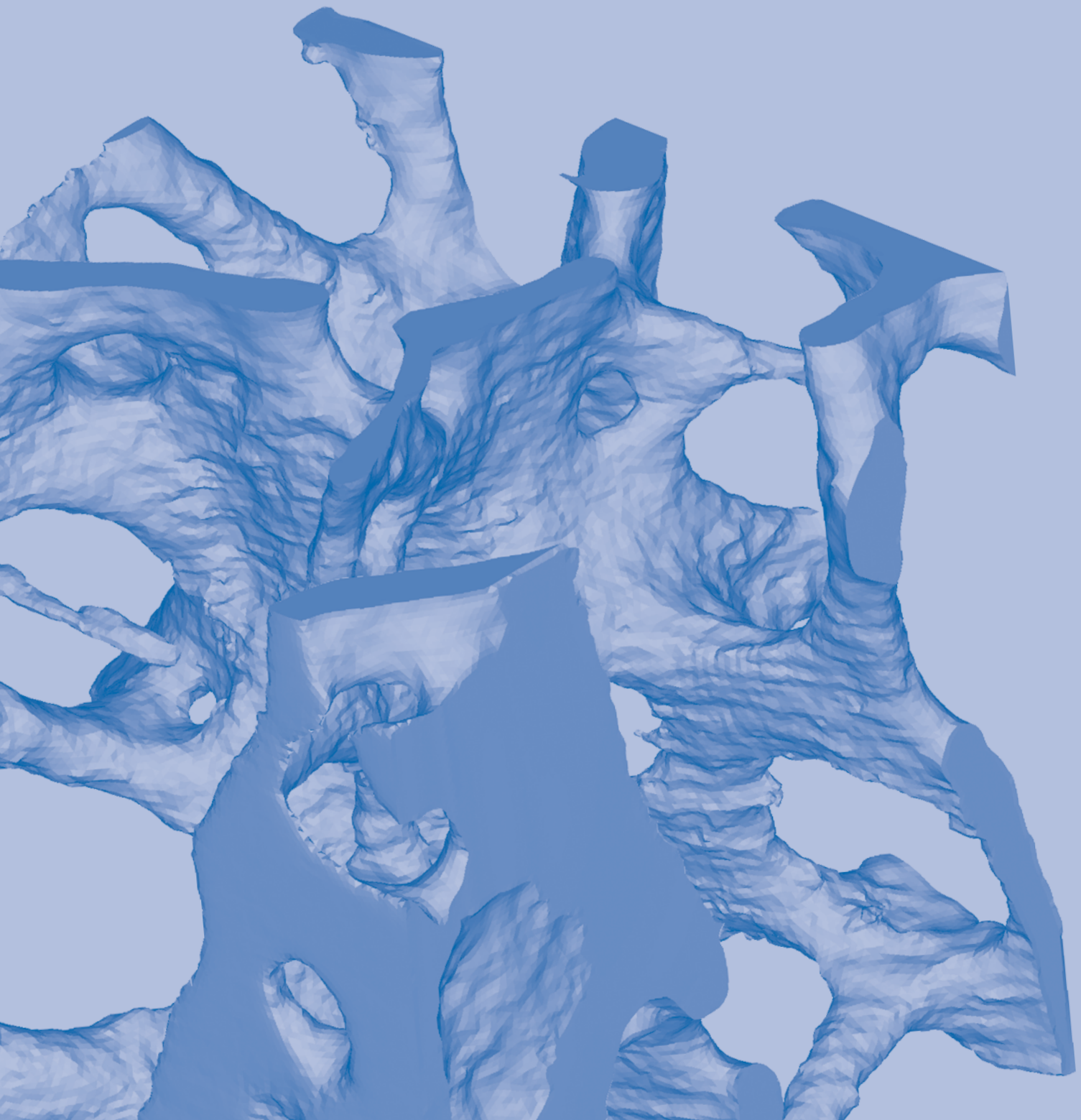
inSiDE

inSiDE • Vol. 8 No. 2 • Autumn 2010

Innovatives Supercomputing in Deutschland



Editorial

- Prof. Dr.
  H.-G. Hegering (LRZ)
- Prof. Dr. Dr.
  Th. Lippert (NIC)
- Prof. Dr.-Ing. Dr. h.c.
  M. M. Resch (HLRS)

Both Germany and Europe have made a leap forward in High Performance Computing and InSiDE can give you a first report on how the future will be shaped. Welcome to this latest issue of InSiDE the journal on innovative Supercomputing in Germany published by the Gauss Centre for Supercomputing.

While this issue goes to print both LRZ@Gauss and HLRS@Gauss are in the final phase of their procurements for Petascale systems. At the same time the results of the second call for projects of the Federal Ministry for Education and Research (BMBF) are eagerly awaited. We will report on the results of this combined hardware and software strategy in our next issue.

After a laborious PRACE-project phase also Europe has made the next step towards an integrated HPC infrastructure and organization. On April 23, 2010, sixteen organizations representing their countries in HPC matters signed the deed that created the PRACE AISBL (Association Internationale Sans But Lucratif) as an international non-profit organization under Belgian law. Since then PRACE has already issued its first call for projects and has started operations. InSiDE reports - as announced in the last issue - on the inauguration of PRACE and on the results of this first call. The first PRACE call clearly shows the need for High Performance Computing in Europe. This is emphasized by the very high quality of the projects submitted.

Projects are what drives our research and development in High Performance Computing. In this issue we present some of the most interesting user projects on German HPC systems. We start with a description of the "Simulation of the unsteady Flow around the Stratospheric Observatory For Infrared Astronomy SOFIA". A further engineering application is the description of the simulation of a scram-jet intake. These two articles show the potential of HPC in aerospace. Sustainable energy is another pressing issue to which HPC can contribute substantially. "Simulation of optical Waves in thin-film Solar Cells" shows how HPC can contribute to a European sustainable energy supply. The close collaboration between research and industry is highlighted in the article "Next Generation Simulation of Cell Factories" which describes how HPC can help to bring the results of biotechnology into our daily life.

Let me highlight three further important collaborations which will enable German centers to stay at the forefront of HPC. JSC@Gauss has formed two alliances with Intel, Partec on the one hand and IBM on the other hand. "Towards Exascale: Foundation of EIC and ECL" reports on the target of this initiative. LRZ@Gauss is working with Intel on multicore architectures as described in "Investigations of Intel's many integrated Core Architecture at LRZ." HLRS@Gauss together with its partners tackles the Green IT issues with partners like IBM in the European GAMES project.

All in all InSiDE will again give you a comprehensive overview of not only HPC in Germany but also in Europe.

Enjoy reading!

Contents

PRACE Research Infrastructure takes off

News

PRACE, the Partnership for Advanced Computing in Europe, is now a European legal entity. On April 23, 2010, sixteen organizations representing their country in HPC matters signed the deed that created the PRACE AISBL (Association Internationale Sans But Lucratif) as an international non-profit organization under Belgian law. PRACE has its initial seat in Brussels, Rue du Trône 98. This formal act was one of the key achievements of the Preparatory Phase project [1] and provides Europe with a HPC service that is second to none in the world. Four additional states joined PRACE in the meantime.

In a festive ceremony on June 9, 2010 PRACE was inaugurated in Barcelona. EC Deputy Director General Zoran Stancic and high level representatives form the Spanish, Catalonian, and Italian governments welcomed the participants. They emphasized the importance of HPC for Europe as a key driver for the development of modern science and technology and for addressing the major challenges of our time like climate change, energy saving, and the ageing population. In spite of the present economic climate only continued investments will ensure Europe's scientific and industrial competitiveness.

France, Germany, Italy, and Spain have each committed to provide supercomputing resources worth at least 400 Million Euro in total to PRACE during an initial period of five years and serve as hosting members. PRACE already started to provide HPC resources and services to European scientists on the IBM Blue Gene/P at the Gauss Centre for Supercomputing (GCS) partner site in Jülich. The first call for proposals was oversubscribed by a factor of five. Out of the 65 proposals by principal investigators from 16 countries, ten projects were granted a total of 328 million core hours. A second call for proposals is open. All applications undergo a single pan-European peer review process with a key focus on scientific or technological excellence to fulfill PRACE's mission as provider of world-class services for world-class science.

The European Commission continues to support PRACE through its Seventh Framework Programme (FP7). It awarded a grant of 20 million Euro to the PRACE membership consortium for a First Implementation Phase project (PRACE-1IP) (2). The project started on July 1, 2010 and has duration of two years. It is designed to support the accelerated implementation of the RRACE Research Infrastructure. The project supports the evolution of the Research Infrastructure by refining and extending the administrative, legal and financial

framework with focus on the specific requirements of industry. To enable leading edge science on novel systems the project assists users in porting, optimizing and petascaling applications to the different architectures and

The PRACE-1IP project advises the PRACE AISBL on procurements of the next generation of systems. Finally, it evaluates promising technologies, especially with respect to energy efficiency, with the ultimate goal to collaborate

deploys consistent services across the hosting sites. The tools and techniques will be selected to have broad applicability across many disciplines. This is accompanied by advanced training in modern programming methods and paradigms, establishing a permanent distributed training infrastructure.

The PRACE brand is already well established in the international HPC scene; extensive dissemination and outreach will be continued. PRACE continues to sponsor the PRACE Award at the annual ISC conference for an outstanding paper by students or young scientists on extreme scalability, scientific breakthroughs using HPC or hybrid systems, or related performance evaluation. PRACE exhibits at ISC and at SC conference, presenting the PARCE AISBL, the achievements of the PRACE projects and of DEISA. PRACE and DEISA are already planning the creation of a global European HPC ecosystem encompassing Tier-0 and Tier-1 resources for 2011 after the end of the DEISA2 project.

with industrial partners to develop products. A corner stone in this work is STRATOS, the PRACE advisory group for Strategic Technologies created in the Preparatory Phase project.

For up-to-date information about the PRACE AISBL and the PRACE-1IP project, the current list of members and latest news visit:
www.prace-ri.eu or
www.prace-project.eu

References

1) The PRACE project received funding from the European Community's Seventh Framework Programme under grant agreement no RI-211528. Additional information can be found under www.prace-project.eu

2) The PRACE-1IP project receives funding from the European Community's Seventh Framework Programme under grant agreement no RI-261557. Additional information can be found under www.prace-project.eu

• Dietmar Erwin

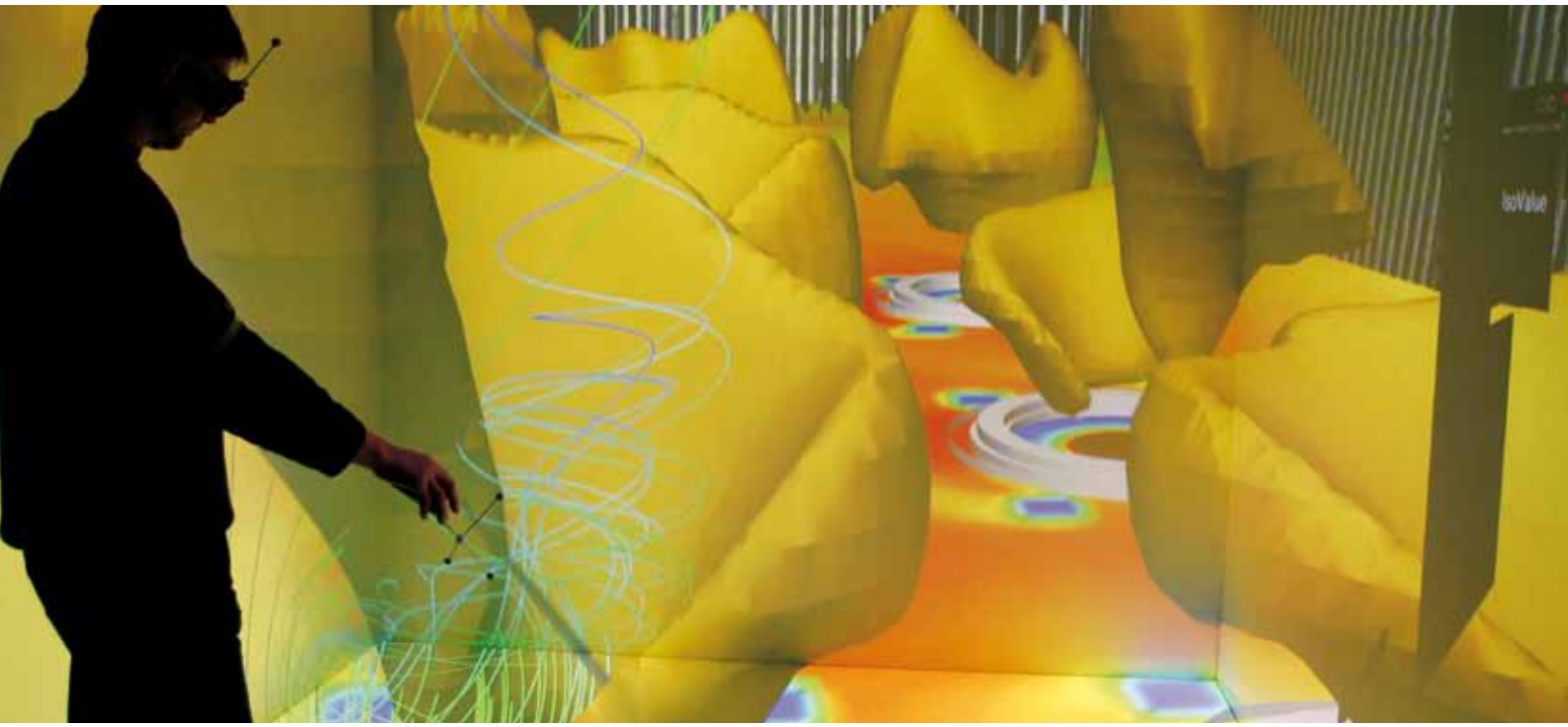
Jülich Supercomputing Centre

PRACE Early Access Projects

Since 1 August 2010, the Partnership for Advanced Computing in Europe (PRACE) offers supercomputer resources on the highest level (Tier-0) to European researchers. The German partner in PRACE is the Gauss Centre for Supercomputing (GCS) - the alliance of the three national supercomputing centres in Garching (LRZ), Jülich (JSC) and Stuttgart (HLRS). Jülich, as a member of GCS and involved in shaping PRACE as well as hosting the only European Tier-0 supercomputer currently available, is dedicating a 35% share of the IBM Blue Gene/P system JUGENE to PRACE within the framework of its commitments. This first Petascale HPC system available to researchers through PRACE is the fastest computer in Europe available for public research.

Proposals for the first PRACE projects on JUGENE were solicited in an early access call, released by PRACE on 10 May 2010 with the deadline 10 June 2010. Emphasis was put on projects that could start immediately with no or only little preparation and that would be able to achieve significant scientific results within an initial grant period of four months. The comparative international peer reviewing process was headed by Prof. Richard Kenway, EPCC.

After scientific evaluation and prioritization, ten out of 65 proposals were accepted in this very competitive process, five from Germany, two from the UK, and one each from Italy, the Netherlands and Portugal. These projects, two each from the fields of Astrophysics, Engineering and Energy, and from Fundamental Physics, and one each from the fields of Chemistry



and Materials, Earth and Environmental Sciences, and from Mathematics and Computer Science, were awarded a total of about 320 million compute core hours. More details on these projects can be found via the PRACE homepage.

From now on PRACE calls for Tier-0 computing time grants will be issued twice every year, the project starting dates being 1 May and 1 November of the year with the respective submission deadlines about 3 months earlier.

http://www.prace-project.eu/hpc-access/page-11/

• Walter Nadler

Jülich Supercomputing Centre

Simulation of the unsteady Flow around the Stratospheric Observatory For Infrared Astronomy SOFIA

The Stratospheric Observatory For Infrared Astronomy SOFIA, a joint project between NASA and DLR, is a 2.5 m reflecting telescope housed in an open cavity on board of a Boeing 747SP aircraft. The observatory operates in the stratosphere at an altitude above 13 km to observe objects in the universe in the infrared region of the electromagnetic spectrum.

The flow over the open port during the observation flights presents some challenging aerodynamic and aero-acoustic problems. In general, the flow over cavities such as the SOFIA telescope port is characterized by unsteady flow phenomena associated with prominent pressure fluctuations caused by amplified acoustic resonances. In the present case, this phenomenon causes unwanted vibrations of the telescope structure and deteriorates the image stability [1].

To investigate the pressure fluctuations in the cavity CFD simulations using the DES and URANS turbulence approach are performed with the Finite-Volume RANS-solver TAU that was developed by the Institute of Aerodynamics and Flow Technology of DLR [3]. The code solves the unsteady, compressible, three-dimensional Reynolds-averaged Navier-Stokes equations on unstructured or hybrid grids.

The objective of these simulations is to improve the telescope's performance by mitigating the amplitudes and changing the characteristic frequencies of the pressure fluctuations inside the cavity. This could be reached by installing passive flow control devices upstream and downstream the cavity port. The simulations show that the investigated measures have a high potential to increase the telescope's pointing stability [2].

Figure 1: View of the SOFIA aircraft during a testflight with the door closed (left picture) and with the door open (right picture)

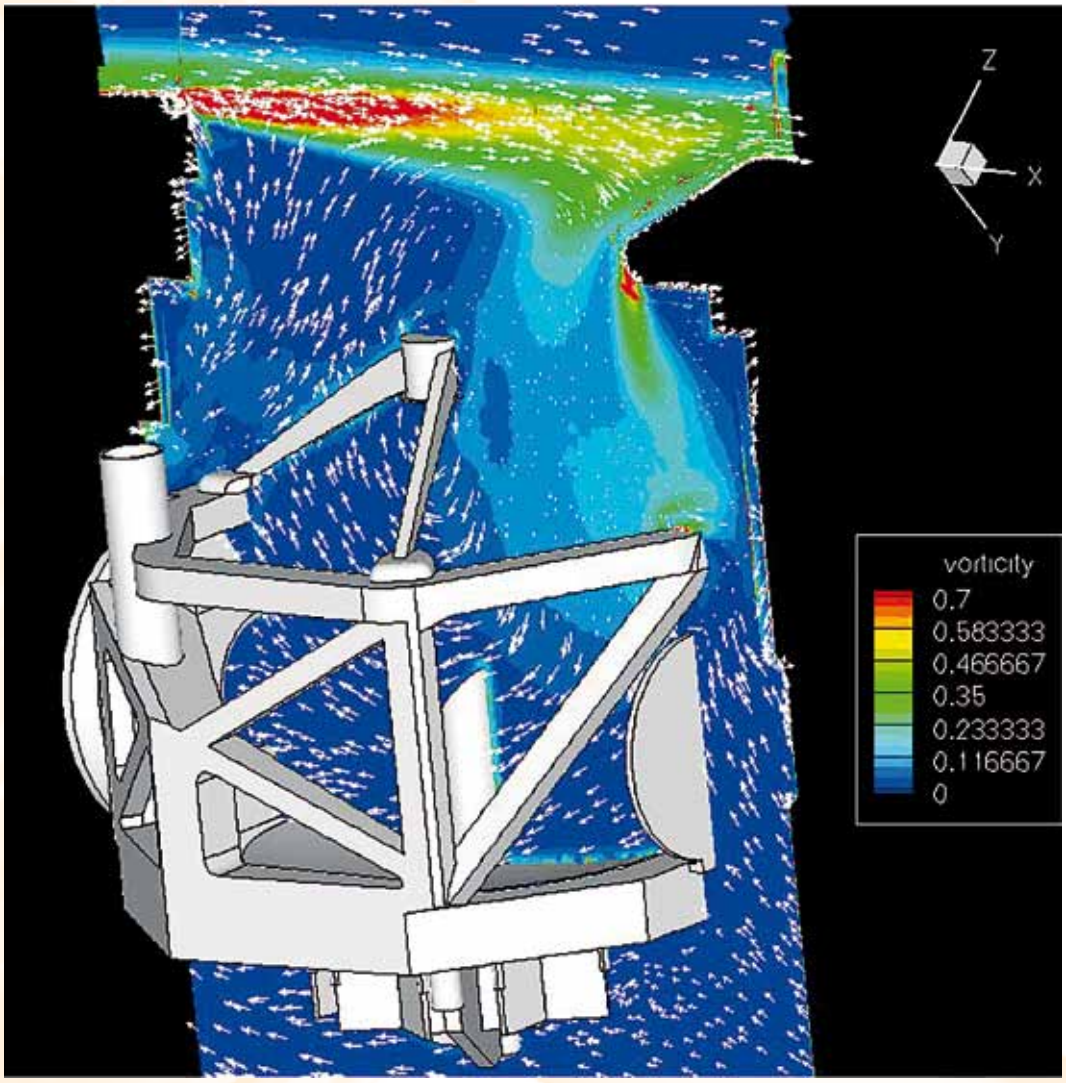


Figure 2: Instantaneous vorticity distribution inside the SOFIA telescope port for cruising speed conditions

The present computations were carried out on the HLRB II at LRZ. The typical grid size for a URANS simulation is about 20 million cells (half model). For parallelization, the domain was equally decomposed into 510 subdomains to utilize 510 processor cores in parallel. A typical CFD computation using the URANS turbulence model consisted of 6,000 physical timesteps with 180 inner iterations per step. One physical time step consumed about 56 s of wallclock time, yielding an overall time consumption of about 48,000 CPUh for one run. The physical time step was chosen to 16.4 ms, resulting in a computational period of 0.1 s under wind-tunnel conditions (1.6 s under flight conditions).

In the DES simulations the time step was decreased to 5.0 ms, 13,000 physical time steps (40 inner iterations per time step) were performed to simulate 0.065 s under wind-tunnel conditions (1.1 s under flight conditions). For DES simulations an additional LES grid block was added in the shear layer zone over the cavity to capture the relevant turbulent structures in this zone. The LES grid block is characterized by a very fine discretization which is blowing up the grid size up to 56 million cells. The computational cost for the DES run was about 65,000 CPUh (parallelization on 1,020 processor cores).

Thanks to these CFD simulations a detailed prediction of the flow field in space and time around the SOFIA telescope is possible (like shown in Fig. 2). In Figure 3 is illustrated the shear layer characteristic over the open cavity port issued from DES simulations, on the left picture by using a passive flow control device, here a porous fence installed upstream the cavity and on the right picture without a fence. In this case the SOFIA aircraft is flying in 41,000 feet with a Mach number of 0.85. Under this flight conditions the shear layer is characterized by large scale "Kelvin Helmholtz" vortices that are passing the cavity. The propagation and the size of these coherent structures have a significant impact on the excitation of pressure fluctuations inside cavity. As the picture in Figure 3 shows the transversal structures are ruptured by the fence, promoting bursting and creating chaotic three-dimensional vortices already at an earlier stage further upstream in comparison to the right picture where no fence is installed. In Figure 4 is illustrated that the shear-layer fence yields a significant decrease of the unsteady loads on the telescope assembly [4]. The repartition of the red zones being an indicator of high pressure fluctuations is reduced.

Outlook

In future studies further passive as well as active flow control devices will be analyzed. More important grid sizes will be necessary requiring a higher processor core number.

Acknowledgments

This project is being conducted on behalf of the DLR, German Aerospace Center and supported with funds from the Federal Ministry of Economics and Technology (BMWi), the state of Baden–Württemberg, and the University of Stuttgart. The support and provision of computing time by the LRZ is also gratefully acknowledged.

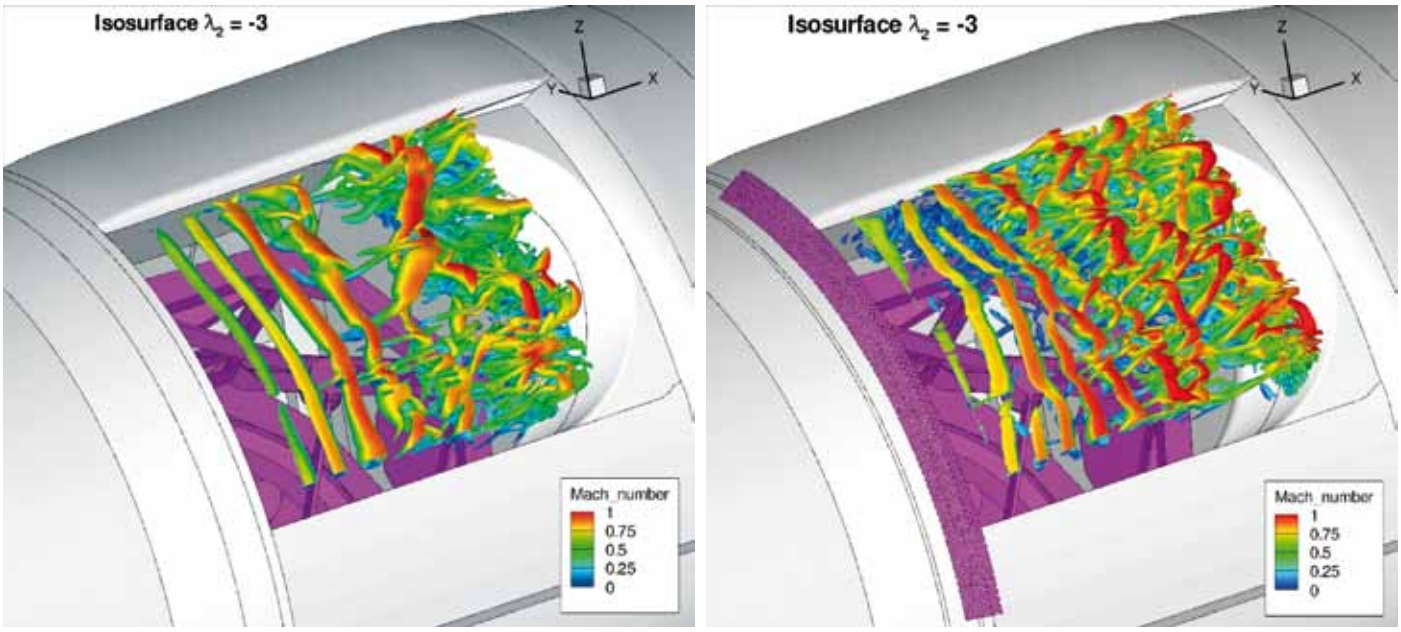


Figure 3: Shear layer characteristic with (left picture) and without a fence installed upstream of the cavity port (right picture)

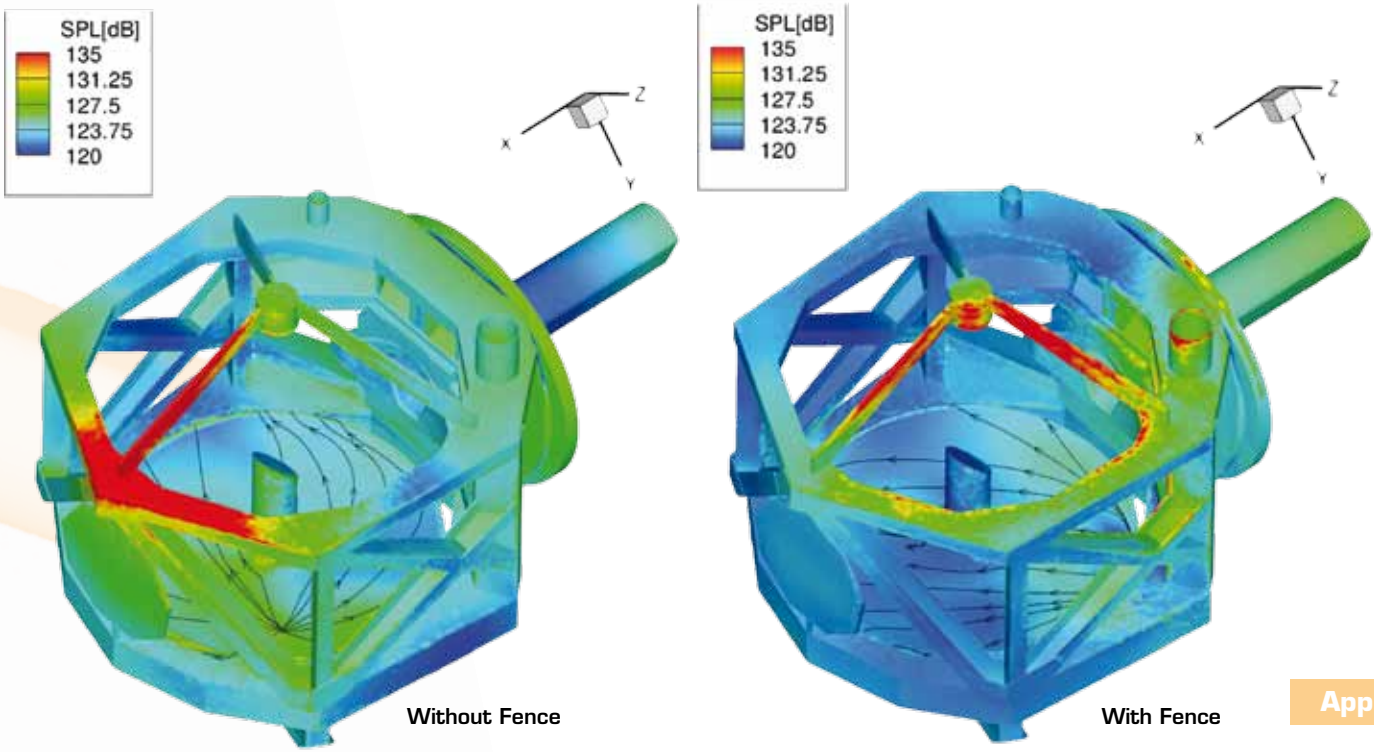


Figure 4: Pressure fluctuations on the Telescope Assembly with (left picture) and without using a fence (right picture)

References

[1] **Schmid, S.**
"Simulation of the unsteady flow around the Stratospheric Observatory For Infrared Astronomy SOFIA", Dr. Hut Verlag, 2010

[2] **Schmid, S., Lutz, T., Krämer, E., Kühn, T.**
"Passive Control of the Flow Around the Stratospheric Observatory for Infrared Astronomy", Journal of Aircraft, Vol. 46, No. 4, July–August, 2009

[3] **Gerhold, T.**
"Numerical Flow Simulation for Aircraft Design, chap. Overview of the Hybrid RANS Code TAU", pp. 81–92. Springer-Verlag, 2005

[4] **Schmid, S., Lutz, T., Krämer, E.**
"Passive control of the unsteady flow inside the SOFIA telescope port by means of a porous fence. In: Third Symposium on Hybrid RANS-LES Methods", Gdansk, Poland, 2009

- Christian Engfer
- Sven Schmid
- Thorsten Lutz
- Ewald Krämer

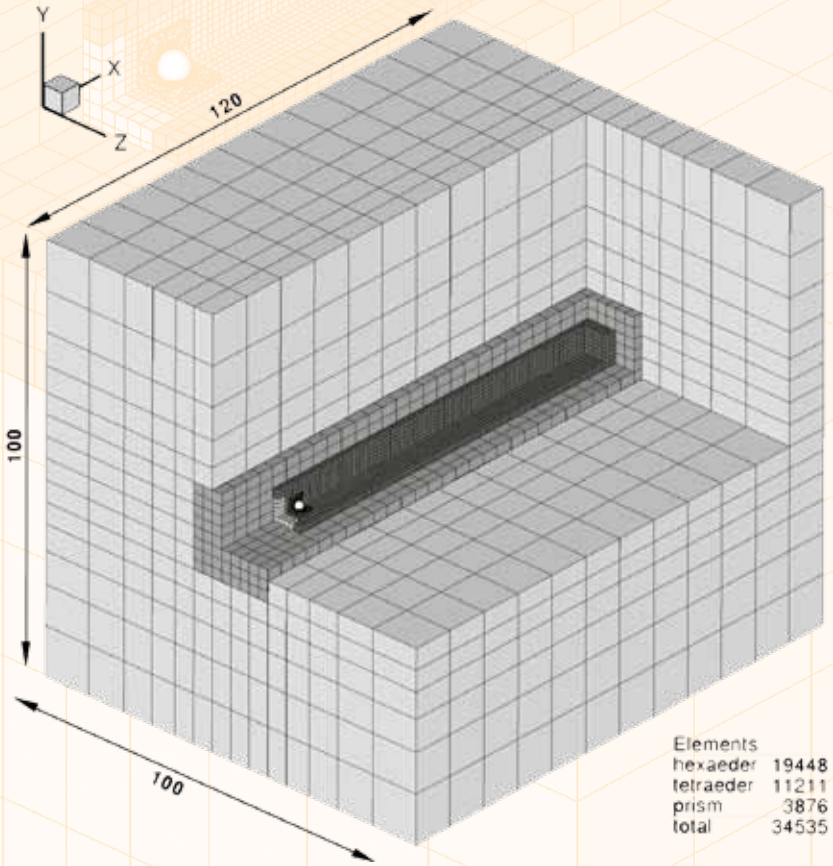
German Sofia Institute and Institute of Aero- and Gasdynamics, University of Stuttgart

High Order large Scale Calculations

This article describes two calculations performed for the numerical simulation of a scram-jet intake at a high Mach number of M=8 on the HLRB II at LRZ. This project is part of the DFG Graduiertenkolleg GRK 1095/2, - conception of a scram-jet demonstrator - and its partners at TU München, RWTH Aachen, Universität Stuttgart and the DLR. In order to pave the way for the simulation of the very complex intake flow field, preparatory calculations were performed to test certain code features and to provide insight into large parallel computations. These calculations are described in this article.

Numerical Code

Since all presented numerical problems contain strongly time dependent instationary phenomena, we are using a special unstructured explicit discontinuous Galerkin code, called HALO (Highly Adaptive Local Operator). This code is of high order of accuracy in both space and time and can handle unstructured hybrid grids with hanging nodes, consisting of tetrahedra, hexahedra, prisms and pyramids in 3D and even polygons in 2D. It is almost fully parallelized and requires only a minimum of inter-processor communication due to its explicit character. A main feature of this code is its time discretization method [1,2,3], that allows a high order



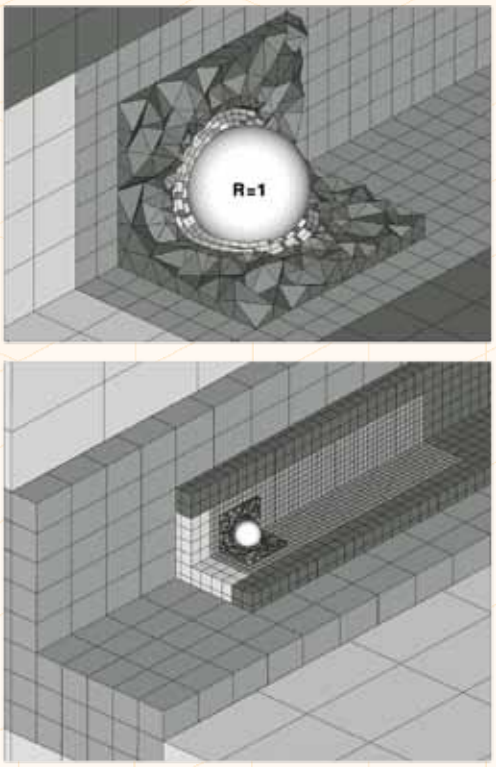


Figure 1: Hybrid block-unstructured mesh for the sphere calculation

time consistent local time stepping mechanism, wherein each grid-cell can advance with its own maximum possible time step. Various equation systems are implemented, such as Euler or Navier-Stokes equations as well as viscous Magnetohydrodynamic (MHD) equations. All presented calculations employ the Navier-Stokes equations. To suppress oscillations at discontinuities such as shocks, artificial viscosity is added to smear the shock profile resulting in stable computations.

Scale-up Efficiency

To test the scaling capabilities of the HALO code, we set up an example for which a perfect load balance was achievable. We were using the so-called manufactured solution technique for the 3D compressible unsteady Navier-Stokes equations: Inserting an analytical function into the Navier-Stokes equations leads to a right hand side that is prescribed as a source term in the numerical code.

The problem was set up with periodic boundaries so that the boundary communication will not differ from the inter-processor communications. The size of the computational problem was increased in parallel with the number of processors for calculation. This way, we kept a constant load in computation as well as in communication.

Table 1 shows the good scale-up efficiency of the HALO code for up to 4,080 processors with a constant load per processor.

Nb. of procs	1	1,000	2,197	4,080
Efficiency (%)	-	99.1	97.8	98.8

Table 1: Scale-up efficiency of the HALO code

The efficiency when calculating on N processors is calculated as the calculation time on one processor divided by the time needed for a calculation on N processors.

Performed Calculations

In order to be able to perform a large calculation of the scram-jet with features like hp-adaption, shock capturing and VMS, preparatory calculations were set up to test certain features and prepare the code for efficient computation on a large number of processors. These calculations tackle physical problems and will be described next.

3D Flow around a Sphere

The laminar time periodic flow around a sphere was set up to test the codes ability to handle unstructured hybrid grids and a p-adaption mechanism in three space dimensions. We were solving the unsteady compressible Navier-Stokes equations with a free-stream Mach number of Ma=0.3 and a Reynolds number of Re=300. The problem was discretized with a block-

Figure 2: Instantaneous λ2 isosurface of the sphere

unstructured grid consisting of prisms for the boundary layer, tetrahedra and hexahedra elsewhere. Figure 1 shows the different grid blocks and dimensions of the computational domain. P-adaption was arranged so that each grid cell was allowed to adapt its polynomial degree between 1 and 5 every 500 time steps.

To demonstrate the calculation results, Figure 2 shows a 3D view of the instantaneous vortex measure$\lambda 2$ for this test case. The color levels indicate the velocity magnitude. Here, one can easily see that the very large cells at the end of the wake cannot provide the necessary resolution and are therefore producing large inter cell jumps of the solution.

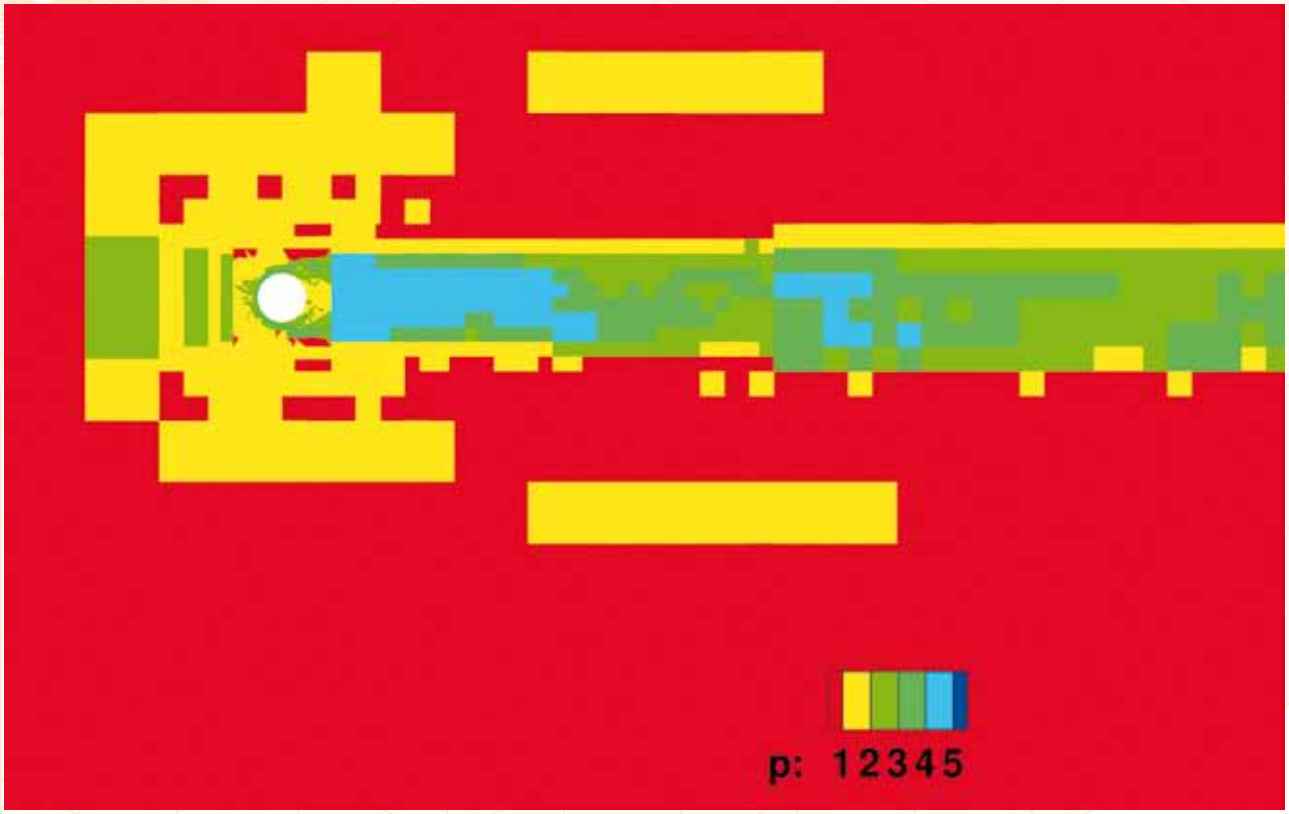
Finally, Figure 3 shows the distribution of the local polynomial degree p at end time tend=1,000.

3D Freestream Injector

This calculation targets shock capturing and p-adaption capabilities of the scheme, as we simulate a Ma=1.4, Re=30,000 injection.

The injection nozzle is designed according to specifications for gas injection devices used in the automotive industry. This problem also contains complex curved geometries that are challenging for high order schemes. The objective of this calculation is the aeroacoustic simulation of the instationary injection process, including the start up of the process. Results can be validated with calculations performed with other codes, both at our institute as well as in industry. Preliminary 2D calculations already provided insight into the necessary grid resolutions and shock capturing strategies. In our case, the problem was calculated on a grid with over 16 million DG degrees of freedom

Figure 3: Distribution of the local polynomial degree at end time of the sphere calculation



(4 million of hexahedral elements) on 500 to 1,000 processors. Figure 4 shows a 2D slice plane of the calculation (density distribution and velocity streamlines) together with an isosurface plot of the density that highlights the development of the flow.

To illustrate the flow development, several streamlines have been added. Please note that the picture presents an early phase of simulation. The geometry itself is rather complex, consisting of four kidney-shaped injection "nozzles" within the cylindrical injector and is assembled with unstructured hexahedra, allowing hanging nodes and polygons at connection surfaces.

Code Performance

The HLRB II provides an easy job performance summary directly in the command line. This tool can provide a first insight into the performance of the calculations and help to determine the computational efficiency of the numerical code. We hereby discovered strong performance differences, depending on the test case. Especially the scale-up tests as well as the sphere calculation where able to perform with up to 700 MFlop/s per processor. The latter performance was discovered for the 4,080 processor scale-up calculation resulting in a total of about 3 TFlop/s. It was found that certain settings of p-adaptivity and mesh structures do have a significant influence on the code performance.

Outlook

The presented calculations did not only show the potential of the HALO code but also gave an insight into the physics they were addressing and helped to improve the code. The code should now be ready to target the full intake simulation.

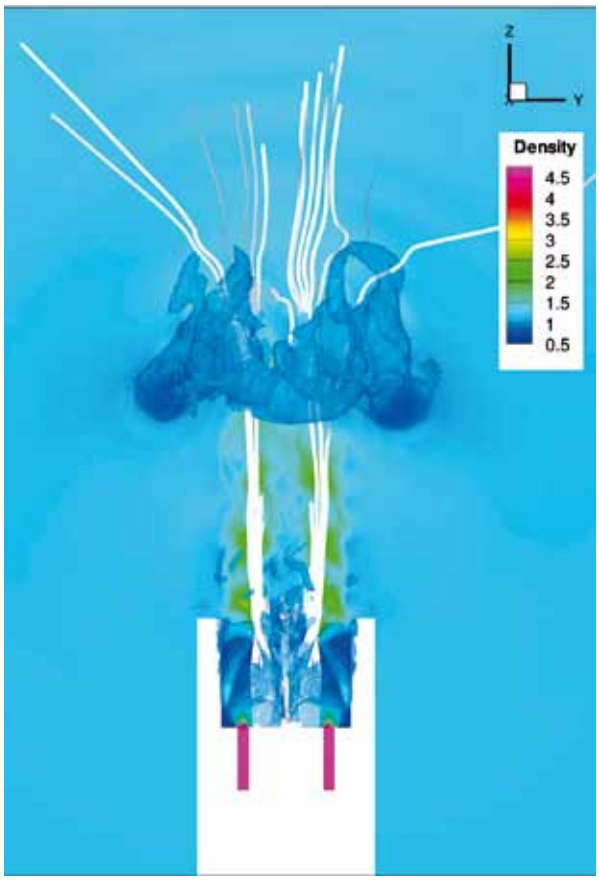


Figure 4: Density distribution and velocity streamlines of the 3D freestream injector

References

[1] **Lörcher, F., Gassner, G. and Munz, C.-D.** A discontinuous Galerkin scheme based on a space-time expansion. I. Inviscid compressible flow in one space dimension. In: Journal of Scientific Computing, Vol. 23, No. 2, pp. 175-199, 2007 DOI= http://dx.doi.org/10.1007/s10915-007-9128-x

[2] **Lörcher, F., Gassner, G. and Munz, C.-D.** A discontinuous Galerkin scheme based on a space-time expansion. II. Viscous flow equations in multi dimensions. In: Journal of Scientific Computing, Vol. 34, No. 3, pp. 260-268, 2008 DOI= http://dx.doi.org/10.1007/s10915-007-9169-1

[3] **Lörcher, F., Gassner, G. and Munz, C.-D.** A Contribution to the Construction of Diffusion Fluxes for Finite Volume and Discontinuous Galerkin Schemes. In: Journal of Computational Physics, Vol. 224, No. 2, pp. 1049-1063, 2007 DOI= http://dx.doi.org/10.1016/j.jcp.2006.11.004

- Christoph Altmann
- Gregor Gassner
- Marc Staudenmaier
- Claus-Dieter Munz

Institute of Aero- and Gasdynamics, University of Stuttgart

Simulation of optical Waves in thin-film Solar Cells

Thin-film solar cells are a forward-looking innovative solar cell technology. The difference of this technology in comparison with waver based solar cells is the manufacturing at low temperature (less than 300° Celsius) and the less material consumption. Both leads to a decrease of production cost in case of high energy prices. Another consequence is the low payback time of less than 1.5 years.

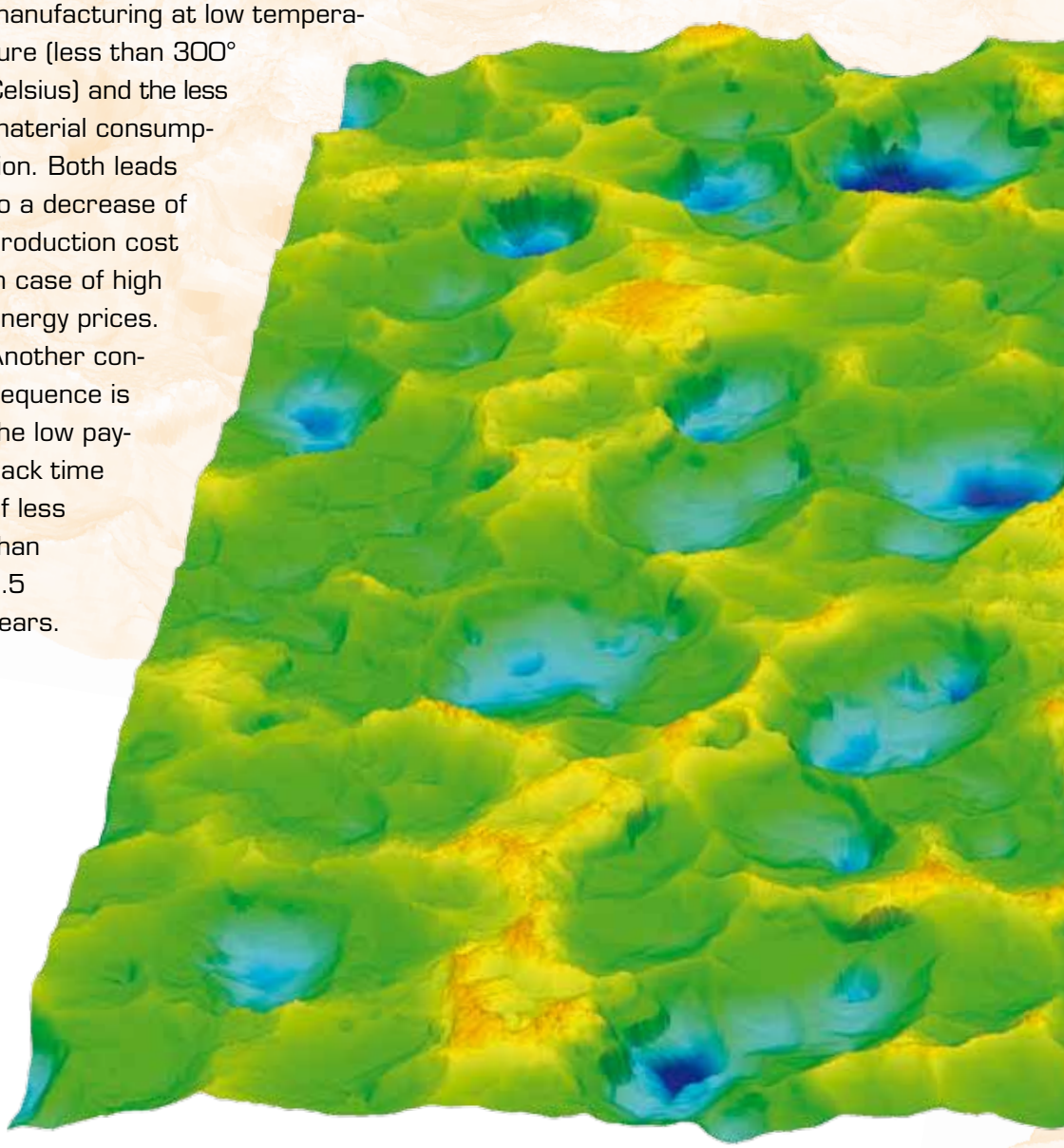
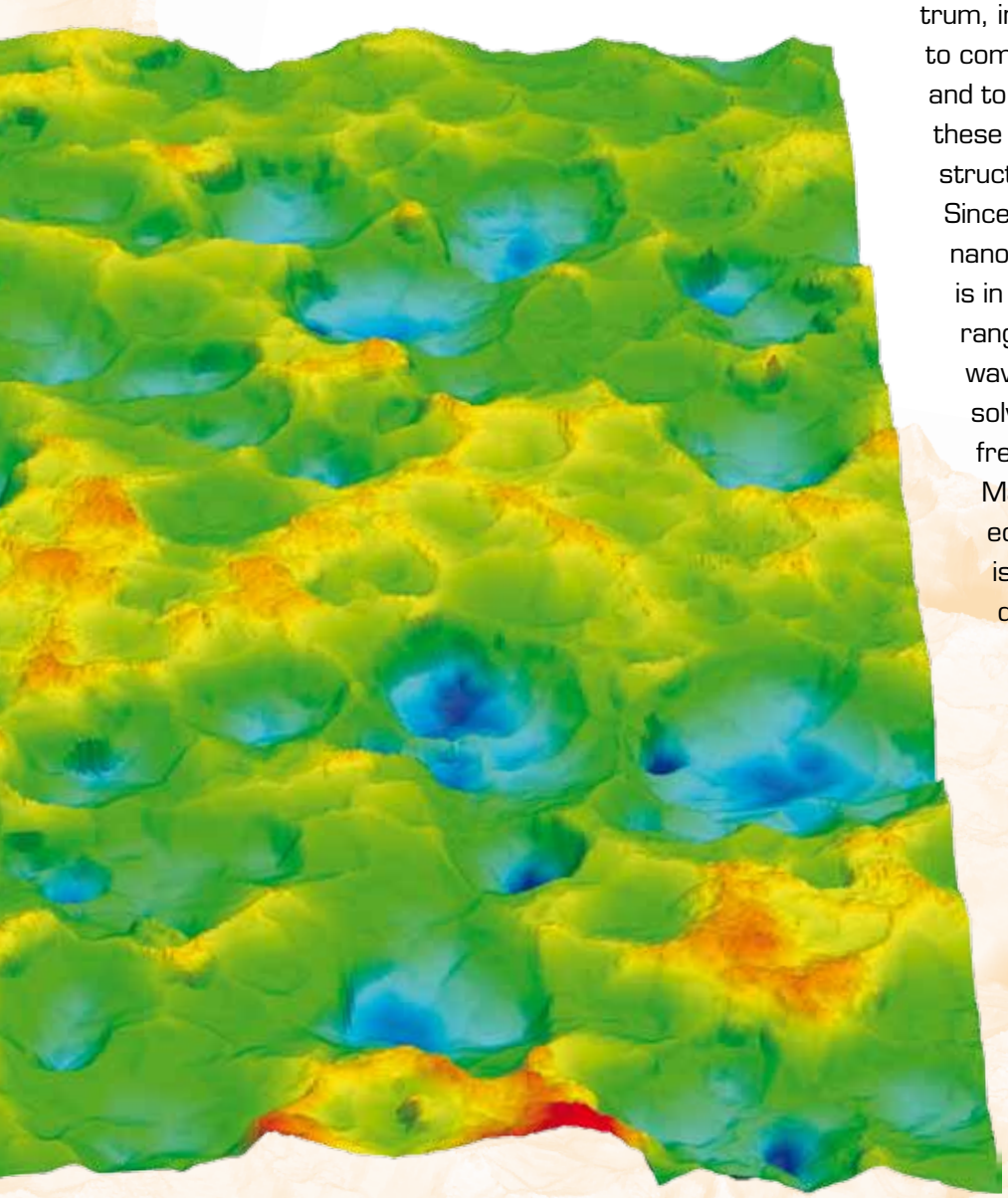


Figure 1: AFM scan of a TCO (ZnO) layer of size 5 µm x 5 µm obtained by magnetron sputtering and post deposition etching (Fraunhofer IST Braunschweig).

However, the efficiency of thin-film solar cells is still low compared to other technologies. It ranges from 3% for organic solar cells up to 8% for amorphous silicon (a-Si:H) solar cells with module size of 5.7 m² and ends at 19% for copper indium selenium (CIS) solar cells. However, only the resources of silicon on earth are high enough for a large scale installation on thin-film solar cells. Therefore, an important topic in research is to increase the efficiency of silicon based thin-film solar cells. This will significantly influence future technologies in the photovoltaic field.

The main reason of low efficiency of silicon based thin-film solar cells is the low absorption of the red light in the solar spectrum. This means that red light is not completely absorbed while passing the thin layer of silicon, but it is reflected at the silver back contact. Totally, this leads to a large amount of light which is reflected by the solar cell. Therefore, an important issue in designing new thin-film solar cells is to capture red light in the solar cell, such that it is mainly absorbed but not reflected. This so-called light trapping can be obtained by suitable nanostructures of the different solar cell layers. One concept is magnetron sputtering with post deposition etching (made by Fraunhofer IST Braunschweig, see Figure 1). However, there are many other concepts to optimize the light trapping in thin-film solar cells using suitable nanostructures. It is important to simulate the propagation of light in thin-film solar cells for the relevant solar spectrum, in order to compare and to optimize these nanostructures. Since the nanostructure is in the same range of wavelengths, solving high frequency Maxwell's equations is required:



$$i\omega\varepsilon\hat{E} = \nabla \times \hat{H} - \sigma\hat{E}, \quad i\omega\mu\hat{H} = -\nabla \times \hat{E}.$$

To this end, we developed simulation software based on the finite integration technique (FIT). As an example, let us consider the simulation of a microcrystalline silicon (μc-Si:H) based thin-film

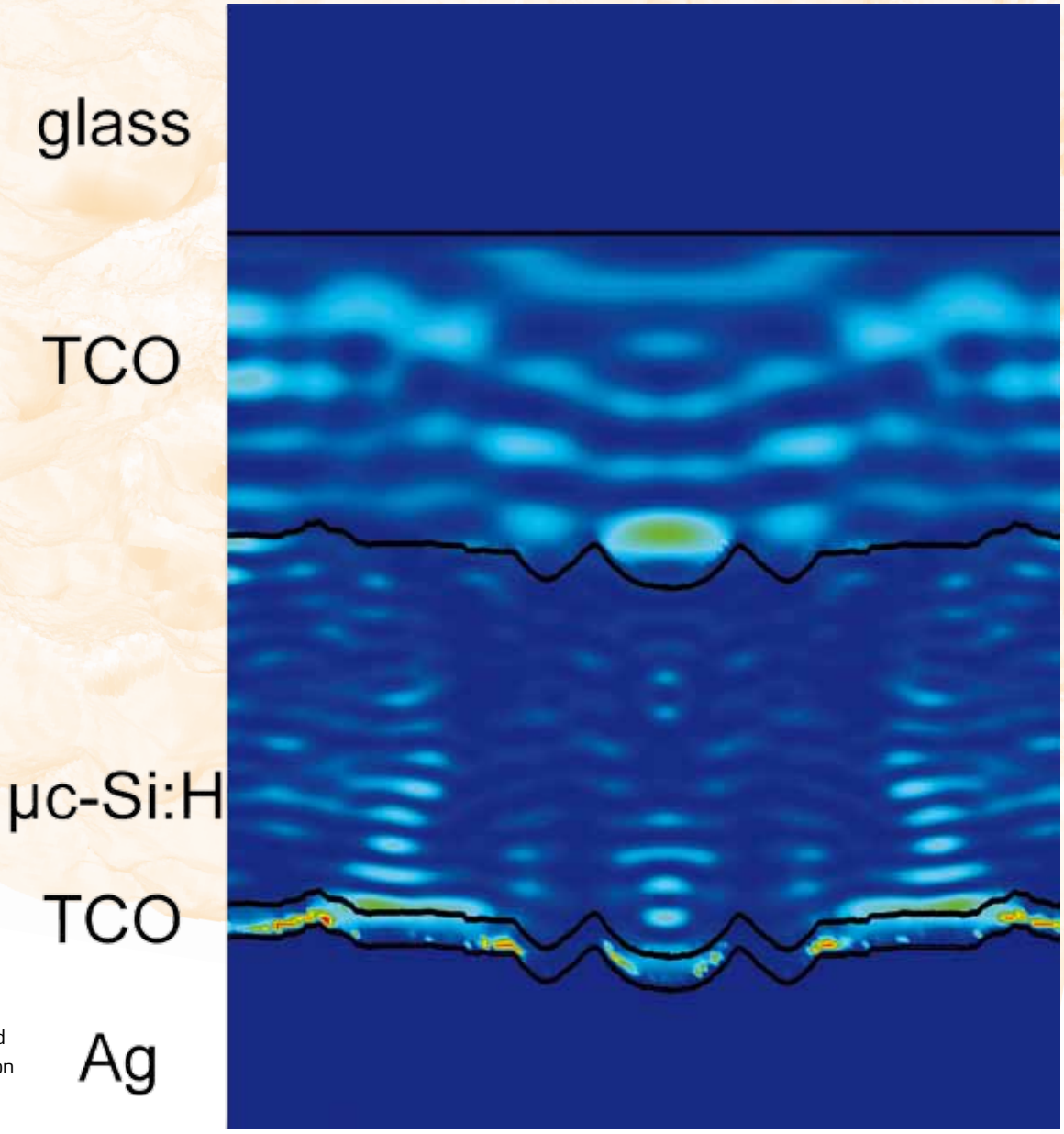


Figure 2: Slice of a 3-dimensional calculated electrical field distribution in a thin-film solar cell.

solar cell. Figure 2 depicts the calculated electrical field distribution in a thin-film solar cell structure. Since only small areas can be simulated, we assume boundary conditions like periodic boundary conditions to model the whole thin-film structure. The solar cell is based on a layer of glass, a layer of transparent conductive oxide (TCO), a layer of µc-Si:H, another layer of TCO and a layer of silver (Ag) at the back of the solar cell. The first layer of TCO and the silver back contact are the + and – contacts of the solar cell. Sunlight impinges at the TCO front contact and propagates through the solar cell. In case of a wavelength larger than 550 nm, absorption of light in µc-Si:H is so small that sunlight is not completely absorbed but reflected at the back contact. This leads to interference effects. These interference effects can be constructive or destructive, and can influence the optical efficiency of the solar cell. To simulate these effects in detail, an accurate calculation of the short-circuit current density is important. The short-circuit current density is calculated using the optical efficiency, which is obtained through solving Maxwell's equations and the solar spectral

irradiance. Only the relevant wavelengths of solar spectrum, which are highly absorbed by solar cell, have to be considered. Therefore, up to 160 simulations are required to calculate the optical efficiency from 300 nm to 1100 nm. This shows simulation of optical waves in thin-film solar cells is a challenging problem. In order to simulate light scattering from the random roughness of AFM scan data (see Figure 1), one has to simulate domains of size up to 5 µm x 5 µm x 3 µm. This requires discretization grids with more than 2.108 grid cells, where six unknowns have to be solved per each grid cell (three electric field components and three magnetic field components). This leads to a linear equation system with more than 109 unknowns. Unfortunately, fast iterative solvers like multigrid fail for high frequency Maxwell's equations. Therefore, we apply an iterative solver based on a time stepping method. This solver requires up to 104 iterations to solve the given linear equation system. Another difficulty is the accurate simulation of the optical properties of silver which has a negative permittivity ε. We have developed a new discretization method to address this issue. Our proposed approach avoids time consuming computations.

To reduce the long computational time for simulating the optical efficiency of thin-film solar cells, high performance computing is necessary. This can be done through domain decomposition of the structured discretization grid and by solving several Maxwell's equations in parallel. Table 1 depicts the performance of the current implementation of the Maxwell solver. Observe that the number of processors is chosen to be proportional to the number of grid cells. By the efficient parallelization of the solver, this leads to nearly the same computational time independent of the number of grid cells. To improve the performance of the implementation one has to implement strategies which circumvent the limitations of the storage bandwidth.

Acknowledgement

The author gratefully acknowledge funding from the Erlangen Graduate School of Advanced Optical Technologies (SAOT), providing AFM scan data by the research group Large Area Coating at Fraunhofer Institute for Surface Engineering and Thin Films IST Braunschweig and helpful discussions with Helmut Stiebig from company Malibu.

• Christoph Pflaum
• Christine Jandl
• Kai Hertel
• Zhabiz Rahimi

Erlangen Graduate School in Advanced Optical Technologies (SAOT)

Number of Processors	Number of Grid Cells in Mio.	Computational Time for 1,000 Iterations
8	28	14 min
16	54	13 min
32	109	14 min
64	219	15 min
128	426	13 min

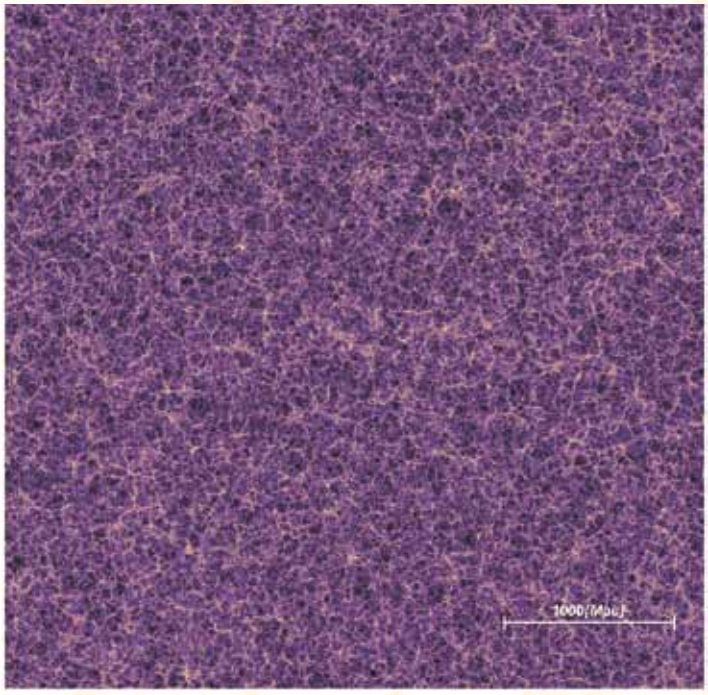
Table 1: Performance of the current implementation of the Maxwell solver on compute cluster machine TinyBlue in RRZE Erlangen (maximum: 672 processors).

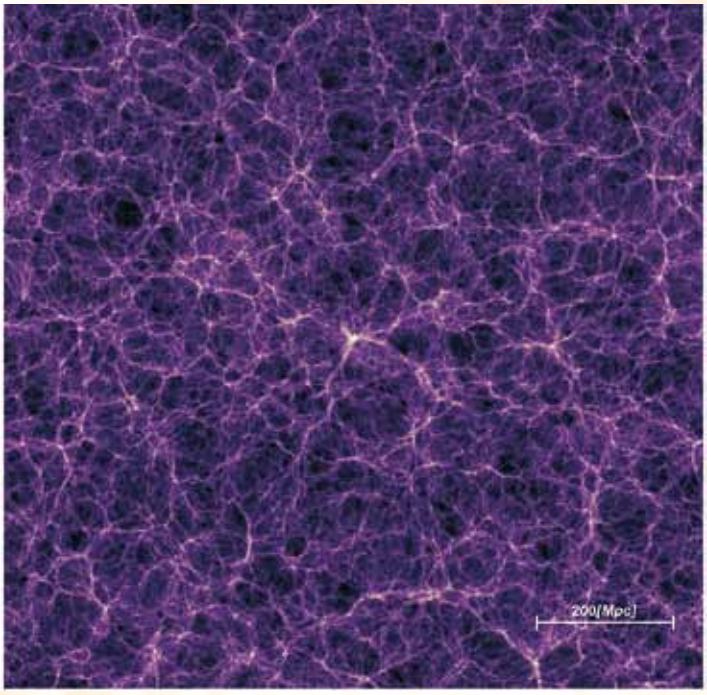
The Millennium-XXL Project: Simulating the Galaxy Population of dark Energy Universes

Modern cosmology as encoded in the leading ΛCDM model confronts astronomers with two major puzzles. One is that the main matter component in today's Universe appears to be a yet undiscovered elementary particle whose contribution to the cosmic density is more than 5 times that of ordinary baryonic matter. This cold dark matter (CDM) interacts extremely weakly with regular atoms and photons, so that gravity alone has affected its distribution since very early times. The other is a mysterious dark energy force field, which dominates the energy content of today's Universe and has led to its accelerated expansion in recent times. In the standard model, this component is described in terms of Einstein's cosmological constant ('Λ'). Uncovering the nature of dark energy has become one of the most actively pursued goals of observational cosmology.

In particular, the arrival of the largest galaxy surveys ever made is imminent, offering enormous scientific potential for new discoveries. Experiments like SDSSIII/BOSS or PanSTARRS have started to scan the sky with unprecedented detail, considerably improving the accuracy of existing cosmological probes. This will likely lead to challenges of the standard ΛCDM paradigm for cosmic structure formation, and perhaps even discover new physics.

One of the most important aims of these galaxy surveys is to shed light on the nature of the dark energy via measurements of the redshift evolution of its equation of state. However, the ability of these surveys to achieve this major scientific goal crucially depends on an accurate understanding of systematic effects and on a precise way





to physically model the observations, in particular the scale-dependent bias between luminous red galaxies and the underlying dark matter distribution, or the impact of mildly non-linear evolution on the so-called baryonic acoustic oscillations (BAOs) measured in the power spectrum of galaxy clustering.

Simulations of the galaxy formation process are arguably the most powerful technique to accurately quantify and understand these effects. However, this is an extremely tough computational problem, because it requires ultra-large volume N-body simulations with sufficient mass resolution to identify the halos likely to host the galaxies seen in the surveys, and a realistic model to populate these halos with galaxies. Given the significant investments involved in the ongoing galaxy surveys, it is imperative to tackle these numerical challenges to ensure that accurate theoretical predictions become available both to help to quantify and understand the systematic effects, and to extract the maximum amount of information from the observational data.

The State of the Art

The N-body method for the collisionless dynamics of dark matter is a long-established computational technique used to follow the growth of cosmic structure through gravitational instability. The Boltzmann-Vlasov system of equations is here discretized in terms of N fiducial simulation particles, whose motion is followed under their mutual gravitational forces in an expanding background space-time. While conceptually simple, calculating the long-range gravitational forces exactly represents an N^2-problem, which quickly becomes prohibitively expensive for interesting problem sizes. However, it is fortunately sufficient to calculate the forces approximately, for which a variety of algorithms have been developed over the years.

This allowed the sizes of cosmological simulations to steadily increase since the early 1980s, roughly doubling the particle number every 17 months and hence providing progressively more faithful models for the real Universe [1, 2, 3, 4]. Such simulations have proven

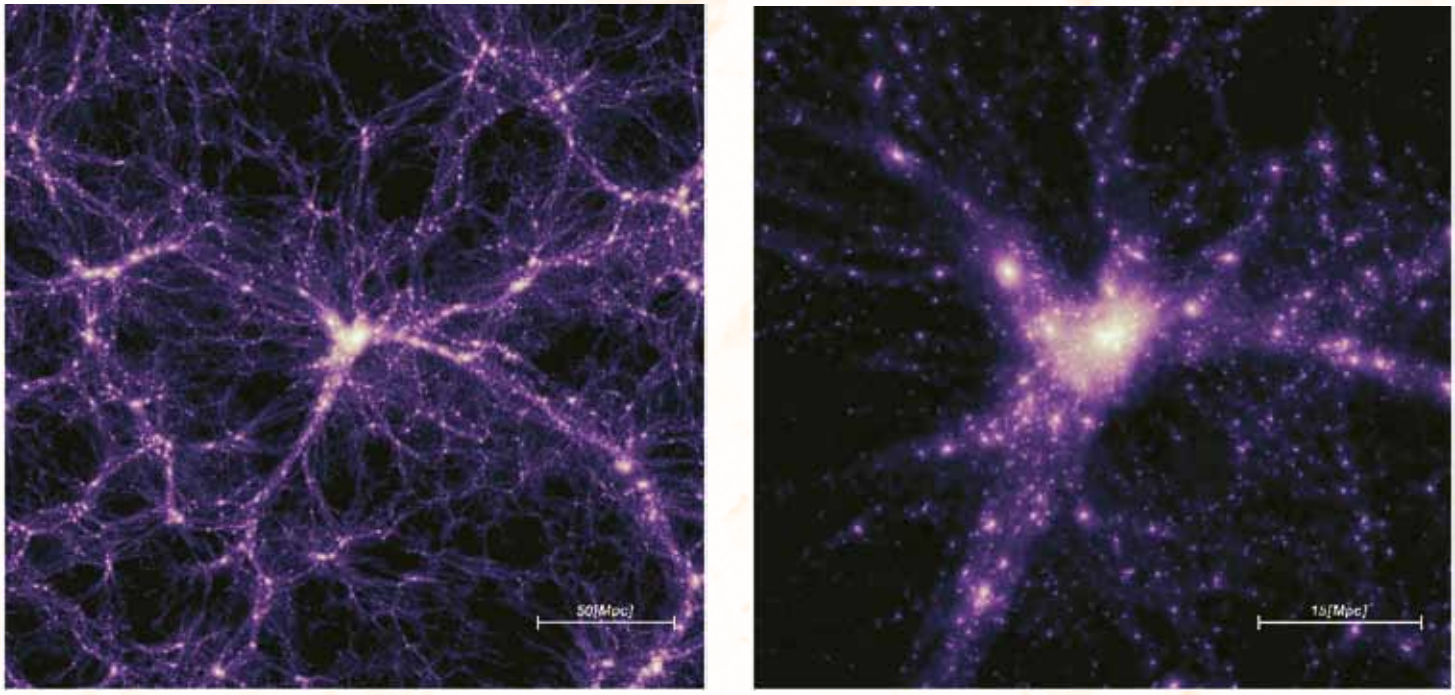


Figure 1: Dark matter distribution in the MXXL simulation on different scales. Each panel shows the projected density of dark matter in a slice of thickness 20 Mpc.

to be an indispensable tool to understand the low- and high-redshift Universe by comparing the predictions of CDM to observations, since these calculations are the only way to accurately calculate the outcome of non-linear cosmic structure formation.

A milestone in the development of cosmological simulations was set by the Millennium Run (MR), performed by our Virgo Consortium group in 2004 [3]. This simulation was the first, and for many years the only run with more than 10^{10} particles, exceeding the size of previous simulations by almost an order of magnitude. Its success was not only computational but most importantly scientific – more than 300 research articles in the fields of theoretical and observational cosmology have used the MR data-products since. The MR has an exquisite mass resolution and accuracy but, unfortunately, its volume is insufficient to get reliable statistics on large scales at the level needed for future surveys.

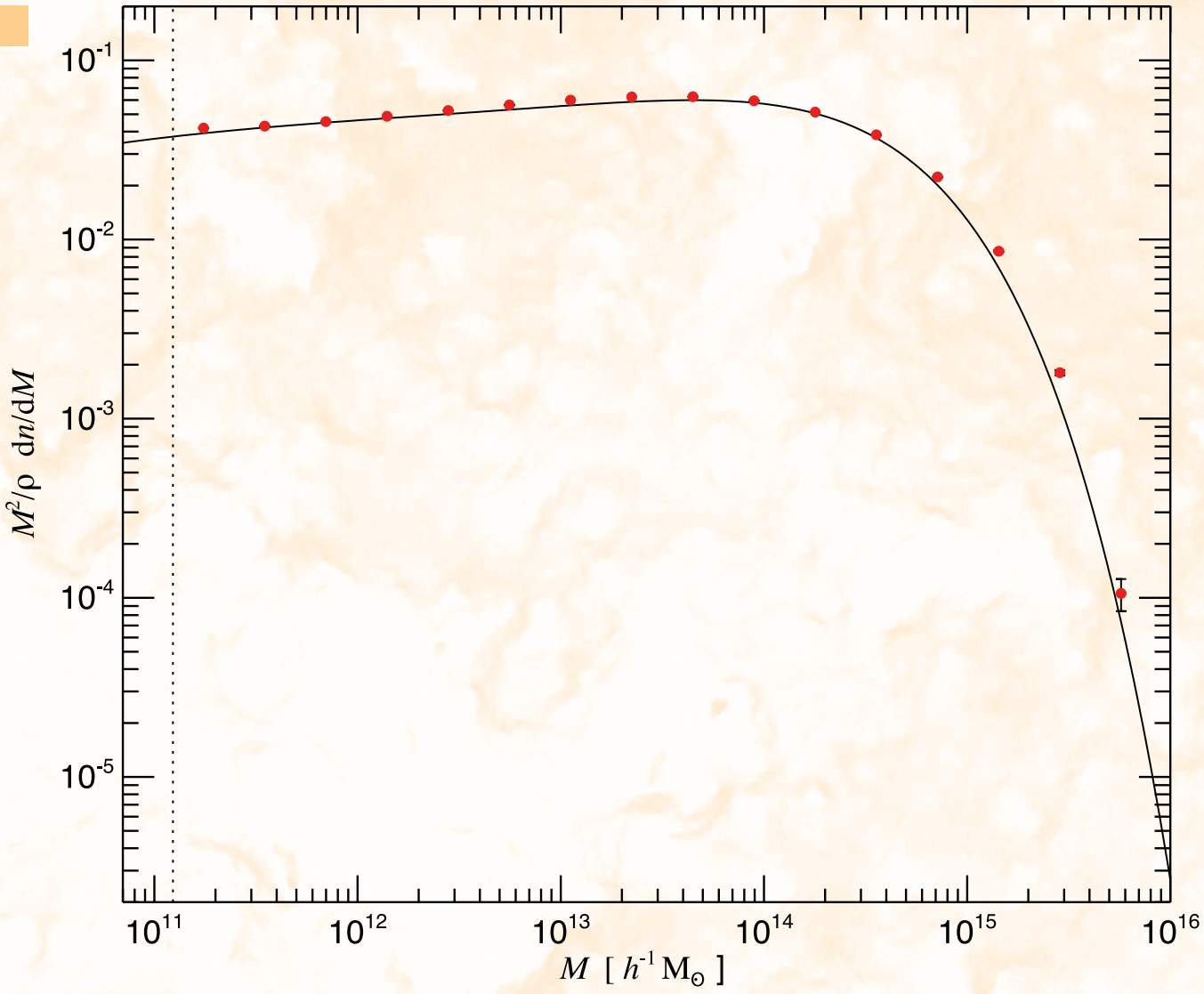


Figure 2: Differential halo abundance as a function of mass at the present epoch in the MXXL simulation. Note the good sampling far into the exponential tail. Apart from the last point, the error bars from counting statistics are smaller than the plotted symbols.

Recently, French and Korean collaborations [5, 6] have successfully carried out simulations containing almost 70 billion particles, but at considerably worse mass and spatial resolution than the MS, which did not allow them to robustly identify Milky-Way sized halos. Also, the need to manipulate and analyze a huge volume of data has proven to be a non-trivial challenge in working with simulations of this size, a fact that has made scientific exploitation of these projects difficult.

We have therefore set out to perform a new ultra-large N-body simulation of the hierarchical clustering of dark matter, featuring a new strategy for dealing with the data volume, and combining it with semi-analytical modelling of galaxy formation, which allows a prediction of all the luminous properties of the galaxies that form in the simulation. We designed the simulation project, dubbed Millennium-XXL (MXXL), to follow more than 303 billion particles (6720^3) in a cosmological box of size 4.2 Gpc across, resolving the cosmic web with an unprecedented combination of volume and resolution. While the particle mass of the MXXL is slightly worse than that of the MR, its resolution is sufficient to accurately measure dark matter merger histories for halos hosting luminous galaxies, within a volume more than 200 times larger than that of the MS. In this way the simulation can provide extremely accurate statistics of the large-scale structure of the Universe by resolving around 500 million galaxies at the present epoch, allowing for example highly detailed clustering studies based on galaxy or quasar samples selected in a variety of different ways. This comprehensive information is indispensable for the correct analysis of future observational datasets.

The computational Challenge

However, it was clear from the start that performing a simulation with these characteristics poses severe challenges, involving raw execution time, scalability of the algorithms employed, as well as their memory consumption and the disk space required for the output data. For example, simply storing the positions and velocities of the simulation particles in single precision consumes of order 10 TB of RAM memory. This figure, of course, is greatly enlarged by the extra memory required by the complex data structures and algorithms employed in the simulation code for the force calculation, domain decomposition, and halo and subhalo finding.

The code we used is a novel version of GADGET-3 we wrote specifically for the MXXL project. GADGET computes short-range gravitational forces with a hierarchical tree algorithm, and long-range forces with a FFT-based particle-mesh scheme [7]. Both the force computation and the time stepping of GADGET are fully adaptive. The code is written in highly portable C and uses a spatial domain decomposition to map different parts of the computational domain to individual processors. In 2005, we publicly released GADGET-2, which presently is the most widely employed code for cosmic structure formation.

Ultimately, memory requirements are the most serious limiting factor for cosmological simulations of the kind studied here. We have therefore put significant effort in developing algorithms and strategies that minimize memory consumption in our code, while at the same time retaining high integration accuracy and calculational speed. The special “lean” version of our code we developed for MXXL requires only 72 bytes per particle in the peak for the

ordinary dynamical evolution of MXXL. The sophisticated in-lined group and substructure finders add a further 26 bytes to the peak memory consumption. This means that the memory requirement for the target size of 303 billion particles of our simulation amounts to slightly more than 29 TB of RAM.

The JuRoPa machine at the Jülich Supercomputing Centre (JSC) appeared as one of the best suited supercomputers within Germany to fulfill our computing requirements, thanks to its available storage of 3 GB per compute core. Still, on JuRoPa the simulation demanded a partition of 1,536 nodes, each equipped with two quad-core X5570 processors and 24 GB of RAM, translating to 12,288 cores in total. This represents a substantial fraction (70%) of the whole supercomputer installation. In fact, normal user operation on such a large partition is still unusual and not done on a regular basis yet. It hence required substantial support on the side of the system administrators at JSC to nevertheless allow us carry out the MXXL production calculation on JuRoPa. In particular, severe problems with the memory usage of our simulation were encountered initially, as the code required essentially all the available physical memory of the compute nodes, leaving very little room for memory buffers allocated by the MPI library or the parallel Lustre filesystem attached to JuRoPa. Fortunately, with help from JSC and ParTec, the software maker behind the MPI software stack on JuRoPa, these problems could be overcome.

For the MXXL simulation, we decided to try a new hybrid parallelization scheme that we recently developed for our GADGET code. Instead of using 12,288 distributed-memory MPI tasks, we only employed one MPI task per processor (3,072 in total), exploiting the 4 cores of each socket via threads. This mixture of distributed and shared memory parallelism proved ideal to limit memory and work-load imbalance losses, because a smaller number of MPI tasks reduces the number of independent spatial domains in which the volume needs to be decomposed, as well as the data volume that needs to be shuffled around with MPI calls. The thread-based parallelization itself was partially done via explicit Posix thread programming, and partially via OpenMP directives for the simpler parallel constructs. The Posix-threads approach has allowed us to achieve effectively perfect parallelizability of the gravitational tree-walk in our code, which dominates the computational cost.

Still, the total CPU time required for the simulation was substantial. This originates from both, the huge number of particles and the very dissimilar time scales and matter densities involved in the problem. Very different density configurations are encountered in the clustered state produced by the simulation, and consequently there are structures featuring a large variety of timescales. Our code tackles this problem by using spatially and temporally adaptive time-stepping, so that short time-steps are used only when particles enter the localized dense regions where dynamical times are short. In addition, we use a new domain decomposition scheme that produces almost ideal scaling on massively parallel computers. In spite of these optimizations, the scope of the

computational problem remained huge. The final production run of MXXL required more than 86 trillion force calculations and took slightly more than 2.7 million CPU hours (~300 years), corresponding to about 9.3 days of runtime on 12,288 cores.

A significant fraction of 15% of the time was however spent for running our sophisticated on-the-fly postprocessing software, notably the group finding, the substructure finding, and the power spectrum calculation, and another 14% were needed for the I/O operations. The long-range force calculations based on 9216^3 sized FFTs consumed only about 3% of the time. We note that the parallel "friends-of-friends" group finding for the 303 billion particles at the final output time took just 470 seconds, which we think is a remarkable achievement. Determining all gravitational bound substructures within these halos took a further 2,600 seconds. It is also during this substructure finding that the peak memory consumption of the code is reached.

Another demanding aspect of the MXXL project lies in the data handling. The production code is able to write data parallel to disk. Despite the huge size of our data sets, the restart files, amounting to 10 TB in total, could be written/read to disk in about 40 minutes on the JuRoPa system. However, of larger practical importance for us is the aggregated size of the simulation outputs produced by the simulation. If an approach similar to the MR was taken for the MXXL project, then the disk space requirement would approach 700 TB for the storage of the dark matter particle positions, velocities and groups catalogues, a prohibitively large amount of data. To cope with this problem, we have therefore carried out a significant part of the post-processing on-the-fly during the simulation, avoiding the need to store the particle phase-space variables for most of the output times on disk. This postprocessing included the identification and characterization of halos and subhalos, and the calculation of the dark matter density field and its time derivative. As a result, the total disk space required for the data products of the simulation could be shrunk to about 100 TB.

First Results and Outlook

In Figure 1, we show the projected density field of the MXXL simulation in a series of thin slices through the simulation volume. This gives a hint about the enormous statistical resolving power of the simulation, and its large dynamic range everywhere in the simulation cube. The gravitational spatial resolution of the simulation translates to a dynamic range of 300,000 per dimension, or formally more than 10^{16} resolution elements in 3D.

We have found 650 million halos at redshift zero in the simulation, binding 44% of all particles in non-linear structures. In total there are more than 25 billion halos in the group catalogues that we produced and that are linked together in our merger trees. In Figure 2, we show the halo mass function at the present epoch, which is one of the most basic and important simulation predictions, describing the abundance of objects as a function of mass and epoch.

The largest cluster of galaxies at z=0 has a mass of 9 x 10^{15} solar masses. Such extreme objects are so rare that they can only be found in volumes as large as that of the MXXL. In Figure 3, we show a map of the time derivative of the gravitational potential of the simulation at redshift z=12. We produced such maps on-the-fly during the simulation in order to study the integrated Sachs-Wolfe effect, which describes distortions in the cosmic microwave background by foreground structures.

The full analysis of the MXXL in terms of its predicted galaxy distributions has just begun, and will still take some time to complete. We can, for example, use the MXXL to predict the spatial distribution of galaxies in the past light-cone of a fiducial observer. Without any replication of the box, we will be able to construct all-sky synthetic galaxy catalogues up to redshift 0.57. These mock observations are crucial for future surveys to study cluster finding algorithms, the influence of projection effects, the evolution of the clustering along the line-of-sight, and most importantly,

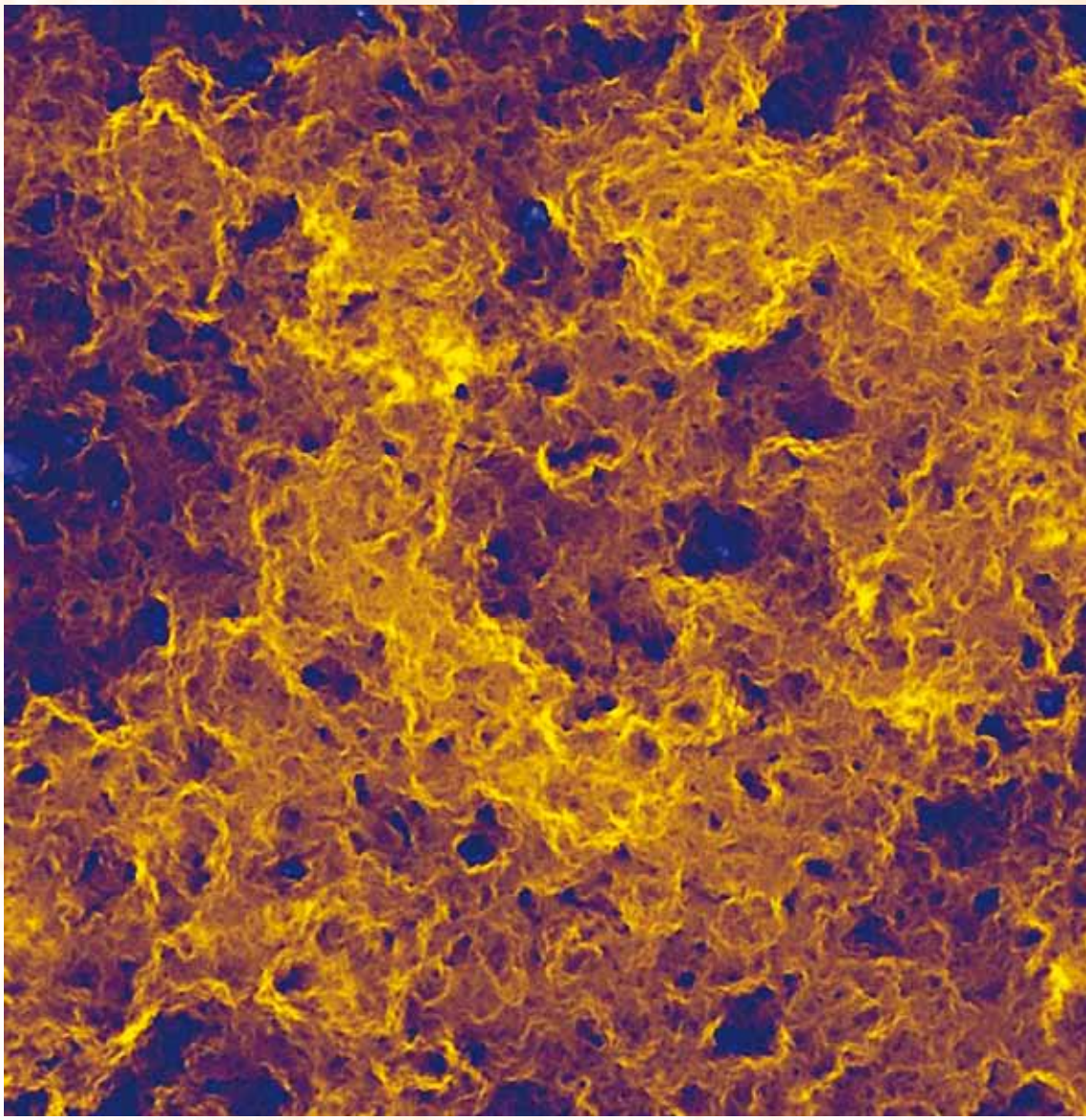
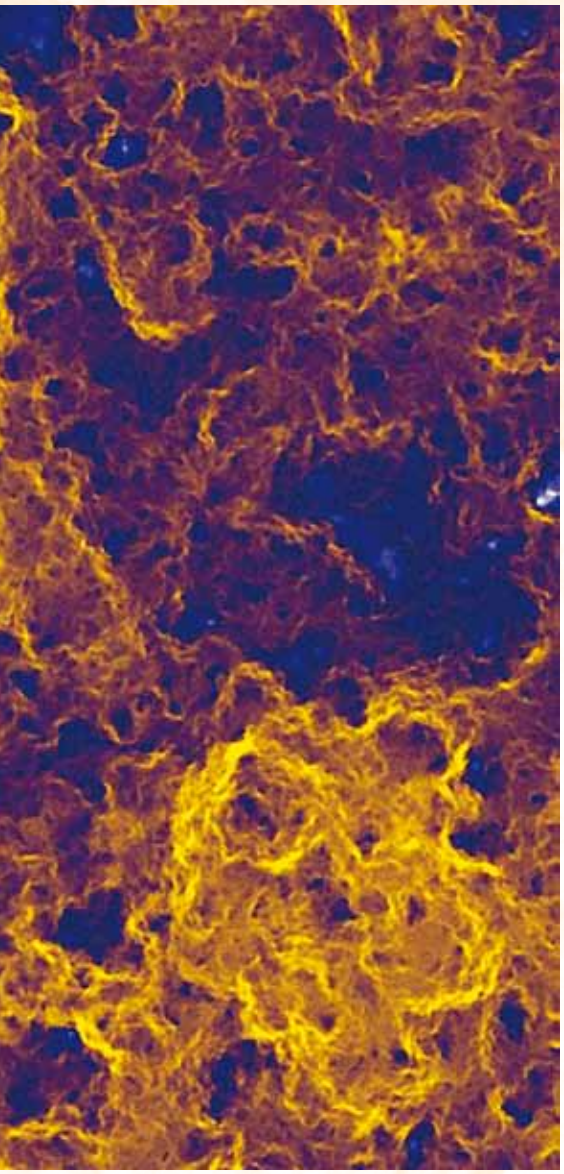


Figure 3: Map of the time-derivative of the gravitational potential at z=12 in the MXXL simulation. Such maps can be used to calculate the integrated Sachs-Wolfe effect along the backwards light-cone to the last scattering surface of the cosmic microwave background.

the systematic effects in the cosmological constraints derived from BAO, weak lensing, or cluster counts. Using a novel scaling technique we just developed [8], it is furthermore possible to cast the simulation results into theoretical models for the galaxy distribution that cover a significant range of cosmologies.

The Millennium XXL is by far the largest cosmological simulation ever performed and the first multi-hundred billion particle run. The scope of the simulation project has pushed the envelope of what is feasible on current world-class HPC resources, but the expected rich scientific return from the project makes it well worth the effort. At the same time, the successful scaling of the cosmological code to well beyond ten-thousand cores is an encouraging sign for computational research in cosmology, which will address yet more demanding problems on future HPC machines.



References

[1] **Davis, M., Efstathiou, G., Frenk, C. S., and White, S. D. M.**
The evolution of large-scale structure in a universe dominated by cold dark matter,1985, Astrophysical Journal, 292, 371

[2] **Navarro, J. F., Frenk, C. S., White, S. D. M.**
1996, The Structure of Cold Dark Matter Halos, Astrophysical Journal, 462, 563

[3] **Springel, V., White, S. D. M., Jenkins, A. et al.**
2005, Simulations of the formation, evolution and clustering of galaxies and quasars, Nature, 435, 629

[4] **Springel, V., Wang, J., Vogelsberger, M., et al.**
2008, The Aquarius Project: the subhaloes of galactic haloes, Monthly Notices of the Royal Astronomical Society, 391, 1685

[5] **Springel, V.**
The cosmological simulation code GADGET-2, 2005, Monthly Notices of the Royal Astronomical Society, 364, 1105

[6] **Teyssier, R., Pires, S., Prunet, S., et al.**
2009, Full-sky weak-lensing simulation with 70 billion particles, Astroparticle Physics, 497, 335

[7] **Kim, J., Park, C., Gott, J. R., Dubinski, J.**
2009, The Horizon Run N-Body Simulation: Baryon Acoustic Oscillations and Topology of Large-scale Structure of the Universe, Astrophysical Journal, 701, 1547

[8] **Angulo, R. E., White, S. D. M.**
2010, One simulation to fit them all - changing the background parameters of a cosmological N-body simulation, Monthly Notices of the Royal Astronomical Society, 405, 143

- Volker Springel[1]
- Raul Angulo, Simon D.M. White[2]
- Adrian Jenkins, Carlos S. Frenk, Carlton Baugh, Shaun Cole[3]

[1] Heidelberg Institute for Theoretical Studies, Heidelberg

[2] Max-Planck-Institute for Astrophysics, Garching, Germany

[3] Institute for Computational Cosmology, University of Durham, United Kingdom

Exotic State in Correlated Relativistic Electrons

Applications

In the recent past condensed matter systems displayed a number of exotic states like unconventional superconductivity in high-temperature superconductors, supersolidity in ^{4}He, and spin liquids in frustrated mangnets, that emerge due to correlations in many-body systems [1]. Of particular interest are spin liquids, where quantum fluctuations preclude ordering from the liquid state to a solid, i.e. to a periodic ordering of the magnetic moments, even at zero absolute temperature. Generally it is expected, that such strong quantum fluctuations arise due to competing interactions that frustrate the formation of an ordered state [1].

A system where quantum fluctuations may play a dominant role is graphene that was recently obtained experimentally by exfoliation of graphite [2]. By means of such a micromechanical cleavage, single layers of carbon atoms with a honeycomb structure, schematically displayed in Figure 1, are produced, that, remarkably, subsist as free-standing two-dimensional crystals. The honeycomb lattice is a bipartite one, i.e. it consists of two sublattices, where the nearest neighbours of the sites of one of them are sites belonging to the other sublattice. Hence, in such a lattice geometric frustration is absent, since e.g. antiferromagnetic order is possible by placing magnetic moments pointing in one direction on one sublattice and pointing in the opposite one on the other sublattice. However, due to the fact that the honeycomb lattice has the smallest coordination number in two dimensions, the effect of quantum fluctuations is the strongest.

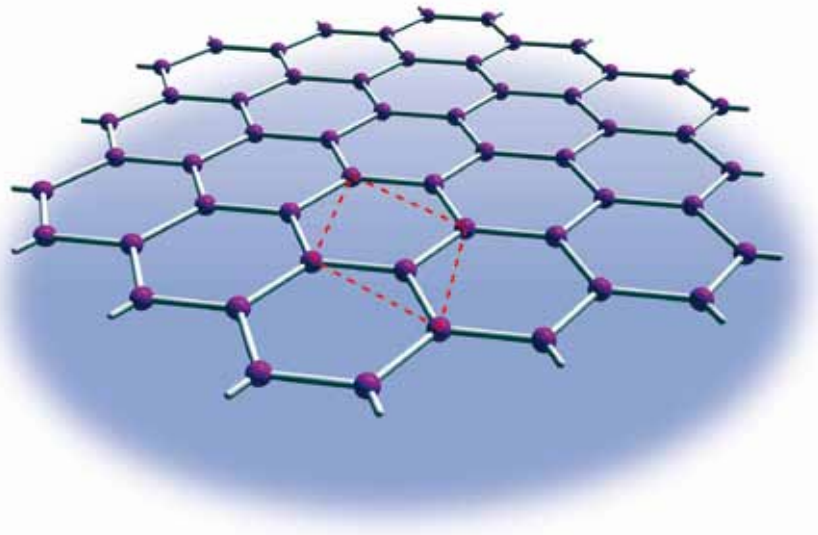


Figure 1: The honeycomb lattice with its unit cell, displayed by the dashed red line. The unit cell contains two atoms, each one belonging to a different sublattice. Therefore, the honeycomb lattice has a bipartite structure, where the atoms of a given sublattice are surrounded by sites of the other one.

A further remarkable feature of the honeycomb lattice appears, when electrons are placed on it. Assuming for the moment non-interacting electrons, the corresponding band structure is as shown in Figure 2. There, it can be seen in the first place, that the electron (particle) states and the hole (antiparticle) states are symmetrically placed around the zero of energy, such that particle-hole (charge conjugation) symmetry is present. The zero of energy corresponds to the Fermi energy when the average density of electrons is unity per lattice site, i.e. for a half-filled band. Moreover, the low energy states around particular points in the two-dimensional Brillouin zone display a relativistic dispersion and can be readily described by Dirac's equation [2]. We therefore, will refer to those points as Dirac points in the following. The relativistic dispersion in the low energy sector at half-filling leads to a vanishing density of states

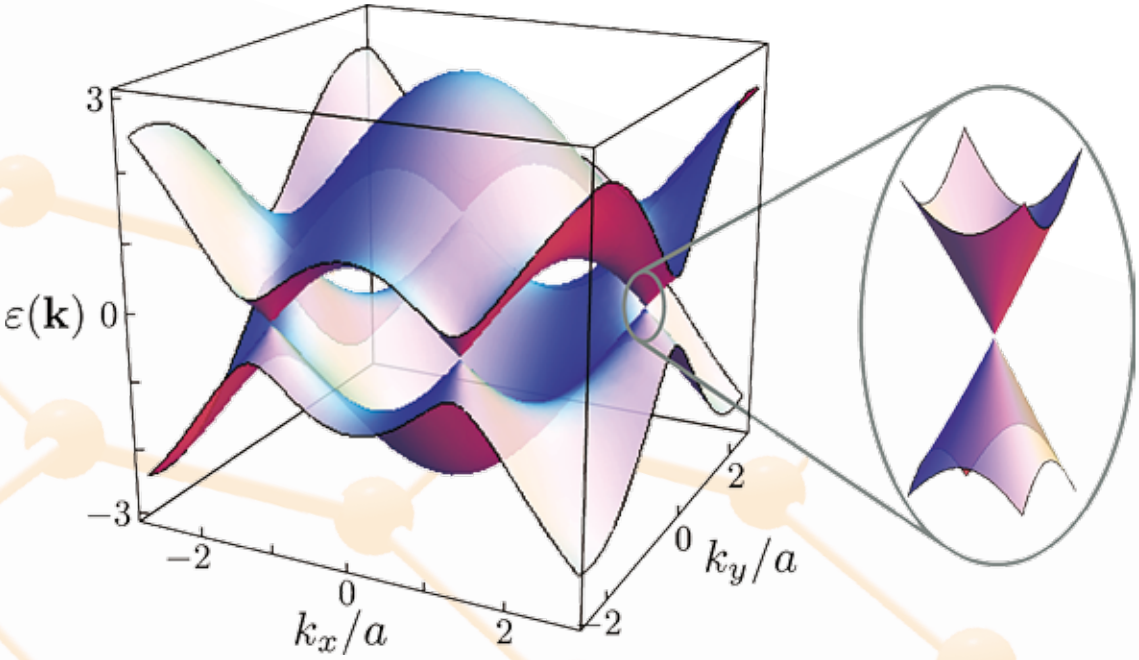


Figure 2: Energy dispersion of the electronic states for non-interacting electrons on a honeycomb lattice. Close to the zero of energy the dispersion corresponds to the one characteristic of relativistic fermions, as displayed by the portion zoomed in.

at the Fermi energy. Therefore, a finite strength for interactions promoting a spontaneous symmetry breaking is necessary, enhancing thus, the role of fluctuations.

In order to study the effects of correlations in electrons on the honeycomb lattice in its most basic form, we consider the Hubbard model, where only an on-site interaction, termed *U*, is present. Such a model is a paradigm for strongly correlated electrons, as in the case of high temperature superconductors [3], and gives an accurate description of ultra-cold fermionic atoms in optical lattices [4,5]. For large values of the repulsive interaction *U*, and at half-filling, the ground-state corresponds to a Mott-insulator, i.e. an insulating state due to interactions in contrast to the metallic state of the noninteracting system. Furthermore, in this limit antiferromagnetic correlations dominate due to Pauli's exclusion principle and the necessity of gaining kinetic energy. Therefore, on a bipartite lattice, they will lead to an antiferromagnetically ordered ground-state. However, as the interaction strength diminishes, a competition between the tendency to order and quantum fluctuations will set in, so that a detailed analysis of correlations is needed to characterize the possible phases.

Needless to say, an unbiased study as delineated above is only possible by numerical means. Among the different methods, Quantum Monte Carlo (QMC) simulations are the most appropriate one since a careful extrapolation to the thermodynamic limit, in this case in two dimensions, is mandatory to determine whether a spontaneous symmetry breaking has taken place. We implemented a projective (temperature $T = 0$) determinantal QMC algorithm in the canonical ensemble that is free of the sign-problem at half-filling [6]. This algorithm allows the calculation of the expectation value of any physical observable in the ground-state by performing an imaginary time evolution of a trial wave function that is required to be nonorthogonal to the ground-state. The value θ reached in the imaginary time evolution corresponds to a projection parameter [6]. For a spin-singlet trial wave function,

we found $\theta=40/t$ to be sufficient to obtain converged ground-state quantities within statistical uncertainty. In the presented simulations, we used a finite imaginary time step $\Delta\tau=0.05/t$. We verified by extrapolating $\Delta\tau\to 0$ that this finite imaginary time step produces no artefacts. The phases described in the following were determined by a finite-size extrapolation to the thermodynamic limit with lattices of $N = 2L^2$ sites with periodic boundary conditions, and linear sizes L in terms of the unit cell containing two sites, with $L \leq 18$. L was taken as a multiple of 3 in order to be able to include the Dirac points in our Brillouin zones, such that the low energy physics is correctly represented. In terms of a simulation of a classical Ising model, in our case with long-range interactions, the largest systems sizes correspond to a lattice with 518,400 sites.

A first insight in the possible phases of the system is obtained by considering the single-particle excitation gap $\Delta_{sp}(\vec{k})$ that we extracted from the imaginary-time displaced Green function (see Ref. [7] for details). $\Delta_{sp}(\vec{k})$ gives the minimal energy necessary to extract one fermion from the system, and corresponds to the gap that can be observed in photoemission experiments. As shown in Figure 3, $\Delta_{sp}(K) = 0$ for $U < U_c \approx 3.6t$, where t is the hopping amplitude in the Hubbard model. The vanishing gap corresponds to a metal, that is commonly called a semimetal (SM) due to the fact that the Fermi surface is in this case reduced to a point. Beyond U_c, the system enters into an insulating phase due to interactions, and hence, as expected for large values of U, the system becomes a Mott-insulator. The values of the gap are obtained via an extrapolation of the QMC data to the thermodynamic limit with energies given in units of t [7].

As explained above, for values large enough of U, one expects long-range antiferromagentic (AF) correlations. We therefore measured the AF spin structure-factor S_{AF} that reveals long-range AF order if $\lim_{N\to\infty} S_{AF}/N > 0$. The results of a finite-size extrapolation are also presented in the phase diagram of Figure 3. AF order appears beyond $U/t \approx 4.3$. Hence, contrary to the usual expectation for a bipartite lattice, AF long-range order sets in later than the insulating phase, leaving an extended window $3.6 \lesssim U/t \lesssim 4.3$, within which the system is neither a semimetal, nor an AF Mott-insulator.

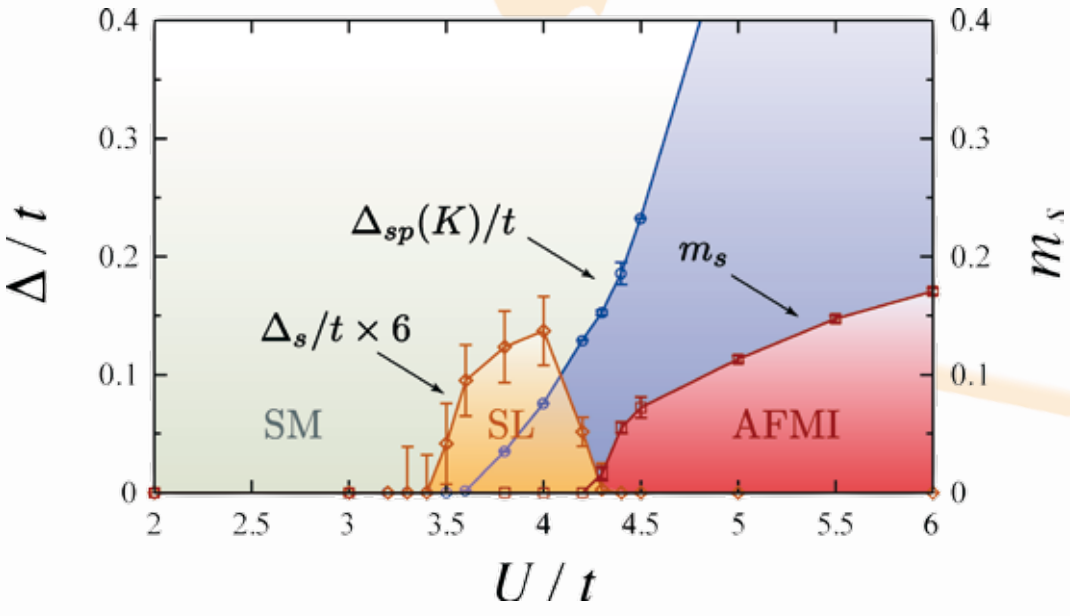


Figure 3: Phase diagram for the Hubbard model on the honeycomb lattice at half-filling. The semimetal (SM) and the antiferromagnetic Mott insulator (AFMI) are separated by a gapped spin liquid (SL) phase in an intermediate coupling regime. $\Delta_{sp}(K)$ denotes the single-particle gap at one of the Dirac points (K), and Δ_s the spin gap. m_s denotes the staggered magnetization whose saturation value is 1/2.

Further details on the nature of this intermediate region are obtained by examining the spin excitation gap, extracted from the long-time behaviour of the imaginary-time displaced spin-spin correlation function [7]. We consider first the spin gap Δ_s in the staggered sector at k = 0, which vanishes inside the AF phase due to the emergence of two Goldstone modes, as well as in the gapless metallic phase. Figure 4 shows finite size estimates of Δ_s for different values of U/t, along with an extrapolation to the thermodynamic limit. A finite value of Δ_s persists within an intermediate parameter regime $3.5 \lesssim U/t \lesssim 4.3$, while it vanishes both within the metallic and the AF phase. We also calculated the uniform spin gap Δ_u by extrapolating the spin gap observed at the smallest finite k-vector on each cluster to the thermodynamic limit. Δ_u is found to be even larger than Δ_s inside the intermediate region (e.g. $\Delta_u = 0.101(8)$ at $U/t = 4$), and vanishes in the metallic and the AF phase [7].

The observation of a finite spin gap rules out gapless phases such as triplet superconductivity as well as quantum spin Hall states. The remaining possibilities can be enumerated by considering the coupling to order parameters that lead to the opening of a mass gap in Dirac fermions, and hence to account for the single-particle gap observed in the QMC data: (i) singlet superconductivity, (ii) a quantum Hall state (QHS), (iii) charge density wave (CDW) order, and (iv) a valence bond crystal (VBC). Our QMC results exclude all those states, as discussed below. Thereby, the intermediate phase is genuinely an exotic state of matter since it cannot be understood at the single particle level within a mean-field theory with a local order-parameter.

Furthermore, since no spontaneous symmetry-breaking is observed, while a spin gap is present, it corresponds to a spin liquid state.

In order to assess, if superconductivity arises in the vicinity of the Mott-transition, we used the method of flux quantization which probes the superfluid density and is hence independent of the specific symmetry of the pair wave function [7]. Let Φ be a magnetic flux traversing the centre of a torus on which the electronic system lies and $E_0(\Phi/\Phi_0)$ the total ground state energy, Φ_0 being the flux quantum. A superconducting state of Cooper pairs is present if in the thermodynamic limit, the macroscopic energy difference $E_0(\Phi/\Phi_0)-E_0(\Phi/\Phi_0 = 1/2)$ is a function with period 1/2. In contrast, a metallic (insulating) phase is characterized by an (exponential) vanishing of $E_0(\Phi/\Phi_0)-E_0(\Phi/\Phi_0 = 1/2)$ as a function of system size. The QMC data is consistent with the vanishing of this quantity in the thermodynamic limit. In addition, we measured singlet superconducting order parameters of (extended) *s*-, *p*-, and *f*-wave symmetry, which turn out to all vanish in the thermodynamic limit [7].

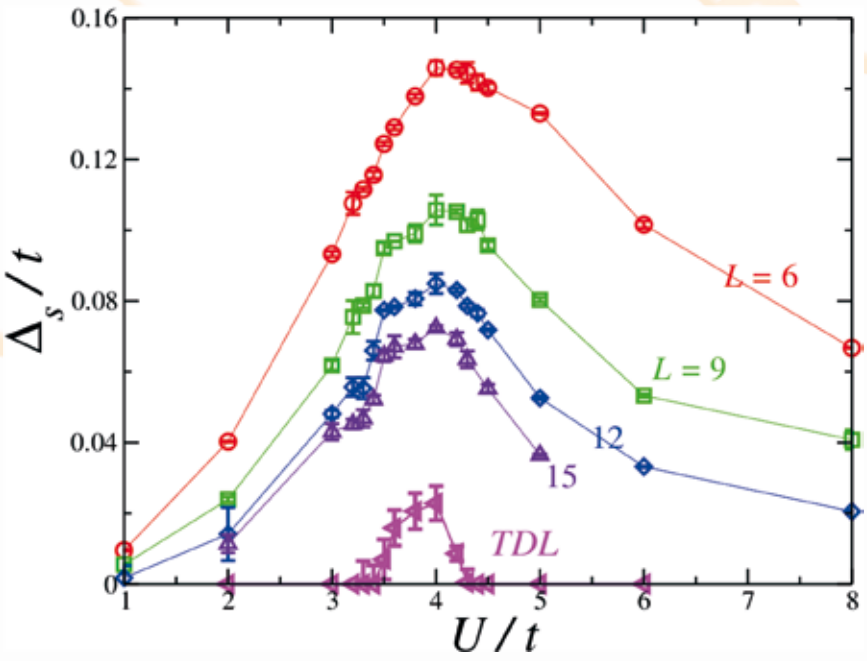


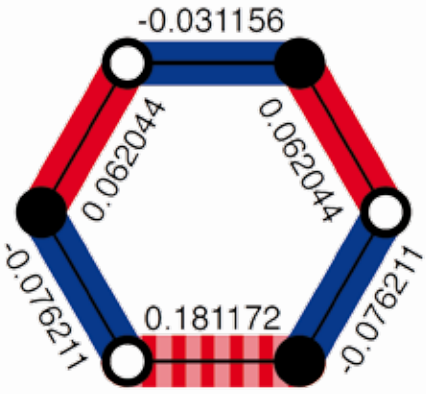
Figure 4: Spin gap in units of t as a function of U/t for various systems sizes. The lowest curve corresponds to the extrapolation to the thermodynamic limit (TDL).

Hence, both flux quantization as well as a direct measurement of pairing correlations in various symmetry sectors lead to no sign of superconductivity.

Both the CDW and QHS trigger a breaking of the sub-lattice symmetry and thereby open a mass gap at the meanfield level. A detailed analysis of the charge-charge correlation functions rules out a CDW. Furthermore, we have found no signature for the presence of (spin) currents in the ground-state. This rules out the breaking of sublattice and time reversal symmetries as required for the QHS [7].

To examine the occurrence of a VBC, we probe for dimer-dimer correlations between a dimer formed by nearest neighbour sites <*ij*> and a distant bond formed by sites <*kl*> [7]. We have found no VBC, neither in the charge, nor in the spin sector. The left side of Figure 4 shows the results of this measurement in the spin sector, i.e. the correlation between singlet dimers at $U/t = 4.0$. The striped bond is the one with respect to which correlations were determined. They are found to be short-ranged, and consistent with the dominance of a resonating valence bond (RVB) state within the hexagons of the honeycomb lattice. This can be seen by comparing the singlet-correlations with those of an isolated hexagon (right side of Figure 4), the classical example of the resonance phenomenon in conjugated π-electrons [8]. Accordingly, we find no long-ranged order from the dimer-dimer structure factors in Fourier space. Our results thus reveal a genuinely exotic state of matter, where no spontaneous symmetry-breaking is observed, while a spin gap is present. It corresponds to a spin liquid RVB state in the intermediate coupling regime in the vicinity of the Mott-transition.

The presence of a spin liquid in the Hubbard model on the honeycomb lattice close to an antiferromagnetic Mott-insulator was hitherto unexpected, due to the bipartite nature of the honeycomb lattice, and hence, the absence of frustration. However, our results indicate that strong enough fluctuations, that develop close to the quantum critical point where AF order sets in, lead to such an exotic state of matter. It could be expected, that such fluctuations would promote some broken symmetry states like superconductivity. However, the vanishing density of states at the Fermi energy



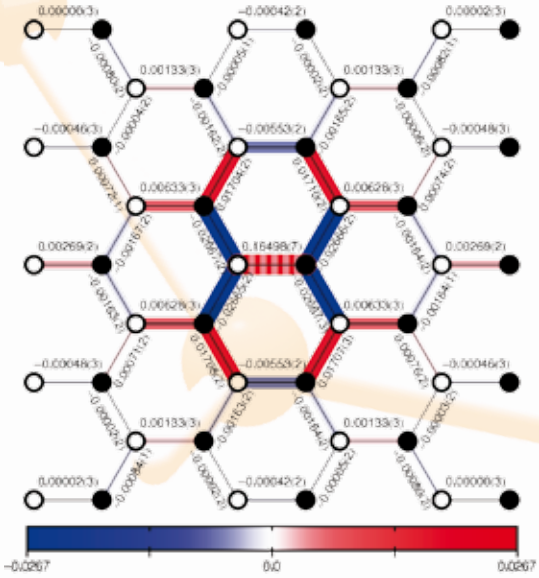


Figure 5: Real space plot of the spin dimer-dimer correlations. Right side: the dimer-dimer correlation function in the spin-channel for a $L = 6$ system at $U/t = 4$. Left side: the same correlation for the isolated Hubbard hexagon also at $U/t = 4$. The reference bonds are dressed with stripes. Numbers in parenthesis indicate the standard error of the last digit.

may be responsible for its absence, since in this case, a finite coupling strength is needed, at least in the BCS-frame.

Having an unexpected realization of a short-range RVB state, it would be highly interesting to explore the consequences of doping, in a spirit rather close to the original scenario proposed by Anderson [3] and Kivelson *et al.* [9] for the cuprates. In particular, for the fully gapped short-range RVB state, the finite spin gap sets the energy scale of pairing in the superconducting state [9]. In this respect, the value obtained for the spin gap is rather promizing. The largest value attained is $\Delta_s \sim 0.025t$ (Fig. 1), that for t in the range of 1.5 to 2.5 eV (in graphene is $t = 2.8$ eV [2]) corresponds to a temperature scale ranging from 400 to 700 K.

Although studies of doping are beyond the power of our quantum Monte Carlo approach due to the sign problem, they could open interesting perspectives e.g. in future experiments with ultra-cold atoms on a honeycomb optical lattice, or with honeycomb lattices based on group IV elements like expanded graphene (to enhance the ratio U/t) or *Si*, where the nearest neighbour distance is expected to be approximately 50% larger than in graphene [10], such that correlations effects are enhanced.

Acknowledgments

We thank L. Balents, S. Capponi, A. H. Castro Neto, A. Georges, M. Hermele, A. L¨auchli, E. Molinari, Y. Motome, S. Sachdev, K. P. Schmidt and S. Sorella for discussions. We are grateful to S. A. Kivelson for thoroughly reading our manuscript and providing important suggestions. F.F.A. is grateful to the KITP Santa Barbara for hospitality and acknowledges support by the DFG through AS120/4 and FG1162. A.M. thanks the Aspen Center for Physics for hospitality and acknowledges partial support by the DFG through SFB/TRR21. S. W. acknowledges support by the DFG through SFB/TRR21 and WE3649. We thank NIC Jülich, HLRS Stuttgart, the BW Grid, and the LRZ München for the allocation of CPU time.

References

[1] **Nature Insight**
Exotic Matter, Nature 464, 175, 2010

[2] **Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S. and Geim, A. K.**
Rev. Mod. Phys. 81, 109, 2009

[3] **Anderson, P. W.**
Science 235, 1196, 1987

[4] **Jördens, R., Strohmaier, N., Günter, K., Moritz, H. and Esslinger, T.**
Nature 455, 204, 2008

[5] **Schneider, U., Hackermüller, L., Will, S., Best, T., Bloch, I., Costi, T. A., Helmes, R. W., Rasch, D. and Rosch, A.**
Science 322, 1520, 2008

[6] **Assaad, F. F. and Evertz, H. G.**
Computational Many-Particle Physics, Lecture Notes in Physics, p. 739, Springer-Verlag, Berlin, 2008

[7] **Meng, Z.-Y., Lang, T., Wessel, S., Assaad, F. F. and Muramatsu, A.**
Nature 464, 847, 2010

[8] **Pauling, L. C.**
The Nature of the Chemical Bond and the Structure of Molecules and Crystals: an Introduction to Modern Structural Chemistry, Cornell University Press, 20th Edition, Ithaca, New York, USA, 1986

[9] **Kivelson, S. A., Rokhsar, D. S. and Sethna, J. P.**
Phys. Rev. B 35, 8865, 1987

[10] **Cahangirov, S., Topsakal, M., Aktürk, E., Sahin, H. and Ciraci, S.**
Phys. Rev. Lett. 102, 236804, 2009

• Zi Yang Meng[1]
• Thomas C. Lang[2]
• Stefan Wessel[1]
• Fakher F. Assaad[2]
• Alejandro Muramatsu[1]

[1] Institut für Theoretische Physik III, Universität Stuttgart, Germany

[2] Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Germany

Combining MPI with OpenMP Parallelism in large Scale QCD Simulations

Combining MPI with OpenMP Parallelism in QCD Simulations

Current HPC systems are equipped with nodes based on multi-core processors which contain two or more distinct cores in the same physical package. Parallel programming of a node can take advantage of its shared memory architecture, where compilers usually support multi-threaded execution of a program: loop level parallelism can be exploited by using compiler directives such as those defined in the OpenMP standard. Inter-node communication is typically implemented with the Message Passing Interface (MPI). MPI and OpenMP can be combined in order to exploit parallelism beyond a single level by using MPI on the coarse-grain level and OpenMP on the fine-grain level. As a consequence users are faced with the question of choosing the best suited programming paradigm for this kind of hybrid computer architecture.

Major HPC vendors like CRAY, IBM and SGI are offering HPC systems with an increasing number of cores per processor and node, while the main memory per core either decreases or remains constant (see examples in the table below) due to the energy consumption and memory costs. This trend opens the possibility to increasingly profit from hybrid parallelization.

	Machine	Processor	Cores/Processor	Memory
CRAY	XT4	AMD Opteron Series 1000	2 or 4 (1 processor per node)	1–8 GByte per node
	XT5	AMD Opteron Series 2000	4 or 6 (2 processors per node)	16–32 GByte per node
	XT6	AMD Opteron Series 6000	8 or 12 (2 processors per node)	32–64 GByte per node
IBM				
	Blue Gene/L	700 MHz PowerPC 440	2	≤ 512 MByte per core
	Blue Gene/P	850 MHz PowerPC 450	4	≤ 2 GByte per core
SGI	Altix 4700	Intel Itanium Series 9000	2 (2 processors per node/blade)	2–6 GByte per core
	ICE	Intel Xeon 5200/5400/ 5500/Series	2 to 4 (2 processors per node/blade)	up to 6 GByte per core
	UV	Intel Xeon Nehalem-EX	4 to 8 (2 processors per node/blade)	up to 4 GByte per core

In this article we compare the performance of a hybrid (i.e. combined MPI and OpenMP) version of the *Berlin Quantum ChromoDynamics Program* (BQCD) [1,2] with the pure MPI case. BQCD is used in benchmarks for supercomputer procurements at our centres, as well as in the DEISA [3] and PRACE [4] projects. The benchmark code is written in Fortran90. It represents QCD as a typical supercomputer application. At a more general level the benchmark measures the performance of a sparse matrix solver which is similar to other application areas.

Machine Overview

We have measured performance on the following machines.

Cray XT5. We have used the XT5 partition of Jaguar at Oak Ridge National Laboratory in the US. This partition has 18,688 compute nodes. Each node contains two hex-core AMD Istanbul processors running at 2.6 GHz. The whole partition has 224,256 cores and a peak performance of 2.3 PFlop/s. The nodes are coupled with Cray's Seastar 2+ interconnect. The network has a 3D torus topology. Jaguar is the fastest computer on the current Top500 list.

IBM BlueGene/P. We have used Babel at IDRIS in France and Jugene at Jülich Supercomuter Centre. The BlueGene/P has quad-core PowerPC nodes running at a relatively low clock frequency of 850 MHz. The BlueGene communication network has a 3D torus topology. The machine at IDRIS has 10 racks (40,960 cores), Jugene at JSC has 72 racks (294,912 cores). Jugene is the fastest European computer on current Top500 list. Its peak performance is 1 PFlop/s.

SGI ICE 8200. We have used Jade at CINES in France. The ICE is a highly integrated cluster. Its nodes have two Intel quad-core processors running at 3 GHz. Jade has 1,536 nodes with *Harpertown* processors and 1,344 nodes with Nehalem processors. The whole machine has a peak performance of 144 TFlop/s. Nehalem processors offer a much higher memory bandwidth which results in an almost doubled performance of BQCD. The Nehalem nodes of Jade are connected with a quad data rate Infiniband network that has a non-blocking hypercube topology.

Setup

We measure the performance of the standard conjugate gradient (CG) solver that is implemented in BQCD. The dominant operation in the solver is the matrix times vector multiplication. This operation involves the communication of boundary elements from neighbouring processes. We are going to a regime where the exchange of boundary elements takes 50 percent of the execution time or even more. The second type of communication needed in the solver is global summation. For the hybrid parallelization all loops were parallelized with the parallel do directive. MPI is called in sequential regions, i.e. only the master threads are involved in the MPI communication. We have used one or two MPI processes per node and one OpenMP thread per core and look at strong scaling for a $48^3 \times 96$ lattice.

Results

Cray XT5 (Figure 1). We have plotted results for pure MPI and 8 threads per MPI process. Pure MPI is faster for smaller numbers of cores. For larger numbers of cores the hybrid runs outperforms pure MPI (the speed-up is 1.25). The break-even point is at 8,192 cores.

IBM BlueGene/P. We made various measurements on the BlueGene at IDRIS and during the Extreme Scaling Workshop 2010 at JSC [5]. The outcome is that on BlueGene/P hybrid runs are always a few percent slower than pure MPI runs.

SGI ICE 8200 (Figure 2). From the figure we see that hybrid runs with 8 threads per MPI processes were favourable in most cases. However the speed-up is moderate. On 2,048 cores the speed-up is 1,06 and on 4,096 cores it is 1.12.

Discussion

We have presented performance result for a straightforward hybrid parallelization approach in which only one thread performs MPI communication. This approach has the advantage that less data have to be communicated but the disadvantage that fewer cores take part in the communication. On two machines, the XT5 and the ICE, this hybrid approach leads to a performance improvement of 10 to 20 percent for large number of cores. On the BlueGene we have seen no improvement. This observation

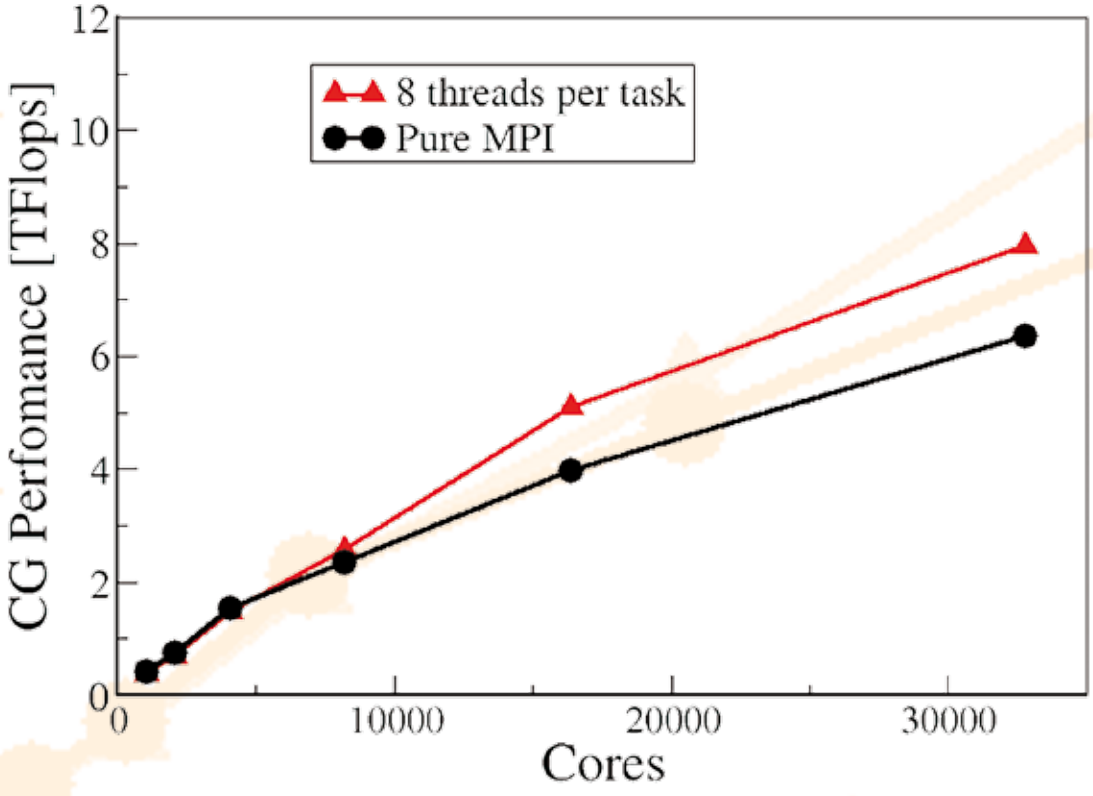


Figure 1: Performance comparision for a $48^3 \times 96$ lattice on the Cray XT5

corresponds to the amount of communication loss on the respective machines. For example, on 4,096 cores the communication loss in the pure MPI case is about 50 percent on the XT5 and the ICE while it is only 15 percent on the BlueGene. It should be possible to speed-up communication in the hybrid approach by having more threads call MPI.

In conclusion we see some acceleration by hybrid parallelisation on current machines. However, the method is not advantageous in general and we have never seen a clear speed-up like 30 percent which was achieved with BQCD on the Hitachi SR8000 at LRZ [6]. Nonetheless, a hybrid parallelization approach might become the method of choice on future many-core processors.

References

[1] http://www.zib.de/stueben/bqcd

[2] Allalen, M., Brehm, M. and Stüben, H. Performance of quantum chromodynamics (QCD) simulations on the SGI Altix 4700, Computational Methods in Science and Technology 14(2) (2008) 69–75 2008

[3] http://www.deisa.eu/science/benchmarking/codes/bqcd

[4] http://www.prace-project.eu/documents

[5] Stüben, H. and Allalen, M. Extreme Scaling of the BQCD Benchmark, in: Mohr, B. and Frings, W. (eds.), Technical Report FZJ-JSC-IB-2010-03 (2010) 31–34

[6] Schierholz, G. and Stüben, H. Optimizing the Hybrid Monte Carlo Algorithm on the Hitachi SR8000, in: Wagner, S., Hanke, W., Bode, A. and Durst, F., High Performance Computing in Science and Engineering, Munich 2004, Springer-Verlag, 385–393 2004

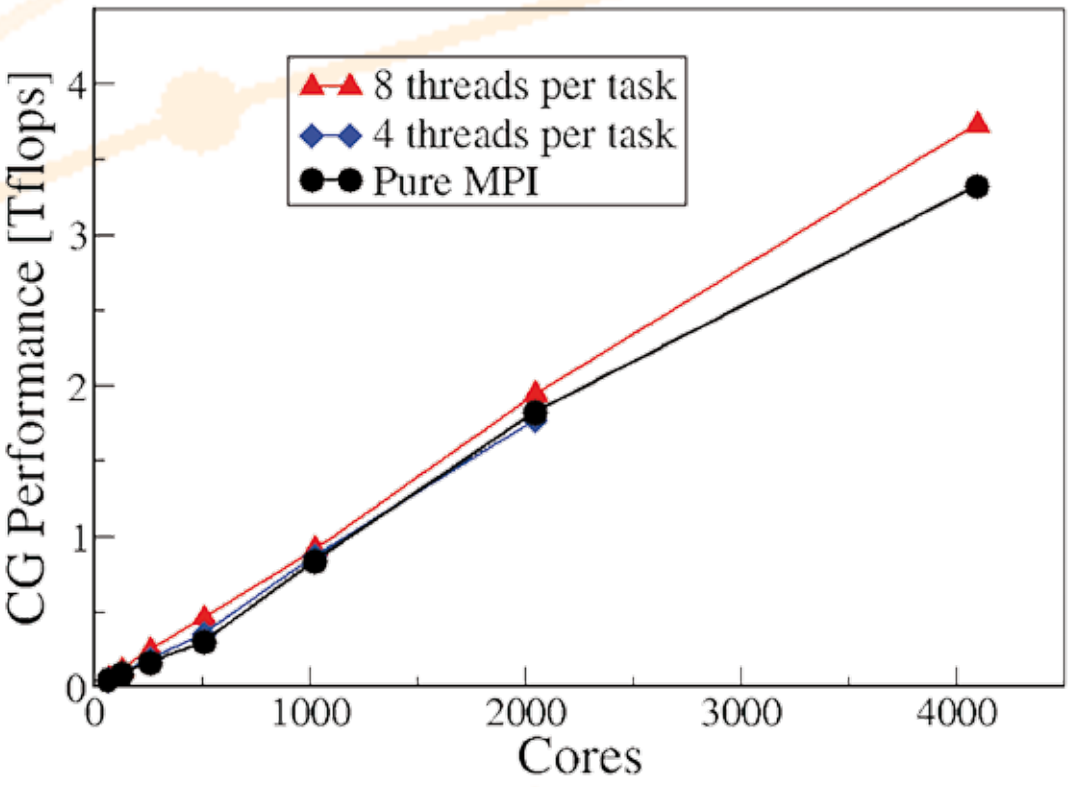


Figure 2: Performance comparision for a $48^3 \times 96$ lattice on the SGI ICE 8200

• Momme Allalen[1]
• Matthias Brehm[1]
• Hinnerk Stüben[2]

[1] Leibniz Supercomputing Centre, LRZ

[2] Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

Next Generation Simulation of Cell Factories

The increase in speed and cost-reduction of genomic data like DNA probes offers in combination with complex network models of biological cells and High Performance Computing new chances for using microorganisms and sugar for the production of chemical commodities. Sugar as a sustainable carbon source avoids the net production of CO_2. Today, microorganism are developed that can even use CO_2 for synthetizing high-value products. This so-called "green chemistry" is applied increasingly in industry, for example for the production of feed additives, biofuels, drugs and other biological goods with microorganisms and mammalian cells. The economic efficiency of these processes, though, depends entirely on the capabilities of the cell factories involved and continuous efforts have to be made to improve their performance. The "Industrial Biotechnology" project aims at accelerating this task by using complex computer models of cellular metabolism and High Performance Computing. The project is a ZIM founded two-year cooperation between the Stuttgart based company Insilico Biotechnology AG and the HPC Center Stuttgart. Computations that need up to several months using scalar processors will eventually be achievable in mere days, putting the beneficiaries far ahead of any competition.

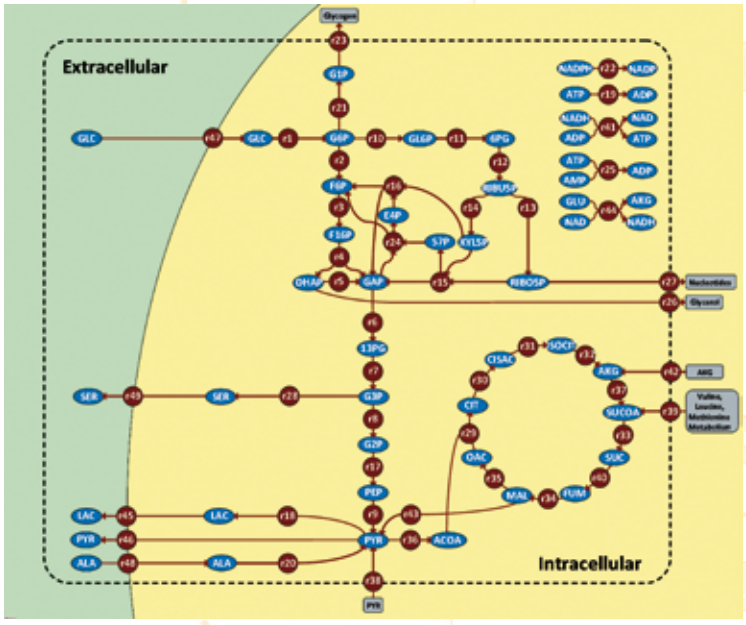


Figure 1: Intracellular metabolite dynamics. The error bars indicate standard deviations of the experimental data.

Predicting the efficiency of such biological processes is an intricate process requiring expertise in both biochemistry and computer science. At first, an initial model of the metabolic networks is developed in laboratory studies. This is followed by an extensive series of simulations to verify the network dynamics of the newly generated model through intracellular data. Finally, newly predicted targets for improved network performance are validated through laboratory experiments.

As an example, the efficiency of "cell factories" like the Escherichia coli bacterium depends on kinetic attributes of biochemical reactions, that are generally unknown and need to be identified by reverse engineering. Simulations with various parameter sets are conducted and differential algebraic equations built from the network topology and kinetic rate equations are solved. The solutions of these equations then yield the trajectories of intracellular compounds. The trajectories, in turn, are fitted to the measured data and later maximized. This allows specifying a function that quantifies the quality of a parameter set and that can be handled with general optimization

techniques. Due to the time needed to solve a single set of equations, combined with the possibly millions of simulations needed to optimize the parameters in both steps of the process, the number of components that can be evaluated in hardware is limited. It is often quoted in literature to not exceed twenty and shows the need for a highly parallel approach to the problem to allow a wide industrial application.

In the area of optimizing real-valued multidimensional functions, an ongoing and vivid development aims to exploit the strengths of highly parallel architectures available today. Function evaluations can be distributed discretely over large number of cores. This allows (i) to cover more points in a given search space and find a presumed optimum faster and (ii) to improve existing solutions resulting at a much higher degree of accuracy. Additionally, several hundred or even thousand instances of the same algorithm provide the capability to either widen the search space or partition the existing one. The benefits are once more reduction of time and/or a higher accuracy. Some methods combine these two approaches by connecting several instances of a given and proven algorithm into a swarm or cloud. This leads to algorithms that are more efficient than the single parts they are made of: each instance could communicate the current local optimum, giving others the chance to adopt their local search parameters and improve their solutions based on the global information available.

Apart from the modularization and optimization of the metabolic networks on current high performance architectures, the use of graphical processing units is targeted to further enhance single CPU performance. Taking into account all possibilities, increasing computational performance is merely unlimited. Combined with modern biotechnology, High Performance Computing will become an increasingly important source for innovation in the life science industries.

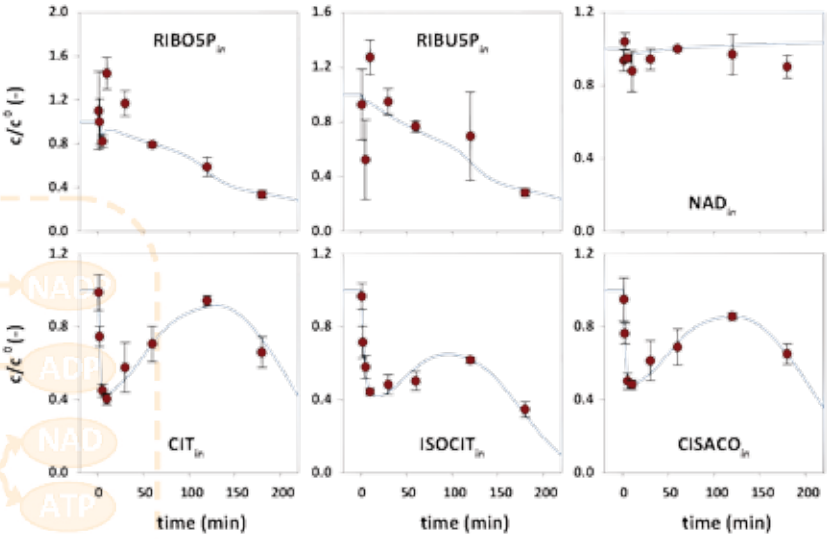


Figure 2: Metabolite time-series data were determined and used to parameterize a dynamic network model of the central carbon metabolism.

Projects

References

[1] **Maier, K., Hofmann, U., Reuss, M. and Mauch, K.**
Dynamics and Control of the Central Carbon Metabolism in Hepatoma Cells, BMC Systems Biology 4:54, 2010

[2] **Hansen, N.**
The CMA Evolution Strategy: A Comparing Review, Studies in Fuzziness and Soft Computing, Vol. 192/2006, pp. 75-102, 2006

[3] **Müller, C. L., Baumgartner, B., Ofenbeck, G., Schrader, B., Sbalzarini, I. F.**
pCMALib: a parallel FORTRAN 90 library for the evolution strategy with covariance matrix adaptation, In: Proc. ACM Genetic and Evolutionary Computation Conference (GECCO'09), Montreal, Canada, July, 2009

[4] **Funded by:**
Bundesministerium für Wirtschaft und Technologie

- Thomas Baumann[1]
- Katharina Benkert[1]
- Alexander Broicher [2]

[1] University of Stuttgart, HLRS

[2] Insilico Biotechnology AG

Gefördert durch:
Bundesministerium für Wirtschaft und Technologie
aufgrund eines Beschlusses des Deutschen Bundestages

GAMES (Green Active Management of Energy in IT Service centres)

Energy consumption and implicit CO^2 emissions of computing and data centres have increased drastically over recent years and are expected to increase even further. Beside the raising costs for energy consumed by IT service centres, people are getting more and more aware about the follow-up of the high demand of energy for the IT sector, like the impact on global warming and CO^2 emissions. As an example, worldwide data centres CO^2 emissions are already equivalent to about half of the total airlines' CO^2 emissions and are expected to overcome the 40% of Total Cost of Ownership of worldwide IT by 2012. Data centre electricity consumption accounts for almost 2% of the world production and their overall carbon emissions are greater than both Argentina and the Netherlands together [1].

Figure 1: Sample of simulation in the design phase

Since computing demand and electricity prices are rising whilst becoming dwindling resources, energy consumption of IT systems and data centre energy efficiency are expected to become a priority for the industry. Despite the fact that many stakeholders have been undertaking significant efforts in delivering new yet more energy effective IT equipment allowing significant cost and energy savings, unfortunately the problem of the energy efficiency of Information Systems as a whole has not been properly addressed so far.

Green Computing is a new discipline and practice aimed at designing and using computing resources in an environmentally-aware way. It was originated more than a decade ago with the main goal of reducing energy consumption of computing resources, yet maintaining a clear focus on the impact on the environment. Although many progresses have been made by Green Computing, making new chips and servers available which undoubtedly consume less energy, in most cases improvements in efficiency are devoured by increasing demand for computing power and capacity, driven by new digitized business processes and services.

The GAMES project [2] aims at developing tools and methodologies to improve the energy efficiency of IT service centres by enabling active management of resources and software equally. Whilst storage hosts can principally reduce the computing frequency to save energy, compute providers and in particular High Performance Computing centres have more dynamic and changing demands towards the infrastructure usage – such as the flexible degree of scale out of a process or the different scope of data access and storage of different applications. Most resources in such an environment do not allow for fast enough adaptation of their energy parameters without affecting the overall performance. What is more, most parameters and relationships between usage and consumption are not even known as yet, e.g. the total energy profile of an application that runs at maximum CPU clock rate for a short time may be lower than that of the same application running at half clock rate for a longer time, depending on the behavioural profile.

Therefore the GAMES project will examine the energy profiles of different applications and systems according to their specific behaviour in more detail, deriving energy profiles from this which indicate how to configure the infrastructure for best energy and performance efficiency. It will expose the profile parameters to enable developers to write energy efficient applications and configure performance according to their needs. GAMES will furthermore develop a monitoring and management system tightly coupled to the resource infrastructure, thus enabling dynamic, flexible and immediate reaction to changing requirements, without affecting the overall execution performance.

In particular, the GAMES project aims at developing a set of innovative methodologies, metrics, Open Source ICT services and tools for the active management of energy efficiency of IT Service Centres. It focuses on the following two aspects:

1. Co-design of energy-aware information systems and their underlying services and IT service centre architectures in order to satisfy users requirements (Quality of Service), service performance, context, addressing energy efficiency and controlling emissions (cp. Fig. 1). A combination of Green Performance Indicators are proposed to evaluate if and to what extent a given service and workload configuration affects the carbon footprint emissions' levels;
2. Run-time management of IT Service Centre energy efficiency, exploiting the adaptive behaviour of the system at run time, both at the service/application and IT architecture levels (including IT components like servers, and storage), whilst considering the interactions with the facility management as well in an overall unifying vision.

In particular, GAMES will advance the current scientific and technology state-of-the-art in energy efficiency for IT service centres in the following domains:

- GAMES will create and make available an integrated methodology (GAMES co-design methodology) for the shared design of "Green IT Service Centres", trading-off Quality

GAMES
Green Active Management of Energy in IT Service centres

of Services, users' business and functional requirements against energy efficiency and emissions;

- GAMES will complement and extend one of the most leading-edge Open Source data centre monitoring tools, namely NAGIOS, with the capacity of assessing, monitoring and controlling, both in a reactive and proactive way, energy costs and emissions (GAMES Energy Efficiency Tool) in real time of alternative yet viable options/ configurations for distributing services among the virtualized machines, workload among servers and storage devices as well as balancing power against heat/ temperature at facility level;

- The GAMES web-tool will be enriched with advanced knowledge-based and information extraction features (GAMES Knowledge & Mining Module) by originally combining data mining, semantic and context management technologies for closely aligning users business requirements for power demand with historical trend and real available resources;

- The GAMES tooling framework will provide an adaptive control feature (GAMES Monitoring & Adaptive Control Infrastructure), matching the planned behaviour with the output dynamically provided by the energy sensing and monitoring infrastructure, the user context and historical patterns, for evaluating if and to what extent the adopted course of actions will contribute to effectively manage the energy efficiency;

- GAMES will define comprehensive energy efficiency assessment metrics (GAMES Green Performance Indicators) as an enabler to combine energy efficiency facility features with IT infrastructure and business/application energy features.

HLRS is participating in two roles within GAMES. First of all HLRS is in the role of a potential end-user of GAMES as a national supercomputing centre with increasing power demands of currently 5 MW for operating the different hardware systems for academia and industry. In this role HLRS also supports the activities in creating the knowledge and information base for the management framework. Additionally HLRS contributes to GAMES as a technology and software provider of highly scalable monitoring solutions and service oriented architecture driven IT solutions with a High Performance Computing focus.

Participants

- Engineering Ingegneria Informatica, I
- Politecnico di Milano, I
- High Performance Computing Center Stuttgart (HLRS), GER
- Technical University of Cluj-Napoca, RO
- IBM ISRAEL - Science and Technology LTD, IL
- Christmann Informationstechnik, GER
- ENERGOECO, RO
- ENELSi, I

References

[1] Kaplan, J. M., Forrest, W. and Kindler, N. Revolutionizing Data Center Energy Efficiency. McKinsey&Company, July 2008

[2] http://www.green-datacenters.eu

• Alexander Kipp

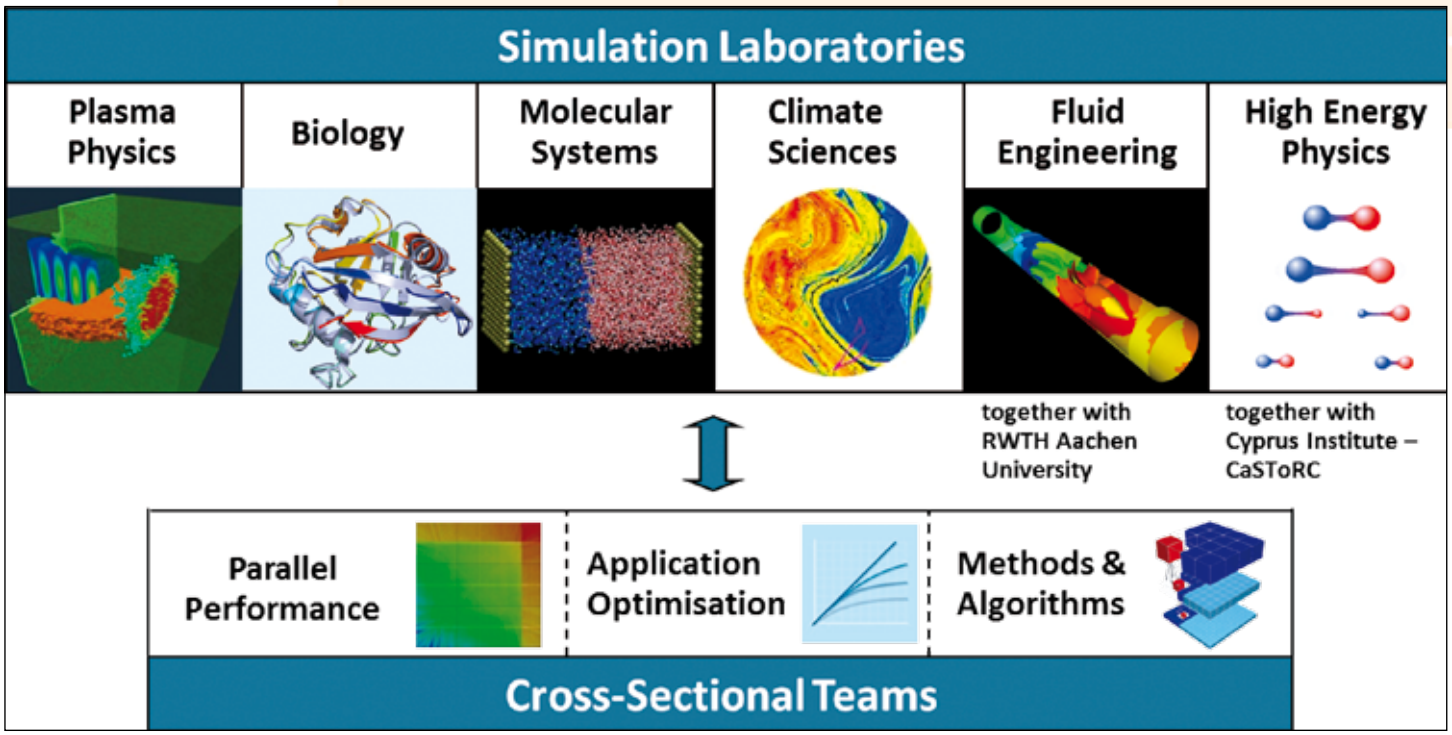
University of Stuttgart, HLRS

Status of Simulation Laboratories at JSC

In recognition of the extreme software challenges facing its HPC users the Jülich Supercomputing Centre (JSC) has begun to set up a new, high-level support structures: Simulation Laboratories. The JSC blueprint defines a simulation laboratory (SimLab) as a community-oriented research and support team, consisting of a core group located at a supercomputer centre and possibly a number of associated scientists outside. A key requirement for the viability and visibility of the SimLab is that as far as possible, the core team should be hired from the community which it is designed to serve and interact with. In this way, the expertise offered can go far beyond the support of a traditional expert advisor, who might typically be able to perform application performance analysis and bottleneck identification, but would not possess the disciplinary know-how to carry out the necessary code restructuring.

The activities which characterize a simulation laboratory will of course vary depending on the make-up and location of its scientific community, but in general will comprise a combination of specialized and more generic tasks, such as:

- Participation in research projects of the community, including active col laboration with user projects and common publications.
- Development and maintenance of software and databases for the whole community.
- Research and development of computational methods needed by major applications on high-end systems.



- Performance analysis, diagnosis and adaptation/optimization of software on highly scalable high-end supercomputers.

- Running training courses in computational methods and in efficient usage of high-performance systems.

- Community building measures such as the organization of community-oriented workshops and construction of web-portals.

Three Simulation Labs based on the above blueprint have now been established at JSC in the fields of Computational Biology, Molecular Systems and Plasma Physics, which have already been actively engaged with user groups from their respective communities over the past year. A fourth SimLab in Climate Modelling will come on stream in 2011, and two more in Fluid Engineering and High-Energy Particle Physics are also planned within the next two years. The SimLabs are currently staffed at around 2-3 postdocs with a similar number of PhD and Master students and/or technicians (in software/programming). In addition to the Simulation Laboratories, JSC also houses a number of Cross-Sectional Teams that focus on the development of generic HPC methods and algorithms, programming tools and petabyte data repositories. These complement the work of the Simulation Laboratories by providing first-hand expertise in computer science and numerical mathematics. A short summary of the JSC SimLab profiles can be found at:

http://www.fz-juelich.de/jsc/simlabs

Over the past year, the first three simulation labs at JSC have endeavored to establish themselves as autonomous research and support units. This has involved a range of community engagement activities, including: short visits of scientists to port and analyse applications to the JSC supercomputers; participation in major German and European HPC proposals; and the staging of a joint workshop in June 2010 to promote HPC activities within their respective scientific communities.

As a next logical step to consolidate these initial community-building activities, JSC has issued its first call to current and potential users of the supercomputers in Jülich to apply for high-level support from the Simulation Labs. Specific support actions covered by this call include: restructuring of computational methods needed to exploit highly parallel architectures, performance analysis and scaling improvement of codes/applications, as well as porting of new codes to the Juropa and Jugene systems. In this 6-month pilot phase, work packages will not exceed 2 person-months of SimLab staff resources. Applicants are also expected to contribute an equivalent amount of manpower to the project, particularly where the work involves major code/algorithm redevelopment. In the longer term, it is expected that project proposals will grow in number, duration and complexity, and will therefore inevitably require some form of ranking via an advisory committee.

• Paul Gibbon

Jülich Supercomputing Centre

Towards EXascale ApplicatTions (TEXT)

With many-core processors offering evermore compute power per socket, and large-scale supercomputers built from these bricks, the long-lasting discussions on parallel programming models are being posed again. Venturing into the realm of Petascale applications, several key questions regarding scalability in terms of memory and processing overhead per parallel instantiation are considered and weighed against the need for portability, readability and maintainability.

The TEXT project is funded by the EC as part of the INFRA-2010 call for two years. The nine partners from Spain, Germany the UK, France, Greece and Switzerland share the vision, that the key component to support high productivity and efficient use of a system is the programming model. Among the partners are four HPC centers, also members of the PRACE collaboration, with JSC having a Petaflops machine in production. The project centers around the MPI/SMPSs, which is part of the StarSuperScalar (StarSs) model developed by Barcelona Supercomputing Center.

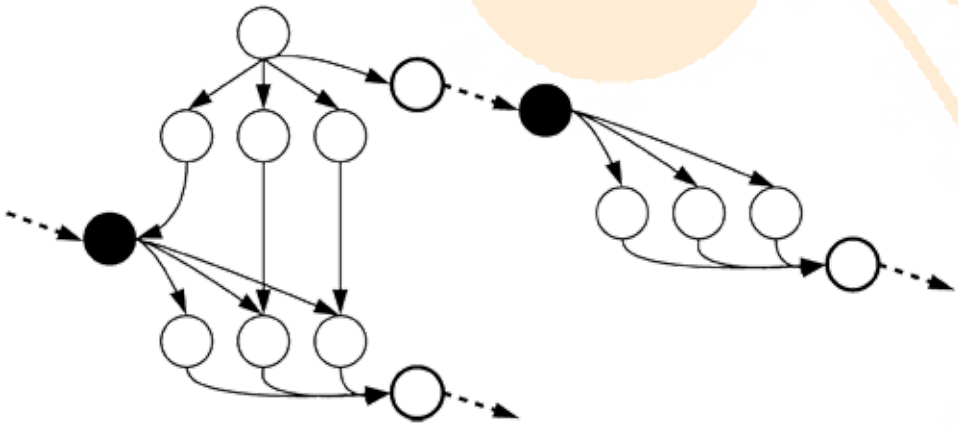


Figure 1: Part of exemplary scheduling graph of SMPSs with communicating tasks

Overview

The TEXT project's technology combines the available scalability of the Message Passing Interface (MPI) across compute nodes with the possibilities of per-node concurrency via asynchronous task of SMP-Superscalar (SMPSs). Given an existing applications using MPI for work decomposition, the programmer may further parallelize the application into so-called tasks using SMPSs. These tasks then are dynamically scheduled by the SMPSs runtime environment to be executed asynchroniously. The runtime generates a graph and efficiently maps ready-to-execute tasks onto the available cores, taking care of dependencies among the tasks. With less MPI ranks running, one has lower connectivity, therefore lower memory overhead for MPI-internal buffers, and potentially bigger message sizes, which together with the SmpSs task model allow for better communication-computation overlap. In the latter case, the MPI communication is handled within an SMPSs-task and scheduled by the runtime, whenever the computed data is available at the sender.

Similar to OpenMP, the programmer annotates her application using pragmas, specifying the functions to be run as task, their input, output and inout parameters, as well as their sizes. An example of a simple function may be:

```
#pragma css task input(SIZE)
inout(v[SIZE])

void compute_vector (float *v,
int SIZE){...}
```

After the programmer has initialized the environment using #pragma css start, any call to the above function will be asigned to a thread and executed asynchronously on the cores of the node. Furthermore, synchronization points may be necessary to wait for and guarantuee a common view on the computed data. The main point however here is simple point-to-point dependency between asynchronously executed tasks, which allow the graph scheduler to more flexibly parallelize independent tasks.

Using a pre-compiler, in our case `SMPSs-cc`, the annotated C or Fortran source code is amended with further administrative code, and finally passed to the native back-end compiler. Using the `-keep` compiler option, the programmer may see the actually generated intermediate source code, which is compiled by the back-end compiler.

Aim of the Project

The StarSs programming model has shown good results in its GridSs and CellSs incarnations for execution in the Grid and on the IBM Cell architecture, respectively. In the TEXT project, we hope to extend the programming model onto the existing MPI-parallel applications, which are important to the compute centers.

These applications chosen have been used already in the context of the PRACE project, and include SPECFEM3d (UPPA), PSC and PEPC (JSC), BEST and LS1 (HLRS) and CPMD (IBM). Based on the MPI-parallel version, combination of the node-local parallelization using SMPSs plus MPI will be investigated using performance analysis.

Each application offers its own challenges, e.g. LS1 being a C++ code has very elaborate class structure and is one of the first C++ codes to be used with StarSs, while BEST uses some of the more intricate features of Fortran95 and Fortran2003. Both of these codes are being ported to MPI+SMPSs.

While the option to keep the intermediate code allows using traditional tools to work, this is cumbersome. Therefore performance and debugging tools will be enhanced to support the special requirements of SMPSs. For example, being able to debug without having to fall back to debugging the intermediate code, or being able to debug using break-points in tasks being generated.

To enhance performance, it will be necessary to evaluate proper chunk sizes of the tasks and estimate the overhead introduced due to dependencies.

The Partners

- Barcelona Supercomputing Center (BSC)
- High Performance Computing Center Stuttgart (HLRS)
- Jülich Supercomputing Center (JSC)
- Edinburgh Parallel Computing Center (EPCC)
- Foundation for Research and Technology Hellas (FORTH)
- University of Manchester (UMAN)
- Université de Pau et des Pays de l'Adour (UPPA)
- Universitat Jaume I de Castellón (UJI)
- IBM Research, Zurich

- Rainer Keller
- José Gracia

University of Stuttgart, HLRS

Towards Exascale: Foundation of EIC and ECL

The Jülich Supercomputing Centre at Forschungszentrum Jülich started two new collaborations with renowned companies to force the research and development on Exascale systems. In March 2010, Forschungszentrum Jülich and IBM signed the contract for a joint "Exascale Innovation Center" (EIC), and, at ISC 2010 in Hamburg, Forschungszentrum Jülich, Intel, and ParTec signed a multi-year agreement to create a new ExaCluster Laboratory (ECL), both located on the campus of Forschungszentrum Jülich. These two endeavours reflect the dualistic concept JSC is pursuing since several years now. On the one hand side, the JUGENE supercomputer based on IBM's BlueGene technology targets highest scalability while, on the other hand, the JUROPA/HPC-FF cluster computer-based on Intel's Nehalem processors and ParTec's ParaStation cluster middle-ware aims at more general purpose applications. Since both architectures provide a large potential for machines beyond the Petaflop range, Jülich's goal is to analyse the prospect to scale these technologies to exascale.

Forschungszentrum Jülich and IBM signing the contract for a joint "Exascale Innovation Center". From the left: David Jursik (Vice President worldwide Deep Computing Sales, IBM Systems & Technology Group), Dirk Wittkopp (Vice President and Director IBM Laboratory Böblingen), Prof. Thomas Lippert (Director of Jülich Supercomputing Centre, Forschungszentrum Jülich), Prof. Sebastian Schmidt (Board of Directors of Forschungszentrum Jülich) (Photo: Forschungszentrum Jülich)

Exascale Innovation Center (EIC)

EIC will develop hardware and software for an exascale supercomputer by the end of this decade. Exascale - which means a thousand times more performance than JUGENE or 1 exaflop/s - is the premier challenge for supercomputing worldwide. With exascale supercomputers it will be possible to perform simulations of unprecedented complexity. However, a great many challenges have to be tackled to achieve this goal. Energy efficiency will be the most prominent challenge to be solved in order to create exascale supercomputers. In a successful collaboration with IBM Germany, the German Research Foundation's (DFG) Collaborative Research Centre for Hadron Physics and JSC developed the supercomputer QPACE, which was declared the world's most energy-efficient supercomputer by the GREEN500 list last November. "With QPACE, we have managed to get a grip on energy consumption," says Prof. Thomas Lippert, director of the Jülich Supercomputing Centre. Further challenges are the development of the chip and processor technology towards the exascale level with respect to hardware and the improvement of algorithms with respect to software in order to run and use a supercomputer of this size. Jülich is contributing its outstanding expertise in the development of algorithms.

According to the schedule, a prototype of the new exascale supercomputer is expected to be available in 2015. Five scientists from the IBM development laboratory in Böblingen and five scientists from Jülich will be collaborating with a team of scientists at the IBM Watson Research Center in Yorktown Heights. The goal is to install an exascale-class system in Jülich by 2019.

ExaCluster Laboratory (ECL)

ECL will explore the key challenges of building computing systems with a thousand times the performance of today's fastest supercomputers. The new lab will initially employ about a dozen researchers and is expected to triple its staff over time.

"The Forschungszentrum Jülich has taken a leading role in driving high-performance computing research in Europe," said Kirk Skaugen, vice president and general manager of Intel's Datacenter Group. "We have chosen to work with Forschungszentrum Jülich and the ParTec Cluster Competence GmbH, because of their strong history of innovation in the area of HPC."

The ExaCluster Laboratory will conduct research into current challenges in systems management software for large heterogeneous supercomputer systems, with a view to scaling this software to reach exaflop/s performance. This will include research on open exascale runtime system software and software tools. The aim of the work on systems management software is to further improve the scalability of ParTec's ParaStation cluster middle-ware and – at the end – to create an OpenSource software stack capable to manage Exascale systems. Specific problems tackled in the first 3-year period of the Laboratory are – among others – the analysis of the necessary measures to improve system-resiliency or known scalability issues with today's MPI-stacks.

Looking at current processor roadmaps it becomes clear that clusters will not be able to cope with proprietary solutions from IBM or Cray in the near future, if they just follow the line of general purpose CPUs. Therefore, the applications of accelerator technologies will be crucial in order to keep clusters competitive in the next years. ECL will investigate innovative technologies to overcome the yet unsolved problem to couple accelerators more tightly to the high-speed interconnects building the spine of today's cluster computers.

- Norbert Eicker
- Willi Homberg
- Sabine Höfler-Thierfeldt

Jülich Supercomputing Centre

At the ISC 2010 in Hamburg, Forschungszentrum Jülich and the companies Intel und ParTec sign the contract for a joint "ExaCluster Laboratory". From the left side: Hugo Falter (ParTec), Prof. Achim Bachem (Board of Directors of Forschungszentrum Jülich), Prof. Thomas Lippert (Director of Jülich Supercomputing Centre, Forschungszentrum Jülich), Kirk Skaugen (Intel Vice President), Raj Hazra (Intel).
(Photo: Forschungszentrum Jülich)

Investigations of Intel's many integrated Core Architecture at LRZ

Supercomputer facilities for HPC are targeting petascale technology, 10^{15} floating point operations per second (flops), but they are also hungry in resources and energy. Some years ago, in order to increase the performance and the number of calculations delivered by a processor, it was feasible to increase the frequency as well as the number of transistors on the processor chip. Today, chip manufactures are facing increasing problems to produce smaller circuitry and put more transistors on a chip (with present 45 nm technology, hundreds of millions already). On the other hand, increasing frequency is not a practical solution either, due to increasing energy consumption, heat-induced leakage currents, and the comparatively low rate of data transfer by RAM memory and disks, which is insufficient to keep up with the CPU's requirements.

How will it be possible to maintain pace with Moore's law, obtain a hundred times more computational power than today's computers, and still be able to pay the electricity bill by using energy efficiently? These challenges are now being addressed by chip manufactures, IT companies and research centres for hardware and software around the world. Leading this trend is Intel, which is now aiming for the so called Intel Many Integrated Core (MIC) architecture.

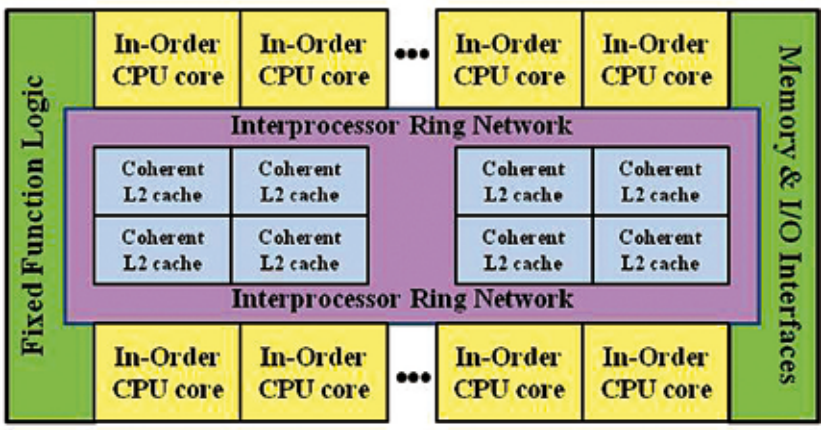


Figure 1: Scheme of Intel MIC architecture prototype

The Leibniz Supercomputing Centre (LRZ) intends to strongly participate in this important area of investigation and therefore, in cooperation with Intel, is starting to enjoy the advantages of Intel MIC technology. Selected LRZ staff is assessing the Intel® MIC technology by investigating how easy or difficult it will be for a normal user from the scientific community to port and to deploy their programs. More specifically, an N-Body cosmological simulation program and several computational kernels were selected for this assessment process; these applications are complex enough to provide a real challenge for the programmer and the architecture. If Intel MIC technology is adequate for applications of this type we are convinced it will be a milestone towards sustained Petascale and Exascale technology. This Intel MIC is one of the first implementations of the 22 nm technology. The first product, whose code name is Knights Corner, was announced during the International Supercomputing Conference (ISC) in Hamburg in May, 2010. This chip will have as many as 50 x 86-based cores [1]. Its principal function will not be to run a standalone program but instead to execute sections of highly parallelizable

code suitable, a concept generally known as application acceleration.

LRZ is evaluating an Intel MIC prototype, which is shown Figure 1. It consists of 32 in-order x86-based processor cores that are augmented by a 512 bit wide vector processing unit (VPU) (see Figure 2). All cores are interconnected by a bi-directional on-chip ring network. This ring network also provides a path for the L2 caches to access memory. The L2 cache is shared between all cores. The instruction decoder of the Intel MIC architecture supports the standard Pentium x86 instruction set as well as new instructions for explicit cache control and vector type operations.

Each core has access to its 256 kByte local subset of a coherent L2 cache. The L1 cache sizes are 32 kByte for instructions and 32 kByte for data. Furthermore each scalar unit supports 4 threads of execution, with separate register sets per thread. Hence thread switching can be extensively used to cover cases such as pipeline stalls and cache misses. Figure 3 shows a block diagram of the 16-wide vector processing unit. A set of new VPU instructions allows a variety of instructions on both integer and floating point data types such as:

- Standard arithmetic operation including fused multiply-add
- Standard logical operations
- Gather and scatter operations

The VPU instructions allow up to three source operands, one of which can directly come from the L1 cache. 8-bit and 16-bit integer and floating point numbers can be directly converted to 32-bit integer and floating point numbers without loss of performance. The VPU supports swizzling of register data in a variety of ways and replication of data from memory across the VPU lanes.

The focus on the assessment of this Intel MIC architecture will be on evaluation of the ease of use of this parallel SMP cache coherent prototype processor platform for HPC.

http://www.intel.com/pressroom/archive/releases/20100531comp.htm

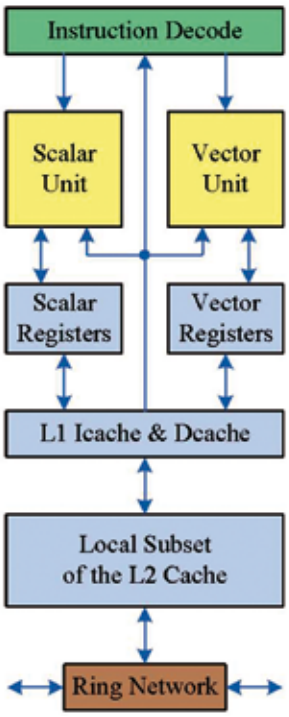


Figure 2: x86-based core and associated system blocks

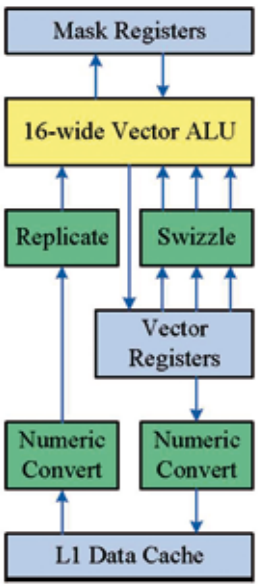


Figure 3: Block diagram of Vector Processing Unit (VPU)

• Orlando Rivera

Leibniz Supercomputing Centre, LRZ

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities (Leibniz-Rechenzentrum, LRZ) provides comprehensive services to scientific and academic communities by:

- giving general IT services to more than 100,000 university customers in Munich and for the Bavarian Academy of Sciences
- running and managing the powerful communication infrastructure of the Munich Scientific Network (MWN)
- acting as a competence centre for data communication networks
- being a centre for large-scale archiving and backup, and by
- providing High-Performance Computing resources, training and support on the local, regional and national level.

Research in HPC is carried out in collaboration with the distributed, statewide Competence Network for Technical and Scientific High-Performance Computing in Bavaria (KONWIHR).

Contact:
Leibniz Supercomputing Centre

Prof. Dr. Arndt Bode
Boltzmannstr. 1
85478 Garching near Munich
Germany

Phone +49-89-358-31-80 00
bode@lrz.de
www.lrz.de

View of "Höchstleistungsrechner in Bayern HLRB II", an SGI Altix 4700
(Photo: Kai Hamann, produced by gsiCom)

Compute servers currently operated by LRZ are

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
HRLB II: SGI Altix 4700 Intel IA64 19 x 512-way	9,728 Cores 39 TByte	62.3	Capability Computing	German Universities and Research Institutes, DEISA
PRACE Prototype SGI ICE Intel Nehalem 8-way	348 Cores 1.1 TByte	3.8	Capability Computing	German Universities and Research Institutes, PRACE
Linux-Cluster SGI Altix 4700 Intel IA64 256-way	256 Cores 1 TByte	1.6	Capability Computing	Bavarian Universities
Linux-Cluster Intel Xeon EM64T/AMD Opteron 2-, 4-, 8-, 16-, 32-way	4,438 Cores 9.9 Tbyte	36.3	Capacity Computing	Bavarian and Munich Universities, D-Grid, LCG Grid
Linux-Cluster Intel IA64 2-, 4- and 8-way	220 Cores 1.1 TByte	1.3	Capacity Computing	Bavarian and Munich Universities
Linux-Cluster Intel Xeon EM64T AMD Opteron 2-, 4-, 8-, 16- and 32-way	3,156 Cores 11.6 TByte	19.1	Capacity Computing	Bavarian and Munich Universities, D-Grid
Linux-Cluster Intel Xeon EM64T AMD Opteron 4- and 8-way	1,188 Cores 2.4 TByte	6.1	Capacity Computing	LHC Tier 2 Computing Grid

A detailed description can be found on LRZ's web pages: www.lrz.de/services/compute

Based on a long tradition in supercomputing at Universität Stuttgart, HLRS was founded in 1995 as a federal Centre for High-Performance Computing. HLRS serves researchers at universities and research laboratories in Germany and their external and industrial partners with high-end computing power for engineering and scientific applications.

Operation of its systems is done together with T-Systems, T-Systems sfr, and Porsche in the public-private joint venture hww (Höchstleistungsrechner für Wissenschaft und Wirtschaft). Through this co-operation a variety of systems can be provided to its users.

In order to bundle service resources in the state of Baden-Württemberg HLRS has teamed up with the Computing Centre of the University of Karlsruhe and the Centre for Scientific Computing of the University of Heidelberg in the hkz-bw (Höchstleistungsrechner-Kompetenzzentrum Baden-Württemberg).

Together with its partners HLRS provides the right architecture for the right application and can thus serve a wide range of fields and a variety of user groups.

Contact:
Höchstleistungsrechenzentrum
Stuttgart (HLRS)
Universität Stuttgart

Prof. Dr.-Ing. Dr. h.c. Michael M. Resch
Nobelstraße 19
70569 Stuttgart
Germany

Phone +49-711-685-8 72 69
resch@hlrs.de
www.hlrs.de

View of the HLRS BW-Grid IBM Cluster (Photo: HLRS)

Compute servers currently operated by HLRS are

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
NEC Hybrid Architecture	12 16-way nodes SX-9 with 8 TByte main memory + 5,600 Intel Nehalem cores 9 TB memory and 64 NVIDIA Tesla S1070	146	Capability Computing	German Universities, Research Institutes and Industry, D-Grid
IBM BW-Grid	3,984 Intel Harpertown cores 8 TByte memory	45.9	Grid Computing	D-Grid Community
Cray XT5m	896 AMD Shanghai cores 1.8 TByte memory	9	Technical Computing	BW Users and Industry
AMD Cluster	288 AMD cores 1.6 TByte memory	3.7	Technical Computing	Research Institutes and Industry

View of the HLRS hybrid NEC supercomputer (Intel Nehalem / NVIDIA S1070 / SX-9) (Photos: HLRS)

JÜLICH
FORSCHUNGSZENTRUM

The Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich enables scientists and engineers to solve grand challenge problems of high complexity in science and engineering in collaborative infrastructures by means of supercomputing and Grid technologies.

Provision of supercomputer resources of the highest performance class for projects in science, research and industry in the fields of modeling and computer simulation including their methods. The selection of the projects is performed by an international peer-review procedure implemented by the John von Neumann Institute for Computing (NIC), a joint foundation of Forschungszentrum Jülich, Deutsches Elektronen-Synchrotron DESY, and GSI Helmholtzzentrum für Schwerionenforschung.

Supercomputer-oriented research and development in selected fields of physics and other natural sciences by research groups of competence in supercomputing applications.

Implementation of strategic support infrastructures including community-oriented simulation laboratories and cross-sectional groups on mathematical methods and algorithms and parallel performance tools, enabling the effective usage of the supercomputer resources.

Higher education
for master and doctoral students in cooperation e.g. with the German Research School for Simulation Sciences.

Contact:
Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich

Prof. Dr. Dr. Thomas Lippert
52425 Jülich
Germany

Phone +49-24 61-61-64 02
th.lippert@fz-juelich.de
www.fz-juelich.de/jsc

View on the supercomputers JUGENE, JUST (storage cluster), HPC-FF and JUROPA in Jülich (Photo: Research Centre Jülich)

Compute servers currently operated by JSC are

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM Blue Gene/P "JUGENE"	72 racks 73,728 nodes 294,912 processors PowerPC 450 144 TByte memory	1,002.6	Capability computing	European Universities and Research Institutes, PRACE, DEISA
Intel Linux CLuster "JUROPA"	2,208 SMT nodes with 2 Intel Nehalem-EP quad-core 2.93 GHz processors each 17,664 cores 52 TByte memory	207	Capacity and Capability Computing	European Universities, Research Institutes and Industry, PRACE, DEISA
Intel Linux CLuster "HPC-FF"	1,080 SMT nodes with 2 Intel Nehalem-EP quad-core 2.93 GHz processors each 8,640 cores 25 TByte memory	101	Capacity and Capability Computing	EU Fusion Community
IBM Cell System "QPACE"	1,024 PowerXCell 8i processors 4 TByte memory	100	Capability Computing	QCD applications SFB TR55, PRACE

Blue Gene Extreme Scaling Workshop 2010

From March 22 - 24, Jülich Supercomputing Centre (JSC) organized the 2010 edition of its Blue Gene Scaling Workshop. Like the last workshop in October 2009, the main focus were application codes able to scale-up during the workshop to the complete Blue Gene/P system JUGENE which consists of 72 racks with a total of 294,912 cores – still the highest number of cores worldwide available in a single system.

Interested application teams had to submit short proposals which were evaluated with respect to the required extremely scaling, application-related constraints which had to be fulfilled by the JUGENE software infrastructure and the scientific impact that the codes could produce. Ten high-quality applications were selected. Participants came from Harvard University and Argonne National Laboratory in the United States, CORIA in France, ETH Zurich and EPFL Lausanne in Switzerland, as well as German groups from Karlsruher Institut für Technologie (KIT), Leibniz-Rechenzentrum (LRZ) in Munich, Zuse Institute Berlin (ZIB), Forschungszentrum Jülich, and the Universities of Marburg, Chemnitz, and Erlangen-Nuremberg.

During the workshop, the teams were supported by JSC parallel application experts, the JUGENE system administrators and one IBM MPI expert; however, the participants shared a lot of expertise and knowledge, too. More than half of the teams succeeded to submit one or more successful full 72 rack jobs during the course of the workshop, and three more, where algorithmic or load balancing issues required the codes to run on a power-of-two number of cores, scaled their applications to 64 racks (262,144 cores). One team "only" achieved to run on 32 racks (131,072 cores) due to a program bug which could not be resolved during the short time of the workshop. A total of 392 jobs were launched using 138.72 rack days of the total 164 rack days reserved for the workshop. This is an 84% utilization which could only be reached because of the extremely good stability of the system and the proactive maintenance of JSC and IBM staff.

Many interesting results were achieved. For example, the team from the Chair of System Simulation of the University Erlangen-Nuremberg worked towards direct numerical simulation of a billion fully resolved rigid bodies immersed in a fluid. The code is

build out of a combination of a lattice Boltzmann fluid simulation and a rigid body physics engine. For a case with densely packed particles representative for sedimentation or segregation processes, they achieved an efficiency which remains over 98% for a domain size of 80x80x80 lattice cells per core (weak scaling) over the whole range of cores.

The CORIA team from France used the YALES2 solver for highly resolved simulations of turbulent flows in complex geometries. YALES2 solves the low-Mach Navier-Stokes equations with a projection method for constant and variable density flows. These equations are discretized with a 4th-order central scheme in space and a 4th-order Runge-Kutta like scheme in time. The efficiency of projection approaches is usually driven by the performance of the Poisson linear solver. During the workshop they simulated an industrial swirl burner with a huge mesh of 21 billion tetrahedral cells. The mesh resolution of 50 µm allowed to have around 1,700 points in each direction of the combustor. Results obtained before the workshop with a 2.6 billion tetrahedral-based mesh are shown in Figure 1.

A team from Argonne National Laboratory and ETH Zurich investigated the scalability of the open-source Nek5000 spectral element code. Nek5000 employs high-order semi-implicit timestepping schemes that decouple the Navier-Stokes equations into independent advection, diffusion, and pressure projection substeps. It was successfully run on all 72 racks. Unexpected issues related to non-power-of-two processor partitioning resulted in non-optimal performance at 72 racks, but excellent scaling was observed for up to 64 racks. Over 71% parallel efficiency is realized for strong scaling from 8 racks to 64 racks, and sustained 172 TFLOPS were achieved. In addition these runs constituted the largest Nek5000 runs to date, with problem sizes exceeding 7 billion grid points.

All experiences and results achieved during the workshop are summarized in the technical report FZJ-JSC-IB-2010-03 available at:

http://www.fz-juelich.de/jsc/docs/autoren2010/mohr2/

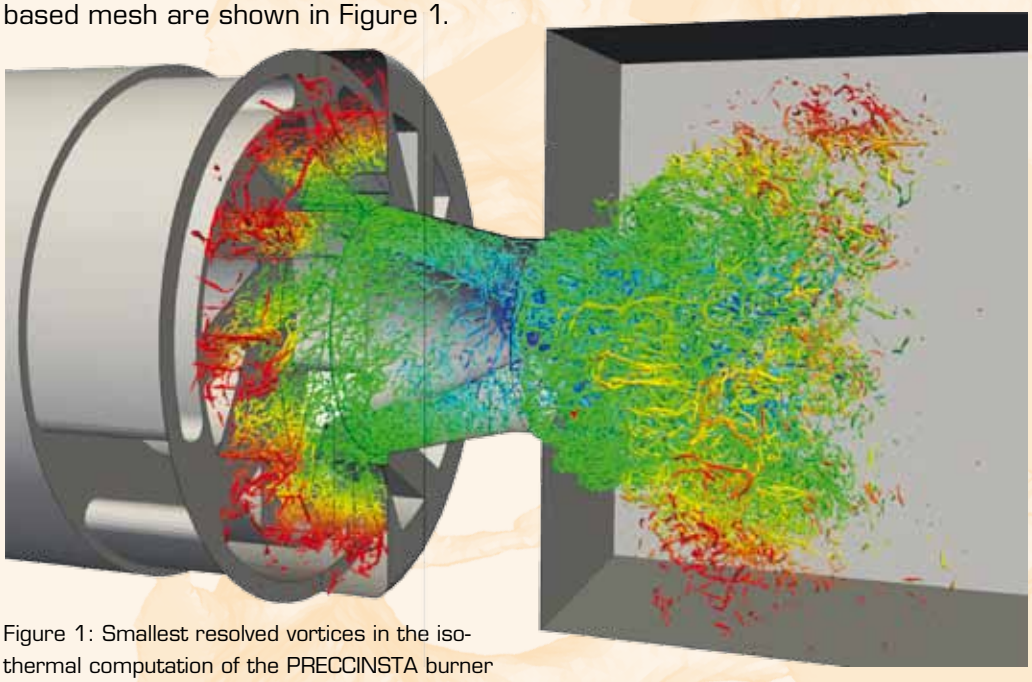


Figure 1: Smallest resolved vortices in the isothermal computation of the PRECCINSTA burner with a mesh of 2.6 billion cells

DEISA Extreme Computing

The DEISA Consortium has continued its work since 2002, first with EU FP6 support from 2004 until 2009, then with EU FP7 support from 2008 until 2011. The recent focus has been on the provisioning and operation of infrastructure services which allow its users to work efficiently within a distributed High Performance Computing environment based on the national HPC services. Through these effective HPC support services DEISA contributes to leading-edge European computational sciences.

The DEISA Extreme Computing Initiative (DECI), launched in 2005, has continually supported the most challenging supercomputing projects in Europe which require the special resources and skills of DEISA.

Through a regular public European Call for Extreme Computing Proposals the most challenging projects are selected, and the most appropriate supercomputer architectures are assigned for each accepted project. Thus DEISA has been opening up the respectively most powerful and suitable European HPC architectures for projects from any EU country, with or without national HPC service. This also mitigates the rapid performance decay of a single national supercomputer within its short lifetime cycle of typically about 5 years, as implied by Moore's law.

DEISA is maintaining teams of experts for operation, technology support, application enabling and application and user support. To advance computational sciences in the supercomputing area, application enabling is of key importance. For this challenging task, DEISA constituted and maintains the Applications Task Force (ATASKF) as a team of leading experts in High Performance Computing. Tasks include scaling of applications, data intensive computing and I/O optimization, workflows, coupled applications, determination of best suited architecture(s) and architecture dependent optimizations.

The number of Extreme Computing proposals received increases from year to year. After supporting 42 scientific projects out of 65 proposals from the DECI call 2008, 75 proposals were received from the call 2009, out of which 51 projects were awarded supercomputing resources and application enabling efforts through DEISA. The DECI call 2010 resulted in a record of 122 applications. Proposals submitted involve researchers from 30 countries - 22 in Europe and 8 from the continents of America, Asia, and Australasia. More than half a billion compute-hours have been requested.

Successful projects are chosen for their potential to achieve groundbreaking scientific results through the use of supercomputers, enabling them to run more detailed and accurate simulations of scientific problems than was previously possible.

A total of 55 new projects could be accepted recently, with an award of more than 90 million processor hours on one or more of the 13 DEISA sites which operate 15 of the Top 100 most powerful supercomputers in the world.

For the first time this year, national supercomputers in Switzerland and in Sweden have also been made available via DECI, in addition to the national supercomputers in Germany, UK, France, Italy, Spain, the Netherlands and Finland.

Through DECI, now in its sixth year of operation, over 180 universities or research institutes with scientists from 35 countries have benefitted. Collaborating researchers from outside Europe come from four other continents. The scientific projects cover major areas of science including Astro Sciences, Earth Sciences, Engineering, Life Sciences, Materials Science, Plasma Physics, Particle Physics and Quantum Chromodynamics.

In a recent press release Jukka Heikkinen, a scientist from Finland who has participated in several collaborative DECI projects, said: "Realization of fusion energy for future requires today a vast amount of simulations of microscopic plasma turbulence at real size of the burning fuel. This is possible only with massively parallelized kinetic particle codes which can be only run in world's top supercomputers. DECI has made some of the best European supercomputers available to the fusion scientists developing the plasma turbulence codes. It has helped, in particular, in characterization of turbulence in several tokamak experiments and in understanding the plasma energy confinement transition at the tokamak plasma edge. In addition to providing flexibility and quickness in distributing the runs on several computers, speeding up thus the science making, the DECI scheme has enforced the harmonization of the user solutions for coding and library use. The latter has made the codes more independent of the computer architecture, helping thus the world-wide integration of the codes."

Scientific achievements through DEISA and especially through DECI have been reported in the booklet "DEISA – Advancing Science in Europe", the DEISA Digest 2008, the DEISA Video, and the recently appeared DEISA Digest 2010. Hard copies are available from the four German DEISA sites HLRS, JSC, LRZ, and RZG, and PDF versions can be downloaded from:

http://www.deisa.eu/publications/results

Acknowledgments

The DEISA Consortium thanks the European Commission for support through contracts FP6-508830, FP6-031513, and FP7-222919

• Hermann Lederer

Rechenzentrum Garching der Max-Planck-Gesellschaft

Figure 1: DEISA Digest 2010 and DEISA Video, available from the German DEISA sites HLRS, JSC, LRZ and RZG

1st SimLab Porting Workshop

From June 9 - 11, the JSC Simulation Laboratories organized a Porting Workshop to promote HPC activities within their respective/target scientific communities.

About half of the 26 selected participants were either FZJ researchers or students at the GRS, while the other half came from other German labs and universities. Their scientific profiles matched one of the research fields covered by the four SimLabs: Biology, Climate, Molecular Systems, and Plasma Physics. Although the participants were distributed across the SimLabs quite evenly, the HPC expertise and experience varied widely from group to group.

The programme included a combination of lectures and hands-on sessions held largely by JSC SimLab staff members. Tutorial lectures on advanced computational methods for HPC provided an opportunity for the SimLabs to present their activities and scientific expertise. Complementary hands-on sessions covered a wide range of practical issues, including accessing the systems, compiling and executing jobs, single core optimization, Jugene's torus network, and HPC tools usage. Additional contributions came from three invited speakers: Frank Schmitz (KIT) presented an overview of the current SimLab status at KIT, while Paolo Carloni (GRS) and Maxim Fedorov (MPI Leipzig), reviewed topical HPC challenges in biology and molecular systems respectively.

Finally, the workshop concluded with round-table discussions within each SimLab group about current supercomputing issues particular to each community. Overall the meeting proved to be a highly effective community outreach event and generated invaluable feedback for the call for high-level HPC support to be announced by JSC later this year.

All tutorial talks and exercises are publicly available at the workshop's webpage:

http://www.fz-juelich.de/jsc/simlab-porting-workshop

LRZ Workshop Molecular Modelling on Supercomputers

The Leibniz Supercomputing Centre held a two day workshop on molecular modelling on supercomputers on April 14 - 15 in Garching with 15 participants from the life science community. The objective of the workshop was to help life science researchers approach challenging problems in computational chemistry and biology. The workshop gave an introduction to two major software packages for molecular modelling. Day one was held by Schrödinger Inc. who presented an introduction to and advanced features of the Schrödinger suite of programs. Each participant was provided with a laptop for the extensive hands-on session by Schrödinger Inc. that had the software and data files pre-installed. Day two dealt with the molecular dynamics simulation software NAMD (NAnoscale Molecular Dynamics) and VMD (Visual Molecular Dynamics). NAMD is a molecular dynamics code that scales to tens of thousands of processor cores and was recently ported to GPGPUs (CUDA). There were interesting discussions how to bridge the different important scales from single molecule first principle quantum mechanics over force-field calculations up to coarse-grained molecular dynamics. The presentations covered many interesting topics, such as ligand binding, visualization of biomolecules and trajectories, and automated generation of configuration files for simulation runs as well as the job submission on the supercomputer of the LRZ.

Summer School on fast Methods for Long-Range Interactions in complex Systems

From September 6 - 10, 2010 Jülich Supercomputing Centre organized a Summer School on Fast Methods for Long-Range Interactions in Complex Systems, which was financially supported by the Wilhelm and Else Heraeus Foundation.

About 30 participants from five countries came to Jülich to learn about modern algorithms which efficiently solve the Coulomb problem and reduce the numerical complexity from $O(N^2)$ to $O(N \log(N))$ or $O(N)$. Ten lecturers from Universities of Bielefeld, Chemnitz, Stuttgart, Wuppertal and the Forschungszentrum Jülich presented state-of-the-art methods, algorithms and implementations of various approaches to tackle the long-range interactions between particles.

The motivation for organizing this Summer School arose from the BMBF funded network project ScaFaCoS (Scalable Fast Coulomb Solver), which aims to develop a scalable library for various fast methods solving the long-range interactions between particles in complex systems. Since different physical problems have different requirements the school covered a variety of algorithms. The spectrum of presentations ranged from simple cutoff methods to Fourier-based methods (P3M), hierarchical tree methods, multigrid techniques and the fast multipole method (FMM). For each method, emphasis was given to the theoretical foundation and derivation, the error control of the approximations and the parallelization.

To get participants acquainted with parallel computing, the first day included a special introduction to MPI followed by a hands-on programming session. Further practical sessions complemented the talks on theoretical foundations and implementation issues of different algorithms in the afternoons, where specific program packages, e.g. the Soft Matter Code ESPResSo or the plasma physics code PEPC were introduced.

17th EuroMPI Conference at HLRS

The 17th EuroMPI was held at Stuttgart, Germany on September 12-15, 2010. About 90 scientists from all over the world met to discuss the most pressing issues in MPI programming.

Bill Gropp from the University of Illinois Urbana-Champaign laid out the issues in exascale computing and presented strategies of how MPI can address these issues.

Jan Westerholm from the Abo Akademi University, Finland addressed the issue of end users and gave examples of how they react to the increase in number of processors.

Jack Dongarra from the University of Tennessee laid out the challenges of extreme scale computing. He dived into the hardware and software issues of exascale computing and issued a call to action at the international level.

Jesus Labarta from the Barcelona Supercomputing Center spoke about details at scale in performance analysis. Drawing from his long experience he presented new approaches to solve performance analysis problems.

On the final day Rolf Hempel from the German Center for Aerospace bridged the gap between programming and end users. He presented a number of activities in the field of interactive visualization of large simulation data sets. Laying out the examples he pointed at the most pressing issue of how to handle the massive amount of data that large scale systems create.

A visit to the Mercedes Museum at Stuttgart gave participants a chance to put their work into an industrial context and to understand the mutual impact of technology and historical development.

The 17th EuroMPI meeting was a success for the community. The organizers would like to thank Edgar Gabriel from the University of Houston and Rainer Keller from the University of Stuttgart for their efforts in putting together an outstanding scientific program. We also thank Neriman Emre and Stefan Wesner from HLRS for taking care of all organizational issues.

Slides of the presentations and some photographs of the event can be found at:

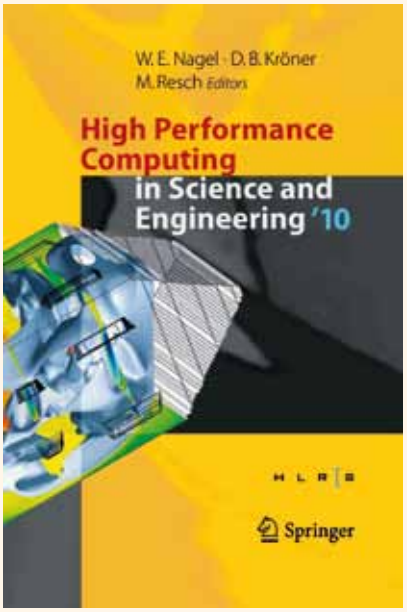
http://www.eurompi2010.org/

Hope to see you next year at EuroMPI 2011 at Santorini, Greece, September 18-21, 2011

Publications in Computational

Nagel, W. E., Kröner, D. B., Resch, M. (Editors), Transactions of the High Performance Computing Center Stuttgart (HLRS) 2010, 1st Edition, 2011, 500 p., Hardcover, ISBN 978-3-642-15747-9, due: November 2010.

This book presents the state-of-the-art in simulation on supercomputers. Leading researchers present results achieved on systems of the High Performance Computing Center Stuttgart (HLRS) for the year 2010. The reports cover all fields of computational science and engineering, ranging from CFD to computational physics and chemistry to computer science, with a special emphasis on industrially relevant applications. Presenting results for both vector systems and microprocessor-based systems, the book makes it possible to compare the performance levels and usability of various architectures. As HLRS operates the largest NEC SX-8 vector system in the world, this book gives an excellent insight into the potential of vector systems, covering the main methods in High Performance Computing. Its outstanding results in achieving the highest performance for production codes are of particular interest for both scientists and engineers. The book includes a wealth of color illustrations and tables. Each project description includes information on the respective scientific background, results achieved and methods used. The book also offers insights into the performance, scalability, and latest enhancements of the applications employed.



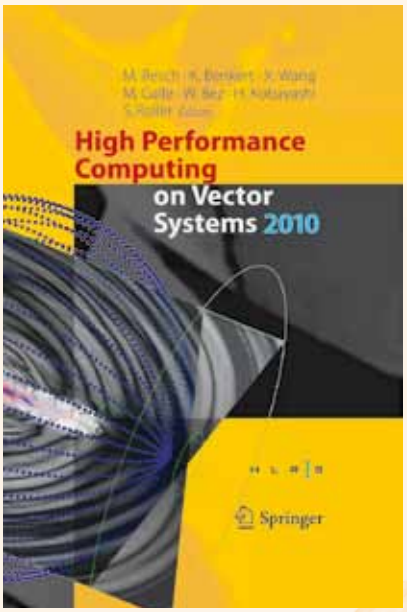
Keller, R., Gabriel, E., Resch, M., Dongarra, J. (Editors), Recent Advances in the Message Passing Interface, 17th European MPI User's Group Meeting, EuroMPI 2010, Stuttgart, Germany, September 12-15, 2010, Proceedings, Series: Lecture Notes in Computer Science, Vol. 6305, Subseries: Programming and Software Engineering, 1st Edition., 2010, XIV, 308 p., Softcover, ISBN 978-3-642-15645-8.

This book constitutes the proceedings of the 17th European MPI User's Group Meeting on Recent Advances in the Message Passing Interface held in Stuttgart in September 2010.

Science and Engineering

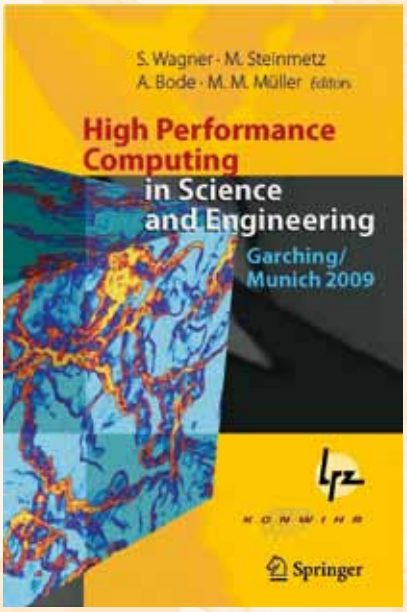
Resch, M., Benkert, K., Wang, X., Galle, M., Bez, W., Kobayashi, H., Roller, S. (Editors),
High Performance Computing on Vector Systems 2010, 1st Edition., 2010, XII, 198 p., Hardcover, ISBN 978-3-642-11850-0, 2010.

This book presents the state-of-the-art in High Performance Computing and simulation on modern supercomputer architectures. It covers trends in hardware and software development in general and specifically the future of vector-based systems and heterogeneous architectures. The application contributions cover computational fluid dynamics, material science, medical applications and climate research. Innovative fields like coupled multi-physics or multi-scale simulations are presented. All papers were chosen from presentations given at the 11th Teraflop Workshop held in October 2009 at Tohoku University, Japan, and the 12th Teraflop Workshop held in March 2010 at the Höchstleistungsrechenzentrum Stuttgart (HLRS), Germany.



Wagner, S., Steinmetz, M., Bode, A., Müller, M. M. (Editors),
High Performance Computing in Science and Engineering, Garching/Munich 2009, Transactions of the 4th Joint HLRB and KONWIHR Review and Results Workshop, 780 p., Hardcover, ISBN 978-3-642-13871-3, 2010.

This book presents selected project results obtained at the current high-performance computer in Bavaria (HLRB II). The research areas covered include Computer Sciences, Computational Fluid Dynamics, Astrophysics, High-Energy Physics, Geo Sciences, Condensed Matter Physics, Chemistry, and Bio Sciences. Thus, the book provides an overview of the broad range of applications requiring high-performance computers to solve challenging numerical problems. Each project description includes information on the respective scientific background, results achieved and methods used. The book also offers insights into the performance, scalability, and latest enhancements of the applications employed.



HLRS Scientific Tutorials and Workshop Report and Outlook

People are more and more fascinated by the performance of GPUs in computational sciences. In June 2010, Dr. Oliver Mangold offered his course **"GPU Programming using CUDA"** the first time at HLRS. We got over 80 registrations and therefore we had to offer a second course in July. On our NEC Nehalem Cluster, the participants could test the GPU performance within several exercises. This series will be continued December 2010 and July 2011.

Another highlight is the **Introduction to Computational Fluid Dynamics**. This course was initiated at HLRS by Dr.-Ing. Sabine Roller. She is now a professor at the German Research School at RWTH Aachen, and with this, the course was held the first time there. In April 2011, it will be presented again in Stuttgart. The presented twice, in March 2010 at HLRS in Stuttgart and in October 2010 at LRZ in Garching.

Tutorials at Scientific Conferences in 2010

Hybrid MPI and OpenMP Parallel Programming. Rolf Rabenseifner, Georg Hager, Gabriele Jost. Half-day Tutorial No. M-02 at Super Computing 2010, SC10, New Orleans, Louisiana, USA, November 13-19, 2010.

Introduction to PGAS (UPC and CAF) and Hybrid for Multicore Programming. Alice E. Koniges, Katherine Yelick, Rolf Rabenseifner, Reinhold Bader, David Eder. Full-day Tutorial No. S-10 at Super Computing 2010, SC10, New Orleans, Louisiana, USA, November 13-19, 2010.

Hybrid MPI & OpenMP Parallel Programming. Rolf Rabenseifner, Georg Hager, Gabriele Jost. Tutorial at the International Supercomputing Conference 2010, ISC'10, Hamburg. Germany, May 30 – June 3, 2010.

Next year, HLRS will have again a number of different courses and workshops. One of the flagships of our courses is the week on **Iterative Solvers and Parallelization.** Prof. A. Meister and Prof. B. Fischer teach basics and details on Krylov Subspace Methods. Lecturers from HLRS give lessons on distributed memory parallelization with the Message Passing Interface (MPI) and shared memory multi-threading with OpenMP. This course will be emphasis is placed on explicit finite volume methods for the compressible Euler equations. Moreover outlooks on implicit methods, the extension to the Navier-Stokes equations and turbulence modeling are given. Additional topics are classical numerical methods for the solution of the incompressible Navier-Stokes equations, aero-acoustics and high order numerical methods for the solution of systems of partial differential equations.

With the **new Petaflop/s system at HLRS in 2011**, a special platforms course will introduce into the details of the new computer.

The **Parallel Programming Workshop**, Sep. 26 – 30 at HLRS, will have 3 parts: The first two days of this course are dedicated to parallelization with the Message Passing Interface (MPI). Shared memory multi-threading is taught on the third day. In the last two days, advanced topics are discussed. As in all courses, hands-on sessions (in C and Fortran) will help users to immediately understand the parallelization methods. The course language is English.

Several 3- and 4-day courses on **MPI & OpenMP** will be presented at different locations all over the year. We also continue our series of **Fortran for Scientific Computing** in January and July 2011, mainly visited by PhD students from Stuttgart and other universities in Germany to learn not only the basics of programming, but also to get an insight on the principles of developing high-performance applications with Fortran.

With **Unified Parallel C (UPC) and Co-Array Fortran (CAF)** in December 2010 and July 2011, the participants will get an introduction of partitioned global address space (PGAS) languages.

The question of PGAS languages and programming on clusters of multi-core nodes was also addressed by HLRS

in three **tutorials at scientific conferences** 2010, see the box on the left page.

In the table below, you can find the whole HLRS series of training courses in 2011. They are organized at HLRS and also at several other HPC institutions: LRZ Garching, NIC/ZAM (FZ Jülich), and CSCS (Manno, CH), ZIH (TU Dresden), TUHH (Hamburg Harburg), and GRS/RWTH (Aachen).

2011 – Workshop Announcements
Scientific Conferences and Workshops at HLRS, 2011
14th Teraflop Workshop (March, not yet fixed)
10th HLRS/hww Workshop on Scalable Global Parallel File Systems (April, 11-15)
5th HLRS Parallel Tools Workshop (July, not yet fixed)
High Performance Computing in Science and Engineering - The 14th Results and Review Workshop of the HPC Center Stuttgart (planned in October)
IDC International HPC User Forum (October, not yet fixed)
Parallel Programming Workshops: Training in Parallel Programming and CFD
GPU Programming using CUDA (HLRS, January 24-25, July 7-8, and December 8-9)
Parallel Programming and Parallel Tools (TU Dresden, ZIH, February 14-17)
Iterative Linear Solvers and Parallelization (HLRS, February 28 – March 4)
Introduction to Computational Fluid Dynamics (HLRS, April 4-8)
Platforms at HLRS (HLRS, April 18-19)
Unified Parallel C (UPC) and Co-Array Fortran (CAF) (HLRS, July 5-6, and Dec. 6-7)
5th Parallel Tools Workshop (TU Dresden, ZIH, July, not yet fixed)
Parallel Programming with MPI & OpenMP (TU Hamburg-Harburg, August 1-3)
Iterative Linear Solvers and Parallelization (LRZ, Garching, September 12-16)
Introduction to Computational Fluid Dynamics (GRS / RWTH Aachen, September 26-30)
Message Passing Interface (MPI) for Beginners (HLRS, October 10-11)
Shared Memory Parallelization with OpenMP (HLRS, Oktober 12)
Advanced Topics in Parallel Programming (HLRS, Oktober 13-14)
Parallel Programming with MPI & OpenMP (FZ Jülich, JSC, November 28-30)
Training in Programming Languages at HLRS
Fortran for Scientific Computing (January 17-21, and July 11-15)
URLs: http://www.hlrs.de/events/
http://www.hlrs.de/organization/sos/par/services/training/course-list/
https://fs.hlrs.de/projects/par/events/2011/parallel_prog_2011/

High-Performance Computing Courses

Introduction to the Programming and Usage of the Supercomputer Resources in Jülich

Date and Location
November 25 - 26, 2010
May 16 - 17, 2011
JSC, Research Centre Jülich

Contents
This course gives an overview of the supercomputers JUROPA and JUGENE. Especially new users will learn how to program and use these systems efficiently. Topics discussed are: system architecture, usage model, compilers, tools, monitoring, MPI, OpenMP, performance optimization, mathematical software, and application software.

Webpage*
http://www.fz-juelich.de/jsc/neues/termine/supercomputer/

Parallel Programming with MPI, OpenMP and PETSc

Date and Location
November 29 – December 1, 2010
JSC, Research Centre Jülich

Contents
The focus is on programming models MPI, OpenMP, and PETSc. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of the Message Passing Interface (MPI) and the shared memory directives of OpenMP. Course language is English.

This course is organized by JSC in collaboration with HLRS. Presented by Dr. Rolf Rabenseifner, HLRS.

Webpage*
http://www.fz-juelich.de/jsc/neues/termine/mpi-openmp/

Intel Ct Training

Date
Tuesday, November 30, 2010 – Wednesday, December 1, 2010 (tentative)

Prerequisites
Good knowledge of C++

Lecturer
Hans Pabst, Intel

Webpage
The content and date of the course are not yet finalized. Please contact: http://www.lrz.de/services/compute/courses

Parallel I/O and portable Data Formats

Date and Location
December 6 - 8, 2010
JSC, Research Centre Jülich

Contents
This course will introduce MPI parallel I/O and portable, self-describing data formats, such as HDF5 and NetCDF. Participants should have experience in parallel programming in general, and either C/C++ or Fortran in particular.

Webpage*
http://www.fz-juelich.de/jsc/neues/termine/parallelio/

GPGPU Programming

Date & Location
Tuesday, December 7 – Thursday, December 9, 2010
LRZ Building, University campus Garching, near Munich

Contents
Heterogeneous GPGPU computing promises tremendous acceleration of applications. This programming workshop includes hands-on sessions, application examples and an introduction to CUDA, CAPS, cuBLAS, cuFFT, the Portland Group Fortran Compiler, pycuda, and R. The intended audience includes scientists which want to port their simulation software to GPGPUs as well as people interested in a short overview of the available programming techniques.

Prerequisites
Participants should have a fair understanding of programming in general and should have knowledge in at least one of the following programming languages: Fortran, C/C++, Python.

Webpage
http://www.lrz.de/services/compute/courses/#GPU

* Due to a web relaunch at Research Centre Jülich in November, the URLs may change. Please start your search at http://www.fz-juelich.de/jsc

and Tutorials

Unified Parallel C (UPC) and Co-Array Fortran (CAF)

Dates & Locations
December 14 - 15, 2010
Stuttgart, HLRS

Contents
Partitioned Global Address Space (PGAS) is a new model for parallel programming. Unified Parallel C (UPC) and Co-array Fortran (CAF) are PGAS extensions to C and Fortran. PGAS languages allow any processor to directly address memory/data on any other processors. Parallelism can be expressed more easily compared to library based approches as MPI. Hands-on sessions (in UPC and/or CAF) will allow users to immediately test and understand the basic constructs of PGAS languages.

Webpage
http://www.hlrs.de/events/

GPU Programming using CUDA

Dates & Locations:
December 16 - 17, 2010
Stuttgart, HLRS

July 7 - 8, 2011
Stuttgart, HLRS

Contents
The course provides an introduction to the programming language CUDA, which is used to write fast numeric algorithms for NVIDIA graphics processors (GPUs). Focus is on the basic usage of the language, the exploitation of the most important features of the device (massive parallel computation, shared memory, texture memory) and efficient usage of the hardware to maximize performance. An overview of the available development tools and the advanced features of the language is given.

Webpage
http://www.hlrs.de/events/

Fortran for Scientific Computing

Date & Location
January 17 - 21, 2011
Stuttgart, HLRS

Contents
This course is dedicated for scientists and students to learn (sequential) programming scientific applications with Fortran. The course teaches newest Fortran standards. Hands-on sessions will allow users to immediately test and understand the language constructs.

Webpage
http://www.hlrs.de/events/

Scientific 3D-Animation with Blender

Date & Location
Thursday, January 13, 2010 –
Friday, January 14, 2010
LRZ Building, University campus
Garching, near Munich

Contents
The course gives an introduction into the high-quality visualization and animation of scientific data using the open source 3D software blender. Based on a real-world example (protein molecule), the course covers all steps of a project: importing the data, cleaning up the geometry, assigning materials, lighting, keyframing and post-production.

Course language
German

Webpage
http://www.lrz.de/services/compute/courses/#Blender

Introduction to the PGAS Languages UPC and CAF

Date & Location
January 19, 2011
LRZ Building, University campus
Garching, near Munich

Prerequisites
Course participants should have a good working knowledge of Fortran and/or C.

High-Performance Computing Courses

Contents

Partitioned global address space languages have emerged as an alternative to other parallel programming models, promising a shorter development cycle due to improved programmability while keeping the performance level on par with MPI. This course introduces the parallel facilities integrated into the Fortran language (coarrays) and the C language (unified parallel C), respectively. A hands-on session allows to experiment with the new concepts, using prototype implementations on LRZ's HPC systems.

Webpage

http://www.lrz.de/services/compute/courses/#PGAS

Introduction to Molecular Modeling on Supercomputers

Date & Location

January 25 - 27, 2011
LRZ Building, University campus
Garching, near Munich

Contents

The course gives an introduction into the simulation of molecules based on several software packages on the supercomputers at LRZ Garching. (Maestro, Desmond, VMD, NAMD, Gromacs). This also includes an introduction to the remote visualization services at LRZ as well as hands-on sessions. The course focuses on biomolecules and targets the life science community. Hands-on sessions featuring example applications are given.

Prerequisites

Participants should have some basic knowledge in programming and life science software.

Webpage

http://www.lrz.de/services/compute/courses/#MolMod

Programming with Fortran

Date & Location

Monday, February 7 -
Friday, February 11, 2011
LRZ Building, University campus
Garching, near Munich
A video conference to other sites can be set up if there is sufficient interest.

Prerequisites

Course participants should have basic UNIX/Linux knowledge (login with secure shell, shell commands, simple scripts, editor vi or emacs).

Contents

This course is targeted at scientists with little or no knowledge of the Fortran programming language, but need it for participation in projects using a Fortran code base, for development of their own codes, and for getting acquainted with additional tools like debugger and syntax checker as well as handling of compilers and libraries. The language is for the most part treated at the level of the Fortran 95 standard; features from Fortran 2003 are limited to improvements on the elementary level. Advanced Fortran features like object-oriented programming or coarrays will be covered in a follow-on course in September. To consolidate the lecture material, each day's approximately 4 hours of lecture are complemented by 3 hours of hands-on sessions.

Webpage

http://www.lrz.de/services/compute/courses/#BasicFortran

Parallel Programming with MPI, OpenMP amd PETSc

Date & Location

February 14 - 17, 2011
Dresden, ZIH

Contents

The focus is on programming models MPI, OpenMP, and PETSc. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of the Message Passing Interface (MPI) and the shared memory directives of OpenMP. The last day is dedicated to tools for debugging and performance analysis of parallel applications. This course is organized by ZIH in collaboration with HLRS.

Webpage

http://www.hlrs.de/events/

Parallel Programming with R

Date & Location

Tuesday, Febuary 15, 2011
LRZ Building, University campus
Garching, near Munich

* Due to a web relaunch at Research Centre Jülich in November, the URLs may change. Please start your search at http://www.fz-juelich.de/jsc

and Tutorials

Prerequisites
Participants should have some basic knowledge in programming with R.

Contents
R is known as a very powerful language for statistics, but it has also evolved into a tool for the analysis and visualization of large data sets which are typically obtained from supercomputing applications. The course teaches the use of the dynamic language R for parallel programming of supercomputers and features rapid prototyping of simple simulations. Several parallel programming models including Rmpi, snow, multicore, and gputools are presented which exploit the multiple processors that are standard on modern supercomputer architectures. Hands-on sessions with example applications are given.

Webpage
http://www.lrz.de/services/compute/courses/#RPar

Visualization of large Data Sets on Supercomputers

Date & Location
Wednesday, February 23, 2011
LRZ Building, University campus
Garching, near Munich

Contents
The results of supercomputing simulations are data sets which have grown considerably over the years, giving rise to a need for parallel visualization software packages. The course focuses on the software packages Paraview, Visit, Vapor and R and their use to visualize and analyse large data sets generated by supercomputer applications which are typically generated in the fields of CFD, molecular modelling, astrophysics, quantum chemistry and similar. Hands-on sessions featuring example applications are given.

Prerequisites
Participants should have some basic knowledge in programming and use of the Linux operating environment.

Webpage
http://www.lrz.de/services/compute/courses/#VisDat

Iterative Linear Solvers and Parallelization

Date & Location
February 28 - March 4, 2011
Stuttgart, HLRS

Contents
The focus is on iterative and parallel solvers, the parallel programming models MPI and OpenMP, and the parallel middleware PETSc. Thereby, different modern Krylov Subspace Methods (CG, GMRES, BiCGSTAB ...) as well as highly efficient preconditioning techniques are presented in the context of real life applications. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of iterative solvers, the Message Passing Interface (MPI) and the shared memory directives of OpenMP. This course is organized by Uni. Kassel, HLRS, and IAG.

Webpage
http://www.hlrs.de/events/

Parallel Programming of High Performance Systems

Date & Location
Monday, March 7 -
Friday, March 11, 2011
LRZ Building, University campus
Garching, near Munich.
RRZE building, University campus
Erlangen (via video conference)

Prerequisites
Good working knowledge of at least one of the standard HPC languages: Fortran 95, C or C++.

Contents
This course, a collaboration of Erlangen Regional Computing Centre (RRZE) and LRZ, is targeted at students and scientists with interest in programming modern HPC hardware, specifically the large scale parallel computing systems available in Jülich, Stuttgart and Munich.

Each day is comprised of approximately 4 hours of lectures and 3 hours of hands-on sessions.

Day 1
- Introduction to
  High Performance Computing
- HPC systems
  at LRZ and in Germany

High-Performance Computing Courses

- Technical aspects of software engineering: development process, usage of libraries, memory management, tools for code documentation, development environments
- Tools: Using Secure Shell, Screen, SVN and Make

Day 2

- Processor architectures: Register, cache, locality, performance metrics
- Overview of accelerators
- Basic features of parallel programming with MPI and OpenMP
- Using the batch systems and software environment at LRZ/RRZE

Day 3

- Architecture independent optimization strategies: unrolling, blocking, dependencies, C++ issues, bandwidth issues, performance projections
- Architecture specific optimization strategies: compiler switches, avoiding cache thrashing, exploiting SIMD capabilities
- Features of current processor architectures relevant for HPC
- Using performance libraries Days 4 and 5
- Parallel algorithms: data parallelism, domain decomposition, task parallelism, master-worker, granularity, load balancing, scalability models
- Parallel Architectures: multi-core, multi-socket, ccNUMA, cache coherence and affinity, tools for handling memory affinity
- Distributed memory architectures, interconnects and their topologies
- Advanced OpenMP
- Advanced MPI

Webpage

http://www.lrz.de/services/compute/courses/#ParallelProgramming

Advanced Topics in High Performance Computing

Date& Location

Monday, March 21 - Wednesday, March 23, 2011
LRZ Building, University campus Garching, near Munich

Prerequisites

Good MPI and OpenMP knowledge as presented in the course Parallel programming of High Performance Systems.

Contents

In this add-on course to the parallel programming course special topics are treated in more depth, in particular performance analysis and I/O.
It is provided in collaboration of Erlangen Regional Computing Centre (RRZE) and LRZ.

Each day is comprised of approximately 5 hours of lectures and 2 hours of hands-on sessions.

Day 1

- Intel Tracing Tools: MPI tracing and correctness checking
- Intel Threading tools for OpenMP correctness checking and profiling
- Profiling on SGI Altix systems: histx and lipfpm
- Introduction to Scalasca

Day 2

- Parallel application performance analysis with Scalasca
- Parallel input/output with MPI-IO

Day 3

- I/O tuning on high performance file systems
- Portability of binary files, big/little endian issues
- Using I/O libraries (pNetCDF, HDF5)

Webpage

http://www.lrz.de/services/compute/courses/#AdvancedTopics

Introduction to Computational Fluids Dynamics

Date & Location

April 4 - 8, 2011

Contents

Numerical methods to solve the equations of Fluid Dynamics are presented. The main focus is on explicit Finite Volume schemes for the compressible Euler equations. Hands-on sessions will manifest the content of the lectures. Participants will learn to implement the algorithms, but also to apply existing

* Due to a web relaunch at Research Centre Jülich in November, the URLs may change. Please start your search at http://www.fz-juelich.de/jsc

and Tutorials

software and to interpret the solutions correctly. Methods and problems of parallelization are discussed. This course is based on a lecture and practical awarded with the "Landeslehrpreis Baden-Württemberg 2003" and organized by HLRS, IAG, and University of Kassel.

Webpage
http://www.hlrs.de/events/

Introduction to the PGAS Languages UPC and CAF

Date and Location
April 12, 2011
JSC, Research Centre Jülich

Contents
Partitioned global address space languages have emerged as an alternative to other parallel programming models, promising a shorter development cycle due to improved programmability while keeping the performance level on par with MPI. This course introduces the parallel facilities integrated into the Fortran language (coarrays) and the C language (unified parallel C), respectively. A hands-on session allows to experiment with the new concepts. Course participants should have a good working knowledge of Fortran and/or C.

Webpage*
http://www.fz-juelich.de/jsc/neues/termine/pgas/

Platforms at HLRS

Date & Location
April 18 - 19, 2011
Stuttgart, HLRS

Contents
The course is focused on the usage of the HPC platforms at HLRS.

Webpage
http://www.hlrs.de/events/

Unified Parallel C (UPC) and Co-Array Fortran (CAF)

Dates & Locations
July 5 - 6, 2011
Stuttgart, HLRS

Contents
Partitioned Global Address Space (PGAS) is a new model for parallel programming. Unified Parallel C (UPC) and Co-array Fortran (CAF) are PGAS extensions to C and Fortran. PGAS languages allow any processor to directly address memory/data on any other processors. Parallelism can be expressed more easily compared to library based approches as MPI. Hands-on sessions (in UPC and/or CAF) will allow users to immediately test and understand the basic constructs of PGAS languages.

Webpage
http://www.hlrs.de/events/

Fortran for Scientific Computing

Date & Location
July 11 - 15, 2011
Stuttgart, HLRS

Contents
This course is dedicated for scientists and students to learn (sequential) programming scientific applications with Fortran. The course teaches newest Fortran standards. Hands-on sessions will allow users to immediately test and understand the language constructs.

Webpage
http://www.hlrs.de/events/

Education in Scientific Computing

Date & Location
August 1 - October 7, 2011
JSC, Research Centre Jülich

Contents
Guest Student Programme Scientific Computing to support education and training in the fields of supercomputing. Application deadline is April 30, 2011. Students of Computational Sciences, Computer Science and Mathematics can work 10 weeks in close collaboration with a local scientific host on a subject in their field.

Webpage*
http://www.fz-juelich.de/jsc/gsp/

inSiDE is published two times a year by The GAUSS Centre for Supercomputing (HLRS, LRZ, JSC)

Publishers

Prof. Dr. H.-G. Hegering | Prof. Dr. Dr. T. Lippert | Prof. Dr. Dr. h.c. M. M. Resch

Editor & Design

F. Rainer Klank, HLRS	klank@hlrs.de
Carina Möhlig, HLRS	moehlig@hlrs.de

Authors

Momme Allalen	allalen@lrz.de
Christoph Altmann	altmann@iag.uni-stuttgart.de
Raul Angulo	reangulo@mpa-garching.mpg.de
Fakher F. Assaad	assaad@physik.uni-wuerzburg.de
Carlton Baugh	c.m.baugh@durham.ac.uk
Thomas Baumann	baumann@hlrs.de
Katharina Benkert	benkert@hlrs.de
Matthias Brehm	brehm@lrz.de
Alexander Broicher	insilico-biotechnology.com
Shaun Cole	shaun.Cole@durham.ac.uk
Norbert Eicker	n.eicker@fz-juelich.de
Christian Engfer	engfer@iag.uni-stuttgart.de
Dietmar Erwin	d.erwin@fz-juelich.de
Carlos Frenk	c.s.frenk@durham.ac.uk
Gregor Gassner	gassner@iag.uni-stuttgart.de
Paul Gibbon	p.gibbon@fz-juelich.de
José Gracia	gracia@hlrs.de
Kai Hertel	kai.hertel@informatik.uni-erlangen.de
Sabine Höfler-Thierfeldt	s.hoefler-thierfeldt@fz-juelich.de
Willi Homberg	w.homberg@fz-juelich.de
Christine Jandl	christine.jandl@informatik.uni-erlangen.de
Adrian Jenkins	a.r.jenkins@durham.ac.uk
Rainer Keller	keller@hlrs.de
Alexander Kipp	kipp@hlrs.de
Ewald Krämer	kraemer@iag.uni-stuttgart.de
Thomas C. Lang	lang@physik.uni-wuerzburg.de
Hermann Lederer	lederer@rzg.mpg.de
Thorsten Lutz	lutz@iag.uni-stuttgart.de
Zi Yang Meng	meng@theo3.physik.uni-stuttgart.de
Claus-Dieter Munz	munz@iag.uni-stuttgart.de
Alejandro Muramatsu	mu@theo3.physik.uni-stuttgart.de
Walter Nadler	w.nadler@fz-juelich.de
Christoph Pflaum	pflaum@informatik.uni-erlangen.de
Zhabiz Rahimi	zhabiz.rahimi@iisb.fraunhofer.de
Orlando Rivera	rivera@lrz.de
Sven Schmid	sven.schmid@gmx.de
Volker Springel	volker.springel@h-its.org
Marc Staudenmaier	staudenmaier@iag.uni-stuttgart.de
Hinnerk Stüben	stueben@zib.de
Stefan Wessel	wessel@theo3.physik.uni-stuttgart.de
Simon White	swhite@mpa-garching.mpg.de

If you would like to receive inSiDE regularly, send an email with your postal address to klank@hlrs.de