generally applicable form in widely disseminated codes. It is important to recognise that while the continuing growth in computer power certainly has a major impact on computational science, by far the greater advances are due to algorithmic developments;
- Maintenance of codes that may contain several hundred thousand lines of Fortran as new advances have to be incorporated from diverse directions;
- Code portability and optimisation for new machines, particularly with novel architectures;
- Interfacing with other codes;
- Incorporating new computational developments such as GRID, middleware, sophisticated databases, metadata, visualisation and the use of different types of architecture for different purposes;
- All types of code, from the large workhorses already mentioned, to a toolbox of simple basic codes for researchers to access as platforms for developing new directions;
- Resolving within the community issues of intellectual property rights arising from such a multi-stage process of code development.

2. *Researcher training and support* is very important and a necessary part of *Cyberinfrastructure*. The need for training is an inherent consequence of the rapid development of the field and the very sophisticated nature of much of the methodology, including numerous approximations, tricks and short-cuts to make the simulations feasible. In many areas it is only in the simplest routine applications that one can use the code as a 'black box' without expert steering. Young researchers having been trained, and code users generally, need continuing expert support and personal contact as research priorities change and codes evolve.

3. *Better code libraries, databases,* input and output standardisation etc. are needed, with the means of access to all sorts of information and personal expertise through websites, newsletters and email lists.

Clearly some organisational structure is needed to implement the three activities described above. This should be on a European level because one country is too small a unit for efficiency and effectiveness. We envisage that the cyber-infrastructure has to be largely managed by the research community itself in each particular field because the circumstances vary so widely across the sciences. But of course this will be with some help from major computer centres and/or European organisations such as CECAM, EMBL etc. These can provide a permanent hub and a home with scientific and organisational support for a particular research community, as well as some technical help. An example from the UK is the CCLRC Daresbury Laboratory and the Collaborative Computational Projects that have been running for over 20 years.

4 DOMAIN AREAS

4.1 WEATHER, CLIMATOLOGY AND EARTH SCIENCES

4.1.1 Summary

There are few more critical areas of understanding that require research than the societal and economic realities that arise from the effect of climate change and weather and the geology and seismology of our planet. Research in the fields of weather, climatology and earth science is of key importance for Europe, for the following reasons:

- Informing and supporting preparation of the EU policy on environment and greenhouse gas emission reductions
- Understanding the likely impact of the natural environment on EU infrastructure, economy and society
- Enabling informed EU investment decisions in ensuring sustainability within the EU and globally.
- Developing civil protection capabilities to protect the citizens of the EU from natural disasters
- Supporting through the EU and ESA joint initiative on Global Monitoring of Environment and Security

The following sub-paragraphs summarise the key motivations

Climate change

The field of climate change research is developing rapidly. However, whilst there is great confidence in the fact that climate change is happening, there remain uncertainties. In particular there is uncertainty about the levels of greenhouse gas emissions and aerosols and, perhaps even more significant, are uncertainties on the degree of warming and the likely impacts. These latter uncertainties can only be reduced by increasing the capability and complexity of 'whole earth system' models that represent in ever increasing realism and detail the scenarios for our future climate. These uncertainties allow the possibility of political biases in the scientific debate: decision-makers need therefore to access scientific expertise and modelling resources that provide the highest level of certainty and credibility.

Oceanography and Marine Forecasting

Our understanding of the oceans is critical to our ability to understanding and predicting climate change. Ocean modelling is an integral and core element of the 'whole earth system' model outlined above and is crucial to reducing uncertainties in climate change and its impacts. Beyond its role in climate, oceanography is also important in other ways and there remain specific challenges. Mankind's activities, including supply of food and energy, transport of goods, etc., are exerting a continuously increasing stress on the open and coastal oceanic environment, a stress that needs to be evaluated and controlled in order to preserve the ocean integrity and available resources. There is a need to be able to build the most efficient modelling and prediction systems to study, understand and predict ocean properties and variations at all scales, and develop economically relevant applications to inform policy and develop services for government and industry.

Meteorology, Hydrology and Air Quality

Weather and flood events with high socio-economic and environmental impact may be infrequent, but the consequences of occurrence can be catastrophic to those societies and Earth Systems that are affected. There is of course a linkage to climate prediction and climate change impacts, if severe meteorological and hydrological events are to become more frequent and/or more extreme. Predicting these low frequency, high impact events a few days in advance with enough certainty and early warning to allow practical mitigation decisions to be taken remains difficult. Understanding and predicting the quality of air at the earth's surface is an applied scientific area of increasing relevance. Air quality can cause major environmental and health problems affecting both industrialized and developing countries around the world (e.g., respiratory difficulties, especially for sensitive people,

adverse effects on the flora and fauna, …). Here also advanced real-time forecasting systems are necessary tools for allowing early enough warning and practical mitigation in the case of pollution crisis.

Earth Sciences

Computational challenges in earth sciences span a wide range of disciplines and have significant scientific and social implications, such as the mitigation of seismic hazards, treaty verification for nuclear weapons, and increased discovery of economically recoverable petroleum resources and monitoring of waste disposal. In the realm of seismic hazard mitigation alone it is well to recall that despite continuous progress in building code, one critical remaining step is the ability to forecast the earthquake ground motion to which a structure will be exposed during its lifetime. Until such forecasting can be done reliably complete success in the design process will not be fulfilled.

All these areas of expertise require increased computing capability in order to provide **breakthrough science**. A programme of provision of Leadership-class computational resources will make it increasingly possible to address the issues of resolution, complexity, duration, confidence and certainty, and to resolve explicitly phenomena that were previously parameterized. Each of the challenges represents an increase at least by a factor of 100 over individual national facilities currently available. A large number of the numerical models and capability-demanding simulations described below will lead to operational applications in other European centres, national centres and in industry.

Societal benefits

Natural disasters claim hundreds of thousands of lives annually and cause vast property losses. To what extent anthropogenic climate change will lead to an increase in occurrence and severity of extreme events and natural disasters is one of today's most important and challenging scientific questions. The countries that will have access to the highest performance in computing will be able to perform experiments which will become the references for future scientific assessments and associated political decisions. Even though Europe has a world-class expertise in climate, oceanography, weather and quality, and Earth sciences modelling issues, European scientists may loose their current prominence if they cannot access the most powerful computing systems.

The economic benefit to society of quantifying the certainty and impact of forecasts, on whatever timescale, is enormous. By providing probabilistic results for impact models to agencies involved in assessing impacts of extreme events or climate change adaptation, mitigation strategies can be developed and the impacts constrained. The societal benefits range from mitigation of high-impact weather by having a more accurate and timely weather nowcasting system to the impact on society from a better air-quality forecast with regard to health, traffic (e.g. fog and rain), agriculture (e.g. ozone influence), etc. Changes in the chemical composition of the atmosphere bear consequences on health and agriculture which both affect the state of the economy.

International back ground

Europe has been able to establish a scientific community that has played a leadership role in climate research, air quality and meteorology and has effectively contributed to international assessments. These have provided the scientific basis for national and international agreements. The European community is already well organised through European projects and collaborative research. The accuracy and details of earth system research and predictions of climate change have been severely limited by the computing resources available to the community. It is to address this problem that Japan has developed a large supercomputing facility, the Earth Simulator in Yokohama, dedicated primarily to Earth system science. Major plans to enhance by orders of magnitudes the computing resources in support of climate research have also been formulated in the United States ("establishing a petascale collaboratory for the geosciences", http://www.isse.ucar.edu/HP_rick/NSA/petascale-science.pdf, 2005).

4.1.2 Challenges in Climate and Earth System research

4.1.2.1 Motivation

There is a vital need for High Performance Computing in order to be able to predict the future evolution of the climate, and answer the key questions for society about the impact of increased greenhouse gases on global warming. Even though there is no doubt that climate is vulnerable to mankind's activity, many questions remain unsolved at the quantitative level. There is a need to better qualify and quantify uncertainty, estimate the probability of extreme events and regional impacts, quantify the feedbacks between climate and biogeochemical cycles such as carbon dioxide and methane, identify the impacts of climate change on marine and terrestrial ecosystems and on societies. All these questions are strongly linked to the amount of computing power available since they ask for increased model resolution, large number of experiments, an increased complexity of earth system models, and longer simulation periods compared to the current state of climate models. Sustained computing power of the order of 100 Tflops to 1 Petaflops or more is required for Europe to maintain its scientific weight in climate change research and international post-Kyoto negotiations.

4.1.2.2 Challenges: description and state of the art

Key questions facing climate change research can be summarized in four key challenges:

1. **The need for very high resolution models for understanding and quantifying prediction of extreme events and for the regional assessment of the impact on society and economy of climate change:**

   The structure and intensity of an extreme climate event depend on complex nonlinear interactions between large scales (several 100 kms) and smaller scales (down to a few kms). Model simulations are required that are able to represent both modification of the larger-scale, global, state (inside which extreme events are developing) and the fine-scale temporal and spatial structure of such events (storms, cyclones, intense precipitation, etc…). Current global climate models have typical grid spacings of hundred kms and are limited in their capacity to represent processes such as clouds, orography effects, small-scale hydrology, etc. High-resolution global models are therefore needed to improve our predictions and understanding of the effect of global warming on high-impact weather events - on seasonal, decadal and century timescales. Concomitantly, the scientific and mathematical description of clouds, of orographical effects and small-scale hydrology will evolve as smaller scales are accessible.

   Another issue is related to the simulation of regional-scale climate features, of crucial importance for assessing impacts on society and economic activities (farming, fisheries, health, transportation, etc), and for which improved regional models, embedded in global climate models, are necessary. Such regional models are also calling for spatial resolution of a few kms. Increasing model resolution down to 1km for regional models or 20 kms for global climate models require increases by factors of 100 to 1000 in computing power compared to the current state.

2. **The need for current climate models to move towards Earth System Modelling;**

   It has become recognized that models must include more sophisticated representations of relevant processes and subsystems, which are important for long-term climate change. In the last few years biological and chemical processes have begun to be included in long-term simulations of climate change, albeit in a simplified way. In addition to the value of being able to predict changes in vegetation and atmospheric composition, it turns out that these additional processes can have quite a marked effect on the magnitude of climate change. For example, the European modelling groups were the first to show that coupling climate to the global carbon cycle accelerates climate change. However, the carbon cycle itself is intertwined with other biogeochemical cycles, such as the nitrogen cycle which will have to be included. Moreover, other processes, such as aerosols and their indirect effect on clouds or interactive vegetation, still need to be better accounted for.

Increase of computing power by a factor of 2 to 10 is required to better account for the complexity of the system. This complexity will need to be included in medium to high resolution models and therefore require a factor of 100 to 10 000 times the current power.

3. **The need for quantifying uncertainty;**

Climate models are inherently imperfect owing to physical processes that are either not completely understood or yet to be adequately represented because of limited computer power. The consensus approach to solving this and related problems is to assume that the uncertainty can be estimated by combining an ensemble of model experiments. Moreover, running ensembles of climate integrations allows accounting for the chaotic nature of climate.

The different scenarios of emissions of greenhouse agents also warrant the analysis of a set of experiments in order to probe the future course of climate. Quantifying uncertainty of future climate change is a key challenge and will require: to run the European climate models in a coordinated way, to investigate the sensitivity of results to parameters and/or initial state. Ensemble experiments are computationally expensive – a factor 10 to 100 for each experiment – but there could be enormous economic benefit as well as an advance in our understanding from improving the reliability of models or estimating uncertainties in forecasts.

4. **The need to investigate the possibility of climate surprises**

In a complex nonlinear system such as the Earth system, minute actions could cause long-term, large-scale changes. These changes could be abrupt, surprising, and unmanageable. Paleoclimatic data indicate the occurrence of such rapid changes in the past. For example, it is important to determine if there are thresholds in the greenhouse gas concentrations over which climate change could become irreversible. The Atlantic thermohaline circulation (THC) might undergo abrupt changes as inferred from paleo-records as well as from some long simulations of future climate. The possible climatic consequences of such a slowdown in Atlantic THC are still under debate. Surprises may also arise from ice-sheet collapse and large amount of fresh water in the ocean.
To investigate such risks requires long simulations on future and past periods with medium to high resolution and various degrees of complexity. A factor 10 to 1000 is required.

4.1.2.3 Roadmap

Climate and earth-system research need both larger resources at the national level and an access to a European machine providing at least 100 times higher performance than the available national level. Such a European facility will be devoted to specific numerical experiments with the highest resolution and complexity as required by the above challenges. Fully accounting for model resolution, number of experiments, complexity of the system and need for long integrations would require an increase of power reaching 10^5 to 10^6, which is completely impossible by 2008 but which indicates that a sustained effort over a longer time is needed. However, an improvement by a factor of 100 with respect to the actual (2006) European situation, will allow significant progress toward our understanding of the above challenges and deliver significant scientific breakthrough.

The requirements for a European computing centre for climate is supported by the "European Network for Earth System Modelling" (http://www.enes.org) which has emphasized a need for at least 60 Tflops sustained as soon as 2008 with a need for a renewal of investment at every 4 to 5 years at least during 10 to 15 years.

4.1.3 Challenges in Oceanography

4.1.3.1 Motivation

Progresses in the field of ocean sciences is strongly linked to the computing power available since they ask for increasingly high model resolutions, large number of experiments, and increasing complexity of the ocean system models. Operational oceanography is a quite novel and fast growing sector of

activity which concerns coastal water quality, fisheries and marine ecosystems, off-shore, military, transport, etc. The development of ocean sciences follows an arduous way, largely because the ocean is hardly observed. Reasons for that are the difficulty for humans to penetrate in the ocean, and the property of salt water to absorb rapidly electromagnetic radiation and light which limits remote sensing observations to the ocean surface. Therefore, understanding the ocean physical and biogeochemical functioning, predicting the evolution of ocean characteristics and of major marine ecosystems at all relevant scales strongly relies on our ability to assimilate available data and to model accurately the processes at stake.

Just like eddies in the atmosphere, ocean eddies have a fundamental role in setting the mean circulation, but also in transporting heat and properties poleward across frontal structures. We also know from observations that high eddy variability exists near boundary currents, and close to complex bottom topography like continental shelf breaks, ridges, seamounts, straits and sills. In addition, the evolution of the ocean biology is clearly linked to the sub-mesoscale vertical motions associated with these mesoscale features. There is thus urgent need to build on recent advances in ocean modelling and to tackle these questions with high resolution models.

4.1.3.2 Challenges: description and state of the art

1. Toward very high resolution models

Not resolving ocean eddies would imply that important elements of the circulation are not resolved, that govern processes (like deep water formation, inter-basin exchanges, or shelf water export) and thus control the circulation strength and variability. An important challenge concerns the possibility to run realistic ocean models at regional and global scales, at resolutions high enough to insure dynamical consistency over a wide range of resolved scales (from eddy to global, from day to decade). The fact that eddies are only a few tens of kilometre in diameter in the ocean puts stringent demands on the horizontal and vertical grid resolution of models which is required to properly simulate ocean eddy statistics (e.g., eddy energy, size and location of eddies, energy cascade due to eddies), and the constraint on the flow of small topographic features.

At the current time, only a few models reached the resolution where eddies seemed to be marginally resolved. The state of art in ocean modelling includes numerical simulation with a 10 km resolution on a global scale. Such resolution is necessary to resolve mesoscale eddies, whose impact on the large-scale circulation is determinant. The next step is to represent scales much smaller than 10km in large domains (i.e. at a basin or global scale). Significant impacts of O(1km)-scales on the large scales have already been demonstrated by process-oriented studies. For the high latitudes, where convection dynamics controls the thermohaline circulation, this impact has been known for a long time.

2. A better estimate of sink and source of carbone by the ocean

Among the greatest challenges in ocean research today are the simulation and understanding of the recent decadal and inter-decadal global ocean variability (including the carbon cycle), because this understanding has an immediate impact on the attribution of recent climate change to anthropogenic causes, as well as on our ability to perform long-term climate predictions. Indeed our knowledge of the global carbone budget is limited by our ability to assess properly the different sources and sinks between the open ocean and the coastal ocean. This thus requires running ocean carbon model with high resolution to represent these small scale phenomenons. Coastal oceanography, because of its coupling to marine biology and sediment transport, is aiming at resolution of 1 km or better over significantly large areas for periods of decades to study the evolution of marine ecosystems in a changing environment.

3. A better understanding and monitoring of marine ecosystems

Another great challenge in ocean sciences is the simulation and understanding of the evolution of regional marine ecosystems over the last decades and their sensitivity to a changing environment. These would represent great advances in the understanding and the monitoring of the marine resources. This requires running models that will be able to simulate biochemical cycles and

"blooms" accurately, and to understand the links between halieutic resources and fine scale ocean circulations. In mid-latitude ecosystems are strongly affected by the vertical eddy velocity that acts as a pump. High resolution is essential, including when needed grid refinement to selectively regional grid refinement (up to 1 km or 1/100e degree) within a global model 5km configuration. This will allow to resolve explicitly the dynamics of specific oceanic regions that are critical for the global oceanic circulation and thus obviate the need for parameterizations (in those regions) that are found to be inadequate by ad hoc process-oriented studies.

4.1.3.3 Roadmap

Running ocean models with sufficient resolution in the horizontal and vertical is, at present, beyond any existing capability, except on areas of small extent and short calculations. Since an increase by a factor of 2 in resolution requires an increase of computer power of 23, resolution alone demands a hundred to thousand-fold increase in computer power. However, we need at the same time to make significant progress in data assimilation, because accurate ocean model runs require accurate initial conditions fields, accurate forcing conditions and accurate internal model parameters. As an example, a global ocean model with a 5km resolution is estimated to require 120 Teraflops sustained and can be run only on a scalar machine with 2 Petaflops of peak performance. Moreover, it would be important to include grid refinements, up to 1 km scale, to resolve explicitly the dynamics of specific oceanic regions that are critical for the global oceanic circulation and thus obviate the need for parameterizations (in those regions) that are found to be inadequate by ad hoc process-oriented studies. Including grid refinement (up to 1 km) in some specific regions will bring to a total requirement of at least 200 Teraflops sustained. (i.e., this would correspond to a dedicated scalar machine with a peak performance of 3-4 Petaflops!)

4.1.4 Challenges: Weather and Air Quality

4.1.4.1 Motivation

The mitigation of high-impact weather achievable by having a more accurate and timely weather now casting system requires that the comprehensiveness of the various models used for describing the development of the atmosphere in the near future will have to be enhanced dramatically. The resolution of the numerical weather prediction models will have to be increased to about one km horizontally to resolve convection directly. In addition, a probabilistic approach will have to be taken to arrive at meaningful warning scenarios.

Similarly, one of the most far-reaching developments resulting from the enhanced capabilities in air quality modelling is the possibility of performing forecasts. Atmospheric dispersion models that aim to simulate the physical and chemical processes in the atmosphere have been used for ozone forecasting. However, such models are unsuitable in many operational settings because they require significant computational efforts and a large volume of input data. The ability to forecast local and regional air pollution events is challenging since the processes governing the production and sustenance of atmospheric pollutants are complex and often non-linear. The availability of increased computational power and the possibility of accessing scattered data on-line with the help of a grid infrastructure, coupled with advances in the computational structure of the models now enable their use in real-time air quality forecasting. Furthermore, this may contribute to the GMES European initiative, which has stated as some priority objectives.

These challenges are being tackled already today, albeit on much smaller scale, because the computing power is about three orders of magnitude smaller than required for the complete solutions. Today it is only possible to solve the problems for a limited number of variables in a reduced area. It is, in principle, known what needs to be done in future but the resources are not available yet.

4.1.4.2 Challenges: description and state of the art

1. **The need for very high resolution atmospheric models and associated forecasting system**

   The resolution of models will have to be increased both in time and space to be able to resolve physical processes directly which today are still parameterised. To arrive at meaningful results, this work also entails the evaluation of the error growth due to the uncertainty of the initial conditions. One option is the computation of multiple scenarios with initial conditions varying within the error space. Furthermore, the data gathered by new, high-resolution observing systems, either space or ground based, need to be assimilated using new techniques such as 4D-Var. Furthermore, the I/O rates of the applications will be in the order of 5 GB/s for the duration of the runs, resulting in files of more than 6 TB. In order to test these ideas, a pre-operational trial of a complete end-to-end system needs to be carried out, checking whether the integrated system consisting of data acquisition, data assimilation, forecast run and product generation can be handled in a sufficiently small time period to allow future operational deployment. Similarly, sets of ensemble forecasts will have to be run as a pre-operational test cases.

2. **High-resolution air quality forecast and cloud-dust-interaction resolving models. Description and state of the art**

   A reliable pan-European capability of air pollutant forecasting of air quality in Europe with high resolution (1 km) becomes essential in informing and alerting the population, and in the understanding of when and why episodes of air pollution arise and how they can be abated. It is well known that the accuracy of an air quality forecasting system depends dramatically on the accuracy of emission data. It is thus timely to develop a system using the four-dimensional variational analysis (4DVAR) to improve simulations of air quality and its interactions with climate change. 4DVAR is an improved way of combining observations valid at different times (satellite data, radiosondes, ground observations, aircraft measurements, photometer data, model data) with background fields from a previous forecast to create a starting point for a new forecast. 4DVAR is computationally very expensive (considerably more expensive than the forecast itself), and requires also a lot of memory.

   In order to be able to provide high-resolution air quality forecasts, an important unresolved question is the inclusion of cloud-resolving models that take into account cloud physics and the role of aerosols in modifying clouds, precipitation and thermal atmospheric structure. Under natural conditions, dust and atmosphere interactions generate complex feedbacks: increased dust load modifies the thermal and dynamic structure of the air; modified atmosphere furthermore changes conditions for the dust uptake from deserts, and so on. In current generation models these processes are still highly simplified. Several efforts in this area must be progressed with the help of supercomputers.

3. **Develop pan-European short range weather and air quality modelling system.**

   Weather and air quality models are commonly used in the forecast mode. Further improvements are still needed which require more computing power. Such improvements include: the representation of clouds and small scale processes, the coupling with the biosphere, atmosphere, aerosols, clouds, chemical reactions in the environment, ocean surface, data assimilation, and the estimation and computation of different sources of emissions. In addition to the requirements of sections 1 and 2, this requires including more physics and chemical species in the models. Short range and very high resolution models need observational data and subsequent data assimilation. This assimilation of new data, in particular radar data for very short forecasts, needs to be resolved, also as a real-time application. This increases the computing requirements by a factor of 10.

4.1.4.3 Roadmap

The national computing resources for solving the challenges outlined above are expected to grow over the coming years. This growth will allow a gradual increase in complexity and resolution of the

various models used and will result in narrowing the gap between what is possible at any given time and what is required. The work will be carried out by established scientific communities consisting of the European national weather services organised in the European Meteorological Infrastructure (EMI), collaborating universities and scientific centres. The European partners have collaborated over many years, often supported by EU projects.

For instance, air-quality forecasts for the area of all EU member states at the highest feasible resolution (1 km) require extensive computing resources, as each of the applications proposed will require 100 – 300 Tflops/s sustained performance. A complete pan European weather and air quality forecast requires high resolution over a wide area, a complete description of the chemistry and aerosols and of their transport and the coupling with the other component of the Earth's system. This will require > Tflop sustained performances. The work will be carried out by established scientific communities : the most relevant initiative is the GMES project (Global Monitoring for Environment and Security), a joint initiative between the EU and ESA to strengthen the acquisition and integration of high-quality EU environmental, geographical and socio-economic data, that will help improve policymaking from local to global level.

4.1.5 Challenges in Earth Sciences

Because solid-earth processes occur on many different spatial and temporal scales, it is often convenient to use different models. A key issue is to better identify and quantify uncertainties, estimate the probability of extreme events through simulation of scenarios and exploration of parameter spaces. For some problems, the underlying physics is today adequately understood and the main limitation is the amount of computing power available. For other problems, a new level of computing power is required to advance our understanding of underlying physics where laboratory experiments can hardly address the wide range of scales involved in these systems, e.g. for example modelling and simulating earthquake dynamics rupturing processes together with high frequency radiation in heterogeneous media. The solid Earth community is preparing itself for massive use of supercomputers by the current (re-) organization of some of the communities through large-scale EU projects, e.g. the Marie Curie Research Training Network SPICE (http://www.spice.rtn.org) and the EC-project NERIES (www.orfeus-eu.org/**neries**/**neries**.htm). Similar developments occur in geodynamics (e.g., TOPOEurope) and in geodynamo. The need for Leadership-class computing is illustrated through four major challenges.

4.1.5.1 Challenge: Earthquake ground motion simulation and seismic-hazard

Motivation

To understand the basic science of earthquakes and to help engineers better prepare for such events, scientists need to identify which regions are likely to experience the most intense shaking, particularly in populated sediment-filled basins. This understanding can be used to improve building codes in high-risk areas and to help engineers design safer structures, potentially saving lives and property. In the absence of deterministic earthquake prediction, forecasting of earthquake ground motion based on simulation of scenarios is one of the most promising tools to mitigate earthquake related hazard. This requires intense modelling that meets the actual spatio-temporal resolution scales of the continuously increasing density and resolution of the seismic instrumentation, which record dynamic shaking at the surface, as well as of the basin models. Another important issue is to improve our physical understanding of the earthquake rupture processes and seismicity. Large scale simulations of earthquake rupture dynamics, and of fault interactions, are currently the only means to investigate these multiscale physics together with data assimilation and inversion. High resolution models are also required to develop and assess fast operational analysis tools for real time seismology and early warning systems.

Description:

Earthquakes are a fact of life in Europe and all around the world. Accurate simulations must span an enormous range of scales, from meters near the earthquake source to hundreds of kilometers across the entire region, and time scales from hundredths of a second—to capture the higher frequencies which

have greatest impact on buildings—to hundreds of seconds for the full event. Adding to the challenge, ground motion depends strongly on subsurface soil behaviour. While providing much useful information, today the most advanced current earthquake simulations are not generally capable of adequately reproducing the observed seismograms. The likely reason is that these models are based on a number of assumptions made largely to reduce the computational effort. There is an urgent need to enhance these simulations and to improve model realism by incorporating more fundamental physics into earthquake simulations. The goal is: (1) to extend by a factor 10 the spatial dimensions of the models; (2) to increase the highest resolved frequency above 4Hz (for structural engineering purposes) implying a 64-fold increase in computational size (size scales roughly as the cube of the resolved frequency); (3) to move to more realistic soil behaviours implying at least a two order of magnitude increase in computational complexity; (4) to incorporate a new physics-based dynamic rupture component at 100 m resolution for realistic wave radiation, and near field risk assessment, implying at least an order of magnitude increase in computation; (5) to invert for both the earthquake source and the geological parameters that necessitates repeated solutions of the forward problem leading to an increase of one to two order of magnitude in computations; (6) to perform probabilistic modelling of seismic events for quantifying uncertainties and exploring earthquake scenarios which implies a 10-50 times increase in computation. These improved simulations will give scientists new insights into where strong ground motions may occur in the event of such an earthquake, which can be especially intense and long-lasting in sediment-filled basins.

Relation to the state of the art:

State of the art can be split in two categories. First, problems, for which the underlying physics is adequately understood and the challenges arise mainly from computational limitations, e.g. simulation of anelastic wave propagation in strongly heterogeneous geological media remains a computational challenge for the new Petaflops technology. Second, problems for which high performance computing resources are required to advance scientific understanding – modelling earthquake dynamic rupturing processes together with high frequency radiation in heterogeneous media is an example of this type of problem. Fully coupled extended earthquake dynamics and wave propagation will remain a grand challenge problem even with the next generation computers.

Roadmap:

Today Gigaflop computing allows simulation of wave propagation up to 1 Hz in small geological basins (30km x 30km x 30km) assuming linear constitutive models and moderate velocity contrasts. State-of-the-art Teraflop simulations with 1.8 billion grid points require 240 processors on the expanded 15.6 Teraflops Datastar supercomputer, 4 days computing with complex intensive data movement between computer nodes, disk and archival storage elements (more than 10 Terabytes of output data), in order to simulate propagation in heterogeneous geological basins of dimensions 300 km x 300 km x 80 km at 200 m wavelength resolution. These Teraflops (0.8-5 sustained Teraflops, 25-35 % of the peak performance, with 80-90 % parallel scaling performance) simulations are providing new insights into where strong ground motions may occur in the event of an earthquake. However these simulations remain limited: the resolved frequencies are still too low for the seismic engineering applications (> 4 Hz); important non linear effects and complex soil behaviours are not yet taken into account; earthquake sources remain simplistic; and uncertainty quantification and data assimilation are not yet reachable. Enhancing the resolution and the physics, making inversion of extended earthquake sources and seismic parameters, quantifying the uncertainties through strong motion scenarios, will push these simulations into the realm of Petaflop computing (~ 0.8-1.5 sustained Petaflops). A key requirement will be the provision of large scale end-to-end data cyber infrastructures to handle, analyze and visualize Petabytes of simulated data for storage. Visualization of very large data sets will be a related important challenging problem.

4.1.5.2 Challenge: High resolution imaging techniques

Description:

The capacity for imaging accurately the Earth sub-surface, on land and below the sea floor, is one of the challenging problems that has important economical applications in terms of resource

management, identification of new energy reservoirs and storage sites as well as their monitoring through time. As recoverable deposits of petroleum become harder to find the costs of drilling & extraction increase; the need for more detailed imaging of underground geological structures has therefore become obvious. Recent progress in seismic acquisition related to dense networks of sensors and data analysis makes it now possible to extract new information from fine structures of the recorded signals associated with strongly diffracted waves. This increase in data acquisition is also inportant for risk mitigation. Imaging accurately seismic rupture evolution on complex curved faulting system embedded in heterogeneous medium, or time-lapse monitoring of volcanoes, proceed in similar fashion.

Motivation:

Seismic imaging of the earth sub-surface has important implications in terms of energy resources and environmental management. While deep-ocean (1000m-2000m) fossil energy resources are going to be under extraction, investigations in complex tectonic zones as foothills structures are crucial because these zones are expected to host reservoirs of future economical interest. With the advent of high resolution and large dynamic instrumentation, the challenge is now to exploit fully fine details of the recorded signals going beyond the first arrivals and exploring late arriving signal associated with strongly diffracted waves. This will open new perspectives in very complex geological settings, as well as the capacity of monitoring through time waste disposals or reservoirs during their exploitation, as well as volcanoes.

Relation to state of the art:

Differential and time lapse seismology, migration, and correlation methods are today explored in order to extract this detailed information. In these imaging techniques, only adjoint methods related to linearized techniques have to date been investigated, and back-projection is the mathematical tool for the image reconstruction. The adjoint methods only allow for local analysis of the resolution and uncertainties. Semi-local analysis would require simulated annealing/genetic algorithms leading to a drastic increase in computer resources we can not yet foresee, without mentioning exhaustive inspections of model space with importance sampling strategy. Because thousands of forward problems should be achieved in an iterative optimization scheme related to the number of sources and receivers, one must investigate techniques for solving efficiently in a combined way these forward problems altogether. Moreover in the forward model, new models must accurately simulate complex wave propagation phenomena such as reflection and diffraction in heterogeneous media with high impedance contrasts, or diffraction by rough topographies at the surface of the earth or at the bottom of the sea, at very high frequencies (10 Hz-40 Hz) where complex attenuation is expected. The bridge between deterministic estimations and probabilistic approaches should be clearly identified and will justify the demanding task of performing wave propagation modelling.

Roadmap:

The time formulation, e.g., wave propagation in time, allows handling 3D imaging problems at the expense of computer time using present-day algorithm technology. For simulations on boxes of 100 km x 100 km x 25 km, with a 10 Hz content, sustained performances range between 1 Teraflops and 10 Teraflops. Moving to more powerful resources will increase the size of the box and/or the maximum frequency. Petaflop computing will improve seismic imaging resolution by using thousands of recorded seismograms. Present-day numerical algorithm know-how should be improved for handling expected biases coming from our initial guess of the Earth image. Petaflop computing will also allow us to tackle the important problem of uncertainties by making use of repeated forward modelling. The frequency formulation, e.g., Helmhotz equation, allows the handling of 2D imaging problems on 1-Teraflop machines. With present-day algorithm technology, matrix decomposition (LU for example) allows efficient image processing by choosing only a few frequencies and by speeding up multi-sources and multi-receivers computations. The 3D imaging problem in the frequency domain is today a challenge both for computer resources and for numerical algorithms. While small models and low frequency analysis are possible on 10 Teraflops machines, the realm of Petaflop computing will provide the possibility of performing such seismic imaging in the frequency domain. This will require access to a large memory/processor ratio and efficient algorithms for direct decomposition of

very large matrices. Moreover achieving load balancing between processors in the frequency domain approach will be a challenge. Unfortunately resorting to iterative methods dims the interest of a frequency formulation compared to the time domain formulation. Seismic Data Processing (SDP) is of paramount importance for imaging underground geological structures and is being used all over the world to search for petroleum deposits and to probe the deeper portions of the earth. Current advances in data acquisitions, multi-component, and multi-attribute analysis have increased the data volume several fold. Processing methods have also changed for high resolution, leading to an increase in the computational effort, which is beyond the scope of actual computer resources. Large data volumes and complex mathematical algorithms make seismic data processing an extremely compute and I/O intensive activity which require high performance computers (1-10 Teraflops sustained) with a large memory.

4.1.5.3 Challenge: Structure and Dynamics of the Earth's Interior

Description:

One of the major problems facing earth scientists is to improve the resolution and the understanding of the earth's interior structure and dynamics. Broadband seismological data volumes, thanks to the federation of digital seismic networks (with European contributions like GEOSCOPE and GEOFON) and the associated distributed data centres like ORFEUS or GEOSCOPE in Europe, is increasing at a faster rate than computational power, challenging both the analysis and the modelling of these observations. So far only a small fraction of the information contained in broadband seismograms is actually used to infer the structure of the earth's interior. Recent advances in high performance computing and numerical techniques have facilitated three-dimensional simulations of seismic wave propagation at unprecedented resolution and accuracy at regional and global scales. The realm of Petaflop computing opens the door to full waveform tomographic inversions that make use of these new tools to enhance considerably the resolution of the earth's interior image. This is a grand challenge problem due to the large number of mesh-dependent model parameters and of wave propagation simulations required during the inversion procedure.

Convection of the solid earth mantle drives plate tectonics and the earth thermal evolution. Mantle convection is dominated by slow viscous creep, involving time scale of hundreds of million of years. Despite the low velocities, the Rayleigh number is of the order of 10^{-7} inducing a quite time-dependent dynamics and small convective scales comparable to the size of the domain. One computational challenge is thus the resolution of convective features of less than 100 km over spherical domain of depth 2900 km and circumference of 40,000 km. Another challenging issue is the resolution of the rapid spatial variations of the physical properties: viscosity is strongly dependent on temperature, pressure and stress, e.g. 6 orders of magnitude with temperature and 2 orders of magnitude with depth. Incorporation of melt-induced compositional differentiation, self-consistent plate-like behaviour (elastic brittle) and composition solid-solid phase change is extremely difficult and computationally demanding. How plate tectonics arise from mantle convection is an outstanding issue.

Motivation:

Seismology is the unique method that can probe the earth interior from the surface to the inner core, as well as its external coupling with the atmosphere and the oceans. Improving the capability to enhance the quality of 3D tomographic images of the earth's interior, with a resolution of the thermal and chemical heterogeneities lower than tens of kilometres, using the continuously increasing data sets of broadband seismological records, is today essential to improve core-mantle dynamical models and our knowledge of the earth's physics. This is also an essential step in order to improve the imaging of earthquake rupture processes using both regional and tele-seismic seismological observations.

Solid Earth internal dynamical processes often take place on scales of tens to million of years. Even with the most advanced observational systems, the temporal sampling of these phenomena is poor. In order to understand these systems, simulations must be carried out concurrent with observations. Mantle convection provides the driving force behind plate tectonics and geological processes that shape our planet and control the sea level. Realistic models of thermo-mechanical mantle convection in 3-D spherical geometry are required to better assimilate mineral physics and seismology

information into the deep earth dynamics. The short-time scale dynamic behaviour will serve as the monitor for stress build-up that loads seismically active regions.

Relation to state of the art:

Numerical 3D simulation of wave propagation at regional and global scales has been achieved recently at unprecedented resolution and will still be improving in the next few years. Using these new developments for non linear arrival time and waveform inversions will lead to a revolution in global and regional tomography in the next decade. Even in the realm of Petaflop computing, this seems an extraordinary computational challenge when facing the hundreds or thousands of model parameters involved here. Taking advantage of the fact that adjoint calculations and time reversal imaging are quite straightforward in seismic inverse problems opens new doors for efficiently computing the gradient of the misfit function and to developing new scalable algorithms for seismic inversions.

Three dimensional numerical simulations of mantle convection with both chemical and thermal buoyancies are today performed both in Cartesian and spherical shell geometries. Numerical simulations that include both melt-induced compositional differentiation and self-consistent plate tectonics-like behaviour have been performed only in two dimensions and in small 3D Cartesian geometries. Incorporating melt-induced compositional differentiation, self-consistent plate-like behaviour (elastic brittle) and composition solid-solid phase changes in high resolution spherical shell models is today a challenging problem that can only be addressed in the realm of Petaflop computing.

Roadmap:

Global wave simulation of body wave phases that explore the earth's core is today a Petaflop challenge problem. This requires global wave simulations, at periods of 1s or less and space resolution at wavelengths of tens kilometres, in 3D anelastic earth models including high resolution crustal models, topography and bathymetry together with rotation and ellipticity. Today front end global seismology simulations run at wavelengths of tens km and typical periods of 5 seconds in 3D earth models with 5 billion of points (2.5 terabytes memory) on 1944 processors of the 38.6 Teraflops Earth Simulator in Japan with a sustained performance of typically 5 teraflops (30% of the peak performance, 93% vectorization ratio). The next generation forward global wave simulations will be simulations at 3 seconds periods and will require 20 Terabytes memory and 0.25 Petaflops sustained performance. Simulations at 1 second periods will require 100 terabyte of memory and 1 Petaflops sustained performance. For these applications, storage is not a challenging issue since only the synthetic seismograms are stored for comparing with the actual observations and eventually inversion. The challenge will be to go for the inversion of complete waveforms using these 3D wave propagation simulation models. However, this will lead to at least one order magnitude increase in the computational requirements.

4.1.5.4 Challenge: Generation of the Earth's magnetic field

Description:

Named one of the enigma's of natural sciences, the extremely involved magnetohydrodynamic simulations of the core-dynamics and the associated external magnetic field are essential to progress in this field. The past seven years have seen significant advances in computational simulations of convection and magnetic field generation in the Earth's core. Although dynamically self-consistent models of the geodynamo have simulated magnetic fields that appear in some ways quite similar to the geomagnetic field, none are able to run in an Earth-like parameter regime because of the considerable spatial resolution that is required.

Motivation:

The history of the Earth's magnetic field variations is engraved in the frozen-in field directions found in most volcanic rocks on Earth (e.g., oceanic crust generated at the spreading ridges). Many of those observable directions are used to derive plate motions of recent times and it is important to understand the constraints of these estimates particularly in times with frequent reversals. On a shorter time scale it is important to understand the phenomenology of magnetic field reversals, not only because

currently the field strength is decreasing steadily with some likelihood of a reversal over the next few thousand years. Understanding the generation of the Earth magnetic field is not only crucial for geophysics, it has strong implications in astrophysics in understanding the magnetism of planets and stars. Besides, geodynamo is one of the challenges of non-linear physics.

Relation to state of the art:

While relevant programs are implemented in parallel, much higher resolution is required to be comparable with the natural conditions. In addition, many realizations are necessary to create stable results for highly-nonlinear processes with strong dependence on initial and boundary conditions. No global convective dynamo simulation has yet been able to afford the spatial resolution required to simulate turbulent convection, which surely must exist in the Earth's low-viscosity liquid core. They have all employed greatly enhanced eddy diffusivities to stabilize the low resolution numerical solutions and crudely account for the transport and mixing by the unresolved turbulence. A grand challenge for the next generation of geodynamo models is to produce a simulation with the thermal and viscous (eddy) diffusivities set no larger than the actual magnetic diffusivity of the Earth's fluid core, while using the core's dimensions, mass, rotation rate and heat flow. This is a grand challenge for the future that requires significant improvements in numerical methods and spatial resolution. In addition, many realizations are necessary to perform ensemble modelling and forecasting.

Roadmap:

The first models were able to simulate magnetic fields quite similar to the geomagnetic fields on Gigaflop technologies with subtle compromise leading to the modification of the equations and a set of parameters far from the earth conditions. In 2005, results for the same set of parameters have been obtained without modifying the physics using 512 processors of the 38.6 teraflops Earth Simulator for 6500 hours. For the first time a dynamo was obtained with a small viscous moment compared with the magnetic one. The challenge is now to achieve the relevant balance for the dynamics of the Earth's core, for which both moments are vanishing. Massive access to Teraflop computing will allow European researchers to investigate the mechanism of these dynamos (only obtained in 2005) and understand their physical principle. Yet the parameters available on such resources are still off by a factor 1 million from the actual geophysical values. Progress achieved over the last few years clearly indicate that an Earth-like solution (for which both moments vanish) could be reached by decreasing the relevant parameter (controlling viscous effects) by a factor 1000 only. Constructing such an Earth-like numerical dynamo model is therefore only realistic in the realm of Petaflop computing. When such simulations will be available, the critical scientific issue will be to interpret the dynamical models in the frame of dynamo theory. This will require Petabytes of storage to describe the 4D (time and space) magneto-hydrodynamic solution.

4.1.6 Requirements for Thematic Area

Computational dimensioning

The computational requirements for the weather, climate and earth sciences applications discussed in this document have one common feature: the urgent need for access to very large computational resources does not stem from a single aspect such as the need to model a larger number objects or to model at a higher resolution. Currently available compute power restricts these applications in several ways. For example, the envisaged advanced climate studies require simultaneously higher resolutions, a more sophisticated representation of processes and ensemble methods to quantify uncertainty. Very similarly, earthquake and ground motion modelling requires higher resolutions, a more sophisticated representation of the physical processes of earthquake source dynamics and quantification of uncertainties on strong motion scenarios.

The need to improve multiple aspects of the application implies very high computational requirements. The requirements are typically a factor of 1000 above what can be run today on the top computational facilities installed in Europe. In absolute terms, the performance requirements of these applications range from 100+ Teraflops sustained to 1 Petaflops sustained, with some of the applications having even higher longer term requirements.

The ratio between sustained and peak performance varies from application to application; factors of 1:10 for scalar architectures and 1:3 for vector processor based systems are typical values in the fields of weather, climatology and earth sciences. The peak performance requirement is therefore 10 Petaflops when considering a scalar architecture, and roughly a third of that for a system based on vector processors.

Special requirements in architecture, software, storage, grid access

Due to the high internal communication requirements of the applications and the continuous need to modify and enhance the model codes, a general-purpose computing system that offers excellent communication bandwidth and low latencies between all processors is required. For most of the weather, climatology and earth sciences applications the amount of computer memory required is generally not higher than that required by applications in other scientific disciplines. However, studies of the structure and dynamics of the earth's deep interior, and high resolution seismic inversion, will require memory sizes approaching 100 terabytes.

To ensure efficient utilisation of the system, an I/O subsystem that supports high transfer rates and provides substantial amounts of online disk storage (1+ Petabytes) is essential. Such online storage needs to be complemented by local offline storage (10+ Petabytes), to enable inputs and outputs to be stored up to 12 months.

A possible long-term storage strategy would be to for each community to develop its own distributed but shared database system based on data-grid technology. The long-term archive could then be held at national facilities. Most of this archive would be communal data available to other researchers rather than private data. Depending on the community, these archives would hold between 20 to 100 Petabytes of data. To implement a grid-based distributed archive system, high speed network links between the European resources and the larger of the national facilities would be a fundamental requirement.

Such a strong link with national facilities would enable the bulk of the pre- and post-processing to be carried out at these facilities. Equally, visualisation and analysis of model outputs would be possible through these network links.

4.1.7 Expected Status 2009

A number of the codes dealing with weather, climatology and earth sciences are well positioned to make efficient use of a very large facility. These are the codes that have already been ported to and tested on the Japanese Earth Simulator or on BlueGene configurations with thousands of processors. About 6 months elapsed time would be needed to adapt these applications to a new system (unless such a system is based on a radically different architecture). Provided that a smaller but architecturally equivalent development system is made available during 2008, a number of applications can be made fit for production use in 2009.

The European facility should have a highly developed user support team that is onsite. This team would provide mainly expertise for architectural dependent optimisation of codes but would as well develop some understanding of the technical aspects of the main applications. An ongoing skill transfer between the scientific community and this support team would be highly desirable.

The following table details for each field the scientific problems that can be progressed given a certain amount of sustained computing capability. Following the initial porting of some codes to the large European facility, a continuous process of code development and optimisation would follow. This process would involve the scientific community, support teams from the national facilities and the above user support team.

Challenge	1TF sustained	10-100 TF sustained	100-1000 TF sustained	>1 PF sustained
Climate Extreme events, Impacts		Modified regime for extreme events (5 to 10 200-year medium-resolution coupled simulations) →		Tropical phenomena (high-vertical resolution O/A coupling, 10cm, 1°A, 0.5°O)
Climate Earth System Modelling		Increase the number of species/variables (→200)	Oceanic CO_2 absorption (1/12° & biogeochemistry)	Oceanic CO_2 absorption (1/20° & biogeochemistry)
Climate Quantifying uncertainties	IPCC-type « crude » estimations	1. Estimate climate sensitivity with "super-parameterizations" (2D at each grid point) 2. Estimate the range of possibilities Ensemble of 1.000 simulations at today resolution with Earth System models	Quantify the probabilities. Ensemble of 1.000 simulations at increased resolution.	Estimate climate sensitivity with high-resolution NH model (Grid-size of ≈3km)
Climate surprises Slow-down of the Gulf Stream	Preliminary studies have shown such events are possible within the next couple of centuries	Define probabilities : More numerous simulations	Better estimation of the time horizon : Longer simulations Increase resolution	Toward real prediction: Increase spatial resolution (1PF and beyond)
Oceanography	Eddy permitting models at the scale of a basin	Global ocean with 10-5 km resolution, and very high resolution 5-1km on limited domains	Global ocean model with 5km resolution with grid refinement	Very high resolution High resolution ocean coupled to atmospheric models
Weather and air quality	Preliminary studies on limited domains	Cloud resolving models and high resolution air quality forecast on limited domains	1km resolution model for the area of EU member state	1km resolution models, including more processes and coupling (aerosols, land surface, …)
Earth Sciences : Earthquake ground motion simulation and seismic hazard	Ground motion simulation up to 1 Hz in small geological basins with anelastic model and simple sources	Ground motion simulation up to 4 Hz in geological basins, and kinematic finite source, Extended source inversion Terabytes output	Ground motion simulation up > 4 Hz in complex geological basins, strong impedance contrast, non linear surface soil behaviour, kinematic finite tens source probabilistic earthquake scenarios 100 Terabytes output	Ground motion simulation 4 Hz, in large complex geological basins, dynamic earthquake source, non linear surface soil behaviour, Hundreds of earthquake scenarios and stochastic approach Dynamic source and velocity inversion Petabytes output
Earth Sciences: Global wave simulation in 3D Earth models and 3D Global tomography	Global surface and body wave simulations at wavelengths of tens km and typical periods of 5 seconds in 3D anelastic earth models	3D long period elastic wave arrival time tomography : mantle heterogeneities Point source (CMT) and extended earthquake source inversion	Global surface and body wave simulations at 3 seconds (0.25 Petaflops sustained performance.) 3D long periods full wave form tomography using elastic wave : mantle heterogeneities and anisotropy Point source (CMT) and extended source inversion	Global surface and body wave Simulations at 1 second periods (100 terabyte of memory and 1 Petaflops sustained performance). Short period full wave form inversion : mantle and core heterogeneities and anisotropy Extended earthquake source inversion
Earth Sciences High resolution imaging techniques	2D elastic wave equation inversion in the frequency domain	3D low resolution elastic wave equation in the frequency domain 2D high resolution elastic wave equation inversion in the time domain	3D low resolution elastic wave equation inversion in the time domain	3D high resolution elastic wave equation in the both frequency and time domain High frequency attenuation effects and impedance contrasts

4.2 ASTROPHYSICS, HIGH-ENERGY PHYSICS, AND PLASMA PHYSICS

4.2.1 Summary

Astrophysics is a field of research in which high-end supercomputers traditionally play a crucial role, mainly because very often modelling and simulations must replace planned and controlled experiments. The most demanding applications deal with systems and structures which span a large range of different length and time scales, and almost always non-linear coupled systems of ordinary and partial differential equations have to be integrated, in three spatial dimensions and explicitly in time, with rather complex material functions as input. The grand challenges that can be foreseen for the coming years range from the formation of stars and planets to questions concerning the origin and the evolution of the Universe as a whole. Moreover, supercomputer resources are needed to evaluate the huge amount of data expected from future space experiments such as the European Planck Surveyor satellite.

In the context of the standard model of ***Elementary Particle Physics***, quantum chromodynamics (QCD) is considered as the fundamental quantum field theory (QFT) of the strong interaction. It describes the highly non-linear interaction between quarks and gluons which creates bound states like the proton or neutron. Because QCD is a strongly interacting non-abelean (i.e. highly non-linear) theory it leads to a great variety of phenomena, including e.g. topologically non-trivial field configurations, phase transitions, chiral symmetry breaking, confinement and a very complex structure of the physical vacuum. Because of its richness and experimental accessibility QCD is the natural framework to study properties which are potentially shared by theories beyond the standard model. In addition, up to renormalization effects and neglecting the small current quark masses, it is a conformal theory and thus the prime candidate to establish an AdS/CFT duality. The latter postulates that string theories on curved space-times in higher dimensions can be equivalent to conformal quantum field theories in ordinary space-time. For all of these reasons QCD is the topic of intense theoretical and experimental research by a large and truly international community involving large European centers like CERN and DESY and integrating collaboration efforts within such as DEISA and the I3HP hadron physics project. In the future, our understanding of QCD, and possibly other QFTs and string theories, will undergo scrutinizing tests at new experimental facilities like LHC at CERN, FAIR at the GSI or the planned ILC (international linear collider). These experiments not only promise to yield a much deeper understanding of the standard model elementary particles and the forces between them, as well as nuclear forces, but are also expected to discover hints for a yet unknown physics beyond the standard model.

In ***Plasma Physics*** the science and technology challenge raised by the construction of the magnetic confinement thermonuclear fusion reactor ITER calls for a major theory and modelling activity, in order for the numerous basic physics phenomena to be better understood, but also for all the relevant models to be integrated in a series of codes able to reliably predict the behavior of both the plasma discharge and the various components of the device, the so-called ‘tokamak simulators’. Both the success of the experiment and its safety rely on such simulators. The quest to realize thermonuclear fusion by magnetically confining a high temperature plasma poses some of the computationally most challenging problems of nonlinear physics. Fusion plasmas are nearly collision-free, and ab-initio modelling of some of the critical phenomena requires a kinetic description of the plasma with a distribution function depending on three space and two velocity coordinates.

4.2.2 Identification of Computational Challenges

In ***Astrophysics***, six areas of grand challenges are identified for which modelling on supercomputers is essential. These are (from small to large mass and length scales):

- the formation of stars and planetary systems (including key questions such as whether planets form by agglomeration or gravitational instability, and how to explain both the rate of star formation and the initial mass function);

- solar and heliospheric physics (including the nature and evolution of solar flares and coronal mass ejections, and their consequences for the Earth's magnetosphere, communication systems and space missions);
- evolution and explosions of stars (including binary interactions, stellar pulsation and convection, 3D radiation-hydrodynamic models of novae, and supernovae);
- Black Hole physics on stellar and galactic scales (including 3D models of the accretion into and relativistic outflows from Black holes, and gamma-ray bursts);
- formation and evolution of galaxies (including the formation of the stellar population in galaxies of different morphological types, and the dynamics of stars and the interstellar medium);
- cosmology and the formation of large-scale structure (including the modelling of dark-matter structures in a dark-energy dominated Universe, and Hubble-volume simulations of galaxies and clusters of galaxies with small-scale processes such as gas dynamics, star formation and supernova explosions).

In ***High-Energy Physics*** the main goal of 'QCD (or more general QFT) on the Lattice' is to solve a QFT rigorously, i.e. to find solutions for the fundamental theories of matter with a precision which rivals that achievable in experiment and that are free from uncontrolled assumptions. This endeavor has many aspects for which the following list of grand challenge problems for HPC gives only an incomplete illustration. Computationally the solution of these problems requires the most powerful supercomputers available at any given time, at least for the mid-term future:

- hadron structure, i.e. the calculation of generalized quark and gluon distribution functions and wave functions;
- QCD at physical quark masses, i.e., light, heavy-light and heavy hadron spectroscopy, resulting in the determination of the quark masses and predictions of the masses of exotic states;
- CKM-physics and CP violation, i.e., to test the standard model through measuring the mixing and decays of different quark flavours;
- matter at non-zero baryon density and high temperatures, i.e., extending the theory by including correlations, studying finite temperature phase transitions and new states of strongly-interacting matter;
- theories 'beyond the standard model', i.e., GUT, supersymmetry and string theory.

The major computational challenges in ***Plasma Physics*** are coming from the need to understand and model complex plasmas on the kinetic and also the continuum level for the design of future controlled-fusion reactors. In particular this includes:

- fundamental plasma processes (including non-linear instabilities in plasmas and magnetic reconnection);
- magnetically confined fusion (including turbulent transport and plasma-wall interactions);
- plasmas far from equilibrium.

4.2.3 The Present Status of Capability Computing

Traditionally, research in ***Astrophysics*** is strongly linked internationally. The European Space Agency (ESA) and the European Southern Observatory (ESO) are examples of very successful institutions. 14 Research Training Networks in astrophysics have been (and are being) supported by the EU under FP5, and 5 more under FP6. 13 out of a total of 51 projects for the DEISA Extreme Computing Initiative were submitted by astrophysicists, asking for a total of more than 10^7 CPU hours in the coming two years. European groups and groups in the USA and Japan collaborate not only on coordinated observational efforts; they work together also on theory and modelling, including the development of computer software, sometimes in friendly competition. European Supercomputer Centers would foster the already existing close collaborations and would allow European astrophysicists to compete successfully with the strong and well-supported groups in the US and in Japan.

For many applications in astrophysics (as well as in other fields) the availability of codes solving the hydrodynamic equations by means of adaptively refined large eddy simulations is highly desirably. Further innovation potential lies in the fields of studies of compressible turbulence and turbulence in self-gravitating fluids. In particular, the development of vorticity in predominantly shock-driven flows is a poorly understood phenomenon with direct relevance for astrophysics and for many questions in mechanical engineering also. Until now, the development of such codes was prohibited by the huge memory requirements expected for production runs, but they will become feasible on systems with several thousand processors.

All of the computational grand challenges identified above are essential parts of large-scale efforts in ground-based and space astronomy. Again, most of these projects are international collaborations, often just within Europe but even more frequently with participating countries from all over the world. These include the ESO telescopes in Chile, the CFHT on Hawaii, recent satellites such as the HST, INTEGRAL and XMM-Newton, and future missions such as ALMA, OWL, SOFIA, COROT, Herschel, Darwin, WST, Planck-Surveyor, LISA, and the Solar Orbiter. European theory groups can only participate successfully in these projects and make use of the huge amount of new information that is expected if they have access to large HPC infrastructures.

In order to achieve the ambitious goals of the new particle-physics experiments in ***High-Energy Physics***, we must be able to interpret the experimental data without uncontrolled theoretical approximations. It is here where we have to rely on simulations of the strong interaction, QCD. The formulation of QCD on a four-dimensional space-time grid as well as its numerical-statistical simulation on supercomputers is called lattice quantum chromodynamics (lattice QCD). Today, this type of computer simulation is the only known method to evaluate QCD *ab initio* in just the energy range where the transition of quarks and gluons of the quark-gluon-plasma phase to protons and neutrons takes place. Due to quantum fluctuations, the strong charge acquires a high value at low energies, and consequently, perturbative calculations do not work. Moreover, in combination with future experiments these computations will show what kind of physics is needed beyond the standard model.

Nuclear and elementary particle physics meet precisely in this energy range where they have the common interest in understanding the nuclear forces from the first principles of the strong interaction. On the one hand, a fundamental theory such as QCD has to be evaluated without the detour via models, if one wants to confirm it or to use its full predictive power. On the other hand, chiral perturbation theory provides a model for the nuclear forces that is consistent and, via the chiral symmetry of QCD, related to other strongly interacting systems, e.g. the pion-nucleon system. These theoretical achievements have to be experimentally validated and should be used to predict the properties of nuclei that are not accessible experimentally, as it is the case for reactions at very small energies, which play a paramount role in stars, as well as the properties of nuclei close to the stability threshold. The computations feasible today for light systems already show that the nucleon forces from chiral perturbation theory lead to realistic values for the ground state binding energies, thus bridging the traditional gap between nuclear and hadron physics. For heavy nuclei the shell model Monte Carlo procedure allows for *ab initio* calculations on the nucleon level, but its application is severely limited by the lack of sufficient computer resources.

Finally lattice QCD can significantly enhance the impact of experimental programs. The extent to which lattice results reproduce experimental observations gives confidence in its predictions for similar quantities which cannot be measured directly by experiment. Also, in a growing number of experimental measurements, the uncertainties in the theoretical predictions from QCD dominate. Thus is makes good sense to supplement the enormous experimental efforts by a related investment in lattice QCD.

In ***Plasma Physics,*** ITER represents a major challenge for the modelling activity, which must now provide for an integration all the relevant models in a series of codes before 2015-2020. The present research activities show that the fundamental processes which, for instance, rule the core plasma behaviour imply fully time-dependent kinetic equations, generally in 2 velocity and 3 space coordinates, coupled to the electromagnetism equations. The relevant timescales cover 4 to 5 orders of magnitude (magneto-hydrodynamics, heat and particle transport, plasma current density profile

evolution). The need to integrate such a core plasma physics with the phenomena ruling the plasma periphery (the so-called plasma-wall interaction issues), being themselves influenced by the local physics and chemistry in a 3D space geometry, and following different timescales, motivates the requirements for centralised computer capabilities well beyond the ones presently accessible to the magnetic fusion community.

The apparently huge required effort in manpower and computer hardware has to be compared to the cost of experiments (4.7 billion € for ITER) which can be utilized much more effectively, and in the future partly perhaps even substituted by modelling if extensive computer simulations become available and if one can demonstrate a truly extrapolatable understanding of the results of fusion devices already now in planning or under construction.

4.2.4 The Quest for High-end Capability Computers

For applications to all of the grand challenges in ***Astrophysics*** listed above, codes are ready for and make use of high-end capability computers as will be demonstrated by means of a few examples.

In ***solar physics***, the simulations have to cover a sufficiently large volume, while resolving small-scale structures as well. A typical and presently feasible simulation needs several 10^5 CPU-hours on an IBM-Regatta system, for a 3D grid of 750 x 750 x 300 grid points, and 100 Gigabytes of memory. For archiving the data of a single production run 15 Terabytes are needed.

At present simulations of supernova explosions of massive stars can only be tackled with reduced dimensionality, but even then require ~10^5 CPU-hours per simulation and, because of the non-local nature of neutrino and radiation transport, only computers with shared-memory architecture can be used efficiently. In contrast, simulations of thermonuclear explosions perform well on massively parallel systems, and the largest computations performed until now used about 1000 hours on 512 processors of an IBM Power-4 system for a single simulation with 1024^3 grid points.

Similarly, state of the art simulations of star formation, of accretion disks around proto-stars and compact objects, supersonic outflows from stellar-mass or supermassive black holes and cosmological simulations, even with simplified physics and often reduced dimensionality, make use of a significant fraction of the time available at all supercomputer centers world wide, including Spain and UK in Europe, and the US and Japan. The VIRGO consortium, the leading effort in computational cosmology worldwide, is an international collaboration based in Germany and the UK. The largest cosmological simulations yet completed by the consortium used up 400,000 CPU-hours on the IBM Power-4 system and produced 23 Terabytes of data, and similar amounts of computer resources are presently used for simulations of two merging black holes in General Relativity at supercomputer centers in the US.

In ***High-Energy Physics*** lattice QCD has matured over the last ten years and has become an indispensable instrument for the investigation of theories of nuclear and elementary particle physics as well as for the planning, evaluation and interpretation of existing and future accelerator experiments. Fundamental quantities of elementary particle physics, such as the strength of the strong force or the masses of quarks, can be determined only with the help of lattice QCD. Apart from the prediction of exotic hadrons, such as the so-called glueballs, or the determination of distribution functions of quark or gluon densities, lattice QCD is regarded as the only method which can lead to an understanding of the properties of the quark-gluon plasma. Today, the solution of lattice QCD is considered as one of the grand challenge problems of computational science. Lattice QCD is a striking example of the use of capability computing because the four-dimensional structure of the problem, on the one hand, requires an extraordinarily amount of computing power but, on the other hand, fortunately guarantees good scalability.

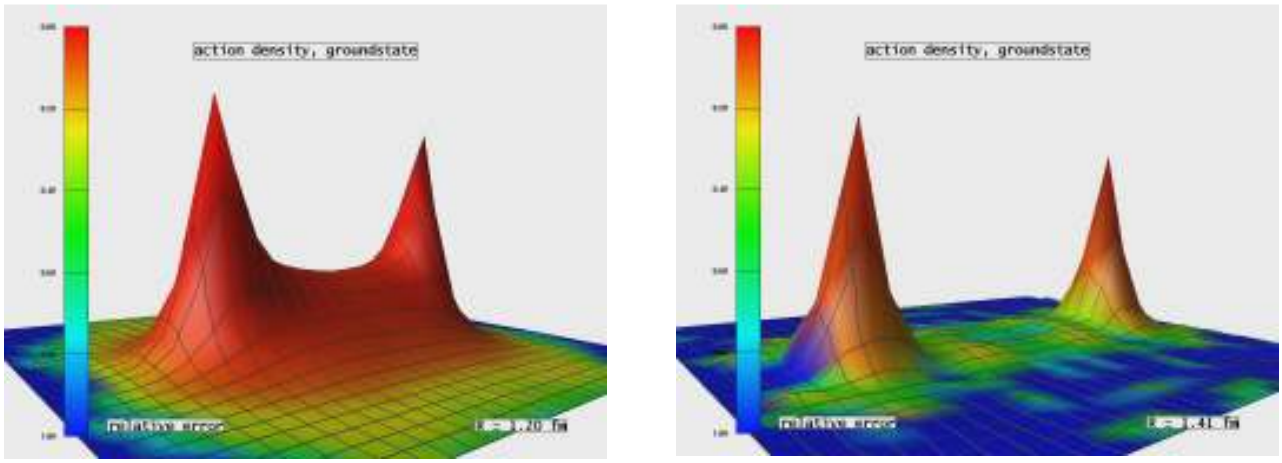


Figure: String-breaking of the QCD Flux Tube

In lattice QCD computations, the lattice should be large enough to accommodate the surrounding pion cloud, while the lattice spacing must be small enough to allow for the resolution of the inner structure of the nucleon. The quark mass range accessible with today's computers poses a major problem: the masses used for the u and d quarks should be small enough to carry out a safe extrapolation to the chiral limit. However, the computational effort grows with a high power of the inverse quark mass; therefore, the most advanced computations today are restricted to unphysical quark masses or approximations which introduce more or less unknown errors, which clearly contradicts the basic idea of an *ab initio* solution. This problem is alleviated by effective low-energy descriptions of QCD like chiral perturbation theory. However, the quark mass ranges in which chiral perturbation theory is well convergent is in most cases not yet reached by lattice QCD. As the computer resources are not yet sufficient for ideal calculations, several different approaches are used which have different specific advantages and disadvantages. Combining their results allows, however, a rather reliable extrapolation to the physical reality. Many groups favour computations using a fermionic action with good chiral properties. Other groups prefer Wilson fermions which break the chiral symmetry and thus lead to large fluctuations, but are computationally often less expensive. Still other groups use staggered fermions which are computationally extremely cheap, but could possibly have a problem of concept. Ideal fermions like overlap fermions or domain-wall fermions require computational resources probably up to two orders of magnitude larger than what is currently used.

The applications of lattice QCD are still strongly limited by the fact that typically one can generate only a few thousand configurations for a given set of lattice parameters. Thus, for some observables of interest the statistical fluctuations are too large to extract a signal. Ambitious future projects involve e.g. the study of hadron resonances or the study of CP violation in K decays, which requires multi-particle final states.

Today, an estimate of the total sustained CPU time dedicated continuously to lattice QCD in Europe amounts to several Teraflop/s. This combined computing effort is generated by many different computing centers and a strongly interwoven system of collaborations.

In the future, even stronger collaboration between European lattice groups will be critical in carrying out the ambitious next generation lattice simulations. European lattice QCD is already preparing for such pan-European projects, setting up a Europe-wide lattice data grid as part of the global international lattice data grid (ILDG), where sites like Edinburgh, Jülich, DESY and IDRIS are involved. At the same time, Europe-wide lattice projects participate successfully in the DEISA DECI-initiative with two highly rated large-scale simulation projects (one on 'Hot and Dense Matter' and one on 'Twisted Mass QCD'), taking advantage from grid-enabled workflows on the DEISA supercluster. As has happened in the past (APEnext, QCDOC, Blue Gene), lattice QCD will drive future developments of highly efficient and highly scalable supercomputer hardware and can help to trigger Europe's efforts in this direction.

High-energy physics has also major computational needs which are not related to lattice QCD. The enormous amounts of experimental data which will e.g. be generated by the LHC experiments has to be analysed and the highly complex detectors have to be simulated in great detail to understand all efficiencies and possible event structures with the required accuracy. This computational work currently involves either event generation or the evaluation of scattering cross sections which can be handled by independent processors. Facilities used to date for this work have either been processor farms or clusters rather than tightly coupled multiprocessors. Future computing resources for this effort will be centred on those provided via the Grid, an approach consistent with the fact that the very large experimental groups work at many different locations in Europe and abroad.

In ***Plasma Physics*** two areas drive the computer requirements of the field. The minimum size of a self-sustained, burning plasma is determined by turbulent transport, driven by the spatial gradients in plasma temperature and density. This transport has so far been estimated and extrapolated essentially using semi-empirical methods. In the near future, however, modelling of gradient-driven turbulence in toroidally confined plasmas has a realistic perspective of covering simultaneously space scales from the ion gyro-radius to the global plasma dimensions and from the eddy turnover to the energy confinement time in single, comprehensive simulation runs.

The physics model to be used (gyrokinetics) is truly first-principle based, with all approximations formulated in mathematically rigorous form, and based on well-satisfied orderings. At least if the present contention – that electron gyroradius-scale modes can be ignored in many relevant situations – holds, such calculations would constitute a truly one-to-one mapping of this aspect of experimental reality onto the computer. In spite of the apparently much larger complexity of this problem than that of ordinary fluid dynamic turbulence, we are in fact here in a more favourable situation, as the ion gyroradius introduces a lower limit to the space scales of interest which is "only" a factor of the order 1000 (in a reactor) smaller than the total plasma dimension, and thereby limits the resolution requirements.

A second area spearheading fusion's need for more computer power, is that of the non-linear development of large scale (magneto-hydrodynamic) instabilities, and in particular their interaction with fast non-thermal particles, such as those produced by the fusion reactions. Experimental investigations in this area will be one of the main aims of the upcoming ITER experiment, and verified *ab initio* models of this process will be needed, to apply, for example, this experience to stellarators.

4.2.5 A Roadmap for Future Capability Computing

In the years ahead, the needs in most fields of ***stellar astrophysics*** and ***cosmology*** will increase by at least a factor of 10 because the spatial resolution of the present simulations is still far from being sufficient to model the interior of stars and their atmospheres, and galaxies and cluster of galaxies in a realistic way. Realistic simulations can only be performed on supercomputers with sustained performance ~100 Teraflop/s, but significant progress can be expected with slower machines in the interim.

For simulations of ***planet and star formation, stellar explosions, astrophysical jets and accretion disks***, solving the radiation-(magneto)-hydrodynamic problems ultimately will need supercomputers with several hundreds of Teraflop/s sustained performance which might become available past 2010. However, significant progress towards the understanding of all these exciting questions can be obtained by reducing the dimensionality of the simulations assuming approximate symmetries or by simplifying the physics of the models. For instance, the simulation of a relativistic jet from an active galactic nucleus in three spatial dimensions, ignoring magnetic fields and transport of radiation, but with high numerical resolution, will require around 10^{20} floating-point operations, but would allow for a first comparison with the excellent observational data.

Lattice QCD requires on the one hand the stochastic simulation of the theory, where ensembles of gluon configurations are generated, and on the other hand the evaluation of many different correlations between quark, gluon, or quark and gluon fields on these configurations to obtain quantities of physical interest. While the second step, the evaluation of correlators, can lead to complex program structures with often very large input-output rates, as well as requirements for data management and

computing power, the stochastic simulation to generate the quantum fields is most costly, but, fortunately, can be carried out on highly scalable systems. Therefore, the successful simulation of lattice QCD requires both general-purpose capability systems (like IBM SMP-clusters or the Hitachi or SGI systems) and highly scalable machines. Prominent examples of the latter are the QCDOC and APEnext machines. QCDOC is a machine designed for QCD, which was developed by an international collaboration involving British and German members, and was the predecessor of the successful IBM Blue Gene leadership-class systems. In Europe, INFN and DESY have developed the APE line of supercomputers with a comparable price-performance ratio. These projects demonstrate that lattice QCD has a broad outreach and will – as one of the main driving forces of supercomputing – continue to be an important forerunner for future technologies in the planned European supercomputer infrastructure.

Lattice QCD is popular in the whole European research area with groups in Austria, Cyprus, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Poland, Spain, Sweden, Switzerland, and UK. The computational requirements voiced by these European groups sum up to more than 1 sustained Petaflop/s by 2009.

The aim of future ***MHD-turbulence*** simulations is to go to 4096^3 grid points and beyond. One envisions results on nonlinear turbulent cascades and their spatial anisotropy due to turbulent or inhomogeneous magnetic fields. The Reynolds numbers one aims at are of the order of 10^4, allowing also the simulation of conditions of crucial importance in engineering, geophysics, space physics and astrophysics. The pseudo-spectral magneto-fluid codes used for this purpose have been shown to scale nicely up to 512 processors (and most likely will scale for many more processors). The total memory needed for such simulations is of the order of tens of Terabytes and the CPU requirements are several 10^6 hours and a sustained Teraflop/s rate of 30 to 50.

In the framework of ***kinetic simulations*** a resolution of 256 x 1282 x 303 grid points (in space and velocity) can be performed on computers with about five Teraflop/s (sustained), whereas just doubling the spatial resolution increases the computer power needed by roughly a factor of 8. The ultimate goal of such simulations is, however, to increase the resolution by even larger factors. Similarly, coupled Vlasov+MHD simulations need computers with about 100 Teraflop/s sustained CPU-power.

4.2.6 Requirements for Astrophysics, High-energy Physics and Plasma Physics

All HPC applications listed in the previous sections are on-going into 2009 and beyond. The new observatories, satellites, accelerators, and fusion experiments which will be operational or will become available by that time will raise the need for largely improved theoretical models and numerical simulations.

- Typical applications are expected to require 70 – 100 Teraflop/s (sustained) over a few months to one year per application in 2007 – 2009, and > 500 – 1000 Teraflop/s (sustained) in the years past 2009. The model parameters which control the computer needs are the 'numerical resolution' (because of the multi-dimensionality and the multi-scale nature of almost all applications), the simulation times (which have to be increased to match the simulated physical systems), and the 'input physics' (taking care of the complexity of the systems and making simulations as realistic as possible).
- Different computer architectures are most cost-efficient for the three communities of this panel. High-energy physics theory applications are highly scalable and need high I/O band-width, but low memory. In contrast, most astrophysics and plasma physics applications need large memory and high single-processor performance. In addition, some applications for which non-local physics (such as transport of photons or neutrinos) is essential, even shared-memory architectures might be needed.
- Grid access will be needed as an infrastructure.
- The panel expressed the strong opinion that a European HPC Centre should offer computers of different architectures to the community. In addition, it should also be a competence and support centre.

4.3 MATERIALS SCIENCE, CHEMISTRY AND NANOSCIENCE

4.3.1 Summary

Materials science simulations are playing an increasing role in almost all aspects of the study of materials: not only traditional materials science, physics and chemistry, but also nanoscience, surface science, electrical engineering, earth sciences, biology and drug design.

By throwing in the atoms and the laws of physics and chemistry into the computer, one effectively performs 'computer experiments' to observe structures, properties and processes and to conduct measurements that cannot be done readily in the laboratory.

Electronic structure theory and statistical physics are the main tools on which a very large part of our understanding of the physical and chemical properties and processes of materials are based. The progress made in the last 20 years, during which those two pillars have been blended into a comprehensive methodology, has enabled the description and physical understanding of "simple" materials. These materials are simple in their *atomic* structure, with relatively few components and of moderate size, simple in their *electronic* structure, and the processes are studied on short time scales. These limitations have hindered our understanding of many technologically important materials, which exhibit a higher degree of complexity.

Overcoming these limitations would impact beyond traditional materials science physics and chemistry in fields such as nanoscience, surface science, electrical engineering, earth sciences, biology and drug design.

An enormous amount has already been achieved, and much more can be tackled with the present methodologies and computer codes given the computer power available in the near future. In addition several lines of research are focusing on better algorithms to deal with more complex situations and yield greater accuracy. The major advances have come, and will continue to come, from improved algorithms and approximations as well as from an increase in computing power

Scientific, technological and societal challenges require the acquisition at the European level of capability computers in the 1-10 Petaflop range. This would allow us to simulate complex systems over the long space and time scales required for quantitatively reliable results and for improved connectivity between experiment and theory. A clear benefit of these developments is the enhanced industrial exploitation potential in a variety of fields relevant to energy production, to electronic and optical communication, and for ceramics for sensors, plastics and complex fluids in the food industry, or hydrogen storage materials. In addition, there are important societal applications that will lead to improvements to cleaner air, sustainable technology, decontamination, nanotechnology and medicine.

To ensure the well-being of European high-end computing in the future, it is crucial to support the research community from which new methodologies arise, including software development. Europe is a world leader in the development and use of atomic-scale materials modelling, and a European top-of-the-range computer service will help to sustain and enhance this leadership position. The importance of this area of research is exemplified by the increased effort in USA, illustrated by the recent National Science Foundation reports on cyberinfrastructure in general, and materials simulation in particular.

4.3.2 Identification of Computational Challenges

Evolution of simulation methodology

In order to sketch the long term perspective and specify Grand Challenges, it is necessary to explicitly identify the simulation methodology, the current hurdles, and detail where a stepwise increase in computer power is both necessary and opens new doors.

Experience has shown over many years that quantum mechanical calculations of the electronic structure are essential when there is a making or breaking of bonds between atoms or a significant

change in the bonding geometry. This is an awesome task for several hundred or even thousands of atoms and a whole raft of special approximations and short cuts have been developed and tested to make it feasible. Many researchers are engaged on new algorithms to extend the methodology in accuracy, system size and complexity of atoms involved, particularly with transition metal compounds or actinide elements. These will play an increasingly important role over the coming decade and beyond.

Large interacting systems of atoms which cannot be broken up into separate pieces for study, as in the usual 'scientific method', are the norm in a very large array of technologically relevant applications. These systems appear in such diverse fields as layered magnetic devices, the transport of biochemicals through pores in cell walls, solutions and biological systems where the surrounding water plays a crucial role or the study of the effect of radiation damage on the durability of materials. This, together with the earlier remark, provide very strong arguments for the provision of Leadership-class Supercomputing.

Beyond the need to simulate larger and more complex systems which inherently involves many atoms or molecules there are more fundamental reasons to use top-of-the-range computing power. One crucial problem is the correlation between the electron motions when these depend strongly on the particular atom or environment, which affect strongly the optical and magnetic properties. Furthermore, the direct simulation of the electrical current including the slow knock-on damage in semiconductor devices will also require very demanding resources.

Chemical processes such as reversible hydrogen storage are controlled not by static energies but by the thermodynamic free energy and long simulation runs are required to sample adequately the thermal equilibrium.

Societal and economic benefits

These simulations have important societal applications in such diverse fields as air cleaning, sustainable technology, decontamination, nanotechnology, medicine, biology, biotechnology, earth sciences and energy production. This is outlined in the challenges detailed below, which have been grouped in three categories:

1. Understanding complex materials
2. Understanding complex chemistry
3. Making progress in Nanotechnology

1. Understanding complex materials

The progress expected in the understanding of technologically relevant materials comprises the determination of purely electronic properties, such as optical properties (optomechanical and optoelectronic properties and optomagnetic processes) or transport properties (giant and colossal magnetoresistance). These properties are central to many devices in the electronic industry.

The simulation of nucleation, growth, self assembly and polymerization is central to the design and performance of many diverse materials such as rubbers, paints, fuels, detergents, functional organic materials, cosmetics and food.

Finally, the relation between process, conditions of use and composition of a material requires a multiscale description of e.g. mechanical properties of materials (nuclear energy production).

As an example, the multiscale simulation of the mechanical properties of irradiated materials would require access to facilities 1000 times larger than those currently available.

These simulations would decisively contribute to the prediction of the lifetime of high performance materials in energy technology, such as high-efficiency gas-turbines ...

2. Understanding complex chemistry

The simulation of technologically relevant chemical reactions and processes, has a huge potential in a variety of fields.

Catalysis is a major challenge in the chemistry of complex materials, with many applications in industrial chemistry. The knowledge of atmospheric chemistry is crucial for environmental prediction and protection (clean air). Improving the knowledge of chemical processing (from soft chemistry including polymers to the atomistic description of combustion) would improve the durability of chemicals. Supra molecular assemblies open new possibilities for the extraction of heavy elements from spent nuclear fuels. In biochemistry, a vast number of reactions taking place in the human body (for example) are not understood in any detail.

Addressing effectively the modelling of these complex chemical systems can only be achieved with a massive upgrade of computing resources.

As an example, the realistic treatment of supported catalytic nanoparticles involving several hundred transition metal atoms would require resources of at least a Petaflop/s. This is a key step in the development of the clean fuels of the future.

3. Making progress in Nanotechnology

The advance of faster information processing or the development of new generations of processors requires the shrinking of devices, which leads unavoidably towards nanoelectronics. Moreover, many new devices, such as nanomotors can be envisioned, which will require simulation of mechanical properties at the nanolevel. Finally, composite high performance materials in the fields e.g. adhesion and coatings will require an atomistic based description of nanorheology, nanofluidics and nanotribology. As an example the description of the complex magnetic and mechano-optical properties of nanodevices components, requires sophisticated computational methods, feasible only on nanosize systems with massively parallel computers in the Petaflop/s range.Scientific Challenges

4.3.2.1 Challenge A1: Understanding Complex Interfaces and Mesoscopic Systems.

Description:

Soft Condensed Matter encompasses polymers, colloids, membranes, amphiphiles and surfactants, synthetic and biological polymers, lipids and proteins as well as organic-inorganic hybrid systems. Classical chain molecules, i.e., polymers in the narrower sense, form only a subgroup of soft matter and serve as a reference for model building. The unifying principles of soft matter systems are structures on mesoscopic length scales from nanometers to micrometers, and typical energy scales on the order of the thermal energy. This huge diversity of scales calls for a multiscale treatment of these systems, although in some cases direct simulation at the atomic or “object” scale (e.g. of hydrodynamics process) yields data which cannot be obtained by standard hydrodynamics methods.

Motivation (Scientific benefits, Social benefits)

Soft matter encompasses a huge variety of technological applications in many different kinds of materials such as rubbers, paints, fuels, detergents, functional organic materials, cosmetics, food, biomembranes, the cytoskeleton and the cytoplasm of living cells, etc. The sufficiently detailed theoretical modelling of highly complex systems such as polymers comprises many orders of magnitude in time and length scales.

Relation to state of the art

The goal of simulation to be permanently monitoring experimental studies requires a dramatic increase in available computer performance. In many cases, a brute force direct approach cannot describe the full set of time scales involved, and, apart from the need for considerably more computational power, it is essential to develop simulation methods that enable the systematic coupling of several length and time scales (multi-scale simulations). With these new techniques, one can expect to go beyond the current limitations – 3,000 CPU years on 1Pflop/s computer would be needed for the molecular

dynamical simulation of an atomically resolved polymer melt.) With these tools, simulations of the dynamics of polymer melt can be conducted at a synthetic level, but will still require a huge amount of computer time. Similar considerations apply for the simulation of simple polyelectrolytes in implicit solvents, multi-scale simulations of systems with relatively short chains or for studies on multi-component systems.

Roadmap

Examples include *the consideration of local ion interactions and explicit solvents (e.g., molecular structure of water) for polyelectrolytes, to which almost all biopolymers belong; the dynamics of realistic polymer melts with branched polymers, the phase behavior of multi-component systems or scale-spanning calculations with realistic dynamics and with conformation changes of smaller and ultimately larger biopolymers.*

Biological systems can control the growth mechanisms of crystals, producing unusual phases, shapes and structures and the organic component often acts as a template for mineral growths. Work on the interaction of biological molecules with materials is urgently required. The development of such models will enable us to screen possible candidates for the manufacture of artificial materials for surgical implants. The complex nature of this interface and the computational expense of quantum mechanical surface/adsorbate calculations are only feasible using Leadersip-class computing facilities.

4.3.2.2 Challenge A2: Multiscale Description of Mechanical Characteristics of Materials

Description:

A major computational challenge is the development of simulation tools that enable the prediction of microstructural and mechanical characteristics of materials resulting from the conditions of their use and material composition. It implies linking mechanisms at the atomic space-time scale (10^{-9}m, 10^{-12} sec) with the component scales (meter, years).

Motivation (scientific benefits, social benefits)

Mechanical properties of materials result from a combination of their elaboration process, chemical composition and conditions of use. Well known examples are the description of welding or understanding the durability of materials when sustaining intense irradiation in nuclear reactors (fission or fusion) or when used for waste storage.

Relation to state of the art

Large databases of properties of microstructural elements (such as point defects, dislocations or interfaces) have been obtained by first principles (FP) or classical molecular dynamics (MD) simulations on systems ranging from a hundred (FP) to millions (MD) of atoms.

The long-term variation of defect structure is obtained by Kinetic Monte Carlo techniques or by rate theory models, using information extracted from these databases. These simulations at the mesoscopic scale yield the defect evolution in shape and size. Microstructural evolutions are taken into account in Discrete Dislocation Dynamics, in order to provide a physically based prediction of mesoscopic mechanical properties of metallic materials. Efficient homogenisation methods are also available to derive macroscopic properties.

Roadmap

With TFlop machines first principles studies of defects are restricted to small defect complexes with selected configurations and zero temperature. The advent of PFlop/s machines will enable:

1. The study from first principles of the extended defects (dislocations) that govern the mechanical properties of materials (an increase by a factor of 5 to 10 of system size, i.e. of a factor of 100 to 1000 of computer time).
2. Finite temperature simulations of defects (involving an increase by a factor of 100 to 1000 in the number of atomic configurations).

To characterize basic mechanisms involved in changes in the microstructure will require thousands of first principles computations on systems of about one hundred atoms (10 to 100 times more computations than at present). The investigation of complex defects such as the structure of dislocations, the evolution of their core under deformation and their interaction with impurities will require larger systems (about 500 atoms) and would be 100 to 1000 times longer. Even larger systems are required to study dislocation kinks or interacting dislocations. Finally the investigation of the diffusion of complex defects at finite temperature will require very long first principles molecular dynamics simulations (100 to 1000 times longer than is currently possible).

New empirical potentials – fitted to FP results – will be needed to better describe these materials, especially to obtain an accurate description of defect energetics and activation barriers of relevant processes. Their use will induce additional costs as compared to present empirical potentials of one to three orders of magnitude.

Dislocation dynamics, which describes the microstructure, will require 3 to 4 orders of magnitude in simulation and a high performance parallelisation of the software to examine realistic samples and deformations. Interface and grain boundary should also be added in simulations. Furthermore, the use of advanced constitutive relations derived from these models will demand extensive computer resources when coupled to mechanical models. The derivation of effective macroscopic properties requires the computation of realistic 3D microstructures by finite element (FE) codes. Results are accurate if a large volume is computed with fine microstructural details. Massive parallel computations and FE simulations provide an efficient but extremely CPU intensive solution.

These materials simulations techniques can be used for studying many other modes of damage to materials, such as irradiation by neutrons, corrosion or fatigue, and the combination of several different stresses. These tools can be used to study structural materials or functional materials. In these problems, the interfaces (e.g. the metal-oxide interface) can play an important part, and this would also require major computer resources.

Finally, the design of hydrogen storage materials (HSM) undergoing reversible hydridation under moderate hydrogen gas pressures, and temperatures not far from ambient requires the capability to compute their free energy of hydridation which sets the hydridation isotherm. These free energy calculations require between 1 and 10 Pflop/s computing power to simulate technologically relevant materials.

4.3.2.3 Challenge B: Realistic Description of Catalysts in Function.

Description:

A major theme of the chemistry component of this area is the study of chemical reactivity and catalysis. The computational power of a Petaflop facility would allow us to address this area with a new degree of realism. First principles simulations of reactions and catalytic processes within the condensed phases and at interfaces are limited by several factors, including the numbers of atoms that can be considered and the timescales of the processes. The present facilities allow a single reaction pathway to be studied for a limited number of atoms and a short timescale.

To develop new catalysts, as well as being able to increase the numbers of atoms and timescale, we need to study many pathways and predict the rates of reactions.

Motivation (Scientific benefits, Social benefits)

The computational design of new heterogeneous and homogeneous catalysts represents a major challenge. Catalysis is a key issue for a cleaner production of chemicals and fuels involving more selective reactions and more effective processes, with reduced waste and energy consumption. Catalysis is critical in the mass development of efficient fuel cells. Key industrial processes, such as selective oxidation propene, and sulfur removal from crude oil involve supported catalytic nanoparticles. . Other processes rely on microporous materials such as zeolites through their combination of shape selective properties and controllable acidity.

Relation to state of the art

Large databases of properties are obtained from FP simulations for molecular reactants, intermediate and products, as well as appropriate atomistic models for reactive solid surfaces (unit-cells containing alternating "slabs" of solids and free space for incoming adsorbates, treated with periodic boundary conditions) .

The determination of reaction rates and pathways requires simulations of many replicas of the systems along the reaction path. In this way, it is possible to obtain the appropriate energetic landscapes, or potential energy surfaces, and localize precisely the key transition states as saddle points on these surfaces.

Roadmap

The realistic description of catalysts of intense technological interest will require the description of reactions of a new degree of complexity, which must take into account the size and structure of the catalysts.

To understand and design optimised catalysts in porous materials for applications such as cracking, one must understand the reactions that occur during catalysis. For example, methanol can be converted by the zeolite ZSM-5 to gasoline with a single pass through the zeolite. Also, the description of polymerisation reactions would help to develop new catalysts containing metals for the production of functional polymers such as polyurethanes, polyethers and polyesters that are not toxic.

To take into account the structural and electronic properties of active nanoparticles carried by spectator nanoparticles and the different processes involved would increase the size of the simulation (by a factor of 10 to 100) and the number of reaction paths (a factor of 10 to 1000). For instance, if we consider Co(Ni)MoS on alumina used in desulfuration of gasoils, in order to address more realistically the complexity of an industrial system, one must evolve from models involving ~200 non equivalent atoms to ~2000. This would imply a 1000 fold increase in computer power for first principles simulations which scale as $\sim N_{atoms}^{3}$, a realistic objective for 2009.

4.3.2.4 Challenge C1: Nanodevices and Nanotechnology

Description:

The fundamental understanding and applications of nanosize metallic or semiconductor particles and well-designed molecules and their ensembles constitute a new and rapidly developing area in materials modelling. Computer simulations can make two substantial contributions to the field. The first concerns the development of a fundamental understanding of the relationship between the structure of the nanoparticles and their electronic, optical, magnetic and transport properties and the second relates to the ability to describe or induce assembly of these particles in a configuration dictated by the fabrication architecture of the nanodevice.

Motivation (Scientific benefits, Social benefits)

Nanosystems are considered key to future developments in new materials and technology. In fields such as information processing (smaller and faster), nanoelectronics (processors beyond current size limitations), nanosystems represent a new frontier. Furthermore, precision engineering, nanoindentation, nanoengineering and nanomotors open doors to challenging applications e.g. non invasive surgery. New composite materials, engineered at the nanolevel, have the potential to perform better in the fields of adhesion and coatings.

Relation to state of the art

The structure, electronic, optical and transport properties of the simplest among individual nanoobjects can be computed with the current capability computers, the most vivid example being the ubiquitous carbon nanotube. The description and design of future nanodevices, will require the description of more complex systems in shape and chemical structure, as well as their assembly. This will require a 100 to 1000 fold increase in computing power, depending upon the property being investigated.

Roadmap

The design of new electronic devices, for example. nanotubes/nanowires-based field effect transistors requires the coupling of atomistic methods, such as out-of-equilibrium transport processes, including electron-phonon coupling with more traditional methods of macroscopic electronic circuit simulation.

The simulation of other devices, such as photosensitive molecular machines, will require the simulation of excited states (using many-body techniques) and mechanical properties of systems made of 20-50 atoms, which represents a 100 to 10000 increase in current computing power.

The design of nanodevices by self-assembly of nanoparticles requires a detailed understanding of the roles of solvent, chemical functionality and physicochemical nature of the substrate and the shape and size of particles on the self-assembly. Ultra-thin films can display structures and properties very different from bulk materials of the same composition. Perovskites such as $(Ba,Sr)TiO_3$ are current candidates for high dielectric constant films in semiconductor devices, zinc oxide, where c-axis-oriented films are of considerable interest for the production of surface and bulk acoustic-wave devices, and, with appropriate doping for electrodes in flat panel displays.

4.3.2.5 Challenge C2: Superconducting and Magnetic Materials

Description:

Materials with complex electronic structure play an important role in many devices including magnetic storage media, sensors or superconductor based devices but often require the use of approaches beyond density functional theory. The development and use of new types of approaches is a major challenge for the study of these important materials.

Motivation (scientific benefits, social benefits)

Simulation of non-collinear magnetization dynamics is essential for the development of future magnetic storage media. For instance, an important role is expected in future quantum electronics and storage technology for carbon nanotubes filled with magnetic materials or molecular magnets.

Similarly, high temperature superconductors might find applications in energy storage and in ultra-fast electronic chips with a wide range of applications.

Relation to state of the art

New approaches which aim to go beyond mean field theories are currently being developed, and rely on very expensive Monte-Carlo approaches, whether in the framework of the so-called Dynamical Mean Field Theory (DMFT) or the full Quantum Monte Carlo Method.

Roadmap

Even after two decades of intensive research it has not been possible to uncover the mechanism underlying high-temperature superconductivity. One of the reasons is that many-body calculations for materials with strong correlations are made exceptionally difficult by the fact that a crystal cannot be described by simply considering a single unit cell with appropriate boundary conditions. Instead, for reaching the bulk limit, one has to simulate supercells of increasing size and extrapolate to infinite size (finite-size scaling). In general, the effort required to solve such a finite-size many-body system increases exponentially with system size. This is reflected in a corresponding increase of computing resources: the storage required for exact diagonalization calculations while in quantum Monte Carlo approaches instead the computing time to acquire sufficient statistics to overcome the fermion sign-problem, which is particularly bad at low temperatures. New approaches like Dynamical Mean-Field Theory allow the bulk limit to be approached more quickly. In its simplest form it works by embedding one unit cell in a self-consistent, dynamical medium. While this approach works fairly well, it makes a too severe approximation to the momentum-dependence of the self-energy. To improve on this approximation it is therefore still necessary to consider clusters.

The simulation of the macroscopic superconducting current of e.g. 10^{23} Cooper pairs in a high-temperature superconducting wire remains a major challenge. The solution of the problem requires simulations for very large model systems. Finite-size scaling procedures, which are based on cluster simulations – especially on quantum Monte Carlo simulations –, have been developed for this purpose. The limits of the currently installed supercomputers have already been reached since the so-called

fermion sign problem of quantum Monte Carlo simulations grows exponentially with the size of the simulation. This problem generally arises during simulation of electrons (fermions) because the fermion's wave function has to be anti-symmetrical upon the exchange of two particles. However, several relatively new methods (quantum Monte Carlo simulations in connection with renormalization-group techniques and with dynamical mean-field theory procedures) are being tested in order to come to a microscopic understanding of pairing mechanisms and superconductivity. This is essential for replacing the empirical search for improved material properties. These calculations cannot be carried out with the currently available supercomputer generation.

Compounds based on rare earth or lanthanide elements display exciting properties resulting from the interplay of charge, spin, orbital, and lattice degrees of freedom. Currently, realistic simulations for strongly correlated systems such as the Mott insulator $LaTiO_3$ are limited to high temperatures. Most of the interesting effects occur, however, below the spin and orbital ordering temperatures. Due to the sign-problem, simulations in this interesting temperature range will require an increase in computer resources by about two to three orders of magnitude. A similar investment in computing power is required to be able to treat more realistic models, i.e. models which contain more of the correlated orbitals explicitly. Petaflop computing will allow solving a whole class of new problems.

4.3.3 Requirements in Architecture, Software, Storage

4.3.3.1 Computational Dimensioning

The relevant parameters helping to quantify a simulation in the field can be roughly defined as follows:

- Size of the system (number of atoms)
- Duration of the calculation for dynamical properties
- Sophistication of the theory due to:
    - The complexity of the system (strongly correlated electrons, importance of van der Waals forces, relativistic effects, magnetism..)
    - The property under investigation (e.g. response function vs static property)

To assess our requirements, we will consider first as a reference a static quantum calculation based upon the standard local density approximation of the density functional theory, of say 1000 atoms on a 10 Tflop/s machine. Including various degrees of sophistication lead to the following estimates:

Method/Problem	*Required increase in computing power*	*Remarks*	*Field of main application*
Large systems (large molecules, dislocations with kinks, etc.)	*Need larger systems 50-100*	*Moderate bandwidth High speed processors*	*Chemistry, Nanotechnology, Materials*
Non-collinear magnetism	*200-500*	*High bandwidth required*	*Electronic industry*
Sophisticated functionals (exact exchange, van der Waals)	*50-100*	*High bandwidth*	*Chemistry, biochemistry*
Strongly correlated systems	*500*	*High bandwidth/ Moderate bandwidth*	*New Materials*
Excited electronic states, transport	*Need smaller systems (factor 10-50). Then, 100-1000*	*Moderate bandwidth*	*Nanotechnology*
Quantum chemistry methods (MP2,CASPT2)	*Smaller systems, then factor 100*	*High bandwidth/high speed*	*Chemistry , new materials and catalysis*
Highly correlated quantum chemistry	*Much smaller systems (100).*		

Transition path sampling	*Factor 10-100*	*Moderate/high bandwidth*	*Catalysis*

These numbers must be (at least) multiplied by the increase in size of the system. Noting that the complexity of the algorithms which can be readily coupled with these methods is cubic, an increase by 10 of the system's size will induce a 1000 fold increase of computer time. Therefore, in parallel to this direct increase of complexity, research must be pursued to devise algorithms with better scaling and a wide spectrum of available properties.

In the field of dynamical calculations, a similar table can be set-up, using now as a reference a calculation of 100 atoms for 10 picoseconds.

Method/Problem	*Required increase in computing power*	*Remarks*	*Field of main application*
Path integral simulations	*100*	*Moderate bandwidth*	*Biology*
Non adiabatic simulations	*100-1000*	*Moderate bandwidth*	*Chemistry*
Full quantum dynamics	*Still very small systems*		

Beside quantum simulations, classical simulations using molecular dynamics or more specialized techniques (Monte Carlo, Kinetic Monte Carlo, Dislocations of other objects dynamics) will also benefit from high performance computing. If we consider a molecular dynamics simulation of 100 million atoms during 10 nanoseconds as a reference, we might infer the following:

Method/Problem	*Required increase in computing power*	*Remarks*	*Field of main application*
Electrostatic effects	*10-100*	*High bandwidth*	*Biology, nanotechnology*
Sophisticated potentials	*10-100*	*Moderate bandwidth High-speed CPUs*	*Chemistry Materials*
Monte Carlo methods	*1-10*	*Moderate bandwidth*	*Soft matter, interfaces*
Objects dynamics (e.g. dislocations)	*10-100 (depending upon object)*	*High bandwidth*	*Metals, soft matter*

One should note that the bandwidth is now somewhat less expensive than very high speed processors.

We have also detailed machine/system requirements according to the computing techniques needed to address different fields in materials sciences .We have in particular identified were a characteristic was critical for a successful adaptation of a given method

Method	Large number of processors	Low latency-fast bandwidth communication fabric	Large processor memory.	Increase of processor speed for long time scale problems (at least 10X):	Software development effort for parallelization and optimisation :
DFT methods , PW and localized orbitals		Critical			Needed
Post Hartree Fock methods		Critical	Critical		Critical
DMFT and related methods [density		Needed			Needed

matrix renormalization]					
Quantum Monte Carlo methods		Needed			Critical
Green's function & direct methods for quantum transport			Critical		Critical
Classical MD (short and long range forces)	Needed	Needed			Needed
QM/MM methods	Needed	Needed			Needed
Discrete Dislocation Dynamics		Needed			Needed
Kinetic Monte Carlo	Needed				
Methods for non linear PDE (multigrid methods)	Needed			Critical	Critical

We could summarize the requirements in the following table:

Estimated sustained performance requirements in TeraFlop/s

Sub Field	*2005-2007*	*2007-2008*	*2009*	*2015*
Nanoscience	*1*	*10-50*	*> 1000*	
Chemistry	*3*	*25-125*	*>600*	
Materials Science	*10*	*10-50*	*>400*	

A rough estimate for 2015 leads to 10pflop/s, but it is hard to predict the evolution for each domain.

4.3.4 Expected Status in 2009

We can expect to reach in 2009, the "virtual laboratory" paradigm, and witness the confluence between large experimental facilities in condensed matter (ILL, ESRF, LLB, DESY, BESSY, etc..) and computing facilities. This will allow studying, elaborating and using materials in many innovative ways and will dramatically increase our knowledge of the behaviour of complex materials, and their mechanical, chemical, optical and magnetic properties. This confluence will also allow a better understanding of many processes such as catalysis, crystal growth, corrosion, friction with wide ranging fundamental (fewer approximations) and technological (more complex systems) impact.

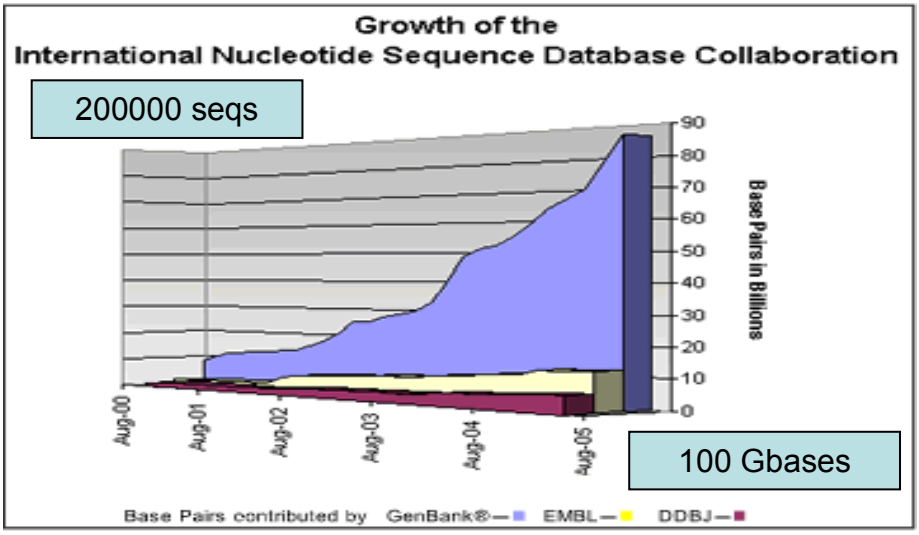
4.4 LIFE SCIENCES

4.4.1 Summary

The summary below is the output of the second meeting of the Life Science panel and several emails for a couple weeks. We should emphasize that there is a complete agreement between the representatives of the four countries in what is noted below.

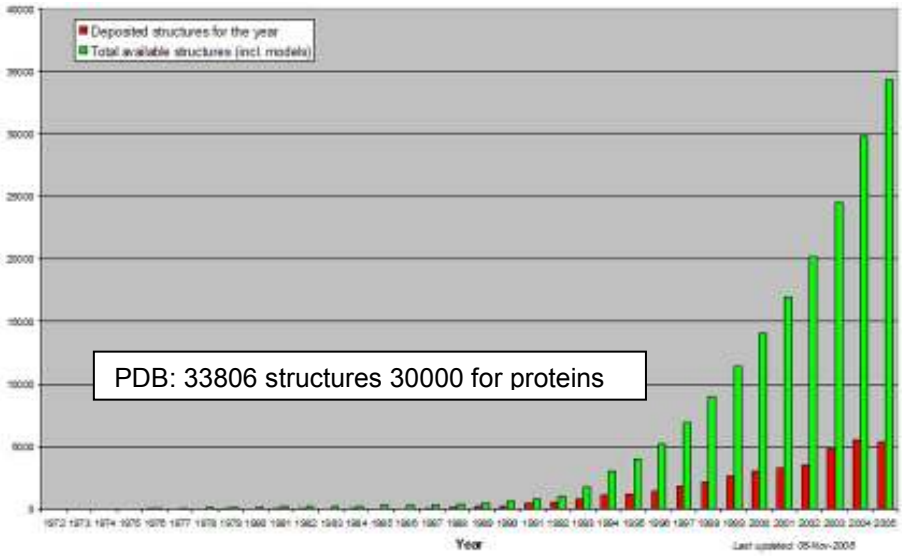
Genome projects have provided an overwhelming amount of data. On the sequence size, databases are increasing in size at an exponential rate and more than 100 Gigabases are now deposited (see Figure 1). Genomes of more than 1000 organisms are available, and mapping individual variability (especially in humans) is being carried out systematically by different international consortia. The annotation and characterization of genes progress less rapidly than sequencing, but databases like UniProt are very close to reach 200000 entries (199607 on Nov 25th). The increase in structure databases is also fast (more than 30000 entries; see Figure 2), but clearly these databases cannot compete with sequence-dabatabes, outlining the importance of theoretical algorithms for structure-prediction.

Fig 1



The initial enthusiasm in genomic projects has been modulated by the fact that translating the amount of data available into biologically-important information has been more difficult than expected. It is probably time to focus most of our effort in the interpretation of the available data. Combining, purely statistical approaches, which have guided research in bioinformatics in the past decade with physical approaches, whose origins can be traced to first-principles of physics and chemistry is probably the challenge for the next years. In this field, Europe is extremely disdvantaged due to the lack of suitable computational resources. For example, considering the growth of sequence and structure databases, we can guess that we should be ready to simulate the structure of around $3x10^5$ proteins by 2008, which even with improvement in current simulation algorithms will require 100 Teraflops to 1 Petaflop machines. Even if structure is available, simulations performed to learn how proteins work are extremely expensive and the cost will increase considering that larger structures are deposited every year in PDB (see Figure 3). In order to simulate most proteins at the atomistic level for a simple millisecond of a single protein we must perform around 10^{23} floating point operations, something just impossible with current computer resources. Processes occurring in the micro-to millisecond time scale, like protein folding, allosteric movements, chromatine reforming,… are just out of the possibilities of computational biologists in Europe.

Fig. 2 **GROWTH OF THE PROTEIN DATA BANK**



US has a strong lead in focusing computer resources in life science, with the centers in San Diego (UCSD, Scripps Institute), Harvard University, University of Illinois at Urbana Champaign, Cornell University and the Thomas J. Watson Research Center (IBM). A total of about 40 Teraflop/s peak CPU-power is available for the US community at various centers including a recent 5.7 Teraflop/s IBM Blue Gene /L eServer at San Diego. Additionally, about 40 Teraflop/s of computing power with a Petabyte-scale data storage is available since October 2004 within TeraGrid that operates a 40 gigabyte per second network and pools computational resources from Argonne National Laboratory, Center of Advanced Computing Research at Caltech, Indiana University, National Center for Supercomputing Applications, Oak Ridge National Laboratory, Pittsburgh Supercomputing Center, San Diego Supercomputing Center and Texas Advanced Computing Center. TeraGrid is used currently by the group of Prof. K. Schulten (University of Illinois at Urbana-Champaign) for the simulation of membrane proteins and the recognition of DNA by proteins. The availability of such large computational resources explains why massive simulations, like those of biological motors by Schulten's group, or the ribosome simulation by Los Alamos's groups are now possible only in the US. More interestingly, the US are reinforcing this scientific policy focusing more and more effort in developing computational centers devoted to computational biology. For instance, IBM has installed a 100 Teraflop/s eServer Blue Gene/L system at the Thomas J. Watson Research Center, dedicated to life sciences. As part of their Roadmap for medical research the National Institutes of Health (NIH) attempts to implement in the USA the core of a universal computing infrastructure for the life sciences. Four centers specialized in the area of biomedical computing have been funded in September 2004. These centers are supported by 5 year research grants with a total of 79.7 Million US$. Funding of three more centers is planed for 2005. The centers aim at "paving a future information highway dedicated for advancing medical research" and training a new generation of multidisciplinary biomedical computer scientists.

In order to stay competitive, computational biology/biophysics/bioinformatics in Europe will need to match the American resources. Otherwise, Europe will be outside of the post-genomic race and should rely in American groups for the deciphering the information hidden in current genomic data files.

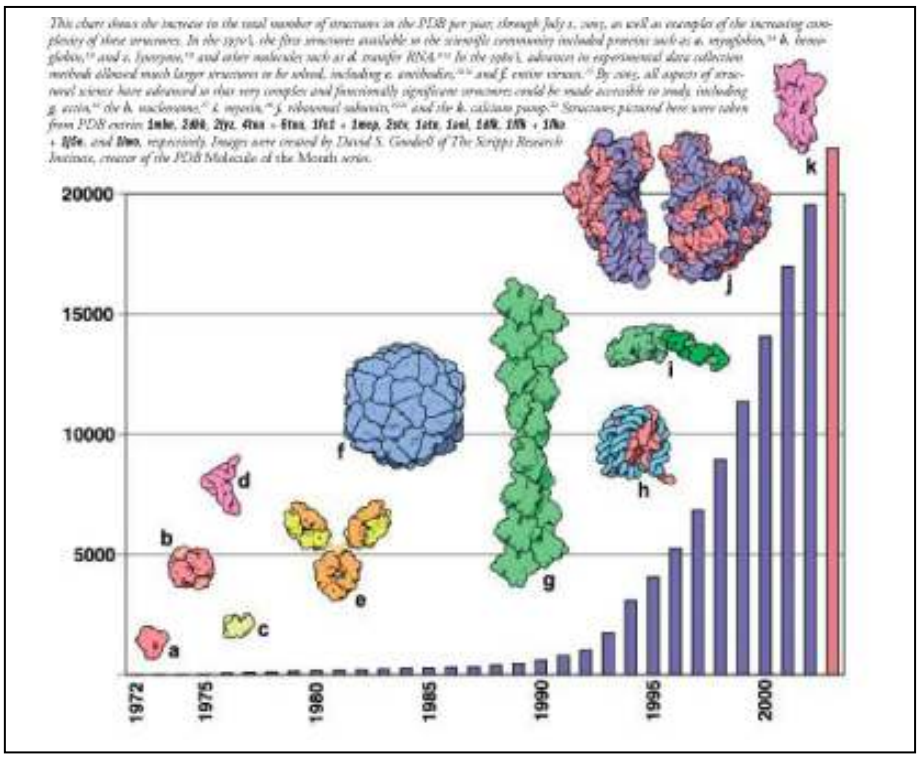


Fig 3. Evolution of the size of new structures deposited in PDB

4.4.2 Petascale Computational Challenges in the Life Sciences

The computational scenario in life sciences should be defined in the next few years by three levels: i) group resources, typically Linux clusters, where most of calculations should be done, ii) national facilities (40-100 Tflop machines) which will support massive calculations and iii) top of the pyramid European Center which will made possible calculations requiring Petaflop performance. Two considerations seem clear for the experts: i) An European supercomputer centre does not reduce, on the contrary increases, the need of smaller national centers, or the need of GRID-based infrastructures, and ii) the biological importance of the problem and not the fitting to a specific computational architecture should guide the use of computers in life sciences.

A European supercomputer centre(s) should support pure supercomputer projects – those of systems biology or atomistic simulations, which demand a fast interconnection between a large number of processors – together with others where fast interconnection between nodes is not as critical as pure CPU power ("massive" projects – those where the entire problem can be broken in pieces and send to individual CPUs, requiring little interchange of data e.g. typical bioinformatics data-mining projects). Finally, the European computer centre(s) should be able to deal with time-critical investigations, such as those related to research focused on epidemic catastrophes or emergent pathologies. While in an "ideal world" a large number (above 100) of medium 64-128 nodes platforms distributed across Europe might be useful for "massive" projects, a supercomputer centre is the best alternative to deal with the three type of projects in play, it being in fact the only solution that can handle both "pure supercomputer" and "time-critical" projects. Other alternatives such as GRID-computing are extremely unlikely to provide a global alternative to supercomputers in biology

4.4.3 Perspective of Petascale Computing for European Life Sciences

As noted above, the gap between Europe and US in the computational resources focussed in life sciences is increasing every year, which might have a very negative impact in our ability to be

competitive in this strategy area. This situation will be especially painful considering the very good positioning of Europe in the field:

i) Europe has very competitive groups in the area, covering a large range of expertises from biophysics simulations to image processing or data-mining. Europe has in fact some of the leading centers for research in bioinformatics, modelling and computational biology

ii) There is a clear commitment of the different States with research in Bio-computation, which becomes clear in the support to different national institutions working in this field and on the existence of several specific programs aimed to grant "state of the art" research in bio-computing.

iii) There is a very strong industrial biotech and biopharma environment. In fact, some of the largest and more active in research pharmaceutical companies are based in Europe, and there is a plethora of other smaller, but also very active companies that with institutional help can jump to the first line. All these small and large companies will largely benefit from a very large supercomputer facility in Europe.

4.4.4 Grand Challenges

Biology is a science in continuous and fast evolution, which makes very difficult to perform reliable predictions in a decade-time (consider for example that the genomic revolution started just five years ago!). The issues that appear in 2006 as Grand Challenges for Biology and where Supercomputer facilities might have a large impact include.

4.4.4.1 Pure Supercomputer Projects

a) *Systems Biology:* These projects are just in the beginnings, but appear as one of the most important issues of research in biology for the next decade. Major European (like EMBL) or National institutions are considering System Biology as the top priority, and are designing entire research programs devoted to this interdisciplinary activity. The main idea of systems biology is to represent using simple models the entire behaviour of cells, tissues, and organs, or to evaluate degradation routes predicting the final excretion product of any drug in any organism. Systems Biology projects start always from the massive information derived from genomic experiments (protein arrays, DNA chips, yeast two hybrid experiments,…) which is used to define the systems nodes, connexions and parameter, generating a giant networks defined with different resolutions and focus in different regions, and where due to the low quality of the experimental information, all nodes and connections have a probabilistic nature that increases even more the complexity of the system. Considering even the simplest cell we cannot expect to identify all the important players from experimental techniques. Simulation will then be the only possibility to detect for example hidden players (processes or proteins that are important only under certain conditions), or to understand the behaviour of the cell, tissue or organs under anomalous conditions. The simulations needed imply massive calculations, impossible to perform in a GRID-based platform, and out of the possibilities of any current computer centre. In any case, Europe should be ready in the next 4 years to host the first "in silico" cell, probably a computational implementation of the "subminimal cells" which are now being experimentally developed. Otherwise, Japan and the United States will take a dramatic advantage in the post-genomic race.

b) *Chromatine dynamics:* Once the genome has been solved, the next question is to determine how the genome regulates its function. Analysis of databases like TRANSFAC reveals that known sequence rules alone cannot explain the differential expression of genes. Physical properties of DNA must then play a key role in modulating gene expression. A particularly interesting issue of gene regulation is the dynamics of chromatine. We know that the DNA of eukaryotes is organized in nucleosomes which are then wrapped forming long fibbers. The organization of DNA in nucleosomes largely modifies the accessibility of transcription factors recognition sites playing then a key role in the regulation of gene function. The understanding of nucleosome dynamics, positioning, phasing, formation and disruption or modifications induced by chemical modifications, or by changes is a great challenge, which would imply extensive molecular

dynamics simulations on the scale of the microsecond for systems that contains from 10^5 to 10^6 atoms.

c) *Large scale protein dynamics:* To understand the physical reasons behind large conformational changes in protein is one of the main objectives of structural biochemistry. For example, fifty years after the discovery that sequence contain all the information for guiding protein folding we are still unable to reproduce this process and it is now clear that such knowledge will arrive only from physical based simulations, where physical force-field will be combined with atomistic molecular dynamics algorithms. The importance of protein folding simulations, which is now the focus of a large number of computational projects like Blue Gene, should not hidden the relevance of the study of other large conformational changes in protein. Particular, due to its biological and biomedical importance big challenges appear in the simulation of protein missfolding, unfolding and refolding (a key element for the understanding of prion-originated pathologies), allosteric changes, and conformational changes induced by protein-protein interactions. All these calculations imply extremely long molecular dynamics simulations (time scale between micro and millisecond) of systems containing 10^4-10^5 atoms. These simulations are possible only in very fast and parallel computer systems, with reasonable fast interconnection between processors.

d) *Protein association and aggregation.* It is clear that proteins are not in the cell are in a crowded environment and are continuously making interactions with other macromolecules. The genomic revolution is showing that such interactions are the responsible for the control of cell function. For example, cascades of protein-protein interactions are the responsible for increasing the expression of a given gene as a response to the binding of a ligand to a membrane receptor. One of the greatest challenges for the next decade will be the simulation of these crowded protein environments. Unfortunately, we are facing a formidable problem since protein-protein interactions are controlled by two mechanisms: i) a Brownian diffusive mechanism, and ii) an atomic detailed docking between macromolecules. Combined methodologies, where Brownian and molecular dynamics are combined with docking algorithms incorporating evolutionary information should be used in massive simulations. To be able to represent “in silico” the formation of the different protein complexes associated to a signalling pathway will be a great biological and computational challenge, which will help to bridge cellular biology with chemistry, opening the door for a better understanding of cellular function and to the generation of new drugs able to interfere in protein-protein interactions.

e) *Supramolecular systems:* Current simulation protocols and computers allow us to reproduce reasonably the interaction of small molecules with small or medium-sized proteins, but the correct representation of protein machines is still out of the possibilities of European groups. American groups, supported by massive computer resource have proven that, under ideal conditions, we can reproduce enzymatic catalysis, or the behaviour of large protein machines (like ATPases, ions channels,…) with more than 10^5 atoms. The challenge will be to analyze systematically how several of these machines (ribosome, topoisomerases, polymerases,…) work. The impact of these challenging simulations will be tremendous, not only for fundamental biology and biochemistry but also for nanobiotechnology application, since we know that molecular catalyzers and protein engines are several times more efficient than the best models developed by humans. Again we are facing big computational challenges since “biass” MD simulations in the range of the microsecond should be made in systems containing up to 10^7 atoms under different external conditions. Only supercomputers can provide answers..

f) *Medicine.* The last objective of all research in Life Science should be the improvement in our life quality. That is, all basic research should be addressed as to provide scientific information that can lead to the development of new, more efficient medical treatments. Genome sequencing, massive genotyping studies are providing us with massive information. But connecting this information with human health is not always simple. Particularly interesting is the simulation of the determinants triggering the development of multigenic-based diseases and the prediction of secondary effects related to bad metabolism of drugs in certain segments of population, or to the interaction of drugs with macromolecules others than their original targets. A great area of work appears for bioinformatics in this area, where the challenge is to combine database analysis with

very CPU-demanding techniques like maximum likelihood models, systems biology approaches, microscopic and mesoscopic simulations.

4.4.4.2 Massive Projects

a) Proteome-scale studies: Dynamome (proteome dynamics), structure prediction/ refinement, protein-docking, binding site prediction. We are facing here problems related to medium to large simulations including: molecular dynamics, docking, Monte Carlo, essential dynamics. The challenge here is to extend these calculations to more than 30000 proteins. A clear example is the generation of a dynamic database of proteins, a project already started in the US by V.Daguet using massive supercomputer resources. Other example will be cross-docking where libraries of 10^5 ligands will be screened against 10^5 possible targets (10^6 if models are used), or binding site comparisons, where we need to perform shape-shape $10^5 \times 10^5$ comparisons of 7- dimensional grids containing in the order of 10^6 points. Clearly, most drug-design studies, those that are guiding the generation of the new family of drugs in all pharmaceutical companies fall within this category.

b) *Genomics-scale simulations.* These include a large series of massive calculations for genome annotation, where pattern-matching techniques like BLAST are used to find relationships between different genomes, and techniques devoted to analysis of massive transcriptomic or proteomic data. Again these are simple calculations, but their cost can be prohibitive by any group as the input data increases. This combined with the time-pressure implicit to any genomic project recommends the use of supercomputer resources

c) *Data mining*. Clearly an increasing problem, as the size of the databases increases with a parallel decrease in their quality (compare Swisprot and TrEMBL). To search for a single database is a problem that can be done by any computer but to cross several of them, screening for a small signal embedded in large amount of noise can be extremely expensive. A prototypical example is the search of biological relationships between proteins or ligands by automatic analysis of the millions of bibliographic entries in MedLine. This, in principle simple procedure becomes in practice impossible for current computers as the deep and flexibility in the search increases and not only abstracts, but the entire text are analyze.

4.4.4.3 Time-critical Studies

a) *New pathologies.* AIDS is the best example of how structural based drug design and simulation techniques Atomistic simulations can be helpful to predict molecular targets for new pathogens and to predict the pathological impact of mutations. This type of calculations, which can help to warn the community against possible pathologies might turn into time-critical projects when the new pathology is already present and society ask the scientist for a fast response which is not possible to achieve without centralized organization able to focus in a single day large resources in a single project. Examples that illustrate that the scenario described above is real include AIDS, Mad Cow Disease. SARS or the Chicken Flu.

b) *Disease evolution.* Again, society wants to have reliable models how the spread of pathologies, Mad Cow Disease Ebola or Chicken Flu are excellent examples. Here supercomputer centre should provide the perfect platforms for epidemiologist to perform fast simulations able to trace the possible evolution of pathologies under different social or environmental conditions, helping then in prevention or isolation of pathologies.

c) *Lethal secondary effects*. Europe has witness in the last decades very important health problems related to secondary, unexpected effects of drugs or food. Alimentary problems like that suffered in Spain with toxic oil, or the systematic appearance of side pathologies related secondary effects of pharmaceutical drugs, which were not detected during clinical trails are good examples. Modelling and systems biology methods are now able to provide very fast first hypothesis on the causes of the unexpected secondary effects, helping then in reducing the cost in lives of this type of health problems.

4.4.5 Inter-play between Biology and other Science-areas

After the two scientific meetings several relationships appear between different areas and Biology. This generates a network of common methodologies that can help to design more rational supercomputer centers. Synergies are especially clear with the following disciplines:

Chemistry. We share many common methodologies, which are in fact those leading to a greater consume of CPU time. Algorithms for performing classical MD simulations in chemical and biological systems are identical. Methods based on QM/MM formalisms are used for both study of reactivity in enzymes or in large chemical cluster. Similarly, Carr-Parinello MD simulation, despite being developed to study large chemical systems is being largely used by biological community to study especially complex enzymatic processes. The common interest between chemistry and biology extends also to other areas like mesoscopic simulations, drug design or docking experiments.

High energy physics. Again, the central core of many of the problems used here is the same than in biology: programs for particle-particle interaction in real or Fourier space. Mesoscopic modelling is also of interest for both communities. This leads us to share with this community a platform of common algorithms. As a community, we were not able to detect such common links with other important areas of physics like astronomy.

Engineering and Climate simulations. Several of the problems faced in systems biology fall actually in the interface between biology and engineering, and in fact, several of the grand challenges presented in the engineering report could fit into the first category of biological super-computer grand challenges. It is however unclear for us our synergies with many other of the projects of interest in the engineering area. We believed we might share common algorithmic interest with engineers working in the climate area, but based on the two meetings we cannot claim this synergy.

4.4.6 Societal Benefits

These are very clear, and are mentioned in several places in the text already. Just to elaborate on points already made:

- genome annotation (which is largely computational)is essential for identifying new drug targets;
- protein fold recognition and structure prediction are essential to allow the engineering of new proteins, as catalysts, biosensors, or nanomaterials (motors, fibers);
- computational biochemistry is a key part of the drug development pipeline, through docking and structure prediction calculations;
- protein misfolding is the basis of several major diseases, such as amyloid fiber diseases (Kreuzfeld-Jacob and mad cow disease);
- emerging pathologies (SARS, avian flu) require very fast computational responses: massive drug design projects were recently performed in a very short time in France by HPC and in the UK by grid computing; several promising lead molecules were identified;
- simulations if disease evolution have obvious implications for public health;
- simulations of biological nanomachines (chromatin, membrane transport systems) are essential to understand the molecular mechanisms of cancer and other diseases;
- computational modelling of the cell will play a role in understanding cell growth and differenciation, and ultimately phenomena such as embryogenesis and development;
- whole-organ modelling will be increasingly a tool for medicine and surgery in humans and animals.

4.5 ENGINEERING

4.5.1 Introduction

The impact of computer simulation in engineering has been significant and continues to grow. Simulation allows the development of highly optimised designs, the investigation of hazards too dangerous to test and reduced development costs. In parallel new scientific investigations are developing understanding in areas such as turbulence and flow control necessary for future engineering concepts.

A clear trend is towards complete system modelling, as higher fidelity information is required and automatic optimisation of the whole system, including interference between sub-components, is required. Such simulations encompass the interaction of different aspects of a system, requiring modelling from multiple disciplines. Examples from the power industry include the determination of material degradation under irradiation and the prediction of ageing due to structural vibrations in flows. Typical problems bring in phenomena at radically different scales. The engineering simulations of today and the future therefore are **inter-disciplinary** and have **multiple scales**, leading to requirements for supercomputing resources which are almost unlimited. The demand for technological solutions to some of the great problems of the modern world, such as cheap, safe and available transport at low environmental cost or low cost power generation with low environmental impact, implies an emphasis on large scale computations to research new concepts and the underpinning science.

The requirement from engineering for European centres for supercomputing is made here through a small number of example applications. These are neither exhaustive nor exclusive but are intended to show some of the potential for advances that large scale computer facilities could enable. The examples are drawn from aerospace, biomedicine, the power industry, the environment and disaster management.

4.5.2 Challenges in Engineering

4.5.2.1 Challenge: Complete Helicopter Simulation for Next Generation Rotorcraft

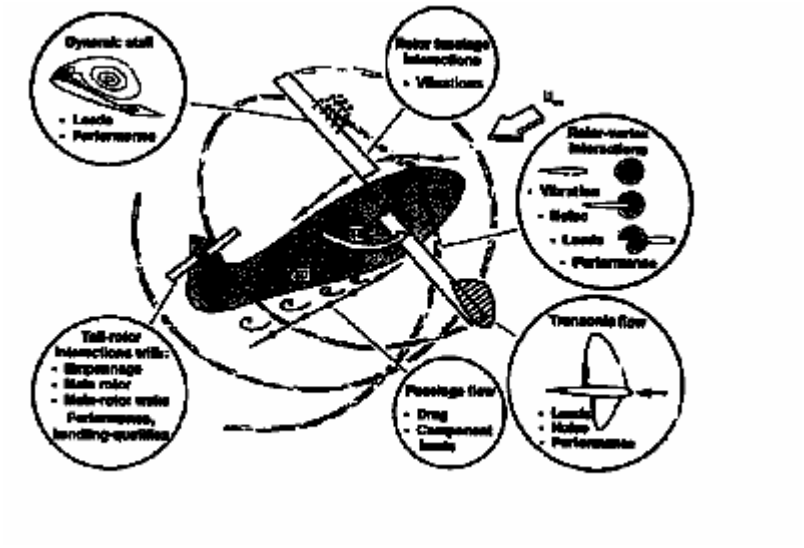
The ability for hover, vertical takeoff and landing gives the helicopter a unique place in the transport system, e.g. with a significant role for rescue operations, police missions, traffic control, and remote transport of goods. However, the helicopter is still a costly machine with high maintenance costs, poor fuel efficiency, moderate ride quality, remarkable noise generation and air pollution. These short comings stem from the complexity of the flow around a helicopter. The complex unsteady aerodynamics encountered features shock waves, separation, vortices and interaction. In addition, the dynamics of the blades and the rotor head assembly are complicated. Thus, the ACARE 2020 vision defines targets for an improved efficiency and ride quality, and a reduced environmental impact of rotorcraft. The European helicopter industry has a strong tradition of innovation in technology and design. Computational Fluid Dynamics (CFD) based simulations of aerodynamics, aeroacoustics and coupling with dynamics of rotorcraft already play a central role and will have to be improved further in the design loop. There has been consistent funding from the European Union to develop the codes necessary for rotorcraft analysis, e. g. DACRO, ECARP, HELISHAPE, HELIFUSE and EROS. The GOAHEAD project recently started to generate a database for the validation of complete helicopter simulations and involves the cooperation of the European helicopter industry (Eurocopter, Agusta-Westland), national research centres, and universities.

Progress has so far been made towards the high fidelity simulation of helicopter components like isolated main and tail rotors in hover and forward flight though researchers are attempting the simulation of the complete helicopter. Grid treatments for moving components, validation for the rotor flows, including dynamic stall, and the fuselage flows have been encouraging. Trimming and computational structural dynamics (CSD) were incorporated into the simulations. Several main

challenges still remain. First, transition is important at the Reynolds' numbers of interest and the practice is to prescribe transition or to use a heuristic stability criterion. Secondly, there is a high demand for grid points and time steps due to the need to preserve structures such as vortices until an interaction of interest happens. Prediction of the acoustic signature of the helicopter is also challenging requiring both high order numerical schemes as well as a significant number of grid points. Grid adaption is likely to be of some use but, due to the complexity of the flow field, a large number of grid points and high order discretisations will still be required.

The main problem with design work is the difficult scaling from model rotors and helicopters to full scale. Froude-scaled and Mach-scaled tunnel tests are possible but then the dynamics of the blades is not well represented, or vice-versa. It is therefore necessary for CFD and CSD methods to tackle the full-scale problem. The benefits of complete system simulations are potentially significant. First, interactional aerodynamics phenomena between main rotor/tail rotor/fuselage play a significant role in the design of a new helicopter, especially due to the non-linear aerodynamics and blade dynamics involved. Secondly, phenomena such as noise generation and vibration require consideration of the complete system. Novel tip shapes and active blade design featuring flaps, high harmonic and fluidic controls have the potential to alleviate some of the adverse noise and vibration characteristics. Thirdly, better prediction of drag requires a detailed simulation of all sources, including skin friction. Fourthly, fundamental work on transition using DNS and LES of model problems is required to allow transition at full-scale to be modelled and possibly predicted by routine calculations. Next hybrid adjoint methods running on parallel computers with simulation and experimental data will be necessary in order to design, analyse and optimise the future-generation helicopters. Finally, new or unusual concepts such as the tiltrotor require detailed assessment to gain widespread acceptance.

The CPU requirements for future simulations are not yet known but are bound to be substantial. Since little experience is available for full-scale/full-helicopter simulations researchers are still in the process of discovering the required grid resolution. An accurate simulation of an isolated rotor on today's supercomputer systems (3 GHz processor) would roughly require 5 million grid points per blade, 2000 time steps for each turn of the rotor, around 5 rotor revolutions until a trimmed state is reached and around 30000 CPU hours. This may well increase as the fidelity of the data required by the researchers and engineers increases. Accurate drag and sectional pitching-moment predictions require viscous flow simulations and this can easily triple the above estimate. Adding a fuselage and a tail rotor will increase this requirement to around 150000 CPU hours. This is due to the bluff body aerodynamics of the fuselage with shed vortices and massive flow separation but also due to the complex interaction between the fuselage and the rotors. Several calculations that are an order of magnitude larger would allow an assessment of issues such as grid dependency and establish confidence levels for smaller simulations on local facilities. The incorporation of flow control devices will potentially boost the computational requirements significantly. Finally, the helicopter system requires a delicate balance between the aeromechanics of the aircraft, the flight control system and the pilot. Simulation of manoeuvres, in contrast to design conditions, becomes an inherently multi-disciplinary problem with contributing modules requiring different treatments when it comes to parallel computing. The aerodynamic analysis will have the lion's share but also a large effort will be required to simulate the effect of the pilot's action and of the control systems on the flight mechanics. This creates a challenging project where different solvers and modelling techniques must come together in a single parallel environment and require a PFlop/s sustained performance.



Schematic of the Helicopter Flow

4.5.2.2 Challenge: Biomedical Flows

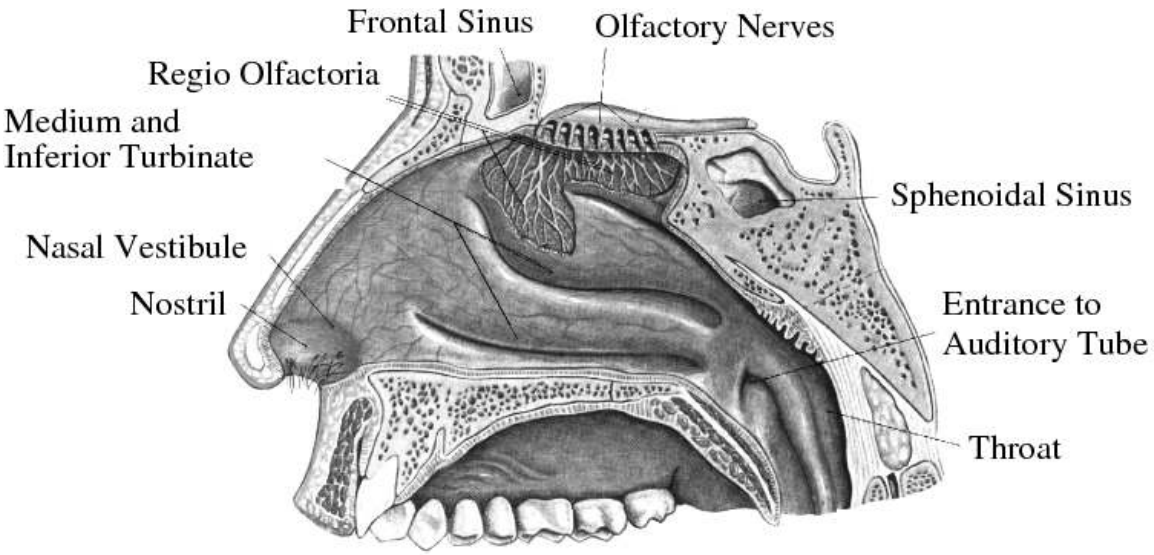
Today biomedical fluid mechanics can improve healthcare in many areas. Among other applications, there is quite an intensive research effort in the field of the human circulatory system, for example blood flow disturbances caused by arterial malformations, the artificial heart or heart valve prostheses are investigated, the respiratory system with the nose flow and the upper and lower airways, and the human balance system, e.g., the fluid mechanics in the cochlea. Although experiments have significantly improved the understanding in the field of biomedical fluid mechanics, numerous questions, the answers of which need a high resolution of the flow field, of the surrounding tissue, or of their interactions, require a detailed numerical analysis of the biomedical problem. The flow problems encountered are characterized by an extremely complex geometry, moving boundaries, locally and globally unsteady flows, fluid-structure interactions, Newtonian and non-Newtonian fluids, and by single and multiphase flows to mention just a few phenomena. Any computational simulation therefore faces an extreme challenge. Similar features defined by quasi discontinuous flow, transition, and intricate configurations also occur in the numerical analysis of flows over aircraft, helicopters, and turbine blades.

Investigations of arterial malformations, including stenoses (narrowed vessel cross section), aneurysms (widening of the cross section, often with balloon-like enlargements) and short-cuts between arteries or an artery and a vein are the subject of many research projects in recent years (including the Esprit project *BloodSim*, FP5 *SimBio*, FP5 *GEMMS*, or the recently started *@neurIST* and *Coast* projects and the German *MediGrid* cluster). The available numerical algorithms and complementary experimental methods have reached a point where small regions of interest can be studied in detail as a showcase with simplified physics (e.g. Newtonian flow). However, to apply these techniques for large scale preventative medical checkups, including all physical effects, computational power is required which exceeds today's largest systems by at least two orders of magnitude.

The flow through the human nasal cavity and its numerical analysis is now discussed in more detail. The nasal cavity has to satisfy a variety of different functions. Besides respiration it is responsible for moistening, tempering, and cleaning the air. These functions are expected to depend strongly on the complex internal geometry of the nasal cavity. Impaired nasal respiration especially under normal breathing conditions, i.e., in everyday situations, is a common and widespread disease, which makes nose surgery one of the most frequently performed operations in the western world. Unfortunately, the success rate, i.e., the number of people having no problems after surgery, is by no means satisfactory. To enhance this situation, it is desired to predict numerically the flow field inside the nose and its relation to geometric changes. This can finally lead to a computer assisted surgery (CAS) tool, that will help on the one hand to understand the flow in the extremely complex human nasal cavity, and on

the other hand to derive criteria for the better prediction of the outcome and to improve the success of nose operations.

The challenge of integrating a flow solver in a CAS tool is driven by the requirement for quick response times. The surgeon will have to modify the nasal cavity geometry several times to improve the fluid mechanics and medical outcome of the surgery. This means the flow field has to be recomputed for any geometrical adaptation. Considering an adequate resolution of the complex three-dimensional nasal cavity geometry is approximately 10 million grid points, an average Strouhal number is approximately $Sr = 0.06$, then to simulate the unsteady human respiration cycle with a turn-around of a couple of hours implies huge computational resources. Currently, a simulation of one respiration cycle on a 1 million cell mesh at $Sr = 0.8$ on an SX8 processor requires about 10 days computing time. This means that the computing performance has to be increased by a factor of roughly 1000 to be able to sufficiently resolve the intricate flow physics needed to determine a convincing solution and in an acceptable amount of time.



Human nasal cavity

4.5.2.3 Challenge: Gas Turbines and Internal Combustion Engines

Combustion represents more than 80% of energy conversion in the world and has a strong impact on the environment (greenhouse gases, pollutant emissions). The objective of combustion studies is to better understand and model physical phenomena to optimise, for example, gas turbines (aero-engines or power generation) or internal combustion engines, in terms of costs, stability, higher combustion efficiency, reduced fuel consumption, near zero pollutant emissions and low noise. Computational Fluid Dynamics (CFD) offers to design engineers the unique opportunity to develop new technical concepts, reducing development costs by avoiding extensive and very expensive experimental campaigns. From an economic point of view, industrial companies involved in propulsion and energy systems are among the biggest employers in the European Union. To give them more efficient and cost effective system designs is crucial support to promote their competitiveness on the worldwide market.

Scientific challenges in gas turbines or piston engines are numerous. First, a large range of physical scales should be considered from fast chemical reaction characteristics (reaction zone thicknesses of about tens of millimetres, 10^{-6} s), pressure wave propagation (sound speed) up to burner scales (tens of centimetres, 10^{-2} s resident times) or system scales (metres for gas turbines). Turbulent flows are, by nature, strongly unsteady and large eddy simulations (LES), where the largest flow motions are

explicitly computed while only the effects of the small ones are modelled,[3] appear as a relevant CFD tool, especially when combustion instabilities or cycle-to-cycle variations in piston engines, occur. Chemistry involves hundreds of chemical species and thousands of chemical reactions, and cannot be handled in numerical simulations without adapted modelling. To design cooling systems requires the knowledge of heat transfer to walls due to convection and radiation.

Usual fuels are liquid, storing a large amount of energy in small volumes (about 50 MJ / kg). Accordingly, two-phase flows should be taken into account (fuel pulverisation, spray evolution, vaporisation and combustion). Solid particles, such as soot, may also be encountered. Interactions between flow hydrodynamics, acoustics and combustion should also be considered as they may lead to strong combustion instabilities and to the system destruction in a short time. Control devices may contribute to avoid these instabilities either based on passive (geometry changes, Helmholtz resonators) or active (actuators) techniques.

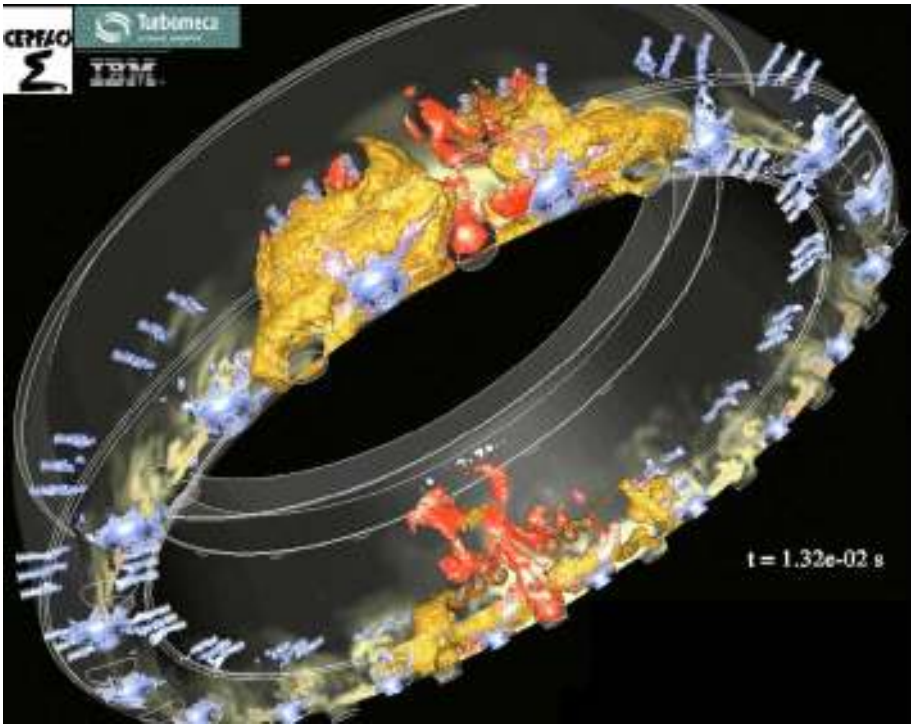
A specific difficulty in gas turbines is linked to the intrinsic deterministic unsteadiness of the flow in compressors and turbines, involving rotor/stator interactions along a large number of rows (up to twenty in industrial compressors). To consider the flow through the whole components is mandatory for design because of blade number off (the minimal flow unit for three or more rows is the whole wheel). The gas temperature profile at the outlet of the combustion chamber should also be compatible with turbine materials and cooling systems. Large eddy simulations have already been successfully performed in realistic gas turbine combustors or piston engines[4] but remain limited today to a part of the whole system and to few cycles, corresponding to short physical times (typically some tens of milliseconds) because of computational costs. These simulations allow the analysis of some physical features but are clearly insufficient for design purposes or to investigate complex interactions inside the whole system or transient phenomena such as ignition or regime changes.

The potential of high-end HPC systems for simulations of combustion systems is illustrated here using a computation performed recently by CERFACS (Toulouse, France) on an IBM BlueGene/L machine (Rochester, U.S.A.). This simulation of the ignition of a helicopter gas turbine (full geometry, 18 fuel injectors, about 20 millions numerical cells) is one of the largest LES simulations in turbulent combustion and took 60 000 CPU hours on 1024 processors[5] (about 2.5 days of wall clock time). Unfortunately, only 4 of the 18 fuel injectors have been ignited at the end of the computation corresponding to only 13.2 ms real time. To reach a steady state regime, 100 to 200 000 additional CPU hours are required. The simulation assumes gaseous fuel burning according to a simple chemistry (no pollutant formation) and radiative heat transfers are neglected. To describe spray combustion would require two times more CPU. More refined physical description could already be incorporated but are out of reach of available computational resources (up to 100 times more CPU hours on the same grid mesh). Even corresponding to a simplified test case, this BlueGene simulation, performed in U.S.A. due to a lack of available resources in Europe, will probably remain for a long time a unique investigation.

[3] Obviously, the flow motions explicitly determined in large eddy simulations should be resolved on the numerical grid mesh. Accordingly, the thinner this grid mesh is (and the higher the computational costs!), the smaller will be the unresolved scales and thus the smaller the importance of sub-grid scale models.

[4] In LES, instantaneous cold and burnt gases regions, that behave very differently in terms of turbulence, pollutant emissions or radiative heat transfers, are identified at the resolved scale level, making combustion modelling easier. This point explains why LES achieves very good results while usual RANS (mean flow description) methods fail.

[5] The AVBP code, developed by CERFACS and IFP in France for gas turbines, furnaces and piston engine applications, runs very efficiently on parallel machines. The speed up factor on BlueGene/L machine has been found to be 4078 on 4096 processors.



Large eddy simulation of the ignition of a helicopter gas turbine. Two hot gas streams (red) are injected in the combustion chamber to ignite the fuel entering the combustor through 18 injectors, apparent on the front face. Dilution jets (groups of three small jets) are also visible. The flow is going from front to back. Only four injectors, surrounding the hot gas injectors on top and bottom, are ignited after 13.2 ms. This computation, performed by CERFACS (Toulouse, France) on an IBM BlueGene/L machine in Rochester (U.S.A.) took 60 000 CPU hours (2.5 days on 1024 processors) in spite of deploying quite simple physical models. The grid mesh contains about 20 millions numerical cells.

4.5.2.4 Challenge: Forest Fires

Regularly forest fires strongly affect south European countries such as Portugal, Spain, France, and Greece and because of climate change, may concern northern regions in the future. The social impact is very important and is concerned with land, buildings, human and animal life, agriculture, tourism and the economy. Accordingly, the development of reliable numerical tools able to model and predict fire evolution becomes of importance in terms of safety and protection ("numerical fire simulator"), fire fighting and could help in real time disaster management.

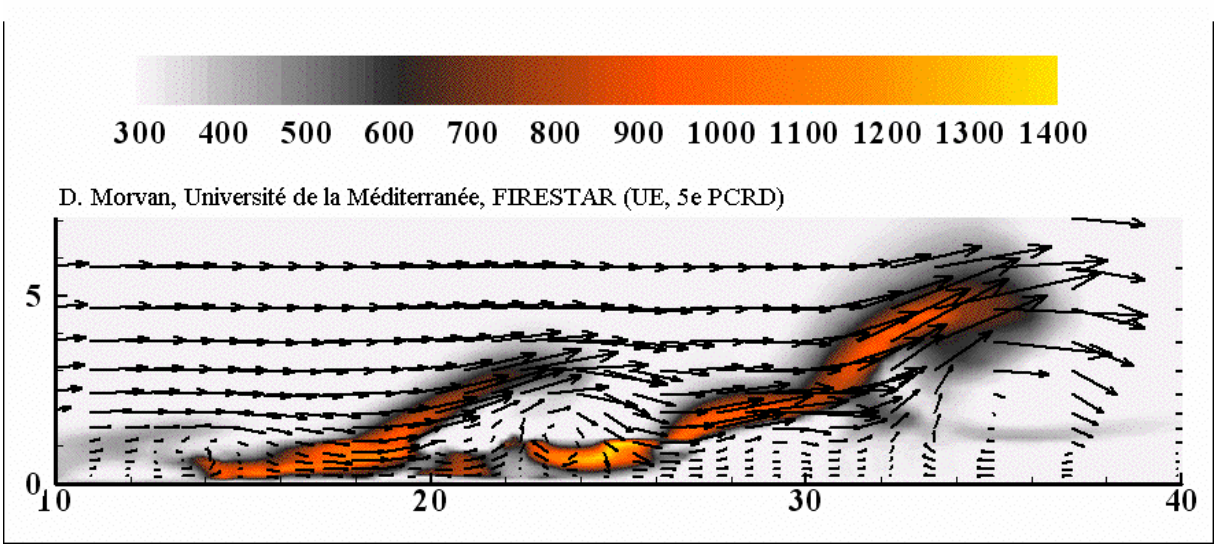
From a scientific point of view, forest fire predictions are very challenging. First, as for other combustion problems (see, for example, gas turbine challenge), a large range of physical spatial and temporal scales is involved. Chemical mechanisms are especially complex (for example, the transformation of wood in burning gases, depends on wood nature, moisture and involves numerous chemical species). Forest fires are also strongly controlled by radiative heat transfer, generally neglected in ordinary combustion computations. These transfers involve long distance interactions and require specific treatments. Buoyancy effects (large scale flames) as well as interactions with the local meteorology (winds, moisture) and the local topography (hills, valleys,…) should also be taken into account and need adapted models when these features are not relevant in burners. The simulation of the fire fighting, for example by dropping fluids without or with retardant, is also a challenging research of crucial importance. Note than the optimisation of the fighting can be recast in terms of optimal control problems, using genetic algorithms.

One of the specific features of forest fires is the size of the physical domains. Up to now, most refined combustion computations are performed over small domains (about tens of centimetres, corresponding to typical gas turbine burners and internal combustion engines) and/or in two-dimensional geometries, but forest fires may develop over tens of kilometres and are intrinsically three-dimensional. Accordingly, such computations need very large meshes, as the local instantaneous flame thicknesses

are about millimetres and typical flame scales about metres, and are out of reach of currently available computers. Because of the physical phenomena involved in forest fires, large eddy simulation, where unsteady large flame and flow motions are explicitly computed, appears as the most promising tool but requires heavy computational resources.

To perform relevant forest fire computations will require significant developments because such numerical simulations are probably now less mature than for gas turbine combustors or internal combustion engines. Most refined simulations have only been performed on two-dimensional geometries. In fact, researchers are faced with some crucial choices: to develop simple propagation models, incorporating quite roughly the key physics but able to run easily on large computational meshes (10 to 100 millions cells[6]) or to optimize the numerical resolution of more complex models. Both approaches should be conducted in parallel with different objectives ("fast prediction tools" versus scientific developments). Also, available fire codes are today probably unable to run efficiently on massively parallel machines and computational developments are needed.

To estimate numerical requirements is very difficult as efficient numerical tools to address forest fires on massively parallel machines are not available today and methods are not mature. Nevertheless, an order of magnitude, based on the known performance of combustion codes in gas turbines or internal combustion engines, assuming some scaling factors, is summarized in the requirement table.



Numerical simulation of a fire propagating from left to right over a shrubland similar to the ones found in Galicia (North-west of Spain) and submitted to a cross wind. This two-dimensional simulation has been performed on a dynamically adapted grid mesh of 235 x 73 points. The flow velocity (vectors) is superimposed on the gas temperature (colors). Dimensions are given in meters. The simulation of the propagation over 40 m required about 14 CPU hours on Intel Xeon 3.6 Ghz processors. Simulation performed by D. Morvan (Université de la Méditerranée, Marseille, France) in the framework of the FIRESTAR program funded by the European Union (FP5). These simulations correspond to the state of the art today and to extend them to three-dimensional configurations over large domains is a challenge for high performance computing.

4.5.2.5 Challenge: Green Aircraft

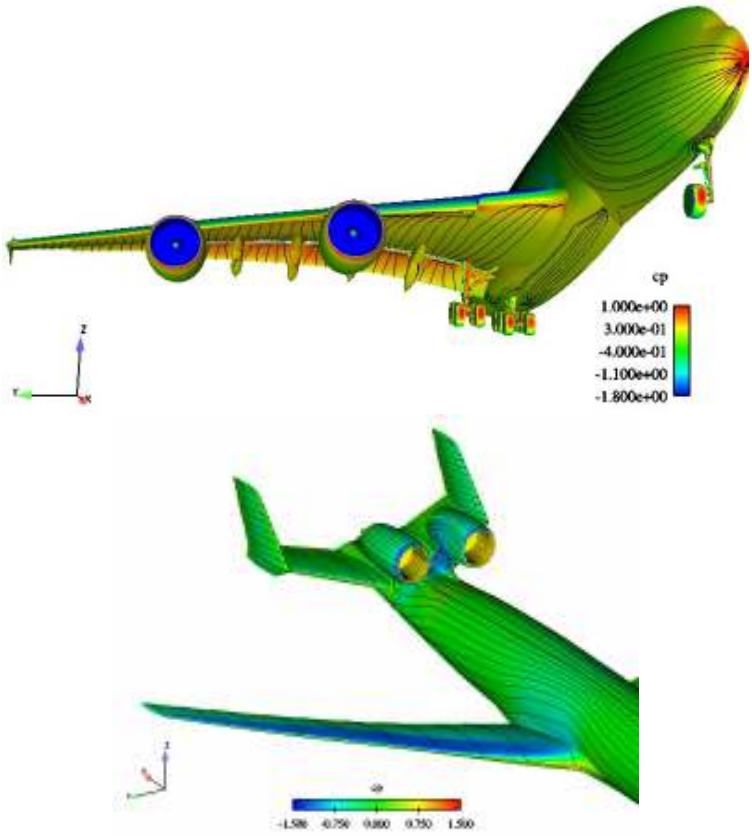
ACARE 2020 provides the politically agreed targets for an acceptable maximum impact of air traffic on people and the environment, while at the same time allowing the constantly increasing amount of air travel. The goals deal with a considerable reduction of exhaust gas (NO_x 80%, CO_2 50%) and noise (10-20 dB). Air traffic will increase by a factor of 3, accidents are expected to go down by 80%. Passenger expense should drop (50%) and flights become largely weather independent.

[6] Adaptative grid techniques with local refinement in the vicinity of the instantaneous flame front would probably allow a drastic reduction in the size of the overall computational meshes as most of the phenomena take place in this region.

It is clear that these targets cannot be achieved by enhancement of existing technology, and that they require essential technology jumps. The "Green Aircraft" is the answer of the airframe as well as engine manufacturing industry. It is supposed to minimize environmental impact and at the same time provide more transport capacity through alternative operational scenarios. In conjunction with the European research community, the aircraft industry has started investigations on the large scale deployment of flow control techniques (hybrid laminar flow, dimples, plasma, etc.) as well as on radically different configurations and design solutions (blended wing/body, etc.). However, it is only by a far more productive high quality numerical simulation and optimisation capability that such a challenging development will be possible. It will be indispensable to be able to compute the real aircraft in operation, including all relevant multi-disciplinary interaction. Major simulation and optimisation challenges are:

- Real-time simulation of aircraft in flight requires full unsteady Navier-Stokes computation of the complete aircraft, based on RSM turbulence models and DES up to LES as well as integrated transition prediction, coupled with the full Finite Element Model of aircraft structures and flight control law interaction.
- Aircraft aerodynamic loads for the whole flight regime requires new methods to investigate in the shortest possible time frame all possible and even unlikely flight situations with respect to the impact on aircraft structures.
- Noise source and impact requires the full development of noise source mechanisms, acoustic radiation and noise impact simulation tools which compute acoustic disturbances on top of aircraft flow.
- Optimisation of aircraft design requires a fully coupled simulation of the flow around a parametrised aircraft configuration and surface shapes covering a reactive structural model within a sophisticated optimisation process.

The computational power needed to cope with full aircraft simulation and optimisation – as described above - is 10^7 – 10^8 higher than today's capability. Only such an increase in speed and respective memory will enable the aircraft industry to develop the necessary aircraft technology and optimise their product from both an economical and ecological point of view. US research and industry is prepared to accept the challenge and develop solutions.



Numerical simulation of aircraft in design (source Airbus).

4.5.2.6 Challenge: Virtual Power Plant

Safe production of high quality and cost effective energy is one of the major concerns of Utilities. This is not only a source of competitiveness for them, but also for their costumers and especially for companies for which energy is a major contributor to their costs. Several challenges must be faced to achieve this goal, amongst which are extending the lifespan of power plants to 60 years, guaranteeing the optimum fuel use or better managing waste. While numerical simulation has been a strategic tool for decades for utilities that produce, transport, distribute and sell energy, these challenges will not be accessible with old generation simulation tools and computers. We need to address very fine simulations in order to provide reference computations for less detailed approaches, to couple various disciplines (fluid-structure interaction, neutronics and thermalhydraulics interaction, …), to do multiscale calculations (from microstructure to macrostructure), and to compute actual (and not simplified) geometries. Petascale machines are required to perform most of these advanced simulations along with a new generation of codes and simulation platforms. Strong collaboration is also required between physics (thermalhydraulics, material science, structural mechanics, chemistry,…), numerics and computer science (including development of parallelism, preprocessing and postprocessing).

The Virtual Power Plant challenge is divided into two sub-challenges :

- Safety and optimization of power plants,
- Design of future plants.

These share common actions. Some of the underpinning work is also briefly presented below.

Multiscale modelling of behaviour and degradation mechanisms of materials

Macroscopic properties of materials are strongly connected to local properties of their components. This is well known for composites but this is also true at a lower scale with a significant effect of microstructure. The ability to predict and control this relationship between microstructure and mechanical properties, damage and durability of structures is a challenge of major importance. This includes the ability to design new materials. In order to address this objective, the challenge is the development of an efficient numerical tool to model microstructure evolution during ageing of current facilities and to predict material and structure degradation. Control of initial and also current yield stress, fatigue limit, stress to fracture of a material submitted to irradiation, thermal ageing, mechanical damage, is a powerful advantage to ensure integrity, safety and reliability. That way the public and safety agencies should significantly increase their confidence in the safety studies, especially for components under very complex aggressive environment (irradiation, ageing of steels and concretes). A computing power of one Pflop/s is needed for handling realistic investigations in this field. This issue is strongly connected to the challenge "Multiscale description of mechanical characteristics of materials" proposed in the Scientific thematic area: MATERIALS, CHEMISTRY AND NANOSCIENCE. In this section, nanoscale is considered, with specific numerical tools including ab-initio computation, Molecular Dynamics, Monte Carlo techniques and Discrete Dislocations Dynamics codes. Those numerical tools are necessary to predict microstructure evolutions induced by ageing, environment, irradiation, but an additional mechanical analysis is required to derive the final mechanical properties from microstructure evolutions. This is covered here, in the ENGINEERING Scientific thematic area.

Improving the Critical Heat Flux Prediction.

Present DNS tools with Interface Tracking Methods (ITM) have already been able to compute a mono-site boiling pool but they still need to be improved to investigate multi-site boiling up to Departure from Nucleate Boiling (DNB). Such computational techniques may provide information on many physical parameters like the bubble departure diameter, the frequency of the bubble detachment, the heat transfer due to vaporization, liquid heating, wall quenching after a bubble departure. The target configuration will represent a significant portion of a Pressurized Water Reactor (PWR) and may approximately have the following features: calculations involving about 1 billion grid points and 5000 processors requiring about 700 Tflop/s for one week. Parametric studies will be necessary to get enough information for modelling purposes.

Distribution of flow and induced vibrations in a group of fuel rod assemblies

Accurately predicting the flow distribution inside a group of fuel rod assemblies while fully modelling the actual geometry of the rods and mixing grids typifies the computational challenge. Today the most refined LES industrial simulations running on 10 Million Cells are limited to a part of a fuel assembly without the representation of the mixing grids. It is estimated that 1 billion cells is necessary to accurately compute a significant part of a group of fuel rod assemblies (including the mixing grids). Moreover, thanks to the progress in CMFD (computational multiphase flow modelling) made in the recent years, a full computation with two-phase flow modelling (with RANS modelling for turbulence) will be within grasp around 2009. This simulation will need 500 Tflop/s for one week.

Lifetime of key components and hypothetical incidental scenarios.

Scenarios are currently studied in order to analyse the impact of hypothetical incidents on the lifetime of some key components. One of the most relevant simulations is the so-called PTS (Pressurized Thermal Shock). These transient (7000 s real time) scenarios involve both single phase or multi-phase flow modelling as well as thermal-mechanics, neutronics and system code coupling. Between 5 to 10 million cells should be suitable for an acceptable description of a complete vessel. We estimate that a single run will need about one Pflop/s for 2.5 months.

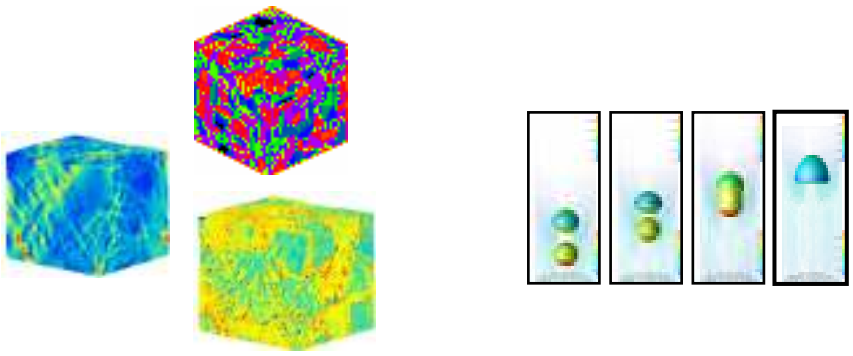
Prediction of the dynamic behavior of a power plant.

In order to predict the dynamic behavior of a power plant, 3D transient numerical simulations are required, coupling neutronics, thermo hydraulics (single phase or two-phase flow) and thermo mechanics. This problem also involves multi-level and multi-model aspects for each physical problem,

since a good predictive simulation must incorporate local as well as global effects (a wide range of events modifying the system are geometrically localized). Today we are able to solve this coupled problem in an industrial way, using a strict explicit time scheme, chaining one model to another. There is a strong need for more detailed and precise simulations and to handle more complex transients. A complete 3D core simulation, including coupling between neutronics, thermalhydraulics and fuel, will need about 500Tflop/s.

Other issues.

Industrial needs require a large number of parametric studies, usually within a rather short period. The very large calculations described above each need a Pflop/s machine in their own right. But, if such parametric studies are urgently required, the need is accordingly much more important. In addition, associated pre- and post-processing must be addressed. This includes efficient visualization tools, acting as an interface for the expert to understand and analyse the computations. The corresponding meshing tools have to be improved. This new generation of codes will need appropriate performance and tuning software tools, adapted numerical multi-level methods, adapted programming paradigms and software architecture. The end user will require a new generation of distributed software simulation environment that allows the coupling of HPC codes from different disciplines, the inter-operability of CAD, (re)meshing, computation and visualization.



Materials and Bubble Simulations

4.5.2.7 Challenge: Complete Launcher Simulation for Next Launcher Generation

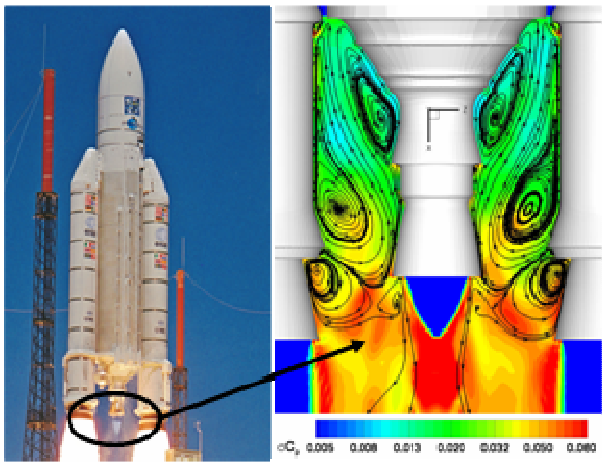
The preparation of new launch systems in Europe, which were required to respond better to the future European institutional needs and to maintain the long-term competitiveness of the European launcher sector, has been a fundamental element of the European strategy for access to space for several years. This concept was recognised on 16 November 2000 by the EU Council and ESA Council at ministerial level jointly in the Resolution on a European Strategy for Space, and reaffirmed by the ESA Council at ministerial level on 15 November 2001 in its Resolution on Directions for the Agency's Evolution and Policy "Space serving European Citizens", and more recently in its Resolution on 2010 perspectives for the European launcher sector, adopted on 27 May 2003. According to the European strategy for access to space outlined in the Launchers Strategic Plan 2006-2013, the European guarantee of access to space will rely on Ariane and Vega, including their possible evolutions, until ~2020, so as to amortise in production any investment made on these two launchers and on their improvements. This date is a consequence of the long development time and level of budget required mainly because the design today still relies largely on semi-empirical methods. Indeed, the development of a new conventional launcher requires more than 10 years at an investment rate of €1 billion per year. Furthermore, rocket based systems cannot be tested gradually but at once. Any failure results in the complete lost of the vehicle and this ought to be the case for almost all new development up to date. The major concerns of the aerospace industry in designing new launchers are discussed below.

One of the most important challenges is the over expansion regime of a future high area-ratio nozzle, in particularly flow separation and consequences on side-loads and thermal loads. The most significant parameter for aerodynamic performance of propulsion engines is the area ratio of expansion of the nozzle. Presently the area ratio is limited by the need for attached flow conditions for all ambient pressures that occur during the ascent of a vehicle. Further increase of the area ratio would result in flow separation for the highest ambient pressures, which, based on present structural properties of nozzles, may lead to mechanical failures due to large side loads generated by asymmetric flow separation in the nozzle. In addition, the exhaust plumes of the vehicle nozzles (main engines or thrusters for control) act as disturbance of the external flow creating an effect that can change the pressure distribution on the vehicle surfaces surrounding the exhaust plume. In the interaction zone a turbulent mixing layer, a re-circulation region and a shock system, plume shock/barrel shock, and reattachment regions with considerable heat loading may be formed. These perturbations, mostly of unsteady character, lead to interaction forces that must be accurately predicted in order to obtain the desired vehicle performance. Simulations of such phenomena in the wind tunnel are not simple. It is difficult, when not impossible, to match all variables at the same time, i.e. flight-path parameters and jet-flow parameters. Here CFD can provide an important contribution in the elaboration of extrapolation models for flight conditions. However, it requires the use of unsteady formulations for both turbulence and chemistry models.

Physical models for propulsion applications include turbulence, chemistry, and boundary-layer transition. Among these, turbulence is the critical item and drives the fidelity of the calculations. Current turbulence models used in propulsion are of the one- and two equation types. The primary requirement from CFD is the prediction of injector performance, combustor wall temperatures and heat loads, overall combustion chamber performance and description of the complex, multiphase environment of the combustion chamber at hot-fire conditions. The complexity of the combustion process occurring in a fluid medium stems from the many physical processes of different types, different natures, different temporal and spatial scales, and different degrees of being describable by deterministic models. Second, all of these processes usually are strongly coupled, making it difficult to simplify the problem. In addition, multiphase reacting flow models are affected by the turbulence and chemistry model used; turbulence and combustion interaction; two-phase flow coupling; the spray combustion model; the vaporization rate; atomization; particle size.

Although there are codes available that contain an impressive array of combustion models, they are prohibitively expensive to use for any realistic three-dimensional geometry and flow conditions. It is envisioned that if a factor of 10 increase in computational efficiency is realized as the next generation solvers become available, then multiphase combustion calculations will become more frequent. Also, the prediction of injector/combustor flows requires a robust and highly efficient numerical platform that can be used for both steady state and transient calculations. In addition, a wide variety of Monte Carlo techniques are required - for tracking of finite-size droplets or particles, equilibrium and finite-rate chemistry models, liquid atomization, droplet collision and break-up models, sub- and super-critical droplet vaporization models, turbulent dispersion models, and turbulent chemistry interaction models. From a numerical point of view, there are unique challenges to the launcher flow simulation. A simple test case includes flow domains with strong contrast and largely different scales, like areas of hypersonic flows in the vicinity of low subsonic to almost stagnating flows; chemically reacting flows with time scale markedly smaller than the fluid dynamic time scale; laminar, transitional and turbulent flows; internal and external vortical flows as well as acoustic loads inducing flow instabilities, among other problems. Numerical schemes of order of accuracy greater than two are necessary to handle wave propagation problems properly. In particular, and from a multidisciplinary point of view, during the ascent of a launcher, knowledge is required of the efficiency of the injectors and the acoustic impact on the stability of the combustion process as well as the knowledge of the fluid-structure interaction process at the nozzles (thermo-mechanical loads) and the resulting impact on the flight mechanics behaviour of the vehicle (control). These require a combination of disciplinary methods with contributing modules requiring different treatments when it comes to parallel computing. The CPU requirements for such simulations are not known but are believed to be extensive. Since little experience is available today, assumptions are made based on present computations. It is estimated that an accurate simulation of the first 20 seconds during the ascent phase on today's supercomputer

systems (featuring 3 GHz processors), accounting for unsteady flows, turbulence modelling via at least LES, thermal-structure and flight mechanic coupling will require 100 million grid points as minimum, demanding 210,000 time steps and hence around 3 $\times 10^6$ CPU hours. This may well increase by a factor up to 40 as the fidelity of the data required by the researchers and engineers increases.



Numerical simulation of the nozzle flow base flow interaction by means of detached eddy simulation (source DLR).

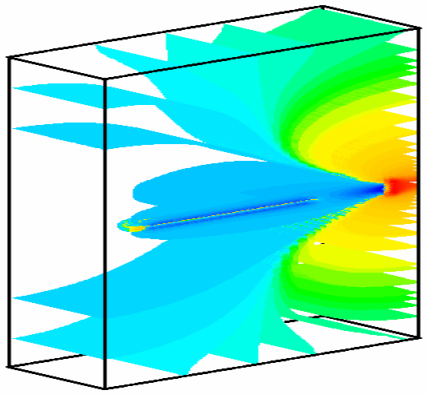
4.5.2.8 Challenge: Improving the Quality of Water Resources

Many areas of the world use groundwater as their main source of freshwater supply. These resources are increasingly being threatened by various sources of pollution. Understanding how this pollution comes about, and designing means of remediation, is critical to the protection of the environment. The same issues arise when one wants to assess the safety of an underground nuclear waste repository, and a closely related problem comes from CO_2 sequestration studies. However, the complexity of the various phenomena involved, as well as the difficulty of directly measuring the properties of the geological medium, make experiments difficult to carry out on a large scale. Numerical simulation then appears as the only way to design and improve new remediation techniques, such as bioremediation, where bacteria that feed on the contaminant are introduced into the medium. This goal requires the ability to run routinely field scale simulations, involving complex coupled phenomena. Its fulfilment will allow progress both in understanding fundamental issues such as more accurate characterization of subsurface properties, as well as more directly applicable results, such as a more reliable assessment of the spreading of contaminant plumes. The public will increasingly expect environmental studies to be carried out with better accuracy: the regulatory threshold for pollutants in drinking water call for accuracy of one part in a million.

An underground aquifer is a heterogeneous porous medium. Quantities of interest are the water velocity and the concentrations of various chemical species as they are carried by the water. Geological heterogeneities in porous media occur over a large range of scales extending from the molecular scale (10^{-10} m) to the field scale (10^3 m). Upscaling (representing quantities on one scale for their use in a large scale) is a fundamental step in almost all hydrogeological studies. The main mechanisms for contaminant transport are convection and mechanical dispersion, which is a macroscopic way of representing the microscopic heterogeneities in the velocity field. These have to be coupled to chemical reactions, or to heat transfer. In extreme cases, chemical reactions can change the physical properties of the medium. Pollution by Dense Non Aqueous Phase Liquids (like chlorinated solvents) leads to multiphase flows. The relevant physical models, as well as the appropriate numerical methods, are still an active area of research. Compositional flows, modelling multicompent, multiphase situations, are still considered too costly, and are replaced by simpler, less detailed models. Hydrogeological models have to deal with the inherent uncertainty in the parameters. They need a way for propagating uncertainty to the computed quantities. Alternatively, one can use inverse methods to estimate these parameters from indirect measurements. A basic question is how to

match the representation of these parameters to the (severely limited) amount of measured data available. The numerical challenges that are yet to be met in a satisfactory way come from sharp fronts at interfaces that need to be finely resolved, strong non-linearities, and complex geometries, calling for unstructured meshes, with heterogeneities varying from metric to kilometric scale. Most hydrogeochemical models lead to coupled sets of partial differential equations, together with algebraic equations. Loosely coupled solution techniques are of limited applicability, so most modern methods discretize the whole system globally. One then needs robust, globally convergent, solution algorithms, and their scalability on petascale computers is still to be assessed.

A finely resolved 3D simulation would need on the order of a thousand grid points in all spatial and temporal directions, and needs to be implicit because of the diffusive nature of the problem. A nonlinear system has to be solved at each time step, with Jacobian matrices that are known to be ill-conditioned due the heterogeneities. It is currently possible to run large 3D models with simplified physics, or more complex physical models with reduced resolution. The overall coupled problem, in realistic 3D situations, is currently out of reach of the fastest computers, mostly because of the large number of chemical systems that have to be solved. Running complex geochemical simulations with thousands of grid points, and a realistic number of chemical species will require orders of magnitude increase in computing power, as well as fundamental advances on several algorithmic fronts: adaptive discretization techniques for advection — diffusion equations, time stepping methods, linear and nonlinear solvers. Mesh generation methods for geological models, and visualization techniques are also part of the required enabling technologies. A typical hydrogeological domain is 30 km x 10 km x 0.5 km. To avoid numerical diffusion, mesh sizes on the order of 1m are needed, which leads to 15 $\times 10^9$ grid cells, with a time step on the order of 1 month.



Flow velocity in the vicinity of a waste disposal area (source: IMFS, CNRS, France)

4.5.2.9 Comments on Simulation Methodology and New Opportunities

The example challenges require advances in physical understanding, numerical methods and software engineering. A selection of the areas where advances are required is given in this section. Multi-scale effects and modelling provide a significant challenge in several areas, including turbulence, combustion, modelling of materials at multiple scales, two-phase flow and heat transfer. On the algorithmic and software side, the emphasis on inter-disciplinary problems places a renewed emphasis on coupling simulation codes for disparate physical models. Finally, a key enabling factor will be the ability of codes to utilise orders of magnitude more processors. The scaling of these codes (for general geometries) will be an important area of research into numerical algorithms and software. Finally, new opportunities to advance the field of uncertainty quantification in simulations will be opened up.

An important practical point for numerical simulations in engineering lies in the fact that several runs are often required to investigate the sensitivity of the results to numerical and physical parameters and to optimise a given device. Accordingly, for a given amount of available computational resources, ten to twenty "small" runs may be preferred to a single "large" run in certain topics.

Turbulence and Transition:

Transition is a significant feature of several challenges, particularly for those from Aerospace and the power industry. These are areas where fundamental studies based on DNS are required to facilitate the development of engineering models. A trend for several of the challenges will be the increasing simulation of turbulence as opposed to modelling. LES for complex flows and geometries will continue to be an active area of research, driving developments in required areas such as high order accuracy on grids for complex geometries, as well as more theoretical considerations such as the validity of LES in the first place. The control of turbulence will be an increasing area of interest. Finally, traditional lines of research such as trends with increasing Reynolds number for building-block geometries will continue.

Combustion:

The scaling up of the Reynolds number to that of realistic flames provides combustion simulations with exciting new opportunities for representing and understanding processes in real engineering systems. As computing power heads towards a Petaflop/s this will become possible.

Multiscale materials:

Fine prediction of highly non-linear mechanical behaviour of metallic materials and concrete require the introduction of microstructural features in modelling. Without technical rupture and the solving of a few-million degrees of freedom problems, no real coupling with Discrete Dislocations Dynamics (DDD) is possible for metallic materials and no coupling with degradation automata is possible for concrete structures. This is even more crucial for the prediction of degradation due to natural ageing, environment or irradiation. An increased computing power is required to address the following features:

- Very complex 3D geometries to represent real microstructures.
- Calibration of complex mechanical formulation by inverse problem resolution
- Weak multi-scale coupling between DDD and Finite Element codes

Two-Phase Flow and heat transfer

The knowledge of “critical heat flux” (namely the Departure from Nucleate Boiling or DNB) is a key challenge for optimisation of heat exchangers where vaporisation of the fluid takes place. Up to now, it has been treated through empirical methods based on experiments. Hundreds of correlations have been established, each of them adapted to a given geometry and given boundary conditions. This phenomenon takes place at very small scales (difficult to address via experimental means) and its prediction will take advantage of Direct Numerical Simulation techniques which become possible with a Petaflop/s machine for simulation of a realistic case. The objective is to avoid empiricism. **Heat transfer** is of primary importance to design cooling systems in gas turbines, piston engines or power plants or to control the charge heating in furnaces or boilers. Convective heat fluxes in thin turbulent boundary layers in the vicinity of walls are not generally well described by usual models. Radiative heat transfer involves long distance interactions and precise models are very demanding in computational resources. Moreover, reacting flows and radiative heat transfers codes may have very different structures. For example, flow balance equations are generally solved on massively parallel machines using physical domain splitting (finite volume framework) while splitting over wavelength range and/or ray direction is more relevant for radiation problems.

Coupling of Different Models/Codes:

For biomedical flows the interaction is used as a transport mechanism. The deformation is much more extreme and the detailed investigation of the nonlinear phenomena is a must for the understanding of blood flow and flows in the respiratory system. Simulating a complete nuclear energy generation cycle requires the coupling of disciplines such as materials, reactor physics, thermal hydraulics, thermal mechanics, fuel management, and geological disposal. Important non-linear effects, a wide range of time scales and influence of existing strong industrial constraints are difficulties systematically arising in all the issues solved in this field : coupling models of vibration and structural

ageing, simulating the degradation mechanisms of materials under irradiation, guaranteeing optimum use of the fuel, coupling neutronics and thermalhydraulics, simulating earthquake resistance of cooling towers, using real geometries to study accidental scenario, predicting the life span of vessels, simulating the geological long term storage of waste material. The improvement of existing models, and qualification of new ones, requires a tremendous increase of computing power.

Scaling and Algorithms:

A general underpinning challenge is likely to be posed by the expansion of future parallel systems beyond ten thousand processors. For fundamental turbulence codes, written to exploit simple geometries and requiring billions of grid points, scaling can be confidently anticipated as a continuation of past trends. For codes which have to deal with complex geometries and order 100 million grid points, scaling becomes more of a question. This may require renewed attention to numerical methods and software engineering. In addition, for challenges like the forest fire and the virtual reactor, the diverse modelling required makes the scaling issue a research question. In addition, attention to extracting more of the theoretical peak performance is an area of significant potential gain for simulation performance.

Uncertainty Quantification:

A general challenge affecting most of the challenges described concerns the treatment of all kinds of uncertainties : intrinsic randomness (such as climate alea, micro-physical noise, ..), lack of knowledge on phenomena involved, variability of spatial/temporal parameters, errors due to unavoidable simplification of models/equations, residuals of the numerical calculation algorithm, scenario ambiguity,….). Demonstrations of the quantification of uncertainties have been carried out and published in the last 20 years in many different industrial, engineering of scientific areas (nuclear accidents and safety, car or airplane mechanics and reliability, environmental pollution and health effects, process metrology, ….). These demonstrate a number of quantification methodologies – deterministic, probabilistic or sometimes possibilistic – be it to modelling of the sources of uncertainties, to propagate them in models, to rank the sensitivity and importance of various sources of uncertainties, or even to identify them through inverse probabilistic algorithms. All these algorithms need to be coupled with the classical deterministic simulation, generating therefore an enormous increase of computing power: a typical uncertainty treatment will multiply by a factor 50 to 100 000 the number of simulations required as compared to a classical (without uncertainties) case; the factor depends widely on the type of question asked (e.g. is it to check low probabilities to exceed a dangerous output under uncertainties, or to rank in a non-linear robust way the sensitivity of a large number of input uncertainties, or is it to get a rough figure of the variability affecting the output) and on the level of optimisation of the uncertainty algorithm (some very efficient algorithms including probabilistic approximation, direct or adjoint differentiation, biased samplings etc.).

5 REFERENCES

[BGS] "Petascale Computational Systems", G. Bell, J. Gray & A. Szalay, IEEE Computer, January 2006

[G-WR] "Petaflop Computing with location Germany in the European Research Area", "Scientific Case" formulated on request of the BMBF by A. Bode, W. Hillebrandt and Th. Lippert, August 2005

[HECRTF] Federal Plan for High End Computing, Report of the High End Computing Revitalization, Task Force, 2005

[ITRS] International Technology Roadmap for the Semiconductor Industry, http://public.itrs.net

[PITAC-CS] "Computational science: ensuring America's competitiveness", report, PITAC, June 2005

[SP-ESR] Strategy Paper on European Supercomputing Resources, joint document BMBF, EPSRC, MENESR, February 2nd, 2005

[SBES] "Simulation Based Engineering Science: revolutionizing Engineering Science through Simulation", Report of the NSF Blue Ribbon panel on Simulation Based Engineering science, February 2006
http://www.nsf.gov/pubs/reports/sbes_final_report.pdf

6 MEMBERSHIP OF INTERNATIONAL SCIENTIFIC PANEL

Area	Country	Title	First name	Last name	Institution	e-mail address	Workshop
Weather, Climatology, Earth Science	DE	Prof.	Geerd-Rüdiger	Hoffmann	Deutscher Wetterdienst	geerd-ruediger.hoffmann@dwd.de	WCE
		Prof.Dr.	Heiner	Igel		igel@geophysik.uni-muenchen.de	WCE
		Prof. Dr	Andreas	Wahner	Forschungszentrum Jülich GmbH - ICG-II: Troposphäre	a.wahner@fz-juelich.de	WCE
	ES	Prof.	Luis	Balairon	INM	balairon@inm.es	WCE
		Prof.	José María	Baldasano	UPC / BSC	jose.baldasano@bsc.es	WCE
	FR		Patrick	Mascart	CNAP	patrick.mascart@aero.obs-mip.fr	WCE
			Pascale	Braconnot	CEA-LSCE	Pascale.Braconnot@cea.fr	
			Fabrice	Dubuffet	CNRS,Lyon	fabien.dubuffet@univ-lyon1.fr	
		Prof	Jean Pierre	Vilotte	Institut de Physique du Globe	vilotte@ipgp.jussieu.fr	WCE
	UK	Prof	Julia	Slingo	University of Reading	swssling@reading.ac.uk	WCE
			Steve	Noyes	Meteorological Office	steve.noyes@metoffice.gov.uk	WCE
			Alan	Dickinson	Meteorological Office	alan.dickinson@metoffice.gov.uk	
		Dr	Vicky	Pope†	Meteorological Office	vicky.pope@metoffice.gov.uk	
		Dr	John	Brodholt	University College London	j.brodholt@ucl.ac.uk	WCE
High energy physics, Plasmas, Astrhophysics	DE	Prof. Dr	Wolfgang	Hillebrandt†	MPI für Astrophysik	wfh@mpa-garching.mpg.de	HEPPA
		Prof. Dr.	Karl	Lackner	Max-Planck-Institut für Plasmaphysik	Karl.Lackner@ipp.mpg.de	HEPPA
		Prof. Dr.	Andreas	Schäfer	Fakultät Physik - Universität Regensburg	andreas.schaefer@physik.uni-regensburg.de	HEPPA
	ES		Jesús	Marco	UC	marco@ifca.unican.es	HEPPA
			Gustavo	Yepes	UAM	gustavo.yepes@uam.es	HEPPA
			José María	Ibañez	UV	jose.m.ibanez@uv.es	HEPPA
			Fernando	Moreno		fmi@iac.es	
			Victor	Tribaldos	CIEMAT	v.tribaldos@ciemat.es	HEPPA
	FR		Edouard	Audit	CEA, SAP	eaudit@cea.fr	HEPPA
			Alain	Becoulet	CEA-DSM	ABECOULET@drfccad.CEA.fr	
			Michel	SPIRO	CNRS	michel.spiro@cnrs-dir.fr	HEPPA
			Xavier	Garbet	CEA-DSM	xavier.garbet@cea.fr	HEPPA
	UK	Prof	Carlos	Frenk	University of Durham	c.s.frenk@durham.ac.uk	HEPPA

		Prof	Richard	Kenway	University of Edinburgh	r.d.kenway@ed.ac.uk	HEPPA
		Dr	Colin	Roach	UKAEA	colin.m.roach@ukaea.org.uk	HEPPA
Life sciences (including bio chemistry)	DE	Prof. Dr.	Ulrich	Hansmann	NIC - Forschungszentrum Jülich	u.hansmann@fz-juelich.de	LIFE
		Dr	Wolfgang	Wenzel	Institut für Nanotechnologie	wenzel@int.fzk.de	LIFE
	ES	Prof.	Modesto	Orozco†	UB / BSC	modesto.orozco@bsc.es	LIFE
		Prof.	Roderic	Guigó	UPF	rguigo@imim.es	LIFE
		Prof.	Alfonso	Valencia		valencia@cnb.uam.es	
		Prof.	José María	Carazo	UAM	carazo@cnb.uam.es	LIFE
	FR		Thomas	Simonson	CNRS	thomas.simonson@polytechnique.fr	LIFE
			Martin	Field	CEA-DSV	martin.field@ibs.fr	
			Olivier	Poch	CNRS	poch@igbmc.u-strasbg.fr	LIFE
	UK	Prof	Mark	Sansom	University of Oxford	mark@biop.ox.ac.uk	LIFE
		Dr	Charles	Laughton	University of Nottingham	charles.laughton@nottingham.ac.uk	LIFE
Chemistry, Materials Science, Nano Science	DE	Prof. Dr.	Heiner	Mueller Krumbhaar	IFF - Forschungszentrum Juelich	h.mueller-krumbhaar@fz-juelich.de	CMSN
		Prof.Dr.	Stefan	Blügel	Forschungszentrum Jülich GmbH - Institut für Festkörperforschung	S.Bluegel@fz-juelich.de	
		Prof. Dr.	Dominik	Marx	Lehrstuhl für Theoretische Chemie -Ruhr-Universität Bochum	dominik.marx@theochem.ruhr-uni-bochum.de	
	ES		Agustí	Lledós	UAB	agusti.lledos@uab.es	CMSN
			Manolo	Yañez	UAM	manuel.yanez@uam.es	CMSN
			Avelino	Corma	CSIC	acorma@itq.upv.es	CMSN
	FI	Prof.	Risto	Nieminen	Helsinki University of Technology	Risto.Nieminen@hut.fi	CMSN
		Prof.	Kai	Nordlund	University of Helsinki	kai.nordlund@helsinki.fi	CMSN
	FR		Gilles	Zerah†	CEA-DAM	gilles.zerah@cea.fr	CMSN
			Yann-Michel	Niquet	CEA-DRFMC	yniquet@cea.fr	CMSN
			Marie Madeleine	Rohmer	CNRS, Labo de chimie quantique de strasbourg	rohmer@quantix.u-strasbg.fr	
			Philippe	Sautet	CNRS, Labo de chimie, ENS Lyon	Philippe.Sautet@ens-lyon.fr	CMSN
	UK	Prof	Jonathan	Tennyson	University College London	j.tennyson@ucl.ac.uk	CMSN
		Prof	Steve	Parker	University of Bath	s.c.parker@bath.ac.uk	CMSN

Engineering	DE	Prof. Dr.	Gerhard	Wellein	U. Erlangen	Gerhard.Wellein@rrze.uni-erlangen.de	Eng
		Prof. Dr.	Christoph	Zenger	Technical University of Munich-Institut für Informatik	zenger@in.tum.de	Eng
		Prof. Dr.		Bungartz	LMU		
		Prof. Dr.-Ing.	Siegfried	Wagner	Universität Stuttgart	wagner@iag.uni-stuttgart.de	
		Prof. Dr.	Norbert	Kroll		norbert.kroll@dlr.de	
		Prof. Dr.	Robert	Sausen		robert.sausen@dlr.de	
		Prof. Dr.-Ing	Wolfgang	Schröder	RWTH Aachen -Lehrstuhl für Strömungslehre und Aerodynamisches Institut	office@aia.rwth-aachen.de	Eng
	ES	Prof.	Javier	Jimenez Sendín	UPM	jimenez@torroja.dmt.upm.es	Eng
			Cesar	Dopazo	UZ	dopazo@unizar.es	Eng
			Christophe	Calvin	CEA-DEN	calvin@soleil.intra.cea.fr	
	FR		Christian	Chauliac	CEA-DEN	christian.chauliac@cea.fr	Eng
			Jean	Gonnord	CEA-DAM	Jean.GONNORD@ANTIGONE.CEA.FR	Eng
			Denis	Veynante	CNRS,UPR288	denis.veynante@em2c.ecp.fr	Eng
	UK	Prof	Neil D	Sandham	School of Engineering Sciences (Aero) - University of Southampton	n.sandham@soton.ac.uk	Eng
		Dr	Ken	Badcock†	University of Liverpool	gnaa36@aero.gla.ac.uk	Eng
	FI	Dr	Roel	Verstappen		verstappen@math.rug.nl	Eng

† Moderator

	Country	Title	First name	Last name	Institution	Position	e-mail address	Country
Country Delegation	DE	Prof. Dr.Dr.	Thomas	Lippert‡	Forschungszentrum Jülich GmbH -Central Institute for Applied Mathematics		Th.Lippert@fz-juelich.de	DE
			Roland	Mader			Roland.Mader@dlr.de	
	ES	Prof.	Mateo	Valero	Barcelona Supercomputing Center	Director	mateo.valero@bsc.es	
		Mr.	Francesc	Subirada	Barcelona Supercomputing Center	Associated Director	francesc.subirada@bsc.es	

		Mr.	Felipe A.	Lozano	Barcelona Supercomputing Center	Business Development Manager	felipe.a.lozano@bsc.es	
	FI	Dr.	Kimmo	Koski	CSC	Managing Director	kimmo.koski@csc.fi	
		Dr.	Leif	Laaksonen	CSC	Development Director	Leif.Laaksonen@csc.fi	
	FR		Jean-Yves	Berthou	EDF		jy.berthou@edf.fr	
			Eric	Chaput	EADS		eric.chaput@airbus.com	
			Guillaume	Alléon	EADS		guillaume.alleon@eads.net ..	
			Daniel	Vervaerde	CEA		daniel.verwaerde@cea.fr	
			Sylvie	Joussaume	CNRS		Sylvie.Joussaume@cnrs-dir.fr	
			Alain	Lichnewsky	MDESR		alain.lichnewsky@recherche.gouv.fr	
			Jean Claude	André	CERFACS		Jean-Claude.Andre@cerfacs.fr	
			Alain	Kavenoky	MDESR		alain.kavenoky@recherche.gouv.fr	
			Michel	Kern	INRIA		michel.kern@inria.fr	
			Jean Pierre	Génin	CEA		genin@dsmdir.cea.fr	
	IT		Piero	Vicini	INFI (Instituto Nazionale di Física Nucleare)		Piero.Vicini@roma1.infn.it	
	NE			to be defined				
	UK	Mr.	Hugh	Pilcher-Clayton	EPSRC		hugh.pilcher-clayton@epsrc.ac.uk	
		Ms	Linda	Sayers	EPSRC		linda.sayers@epsrc.ac.uk	
		Prof.	Volker	Heine	Cavendish Laboratory, University of Cambridge		vh200@cam.ac.uk	
		Dr	Martyn	Guest	CLRC, Daresbury Laboratory		m.f.guest@dl.ac.uk	
European Comission	PT		Maria	Ramalho	European Comission		vh200@cam.ac.uk	EU
	IT		Lorenza	Saracco	European Comission		Lorenza.SARACCO@cec.eu.int	EU
	GR		Georgios	Amanatidis			georgios.amanatidis@cec.eu.int	EU
ESFRI	ES		Carlos	Alejaldre	MEC		dgpt@min.es	EU
	FI		Heikki	Mannila			Heikki.Mannila@cs.helsinki.fi	EU
e-IRG	NE		Patrick	Aerts	Netherlands National Computing Facilities Foundation		aerts@nwo.nl	EU
Other	GE		Walter	Zwieflhofer	ECMWF		Walter.Zwieflhofer@ecmwf.int	

‡Head of German Scientific Delegation