1. European Computational Science Forum: The "Lincei Initiative": from computers to scientific excellence

2. Principal Applicant

Berend Smit
CECAM/Ecole Normale Supérieure de Lyon
46 allée d'Italie, 69364 Lyon Cedex 07, France
Tel: +33 472728637 FAX: +33 47278636
E mail : B.Smit@CECAM.fr www.CECAM.org

3. Keywords:

Ab-initio calculations; molecular simulations; high-performance computing; cyber-infrastructure

4. Abstract

The aim of this Forward Look is to develop a vision on how Computational Sciences will evolve in the coming 10 to 20 years. Based on a scenario of how this field will evolve and on the needs of the scientific community, a strategy will be presented aiming at structuring software and hardware support and development at the European level. The initiative of this Forward Look is taken by the atomic-scale materials simulation community, but the outcome will be of interest to other Computational Science communities as well. Implementation of the recommendations of this Forward Look should lead to an advanced cyber-infrastructure that allows Europe to maintain its leading position in this field.

5. The case for a Forward Look

Rationale:

Computational Sciences and computer simulations in particular, are playing an ever growing role in fundamental and applied sciences. It is clear, that this success is based to a large extent on the development of computers as computing and storage tools. In additions, the internet has also changed the way we can organize science. In addition to a European (super-)computer infrastructure, the scientific software developed by the various European groups is becoming increasingly important from a strategic point of view. Whereas ten years ago every research group could develop its own research codes, nowadays the complexity and sophistication of these methods in use have grown to such an extent that many groups rely on the availability of this scientific software. For example, in the field of (ab-initio and molecular) Materials Science, the four most important (European) codes have been used as "research instrument" in over 2000 publications per year in international journals.

To ensure not only that Europe will maintain a leading position in the field, but also that as many European research groups as possible take advantage of this unique position, it is essential to develop a strategy to ensure an efficient use of resources at the European level. Many HPC (high performance computing) centers give support for local software

development, but since most of the software is used across Europe, this support could be even more efficient if it were coordinated at the European level.

The concept of a cyber-infrastructure has been proposed (originally in the US) that includes all the necessary developments which have to take place, in order to provide the means to take full advantage of the possibilities offered by the rapid growth of hardware and of the grid. Code developers across Europe should be able to work in a collaborative and integrated way: sharing data, building specific codes in addition to existing ones, using remote facilities and graphic tools. Such an infrastructure will be distributed and virtual, and its realization needs a specific action being taken at a European level: this action will result from an analysis of the actual needs of the scientists involved and the vision they have on how these can be met. This vision will be the main outcome of the present Forward Look, the recommendations of which will be addressed to the European Research Councils, ESF and the EC in order to be implemented.

More importantly, the availability of such a cyber-infrastructure should create new opportunities for science. For example, combining the expertise of the psi-k and MolSimu communities allows us, for example, to link electronic structure calculations with techniques to compute a macroscopic phase diagram of a material. A well designed European cyber-infrastructure can deal with the complexity of the various codes and the fact that many different European groups have to contribute.

State of the Art:
Computational approaches are becoming an increasingly important tool in modern materials research (including biomaterials). The complexity of the modern codes has caused a transition. A few years ago, each computational group had its own “home-brewed” software. At present, an increasing number of groups rely on the availability of such codes. In the field of computational material science, Europe is playing a leading role. Importantly, this field is to a large extent supported by two important research networks, the ESF programmes Psi-k and SIMU, which represent over 20 European countries and involve over 1000 research groups. To this one can also add the research groups working in the closely related field of modeling of quantum molecular processes. The combined activities of these groups will become even more visible in connection with the realization of the planned European Super Computer Center, which would multiply the present super-computer power by a sizable factor in the next 5 years. But one should not forget that important research is also carried with the same software, running on a simple desktop PC, and these researchers should also benefit from this cyber-infrastructure.

Another development is that the range of users is broadening: they range from theoreticians developing improved algorithms and new types of applications, to experimentalists who need to use simulation to interpret complex data. Moreover, it is not just a matter of distributing software: essential elements of the success of a cyber-infrastructure are appropriate training and user support geared at the high level of expertise needed to use the software reliably.

One can argue that this situation is very similar to neutron or X-ray facilities. Also in this domain Europe has gained competitive advantages by creating an experimental research infrastructure at the European level. Such an infrastructure has to be created for software, and what such an infrastructure should look like to match the needs of the scientific community will be one of the central themes in this Forward Look.

Context

As High Performance Computing requires large-scale investments in supercomputers, many national organizations in the EU justify these investments with the increasing importance of computational science.[1] Support to software development has remained national and limited (e.g. the CCP projects in the UK), although giving rise to interesting results.[2] Given the larger scale, the US has taken the initiative to introduce the concept of cyber-infrastructure as the platform to further enhance the impact of computational approaches in a large number of fields[3]. In addition, it is important to mention the ESFRI discussion, which is currently taking place, and the possible creation of a European supercomputer center, making a clear vision, by the scientists involved, on how software issues should be structured on a European level, essential and urgent.

Clearly, novel technologies, such as, for example, the grid, the connection of data-bases, the connection of remote graphic tools, or the connection to remote experiments, can have a large impact on the way in which science is done. At this point we would like to emphasize that although materials science has taken this initiative, the problems go much beyond the boundaries of this field, and other communities will be invited to explain their views: Atomic and molecular physics and chemistry, astrophysics, plasma physics, high-energy physics, and various engineering research areas, all are facing similar problems, although they have been acting separately until now.

The need for more structure in the area of computational sciences has already been addressed before. For example, one of the conclusions of the EUPRO 2003 annual meeting in Lyon was that some thoughts should be spent on what is going to be the future of computational physics, in particular the organization of the funding at the European level concerning all the science ahead and the software needed to be developed and distributed. The EUPRO meeting had also insisted on the existing plan of the NSF (blue ribbon panel report on cyber-infrastructures). As a follow-up a workshop supported by ESF-PESC was organized in the Accademia dei Lincei, in Rome (12-13 July 2005) and the present proposal is based on the conclusions of this workshop.

Questions to be addressed by the Forward Look

---

[1] http://www.scidac.org/

[2] See for example, High End Computing Terascale Resources, HECToR. Scientific Case, EPSRC, UK (2005) http://www.epsrc.ac.uk/ResearchFunding/FacilitiesAndServices/HighPerformanceComputing/HECToR/default.htm

[3] Revolutionizing Science and Engineering Through Cyberinfrastructure:Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure (January 2003) www.communitytechnology.org/nsf_ci_report

What are the key scientific questions that will be addressed in the coming 5-10 years?
The start of the discussion on a desirable cyber-infrastructure should be a plausible scenario of how computational material science will evolve in the coming years.

Are the needs for atomic scale materials simulations different from other fields of computational science?
To have an adequate focus for the Forward Look, the idea is to use atomic scale materials science as a reference point. However, many issues that will be addressed in the Forward Look will be of interest to other fields of computational science, such as astrophysics, quantum chemistry, computational fluid dynamics, bioinformatics, and particle physics. It is therefore important not only to obtain input from these fields, but also to have active participation.

What is the expertise needed to set up this infrastructure?
Setting up such an infrastructure is a multidisciplinary activity which could benefit from state of the art developments in hardware (grid, middelware, etc) and computer science (integrated code development, good practice, multiscale-multi code issues). What are the emerging technologies and how can they be used?

How should the European Computer Infrastructure, both in hardware and software, be organized such that Europe has a competitive advantage?
A software infrastructure will most likely not be a central building, but a distributed (virtual) cyber infrastructure. In addition, the availability of such a cyber-infrastructure may create opportunities for new science, for example, by coupling computational methods to data-bases. How to set up such a cyber infrastructure? How to ensure that the services can be used by the European scientific community at large (importance of training and tutoring)? How to build on existing groupings, such as the Psi-k and Molsimu and other communities?

Objectives:
The outcome of the Forward Look will be a vision on how the field of computational science will evolve and how this field should be organized such that Europe will be taking advance fully of its unique position in the coming 10 to 20 years.

The Forward Look instrument will be used for the following activities:

- Planning meetings and interviews to prepare the workshops
- 4 Workshops, in which to address the points raised in "Questions to be Addressed" (see previous section):
  1. *"Challenges in Computational Material Science"*
  2. *"The needs of the various computational communities"*
  3. *"How can advances in computer science and technology be used in a cyber-infrastructure?*
  4. *"Towards a Cyber-infrastructure"*
- The Forward-Look Conference. The large Forward-Look Conference will be asked to discuss the conclusions and recommendations of the provisional Forward-Look

report, which will be made available prior to the conference. The format will be a mix of plenary meetings and small workshops.

- The final printed Forward-Look Report in which the conclusions and recommendations of the provisional report are modified according to the outcome of the Forward Look Conference. This document will present the analysis of the needs in the various fields and make recommendations to the actors

Target groups;
The primary target groups are funding agencies for research and research infrastructure at the national and European level (EC, RI unit of ESF, ESFRI) and the European HPCN centers and computer centers at universities.

Budget proposal (in Euros)
The proposed budget (180 k€) will be used for the following:

a. Scientific Officer: to support the Organization Committee
b. 6 Planning/OC meetings
c. Interviews
d. 4 Workshops (20 participants each)
e. 1 large Conference
f. Documentation

Additional support (ca. 40k€) will be given by CECAM, Psi-k, MolSimu, Juelich (Germany), CSCS (Switzerland) Accademie de Lincei (Italy), and Daresbury. (UK). In addition to this financial support, CECAM , Juelich, Daresbury, CSCS, and Accademia dei Lincei have already agreed to take care of the local organization of the workshops or conferences. The requested ESF support is 140k€.

Tentative Organising Committee:
The Forward Look is coordinated by CECAM and the organizing committee has representatives from the various communities (Psi-k and Simu), the HPCN centers, and those involves in the development and support of computer codes such as the CCPs.

CECAM[4] is presently supported by 14 organizations from 8 different (European) countries. The mission of CECAM is to facilitate dissemination, education, and networking activities for the scientific community interested in particle-based simulations at the European level. CECAM houses every year 20-25 workshops attracting 600-700 international scientists. The workshops organized in the context of this Forward-Look will therefore obtain a unique exposure to the international scientific community.

Proposed members are

1. Prof. Berend Smit (CECAM, Lyon France) chair
2. Representative from the Psi-k community: Professor Volker Heine (Cambridge, UK)
3. Representative from the Simu-community; Professor Giovanni Ciccotti (Rome, Italy)

[4] www.CECAM.org

4. Representative from the HPCN centers: dr. Marie-Christine Sawley (Manno, Switzerland)
5. Representative from the code development centers: Prof. Paul Durham (Daresbury, UK)
6. Representative from the code development centers: Prof. Franco A. Gianturco (Rome, Italy)

Tentative Schedule of activities and their format

The entire forward look is expected to take approximately 12 months. The following planning is envisioned

- Month 0: kickoff planning meeting
- Month 1-6: Forward Look preparation phase (4 workshops using the following structure)
    - Planning meeting
    - Interviews
    - Discussion document
    - Workshop, Workshop report
- Month 7-11: Provisional Forward Look Report
- Month 12: Forward Look Conferences
- Month 13-16 : Publication of Forward Looks

6. Appendix: a short CV of up to 3 pages of the Principal Applicant should be included.

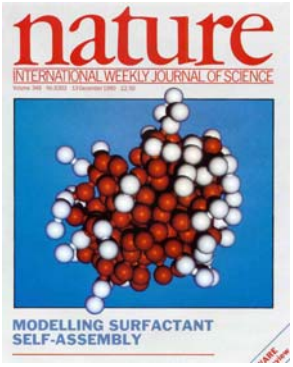
CECAM

CECAM is a top-level institute based in Lyon that is supported by 14 European research organizations. In addition to fostering ground-breaking research in the area of molecular simulation, CECAM aims at facilitating dissemination and networking activities. For example, each year, 20-25 workshops and tutorials are being organized at CECAM by world experts. Many of the ideas of algorithms that are now commonly used in molecular simulation software have originated from workshops that CECAM has organized over the last 30 years.

Smit, Berend born November 8th, 1962

1990	PhD (cum laude) Chemistry, Utrecht University (NL)
1988-1997	Shell Research Amsterdam (NL)
1997-…	Professor, University of Amsterdam
1998	Gold medal of the Royal Dutch Chemical Society
2004-…	Elected Director of CECAM

Berend Smit co-authored with Daan Frenkel the book *Understanding Molecular Simulations* (Academic Press, 2nd edit., 2002). Together with the first edition this book has been cited over 1500 times and is used as the text book on molecular simulations in many universities. A quotation of a review "… this book brilliantly lays down the scientific foundations of the simulational approach…" (Physics World, 1997). In 2003 a Chinese translation appeared. On the basis of this book the authors have organized the Amsterdam Molecular Simulation School which has attracted over last 5 years 300 (young) scientists from all over the world.



Research Activities: The research expertise of the Smit group is the development of novel computational methods to simulate complex fluids, such as configurational-bias Monte Carlo. In the area of membrane simulations the group of Smit has developed mesoscopic models to simulate phase transitions in membrane and to investigate the effect of alcohol or peptides on the properties of a membrane. The work of Smit has been published in over 160 publications of which 6 in Nature and Science. Prof Smit has been awarded the gold medal of the Royal Dutch Chemical Society, which is the most prestigious award given annually to a chemist in the Netherlands. His simulation of the self-assembly of micelles: featured on the cover of Nature.

Publications: over 170 with a total of over **5500 citations**

Ten key publications:

1. Siepmann, J. I.; Karaborni, S.; Smit, B., *Simulating the critical properties of complex fluids*. Nature, **365**, (1993), 330.
2. Karaborni, S.; Esselink, K.; Hilbers, P. A. J.; Smit, B.; Karthauser, J.; Os, N. M. v.; Zana, R., *Simulating the self-assembly of gemini surfactants*. Science, **266**, (1994), 254
3. Smit, B.; Maesen, T. L. M., *Commensurate ``freezing'' of alkanes in the channels of a zeolite*. Nature, **374**, (1995), 42.
4. Karaborni, S.; Smit, B.; Heidug, W.; Urai, J.; Oort, E. v., *The swelling of clays: molecular simulations of the hydration of montmorillonite.* Science, **271**, (1996), 1102
5. Vlugt, T. J. H.; Krishna, R.; Smit, B., *Molecular Simulations of Adsorption Isotherms for Linear and Branched Alkanes and their Mixtures in Silicalite.* J. Phys. Chem. B, **103**, (1999), 1102
6. Beerdsen, E.; Smit, B.; Dubbeldam, D., *Molecular Simulation of Loading Dependent Slow Diffusion in Confined Systems*. Phys. Rev. Lett., **93**, (2004), art. no 248301.
7. Rekvig, L.; Hafskjold, B.; Smit, B., *Chain Length Dependencies of the Bending Modulus of Surfactant Monolayers*. Phys. Rev. Lett.. **92**, (2004), art. no:116101.
8. Dubbeldam, D.; Calero, S.; Vlugt, T. J. H.; Krishna, R.; Maesen, T. L. M.; Beerdsen, E.; Smit, B., *Force field parametrization through fitting on inflection points in isotherms.* Phys. Rev. Lett., **93**, (2004), art no. 088302.
9. Venturoli, M.; Smit, B.; Sperotto, M. M., *Simulations of Protein-Induced Bilayer Deformations, and Lipid-Induced Protein Tilting*, Biophys. J., **88**, (2005), 1778.
10. Dubbeldam, D.; Calero, S.; Beerdsen, E.; Smit, B., *Molecular path control in zeolite membranes*. Proc. Natl. Acad. Sci. USA, **102**, (2005), 12317–12320.