EESI – Working Group 3.3

WP3 Task 3 - Fundamental Sciences - Chemistry, Physics

Chair: CECAM-FZJ (Godehard Sutmann)
Vice-Chair: CEA (Jean-Philippe Nominé)
Kick-off Summer 2010

1st Expert Meet.
Paris
Jan. 2011

2nd Expert Meet.
Brussels
April 2011

WP3&WP4
Bologna
Feb 2011

IESP San Francisco
April 2011

Final meeting Autumn 2010
10-11 October Barcelona

EESI - Paris
June 2011
Scientific Domains in WG 3.3

- **Fundamental sciences:** Physics, Chemistry, Material Sciences, Astrophysics

[Diagram showing domains and sub-domains related to scientific research areas]
Composition of WG 3.3

- Description of the scientific and technical perimeter of the WG
  - Address science drivers and grand challenge problems in the fields of physics and chemistry
  - How are communities/science organizations going to prepare future software issues?

- Workgroup composition:
  - Astrophysics
  - Laser- / plasma-physics
  - Fusion
  - Material Sciences
  - Quantum Chemistry
  - Soft Matter Physics
  - Software Engineering and Algorithms
<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
<th>Country</th>
<th>Area of Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volker Springel</td>
<td>Garching, MPI Astrophysik</td>
<td>Ger</td>
<td>Astrophysics</td>
</tr>
<tr>
<td>Romain Teyssier</td>
<td>ETH Zürich</td>
<td>Sui</td>
<td>Astrophysics</td>
</tr>
<tr>
<td>Maurizio Ottaviani</td>
<td>CEA</td>
<td>Fra</td>
<td>Fusion</td>
</tr>
<tr>
<td>Luis Silva</td>
<td>Universidade Tecnica de Lisboa</td>
<td>Por</td>
<td>Laser Plasma Interaction</td>
</tr>
<tr>
<td>Alexei Removich Kokhlov</td>
<td>Moscow State University</td>
<td>Rus</td>
<td>Soft Matter</td>
</tr>
<tr>
<td>Alessandro Curioni</td>
<td>IBM Research - Zurich</td>
<td>Sui</td>
<td>Materials Sciences</td>
</tr>
<tr>
<td>Gilles Zerah</td>
<td>CECAM - CEA</td>
<td>Fra</td>
<td>Materials Sciences</td>
</tr>
<tr>
<td>Nicola Marzari</td>
<td>University of Oxford</td>
<td>UK</td>
<td>Materials Sciences</td>
</tr>
<tr>
<td>Adrian Wander</td>
<td>STFC Daresbury</td>
<td>UK</td>
<td>Materials Sciences</td>
</tr>
<tr>
<td>Mike Payne</td>
<td>University of Cambridge</td>
<td>UK</td>
<td>Quantum Chemistry</td>
</tr>
<tr>
<td>Thierry Deutsch</td>
<td>CEA</td>
<td>Fra</td>
<td>Quantum Chemistry</td>
</tr>
<tr>
<td>Mike Ashworth</td>
<td>STFC Daresbury</td>
<td>UK</td>
<td>Methods and algorithms</td>
</tr>
<tr>
<td>Thomas Schultess</td>
<td>CSCS</td>
<td>Sui</td>
<td>Materials Sciences</td>
</tr>
<tr>
<td>Pieter in t'Veld</td>
<td>BASF</td>
<td>Ger</td>
<td>Soft matter</td>
</tr>
</tbody>
</table>

Ger: 2  Sui: 3  Fra: 3  Por: 1  Rus: 1  UK: 4  Total: 14
Example for codes in WG 3.3

- **Quantum Chemistry / Material Science**
  - AbInit, BigDFT, CASTEP, ONETEP, CP2K, CPMD, Quantum-Espresso, Wannier90, Octopus, GPAW, Crystal, Dalton, Turbomole, Columbus

- **Molecular Dynamics (Soft Matter)**
  - DL_POLY, Gromacs, Espresso, LAMMPS, NAMD

- **Laser / Plasma**
  - TORB, ORB5, Euteurpe, ELMFIRE, GYSELA

- **Astrophysics**
  - Gadget, AREPA, PKDGRAV, Pluto, RAMSES

*List not complete...*
Scientific activities and software issues

- **scientific activities**
  - astrophysics: large scale structure of the universe
  - fusion: ITER facility
  - plasma: cold plasmas, magnetic plasmas
  - material sciences: catalysis, cracks, magnetic properties
  - soft matter: polymers, membranes for fuel cells, self-aggregation
  - algorithms: fault tolerance, energy efficiency, locality, optimal order algorithms

- **potential need for exascale performance proved**
  - Material Science and Quantum Chemistry have potential for sustained PetaFlop applications at present
  - > 1PFlop/s sustained performance on Jaguar
### Applications running at scale on Jaguar @ ORNL

**Fall 2009**

<table>
<thead>
<tr>
<th>Domain area</th>
<th>Code name</th>
<th>Institution</th>
<th># of cores</th>
<th>Performance</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>DCA++</td>
<td>ORNL</td>
<td>213,120</td>
<td>1.9 PF</td>
<td>2008 Gordon Bell Prize Winner</td>
</tr>
<tr>
<td>Materials</td>
<td>WL-LSMS</td>
<td>ORNL/ETH</td>
<td>223,232</td>
<td>1.8 PF</td>
<td>2009 Gordon Bell Prize Winner</td>
</tr>
<tr>
<td>Chemistry</td>
<td>NWChem</td>
<td>PNNL/ORNL</td>
<td>224,196</td>
<td>1.4 PF</td>
<td>2008 Gordon Ball Prize Finalist</td>
</tr>
<tr>
<td>Materials</td>
<td>OMEN</td>
<td>Duke</td>
<td>222,720</td>
<td>860 TF</td>
<td></td>
</tr>
<tr>
<td>Chemistry</td>
<td>MADNESS</td>
<td>UT/ORNL</td>
<td>140,000</td>
<td>550 TF</td>
<td></td>
</tr>
<tr>
<td>Materials</td>
<td>LS3DF</td>
<td>LBL</td>
<td>147,456</td>
<td>442 TF</td>
<td>2008 Gordon Ball Prize Winner</td>
</tr>
<tr>
<td>Seismology</td>
<td>SPECFEM3D</td>
<td>USA (multiple)</td>
<td>149,784</td>
<td>165 TF</td>
<td>2008 Gordon Ball Prize Finalist</td>
</tr>
<tr>
<td>Combustion</td>
<td>S3D</td>
<td>SNL</td>
<td>147,456</td>
<td>83 TF</td>
<td></td>
</tr>
<tr>
<td>Weather</td>
<td>WRF</td>
<td>USA (multiple)</td>
<td>150,000</td>
<td>50 TF</td>
<td></td>
</tr>
</tbody>
</table>
Materials Science: first principles design of materials

- e.g. energy conversion
  - chemical energy to electricity (fuel cells)
  - sunlight to electricity (photovoltaic cells)
  - nuclear energy to electricity (nuclear fission or fusion)

- e.g. energy storage
  - batteries
  - supercapacitors
  - chemical storage in high energy density fuels (hydrogen, ethanol, methanol)

- e.g. energy use
  - solid state lighting, smart windows, low power computing, lightweight materials for transporting
Materials Science: first principles design of materials

- High throughput design

  "Once a material property can be calculated accurately from quantum simulations, it becomes straightforward to systematically explore hundreds or thousands of compounds for improved performance."

  N. Marzari

- Acceleration of invention and discovery in science and technology
  - reduce costs
  - reduce interval of time-of-discovery to time-to-commercialisation
WG3.3 Science Drivers

- **Soft-Matter Research**: overcome time- and length scales for device simulations

  - Catalysis
    - temperature effects
    - non-equilibrium
    - chemical reactions

  - Self-organization and self-assembly of nano-structures
    - length- and time-scale wall (e.g. for polymers timescales ~N²)

  - Coupled device modeling
    - Fuel cells: device modeling with explicit system description
      -> multiscale modeling for gas-liquid-solid + chemical reactions
WG3.3 Science Drivers

- **Plasma-Fusion Research**: ITER - magnetic fusion for energy production
  - Characterisation of plasma dynamics
  - Understanding energy and particle confinement
  - Predict instabilities and their consequences
  - Challenges:
    - Spatial domains: electron scales (\(\Omega m\)), micro-turb. (mm), machine scal. (m)
    - Time-scale separation: energy confinement time (s), microturbulence (\(\Omega s\))
    - Fully electromagnetic simulations, ions + electrons
    - Time scale simulations up to \(~1s\) (1ms at present)
    - Spatial resolutions down to mesh sizes \(10^5\)
      -> electronic scales
    - Memory limitations at present (need for fat nodes)
WG3.3 Science Drivers

- **Astrophysics**
  - Dark energy, large scale structure and cosmology
    - From galaxies to Hubble volume
    - Large n-body simulations ($16384^3$ part. for resolution)
      current limit $8192^3$ particles – memory limits
  - Physics of clusters and galaxy formation
    - very inhomogenous systems, different time scales – load balancing
    - need for lots of simulations to explore variance
  - Physics of planetary formation
    - from molecular clouds to planetesimals
    - time- and length scales – need for multiscale modeling
    - new physics with MHD and dust-gas coupling
  - Stellar interiors and model of the Sun
    - Supernovae mechanisms, compact stars
    - goal: global model of stellar structure & dynamics
    - MHD & radiative transport
Main methods used in application fields:

- Particle methods
  - e.g. molecular simulations, particle based hydrodynamics
  - PIC, MD, Brownian Dynamics, Monte Carlo
  - Long range- / non-local interactions

- Mesh-based methods
  - e.g. Navier-Stokes, MHD
  - Adaptive mesh-refinement
  - Multigrid
  - FFT

- Ab initio / electronic structure calculations
  - Linear algebra (e.g. Eigenwert solver, Cholesky, matrix-vector)
  - Wavelets
Goals to be achieved

- reduce time-to-solution / time-to-market
- reduce / optimize energy-to-solution
  - new metric for job-accounting foreseen: (time-to-solution) x (energy)
- Algorithmic targets:
  - seek for optimal complexity (O(N))
    * Fast Multipole Methods
    * Decomposition methods
    * Multigrid
    * H-matrices
  - seek for locality (reduce data movement and therefore energy)
    * communication-friendly or -avoiding algorithms
    * time-scale splitting schemes
    * real space methods
    * wavelets
Preparing for the next steps: **Short term perspectives**

- **Modularisation of codes**
  - share components of codes between different groups

- **Most groups start thinking in terms of**
  - extending codes to hybrid: MPI + OpenMP / P-Threads
  - writing codes or parts of codes for: GPU
  - planning extensions for codes in multi-stage parallelism MPI + OpenMP + accelerator (GPU, FPGA,...)
Preparing for the next steps: Long term perspectives

- The exa-scale challenge – the threefold way
  - Strong- / weak scaling
  - Multiscale (horizontal / vertical)
  - Ensemble simulations
WG3.3: strong- / weak-scaling

- Strong- / weak scaling

  - Common opinion of experts:

    - do not build up on existing codes
    - rewrite legacy codes
    - adjust / choose algorithms for exa-scale hardware
    - address hardware specific features and design special algorithms for specific hardware features

    - implies several man-years for redesign of functional units in programs

  New design and optimal implementation of codes


- Multiscale Simulations

  - partial solution to escape from the dilemma of hyper-scaling
    * strong scaling not possible for a lot of codes
      (e.g. for some commercial codes or quantum chemistry)

  - solution to weak scaling (WS) problem, since WS not always
    * desired (e.g. in Quantum Chem. not everything is “worth” to be calculated in full precision)
    * or possible (e.g. non-linear increase in memory consumption)

  - combine codes and run concurrently

**Horizontal Multiscale**
Multiscale Simulations

- not all experts need / want full exascale performance in a single program application
  
  - solution is in running different codes coupled simultaneously
  - run codes concurrently on smaller number of nodes
  
- **chance for a survival of legacy codes**
  
  - codes should be coupled through *standard interface*
    (this has to be developed and agreed for in the community)
  
- develop multiscale simulation codes which can be used as *plug & play*
WG3.3: Ensemble Simulations

- Ensemble Simulations
  - “perfect parallel scaling”
    - used to increase statistics
    - verification of methods
    - method towards fault tolerant computing
Software issues for Fundamental Sciences

- Fault tolerant and energy efficient algorithms
- Software support to measure or estimate energy
- Data locality
- Optimal order algorithms
- Mesh generation
- Algorithms with low communications
- Standard interfaces for multiscale simulations
- Support of several executables in one job
- Parallel I/O
- Load-balancing