

# Service Infrastructures for Research: Supercomputers, Grids, and Clouds, and the DEISA Example

Wolfgang Gentzsch, Hermann Lederer

Rechenzentrum Garching der Max-Planck-Gesellschaft  
Max-Planck-Institut für Plasmaphysik  
Boltzmannstr. 2  
D-85748 Garching  
gentzsch@rzg.mpg.de  
lederer@rzg.mpg.de

**Abstract:** In the following, we will summarize the latest developments in the areas of Grid and Cloud Computing. With the aid of the EU funded project DEISA (Distributed European Infrastructure for Supercomputing Applications), we will explain design, development, and use of e-Infrastructures, and application enabling and implementation. The second part will briefly present three science applications and their results achieved on the DEISA HPC infrastructure, in the areas of Fluid Turbulence, Lattice Quantum Chromodynamics, und Multiprotein Complexes. We will also explain the implementation of the application codes on and optimization for the distributed DEISA HPC Infrastructure.

## 1 Introduction

During the last decades, computer simulation has been established as the 3<sup>rd</sup> pillar of sciences, next to theory and experiment. Scientists are now able to access modern compute and data service infrastructures and run their grand-challenge applications on the best suited supercomputer.

High-speed networks transport data at the speed of light, middleware manages distributed resources in an intelligent manner, portal technology enable secure, seamless, and remote access to resources, and sophisticated numerical methods approximate the underlying mathematical equations in a highly accurate way. With the convergence of these core technologies into one complex service oriented architecture, we see the rise of large compute and data grids currently being built and deployed by e-Infrastructure initiatives such as DEISA.

## 2 Workshop Programme

In this workshop, we will present the latest developments in the areas of Grid and Cloud Computing. With the aid of the EU funded project DEISA (Distributed European Infrastructure for Supercomputing Applications), we will explain design, development, and use of e-Infrastructures, and application enabling and implementation.

The second part of the workshop will present three science applications and their results achieved on the DEISA HPC infrastructure, in the areas of Fluid Turbulence, Lattice Quantum Chromodynamics, und Multiprotein Complexes. We will also explain the implementation of the application codes on and optimization for the DEISA HPC Infrastructure.

- Wolfgang Gentzsch, DEISA: Building and operating research e-Infrastructures, Grids, and Clouds
- Andrea Vanni, CINECA: Infrastructure operation & technologies deployment in DEISA
- Mariano Vazquez, BSC: Enabling applications for the European supercomputing ecosystem
- Joerg Schumacher, University Ilmenau: Cloud09: Cloud formation in moist convective turbulence
- Istvan Montvay, DESY: Lattice Quantum Chromodynamics – Quantum Field Theory on the Lattice
- Michael Martinez, EML-Research: Musiprol: Multiscale simulation of membrane-associated multiprotein complexes

The following chapters provide a summary of the six individual workshop contributions.

## 3 Building and operating e-Infrastructures, Grids, and Clouds

High-speed networks transport data at the speed of light, middleware manages distributed computing resources in an intelligent manner, portal technology enable secure, seamless, and remote access to resources, and grand-challenge applications, data, and sophisticated numerical methods approximate the underlying mathematical equations in a highly accurate way. With the convergence of these core technologies into one complex service oriented architecture, we see the rise of large compute and data grids currently being built and deployed by e-Infrastructure initiatives such as DEISA, EGEE, NAREGI, and TERAGRID. While we master most of the technology aspects of these infrastructures, we still face a number of social, mental, cultural, and legal challenges.

While grids and virtualization provide the ‘plumbing’ to enable seamless access to distributed resources, clouds on the other hand denote services on a pay-per-use basis. Grids stand out because of their flexible, dynamic, feature-rich resources and thus are complex by their very nature. Cloud applications will likely follow similar strategies as grid-enabling ones. Just as challenging, though, are the cultural, mental, legal, and

political aspects of clouds. Building trust and reputation among the users and the providers will help in some simple scenarios. But it is still a challenge to imagine users easily entrusting their corporate assets and sensitive data to cloud service providers.

One promising example of an HPC grid is the Distributed European Infrastructure for Supercomputing Applications, or DEISA. DEISA's Extreme Computing Initiative (DECI) is successfully offering millions of supercomputing hours to the European e-Science community, at their finger tip, and helping scientists gain new scientific insights. Why is DECI so successful? Several reasons: because it has a very targeted focus on specific, long-running, supercomputing applications; many of the applications just run on one single system; it has user-friendly access to resources, also through DESHL and UNICORE; its coordinating function gives consortium partners full autonomy; and because there is an application task force (ATASKF) that helps users port their applications to the supercomputing infrastructure.

This presentation begins with a short introduction into architecture, components, applications and benefits of e-infrastructures. We will elaborate on some of the most obstructive barriers for building these infrastructures and for their wider acceptance. We will then discuss some guidelines which might help in developing sustainable infrastructures for research and industry. Finally, we will look at several important aspects beyond technological issues, such as sharing of resources, sensitive data, grid- and cloud-enabling applications, open source, liability, licensing, and intellectual property, which can prevent further development and acceptance of these new technologies. Finally, we will elaborate on the main differences between grids and clouds, with the aid of the DEISA experience, and discuss what grids can learn from clouds such that they become more user-friendly, or even become an HPC Cloud.

## **4 Infrastructure operation & technologies deployment in DEISA**

DEISA is a stable, resilient, available and serviceable leading edge distributed infrastructure. Technologies WP aims to keep updated, enhance and further empower the existing DEISA infrastructure by transposing latest available technologies into improved or additional useful services based on identified user requirements. On the other hand Operation WP operates day by day for guarantee a working infrastructure by resolving infrastructure related problems installing on production new software proposed by Technologies WP and identifying, collecting and proposing turnkey operation best practises for the integration of new HPC sites creating European HPC eco-system.

## **5 Enabling applications for the European supercomputing ecosystem**

DEISA, the Distributed European Infrastructure for Supercomputing Applications, is a consortium of leading national Supercomputing centres that aims at fostering the pan-European world-leading computational science research. DEISA deploys and operates a persistent, production quality, distributed supercomputing environment with continental scope. It aims at delivering a turnkey operational solution for a future European HPC ecosystem. In order to make this infrastructure a real tool, available to the scientific community, one of the main DEISA efforts is to adapt European scientific applications to run efficiently on this infrastructure. These are the duties of the Applications Work Packages which are in charge of enabling the applications, make them ready-to-run in production, and select a small group of applications to go beyond the pure enabling, by identifying bottlenecks and enhancing parallel scalability.

## **6 Cloud09: Cloud formation in moist convective turbulence**

In the DEISA project Cloud09 a detailed investigation of moist convective turbulence in a laterally extended shallow layer was conducted without usage of small-scale parameterizations of cloud and turbulence physics. Our central goal was to understand the differences between the well-known dry and the less-explored moist turbulent convection case, such as the impact of phase changes on the turbulent transport through the layer. Moist turbulent convection was studied therefore in a set of equations that extend the classical dry convection Boussinesq system in a first significant step. The high-resolution numerical experiments allowed a comprehensive analysis of the large-scale cloud formation in moist convection and its tight relation to the local fine-scale mechanisms of heat (or buoyancy) transport. In close connection to this point, the geometric properties of the generated cloud patterns were explored and compared with existing results from Large Eddy Simulations which use subgrid-scale models.

This computational study showed that the shallow moist convection model with a linear thermodynamic equation of state generates a variety of cloud patterns, starting from a cloud-free layer, via isolated clouds to a closed cloud layer. It has been demonstrated that the cloud cover is tightly connected to an enhanced turbulent buoyancy flux across the convective layer and an asymmetry of the vertical turbulent velocity fluctuations.

## **7 Hadron spectrum of QCD with one quark flavor**

After a short introduction to QCD lattice simulations the motivation of investigating a model quantum field theory with a single quark flavour is discussed. The results of the DEISA Project Nf1 on the hadron spectrum and the possible spontaneous CP-symmetry breaking are reviewed.

## **8 Musiprol: Multiscale simulation of membrane-associated multiprotein complexes**

Membrane-associated multiprotein complexes play a critical role in many biological processes. The modelling and simulation of such complexes is very challenging for computational studies due to their high number of degrees of freedom and their heterogeneity. A multiscale approach is taken using both all-atom and coarse-grain models of biomolecules.

We will illustrate our approach by means of two application examples: (i) the conformational dynamics of cytochromes P450 drug metabolizing enzymes bound to a phospholipid membrane, and (ii) the mode of binding of transient anchored proteins, which are bound to the membrane through an hydrophobic compound. Free energies calculations of the extraction of a hydrophobic compound from a membrane have been computed.

The all-atom calculations are performed with freely available molecular dynamics codes: (i) NAMD for its efficient parallelization until 2000 CPUs of a system containing 150 000 atoms on the BlueGene at IDRIS, (ii) GROMACS which allows parallel simulations of independent runs needed for free energy calculations. A total 5 microseconds was run on the CRAY machine at the CSC center.

The authors gratefully acknowledge computational resources from the DEISA Extreme Computing Initiative ('Musiprol' project) and the support of the German Federal Ministry of Education and Research (BMBF) HepatoSys network on systems biology of the hepatocytes ('EndoSys' project) as well as the Klaus Tschira Foundation.