D3.5 Working Group Report on Fundamental Sciences (Chemistry, Physics)

CONTRACT NO	EESI 261513
INSTRUMENT	CSA (Support and Collaborative Action)
THEMATIC	INFRASTRUCTURE

Start date of project: 1 JUNE 2010 Duration: 18 months

Name of lead contractor for this deliverable: JUELICH

Authors: Godehard Sutmann (JSC / CECAM) and Jean-Philippe Nominé (CEA)

Contributors to the report: M.Ashworth (STFC Daresbury), A.Curioni (IBM), T.Deutsch (CEA), A.Khokhlov (Moscow State Univ.), L.Lellouch (CNRS, CERN), N.Marzari (Univ. Oxford), U.Meißner (Univ. Bonn / FZ Jülich) M.Ottaviani (CEA), M.Payne (Univ. Cambridge), Martin Plummer (STFC Daresbury), T.Schulthess (CSCS Manno), L.Silva (Univ. Lisbon), V.Springel (HITS Heidelberg), R.Teyssier (ETH Zürich), G.Zerah (CEA / CECAM).

Name of reviewers for this deliverable: Stephane Requena (GENCI)

Abstract: This deliverable provides an overview of computational demands and challenges in the field of fundamental sciences (physics and chemistry) and positions the role and competitiveness of European research and software development for scientific applications.

Release N° 1

Due date of deliverable:	Mai 31, 2011
Submission date:	July 01, 2011

Internal release N° 2

Due date of deliverable:	November 30, 2011
Submission date:	December 13 2011
Publication date:	December 22 2011

Project co-funded by the European Commission within the Seventh Framework Programme (FP7/2007-2013)		
Dissemination Level		
PU	Public	X

Table of Contents

Glossary

Abbreviation / acronym	Description
AGN	Active Galactic Nuclei
ARL	US Army Research Laboratory
ASC	Advanced Simulation and Computing
ASH	Anelastic Spherical Harmonic
BLAS	Basic Linear Algebra Subprograms
BSM	Beyond the Standard Model
CCP	Collaborative Computational Projects
CEA	Commissariat à l'Energie Atomique
CECAM	Centre Européen de Calcul Atomique et Moléculaire
CERN	European Organization for Nuclear Research
CLS	Coordinated Lattice Simulations
CNRS	Centre National de la Recherche Scientifique
CPU	Central Processing Unit
CRA	Collaborative Research Alliance
CSCS	Swiss National Supercomputing Centre
CSE	Computational Science and Engineering
CSED	STFC Computational Science and Engineering Department
CTMC	Chemical Theory, Models and Computational Methods
DCSE	distributed Computational Science and Engineering
DFT	Density Functional Theory
DL	Daresbury Laboratories
DPD	Dissipative Particle Dynamics
EFT	Effective Field Theory
ELI	Extreme Light Infrastructure
EMM	Enterprise for Multiscale Materials
EPR	Electron Paramagnetic Rseonance
FAIR	Facility for Antiproton and Ion Research
FD	Finite Differences
FE	Finite Elements
FTE	full time equivalent
FZJ	Forschungszentrum Jülich
GENCI	Grand Equipement National de Calcul Intensif
GPGPU	General Purpose Graphics Processor Unit
GRAPE	Gravity Pipeline Engine
HEP	High Energy Physics
HiPER	High Power Laser for Energy Research

Abbreviation / acronym	Description
HITS	Heidelberg Institute for Theoretical Studies
HP2C	Swiss Platform for High Performance and High Productivity Computing
HPC	High Performance Computing
IESP	International Exascale Software Project
IFERC	International Fusion Energy Research Centre
IP	Intellectual Property
ITER	International Thermonuclear Experimental Reactor
JET	Joint European Torus
JWST	The James Webb Space Telescope
KIT	Karlsruhe Institute for Technology
LB	Lattice Boltzmann
LFT	Lattice Field Theory
LHC	Large Hadron Collider
LOFAR	The LOw Frequency ARray (a multi-purpose sensor array)
LQCD	Lattice Quantum Chromo Dynamics
MC	Monte Carlo
MD	Molecular Dynamics
MDGRAPE	Molecular Dynamics Gravity Pipeline Engine
MEMEPhys	Multiscale Mechanistic Model of the PEMFC electrochemical processes
MHD	Magneto-Hydrodynamics
MMM	Multiscale Materials Modelling
MSU	Moscow State University
NMR	Nuclear Magnetic Resonance
NPEFT	Nuclear Physics from Effective Field Theory
NSF	National Science Foundation
ONIOM	Our own N-layered Integrated molecular Orbital and molecular Mechanics
OU	Oxford University
PEMFC	Polyelectrolyte Membrane Fuel Cell
PIC	Particle in Cell
PRACE	Partnership for Advanced Computing in Europe
PRISM	Polymer Reference Interaction Site Model
PRMAT	Parallel R-matrix Program
QCD	Quantum Chromo Dynamics
QCDSF	QCD Structure Function
QED	Quantum Electrodynamics
QM/MM	Quantum Mechanical / Molecular Mechanics
QUB	Queen's University Belfast

Abbreviation / acronym	Description
ReaxMD	Reactive Molecular Dynamics
RHIC	Relativistic Heavy Ion Collider
RIKEN	Rikagaku Kenkyusho (natural science institute in Japan)
RISM	Reference Interaction Site Model
SimLab	Simulation Laboratory
SKA	Square Kilometre Array
SLA	Service Level Agreement
SM	Standard Model
STFC	Science and Technology Facilities Council
TDGL	Time Dependent Ginzburg-Landau
TDRM	Time Domain Reflectometry Microcomputer
TDSE	Time Dependent Schrödinger Equation
UCL	University College London
UKQCD	The United Kingdom QCD
UK-RAMP	UK R-matrix Atomic and Molecular Physics HPC Code Development Project
UKRmol	UK molecular R-matrix suites
UVSQ	Université de Versailles Saint-Quentin-en-Yvelines

1. Executive summary

This report provides an overview of current trends, future perspectives and needs in compute resources and software development in the field of fundamental sciences, i.e., astrophysics, cosmology, nuclear-, hadron- and high energy physics, plasma physics, fusion research, quantum chemistry, atomic physics, soft matter research and materials sciences. The report is the result of two expert meetings, held in January 2011 in Paris and April 2011 in Brussels, as well as a questionnaire addressed to the group of experts involved into the working group.

As Robert-Jan Smits, Director General for Research and Innovation of the European Commission, has noted, Europe today suffers from an innovation gap in relation to the United States and Japan, with newer rivals such as China "quickly catching up" to the leaders. New insights provided by fundamental science are the starting points for innovation, and high performance computing (HPC) has firmly established itself as a powerful accelerator of scientific innovation.

Through PRACE and other initiatives, Europe has made substantial investments in petascale high performance computing hardware systems and has plans in place for further investments in hardware. Developments in parallel software are needed to produce strong returns on these investments by making the hardware systems highly productive for Europe's scientific and research communities.

A recent study noted that "in the present era of petascale computing and in the exascale era that will begin before 2020, sustained performance leadership on real-world HPC applications and workloads will be determined far more by software advances than by hardware progress...If Europe quickly begins to provide appropriate, sustained investment and support for parallel software development, Europe can become the global innovation leader in scientifically and economically important domains in which Europe already has substantial strengths."[1]

A summary of the survey and expert meetings is that Europe has a very strong position in the global scientific community and is leading in several fields of astrophysics/cosmology, nuclear/hadron physics, fusion research and materials sciences. European software development of community codes is very advanced in materials science, quantum chemistry, astrophysics, nuclear/hadron physics and plasma/fusion physics. The newer field of multiscale simulations will have a strong impact on future scientific leadership and competitiveness. With appropriate support, Europe would be well positioned to develop pioneering multiscale software frameworks.

One challenge is that computational needs are very diverse in the specific scientific domains, ranging from strong-/weak-scaling demands in astrophysics and materials research, to task level parallelism in multiscale simulations with $O(10^3)$ cores in every task, and to replica exchange and ensemble simulations in soft matter research, materials design and statistical physics.

On the competitive front, large scale initiatives to develop application software are in place and growing in the U.S., with the goal of creating or significantly adapting codes for next generation supercomputers. Planned co-design centres in the U.S. will likely take software development for exa-scale computing a large step forward. In addition, Japan and China have the world's first and second most powerful scientific computers as of this writing (www.top500.org) and appear determined to continue investing heavily in hardware and software development. China, in particular, has the advantage of not having to advance legacy software.

At a minimum, Europe has to establish support actions so as not to lose further ground to competing nations. We recommend that Europe go beyond this to pursue leadership in important areas of

[1] *Financing a Software Infrastructure for Highly Parallelised Codes: IDC Final Report for the DG Information Society of the European Commission.* July 2011

exascale-capable scientific software. Examples of current initiatives in this direction are HP2C in Switzerland and Simulation Laboratories in Germany. More broadly, Europe's scientific community organizations should play a more important role by interacting with application and developer groups and with grant organizations to develop a parallel software strategy for the exascale era.

Education and training in high performance computing also has to be improved to prepare for software development on more complex hardware architectures and to train next generation scientific programmers, as well as to map more complex physical scenarios involving multi-physics and multiscales to the more complex, heterogeneous hardware layouts that future-generation HPC systems are expected to exhibit. Fault tolerance and energy efficiency should not only be considered on the system programmers and hardware level, but also on the application software level, since the design and implementation of energy efficient algorithms for applications have a strong impact on the efficient use of future architectures.

2. Introduction

Physics and chemistry applications have exploited supercomputer capabilities since the inception of high performance computing and often have been a driving force for improvement and further development of algorithms and architectures. Due to the high complexity of the simulated natural systems and the associated algorithms, along with the need for gathering large and/or long data sets, the fundamental sciences have always been able to take advantage of the largest compute systems in the world. Therefore it might not be surprising that the codes with the highest performance typically originate from physics and chemistry. This has been reflected in the annual Gordon Bell prizes, which are awarded to the highest-performing applications (the 2010 applications that exceeded petaflops performance were in chemistry and materials sciences).

Extending the computational scale 1,000-fold from petaflops to exaflops gives rise to new challenges for applications, and also to the interplay between applications and hardware. Although no exascale computer exists yet, it is reasonable to expect that several concepts of traditional programming will have to be abandoned. Since 2000 there has been a nearly exponential increase in the number of compute cores (considering the Top 20 supercomputers on the semi-annual Top-500 list). A simple extrapolation from present technology suggests that multi-core technology will increase the concurrency on a compute node to $O(10^3)$ or even $O(10^4)$, while the number of nodes increase to $O(10^5)$ or $O(10^6)$. Although it might be hoped that the system memory will increase by O(100), this still means that the memory/core will strongly decrease. Therefore, to use this large number of compute cores most efficiently will require both a proper design of numerical algorithms and a proper mapping of memory distribution and efficient memory access.

A simple extrapolation of the energy consumption of an exascale machine from present petascale machines would point to gigawatts. In practice, a tolerable limit is in the range of 20 megawatts. Therefore, (i) technology has to be improved to meet the energy constraint, and (ii) energy efficient algorithms have to be developed for both system software and application software. In the latter case, the benchmark will be not only performance, but also energy consumption, which might shift the traditional *time-to-solution* metric, which is usually minimized towards the metric (*energy* x *time-to-solution)*. In addition, fault tolerant algorithms will be of high importance, since it is anticipated that the mean-time-to-failure of a system will be decreased to O(1 day), necessitating hardware support to handle errors, as well as library support (e.g., fault tolerant MPI) and fault-tolerant numerical algorithms (e.g., data reconstruction by backward integration).

In summary, the transition from peta- to exascale calls for enabling simulations of natural systems with a thousand times more degrees of freedom, a thousand times more complexity, or a thousand times longer time-scales, enabling significant breakthroughs in multiple scientific fields. But these achievements will require a large effort to adapt present codes or applications for efficient use on exascale machines, or else re-designing and re-writing codes for different sets of applications.

Fundamental science has a strong focus on the understanding of mechanisms, interrelations and governing principles in nature. This general description includes sub-disciplines which might seem far removed from technological requirements or societal benefits, such as high energy physics or cosmology but which rely heavily on computational resources, are strong drivers for algorithm- and method-development which also has relevance for industrial research. Hence, fundamental sciences should not be thought of as isolated disciplines but as research directions. They explore possibilities that may advance our current understanding of the world and point towards practical uses serving societal needs.

Categorizing the fundamental sciences may be done by disciplines and domains, or by the time- and spatial scales which characterize the fields. Picture a two-dimensional map, where time- and length scales are defined by the characteristic scales of the objects under study, such as elementary particles, charged particles, atoms, molecules, aggregates or large scale structures which exhibit a range in spatial scales from $<10^{-10}$ m to $>10^{+20}$ m, and time scales from $<10^{-15}$ s to $>10^{+15}$ s. This

exercise produces fields with more than 30 orders of magnitude differences in space and time. From the domain side, this diagram is spanned by quantum chromodynamics/high energy physics, nuclear physics, laser-plasma physics, nuclear fusion research, quantum chemistry, soft matter research, materials sciences and astrophysics/cosmology. Although there is a large diversity in length and time scales and therefore also different methods and paradigms in the domains, various domains are interlinked by multiscale simulation approaches that aim, for example, to bridge the gap between a microscopic atomistic description and a macroscopic continuum description of matter.

Since time- and length-scales often characterize the physical systems under consideration, large efforts are often concentrated on the extension of time scales, and/or length scales, i.e., extending the number of particles or number of degrees of freedom comprising the system, in order to observe new phenomena emerging in the system, or to increase statistics and therefore accuracy. Since the computational effort can often increase non-linearly with the system size, a large effort is spent finding the best implementation for a given algorithm or to reduce the complexity of the numerical algorithms which solve the problem.

From a computational viewpoint, the efficiency (in relation to the computer's peak performance) of parallel algorithms strongly depends on the data reuse, data locality and also on the complexity of the solver. Algorithms which, e.g., extensively apply matrix-matrix multiplications are good candidates for codes with a high numerical efficiency. If it is possible to partition these algorithms into loosely coupled or embarrassingly parallel tasks over a large set of processors, it is possible to achieve performance close to the peak performance of the computer. Examples may be found in materials science and chemistry, e.g., in DFT for calculations of response functions in the random phase approximation[I1] or Monte Carlo simulations of spin systems[I2].

On the other hand, efficiency may be also related to scalability, either with strong or weak scaling. To scale up to the largest number of available processors requires a small enough ratio between communication and computation, and therefore requires either purely local communication between neighbor processes, or a heavy computational load. In the case of strong scaling, there is also the requirement for a good load-balancing of work on each processor to guarantee minimal waiting times and communications overhead.

Methods used in fundamental sciences often explore particle or mesh methods (or combinations of both). For example, in statistical physics, physical chemistry, plasma or astrophysics applications, particle methods often are used to explore molecular dynamics, Brownian dynamics, Monte Carlo or particle-in-cell (PIC) methods. The scalability of these methods strongly depends on the computational load and the locality of the algorithm. PIC codes often show a better scalability, due to their locality, than molecular dynamics methods, which include long range interactions requiring information exchange over the whole system. Grid-based methods, as applied for example in hydrodynamics or magneto-hydrodynamics, include finite element or finite volume methods work on structured or non-structured adaptive meshes. Applying solvers such as multigrids introduces problems on large scale platforms, e.g., load-balancing in terms of grid partitioning and maintaining a balanced workload over the different grid levels.

The development of fast and efficient algorithms has always been important for the fundamental sciences. Therefore, exascale computing will impose new challenges and requirements for algorithm development and simulation codes. This will be needed to improve *time-of-discovery* to *time-to-market* by using exascale architectures in design and device studies. The most efficient usage of these machines will require, on the one hand optimal algorithms like Fast Multipole methods, multigrid or H-matrix methods; and on the other hand, the development of local algorithms such as real space methods that may, for example, apply wavelets. The combination of optimal order and data local algorithms is a promising approach to achieve energy efficient and fast methods.

Not all disciplines in the fundamental sciences will be able to efficiently use exascale machines, either in terms of strong or weak scaling. This might be due to a super-linear increase of memory consumption when up-scaling the problem size, or to inherent limitations in the parallelism of a given algorithm. Therefore, the efficient usage of exascale machines might be considered on different levels

of parallelism. Strong and weak scaling up that exploit the entire computer will certainly be exceptional in most fields. Even where it is possible, to achieve this goal a very large effort has to be made in the development and adjustment of codes. It is certainly necessary to rewrite and redesign codes to achieve strong scaling. This implies that legacy codes should not be considered as candidates for highly scalable applications on exascale machines. New algorithms and specialized implementations are required to achieve optimal scaling on $O(10^6\text{-}10^9)$ cores that are expected to have much less memory/core than on current Tier-0 architectures. The most flexible implementation would, therefore, not build on existing codes but on new frameworks, using specialized and optimized libraries and modules for compute-intensive and communication-intensive parts of the program.

An alternative to strong/weak scaling on whole exascale machines is to use task-parallelism and to partition nodes and cores into groups, each running different tasks or executables. This scenario would reflect the needs of a horizontal multiscale approach, where simultaneously different levels of detail are considered in one simulation and consequently appropriate methods are devoted to each resolution level, e.g., quantum chemistry methods for electronic details and bond breaking, molecular dynamics for particle motions, and finite elements for mechanical and structural quantities. Therefore the efficient use of a massively parallel architecture is distributed into different tasks, which might not all have a sufficient scaling behavior for implementation across the whole machine. This approach would also apply to important commercial codes, where companies are not prepared to invest sufficient money to adjust a code design for the efficient use of a whole exascale machine. Similarly, this might also apply to community legacy codes, which could be run on parts of an exascale machine within such a multiscale simulation.

On the other hand, weak scaling could, in principle, be a solution for programs which do not exhibit strong scaling properties. However, the most detailed description on an electronic level is often not necessary for a large system and would introduce an additional time overhead. Furthermore, weak scaling is sometimes not possible due to a super-linear increase of memory requirements. Since the memory/core will be reduced on an exascale machine, compared with current Tier-0 architectures, weak scaling would not be a general solution for various applications. The task parallel multiscale simulations therefore could alleviate the problems associated with both the strong and weak scaling approach.

A third approach to parallelism relies on embarrassingly parallel approaches, such as ensemble simulations, running the same system with different initial conditions, or tightly coupled simulations, like replica exchange simulations, used to study rare events or to optimize structural properties of very complex systems. Accordingly, this type of simulation guarantees an efficient usage of large scale parallel machines while reducing simulation times almost linearly to study complex parameter spaces or to increase statistics.

Following these considerations, the transitions from peta- to exascale machines offer a great opportunity to increase the efficiency of simulations and tackle new problem dimensions. This can provide new insights (e.g., astrophysics, nuclear physics), enable the numerical verification of theoretical concepts (e.g., high energy physics, cosmology), validate concepts (e.g., fusion research), make predictions for experimental findings in complex setups (e.g., plasma physics) or establish the basis for new approaches of material design, discovery and device modeling (e.g., soft matter, materials sciences).

In the present report, fundamental science is represented by nuclear/hadron/high energy physics, astrophysics, cosmology, soft matter physics, quantum chemistry, plasma/fusion physics and materials science. Although the field is actually much broader, this selection of disciplines provides a good profile of the time and length scales appearing in the fundamental sciences, as well as the different methods and simulation software used by the communities. The report aims to provide an overview of the current status and perspectives of the disciplines in view of the transition to exascale computing. In addition, cross-disciplinary aspects are considered, as well as the general European position within fundamental sciences. The report also aims to discuss educational and societal needs and tries to estimate the costs for the transition from the peta- to the exascale era.

List of Experts of Work Group 3.3: Fundamental Sciences

Name	Organization	Country	Field / Domain
Godehard Sutmann	CECAM / JSC	Germany	Chair
Jean-Philippe Nominé	CEA	France	Vice-Chair
Mike Ashworth	STFC	United Kingdom	Methods and Algorithms
Alessandro Curioni	IBM	Switzerland	Algorithms
Thierry Deutsch	CEA	France	Materials Sciences
Pieter in t'Veld	BASF	Germany	Soft Matter
Alexei Removich Kokhlov	MSU	Russia	Soft Matter
Laurent Lellouch	CNRS/CERN	France	High Energy Physics
Nicola Marzari	University of Oxford / EPFL	UK / Switzerland	Material Sciences
Ulf Meißner	University of Bonn / FZ Jülich	Germany	Hadron/Nuclear Physics
Maurizio Ottaviani	CEA	France	Fusion
Mike Payne	University of Cambridge	United Kingdom	Quantum Chemistry
Thomas Schulthess	CSCS	Switzerland	Material Sciences
Luis Silva	Technical University of Lisbon	Portugal	Plasma Physics
Volker Springel	HITS	Germany	Astrophysics
Romain Teyssier	ETH	Switzerland	Astrophysics
Adrian Wander	STFC	United Kingdom	Materials Sciences
Gilles Zerah	CEA	France	Condensed Matter

3. Science Drivers and Grand Challenges in Physics and Chemistry

3.1. Astrophysics and Cosmology

3.1.1. Computational Cosmology

The next decade will see the advent of large-scale cosmological experiments, both from the ground (LSST) and from space (Euclid and WFIRST). The science objective of these very ambitious galaxy surveys is to determine the exact nature of dark energy, probably the main theoretical challenge to the standard model in physics. The technique used by these various projects is to measure to great accuracy the growth of matter fluctuations in the universe (using weak-lensing) or the positions of the Baryons Acoustic Oscillations (using galaxy clustering), and to relate these to the various cosmological parameters, especially the ones describing dark energy. One key step in determining these parameters is to compute reliably the clustering of dark matter for a given set of cosmological parameters. Since gravitational dynamics is a complex non-linear problem, we need to use large N body simulations, covering the same volume covered by these surveys, and the same galaxy mass range detected by these surveys. Practically, we need to use a box size of 6000 Mpc/h (up to the cosmological Horizon), with *at least* 100 particles per halos of size L*/10, where L* is the luminosity of the Milky Way. This translates into a prodigious number of particles, namely $N=16384^3$ or 4 Trillion bodies. State-of-the-art N body solvers (GADGET, PKDGRAV or RAMSES) usually require 200 bytes per particle. With 2 GB per core, taking into account memory overheads (x2), this will require a 750000 core machine. Such extreme simulations will be of great value for preparing and exploiting forthcoming dark energy experiments. Data compression of the simulation results will be a critical component of such projects. On the fly halo and sub-halo finding algorithms will be used to compress raw N-body data into a more practical form that will be post-processed later to generate mock galaxy catalogues and gravitational shear all-sky maps.

3.1.2. Galaxy Formation and the Cosmological Context

Galaxies exist in a bewildering variety of shapes and sizes, ranging from dwarf satellites in our own Milky Way, to grand design spirals and huge elliptical galaxies. Our understanding of galaxy formation is still extremely sketchy at best, even though a basic paradigm for a theory of galaxy formation exists (“hierarchical galaxy formation”). The fundamental problem is that galaxy formation involves a blend of different physics that is non-linearly coupled on a wide range of scales, leading to extremely complex dynamics. Specifically, of primary importance are the self-gravity of dark matter and gas, hydrodynamical shocks and high Reynolds number turbulence, radiative cooling processes in the gas, radiative transfer, star formation and evolution, non-gravitational energy input into diffuse gas by supernovae or black hole accretion, and magnetohydrodynamics. The set of partial differential equations describing this blend of physics is well known, but largely inaccessible by analytic techniques. For this reason, HPC simulation techniques have become the primary avenue for theoretical research in galaxy formation. This is also helped by the fact that the current standard model of cosmology precisely specifies the initial conditions of cosmic structure formation at a time briefly after the Big Bang. It becomes then a computational problem par excellence to try to evolve this initial state forward in time, staying as faithful to the physics as possible.

Current state-of-the-art hydrodynamic simulations of galaxy formation reach in the best cases, of order 1 billion resolution limits in the gas, and an equal number of bodies for the dark matter. The most expensive runs in the field have thus far consumed up to a few million CPU hours, and the set of physics they include is typically restricted to gravity, hydrodynamics, radiative cooling, and a sub-resolution treatment of star formation and feedback processes. It is imperative to dramatically advance these simulations to allow proper interpretation of upcoming observational data in cosmology. For example, new radio telescopes like LOFAR or SKA are bound to revolutionize our understanding of the high redshift Universe, providing 21cm tomography of the epoch of cosmic reionization. The successor of the Hubble space telescope, JWST, as well as extremely large optical telescopes of the 30m class here on Earth will peer back in time to observe the infancy of galaxy formation in unprecedented detail, yielding insights into the formation of the first generation of objects. Large galaxy surveys that are currently under way or are to commence in this decade (such as Pan-STARRS[AC5], Big-BOSS[AC6], etc.) will drastically improve the statistical constraints on galaxy formation and evolution. Finally, the astrometric GAIA mission[AC4] will, in about five years, completely transform our understanding of the fine-scale structure of our Galaxy and of the Local Group, making near-field cosmology a reality.

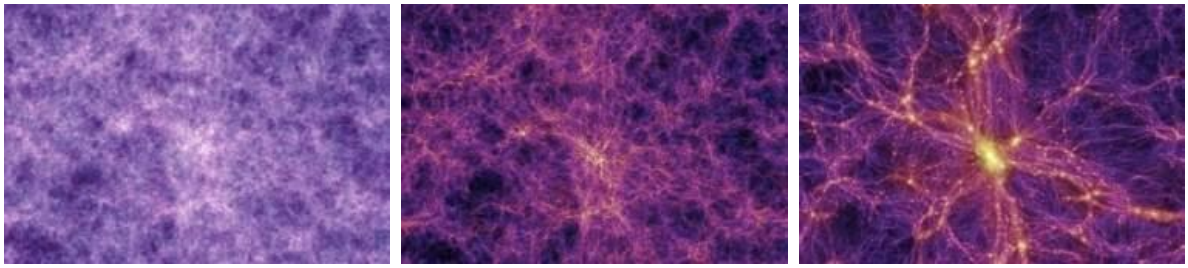

Figure 1: details of the millennium simulation at different redshift for the same co-moving area [AC3]

To address the reionization problem, future simulations need to resolve a complete census of the source population of star forming galaxies, down to halo masses of a few times 10^8 Msun, in a sufficiently large volume of at least 100 Mpc/h. This needs to be combined with an on-the-fly treatment of radiative transfer, something that is not possible at adequate resolution with current computational techniques. In fact, cosmological galaxy formation codes that are able to efficiently combine radiative transfer and hydrodynamics and still show good scaling to 10^4-10^5 cores and beyond do not exist yet, but work in this direction has started, especially in the RAMSES and AREPO codes. The scientific goal to make detailed and reliable multi-frequency predictions for the phase transition of cosmic reionization will require very substantial computational resources that are only delivered by of order $O(10^5)$ cores.

A major goal in theoretical cosmology is to arrive at simulation models that can successfully explain the detailed structure and morphology of spiral galaxies like our own Milky Way. However, the best present calculations largely fail to reproduce the observed prevalence of pure disk galaxies; instead they form overly luminous galaxy bulges with anemic disks. The disks themselves are still hopelessly under-resolved to allow a faithful study of the formation of spiral arms, stellar bars, or thin and thick disk components. Also, numerical studies of stellar migration in the disk though scattering on spiral waves, and the build up of the stellar halo through the disruption of infalling satellites, are equally frustrated by a lack of sufficient resolution. The best state-of-the-art simulations of Milky Way-sized objects are now reaching a mass resolution of 10^3 Msun in the dark matter, resolving the Galaxy with a few billion particles. A resolution increase by a factor ~1000 will allow Milky Way models that come close to being able to represent every single star by a simulation particle. Such simulations will be an extremely powerful tool for the full exploitation of the GAIA data, but they require one to several hundred billion particles and a machine with at least 10^5 cores (assuming 2 GB/core).

3.1.3. Star Formation in the Galactic Context

Understanding the formation and evolution of galaxies requires understanding how they convert their diffuse gas, obtained from large cosmological reservoirs, into stars. This proceeds through the compression of gas into dense and cold clouds, thermal and magneto-gravitational instabilities forming pre-stellar cores, and the regulation by interstellar turbulence and by the radiation field and explosions of young massive stars, causing winds and outflows out of galaxies. Modern numerical simulations can probe the infall of gas onto galaxies, or the conversion of gas into star-forming regions. However, the triggering and regulation of star formation depend on so many processes that act in concert on very different scales (from individual explosions to extra-galactic outflows) that a thorough modelling cannot be performed yet. How galaxies preserved massive gas reservoirs to continue star formation in today's 13-billion-year-old Universe remains a mystery in modern cosmology. Adaptive mesh refinement techniques (e.g., the RAMSES code) offer stronger and stronger scaling performances when large dynamical ranges are explored, and can encompass thorough multi-physics modelling. As of today, they can be employed to resolve the physics of star-forming clouds inside galaxies, or the growth of galaxies inside cosmological models, using up ~10000 CPUs on petascale computers. The use of such techniques on exascale computers will enable us to directly resolve, in self-consistent models, all the important scales in the gas accretion and star formation history of galaxies, from individual star-forming cores and supernovae explosions to large-scale extragalactic outflows.

3.1.4. Planet Formation

Planets form in turbulent gaseous protoplanetary discs rotating around newly born stars. Understanding how they form is a major challenge in modern astrophysics and requires a detailed knowledge of the properties of the flow in these disks. Today, such simulations largely rely on local simulations (only a small volume of the disk is considered) with typical resolution being at most 200 cells per disc scale height. Exascale computers will provide the opportunity of performing such simulations on the global scale, covering the entire planet formation region (located from 0.5 to 10 astronomical units from the central star). Using 16384 GPGPUs, the total resolution of such MHD simulations will be up to 8192x2048x2048, divided in equal blocks of 128^3. Given the current estimate of CPU time required to update each cell with a shock capturing scheme (about 10 micro-seconds) and today's typical estimates of the speed-up obtained on GPGPUs for such a scheme (of the order of 50 for blocks of 128^3), these give a total simulation time of 10 million GPGPU-hours, which can be converted to about 20 days of wall-clock time (5 million time steps are required to cover the 300 dynamical times that are required to study the turbulence properties). Such a simulation environment will provide a fantastic laboratory which will enable us to study the conditions under which planets forms.

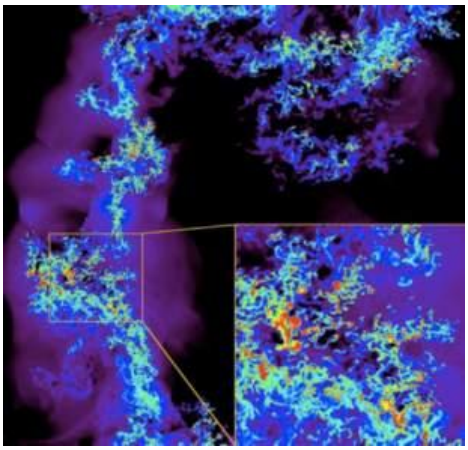

Figure 2: Simulation of interstellar turbulence

3.1.5. Stellar Magnetism

Understanding global solar/stellar convection, turbulence, rotation and magnetism and their nonlinear coupling requires the latest class of supercomputers available in order to compute the largest range of spatio-temporal scales possible and improve the realism of solar/stellar 3-D magnetohydrodynamics (MHD) numerical simulations. For instance, in the Sun it is necessary to couple the convective envelope to its deep radiative interior in order to describe with the highest fidelity possible the solar

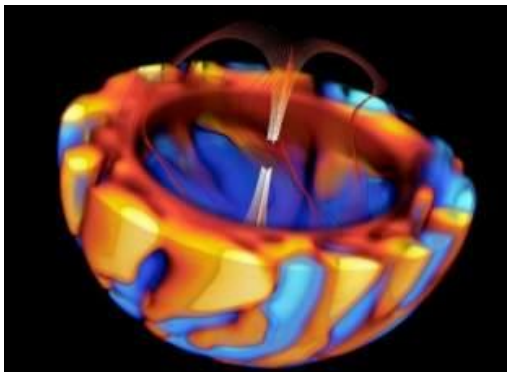

Figure 3: ASH simulations of the velocity field and magnetic field on the sun[AC2].

dynamo and tachocline (an intense zone of shear at the base of the convection zone thought to be at the origin of the large scale magnetic field responsible for the 11-yr activity cycle), the excitation and generation of internal waves and turbulent convective penetration and magnetic pumping. All these processes require a high radial resolution of about 5000 points, in order to model the small scale dynamics present in stratified magnetized turbulence. Further to model on a full sphere the solar surface small scale granulation one needs at least to reach spherical harmonic degree of $l_{max} \geq 6000$. Overall a resolution of 5000x9000x18000 is required at least to model a realistic Sun over at least several decades if one wishes to solve the puzzle of the 11 year cycle. This constitutes a grand challenge both in resoluton and timescale for exaflops (50,000+ cores) supercomputers. Given the global geometry of the problem one approach is to use about 1000 cores for the radial direction (via a domain decomposition) and about 500 to 1000 cores in the horizontal direction (using a spectral decomposition of the modes). Fat node supercomputers would be better suited to such solar physics grand challenge. The next stage will be to couple the interior solar dynamics to its environment via a wind or a magnetosphere.

3.2. Particle, Hadron and Nuclear Physics

3.2.1. Background

The Standard Model (SM) of particle physics is one of the towering achievements of contemporary physics. It unifies three of the fundamental interactions of nature that are based on quantum field theories, namely the strong, the electromagnetic and the weak interactions. Despite its enormous successes, many open questions remain, that are tied to a variety of research directions. These include probing higher energies to reveal new physics phenomena beyond the Standard Model (BSM). Such BSM physics is also expected to explain the shortcomings of the SM. This is the area of particle (high-energy) physics. Here, one usually differentiates between the energy frontier and the intensity (precision) frontier. While the former can be characterized by the production of ever heavier particles with ever increasing collision energies, complementary to this one can look for virtual excitations of massive particles at fixed energies, requiring as much luminosity (collision rate) as possible to achieve a very high precision. Closely related to it is the field of hadronic physics, that tries to explain how the strongly interacting particles are generated from the underlying gauge theory of quarks and gluons, Quantum Chromodynamics (QCD). Naturally connected to this field is the area of nuclear physics, that tries to explain the generation of atomic nuclei (the baryonic matter in the universe) as well as matter under extreme density or temperature conditions, such as found in the core of neutron stars or the early stages of the history of the universe as reconstructed in ultrarelativistic heavy-ion collisions. All this research is tied to billion Euro investments in facilities like the LHC or FAIR.

3.2.2. Challenges in Particle Physics

In high-energy physics, there are three priority research directions for numerical simulations in the next decade (see also [HN1]):

Searching for BSM physics: The search for new physics at the intensity frontier requires performing ever more stringent test of Standard Model predictions for rare processes. Abundant high-precision data from mesons containing light and heavy quarks have and will become available. To find the traces of BSM physics, one has to calculate such decays to a precision that is significantly below the experimental uncertainty, which would then allow to reach energy scales beyond the reach accessible at the LHC. Complementary to this, precision lattice QCD (LQCD) calculations at the energy frontier will be required to interpret possible signals of new physics as most of the underlying models of new physics are intrinsically non-perturbative.

Testing QCD at the sub-percent level: Two of the three components of the SM have been verified with great precision. Tests of QED attain the extraordinary precision of a few parts in 1010, while the electroweak sector is tested on the per mille level. By contrast, QCD has been tested much less accurately. At both high energies, where perturbative methods can be applied and at low energies, that are intrinsically nonperturbative, verification on the level of 5% to 10% is typical. A grand challenge for theory is to bring this accuracy below 1%. This would be an extraordinary achievement – precise numerical control over a non-perturbative quantum field theory. Such sub-percent calculations of QCD are also a necessary ingredient for the tests of BSM physics mentioned before, and it would also tell us how to simulate accurately other strongly coupled quantum field theories.

Simulating possible theories of BSM physics: The LHC is probing physics at the energy frontier. By the dawn of the exascale era, the LHC will have run for several years and hopefully will have provided tantalizing evidence of BSM physics. Theoreticians have already produced an overwhelming array of candidate models. To really understand the options and make a definite discrimination between experimental signatures requires investigations of lattice field theory beyond QCD. Two important classes of BSM theories are supersymmetry and those that are nearly conformal (invariant under rescaling of lengths). Computations in both classes of theories are significantly more challenging than in QCD. The ultimate aim is to allow a screening of the viability of the various models in view of the experimental hints from LHC and elsewhere. Once a particular model proves viable, one would then undertake more detailed computations. In such a way, the interplay of theory and experiment could allow one to zoom in into the correct interpretation of BSM physics.

3.2.3. Challenges in Hadronic Physics

Hadronic physics is the investigation of the spectrum and interactions of the strongly interacting particles made of quarks and gluons. In this field, two major challenges exist:

The spectrum of QCD: Computing the bound state spectrum of QCD is vital to claim a complete description of the strong interactions - the emergence of structure from QCD defines one of its grand challenges. A major breakthrough has been achieved recently by the spectrum calculation of the low-lying hadrons made from up, down and strange quarks at almost physical quark masses by the BMW collaboration [HN2] (see figure).

Still, almost all hadrons are resonances and thus decay. While methods are being developed to extract resonance properties from lattice QCD simulations, most of these require multiple lattice volumes that require extreme computing resources. Also, the role of gluons and the inclusion of very noisy isoscalar (disconnected) quark contributions are required to make direct contact to the many existing and upcoming data on hadron resonances from various facilities world-wide.

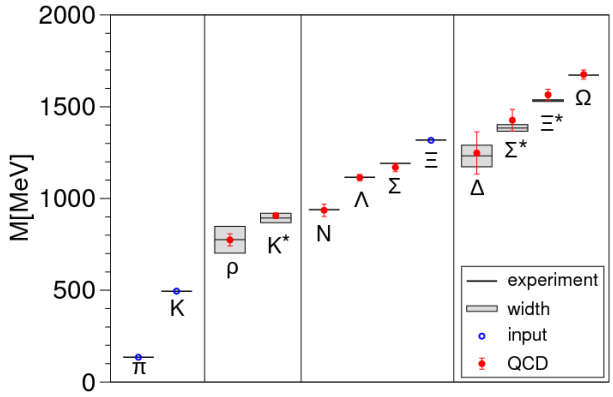

Figure 4: The spectrum of the low-lying hadrons made from up, down and strange quarks as calculated by the Budapest-Marseille-Wuppertal collaboration.

How QCD makes a proton: Protons and neutrons (denoted collectively as nucleons) are the building blocks of the baryonic matter in the universe. Made themselves from quarks and gluons, LQCD calculations are able to derive the structure and properties of nucleons from QCD, see the figure for the quark distributions in a proton. Of highest interest are the flavor-singlet and sea-quark contributions as well as an understanding of how gluons contribute to such a fundamental property like the nucleon spin. In particular, so called disconnected contributions, that are plagued by severe noise problems, will have to be computed precisely if one wants to calculate proton and neutron properties separately. In addition, the investigation of short-distance physics in hadrons requires very fine lattice spacings. All these tasks require exascale computing.

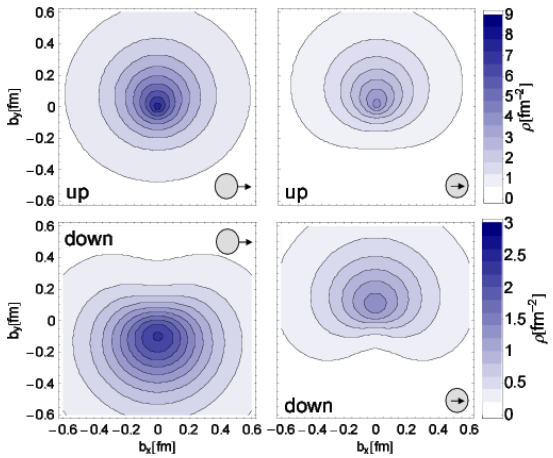

Figure 5: The density of unpolarized quarks in the transversely polarized proton (left) and the density of transversely polarized quarks in the unpolarized proton (right) as calculated by QCDSF.

3.2.4. Challenges in Nuclear Physics

For the area of nuclear physics, the following grand challenges have been identified:

How QCD makes nuclei: Besides from hadrons, structure in QCD emerges also in the form of atomic nuclei. While many properties of nuclei and nuclear matter have been understood and are being investigated in phenomenological models and effective field theories (EFTs), deriving the formation of nuclei directly from LQCD remains a venerable challenge [HN3]. The computational challenges are three-fold, namely the large noise problems for the Greens functions of nuclei, the very small binding energies compared to the total mass of the systems and the large number of Greens functions to be computed due to the combinatorics of quark field contractions. This field is only developing now and will require extreme computing to achieve maturity and become a comprehensive program. For example, it has been estimated that the ab initio calculation of the helium nucleus, that consists of two protons and two neutrons, will require about one sustained exaflop in computational resources [HN4].

Precision calculations of nuclei: EFT has been successfully applied to derive the forces between two, three and four nucleons. This framework can be combined with Monte Carlo (MC) methods to perform ab initio calculations of atomic nuclei and their reaction dynamics (so-called nuclear lattice simulations). These simulations profit form the approximate SU(4) of the nucleon-nucleon interactions and are therefore much less plagued by the sign problem as lattice QCD computations are. Presently, nuclei up to atomic number 12 have been calculated. A spectacular recent success was the first ever calculation of the Hoyle state in carbon-12 that is required for the generation of heavy elements in stars, making carbon-based life possible [HN5]. With exaflop capabilites, this scheme will allow to provide calculations of spectra and reactions up to atomic number 40. Other approaches will allow to bridge the gap to higher mass numbers and more complicated reactions, just to mention variational MC methods, the no-core-shell model and the attempts to derive density functionals from microscopic nuclear forces. Such type of calculations also require large memory. These calculations will allow for an understanding of very proton- or neutron-rich instable nuclei and the generation of elements through the r-process in stars. They also form the theoretical backbone of the various radioactive beam facilities world-wide.

Phases of QCD and its thermodynamics: Predicting and exploring the phase diagram of QCD will not only deepen our understanding of the interplay of quark confinement and chiral symmetry restauration, but also leads into the domain of very novel and exotic forms of matter like e.g. the color glass condensate (the ``perfect liquid''). First, finite temperature calculations for zero net baryon number density will have to be performed with chiral fermion discretizations. This can only be done with exascale computing resources. Second, a major challenge is presently the exploration of matter at finite baryon density, which is requires new methods to overcome the sign problem in Monte Carlo simulations with finite chemical potential. In particular, the existence or non-existence of the tri-critical point in the QCD phase diagram is of highest interest. Such calculations will also set strong constraints on the development of models of the high-density regime of strongly interacting matter.

Investigating the perfect liquid: Experiments at RHIC and the LHC have established an almost-perfect liquid picture of the strongly interacting quark-gluon matter as produced in violent heavy-ion collisions. Calculation of transport coefficients and in-medium hadron spectral functions will allow to quantify the relation between the successful modeling of heavy-ion collisions based on the data from RHIC and the LHC and QCD. One of the outstanding problems to be solved is to calculate the time required to form the strongly interacting quark-gluon plasma after the onset of the collision. The full description of the initial state of a heavy-ion collision and subsequent thermalization of matter requires methods that self-consistently describe varies time-scales that describe very different phenomena (saturation, chromo-instabilities, etc.). Contact to the experimental situations can only be made possible with exascale computing, as the figure from an actual collision at the LHC indicates.

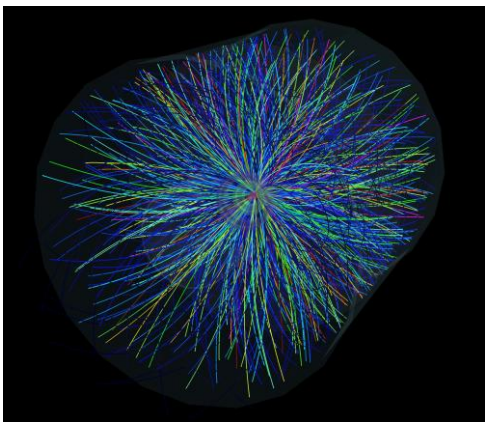

Figure 6: The first heavy-ion collision measured with the ATLAS detector at the LHC. The large number of hadrons emerging from the violent heavy-ion collision reveals the features of the almost perfect quark-gluon formed.

3.2.5. Required Computational and Infrastructure Resources

Hardware requirements: Lattice field theory (LFT) calculations can exploit a broad range of hardware platforms. Central to LFT are extreme scale capability machines. The generation of gauge field configurations is the single most CPU extensive task in LFT simulations. The calculations of physical quantities from these ensembles require smaller partitions and can be carried out on workstations clusters or farms GPUs. The substantial requirements of lattice ensemble generation are expected to be well suited to future exascale computers. These calculations require a challenging balance between floating point capability and communication bandwidth and latency. However, given the steep

scaling of computation cost with lattice spacing, the total memory requirements grow slowly with computing performance. Current trends in petaflop systems appear efficiently matched to these LFT calculations. Even with economical mesh-like layers, a regular problem like LFT achieves a good performance, and this trend is expected to continue to the exaflop scale. Thus, we expect that LFT will continue to be one of the leading application areas demonstrating the transformative potential for extreme scale computing capability.

Software infrastructure: As the complexity of both numerical algorithms and the computer hardware grows, it becomes increasingly difficult for LFT collaborations to keep up. A move to community-designed and maintained software that is accessible, well-documented and verified might thus be desirable. This has not yet been initiated on the European level, though there exist large transnational collaborations which share projects, codes and/or gauge field configurations. While is lack of European-wide software development has its drawbacks, it also has advantages in a field where highly optimized algorithms and codes are essential to produce important science in a timely fashion. It fosters a diversity and flexibility of computational approaches which may be better suited to solve particular problems and which are necessary to validate the results obtained with any given approach.

3.3. Magnetized Plasmas and Nuclear Fusion

Magnetised plasmas are almost ubiquitous. In the Universe, matter is observed prevailingly in the plasma state. Likewise, magnetic fields of different strengths and configurations occur in stars, planets, as well as in the interstellar and in the intergalactic medium. In the laboratory, magnetised plasmas are produced for research purposes and in view of various applications, and in particular in the context of thermonuclear fusion research.

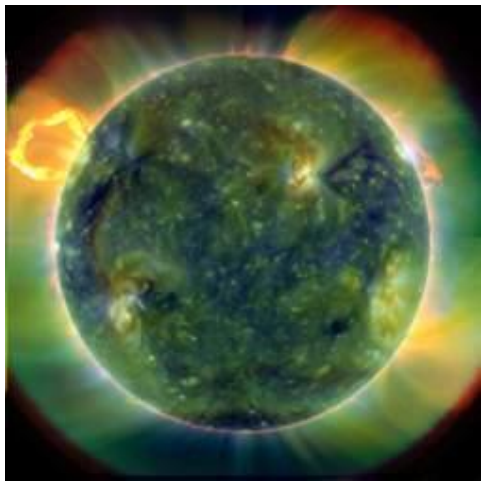

Figure 7: Multi-wavelengthImage of the Sun from the Solar Dynamics Observatory.

Understanding the dynamics of magnetised plasmas is of primary importance for the understanding of our world. Theories involving magnetised plasma as key ingredients were advanced in the attempt to solve such fundamental problems as the generation of magnetic fields during the early stage of the Universe and the formation of stars. More closely to us, solar physics is a very active field of research, with new satellites becoming operational, and with increased interest in the numerical simulations of the complex solar dynamics and the impact of solar phenomena on the terrestrial environment. Specifically, the present-day theoretical challenge lies in the need to resolve the equations of magneto-hydrodynamics in the turbulent regime across a large portion of the Sun, from the thin (in relative terms) boundary layer, called the tachocline, marking the transition between the inner region dominated by radiation and the outer convective zone ruled by hydrodynamics, up to the solar surface with its rich phenomenology. Even more ambitiously, future simulations aim to target the full 11 year solar cycle, including a treatment of the coronal phenomena and the generation and time-variation of the solar wind. In parallel, the interaction of the solar wind with the geomagnetic environment is also a subject of study of great interest. Besides their interest as natural phenomena (aurorae, etc.) magnetospheric perturbations are also of practical interest since they can affect satellites, disrupt telecommunications, and occasionally affect power grids. The combined study of observation and simulation of the combined Sun and Earth magnetoplasma system is often referred to as the field of Space Weather.

Magnetic fusion research seeks to reach thermonuclear conditions by containing plasmas with strong magnetic fields in suitably designed devices. Burning plasma conditions, at the density achievable in present-day devices, requires heating the plasma to very high temperatures (of the order of a hundred million degrees), and correspondingly high gradients (10-20 million degrees per meter). At these

conditions, energy and particles are lost by the device through turbulent convection. Understanding transport of energy and matter is among the key questions of this field of research, and one of great practical interest, since the efficiency of the eventual power station would depend on the ratio of the produced energy to the energy required to operate the device.

In both fields, the computational challenges stem from the very high spatial scale separation required to model any significant portion of the system, to the extremely long integration times.

Fusion devices are several meters across, whereas turbulent structures occur at the millimetre scale and significant magnetic disturbances at the scale of at least several centimetres. In a machine like ITER, encompassing all the important phenomena underlying energy losses would require simulations of several thousand grid points, in the two directions perpendicular to the magnetic field (less in the parallel direction). Furthermore, in fusion devices, plasma collisions are so rare that the mean free path along the field lines can be longer than the characteristic macroscopic scale. Modelling transport along the field lines by a fluid closure is problematic. The modern trend is to use a reduced form of the kinetic equation for each plasma component (electrons and ion species). At best, this requires a resolution in the hundreds for two (parallel and perpendicular) velocity directions.

In the case of ITER, the smallest time scale to be resolved is the one associated with millimetre-size vortices, which is typically of the order of a microsecond. The energy confinement time can be of the order of a second. Thus, depending on the time-advancing algorithm, a significant simulation to steady-state energy balance would require tens of millions of time steps.

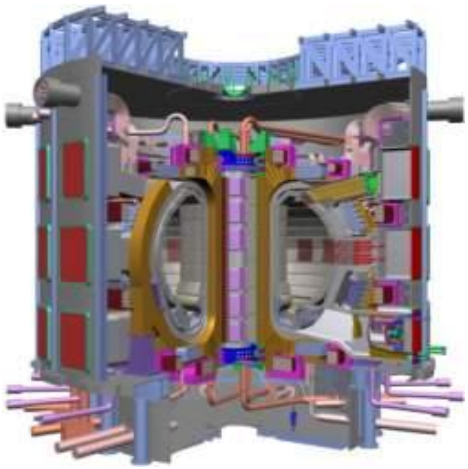

Figure 8: Schematic of the ITER fusion facility.

Solar physics simulations are not less challenging. Solar plasmas are characterised by a variety of inter-playing phenomena, with a wide range of plasma parameters. Magnetism originates as a result of a yet ill-understood dynamo mechanism in the convective region. This couples to the Sun surface where events such as protuberances, flux emergence, solar flares occur, and where the solar wind is generated. Turbulence at high Reynolds number is the rule, and spatio-temporal scale separation is extreme (8-10 orders of magnitude). Present research aims at MHD simulations of at least ten thousand grid points in each spatial direction for very long integration times.

The possible economic benefit from top level computing capacity in the field of magnetised plasmas is self-evident. Achieving thermonuclear fusion as a possible future energy source is a great dream of mankind. It requires very large investments, and, as in many other fields, computer simulations are an effective way to steer applied research in the right directions. Europe is the World leader in magnetic confinement with the Joint European Torus (JET) as the top running experiment and as the partner contributing the largest fraction of the construction costs of the international project ITER. Solar physics modelling, coupled with observation can anticipate disrupting solar activity and contribute to the strategy to protect investments in satellites and telecommunications. Europe is also the leading partner in several research satellites, Ulysses, SoHO, Cosmic Vision Solar Orbiter.

Forefront computing capabilities are an essential cost-effective mean to back all these investments and to strengthen or maintain leadership in these fields.

3.4. Atomic Physics

Atomic physics addresses the quantum dynamics of electron and positron collisions with atoms and molecules, plus field/multiphoton interactions with atoms and molecules in laser fields, including strong fields and ultrafast laser pulses. These are the fields of interest of (among other projects) the UK-

RAMP consortium (Queen's Belfast, UCL, Open University, STFC Daresbury). The Daresbury PRMAT suite for 'R-matrix' ab initio electron atom scattering is being developed under UK-RAMP, through HECToR DCSE projects and general CSED SLA support. The 'outer region' modules of this code are being adapted for the UKRmol molecular R-matrix suite (UCL, OU, supported by DL) while the 'inner region' modules also form the basis of the field-atom TDSE codes TDRM and RMT (developed at QUB): RMT is a multi-electron extension to the world-leading HELIUM code for two-electron systems in arbitrary pulses. In the next phase of UK-RAMP, RMT technology will be applied to UKRmol for ab initio treatment of molecules in arbitrary pulses. In addition the POSITRON suite of codes (Nottingham and Daresbury) treats positron molecule scattering and annihilation, and interactions between antihydrogen and small molecules, using explicitly correlated leptonic wavefunctions.

Scientific grand challenges in atomic scattering include accurate treatment of atoms with open d-shells (with relativistic 2-body terms to be included), and inner region double-continua for intermediate energy collisions. Applications are in astrophysics (the iron peak elements) solar and atmospheric physics, engineering ('clean' Mo lighting and laser-produced plasmas using Sn) plus diverse theoretical atomic collision data for the ITER and HiPER projects for fusion and laser–ignited fusion. Accurate treatment of double-continua is also essential for understanding inner shell excitation and subsequent relaxation (and light emission) of complex atoms by laser pulses. Electron scattering by molecules has similar astrophysical and atmospheric physics applications while the characterization of temporary anion states (resonances) in large molecules is crucial for the understanding of radiation damage in biological systems. The planned extension of RMT to UKRmol is a grand challenge leading to the understanding and coherent control of molecular processes in laser fields. Positron scattering introduces new interaction/reaction paths for molecules with the addition of annihilation and positronium formation processes. With the CERN ALPHA project having very recently produced trapped antihydrogen, theoretical data on antihydrogen interactions (eg, He is an 'impurity' which can lead to postronium formation and break-up of the antiatom) is particularly needed. Computationally, the grand challenges, apart from introduction of new science into the programs, include extension (and in certain cases, introduction) of parallelism at both multicore and peta/exa-scaling levels to cope with the complexity of the multi-channel, multi-electron wavefunctions and interactions: propagation, Hamiltonian diagonalization, (complex) operator matrix construction and extraction of scattering parameters, together with defining the accuracy of variational calculations that do not intrinsically involve a minimum principle.

3.5. Soft Matter Physics

Soft matter (or soft condensed matter) is an acronym for an increasingly important class of materials such as colloids, liquid crystals, surfactants, foams and gels, granular materials, polymers, polymer nanocomposites and complex macromolecular assemblies, proteins and other biological systems, including membranes, cells and tissue. Typically, these materials are built of organic molecules and they are often either amorphous or can self-assemble from structural elements. There are often many levels of topological and geometrical complexity with hierarchical, supramolecular structures that can be cooperative, far from equilibrium, and sensitive to external conditions. Computer simulations have proven to be an indispensable, if not the most powerful, tool to understand properties of these complex materials and link theoretical models to experiments.

3.5.1. Grand scientific challenges

Soft materials share an important common feature in that their behavior is determined by a subtle interplay between strong entropic effects (due to large thermal fluctuations) and relatively weak interactions among molecular or supramolecular components that occur at an energy scale comparable to room temperature thermal energy. Another common essential feature of the soft matter systems is the huge span of characteristic lengths and time scales, from the subatomic, over the atomic and mesoscopic to the macroscopic. As a result, an enormous range of length and time scales has to be covered and the computational tools required to describe them are extremely diverse.

Particularly difficult problems arise when describing the dynamics of soft materials. The dynamic processes in these systems involve a whole hierarchy of characteristic times, ranging from the very fast scales (e.g., single-bond vibration) to the very slow ones. Characteristic temporal scales can easily span 10 orders of magnitude. Due to new heavily customized massively parallel supercomputers and large scale distributed computing, one can now simulate molecular systems with up to millions of atoms for periods on the order of hundreds of nanoseconds, using atomic-level molecular dynamics (MD) coupled with theoretically-driven, accelerated MD schemes (e.g., hyper- and metadynamics, parallel replica and replica-exchange schemes). However, far longer time scales – microseconds, milliseconds and beyond – are needed to simulate such "slow" processes as transitions from a low viscosity liquid to an amorphous solid (solidification), directed crystallization under the influence of soft molecular assemblies, rare events, crack propagation, development of dislocations, enzyme kinetics, nucleation and protein folding. Enhanced simulation methods or sampling techniques are also required for effective exploring "rough" (highly frustrated) energy landscapes with a multitude of local minima separated by high energy barriers and modeling the processes that control the self-assembling of macromolecules and lead to the formation of partly ordered metastable structures and heterogeneous multiphase materials. The development of hybrid (multiscale) simulation strategies that combine complementary theoretical approaches and thereby bridge the different spatiotemporal scales, opens the way to considerably extending the range of problems that can be solved.

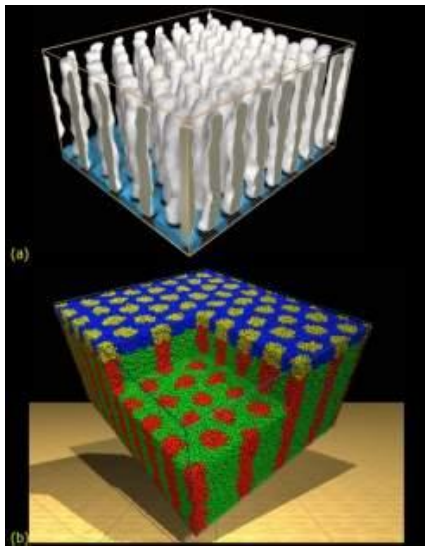

Figure 9: Hexagonal cylindrical phase of diblock copolymers in thin films formed through template-guided self-assembly.

A rigorous simulation of chemical reactions is facing even more serious challenges due to the time scale problem. This challenge has two major aspects. First, reliable quantum chemical calculations based on e.g. density functional theory (DFT) have to be implemented in such a way as to retain DFT's accuracy while performing MD simulations for soft condensed materials with tens or even hundreds of thousands of atoms. This is related to the so-called O(N), or linear scaling, problem. Second, there is currently enthusiasm about classical reactive MD methods (in particular, ReaxMD) with which it is possible to speed up the calculation by several orders of magnitude as compared to first-principles calculations. However, ReaxMD requires the development of automated methods of generating reactive force fields for a broad range of materials by incorporating quantum-calculated information and training reactive models appropriately.

A rigorous treatment of many interesting phenomena that occur away from equilibrium is another great challenge. Far-from-equilibrium behavior is ubiquitous in the soft matter world. It arises across the entire spectrum of condensed-matter and materials physics in a host of problems of fundamental interest and is intimately connected to cutting-edge materials processing technologies. The field is vast, and it is unlikely that any particular organizing principle will work for all far-from-equilibrium soft-matter systems. Nonetheless, there is great value in identifying classes of systems that might have common underlying physics or that might be tackled by common methods.

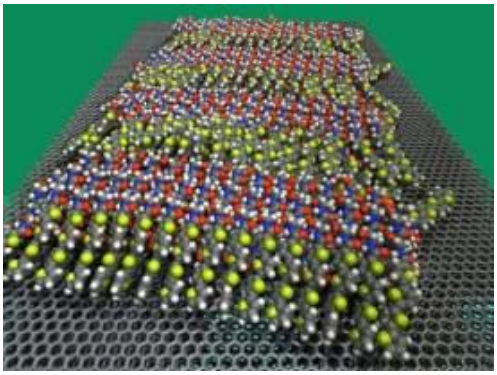

Figure 10: Thiophene-peptide block copolymers ("molecular chimeras") self-assembled on a graphite surface.

Studying the interfaces and nanobiointerfaces between soft matter (organic molecules such as polymers and proteins, biomembranes and cells) and hard matter (minerals, metals, semiconductors) is a rapidly growing research area in which simulations are likely to become increasingly valuable. Important applications here include the biomimetic design of materials with unique properties, the drug design, the biomineralisation, the processes of biorecognition and adaptation, etc.

Many of the problems in soft matter science are addressed successfully in practice. With currently available petaflop-level computing, the possibility exists to achieve predictive capabilities to manipulate microstructures and interfaces to enable the design and development of advanced soft materials, such as polymer nanocomposites, ion-exchange membranes for fuel cell applications, organic optoelectrical materials, polymer-supported catalysts, bio-inspired functional supramolecular structures that combine synthetic and biological parts, etc.

The advent of exascale computing provides a new opportunity to design advanced soft materials through a combination of electronic structure, atomistic, and mesoscale simulations. However, the availability of increased computing power will not be sufficient in many cases. What is required is the application of accelerated simulation methodologies and the development of new algorithms/codes for relevant physical phenomena, optimized to extreme-scale computing architectures.

3.5.2. Focal problems

Leadership-class petascale machines and ultra-high performance special-purpose supercomputers designed for MD simulations (such as ANTON) dramatically increase the speed of calculations, making possible the simulation of molecular systems at an atomic level of detail for periods on the order of a millisecond. Very-large-scale fully atomistic MD simulations of say, 1012 atoms (i.e., at the micron scale) up to the macroscopic observation time of millisecond and more may well be feasible in the not-too-distant exaflop era. Such brute-force simulations can be considered as a real computer experiment that can closely mimic the problem at hand on engineering scales. But to realize this potential, several scientific and computational problems must be solved.

Within the past decade, priority has been given to developing truly multiscale methods for modeling materials properties, with seamless integration of atomic scale, intermediate length scale, and continuum methods. Similar multiscale approaches are needed to treat materials physics involving time evolution on a wide range of timescales. While great progress has been made, additional breakthroughs are needed. The seamless aspect can be achieved by running large-scale simulations with an atomistic method that directly overlaps the spatiotemporal scale at which a mesoscale/continuum method is used. This will enable direct verification of the coupling algorithms used between different methods and will thus create a direct predictive modeling capability. One of the related problems is to increase the scalability of existing/future codes with respect to the exascale resources. More efficient, accurate, and broadly applicable first-principles methods must be developed

to study chemical reactions, dynamics, effects of thermal fluctuations, and excited states of materials. Particular difficulties are connected with models that are capable of capturing the essential physics of materials with strongly correlated electrons. The density matrix renormalization group (DMRG) formalism, the phase retrieval methods as well as new forms of linear-scaling quantum Monte Carlo have promise for significant impact in this field. Unfortunately, these computational schemes are very time-consuming. Both in the numerical study of simple models and in first-principles simulations, the development of new efficient accelerated molecular-dynamics methodologies (e.g., for automatic identification of characteristic motions and rare events in simulations) is needed. In all of these problems, the necessary ingredient for continued success is collaborations with applied mathematicians and computer scientists to create codes and algorithms that will scale to the largest machines in order to effectively utilize computational resources at the highest end.

3.5.3. Computational and algorithmic challenges

From an algorithmic viewpoint, the challenges include: (i) Developing massively-parallel linearly scalable methods and algorithms for classical MD, DFT and first-principles MD calculations showing linear-scaling performance up to 1 M CPUs (spatial scalability is more or less ok, but temporal scalability is difficult). A number of linear scaling methods for Kohn-Sham DFT have appeared in the literature (orbital minimization, density matrix minimization, Fermi operator expansion, divide-and-conquer strategy, dual-level hierarchical scheme for linear-scaling divide-and-conquer correlation theory, subspace iteration method). With a large number of processors, in general, the divide-and-conquer class of algorithms should be an optimal computational strategy. Further improvement can be achieved by distributing individual atoms among multiple CPUs, using many sophisticated parallelization schemes involving sparse linear algebra, FFTs, grid kernel routines, etc. (ii) Developing reactive MD with quantified accuracy, so dynamics can be modeled at the micron/microsecond scale with a chemical accuracy of 1 kcal/mol. (iii) Adaptivity is the key for applications to effectively use available resources whose complexity is exponentially increasing. (iv) The development of compilers for hybrid architectures.

H/W and S/W related problems: (i) Techniques for which communication is minimal can efficiently address new (hybrid) architectures (e.g., GPGPU). This imposes the development of "data locality principles" (divide-and-conquer strategy). (ii) Transition to hybrid/heterogeneous parallelism to expose scalability in algorithms: combination of general purpose H/W + GPGPU + inter-chip/intra-chip parallelism + new interconnect topologies; effective combined MPI/OpenMP parallelization techniques have to be developed (this can be achieved e.g. with object-based parallelization like that realized in Charm++, CnC, Plasma, Fork-join parallel BLAS).

3.5.4. General trends

1. It is expected that with exascale computing resources, the focus of simulations will shift from a qualitative description of basic phenomena to the optimization and quantitative prediction of soft materials properties, including the description of microstructural evolution and chemistry-driven problems. First-principle simulations will play an increasing role in these areas. Furthermore, one can expect that the main accent will gradually shift from materials simulation to device modeling.
2. Today’s petascale computers and the next-generation exascale computers tend to be heterogeneous, with stacks, nodes, and cores (processors) at three different levels. These hardware developments pose daunting challenges for software developers.
3. It would be desirable to have a more modular/cellular architecture, with building blocks that can be replicated ad infinitum as necessary, and later reused. The blocks should have low power/high density characteristics, should also be cheap to build and easy to interconnect.
4. For ultra-high-performance computing, the best way would be a combination of General Purpose H/W with application specific integrated circuits (such asMDGRAPE-3 or ANTON).
5. The application of reconfigurable hardware architectures based on FPGAs for soft matter simulations represents another promising trend in hardware/algorithm development.

Software currently used in soft-matter simulation

1. Large-scale atomistic MD simulations: LAMMPS, DL_POLY, NAMD, GROMACS, AMBER, reactive molecular dynamics (ReaxFF MD); lattice and off-lattice Monte Carlo methods.
2. Mesoscale coarse-grained particle-based simulation methods: Dissipative particle pynamics, smoothed particle hydrodynamics, highly discretized particle dynamics (e.g. lattice Boltzmann method, multi-particle collision dynamics), patchy particles methods.
3. Ab initio (Car-Parrinello/Born-Oppenheimer) MD simulations: CPMD, Quantum-ESPRESSO, ABINIT, CP2k, CASTEP, NWChem, VASP, Qbox, CONQUEST, hybrid QM/MM methods.
4. Coupled atomistic/coarse-grained simulation methods, hybrid particle-field simulation methods, hybrid particle-based/continuum dynamics simulation methods. Field-theoretic (self-consistent field, dynamic density functional theory, RISM/pRISM) methods: complex Langevin dynamics, time-dependent Ginzburg-Landau (TDGL) method, macroscopic Cahn-Hilliard (CH) and Cahn-Hillard-Cook (CHC) models, commercial S/W like MesoDyn, Mesotek+, Palmyra.

3.6. Quantum Chemistry

3.6.1. Background

In less than twenty five years, first principles or ab initio quantum mechanical simulation techniques have progressed from modelling either a few electrons or one or two atoms to modelling systems containing many hundreds of atoms. Using such calculations, it is now possible to make reliable predictions about the structures and properties of many atomistic systems. There have been numerous applications of such first principles calculations to study surfaces, point defects such as vacancies and impurities, extended defects such as grain boundaries and dislocations, catalysis and many, many other applications. In recent years, these techniques have been extended to predict a wider range of physical and chemical properties and to predict theoretical values for experimentally measured spectra such as optical, nuclear magnetic resonance (NMR) and electron paramagnetic resonance (EPR) and many others. Given the predictive power of quantum mechanical techniques, it is not surprising that these methods are already widely used in industry.

A very diverse range of quantum mechanical based computational methods have been developed within chemistry and physics. Researchers in these fields were early adopters of parallel computing in the early 1990s as they required the increased computational speed and memory available on these machines to be able to apply their techniques to relevant systems. It is important to note that progress has been driven not only by increasing computer power but also, and usually more importantly, by improved theoretical and numerical approaches. This intense effort over many years to improve methodologies does, however, mean that there is limited scope for continuing to develop existing methods up to the exascale in the same way as has been done in the past – primarily by allowing application to larger systems. For instance, calculations on large systems often require large amounts of interprocessor communication across the whole machine and this will simply not be possible on exascale machines. The cost of most conventional quantum mechanical computations scales at least as the cube of the number of atoms or electrons in the system so, for instance, moving from petascale to exascale computing would only double the lengthscale in a simulation of a bulk material. This is clearly not a good use of computing power and it is also totally unnecessary given the existence of linear scaling techniques such as ONETEP. Given the diversity of ab initio methods and techniques available there are, of course, exceptions to this rule so, for instance, quantum Monte Carlo methods are known to scale exceptionally well on massively parallel computers and should have little difficulty exploiting exascale resources. Some first principles calculations offer a degree of almost embarrassingly parallelism such as linear response calculations of phonon spectra for which each phonon wavevector is an independent calculation. This allows some codes to use of thousands of processors efficiently for some tasks but even here there is a limit to the number of phonon wavevectors needed and thus a limit to the number of processors that can be efficiently used. Thus, it is believed that of the quantum mechanical methods and approaches that have been very successfully developed to exploit terascale to petascale compute resources very few can be developed to the exascale. While this might appear to sound the death knell for quantum mechanical simulations, fortunately this is not the case and, on the contrary, exascale resources will provide the launch pad for

a sea-change in the way that quantum mechanical simulations are exploited in the future. The goal should be to move the molecular design process to the computer in the way that many other design processes in science and engineering have already moved. This will generate an enormous societal and economic benefit - one that Europe is well positioned to exploit given Europe's ownership of the vast majority of computer codes in this field but one which will only be achieved if it moves forward rapidly.

3.6.2. Ab initio materials discovery

The cost of quantum mechanical calculations has, in the past, generally restricted their use to look at individual systems or problems, usually those where experiment has identified interesting behaviour. The large number of mature quantum mechanical approaches which can accurately determine physical and chemical properties of systems containing up to a few hundred atoms using terascale to petascale resources combined with the availability of exascale compute resources will allow us to use the computer to discover new molecules and materials with specified properties needed for a particular industrial application. This ultimate promise of computational materials design has often been cited as a reason for supporting the development of quantum mechanical simulation codes in the past but, until now, the computational resources necessary to realise this vision have not been available. Indeed, the first stage of this paradigm shift is already taking place as exemplified by work on ab initio crystal structure prediction and in the materials discovery work of a number of researchers, perhaps most notably Gerd Ceder. For instance, with exascale resources it will be possible to ask the computer to discover a material which has a particular dielectric constant and lattice parameters that are matched to another material to that it can grow epitaxially on the other material. In the process of searching for the required material, the search will also discover many other materials with different properties and it is crucial that this information is archived and made available to any other researcher for whom it may be exactly the material they require for a different application. Over the next few years, a massive increase in the number and size of such ab initio databases of materials and molecular properties is expected and it will be critical that these databases mesh seamlessly with existing databases of experimental properties and that suitable search tools are available to allow any researcher easy access to the information in the databases.

3.6.3. Ab initio design of fabrication

Ab initio discovery of materials will become routine with exascale compute resources. However, this discovery process will only have an impact if the materials discovered can be fabricated. This will almost invariably require the material to be produced within a complex multi-facetted product and most often, for instance in the case of solar power, will only have any appreciable economic and societal impact if the product can be deployed on a massive scale which will only be possible by using low cost fabrication techniques. Ab initio quantum mechanical simulations must therefore be developed so that they can be used not only to discover new materials and systems but also to help design suitable fabrication paths for creating them. Ab initio determination of fabrication paths is much more challenging that materials discovery but, remarkably, there are already examples of success. For instance, MEARS Technologies Inc. have used quantum simulation techniques to develop a process for producing atomically precise layers using only standard industrial fabrication tools. Nanotechnology offers the promise of far higher packing densities of devices which have much lower energy consumption than present electronic devices. However, this promise will only be fulfilled when the challenge of mass fabrication of nanoscale components is solved. While considerably more challenging than materials discovery there are elements of the fabrication process that can be addressed simply by extensions of the techniques required for materials discovery. For instance, showing that a series of epitaxial layers of different materials have lower free energies than all other arrangements of the same atoms at a chosen fabrication temperature for particular chemical potentials of each element will show that this structure will be preferentially formed under these fabrication conditions (temperature, chemical potentials). Subsequently showing that the free energy of the fabricated structure at all temperatures between room temperature and the fabrication temperature is lower than the free energy of any structure produced by displacement of one or more atoms will prove the long term stability of the fabricated structure. However, growth processes are complex and this information about preferred structures and their stability will have to be accompanied by an understanding and control of growth mechanisms. These mechanisms can be studied directly using ab initio through long timescale molecular dynamics simulations or indirectly by first obtaining a fine

scale sampling of the potential energy surface and subsequently applying transition rate theory to this surface. The latter approach is intrinsically parallel and immediately amenable to exascale computing, the former approach is discussed in the following section.

3.6.4. Dynamics

A single long timescale ab initio simulation cannot be efficiently performed on exascale computers because of the inability to scale ab initio calculations to the exascale. However, it should be possible to access long timescales on exascale computers by combining the results of large numbers of parallel dynamical simulations run on un-correlated configurations. Considerable work needs to be done to prove that this approach does correctly model long timescales. There is considerable further potential for enhancing the efficiency of this approach by collating information from the parallel simulations and using this in a metadynamics type approach to prevent any of the simulations revisiting configurations that have been explored already. Once suitable protocols for this approach have been developed, these need to be encapsulated in expert systems that can be applied to any complex system. These technical developments will, effectively, bring long time dynamics within the reach of ab initio simulations thus allowing the predictive power of these approaches to be extended to a much wider range of systems and phenomena than is currently possible. Perhaps most notably, it will become possible to perform predictive ab initio simulations of biological processes which will identify new pathways for therapeutic intervention which circumvent the problems confronting conventional approaches to drug development.

3.7. Materials Sciences

3.7.1. Introduction

Condensed matter physics and chemistry try to describe the matter of the scale of atoms, as well as their properties, in order to build new devices, new molecules and study new phenomena. The basic equation is known: The Schrödinger equation, which needs to be solved in order to simulate the atomic systems. The problem is that this equation using only two-body interactions comes from intrinsically a many-body theory which considers the wavefunction of N electrons. It is really difficult to avoid the many-body wavefunction, which complicates considerably the resolution of the Schrödinger equation.

For less than 10 electrons, the Schrödinger equation can be solved exactly. For larger systems one is easily confronted with computational limits as the problem scales exponentially. For more electrons, the theory which is widely used is the Density Functional Theory (DFT), which has no parameter (so it has a predictive power) but uses some approximation for the electron-electron interaction. To simulate a device, a model is required for this device. There are phenomenological-type models, but there exist also physical models which consider the problem on the level of the atomic scale.

One example is the modeling of a fuel cell for which a model has been developed using irreversible thermodynamics. This model, MEMEPHYS, needs to consider the chemical reactivity of the species and other related quantities. To calculate these quantities, accurate ab initio calculations are required taking into account as much as possible the correct physics. A future scenario will consider multiscale modeling of materials to take into account a whole device which will be based on ab initio calculations (simulation) in order to be coherent and predictive, i.e. ab initio simulation and modeling. What is required in practice is to simulate atomic systems up to about $O(10^4)$ atoms using ab initio methods in order to build modeling tools.

3.7.2. Simulations in Material Sciences and Chemistry

Doing simulations, one main goal is to save money (experiences) and human-time, as well as giving hints to experimentalists and to predict the physics. Therefore scenarios like predicting the structure of materials and molecules in the presence of defects, the dynamics or evolution and their properties should be considered.

There are many properties of materials which need to be calculated. To do simple, if the property depends strongly on the many-body interaction (we can reduce easily to a pair interaction between

atoms), then this property is rather difficult to simulate and accurately to predict. One example is the case of superconductivity. Since this is a property which depends strongly on many-body interactions between electrons and atoms, at the present stage, simulations are able to predict some qualitative results but it is far away from a quantitative comparison. There is no robust method available up to now for a quantitative prediction.

Another example for a physical property is the electronic transport as for instance the electron mobility. In order to compare quantitatively with experiments, temperature (phonons) needs to be considered which increases strongly the cost of calculations since interactions between electrons and phonons have to be considered properly, but which have not the same locality.

To simulate a device for thermo-electricity or photovoltaics, different properties have to be combined (electronic and thermal transport) and materials or molecules will be determined in function of their properties. Therefore, a systematic search of atomic systems needs to be performed in order to select the right candidate materials. This means that methods are needed which explore efficiently the space of atomic configurations with high accuracy. Properties are strongly related to stable atomic structures. Determining the more stable structure is often not as simple. Furthermore, defects need to be calculated in order to study the dynamics (kinetics), the aging or the evolution of atomic systems as materials or molecules.

Statistical physics is the second theory which determines the kinetics of the atomic systems. The Schrödinger equation predicts the interaction and the fast dynamics. Statistical physics describes the influence of temperature and aging and it needs to consider the space of atomic configurations and to integrate over it. There are methods like molecular dynamics to calculate time averages or methods like Monte Carlo which sometimes are more efficient to calculate ensemble averages and explore faster the configuration space.

To do simple, a kind of Catia software is needed to build and predict new structures and properties of materials or molecules. One goal would be to have a tool to build atomic structure and to get an idea or a prediction of their stability, aging and calculated physical and chemical properties. Even more, we would like to have a given property, e.g. photovoltaics, and determine the best candidates of molecules or combinations of different molecules for these properties under given physical and thermodynamic conditions.

3.7.3. Grand Challenges in Material and Chemistry

Growth and Kinetics of Materials

Given a material, the growth of it and more widely the calculation of the dynamics or kinetics is a great challenge which needs an exhaustive and efficient exploration of atomic configuration space as molecular dynamics or statistical methods. A large number of calculations are needed for each material or given molecules. This challenge concerns e.g. protein folding or aging of materials for which the concurrence of elementary mechanisms has to be studied. To understand dynamics, methods like kinetic Monte Carlo methods or other methods using the elementary mechanisms as inputs are needed. To give an estimate for this scenario, the calculation of systems consisting of about O(1000) atoms and O(100,000) evaluations of forces are required, where one evaluation costs between 1 and 10 TFlops depending on the nature of atoms. Therefore the required compute capacity which is needed will be $10x10^{15}x10^{5}$=1 ExaFlops, which is not necessarily done in strong- or weak-scaling but can be split into different replica simulations. Since there is a vast number of possible atomic systems (i.e. materials or molecules) to study, there will be a huge demand of compute power to be foreseen in order to tackle this computational challenge.

Finite Temperature and Excited State ab initio Simulations

The second kind of computational challenge concerns the determination of material properties not only at zero temperature but also at room temperature or even higher. For an illustration consider the color of a material. At the present time, the error of finite temperature calculations is in the range of about 100meV which means that it is not possible to distinguish between blue and green color.

Both the calculation of excited states as well as combining different physical and chemical properties are challenges which are very important for the physics today. To give an illustration, consider the field of organic photovoltaics. Using two different molecules (acceptor and donor of electrons), it is first required to study the structure of the mixing. Then it is needed to determine the characteristics of the intensity over voltage as function of the light intensity. To do that, the best way is to have a model and

to simulate the main basic features: production of an exciton (electron-hole pair), the separation between electron and hole in function of an applied electric field, the diffusion and recombination of exciton and the collect of electrons and hole. Altogether this means to have a model and quantitative simulation at the atomic levels.

The more time-consuming part are the ab initio calculations. The modeling part is not so CPU-time-consuming. With the combination of simulation at atomic levels and modeling, the modeling up to the scale of devices should be feasible. The modeling can have many levels of descriptions, linked together or not. The key point is the modeling has to be done on a physical basis and not on a phenomenological one in order to have predictive power for materials properties.

The calculation of properties is more time-consuming. Each calculation takes O(10)-O(100) TFlops. 1000 calculations of properties seem to be reasonable for an almost exhaustive study, which sums up to a cost between 10 and 100 Pflops. Besides the computational challenges, there are many algorithms and physics to develop even if the equations are well known.

Therefore, to model a device or a class of problem, the determination of the structure as well as the calculation of properties are required in order to compare with experiments. Estimations for the computational costs are in the range of one exaflops. In addition to the computation of physical properties of materials it will be an important task to build databases, from which data will available for materials design and device studies. Therefore, database as well as data mining will be an important field in materials science in the future.

3.7.4. Towards Exa-scale Computing in Materials Science

From the viewpoint of the user, the time-to-solution, the robustness of a method and the simplicity of usage of software are the three main criteria (especially the first two of them) which count. At the present time, ab initio code can be used on CPU and GPGPU for about 500 atoms. Doing more is not tractable for the user because the scaling of $O(N^3)$ as function of the number of atoms discourages to consider bigger systems for a small improvement. The cubic behavior is sensitive for systems larger than 500 atoms and limits current simulations to this size.

Combining MPI+OPENMP+GPGPU can give a order of magnitude improvement, but the main advance will come from algorithmic improvements, especially from robust (or pragmatic) order O(N) methods. At the present stage, the ab initio code BigDFT using wavelet basis sets, PEtot and Qbox using plane wave basis sets have already some version combining MPI+OPENMP+GPGPU.

With order O(N) methods, considering bigger systems should be possible in a reasonable time-to-solution if the crossover point between order N and N^3 calculations is not so far to 1000 atoms, which is dependent on the system under consideration. It is seems also quite possible to have order O(N) methods to determine properties. The uncertainty concerns the accuracy of the calculations. Finally concerning the calculation for strong correlation effects, many algorithmic improvements need to be done, which is quite challenging.

In conclusion, materials sciences will need sustainable petaflops/s to do more than 99% of the calculations but also needs exaflops/s capacities to tackle the grand challenges of the upcoming years.

3.8. Laser-Plasma Physics

Particle-in-cell simulations are one of the most widespread and common numerical tools in plasma physics. The key algorithm was first proposed and developed in the late 50s' by John Dawson, then at Princeton University, and Oscar Buneman in Stanford (for a review see [LP1]). It was one of the first algorithms to take fully advantage of computers to model the temporal evolution of a many body systems. Particle-in-cell simulations play an important role in many topics in plasma physics, whose underlying physics is highly nonlinear and depends on the self-consistent trajectories of individual charged plasma particles. PIC codes are used in almost all areas of plasma physics, such as fusion energy research, plasma accelerators, space physics, astrophysics, ion propulsion, and plasma processing, and many other areas. PIC algorithms are also used in cosmology, astrophysics, accelerator physics, and semi-classical quantum simulations.

The holistic comprehension of plasmas, in particular in conditions which are highly nonlinear such as those associated with intense sources of beams or radiation (such as lasers) is highly complex, and it requires *ab initio* fully kinetic simulations of which the PIC method is one of the most effective possibilities. In the PIC method the trajectories of a large number of particles are solved self-consistently using forces calculated from field equations solved on a grid. When modeling plasmas, the field equations are the Maxwell's equations for the electric and magnetic fields. These models work at the most fundamental, microscopic level. As a result, they are the most compute intensive model in plasma physics.

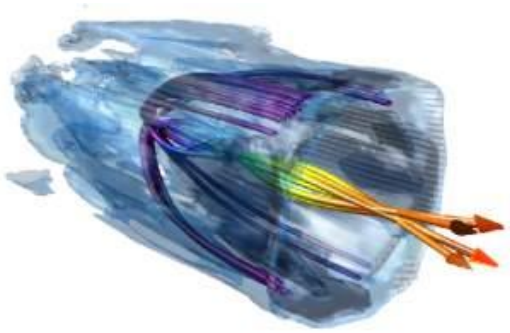

Figure 11: Results from a LWFA simulation (OSIRIS) showing electrons being trapped and accelerated in a wakefield.

Unlike molecular dynamics codes, where particles interact as binary pairs, particles in PIC codes interact via fields calculated on a grid. PIC codes are possible whenever there is some (mean) field equation describing the fields in terms of (particle) sources. There are a variety of PIC codes in common use, differentiated by the kinds of forces retained in the model. The simplest is the electrostatic force, described by a Poisson equation, and the more complex are the fully electromagnetic models. PIC codes are increasingly being used to validate reduced plasma descriptions, such as fluid models. Moreover, in fully electromagnetic relativistic PIC codes the equations to be solved are covariant, and, if properly implemented, the fully relativistic PIC model is also covariant thus allowing for simulations to be performed in any Lorentz frame, as recently demonstrated by simulations in Lorentz boosted frame for plasma-based accelerator [LP2] (Figure 1) for one of the state-of-the-art PIC codes now used in production [LP3].

PIC codes generally have three important procedures in the main iteration loop. The first is the deposit, where some particle quantity, such as a charge or current, is accumulated on a grid via interpolation to produce a source density or current density. The second important procedure is the field solver, which solves Maxwell's equations or a subset to obtain the electric and/or magnetic fields from the source densities. Finally, once the fields are obtained, the particle forces are found by interpolation from the grid, and the particle coordinates are updated, using Newton's second law and the Lorentz force. The particle processing parts dominate over the field solving parts in a typical PIC application.

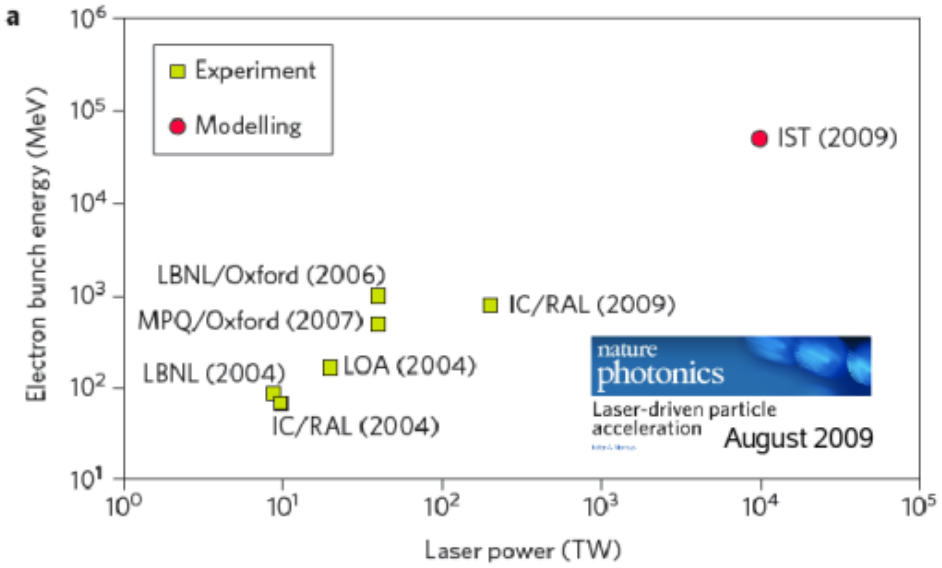

Figure 12: Highlight of the results of boosted frame simulations in the modelling of plasma-based accelerators (reprinted from P. Norreys, Nature Photonics 2009).

In plasma physics, some of the most demanding scientific and computation grand challenges are closely tied with recent developments on ultra intense laser technology and with the possibility to explore with fidelity astrophysical scenarios, previously not possible due to limitations in computing power. The main scientific challenges are on (i) plasma accelerators (either laser or beam driven) and the associated advanced radiation sources based on these accelerators, which hold the promise of providing secondary sources for bio imaging and medical therapy, (ii) Inertial fusion energy and advanced concepts with ultra intense lasers, which can lead to the demonstration of nucler fusion ignition in the laboratory, and (iii) collisionless shocks in plasma astrophysics, associated with extreme events such as gamma ray bursters, pulsars and AGNs.

These are topics of relevance not only from a fundamental point of view but also in terms of potential direct economic benefits. For instance, research in plasma accelerators will lead to a new generation of more compact and cheap particle and light sources, a topic where Europe is clearly leading due to the large scale pan European laser projects (e.g. Extreme Light Infrastructure -ELI [LP4] and High Power Laser for Energy Research – HiPER [LP5]) and the national efforts on the development of laser based secondary sources (e.g. in Germany, France, U.K.), and the exploration of an alternative path for nuclear fusion is critical for a sustainable energy production, which is the driving force for economic growth.

PIC codes are the critical tools not only to advance the knowledge of these systems but also to perform comparisons with experiments and with observations. It is important to stress that many of the multidimensional features of these scenarios remain totally unexplored due to the complex and highly nonlinear nature of the underlying plasma microinstabilities that preclude a detailed and exhaustive theoretical treatment, and that until very recently could not be appropriately explored with numerical simulations due to the lack of the appropriate tools and of the extreme computing power required to fully resolve these conditions. The PIC methods extend far beyond the area of plasma physics and astrophysics; the study of the most advanced particle-based simulation methods is at the forefront of supercomputing now being routinely used to test the largest supercomputers in the World.

Particle-in-cell simulations are particularly suited to take full advantage of the most recent developments in High Performance Computing as highlighted the strong scaling results recently obtained in Jülich (see Figure 2) [LP6] demonstrating excellent strong scaling scalability, thus showing the maturity of the PIC technique to take advantage of these extreme resources. Moreover, there is as strong effort in the community of PIC codes to tap into the full power of new

architectures/configurations, (e.g GPUs, SIMD/vector units, etc) with some important milestones already achieved that demonstrate the ability to use these algorithms in the high end machines and in a wide variety of hybrid architectures, the paradigmatic example being the code VPIC that has been extensively used in the Roadrunner computer in Los Alamos.

PIC simulations are also at the forefront of science as demonstrated by recent high impact results that have leveraged on the combination of the PIC algorithms, running in some of the largest supercomputers in the World, with experiments, for instance in plasma-based accelerators [LP7,LP8,LP9], intense radiation sources [LP10], or exploring novel ideas such as laser amplification schemes using plasmas [LP11], or astrophysical scenarios [LP12].

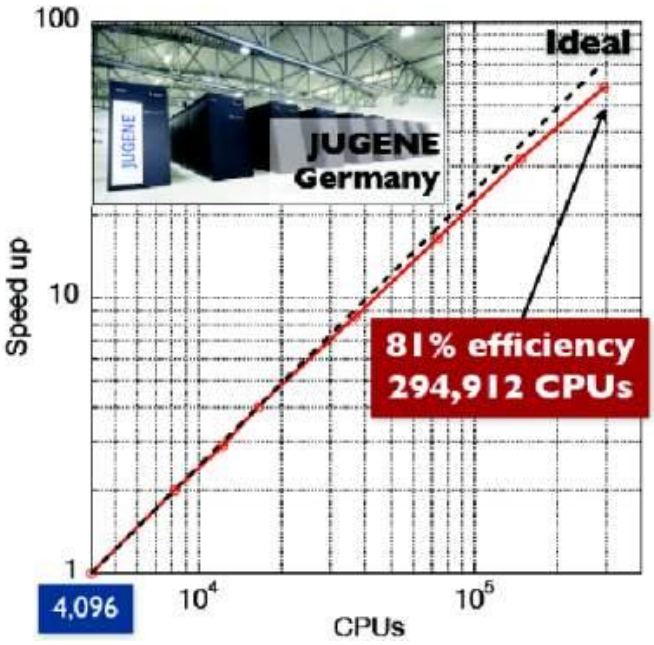

Figure 13: Strong scaling of OSIRIS for a typical production run on the largest civilian supercomputer in the World, obtained in the Extreme Scaling Workshop (Jülich, November 2009) demonstrating unmatched strong scaling in a PIC code.

3.9. Cross-Cutting Issues

3.9.1. Multiscale Simulations

One of the main challenges in physics and chemistry is the accurate description of complex systems on large time and spatial scales. The most accurate description would rely purely on quantum ab initio methods, but due to the computational complexity, these methods are usually restricted to a small number of atoms and short time scales. For instance, considering the dynamical evolution of polymer systems, consisting of chain-molecules of several thousands of monomers, where the relaxation time increases quadratically with the number of monomers, would lead to numerically unsolvable problems, even on exascale machines. Therefore both the number of degrees of freedom as well as the time scale involved, i.e., the number of time-integration steps, are intractable for most accurate methods. Interesting phenomena like self-aggregation, structure formation, charge transport through porous media or solvation of large molecules (to name a few) would be excluded due to the complex time evolution and system size. In fact, self aggregation phenomena occur on time scales of ?s-secs, which is a factor of $O(10^9\text{-}10^{15})$ longer than the time to resolve intra-molecular bond-vibrations in a simulation. Therefore the application of computationally expensive methods, like ab initio methods, is not applicable. In addition, these methods usually do not show a strong scaling behavior up to $>10^4$ compute cores, so that even exascale machines cannot solve this problem.

To overcome these limitations the usual way is to coarse grain the description of the systems, e.g. neglecting electrons, polarization effects, detailed atomistic properties or in making the transition from

an atomistic description to a field description. Traditionally these levels of description were used in their own fields, e.g. a force field description in molecular dynamics or continuum field descriptions in fluid dynamics. However, it is also possible to combine these levels of descriptions in different multiscale descriptions in a simulation setup.

The one way, which is *vertical multiscale*, takes the approach to simulate a small part of the system or single system components on a very detailed level, in order to derive parameters, like charge distributions in a molecule, bond distances, dipole moments, which then enter into a coarser description of the system, e.g. to derive force field parameters. Having extracted these parameters makes it possible to calculate on a coarser level a larger portion of the system, from where e.g. dielectric constants or transport properites might be derived, which enter into a next coarser level of description. This hierarchical way may be continued up to a coarsest level. Consequently this approach provides a consistent description of a system with parameters derived through a cascade of coarsening steps.

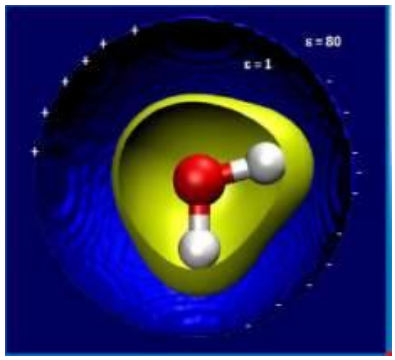

Figure 14: Schematic for an implicit solvent model

The other approach, which is *horizontal multiscale*, considers different levels of description simultaneously in the same simulation. Usually these simulations follow a top-down approach, simulating on a coarse level, but identifying regions where it is necessary to simulate on a finer level for a proper description. This approach is followed for example in the simulation of crack propagation in solids, where the bulk system might be described on a force-field level, but the tip of the crack, where energy hot-spots occur and molecular or covalent bonds are broken, a more accurate description on the *ab initio* level is chosen.

These descriptions will be of particular interest for exascale machines in future. From a physical challenge aspect, this provides the methodology to consistently describe complex materials and dynamical processes in samples. In connection with the ab initio design of new materials, this gives the opportunity to simulate larger samples or even devices for which an experimental setup would be both more time and cost consuming.

Exascale opportunities for the *vertical multiscale* approach are based on the fact that systems may run independently on a subset of processors of a machine in order to build a data base for simulations at different thermodynamic state-points. Since the parameters, derived for a coarse grain description of a system are in general not transferrable between different thermodynamic state points, modeling of force-field parameters has to be done under various conditions, e.g. variation of temperature, pressure, chemical potential. Since these simulations are independent, the generation of such data-bases can be considered as embarrassingly parallel. With exascale machines the necessary compute capacity would be established to systematically setup data bases for (i) properties of materials under various conditions and (ii) force field parameters for these materials. Exploring this systematically will enable to screen materials with desired properties and to test them in multiscale simulations under various conditions. This approach is not yet very much developed in the physics or chemistry community. Data base research and high throughput computing is more established e.g. in sequence analysis in biophysics. Therefore, support actions would be necessary to establish publicly accessible data bases from which the scientific community could strongly benefit in complex system and device modeling.

Horizontal multiscale approaches can be considered on various levels. E.g. in solvation models the solvent molecules around a solute are often reduced to a continuum description, therefore entering into the calculation of molecular properties as a parameter description or as a known function and therefore dramatically reduce the number of degrees of freedom in a calculation. This might be a valid approach, also for exascale computing, if the solute itself has complicated pathways for its structures or active centers, which require large amount of compute resources. On a larger length scale, atomistic and particle or continuum descriptions of a coarser system level are coupled, where continuum is e.g. modeled by finite elements (FE) or Lattice Boltzmann (LB) and particle based

approaches by e.g. dissipative particle dynamics (DPD). This approach has the computational advantage that both methods, atomistic and finite element calculations do scale on parallel machines. An important issue of this multiscale approach is to balance the load between the processes running on different partitions on the machine where simultaneously different methods are executed. A technical problem concerns the proper coupling between the various levels of description. It is not yet clear, which combinations of properties can be conserved when doing the transition from a fine to coarse level (and back). E.g. short wavelength phonons will be reflected at the boundary when the coarser level only supports long wavelength phonon propagation. Therefore, systematical research should be supported to establish a consistent description of a multiscale coupling framework.

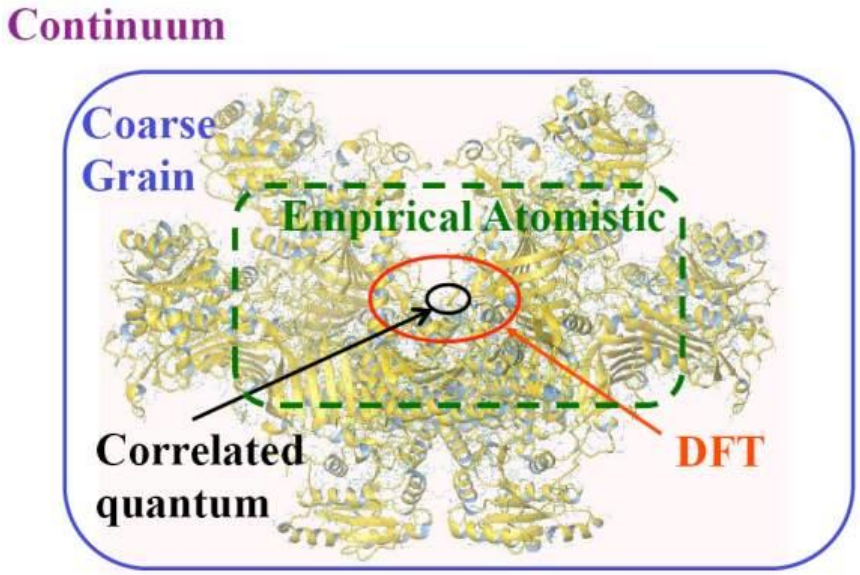

Figure 15: Schematic for a multiscale simulation scenario with high level quantum ab initio calculations, like correlated electron computations, DFT, force-field calculations and coarse grained solvent simulations

In the atomistic region this scheme can then be combined with QM/MM approaches, where active regions can be resolved on a finer time- and length-scale. Current approaches use e.g. the ONIOM model, or the Learn-on-the-Fly approach. The latter one is of particular interest, since it reduces strongly the explicit calculation of quantum mechanics part by re-fitting classical potentials to the particles in the subsystem of interest during the simulation and therefore moves a large amount of work into the molecular mechanics part, which shows good scaling properties on parallel machines. However, these approaches still have to be improved in terms of (i) understanding and correcting the error propagation away from the QM/MM region, (ii) correcting discontinuous properties over the QM/MM boundary, which might be reduced by introducing transition regions, (iii) solving the problem of automatic detection of interesting subsystems which have to described on a finer level, i.e. removing the user interaction to define, where the *hot spots* in a simulation are located.

Multiscale simulations as described so far are of particular interest for a large part of the scientific community with respect to exascale, since it is a pathway to use ab initio legacy codes which have a long history of development, including several 100 man-years of work for various codes. Therefore, rewriting and optimizing such codes would (i) take a tremendous effort in terms of time and money and (ii) would also imply IP conflicts for various commercial code packages. Therefore, the practical use of vertical multiscale simulations would prescribe on the QM level the use of an ensemble like approach, where a number of simulations run simultaneously on different partitions of an exascale machine. At present it is not clear how the architecture of such a system will look like, but assuming a concurrency of $O(10^3\text{-}10^4)$ cores on each node, it seems to be most likely that within a vertical multiscale scheme QM codes might run on single nodes, where a large set of nodes can be used simultaneously in order to calculate different properties or material compositions in order to fill a data base for coarser level

simulations. Therefore work could be spent into such codes to move from MPI parallelization to thread based parallelization on a shared memory subsystem. On the other hand MM or FE simulations show a better scaling behaviour and may use multi-node partitions on exascale machines. Usually MM codes do not have complexity such as QM codes and to use exascale machines most efficiently it is most likely that these codes should be re-designed in terms of optimized modules for inter-node communication and local and energy efficient algorithms.

In the other scenario of horizontal multiscale QM, MM and FE run simultaneously with a different work-load, depending on the number and complexity of hot spots in the simulation. This induces a critical demand on good load-balancing schemes in order to use a compute partition most efficiently. Assuming that the QM part is calculated by one of the legacy codes, a precondition is that the system software supports more than one executable in a submitted job, since MM and FE should be considered as seperate executables and not necessarily integrated into a library. For this it is necessary to develop a multiscale simulation framework, where e.g. any electronic structure code, classical MD, LB, DPD or continuum FE code might be coupled together. This framework (i) has to manage the software level, i.e. executing different codes simultaneously in a coupled way, therefore also providing an interface for data transfer between the codes and (ii) providing an automated execution of the physical problem, i.e. setting up all the different multiscale regions by identifying energetic hot spots or transition events and bulk like properties.

3.9.2.Fault Tolerance and Energy Efficiency Algorithms

Molecular simulations have a growing impact on scientific developments, also due to the exponential increase of the computational resources we experienced in the past 20 years.

In the past 10 years, the driver of this exponential increase was the advent of massive parallelism. Most of the effort in the community focussed on re-engineering algorithms and methods to extend their scalability and the efficiency, with the aim of maximizing the exploitation of the computational power (FLOPS) available. This is reflected by the numerous Gordon Bell award winners [FT1] associated with atomistic and molecular simulations since 2000. On the other hand, massive parallelism spontaneously leads to a higher probability of failure and energy consumption, assuming that the individual components are characterized by both constant failure rates and constant energy efficiency. So far, these issues have not turned into a show stopper for any of the numerous computational fields.

However, this is likely to change dramatically in the next five years as we move to exaflop machines, and both energy and fault tolerance will become fundamental issues. In fact, even when assuming the currently lowest failure rates (0.02 failures/TF/month on IBM Blue Gene/P), software running on a hypothetical exaflop architecture will have to handle approximately 1 failure per minute of elapsed time. Extrapolating from the most energy-efficient architecture available today (IBM Blue Gene/Q prototype, from the Green500 list [FT2]), an exa-flop architecture will need ~0.6 Gwatt (!!) to run current “optimized” algorithms and methods.

Note that, even with petascale supercomputers, atomistic simulations still are far from being able to model the real world in terms of both complexity and reachable time scales. Therefore is not an option to simply avoid the issues of fault tolerance and energy efficiency by limiting the use of such simulations to a (sub)-petascale capacity regime. It is necessary to dedicate time and attention to these two big fundamental issues today, in order to be able to exploit the full capabilities of the coming exascale supercomputers and to increase both the reliability and the impact of atomistic and molecular simulations.

Electricity costs impose an increasing strain on the budget of data and computing centers and are bound to limit – at all scales – the cost effectiveness and applicability of simulations. Moreover, energy dissipation causes thermal problems. Most of the energy consumed by a system today is converted into heat, resulting in wear and reduced reliability of hardware components. For these reasons, energy has become a leading design constraint for computing devices in recent years, and its importance will even grow in the years to come. Hardware engineers and system designers explore new directions to reduce the energy consumption of computing systems. On the other hand, energy consumption is already being dominated – and will be even more so – by data movement (from and to the memory,

on the interconnect and I/O subsystems) and not by the execution of operations [FT3]. As algorithms and their implementation dictate the computational intensity, the data locality and the need for inter-processor communication, it becomes clear that energy efficiency will be not only a system or library issue, but will demand specific work at the algorithm/application level.

Although these issues are incredibly topical, only very few groups have recognized their importance and have started to focused on developing algorithms to tackle the issues of fault tolerance and energy efficiency. Note that the two issues are tightly connected, and that the fault-tolerance characteristics of an algorithm directly determine its energy usage.

So far few algorithms have been developed to handle faults in machine hardware, such as, for example, the implementation of the multiple walkers in Metadynamics [FT4], or parallel schemes in Monte-Carlo sampling [FT5] as well as finite differences schemes [FT6]. All these implementations rely on requiring information only from the implementation's own memory (independent task), delegating the communication between the different tasks to socket-like structures.

Another important way of handling fault tolerance and improving the energy efficiency is the use of iterative rather than direct algorithms, because in case of hardware failures iterative algorithms can handle faults more easily. At the same time, exploiting data locality will become an even more important issue. Examples along these lines can be found in [FT8], where a low-complexity iterative solver is proposed to solve dense linear systems such as those we find in electronic-structure methods, or in [FT9], where a novel orthogonalization procedure with enhanced data locality is proposed.

These represent only a couple of examples, and is clear that similar work will be needed to analyze and re-engineer the numerous computational "dwarfs" needed in atomistic simulations having fault tolerance and energy efficiency as guiding principles.

To conclude, the issue of energy-efficient and fault-tolerant algorithms is definitely a topic that scientists became aware of only very recently [FT4-FT10], and the development of energy-efficient algorithms in the field of molecular simulations it is something the entire community should devote a significant effort to, in order to enable relevant applications for exascale computing.

4. Societal Aspects

4.1. Needs for Education and Training

The transition from peta- to exascale will present the scientific community with various problems and will be very demanding in terms of new programming schemes, efficient mapping of algorithms to hardware, and knowledge of memory hierarchies and data access over different levels of hardware implementations. In order to support the scientific communities with porting and optimization or re-design of simulation codes for a specific exascale machine, there is a large need for a generation of computational scientists who have both a deep knowledge in scientific disciplines, in software design and hardware architectures. In order to keep the scientific community on track, there is consequently the need for

- Master's level and PhD programs
- training of scientists at the postdoctoral or senior scientist level.s

In Europe, traditionally the education of scientists focuses on disciplines like physics, chemistry or computer sciences, with specializations within the disciplines at the master's or PhD level. There are ongoing developments to establish cross-disciplinary educational programs in scientific computing. Examples for this are the *Master of Philosophy in Scientific Computing* in Cambridge [ET1], the *Scientific Computation Masters* in Nottingham [ET2], and the *Master Scientific Computing* in Utrecht [ET3], which are consecutive programs, or the *Master of Science in Simulation Sciences* at Aachen [ET4], which is a non-consecutive program.

These existing master's programs (among others) are certainly a very good step forward in linking the scientific disciplines to the computational science and high performance computing community. However, the transition to the exascale era demands higher-level education of computational scientists who must be equipped with knowledge in scientific disciplines, software design and engineering, parallel languages, performance analysis and optimization tools, as well as hardware architecture.

The number of universities which currently offer master's programs in scientific computing is, however, still small and several European countries still do not offer such programs. In order to stimulate the development of interdisciplinary master's programs in a larger number of European countries, support of Erasmus Mundus-like programs could be a step forward. In this model, an educational program is distributed over different European universities, which may provide courses on a specific topic at a high level; In this way, the number of countries offering scientific computing master's programs could be enlarged without necessarily offering a complete curriculum in every country. Interdisciplinary studies could profit from master's programs organized along these line, in that specific topics could be delegated to highly specialized groups in Europe, which would form a network of *educational hot spots*.

Furthermore, in analogy to co-design centers, which aim to develop simulation codes in close cooperation between hardware vendors and system architects, system software developers and scientists, one possibility would be to extend this concept to co-education centers, which aim to provide profound knowledge in the various disciplines of hardware, software and parallel computing. This type of educational program would probably be organized as a non-consecutive program and could be an additional (e.g., one-year) program after a master's program in scientific computing. This could provide a perfect basis to learn advanced programming.

In addition to educational programs for students on the master's level, there is an additional strong need for training of postdoctoral and senior scientists, who have to follow the trends and developments of hardware architectures, software and optimization and tuning of codes. This is especially important if concepts like energy efficiency, data locality and fault tolerance become more and more important in the near future. Therefore, training programs have to be established for scientists coming from various disciplines who aim to distribute specific knowledge in porting,

optimizing and re-designing simulation codes. Such training programs could be established at European Tier-0 centers that provide access to the highest class parallel computers in combination with experts in hardware and system software. A combined effort of HPC centers, hardware vendors and software developers would be preferred for the most efficient training of researchers for the transition to exascale computing.

4.2. Costs and Efforts

In order to judge about potential costs which are related to exa-scale computing in Europe one may consider some preliminary estimates. On the one hand costs are understood as investments in the case of developing an exa-scale program in Europe, whereas on the other hand we can consider the likely impact of not investing, and estimate the losses in competence and competitiveness in the case of abandoning the exa-scale efforts.

4.2.1. Costs and efforts estimate – Software and Co-design centers

As outlined below, the direct financial effort to move fundamental sciences to exa-scale computing is related to adjusting, re-writing and re-designing simulation software.

For codes, which will be designed to run at full scale, a complete re-design is foreseen in order to exploit all capabilities of exa-scale machines. Depending on the functionality and complexity of codes, there may be large differences in time-scales for a re-design. A very rough estimate would be given by estimating about ten man years for an existing code (although there are e.g. quantum chemistry codes which certainly are more demanding). Taking into account that there is not only a straight forward re-implementation of existing software but research effort involved, which is closely coupled to development in system software, communication libraries and hardware, the time will be certainly longer. As outlined above, domain specific co-design centers could strongly accelerate and optimize the development process in view of simulation code but also in optimizing system and communication software, as well as the layout for hardware. This way of co-design is certainly most promising in a stage when architectural details for hardware components for exa-scale are specified and first prototypes exist, so that there can be a cooperative work between software and hardware developers. Current developments, pointing into this direction are e.g. (i) the *Cray Exascale Research Initiative*[CE3], involving Edinburgh Exascale Technology Center [CE1] and the Swiss National Supercomputer Center [CE2] CSCS including the HP2C initiative, (ii) the Exascale Innovation Centre[CE4], a cooperation of IBM and FZJ at Jülich (iii) the Intel Exascale Cluster Laboratory, a cooperation of Intel, ParTec and FZJ at Jülich[CE4,CE5] or Intel Exascale Computing Research Laboratory at CEA, Genci and UVSQ [CE6].

A cost estimate could be given on a basis of current agreements between HPC centers and hardware vendors, which prescribe a bilateral investment of about 5 FTE's from each side (i.e. about 10 FTE's) per year [CE7]. Establishing co-design centers for different domains, e.g. plasma physics, fusion research, astrophysics, high energy physics, material sciences (including hard and soft materials), condensed matter physics or quantum chemistry, would imply an annual investment of about 70 FTE's.

For the case of Simulation Laboratories, as outlined above, the current structure involves about three senior scientists per scientific domain plus two technical staff members and PhD students. Since the current implementation is based on a mainly national user community, it is expected that SimLab's have to be enlarged or to be distributed over different Tier-0 and Tier-1 centers. A more realistic estimate would therefore consider about 8 scientists per SimLab plus technical staff members, which would also sum up to about 70 FTE's per year.

This combination of co-design centers and SimLab's would allow for the development and optimization of simulation software for strong and weak scaling applications on full scale on exa-scale machines as well as moving the broader scientific community to peta-scale computing.

Resources	Human Resources by year → 2015	Integrated (4 years) Provisional Costs 2012 → 2015	Human Resources by year 2016 → 2020	Integrated (4 years) Provisional Cost 2016 → 2020
Co Design Center	70	28 000	70	35 000
Simulation Labs	70	28 000	70	35 000
Total resources	140		140	
TOTAL in k€		**56 000**		**70 000**

Human resources are given is given in Man * Year.

Provisional costs are given in K€ with a yearly flat rate of 100k€/FTE

4.2.2. The case of not having exascale in Europe

If Europe did not establish its own exascale program, the compute resources for cutting-edge computational science would be located most probably in the U.S., China, or Japan. The implications of this include losing leadership in various fields of research, such as the new promising fields of computational materials design and device design. This would (i) lead to the migration of leading experts abroad; (ii) the probable location of spinoff industrial firms and new jobs to locations outside of Europe. The same is true for plasma physics, high energy physics, astrophysics and quantum chemistry, where Europe has an outstanding position in scalable codes and expertise. Not investing in exascale software development in these fields risks losing competitiveness in areas of significant importance for the scientific and industrial competitiveness of Europe.

In addition, there will be a risk of losing young researchers to the U.S., Japan or emerging countries like China. In recent years, for example, approximately 100 PhD theses per year were connected to simulations performed within projects on the IBM Blue Gene/P system JUGENE in Jülich. Losing one of the world's leading positions in HPC computing, Europe would risk a migration of researchers abroad. in the number of FTE's lost per year would be easy to estimate, but the damage for Europe's competitiveness over the next several decades might be inestimable.

4.3. Societal and Scientific Community Aspects

4.3.1. Support Actions for the Scientific Community

In close relation to the educational actions which have to launched in order to support and develop the next generation of computational scientists, actions have to be planned to support the scientific community's ability to run efficiently codes on exascale architectures. As outlined above, there have to be various approaches to use exascale machines, given the fact that there is a large diversity of currently existing community codes and also the need to develop a new software stack for highly scalable scientific applications.

Developing a complete set of efficient codes for petascale architectures is equivalent to at least several hundred person-years of effort and might therefore be infeasible.. However, it is possible to analyze all those codes in some detail in terms of requirements for processors, communications, memory access, data structures, algorithmic efficiency, and other attributes. In other words, it is entirely feasible to construct a landscape of requirements for successfully porting existing codes to exascale machines,and from this to define community- and code-specific priorities.

The second class of applications exhibit strong or weak scaling up tofull scale. Since it might be anticipated that architectural details of the exascale machines will change dramatically in terms of memory size per core, memory hierarchies, CPU and accelerator design, this class of applications needs a major re-write and re-design in terms of data structures, algorithms, memory layout, communication and task parallelism.

For both classes, there is a strong need for user and software developer support. For porting, optimizing and adjusting existing software, special support actions, such as bilateral cooperation between HPC centers and developers or establishment of simulation laboratories, is needed.

An existing example for the first support action is the *Swiss Platform for High Performance and High Productivity* (HP2C), where, at a computer center, a scientific core group is established which consists mainly of computational mathematicians and computer scientists, who work closely together with application developers and domain scientists, and who possess the necessary know-how in hardware, development environments and tools. Furthermore, financial support is given to application developers, who are directly embedded into the domain science research groups. In addition to the direct support of software development, this initiative intends to establish a network of research groups who will benefit from knowledge exchange and distribution and scientific cooperations within the network.

An example for community-oriented *Simulation Laboratories* (SimLabs) is the model followed by Forschungszentrum Jülich [S4], Aachen University [S5] and Karlsruhe Institute of Technology [S6]. Though similar to HP2C, the philosophy is to establish core groups of different scientific domains at the HPC centers. These groups work in close cooperation with national or international research groups and provide support in terms of code development, porting and optimization, as well as algorithm development for specific tasks. In this model, it is required that the research groups contribute with in-kind matching funds, such that one member of the research group (e.g., a PhD student) forms a link between the HPC center and the group and invests time equivalent to that invested by members of the core group at the HPC center.

Both types of support actions are intended to support individual research groups, as well as to support existing scientific communities or strengthening scientific networks. These types of support actions are not only needed for applications running at full scale on exascale machines, but also to support groups with less ambitious applications. Considering the fact that in the exa-scale era, computations at petascale have to be the *standard case*, there will be an increasing efficiency gap between full scalable applications and those which can use only a small portion of HPC platforms, if no specific support actions are undertaken from now on. In perspective, computational capacities currently available on Tier-0 platforms will be moved to Tier-1 and Tier-2 platforms in future. If the scientific community is not investing enough effort in further development and improvement of current simulation codes, future Tier-1 and Tier-2 architectures will only be used below their potential capabilities. As noted earlier in this report, leadership in HPC and in the scientific and research fields that depend on HPC will be determined much more by software advances than by hardware progress.To ensure strong returns-on-investment from Europe's hardware systems, support actions driven by HPC centers (e.g., Tier-0 and Tier-1) should not only address the most scalable parallel codes but also less scalable codes that may be of equal or even great scientific and economic importance..

4.3.2. Role of Community Organizations

Traditionally, community organizations play an important role in forming networks between scientific groups, disseminating knowledge within specific scientific domains, and identifying demanding and grand challenge problems in the domains. Organizations like the *Centre Européen de Calcul Atomique et Moléculaire* (CECAM) or the network *Psi-k* organize workshops, tutorials, conferences and visitor programs. Software activities start to emerge, e.g., the scientific software development group at CECAM [S1], but community organizations still have a stronger focus on disseminating information on community codes (see e.g. [S2]). A direct link to software development for domain specific problems is still exceptional. Positive examples of software-oriented organizations are the *Collaborative Computational Projects* (CCP, [S3]) in the UK. These include twelve projects. among which are *Electronic Structure of Molecules* (CCP1), *Continuum States of Atoms and Molecules* (CCP2), *Computational Studies of Surfaces* (CCP3), *Computer Simulation of Condensed Phases* (CCP5), *Molecular Quantum Dynamics* (CCP6), *Electronic Structure of Solids* (CCP9) and *Computational Plasma Physics* (CCPP). Flagship codes are developed within these projects, and software is both

maintained and distributed. In addition to this, schools are organized to train users to apply the codes. In order to identify the needs for future code development and to channel the activities, each project is led by a chair and assisted by a working group. To coordinate the activities of all the CCP's, a steering panel, including all project chairs, was formed.

This type of community organization could be established on a European level in order to strengthen and to focus initiatives for exascale software development. Existing community organizations, like CECAM, could lead such initiatives on the scientific level. In view of the fact that software development for leadership computer architectures will need a large effort, scientific communities should be strengthened by being linked to community-specific co-design centers, where future software is developed in close contact between the scientific communities, hardware and system software developers.

4.4. European Competitiveness, Strengths and Weaknesses

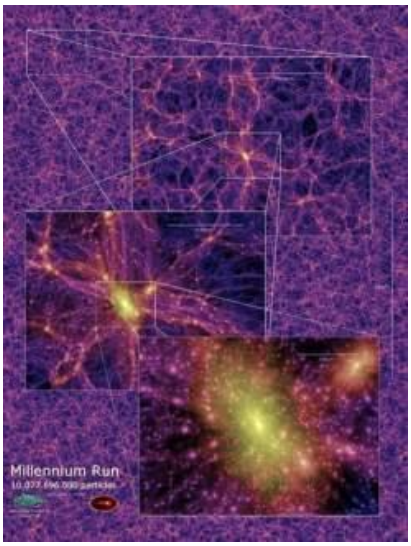

Figure 16: Different detail views from the Millennium simulation with more than 10 billion particles.

European fundamental sciences have a high reputation worldwide and are definitely competitive, if not always leading, in scientific domains such as high energy physics, plasma and fusion research, astrophysics, cosmology, material sciences, quantum ab initio methods and soft matter physics. Especially the field of computational material sciences and quantum ab initio chemistry has a long history in Europe which is reflected in a well-organized community (e.g., organizations like CECAM, Psi-k or CCP). This organizational structure strongly supports the efficient dissemination of information and software across European as well as non-European groups. It is appropriate to state that Europe has certainly the strongest track worldwide in ab initio codes. Amongst others there are codes like CPMD[EC1], CP2K[EC2], specially suited for Car Parrinello ab initio MD simulations, VASP[EC10] for ab initio MD with an exact evaluation of the instantaneous electronic ground state at each MD-step, DFT plane wave codes like Turbomole[EC3], QuantumEspresso[EC4], CASTEP[EC5], ONETEP[EC6], FLEUR[EC7], abinit[EC8], ADF[EC9], electronic structure codes using wavelet basis sets like BigDFT[EC13], DFT methods based on real-space uniform grids and multigrid methods or atom-centered basis-functions like GPAW[EC14] or highly accurate quantum chemistry packages for molecular electronic structure calculations based on multiconfiguration-reference CI, coupled cluster and associated methods like Molpro[EC12], Columbus[EC11], to name a few. Some of the codes, like CPMD, CP2K, BigDFT, VASP or GPAW, are efficiently parallelized and currently run well on petascale architectures.

Parallel codes for classical MD simulations in Material Sciences, Statistical Physics or Soft Matter Research include DL_Poly[EC23], DL_Meso[EC24], Gromacs[EC25], Moldy[EC41], IMD[EC37] or ESPResSo[EC26].

In other domains, such as astrophysics and cosmology, there is as well a variety of high performance parallel codes developed in Europe, like GADGET[EC15] for cosmological N-body/SPH simulations, RAMSES[EC16] for self-gravitating magnitized fluids, AREPO[EC17] based on a moving unstructured mesh, PLUTO[EC18] for astrophysical and high Mach number flows for Newtonian, relativistic, MHD or relativistic MHD fluids, the tree-code PKDGRAV[EC21] or NBody[EC20] a particle based numerical simulation code for gravitational systems. These codes were used in world leading, large scale astrophysics simulations, like the Hubble simulation[EC18], where the formation of clusters of galaxies was examined, the Horizon simulation[EC20] or Millennium simulation[EC19] for understanding the structure formation in the early universe as well as dark matter simulations, e.g., the Aquarius project[EC22].

Currently it seems that Europe is still in a world- leading position with ab initio and astrophysical codes. The underlying software packages in these domains are rather complex and often have a long history of development. This implies a sustainability in code development, maintenance and porting to new architectures. Compared to the U.S., for example, Europe has been in a better situation regarding permanent positions for scientists and technical staff members. This has allowed European organizations to invest in projects which need long-term development and guidance within research groups. Therefore, codes like quantum ab initio codes, which often contain hundreds of man-years in programming and development, were only possible to realize in Europe. The U.S. has made up for this comparative deficiency by excelling in complex application software and, nowadays, also highly efficient and parallel community codes such as NWChem[EC27] or GAMESS[EC28] in quantum chemistry, LAMMPS[EC33], SPaSM[EC34] in classical particle simulations or ENZO[EC29], GASOLINE[EC30] or FLASH[EC31,EC32] in astrophysics and cosmology.

Fusion for energy is also a domain where Europe is very active. The largest codes in magnetic fusion research solve the gyrokinetic (GK) equation in a global (full torus) domain. The leading European GK global codes employed for turbulence simulations are GYSELA [EC38], which solves the p.d.e with a semi-Lagrangian method, and ORB5 [EC39], a PIC code. GENE [EC40] is the leading local (the domain is a portion of the torus) GK code, based on the Eulerian approach with explicit time advancing. Other codes are ELMFIRE[EC44] (PIC), EUTERPE[EC45], GKW[EC46]. A variety of codes solving fluid models is also employed. Noteworthy fluid codes are those solving the MHD model in realistic tokamak geometry. Of these, JOREK[EC42] and XTOR[EC43] are the European leaders.

Lattice field theorists organize in collaborations, which are centered around common scientific objectives, common theoretical formulations, common codes and/or common supercomputers, though the latter is becoming more rare because of the size and cost of the computers required. Some European scientists belong to more than one collaboration.

The main European collaborations in Particle, Hadron and Nuclear Physics are:

- The Alpha collaboration [EC47], which regroups approximately 30 scientists (including students and postdocs) from 7 European countries and Brazil. Different subgroups focus on different topics in the fields of “Searching for BSM physics” and “Testing QCD at the sub-percent level”. There is a fair amount of overlap between the members of CLS (see below) and Alpha.
- The Budapest-Marseille-Wuppertal collaboration [EC48], which regroups about 15 scientists (including students and postdocs) from the 3 European countries implied by the collaboration's name. The fields of focus are mainly “Searching for BSM physics” and “Testing QCD at the sub-percent level”. A subgroup with additional scientists in Budapest and the University of California San Diego have produced work in the field “Simulating possible theories of BSM physics”. Another subgroup with additional scientists in Budapest also works in fields “Precision calculation of bulk thermodynamics of strongly interacting matter” and “QCD phase structure and nonzero net baryon number density”.
- The Coordinated Lattice Simulations (CLS) collaboration [EC49], which is a community effort, launched in 2007, whose aim is to bring together the human and computer resources of several teams in Europe interested in lattice QCD. Currently the CLS regroups about 30 scientists (including students and postdocs) from 4 European countries. Codes and gauge field configurations are shared.
- The QCD Structure Function (QCDSF) collaboration, which has approximately 20 scientists (including students and postdocs) mainly from Germany and the UK. Codes, computing resources, gauge field configurations and projects are shared.
- The HotQCD Collaboration, which has over 20 scientists (including students and postdocs) from the US and from Germany. It shares codes, computer resources and gauge field configurations.
- The United Kingdom QCD (UKQCD) collaboration [EC50] is a British collaboration whose purpose is to procure and jointly exploit computing facilities for lattice field theory calculations. It regroups over 40 scientists from 8 British universities. It has close ties with the US Riken-Brookhaven-Columbia (RBC) [EC51] collaboration, with which it shares supercomputers, codes, gauge configurations and projects. It also has members in common with the High Precision QCD

(HPQCD) collaboration which includes US scientists too. It further collaborates with the QCDSF collaboration discussed above.

- The “Nuclear Physics from Effective Field Theory” (NPEFT) collaboration is newly emerging with scientists mostly from Europe and the US that perform nuclear lattice simulations. It intends to make its gauge configurations publicly available.

As attested by the results and publication records of these collaborations and by the many invitations of their members to give plenary talks at international conferences in particle and nuclear physics, Europe is a world leader in Lattice Field Theory. Its areas of strength are all of the fields discussed above, except for “From QCD to nuclei” where the US and Japan are playing a leading role.

Considering weak points of European software development, this is related to multiscale modeling and frameworks for multiscale simulations. Although there is a trend in Europe to strongly develop this area of research, it has already been a long-term focus of U.S. research. In addition to this, a general difference to other countries is that in Europe the value of simulation has not had such strong recognition, and there is tendency to consider simulations as less important than observations or experiments. This is different in countries such as the USA, where a very strong driving force for the development of simulation infrastructure (and hardware) was (and still is) the ASC program [EC35,EC36]. Through this program, supercomputing has a direct link to national security and therefore a very high significance.

A very recent initiative, driven by the U.S. Army Research Laboratory (ARL) is establishing two Collaborative Research Alliances (CRA) as part of the Enterprise for Multiscale Materials (EMM), which aims to develop multiscale modeling and simulations in materials sciences and to bring together government, industrial, and academic institutions to address some of the fundamental scientific and technological questions.

Programs like ASC and initiatives driven by ARL have certainly improved the employment situation for scientific software developers and the competitiveness of the U.S. This tendency can be further observed in the new U.S. initiatives (including Exascale software center and co-design centers), in which large centers are planned where software development has a high significance. Therefore, a turning point for Europe in terms of scientific application software can be foreseen, if no support actions are launched.

4.5. Potential and on-going non-European cooperations

Similar to the idea of the PRACE project, exascale computing in Europe should not be based on national initiatives only. In the near future, the financial investments will reach such a dimension that national initiatives will most probably not be able to keep pace for future developments and investments in order to be on the same level as countries like the U.S. or China. On the other hand, application software for exascale machines has to be planned and developed in line with hardware, e.g., in co-design centers, in order to be available in time for the availability of exascale hardware. Therefore, not only for hardware but also for domain specific software development it will be essential to have collaborations within Europe and, equally important, with non-European countries.

Within several domains, there are already established or planned collaborations for software development. For example, in astrophysics and cosmology the *Virgo Consortium*[C1] is an international initiative which has a core membership of about a dozen scientists in the UK, Germany, Netherlands, Canada, the USA and Japan. This aims to carry out state-of-the-art cosmological simulations. The research areas include the large-scale distribution of dark matter, the formation of dark matter haloes, the formation and evolution of galaxies and clusters, the physics of the intergalactic medium and the properties of intracluster gas. Codes developed within this consortium are e.g. HYDRA[C2] and GADGET[C3]. The FLASH code[C4], which will be further developed for exascale simulations in the FLASH co-design center[C5] within this consortium.

Another large-scale astrophysics international ongoing project is the Square Kilometer Array project (SKA,[C6]) with headquarters in Jodrell Bank observatory in Chesire, UK. This project aims to develop a radio telescope which will have a total collecting area of approximately one square kilometer. It will operate over a wide range of frequencies and will require very high performance central computing and long distance links with a capacity greater than current worldwide internet traffic. This very demanding and ambitious project, which includes the development of hardware and software, is based on a cooperative effort of 20 European and non-European countries, amongst which are Australia, Brazil, Canada, China, India, Japan, Korea, New Zealand, Russia, South Africa and the U.S.

In many other domains there are international cooperations, based on individual contacts between research groups and often initiatiated via individual contacts to university research centers, e.g., for Materials Science Boston/Oxford/EPFL Lausanne[Marzari], Oak Ridge/CSCS Manno [Schulthess], Dublin/Boston [Coker] or Multiscale Modeling Harvard/EPFL Lausanne [Kaxiras], to name a few. Other cooperations are based on co-development efforts of community codes, like GADGET or AREPO codes (cosmology - Heidelberg/Germany and Harvard University/US), ENZO, FLASH, PKDGRAV and GASOLINE codes (astrophysics - Zürich/Swiss, Seattle/US and MacMaster/Canada) or the CASTRO code (cosmology – Göttingen/Germany and LBNL Berkeley/US). Future efforts are planned for the RAMSES code (cosmology – Zürich/Swiss, Saclay/France and Berkeley/US).

The emerging field of multiscale simulations is more advanced at present in the US, but Europe is very active in developing mathematical formalism and algorithms to couple different levels of description together into one simulation[C7]. The field of QM/MM[C10] is already well represented in Europe, as is the development of particle-continuum descriptions[C9] and atomistic/coarse-grain descriptions [C8]. This development is supported by a common initiative between the U.S. and Europe which since 2002 has organized the biennial MMM conference series (Multiscale Materials Modeling,[C11]). This has been held in London, Los Angeles, Freiburg and Tallahassee with a growing number of participants.

Other worldwide collaborations include the nuclear fusion community, which is centered around the ITER facility[C12]. As part of the so-called “Broad Approach” agreement between Europe and Japan, the HPC centre IFERC (International Fusion Energy Research Centre) will be built in Japan and it will be dedicated to large scale simulations of fusion plasmas. At present, a computer system which is very similar to the planned one is installed at Forschungszentrum Jülich, Germany and provides compute resources to the fusion community. The ITER project has already established international collborations within the fusion community, but regarding the future compute infrastructure in Japan, the compute power will strongly increase and ITER should form and support more international cooperations on computational fusion research.

Interesting developments for fundamental sciences furthermore exist at RIKEN in Japan[C14], which has earned worldwide renown for its development of special-purpose architectures, e.g., the MDGRAPE architecture[C15,C19], which was built to calculate non-bonded interactions in molecular dynamics simulations and was used to biophysics[C17], astrophysics[C16] and other statistical domains[C18].

In the U.S., there are various operational programs under NSF which aim to develop computational and software infrastructure for the fundamental sciences. For Europe this could be a potential target for joint research programs, but also software development and implementation programs. At NSF, these programs currently cover (among others): Software Infrastructure for Sustained Innovation[C20], Bilateral programs in the Materials World Network (MWM)[C21] and International Materials Institute (IMI)[C22], Multiscale Modeling and Computation[C23], Chemical Theory, Models and Computational Methods (CTMC) supporting the discovery and development of theoretical and computational methods to address challenges in chemistry, with emphasis on emerging areas of chemical research[C24], Theoretical Elementary Particle Physics[C25], Physics at the Information Frontier (PIF), addressing data-enabled science, community research networks, and new computational infrastructure as well as next-generation computing[C26] and Computational Mathematics, supporting mathematical research in areas of science where computation plays a central and essential role, emphasizing design, analysis, and implementation of numerical methods and algorithms, and symbolic methods[C27].

Russia is another country which will strongly develop national HPC resources and will enter into the supercomputing league. Funding organizations, which might be of interest for international support programs include Ministry of Education and Science[C29], Russian Foundation for Basic Research[C30], Russian Corporation in Nanotechnologies[C31], Skolkovo Foundation[C32] as well as Russian scientific organisations in the field of fundamental sciences: Moscow State University[C33], Russian Academy of Sciences[C34], Kurchatov Institute[C35].

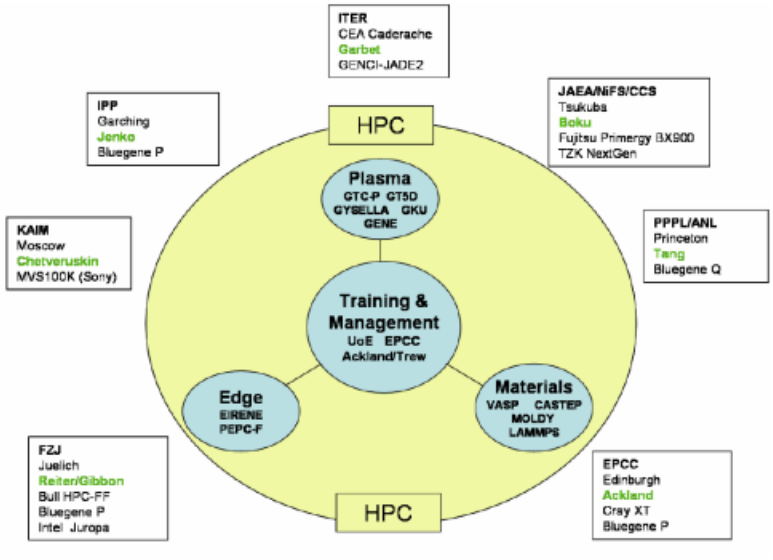

Figure 17: Schematic of the organization of the G8 project NuFuSE

There are already various EU programs launched, which support cooperation between Russia and the EU. Amongst others, there is the Initiative of the G8 Research Councils on Multilateral Research Funding, which fosters the development of application software towards exascale computing[C28]. One project (NuFuSE – Nuclear Fusion Simulations at Exa-scale), which was submitted by a consortium consisting of members from France, Japan, US, UK, Germany and Russia, aims to model and simulate key issues for nuclear fusion: the nuclear plasma, the interaction between plasma and vessel and the material of the vessel itself. This project is a very good example for coordinated efforts among the G8 countries, distributed over the world and putting efforts together to develop models and simulation tools for solving grand challenging problems which are relevant for worldwide interests, in this case. the global energy supply problem.

Certainly, there should also be collaborations established with Chinese scientific organizations and researcher groups. Having already installed some of the world's most powerful computers, China will invest money and manpower into future compute initiatives and it might be anticipated that China is also going to play a prominent, and perhaps dominant, role in exascale computing in the future. At present there are various visits from Chinese politicians and representatives of computer centers to Europe and the U.S., as well as vice versa, in order to establish closer contacts and to discuss common interests. Therefore, funding organizations should also consider including China in bilateral and international support calls.

5. Conclusions

The fundamental sciences play a crucial role not only in advancing human knowledge but in maintaining intellectual and economic competitiveness. Europe has a very strong standing in the fundamental sciences. In order to stay competitive with the U.S., China and Japan, support actions are needed for application software development. Specific software solutions are needed for scalable applicationsthat efficiently exploit the capabilities of exascale machines and thereby ensure strong returns on the very substantial investments Europe plans to make in these machines and their existing petascale predecessors. Porting and further development of existing community codes will be a feasible variant for multiscale simulation environments and parallel replica or ensemble simulations. Further development of multiscale frameworks should be enforced to stay competitive and to strengthen fields such as materials sciences and nuclear fusion research, which not only have high impact for fundamental knowledge but also for Europe's economy and society – for example by providing the necessary knowledge for designing materials for energy storage, supply and efficient usage.

To meet the requirements of code development for tackling grand challenge problems in domain specific fields, efficient algorithms and software solutions have to be developed that exhibit "awareness" of memory efficiency, memory locality, optimal order, fault tolerance and energy efficiency. Development of software should be supported by domain specific units, as already happens with the HP2C initiative or the Simulation Laboratories. Importance should also be given to co-design units or centers, similar to those existing or planned in the U.S. Current exascale laboratories, which are common initiatives between vendors and research centers of universities, could guide activities towards co-design.

Current research activities include various non-European partners, mainly in the U.S., but increasingly also in rising nations such as China. To stay competitive, Europe should strengthen international projects and should seek synergies with non-European main players in HPC or software development, including hardware specific developers (e.g. RIKEN, Japan) or system- and application-specific software developers (e.g. Exascale Software Center, U.S., or co-design centers, U.S.). China will also play a strong role in computation-based fundamental sciences in the near future, and Europe should seek for common links and projects.

Europe should develop strong and focused initiatives towards exascale computing, including specific training and education programs, which aim to set up or strengthen master's and doctoral programs that provide domain-specific scientific backgrounds and also have a strong focus on mapping scientific problems onto high performance computers, including parallel programming, algorithm layout for fault tolerance, energy efficiency and data locality, as well as an introduction to modern hardware architectures.

In summary, Europe already has world-class strengths in important areas of software development support the fundamental sciences, but in a number of areas leadership has migrated to the U.S. or other nations. Support actions are needed to maintain and advance Europe's standing in parallel software development needed by the fundamental sciences. The alternative is an almost certain loss of capabilities that are indispensable for scientific and industrial innovation.

6. Bibliography

[HN1] http://extremecomputing.labworks.org/highenergyphysics/index.stm

[HN2] S.Dürr *et al.*, "Ab-Initio Determination of Light Hadron Masses", Science **322** (2008) 1224-1227.

[HN3] M.J.Savage, "Nuclear Physics from Lattice QCD", [arXiv:1110.5943 [nucl-th]].

[HN4] http://extremecomputing.labworks.org/nuclearphysics/report.stm

[HN5] E.Epelbaum, H.Krebs, D.Lee, U.G.Meißner, "Ab initio calculation of the Hoyle state", Phys.Rev.Lett. **106** (2011) 192501.

[I1] A.V.Kozhevnikov, A.G.Eguiluz and T.C.Schulthess, *Towards first-principles electronic structure simulations of excited states and strong correlations in nano- and materials science,* Proceedings of SC2010, New Orleans, Louisiana (2010).

[I2] M.Eisenbach, C.G.Zhou, D.M.Nicholson, G.Brown, J.Larkin and T.C.Schulthess, *A scalable method for ab initio computation of free energies in nanoscale systems,* Proceedings of SC2009, Portland, Oregon (2009).

[AC1] R.Teyssier et al., Astron. Astrophys., 497, 335 (2009).

[AC2] A.Strugarek, A.S.Brun and J.P.Zahn, *Magnetic Confinement of the Solar Tachocline: II. Coupling to a convection zone, Astron. Astrophys.*; http://www.stars2.eu/A_Task_Content/TaskA2/MagConfTachoII_lowres.pdf

[AC3] V.Springel, S.D.M.White, A.Jenkins, C.S.Frenk, N.Yoshida, L.Gao, J.Navarro, R.Thacker, D.Croton, J.Helly, J.A.Peacock, S.Cole, P.Thomas, H.Couchman, A.Evrard, J.Colberg, F.Pearce, *Simulating the joint evolution of quasars, galaxies and their large-scale distribution,* Nature ***435****, 629-636 (2005).*

[AC4] http://www.mpia.de/GAIA/

[AC5] http://pan-starrs.ifa.hawaii.edu/public/

[AC6] http://bigboss.lbl.gov/

[LP1] J. M. Dawson, *Particle simulation of plasmas*, Rev. Mod. Phys. **55**, 403 (1983).

[LP2] S. F. Martins et al., *Exploring LWFA regimes for near term lasers using particle-in-cell simulations in Lorentz boosted frames*, Nat. Phys. **6**, 311 - 316 (2010)

[LP3] R. A. Fonseca et al., *OSIRIS: A Three-Dimensional, Fully Relativistic Particle in Cell Code for Modeling Plasma Based Accelerators*, Lect. Notes Comput. Sci. **2331**, 342–351 (2002); R. A. Fonseca et al., *One-to-one direct modeling of experiments and astrophysical scenarios: pushing the envelope on kinetic plasma simulations*, Plasma Phys. Control. Fusion **50**, 124034 (2008).

[LP4] http://www.extreme-light-infrastructure.eu/

[LP5] http://www.hiper-laser.org/

[LP6] http://www.fz-juelich.de/jsc/bg-ws09/

[LP7] I. Blumenfeld et al, *Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator,* Nature 445, 741-744 (2007).

[LP8] S. P. D. Mangles, et al., *Monoenergetic beams of relativistic electrons from intense laser–plasma interactions*. Nature 431, 535–538 (2004); C. G. R. Geddes, et al. *High-quality electron beams from a laser wakefield accelerator using plasma-channel guiding*. Nature 431, 538–541 (2004); J. Faure, et al., *A laser–plasma accelerator producing monoenergetic electron beams*. Nature 431, 541–544 (2004); J. Faure et al, *Controlled injection and acceleration of electrons in plasma wakefields by colliding laser pulses*, Nature 444, 05393 (2006).

[LP9] B. M. Hegelich et al, *Laser acceleration of quasi-monoenergetic MeV ion beams*, Nature **439**, 441-444 (2006).

[LP10] S. Kneip et al., *Bright spatially coherent synchrotron X-rays from a table-top source*, Nature Physics **6**, 980-983 (2010).

[LP11] R.M.G.M.Trines et al, *Simulations of efficient Raman amplification into the multipetawatt regime*, Nature Physics **7**, 87–92 (2011).

[LP12] Anatoly Spitkovsky, *Particle Acceleration in Relativistic Collisionless Shocks: Fermi Process at Last?*, Astrophysical Journal Letters **682** L5 (2008); S. F. Martins et al., *Ion dynamics and acceleration in relativistic shocks*, Astrophys. J. Lett. **695**, L189-L193 (2009).

[FT1] Gordon Bell winners, http://en.wikipedia.org/wiki/Gordon_Bell_Prize

[FT2] Green 500 list, http://www.green500.org/

[FT3] P. Kogge at al, www.exascale.org/mediawiki/images/6/6e/Sc09-exa-panelkogge.pdf

[FT4] P. Raiteri et al., J. Phys. Chem. B **110**, 3533-3539 (2006)

[FT5] C. Lindemann, et al., IEEE Trans, Reliability **44**, 694-704 (1995)

[FT6] Y.-W. Chen, K.-L.Chung, Information Processing Letters **79**, 11-16 (2001); A. Geist, 2002, http://www.csm.ornl.gov/~geist/Lyon2002-geist.pdf

[FT7] G. Fagg et al., Parallel Computing **27**, 1479-1495 (2007)

[FT8] C. Bekas and A. Curioni, Computer Science – Res. & Develop. **25**, 187-195 (2010)

[FT9] C. Bekas and A. Curioni, Computer Physics Commun. **181**, 1057-1068 (2010)

[FT10] S. Harizopoulos et al., In Proc. Fourth Biennial Confe. Innovative Data Systems Research (CIDR), Asilomar, CA, January 2009.

[EC1] http://www.cpmd.org/

[EC2] http://cp2k.berlios.de/

[EC3] http://www.turbomole.com/

[EC4] http://www.quantum-espresso.org/

[EC5] http://www.castep.org/

[EC6] http://www2.tcm.phy.cam.ac.uk/onetep/

[EC7] http://www.flapw.de/pm/index.php?n=Main.AboutFleur

[EC8] http://www.abinit.org

[EC9] http://www.scm.com/

[EC10] http://cms.mpi.univie.ac.at/vasp/

[EC11] http://www.univie.ac.at/columbus/

[EC12] http://www.molpro.net/

[EC13] http://inac.cea.fr/L_Sim/BigDFT/

[EC14] https://wiki.fysik.dtu.dk/gpaw/

[EC15] http://www.mpa-garching.mpg.de/gadget/

[EC16] http://irfu.cea.fr/Projets/Site_ramses/RAMSES.html

[EC17] http://www.mpa-garching.mpg.de/~volker/arepo/

[EC18] http://www.mpa-garching.mpg.de/Virgo/hubble.html

[EC19] http://www.mpa-garching.mpg.de/galform/millennium/

[EC20] http://www.ast.cam.ac.uk/~sverre/web/pages/nbody.htm

[EC21] http://hpcc.astro.washington.edu/faculty/trq/brandon/pkdgrav.html

[EC22] http://www.mpa-garching.mpg.de/aquarius/

[EC23] http://www.cse.scitech.ac.uk/ccg/software/DL_POLY/

[EC24] http://www.cse.scitech.ac.uk/ccg/software/DL_MESO/

[EC25] http://www.gromacs.org/

[EC26] http://espressomd.org/wiki/Main_Page

[EC27] http://www.nwchem-sw.org/index.php/Main_Page

[EC28] http://www.msg.ameslab.gov/gamess/

[EC29] http://lca.ucsd.edu/software/enzo/v1.0.1/

[EC30] J.W.Wadsley, J.Stadel and T.Quinn, New Atron. **9**, 137-158 (2003)

[EC31] http://flash.uchicago.edu/website/groups/home.py?submit=codegroup.txt

[EC32] A.Dubey, K.Antypas, M.K.Ganapathy, L.B.Reid, K.Riley, D.Sheeler, A.Siegel and K.Weide, Parall. Comp. **35** 512--522 (2009)

[EC33] S. Plimpton, J Comp Phys, **117**, 1-19 (1995); http://lammps.sandia.gov/

[EC34] D. M. Beazley and P. S. Lomdahl, Comput. Phys. **11**, 230 (1997).

[EC35] http://nnsa.energy.gov/asc

[EC36] http://www.sandia.gov/NNSA/ASC/about.html

[EC37] J. Stadler, R. Mikulla, and H.-R. Trebin, *IMD: A Software Package for Molecular Dynamics Studies on Parallel Computers*, Int. J. Mod. Phys. C **8**, 1131-1140 (1997); http://www.itap.physik.uni-stuttgart.de/~imd/index.html

[EC38] http://www-ljk.imag.fr/membres/CHANT/Documents/GRANDGIRARD_IHP.pdf

[EC39] http://crpp.epfl.ch/page-50801-en.html

[EC40] http://crpp.epfl.ch/page-50803-en.html

[EC41] K.Refson, *Moldy: a portable molecular dynamics simulation program for serial and parallel computers,* Comp. Phys. Comm. **126**, 310-329 (2000).

[EC42] G.T.A. Huysmans and O. Czarny, Nucl. Fusion **47**, 659-666 (2007).

[EC43] Lütjens H., Luciani J. F., *XTOR-2F: a fully implicit Newton-Krylov solver applied to nonlinear 3D extended MHD in tokamaks.* J. Comp. Phys. **229,** 8130-8143 (2010).

[EC44] http://www.tkk.fi/Units/AES/ELMFIRE/

[EC45] X.Sáez, A.Soba, E.Sánchez, R.Kleiber, F.Castejón and J.M.Cela, *Improvements of the particle-in-cell code EUTERPE for petascaling machines*, Comp. Phys. Comm. **182**, 2047-2051 (2011).

[EC46] http://www.gkw.org.uk/tikiwiki/tiki-index.php

[EC47] http://www-zeuthen.desy.de/alpha/

[EC48] http://www.bmw.uni-wuppertal.de/Home.html

[EC49] https://twiki.cern.ch/twiki/bin/view/CLS/WebIntro[EC50] http://ukqcd.swan.ac.uk/mediawiki/index.php

[EC51] http://rbc.phys.columbia.edu/

[CE1] http://www.epcc.ed.ac.uk/research-collaborations/activities/exascale/

[CE2] http://www.hpc-ch.org/blog/2009/12/08/cscs-and-hp2c-involved-in-cray-exascale-research-initiative/

[CE3] http://investors.cray.com/phoenix.zhtml?c=98390&p=irol-newsArticle&ID=1360843&highlight=

[CE4] InSiDE, Vol. 8, (2010); http://inside.hlrs.de/htm/Edition_02_10/article_13.html

[CE6] Exascale Computing Research Laboratory
http://www.exascale-computing.eu/
http://www-hpc.cea.fr/en/collaborations/docs/exatec.htm

[CE7] In the case of EIC and ELC at FZJ.

[S1] http://www.cecam.org/initiatives.html

[S2] http://www.psi-k.org/codes.shtml

[S3] http://www.ccp.ac.uk/

[S4] http://www.fz-juelich.de/ias/jsc/EN/Expertise/SimLab/simlab_node.html

[S5] Within the Jülich-Aachen Research Alliance JARA-HPC; http://www.jara.org/en/research/jara-hpc/research/?S=0

[S6] http://www.scc.kit.edu/forschung/5960.php

[C1] http://www.virgo.dur.ac.uk/

[C2] http://hydra.sussex.ac.uk/

[C3] http://www.mpa-garching.mpg.de/gadget/

[C4] A.Dubey, K.Antypas, M.K.Ganapathy, L.B.Reid, K.Riley, D.Sheeler, A.Siegel and K.Weide, Extensible Component Based Architecture for FLASH, A Massively parallel, Multiphysics Simulation Code, Parall. Comp. **35**, 512-522 (2009); B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, D. Q. Lamb, P. MacNeice, R. Rosner, J. W. Truran and H. Tufo, FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes, Astrophys. J. Suppl. Ser. 131, 273334 (2000); http://flash.uchicago.edu/site/flashcode/

[C5] http://flash.uchicago.edu/site/index.shtml

[C6] http://www.skatelescope.org/

[C7] J.Grotendorst, N.Attig, S.Blügel and D.Marx (Edts.), Lecture Notes on *Multiscale Simulation Methods in Molecular Sciences*, Winter School, 2-6 March 2009, Forschungszentrum Jülich (2009).

[C8] M.Praprotnik, L.Delle Site and K.Kremer, *Adaptive Resolution Scheme (AdResS) for Efficient Hybrid Atomistic/Mesoscale Molecular Dynamics Simulations of Dense Liquids*, Phys.Rev. E **73**, 066701 (2006).

[C9] G. Giupponi, G. De Fabritiis and P. V. Coveney, *A coupled molecular-continuum hybrid model for the simulation of macromolecular dynamics*, Int. J. Mod. Phys. C **18**, 520 (2007).

[C10] G.Csányi, T.Albaret, M.C. Payne and A. De Vita, *"Learn on the fly": a hybrid classical and quantum-mechanical molecular dynamics simulation*, Phys. Rev. Lett. **93** 175503 (2004).

[C11] http://www.mmm2010.de/aim-and-scope-of-the-conference/mmm-conference-series/

[C12] http://www.iter.org/

[C13] Council Decision 2007/614/Euratom;
http://europa.eu/legislation_summaries/energy/nuclear_energy/i23046_en.htm

[C14] http://www.riken.go.jp/engn/

[C15] M.Taiji, T.Narumi and A.Konagaya, MDGRAPE-3: A petaflops special-purpose computer system for molecular dynamics simulations, Adv. Parall. Comp. 13, 669-676 (2004)

[C16] Fukushige, T., Taiji, M., Makino, J., Ebisuzaki, T., and Sugimoto, D., A Highly-Parallelized Special-Purpose Computer for Many-body Simulations with An Arbitrary Central Force: MD-GRAPE, Astrophys. J., 468, 51-61, (1996).

[C17] G.Kikugawa, R.Apostolov, N.Kamiya, M.Taiji, R.Himeno, H.Nakamura and Y.Yonetawa, Application of MDGRAPE-3, a special purpose board for molecular dynamics simulations, to preiodic biomolecular systems, J. Comp. Chem. 30, 110-118 (2008).

[C18] B.G.Elmegreen, R.H.Koch, M.E.Schabes, T.Crawford, T.Ebisuzaki, H.Furusawa, T.Narumi, R.Susukita and K.Yasuoka, *Simulations of magnetic materials with MDGRAPE-2,* IBM J. Res. Dev. **48**, 199-207 (2004)

[C19] D.Madina and M.Taiji, Circuit and physical design of the MDGRAPE-4 on-chip network links, Proceedings of the 2008 international workshop on System level interconnect prediction, (2008).

[C20] SI^2 – Software Infrastructure for Sustained Innovation;
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503489&org=DMR&from=home

[C21] Bilateral program MWN;
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12820&org=DMR&from=home

[C22] Materials World Network and International Materials Institute (IMI);
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5328&org=DMR&from=home

[C23] Multiscale Modeling and Computation; http://www.nsf.gov/pubs/2009/nsf09017/nsf09017.jsp

[C24] Chemical Theory, Models and Computational Methods (CTMC);
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503420&org=CHE&from=home

[C25] Theoretical Elementary Particle Physics;
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503535&org=PHY&from=home

[C26] Physics at the Information Frontier (PIF);
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=6681&org=PHY&from=home

[C27] Computational Mathematics;
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5390&org=DMS&from=home

[C28] http://www.dfg.de/en/research_funding/international_cooperation/research_collaboration/g8-initiative/

[C29] http://eng.mon.gov.ru/

[C30] http://www.rfbr.ru/rffi/eng/

[C31] http://www.hugeinformations.co.cc/2011/05/russian-nanotechnology-corporation.html

[C32] http://i-gorod.ru/en/ ; see also http://web.mit.edu/newsoffice/2010/ia-skolkovo-0624.html

[C33] http://www.msu.ru/en/

[C34] http://www.gpi.ru/eng/index.php

[C35] http://www.kiae.ru/100e.html

[C36] http://aliceinfo.cern.ch/Public/Welcome.html