Computational Challenges and Needs for Academic and Industrial Applications Communities

Breakout Session 2

IESP, Saclay, June 28, 2009

Group #2 Mission

- Agree on a plan
- Assess application fields
- Determine (software) requirements and needs of different application fields
  - Need help of experts
- Set up roadmap of requirements
  - Need help of experts

Intended procedure

- Build upon pre-existing studies, e.g.,
  - PRACE surveys
  - Extreme Scale workshops in the U.S. DOE
  - Hartree Center Workshops at Daresbury
  - HET scientific cases
- Refine group notes with off-line research by panelists
  - Consult as needed with experts, by mid-September
  - Complete before 18-20 October workshop in Japan

Question battery for experts

- Process scalability for today’s state-of-art runs
- Asymptotic complexity of today’s state-of-art algorithms
- Quantitative characterization of architectural resource requirements (flops, storage, I/O, communication, synchronization, etc.)
- Known bottlenecks (e.g., solver, load balancer, adaptation method, etc.)
- Software dependencies

Scope for Group 2 discussions

- Assume evolutionary path from present to expected 2018 arrival of exaflop/s
- Assess current state of the art of (selected representative) legacy applications
- Identify in advance steps required to use exascale machines, to promote concurrent software development of important applications with hardware development
- Assess points of need from experts (gaps analysis)
- Assess potential limits on scalability
- Assess distinct architectural domains of applications (e.g., flop/s per byte of storage, flop/s per bytes/sec of transfer, etc.)

Applications considered (not comprehensive, not ordered)

- Climate change
- Meterology
- Materials science
- Biology
- Plasma physics/fusion
- Geophysics
- Fluid dynamics
- Structural mechanics
- Electromagnetics
- Aerodynamics
- Combustion
- Lattice quantum chromodynamics
- Biophysics
- Astronomy/cosmology
- Molecular dynamics
- Video processing
- Chemistry
- Nuclear engineering/fission
- Epidemiology
- Nanotechnology/microelectronics
- Emergent sciences (e.g., social, networks, etc.)

CS techniques represented

- Parallel code optimizations
- Parallel programming abstractions and support
- Adaptive optimizations (load balancing, etc.)
- Post-processing of large data sets
- Coupling multiple models (control and data aspects)

Math techniques represented

- Partial differential equations
- Solution adaptive methods
- Molecular dynamics modeling and algorithmics
- Particle methods (distinct from classical MD in terms of inhomogeneous distributions, e.g., cosmology)
- DFT/electronic structure
- (hybrid) Monte Carlo
- Stochastic optimization
- Scalable implicit solvers (linear and nonlinear)
- Eigenanalysis techniques
- FFTs and other fast transforms
- Smoothed particle hydrodynamics
- Agent-based methods
- Coupling multiple models (stability aspects)
- Data assimilation/fusion of observations and models
- Data mining
- Uncertainty quantification

ACADEMIC AND INDUSTRIAL APPLICATIONS COMMUNITIES

Weather, Climate 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

Astro physics	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.
Element ary 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 centres worldwide. 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.

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.
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 Dyn.	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	Correct representation of protein machines. 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	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).

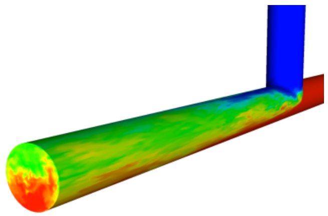
Engineering continued

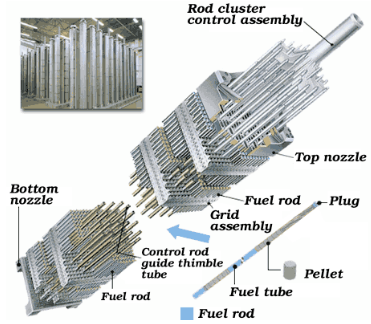
Forest Fires	The development of reliable numerical tools able to model and predict fire evolution is critically important in terms of safety and protection ("numerical fire simulator"), fire fighting and could help in real time disaster management. The social impact is very important and is concerned with land, buildings, human and animal life, agriculture, tourism and the economy.
Green Aircraft	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 optimization 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.

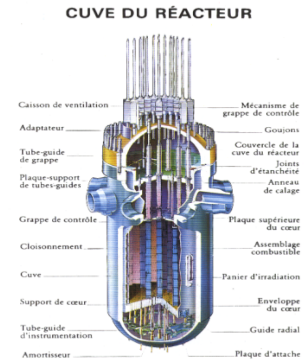
COMPUTATIONAL CHALLENGES AND ROADMAPS

Code_Saturne – HPC Roadmap: *application examples*

2003	2006	2007	2010	2015
Consecutive to the Civaux thermal fatigue event Computations enable to better understand the wall thermal loading in an injection. Knowing the root causes of the event ⇒ define a new design to avoid this problem.	Computation with an L.E.S. approach for turbulent modelling Refined mesh near the wall.	Part of a fuel assembly 3 grid assemblies	9 fuel assemblies No experimental approach up to now Will enable the study of side effects implied by the flow around neighbour fuel assemblies. Better understanding of vibration phenomena and wear-out of the rods.	The whole vessel reactor







Computations with smaller and smaller scales in larger and larger geometries
⇒ a better understanding of physical phenomena ⇒ a more effective help for decision making
⇒ A better optimisation of the production (margin benefits)

10^6 cells **3.10^{13} operations**	10^7 cells **6.10^{14} operations**	10^8 cells **10^{16} operations**	10^9 cells **3.10^{17} operations**	10^{10} cells **5.10^{18} operations**
Fujistu VPP 5000 1 of 4 vector processors 2 month length computation	Cluster, IBM Power5 400 processors 9 days	IBM Blue Gene/L « Frontier » 8000 processors # 1 month	**30 times** the power of IBM Blue Gene/L « Frontier » # 1 month	**500 times** the power of IBM Blue Gene/L « Frontier » # 1 month
# 1 Gb of storage 2 Gb of memory	# 15 Gb of storage 25 Gb of memory	# 200 Gb of storage 250 Gb of memory	# 1 Tb of storage 2,5 Tb of memory	# 10 Tb of storage 25 Tb of memory
Power of the computer	Pre-processing not parallelized	Pre-processing not parallelized Mesh generation	… ibid. … … ibid. … Scalability / Solver	… ibid. … … ibid. … … ibid. … Visualisation

Weather, Climate and Earth Sciences: Computational Challenges

Weather, Climate and Earth Sciences: Roadmap

Weather, Climate and Earth Sciences: Experts

- André, Giovanni (climate change)
- ECMWF, NCAR (meteorology, oceanography)
- Total, Calendra, Japan/US TBD (solid earth sciences)

Astrophysics, HEP and Plasma Physics: Computational Challenges

Astrophysics, HEP and Plasma Physics: Roadmap

Astrophysics, HEP and Plasma Physics: Experts

- W. Hillebrandt (MPI), E. Audit (CEA-Saclay) – Astrophysics
- Kalé, Quinn – Cosmology
- T. Lippert, R. Kenway, A. Ukawa, S. Gottlieb – LQCD
- S. Günter, T. Sato, E. Sonnendruecker – Plasma physics

Materials Science, Chemistry and Nanoscience:
Computational Challenges

Materials Science, Chemistry and Nanoscience:

Roadmap

Materials Science, Chemistry and Nanoscience: Experts

- R. Parinello, G. Martyna, Binder, G. Zerah, K. Terakura– materials science/complex materials
- T. Dunning, J. Pople, K. Hirao, D. Marx – chemistry
- T. Schultheis, T. Deutsch, S. Bluegel, Cuniberti – nanomaterials

Life Sciences:
Computational Challenges

Life Sciences:
Roadmap

Life Sciences:
Experts

- K. Schulten, A. Grubmueller, T. Simonson, R. Lavery, – molecular biology
- ?? – systems biology
- Masella

Engineering:
Computational Challenges

Engineering:
Roadmap

Engineering:
Experts

- Kaneda, C. Rossow, E. Chaput – aeronautics
- Schroeder, P. Moin – turbomachinery
- Fournier – vessels
- M. Heath – rockets

SUMMARY