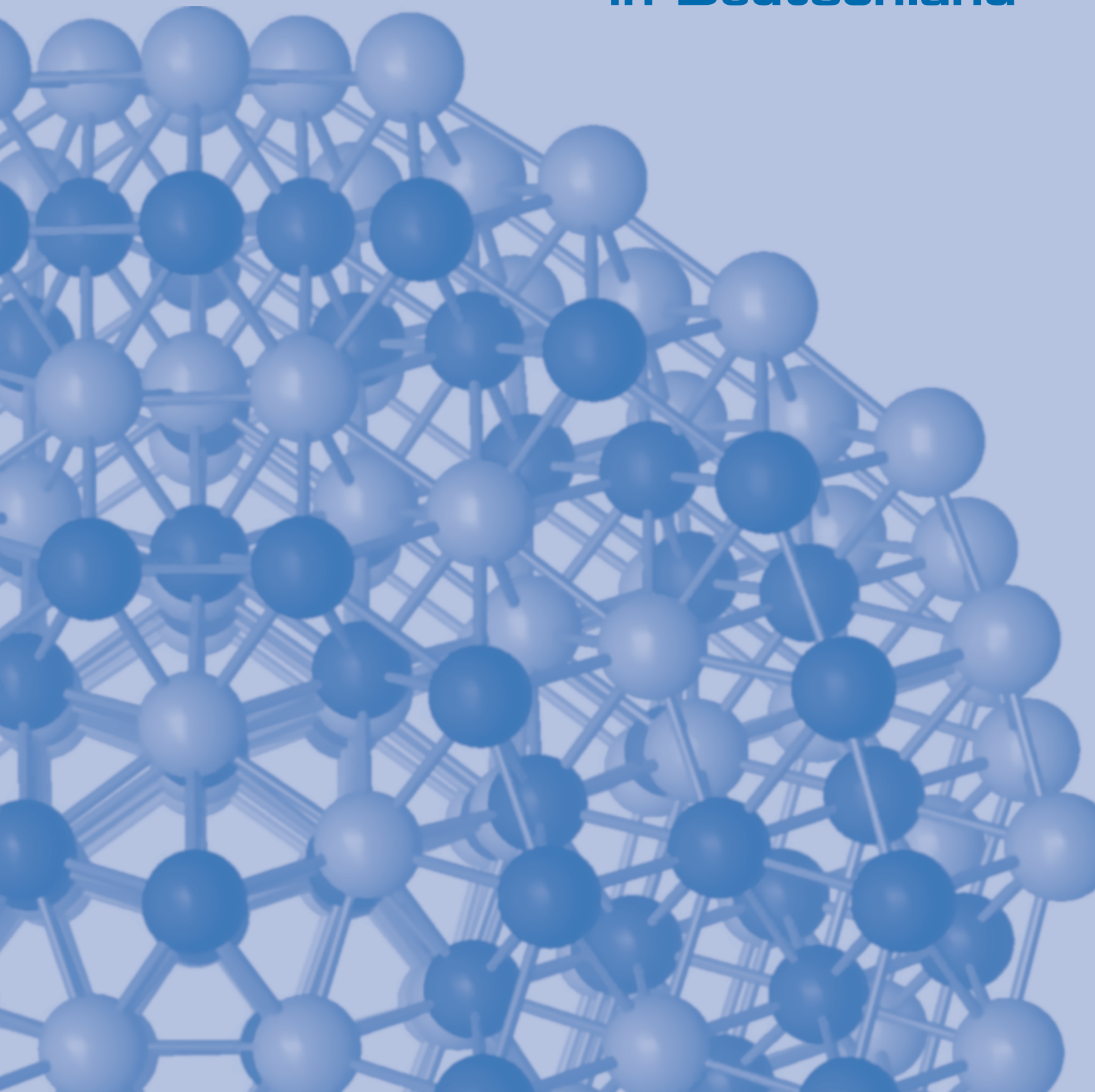
inSiDE

inSiDE • Vol. 6 No. 2 • Autumn 2008

Innovatives Supercomputing in Deutschland



Editorial

The world of Supercomputing is increasing its pace and Germany is keeping its leading position. The second step of the German national strategy has been made. During ISC' 2008 at Dresden the Gauss Centre for Supercomputing (representing the three national centres at Jülich, Munich and Stuttgart) and eleven regional and topical centres signed a memorandum of understanding for the future collaboration in the Gauss Alliance. At the same time an agreement was reached for the funding of hardware at regional centres within the new constitutional framework of Germany.

At the same time as Germany sets up its national concept the European HPC community is moving quickly towards new strategies and forms of organization. PRACE, the partnership for Advanced Computing in Europe, has successfully initiated its work towards a European HPC infrastructure. In July the project has agreed on 6 prototype installations in Europe described in this issue.

During the discussions both at the German and the European level applications were the key issue. Hence in this issue of inside you will find a very big application section. Eight contributions from users all over Germany present the variety of simulations that are currently run on German HPC systems. At the same time these papers give a good insight into the future needs for supercomputing in Germany.

The project part is another reflection of the strong conjunction of German and European High Performance Computing. The first project described is part of the German initiative of excellence. The cluster for simulation technology (SimTech) started 9 months ago with a focus on general problems of simulation. DEISA is reporting in another article on the application side of the project. The second phase of the Bavarian regional network for technical and scientific HPC (KONWIHR II) is presented. The project section concludes with the description of a hardware project based on the Cell Broadband Engine.

As usual, this issue includes information about events in supercomputing in Germany over the last months and gives an outlook on workshops in the field. Readers are invited to participate in these workshops.

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

Contents

Gauss Alliance founded

During the ISC 2008, an agreement was signed to found the Gauss Alliance, which will unite supercomputer forces in Germany. The Gauss Centre for Supercomputing (GCS) and eleven regional and topical high-performance computer centres are participating in the alliance, thus creating a computer association that is unique worldwide. With the foundation of this German HPC alliance, the partners have given their longstanding close co-operation a legal basis for the future.

The signatories are:
Gauss Centre for Supercomputing (GCS), Center for Computing and Communication of RWTH Aachen University, Norddeutscher Verbund für Hoch- und Höchstleistungsrechnen (HLRN) consisting of Zuse Institute Berlin (ZIB) and Regionales Rechenzentrum für Niedersachsen (RRZN), Center for Information Services and High Performance Computing (ZIH) Dresden, Regionales Rechenzentrum Erlangen (RRZE), Center for Scientific Computing of the Johann Wolfgang Goethe University Frankfurt (G-CSC), Rechenzentrum Garching (RZG) of the Max Planck Society and the IPP, Deutsches Klimarechenzentrum (DKRZ), Steinbuch Centre for Computing (SCC), Deutscher Wetterdienst (DWD) and Paderborn Center for Parallel Computing (PC2).

PRACE Project: First Milestones reached

PRACE, the Partnership for Advanced Computing in Europe, selected a comprehensive set of promising architectures for which the project will deploy six prototype systems at partner sites starting in 2008. The systems will be used to evaluate the architectures with respect to application performance and scalability, development of a benchmarks suite for the future production system, and the assessment of power consumption and total cost of ownership. To meet the challenging schedule for the selection of prototypes several of the eight work packages of the PRACE project worked in parallel: Work package 6 analyzed key scientific applications currently in use at the partner sites and mapped them to suitable architectures. Work package 7 conducted a market survey and interviewed vendors of systems that were likely to reach the one Petaflop/s level in 2009/2010. PRACE issued an internal call for proposals for the installation of prototypes at the partner sites and to commit the funding required to complement the EC contribution. As a result, six prototypes were selected by the PRACE Technical Board: two MPP systems, one thin node cluster, one fat node cluster, and two advanced hybrid systems. François Robin, CEA, the leader of work package 7 stated: "Our objective was to build the best set of prototypes for preparing a timely and seamless deployment of production systems in 2009/2010 – not to attempt to select the best individual prototypes."

The PRACE Management Board approved the proposed set of prototypes and decided on their placement and funding in its session on May 29, 2008 in Amsterdam. PRACE presented the proposal to the European Commission and a panel of external reviewers on July 19, 2008. The proposal and the associated amendment to the PRACE contract was formally approved on August 29, 2008. In detail, PRACE partners will evaluate the following systems:

BSC (Barcelona Supercomputing Center, Spain) installs a hybrid prototype combining IBM Cell and Power6 processors. The Cell processors are used for computation; the Power6 processors for service.

CEA (French Atomic Energy Commission, France) and **FZJ (Forschungszentrum Jülich, Germany)** jointly use Intel Nehalem/Xeon processors in their systems. Two shared-memory multi-processors (thin node clusters) will be distributed over the two sites; a prototype produced by BULL at CEA, a larger system of the same architecture at FZJ.

CSC (The Finnish IT Center for Science, Finland) and **CSCS (Swiss National Supercomputing Centre, Switzerland)** jointly evaluate the CRAY XT5 architecture. This Massively Parallel Processing (MPP) prototype will be installed at CSC's facilities.

FZJ (Forschungszentrum Jülich, Germany) provides its already installed IBM Blue Gene/P system, as a Massively Parallel Processing prototype.

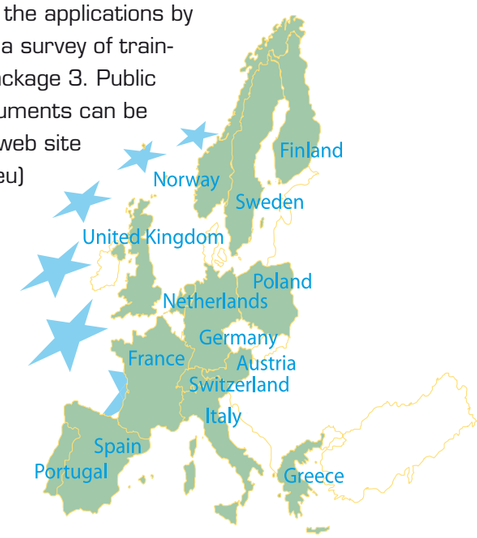
HLRS (High Performance Computing Center Stuttgart, Germany) offers a NEC SX-9 and an x86 based cluster as a hybrid prototype.

NCF (Netherlands Computing Facilities Foundation, The Netherlands) evaluates the IBM Power6 architecture, a shared-memory multi-processor (fat node cluster). This prototype which will be installed at SARA Computing and Networking Services facilities in Amsterdam.

In particular, these prototypes will be used to evaluate software for managing the distributed infrastructure, to stress various architectures in near-production situations, to prepare benchmarks for future Petascale systems, to allow a better understanding of user requirements, to scale and optimize libraries and codes, and to define technical requirements and procurement procedures for the PRACE Petaflop/s production systems expected in 2009/2010.

Although the review focused on the prototypes and the associated contractual issues only, it is worth mentioning that the external reviewers explicitly praised two documents produced by the consortium: The analysis of the applications by work package 6 and a survey of training needs by work package 3. Public versions of both documents can be found at the PRACE web site (www.prace-project.eu)

The PRACE project receives funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° RI-211528.



Numerical Simulations of the Ignition of Droplets in a Convective Gas Flow

Combustion is an old technology, which at present provides about 90 % of our worldwide energy support. An optimization of combustion processes in view of a high efficiency of the engine processes and therefore a minimization of the consumption of fuel is extremely important. Furthermore, during the combustion process the formation of products like carbon dioxide as well as the formation of pollutants like carbon monoxide and nitric oxide cannot be avoided. Therefore, the optimization of the combustion processes also includes the minimization of the amount of exhausted pollutants. Besides the combustion of coal, natural gas and wood, the combustion of liquid fuels plays a major role in many practical applications from industrial burners to diesel engines and aero-engines. Several innovative combustion concepts rely on spray combustion of liquid fuels. For gas turbines the lean premixed prevaporized (LPP) combustion of liquid fuels represents a promising technology to minimize pollutant formation. Another concept to reduce the formation of pollutants and to increase the efficiency in the case of internal combustion engines is given by the direct injection of liquid fuel into the cylinder of internal combustion engines (lean direct injection (LDI)). For the combustion of liquid fuels the process of the ignition of the fuel plays an important role. Examples are the ignition timing in Diesel engines or the unintentional autoignition of parts of the liquid fuel spray in mixing ducts, which can cause damages of the combustion device. Therefore, a detailed understanding of this process is necessary. Detailed investigations and analyses of combustion processes improve the understanding

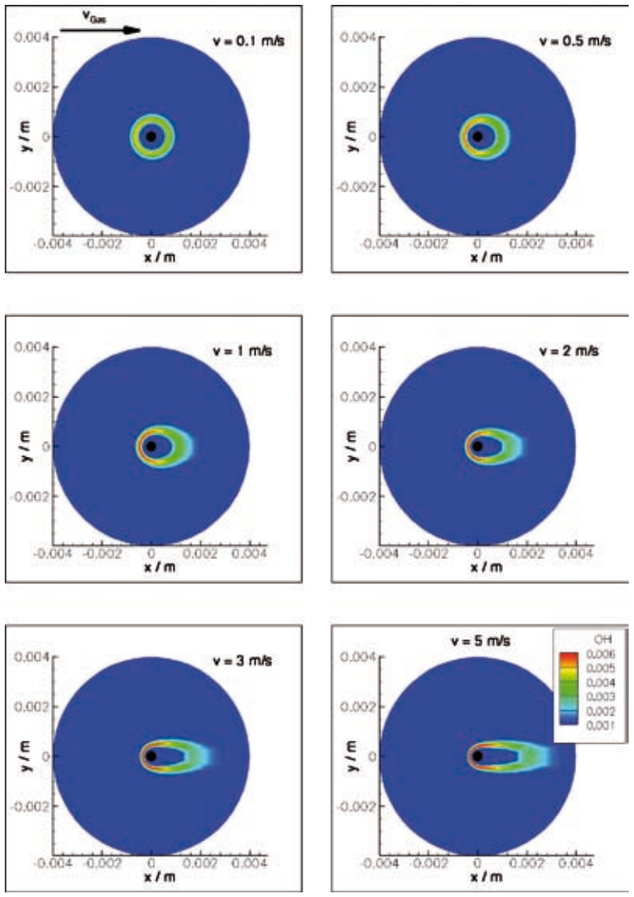


Figure 1: Change of the flame structure with increasing gas velocity: The spatial profile of the mass fraction of the OH radical during the combustion of the droplet is shown (methanol, Tg = 1400 K, p = 7 bar, rD(0) = 200 µm, t = 9 ms, gas flow from the left to the right)

of the underlying physical and chemical processes. This understanding is necessary for a reliable description and modeling of combustion processes which helps to optimize technical combustion systems. An improved understanding of the basic physical and chemical processes is achieved by investigating well-defined, representative physical systems. Therefore, the ignition of single fuel droplets surrounded by a gas atmosphere is studied, which represents the most elementary system consisting of liquid and gas phase and is the basic element of a liquid fuel spray.

Ignition of Fuel Droplets in a Convective Gas Flow

If liquid fuel droplets are surrounded by a high temperature air environment liquid fuel vaporizes and the vaporized fuel mixes with the surrounding air. If the temperature of the surrounding is high enough (depending on the investigated fuel) the mixture of fuel and air ignites without the need of an external ignition source. This auto-ignition process is investigated for single methanol and n-heptane droplets by performing transient numerical simulations starting with the exposition of cold liquid fuel droplets to a high temperature air environment. In technical systems fuel droplets are usually exposed to an external gas flow. Therefore, the influence of an external air flow on the ignition process of the fuel droplets is investigated.

Mathematical modeling is performed by solving the governing system of conservation equations for mass, momentum, energy, and species masses both for the liquid and the gas phase in a fully coupled way. The studied configuration is assumed to have a rotational symmetry. A vaporization model based on the assumption of local phase equilibrium accounts for the coupling of the liquid phase and the gas phase. One major problem in modeling combustion processes is the fact that the chemical kinetics involves a large number of chemical species, where for each species a conservation equation has to be solved. Furthermore, the kinetics introduces a large number of different time scales, causing an extreme stiffness of the resulting partial differential equation systems. The chemical kinetics of methanol and n-heptane are modeled by detailed reaction mechanisms comprising 23 chemical species and 166 elementary reactions for methanol and 56 chemical species and 570 elementary reactions for n-heptane.

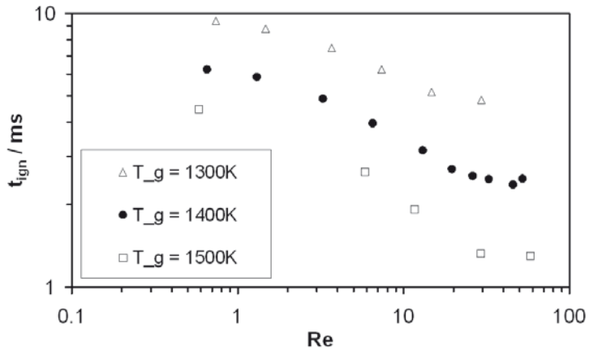


Figure 2: Dependence of the ignition delay time for a methanol droplet on the droplet Reynolds number (p = 7 bar, rD(0) = 200 µm)

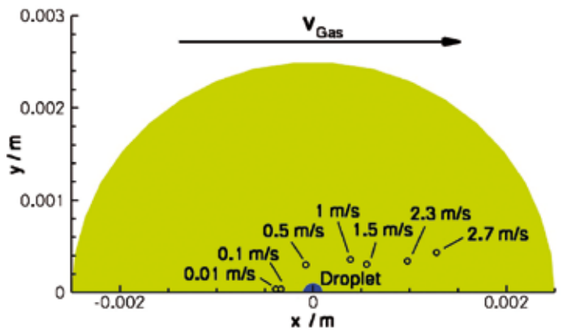


Figure 3: Location of the ignition for different gas velocities (n-heptane, Tg = 1200 K, p = 7 bar, rD(0) = 100 µm)

Analysis of the Ignition Process

Based on the model presented above the auto-ignition of single methanol and n-heptane droplets in a laminar convective environment has been simulated for varying conditions of ambient gas temperature and flow velocities. For the auto-ignition of a fuel droplet it is important to know how the ignition process is influenced by the gas flow. As can be seen in Figure 1 the flow has a significant impact on the flame shape, which is visualized by the OH-concentration (a radical formed in the reaction zone as an intermediate during the combustion process). With increasing gas velocity the flame shape differs increasingly from spherical shape.

In several technical combustion concepts like LPP combustion in gas turbines or LDI in internal combustion engines liquid fuels are injected into a mixing chamber by an injection system. In these applications the unintentional auto-ignition of the liquid fuel droplets has to be avoided. The physical parameter to quantify the propensity to auto-ignition is the ignition delay time. Therefore, the influence of the gas flow on the ignition delay time of a liquid fuel droplet is investigated. Figure 2 shows the dependence of the ignition delay time on the droplet Reynolds number. The droplet Reynolds number is defined as $Re = \rho_{air} * v_{air} * d_D / \mu_{air}$ with d_D as the droplet diameter and ρ_{air}, v_{air}, and μ_{air} as the density, the velocity, and the viscosity of the ambient air, respectively. A significant decrease of the ignition delay time with increasing flow velocity can be observed. As one can see, a variation of the flow velocity by an order of magnitude has a comparable effect on the ignition delay time to a variation of the gas temperature by +-100 K.

Beyond the above analysis of global characteristics of the ignition process numerical simulations provide a deeper insight into partial processes and local phenomena of the ignition process. This detailed local analysis is required for a detailed understanding of the basic physical and chemical processes, like vaporization, molecular transport and chemical kinetics and their interaction in the case of droplet ignition. The gas flow does not only influence the ignition delay time, but also affects the ignition location. The local point of ignition moves gradually from the upstream region in the case of lower velocities to the wake of the droplet in the case of higher flow velocities. This behavior is illustrated by Figure 3, in which the local points of ignition for different gas velocities are shown. In all investigated cases up to a flow velocity of 2.7 m/s the formation of a flame, which envelops the droplet, is observed (cf. Figure 1). If the point of ignition is located in the wake of the droplet a so called upstream flame propagation results. This means a propagation of the reaction zone from the wake to the upstream region of the droplet. The upstream flame propagation is shown in Figure 4 by the local temperature field around the droplet. The left figure shows the temperature field at the end of the ignition. The arrow indicates the direction of the subsequent propagation of the flame zone. The central and the right figure show the propagation of the reaction zone and the formation of the flame. In the right figure the development of the envelope flame (cf. Figure 1) can be seen.

These investigations provide a first insight into the complex processes governing the combustion of fuel sprays. Large numbers of such studies are needed both to cover the whole parametric range of interest (droplet sizes, ambient gas temperatures, pressures, fuel mixtures, etc.) and to build up large libraries of single droplet combustion events which can be used subsequently in hierarchical models for turbulent spray flames. Therefore, the speed-up of the numerical solution methods for the governing equations is subject of future work.

References

[1] **Stauch, R., Maas, U.**
Ignition of Droplets in a Laminar Convective Environment. In: W.E. Nagel, D. Kröner and M. Resch, editors, High Performance Computing in Science and Engineering '07, pp. 241–253. Springer, Berlin, 2008

[2] **Stauch, R., Maas, U.**
The ignition of methanol droplets in a laminar convective environment, Combustion and Flame 153:45-57, 2008

• Rainer Stauch
• Ulrich Maas

University of Karlsruhe (TH), Institute for Technical Thermodynamics

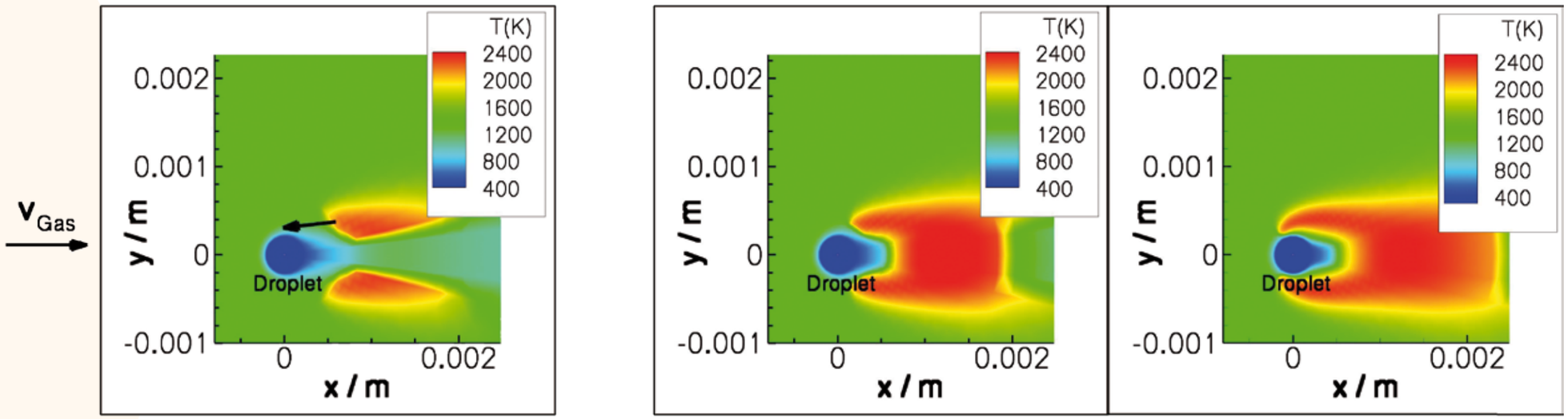


Figure 4: Upstream propagation of the reaction zone after an ignition downstream of the droplet (methanol, Tg = 1400 K, p = 7 bar, rD(0) = 200 µm, v = 3 m/s, left: t1 = 2.80 ms, central: t2 = 2.92 ms, right: t3 = 3.04 ms)

Applications

Coupled Simulations in Heterogeneous High Performance Environments

With increasing computational power, coupled simulations incorporating more than one single physical model or more than one single scale become more and more common. By taking different physical effects into account, simulation models get much closer to reality, and the accuracy of the model simulation can be greatly improved. For examples in multi-physics simulations you might think of climate prediction, which is based not only on atmospheric modelling but also on simulated models of the oceanographic behaviour and icing simulation. Another widely known example is Fluid-Structure Interaction (FSI), used in engineering applications like plane development as well as in medical applications like blood flow simulation. An example for multi-scale simulation is the direct simulation of aero acoustics, where very small structures in the flow, generating the noise, have to be resolved and at the same time the propagation of the sound waves with much longer wave-length over large distances. Many more multi-physics / multi-scale simulations are deployed and increase the insight to phenomena observed in the real world. Because of the growing computational power, these more complex simulations become ever more feasible and thus gain increasing share in computational science.

Generally, the typical setting for multi-physics and multi-scale applications is a simulation code consisting of at least two modules for different physical models plus a coupling code. The coupling step consists of compute parts as well as steering parts. It takes into account all data handling between the modules, e.g. interpolation in space, data transfer, and sub-cycling in time. Each of the interacting modules in these coupled simulations have their very own numerical kernel, where most of the computational time is being spent. Often, these kernels have formerly been stand-alone codes. Those kernels usually have very different properties and requirements, thus resulting in a heterogeneous overall computational model. On the other side of the computation, the hardware which is used to compute the simulations is getting differentiated as well. Higher performance can not be achieved anymore by increasing clock frequency of processors, as it used to be for the last decades. Instead, different approaches have to be found to further increase the performance and take advantage of modern down sized circuits. Upon approaching 3 GHz of clock frequency we saw the rise of multi-cores in commodity processors and no further increasing of the single core performance. Furthermore, there is a wide variety of specialized processors emerging to still increase the hardware performance. For example the first installation providing 1 PetaFLOP peak performance in the Top500 list, the IBM Roadrunner at the Los Alamos National Laboratory combines AMD Opteron dual core processors with the supplicant power of Cell Broadband Engine processors. Another observed trend is the usage of graphic processing units in high performance computing or even more specialized hardware designs like the MD-Grape technology especially targeted at molecular dynamic simulations. As a matter of fact, with the limit of easily increased computational power by increasing the clock frequency of scalar processors, we are now facing a heterogeneous computing environment with many different architectures providing each its own set of features and capabilities.

The challenge in this diverse environment of modern scientific computing is to map the heterogeneity of the applications onto the heterogeneity of the computing hardware. We are no longer faced only with parallelisation and distribution of workload but also with different qualities attached to both, the computations and the computers. To take full advantage of the power provided by the computing systems, we have to use their features properly. For simulations involving different numerical kernels this results in the need to map the requirements of those kernels onto the available architectural features. Only by adapting the implementations to the new processor designs it will be possible to gain high sustained performance for each of the involved model kernels and to reach a shorter time to solution for the overall simulation. At HLRS we have very different architectures at hand in a heterogeneous network of different clusters. Especially with our NEC installations, a Linux cluster with Intel 64 Xeon CPUs (cacau.hww.de) and the NEC-SX8 installation with vector processors with its frontend, a shared memory NEC-TX7 i9510 cluster (IA64 cluster). The coupling of the High-End system to its frontend is tight, the coupling to the external Linux cluster is loose, as it is typical for current state-of-the-art installations. This setup is used for hybrid simulations using more than one resource at a time. One example of applications that is being run using more than a single resource in our network is fluid structure interaction, for example in

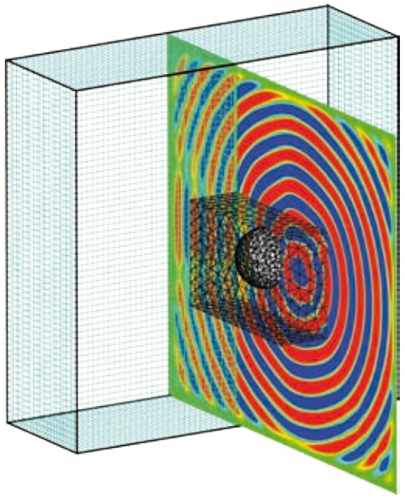


Figure 1: Aero-acoustics simulation of noise generation in the proximity of a sphere. Unstructured domain around the obstacle, structured domain for the far field by courtesy of Institut für Aero- und Gasdynamik

the calculation of hydropower turbines. In this setup the flow is computed using the Fenfloss program, developed by the Institute of Fluid Mechanics and Hydraulic Machinery (IHS), and the structural mechanics part is simulated by the commercial FEM solver Abaqus. The coupling is provided by a third application, MpCCI which is developed and licensed by the Fraunhofer Institute and allows general grid mapping and coupling of different programs. Adapters for common commercial applications are available, allowing you to choose the most appropiate one. If a commercial tool like in this case Abaqus is used, it is typically not available on the high end system, as the vendors usually do not support these platforms. But even if an otherwise available third party tool is used, it normally is tuned for Linux clusters only. On the other hand, the IHS in-house code Fenfloss which is run to simulate the fluid flow, delivers 30-50% of peak performance on the NEC-SX8 vector platform. Since the fluid part of the simulation usually requires an order of magnitude larger computational power than the structural mechanics part, it is the dominating part of the simulation and should not be forced to run on a standard Linux cluster by the much less compute-intensive structural solver. Because of the needed computational

power, it is essential to run Fenfloss in the most effecient available way. Forcing the fluid part to be run on a Linux cluster, just because the structural mechanics part is not available anywhere else, means to reduce the efficiency to the least common denominator. Using MpCCI to couple modules running on different machines, is a method to increase the efficiency of fluid structure simulations.

Direct aeroacoustic simulation of noise generated by flows is another application, where the different architectures can be utilized. The KOP3D program, developed by the IAG, was designed to tackle this multi scale problems by allowing the simulation to be split in different domains. The flow around an obstacle with complex geometry is discretized using a fine unstructured grid, allowing for resolving small flow phenomena and approximate arbitrary geometries. Farther away structured grids can be used allowing additionally for much larger time steps. To simulate the soundwave propagation in the far field it is sufficient to solve the linearized Euler equations with a lower resolution. The dominating part for the coupled simulation is the structured far field since it requires to solve for a very large computational domain. The bottleneck in the simulation on the High-End system on the other hand is the solver for unstructured domains since it does not vectorize well. The behaviour of the domains is very different on the available architectures. The structured grid on the other hand can be solved in the most expensive routine with almost 14 GFLOPS per processor. As huge domains for the far field have to be solved, the usage of the NEC-SX8 is very desirable for these domains. On the other hand, the unstructured part is very inefficient on this architecture (less than 100 MFLOPS per processor), but much better suited for the Linux cluster environment. Thus, there is no good in forcing either the one or the other part of the simulation to a non-fitting machine. The aim is to use the high-end system NEC SX-8 as well as the Linux cluster, each for the module that fits optimal. In this setup the coupling itself is done by the application and not by a third party tool. To run the different domains in one simulation on distributed machines, the

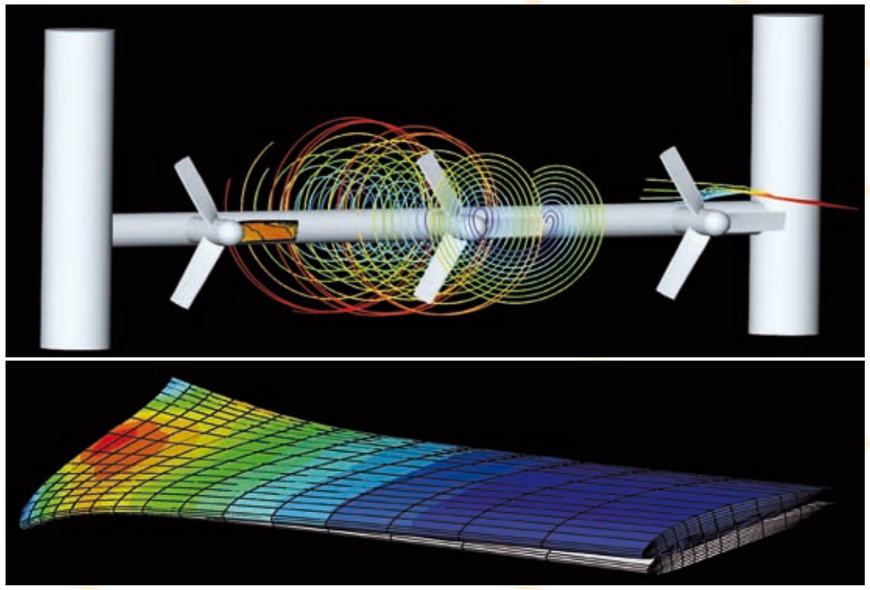


Figure 2: Interaction of fluid and structure in a tidal power turbine. On the left the pressure distribution by the flow on the blade of the rotor is shown, in the middle the streamlines in the relative (rotating) co-ordinates are visualized and on the right the streamlines are traced in absolute coordinates.

Figure 3: Here the van Mises stresses are shown as colors, alongside the deformation caused by them in one of the rotor blades. Both figures kindly provided by Institut für Stömungsmechanik und Hydraulische Strömungsmaschinen.

	1xIA64	1xSX	1xSX + 1xIA64
UNSTRUCT:	2,993.67	**7,745.82**	3,019.24
STRUCT:	**23,887.32**	2,870.66	2,869.46
KOP:	1,012.37	321.15	554.21
waiting:	0.00	0.00	164.20
KOP calculating time:	1,012.37	321.15	390.02
Total CPU time:	27,893.35	10,937.62	6,278.72
Total elapsed (sec):	27,924.78	10,966.23	3,207.09
Total elapsed (h):	7:45'	3:03'	0:53'
Relative Price	1	1,57	0,58

Table 1: Efficiency of the coupled simulation in an heterogeneous hardware environment w.r.t time and cost

PACX-MPI library is used. This library transparently enables the use of different clusters in a MPI parallel application. It is transparent since the application calls MPI routines in the usual way. The application does not care of the distinct clusters or networks. The function of the PACX-MPI is to catch up these MPI calls, check whether the communication remains within the same machine – in this case it hands the call over to the native MPI installation here – or is a cross-cluster communication – in which case it opens a communication channel to the other side. The advantage from users perspective is its transparency, from efficiency point of view, it is the possibility to use the native (i.e. highly optimized) MPI libraries on each cluster, and from networking point of view the possibility to handle different network types and protocols on the one side, the other side, and the cluster-interconnect. On the other hand, there is an overhead due to the additional processes needed for the inter-cluster communication. Nevertheless, the effort pays off. Table 1 shows the elapsed time for the unstructured, the structured and the coupling parting as well as the waiting times. Additionally, the relative price in Euro for each simulation is given. The columns compare the simulation using the IA64 cluster only (column 1xIA64), the vector computer NEC SX-8 only (column 1xSX), and the coupled simulation using both sides for the according modules (column 1x SX + 1x IA64). The first simulation, using the IA64 cluster only, needs the longest turn-around time of nearly 8 hours for the entire simulation. Using the high-end NEC-SX8 vector system, the turn-around time is substantially reduced to around 3 hours. But the relative price is about 60% higher. Using the vector system for the structured part, the IA64 cluster for the unstructured, the turn-around time for the entire simulation reduces to less than one hour, and even the relative price reduces by 40%. Thus, both measures that are relative from users point of view, i.e. turn-around time as well as total cost – argue for the coupled simulation.

The examples above show that with the upcoming of multi-physics and multi-scale simulations the simulation codes are no longer monolithic and homogeneous. They tend to consist of modules, plugged together according to the needs of the current application. We observe a non-homogeneous behaviour of those modules w.r.t. hardware architecture. At the same time, the architectures become heterogeneous as well. The art for the next future is to develop tools and methods to map the modules and their non-homogeneous requirements to the according parts in the hardware environment. To ease the usage and to increase the efficiency, future systems have to be much tighter coupled w.r.t. network, management, scheduling, but also tools like debuggers and performance analysers. This will be the important topic for HLRS in the next future.

• Sabine Roller
• Harald Klimach

University of Stuttgart, HLRS

Ab Initio Simulations of the Initial Oxidation of CoCr Alloy Surfaces

The oxidation of a cobalt chromium (CoCr) alloy in contact with air is characterized by the spontaneous formation of a thin oxide film which protects the surface. The oxide layer governs the interactions of the alloy with the external environment and is responsible, in particular, for the bioactivity of CoCr implant surfaces in the human body. Understanding and predicting these interactions requires a precise knowledge of the structure and composition of the oxide layer.

The aim of our work is to investigate the mechanisms of initial formation of the native oxide layer formed on a CoCr alloy by first-principles molecular dynamics (FPMD). Examples of applications of this technique to the study of the initial oxidation reactions of metal and semiconductor surfaces can be found in [1,2] and references therein.

The first problems to face in the case of CoCr concern surface composition, segregation behavior in vacuum and in the presence of oxygen, and selective oxidation of the components. The interaction of a CoCr surface with oxygen molecules occurs via a chemisorption process, characterized by a highly exothermic dissociation of the O-O bond and stable adsorption of the separate oxygen atoms on the surface. Formation of an oxide layer takes place if adsorbed oxygen atoms can exchange place with underlying metal atoms and be incorporated below the metal surface. Driving forces for this metal/oxygen place-exchange process are in principle available due to generally high heat of adsorption of oxygen on metals. However, the mechanistic details of the actual exchange process and especially the atomistic structure of the oxide forming on CoCr are widely unkown.

Computational details

Our simulations are performed with a first-principles molecular dynamics (FPMD) approach based on the Density-Functional Theory (DFT) [3]. The technique allows us to investigate the dynamical evolution of a system composed by atom nuclei and electrons at the quantum level of precision. However, the DFT formalisms, when applied to systems containing about 100 atoms in the unit cell, are computationally very demanding and necessarily require the use parallel computer architectures and fast interconnections between the CPUs.

Indeed, calculation of the electronic structure and a full geometry optimization of such a system necessitates about 700 CPU-hours on a SGI-Altix platform to be completed, while one picosecond of simulated first-principles molecular dynamics requires about 5,000 CPU-hours.

The highly optimized parallel code used for completing the present work presents very good scaling behavior using medium-size CPU partitions (up to 64 computing nodes).

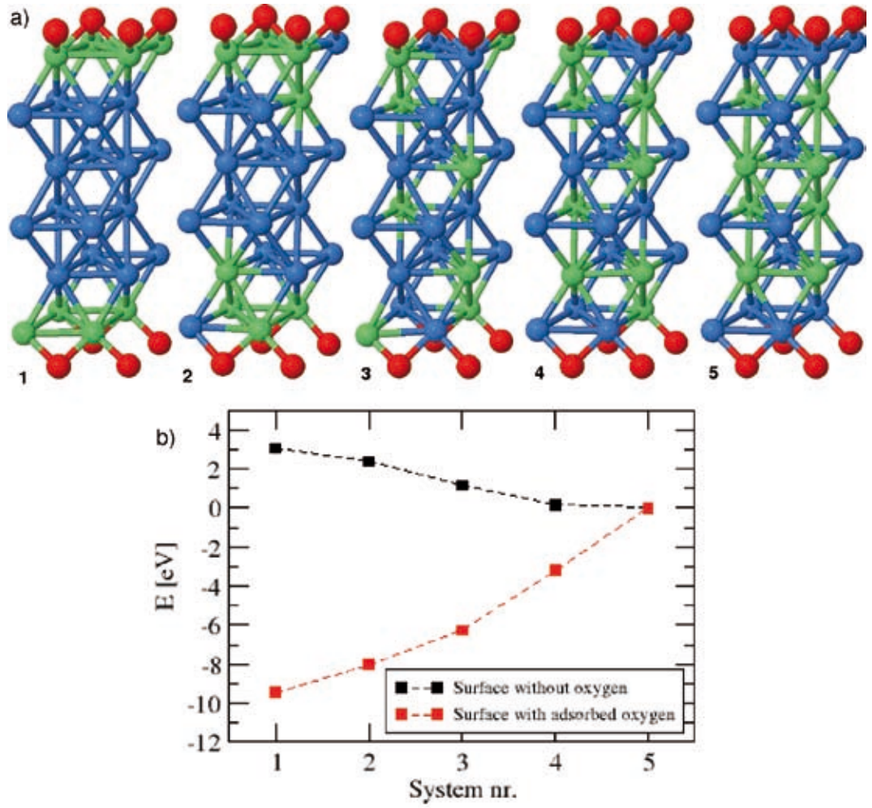


Figure 1: (a) Different CoCr alloy models with oxygen adsorbed on hollow sites (Co in blue, Cr in green, O in red) for a 2x2 (0001) surface unit cell; (b) Ground state energy differences for the five models calculated in absence and in presence of an oxygen ad-layer at a coverage of 1ML. The energy values are relative to system 5.

CoCr alloy composition

Mixing of two metal elements to form an alloy permits to tune to a certain degree the thermal, electrical and mechanical properties of the resulting material. In particular, the surface reactivity of an alloy, often mediated by an oxide layer forming in contact with the atmosphere, may differ substantially from the reactivities of its separate components. This is due to the peculiar surface electronic properties resulting from selective segregation processes of the alloying elements near to the surface. These issues are addressed here by a set of preliminary calculations of the surface properties of different models of a CoCr alloy at the typical technological Cr concentration of 30 at.% (see also Ref. [4,5]). The surface has been modeled by a (2x2) surface cell with a metallic slab of six atomic layers separated by a vacuum gap of the same thickness. This cell is repeated periodically in each direction. A series of five systems of equal crystal structure (HCP) and composition (Co:Cr=2:1), but with different arrangements of Co and Cr atoms on the lattice sites have been considered, as shown in Fig. 1a.

The first system corresponds to a full segregation of Cr to the surface layer. Going from the first to the fifth system, the Cr atoms mix with Co in the inner part of the slab and Co atoms move to the surface, up to pure Co surface segregation. In the case of bare metal surfaces, the total energy of the systems decreases from system 1 to system 5 by about 3.1 eV (Fig. 1b).

This is due to a combination of the smaller surface energy of Co(0001) with respect to hcp-Cr(0001) and of the favorable enthalpy of mixing of the alloy with respect to the separate components. The situation changes completely in the presence of 1 monolayer (ML) of oxygen adsorbed on hollow sites on the alloy surface. Namely, due to the higher heat of adsorption on chromium, full segregation of Cr atoms to the surface layers becomes energetically favorable with respect to the other systems, with a total energy gain of about 5 eV going from system 5 to system 1 (Fig. 1b). This indicates the presence of a strong thermodynamical driving force for an oxidation-driven segregation of Cr atoms towards the surface of the alloy. However, given that a surface segregation of Co atoms seems to be preferred in the case of bare surfaces, how the actual diffusion of Cr atoms will take place during the oxidation of an initially bare alloy surface remains to be investigated.

This issue requires larger model systems and an analysis based on dynamical simulations of the oxidation reactions. We choose a model system of a $Co_{0.67}Cr_{0.33}$ alloy consisting of 6 atomic layers with 12 atoms in each layer. This model contains 3 Cr atoms in the surface layer and 4 Cr atoms in the subsurface layer. On this system we studied the adsorption of oxygen and the formation of a thin oxide film, up to a maximum of 1ML coverage.

Molecular dynamics simulations

Mechanism of O_2 dissociation

Oxidation at room temperature takes place via a dissociative adsorption of oxygen molecules on the metal surface. In a number of first principles molecular dynamics (FPMD) simulations, this so-called "hot-atom" dissociation mechanism has been observed to lead to spontaneous O_2 dissociation on Co, Cr, Al, Ti, TiN, and on Si. A similar behavior is observed in the CoCr model systems considered in the present work. As starting condition we choose to place the oxygen molecule with zero velocity at a distance of 0.3 nm over the fully relaxed surface. We tried this with different orientations for the O_2 molecule as starting position, but found no substantial differences in the results of the simulations. Figure 2 shows snapshots from a FPMD simulation of the adsorption of O_2 on CoCr(0001).

Initially, the molecule is quickly adsorbed on the surface, where it dissociates spontaneously and the two O atoms arrange on the most favorable adsorption sites. The dissociation is activated by electron donation from the surface atoms into the unoccupied states of the antibonding orbitals (π^* and σ^*) of O_2. Occupation of these orbitals leads to an increase of the O-O bond and to final splitting of the bond. A detailed analysis (Fig. 2b-c) reveals that the integrated spin density of O_2 at the beginning of the dynamics is not exactly 2.0 (due to the 2 unpaired electrons in O_2) but about 1.75 electrons, indicating a partial charge donation from the surface into the π^* orbital of the molecule (see also Ref. [1]). During adsorption, further electron donation into the π^* orbital leads eventually to a full quenching of the magnetic moment after about 0.3 ps. The O-O distance gradually increases until at 0.16 nm also the σ^* antibonding orbital of the adsorbed molecule becomes partially populated, leading to evident oscillations in the spin density. The O-O bond dissociates as a consequence of this event and the O-O distance increases abruptly in a way which is indicative of a non-thermal, "hot atom" mechanism. The results of this first simulation are starting point for the further adsorption of O_2 molecules which are placed over the fully relaxed surfaces obtained after each adsorption event.

Dynamics of the oxidation of CoCr(0001)

We performed a total of five consecutive FPMD simulations of the adsorption of oxygen on a CoCr surface, each time starting with one or two new O_2 molecule placed near the relaxed surface obtained in the previous simulation. After the dissociative chemisorption of the first O_2 molecule (Fig. 2a), a second molecule dissociates in the same way, and, at the end of the simulation, four O atoms are stably adsorbed in hollow surface sites, as shown in Fig. 3a. The final snapshots of the subsequent adsorption of further O_2 molecules are illustrated in Fig. 3b-d.

Surprisingly, up to the adsorption of four O_2 molecules there are no substantial changes in the structure of the CoCr slab. No place exchange takes place between the surface- and oxygen atoms, and a flat oxygen-layer (0.67 ML) forms on a nearly perfect surface, similarly to what observed on a pure bcc-Cr(110) surface in a separate set of simulations. Only after the adsorption of a fifth molecule we observe a Cr atom escaping from the surface layer: the molecule binds without dissociating

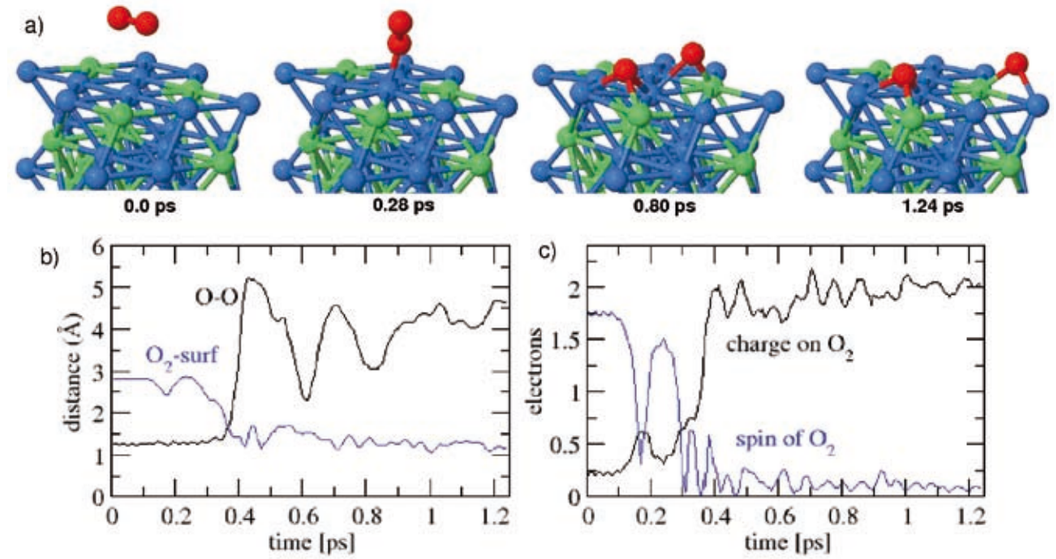


Figure 2: (a) Snapshots of a molecular dynamics simulation of the adsorption of O_2 on CoCr(0001). The model system represented here counts 3 Cr atoms in a 12-atom surface unit cell. Only the upper 3 layers of the slab are shown; (b-c) Analysis of the O_2 adsorption on CoCr: (b) Evolution of the O-O distance and the height of the center of mass; (c) Spin density and atomic charge of the O_2 molecule. Color code as in Fig. 1

on-top of this Cr atom, which was pulled out from the surface by 0.22 nm, forming a structure that closely resembles the tetrahedral Chromate-ion $(Cr_2O_4)^{2-}$. The energy released during the adsorption and dissociation of a sixth O_2 molecule allows a Co atom to escape from the surface layer by about 0.21 nm, and an oxygen atom binds on-top on it, in a tetrahedral structure (Fig. 3d). The so-obtained structure presents no further adsorption sites for incoming oxygen molecules, which are rejected by surface, while the O_2 adsorbed in molecular form still does not dissociate. This can be attributed to a reactivity loss of the partially oxidised surface with respect to the bare surface. Indeed, charge transfer into the orbitals of an incoming O_2 molecule decreases with increasing coverage due to electron depletion in the d-orbitals of the surface atoms. Moreover, also the repulsion induced by the negatively charged adsorbed oxygen atoms inhibits further reactions with incoming oxygen molecules.

To activate these reactions we increased the temperature of the system up to 900 K through simulated annealing. The heat causes the atoms to move off their initial positions and find a configuration with lower energy than the initial one after quenching the system to zero temperature. The structures of CoCr with 1 ML of oxygen before and after annealing are shown in Fig. 3d and 3e. The annealing process induces large transformations of the structure: not only does the O_2 molecule on top of the Cr atom dissociate, but further atoms escape from the surface layer causing formation of a thin oxide structure. In particular the Cr atoms of the uppermost layer seem to escape the surface preferentially and to become coordinated by incoming oxygen. This is in accordance with experiemental data, which predict that oxidation of CoCr takes place with formation of Chromium oxide Cr_2O_3 with only a minor contribution of Co-oxides [6].

In conclusion, our FPMD simulations reveal that formation of a thin oxide layer through place exchange between oxygen and metal atoms starts only at high oxygen coverages (0.83 ML) and is thermally activated. This seems to be a characteristic of the CoCr system, which was not observed in the case of e.g. Al or Si, where nucleation of the oxide layer initiates spontaneously at 0.5 ML. This occurrence requires further FPMD simulations on CoCr in order to achive higher oxygen coverages.

Furthermore, the exact mechanism of selective oxidation has to be understood in a more detailed way. Future work will be devoted to a precise analysis of the process of bond-breaking of the Cr and Co atoms in order to become coordinated to oxygen and to escape from the surface.

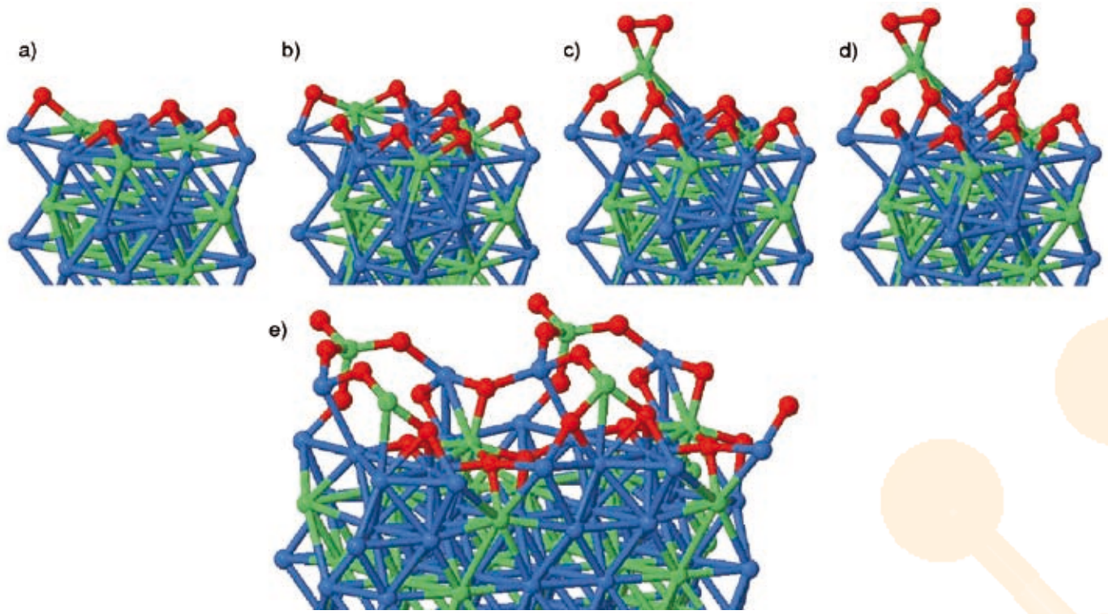


Figure 3: Final snapshots of five consecutive FPMD simulations of the initial oxide formation on a CoCr(0001) surface. The oxygen coverage increases from 0.33 ML (a) to 1 ML (d-e). The structure in (e) has been obtained after annealing of the structure in (d). (Color code as in Fig. 1)

References

[1] **Colombi Ciacchi, L., Payne, M. C.**
Hot-atom O_2 dissociation and oxide nucleation on Al(111), Physical Review Letters 92, 176104, 2004

[2] **Colombi Ciacchi, L., Payne, M. C.**
First-principles molecular-dynamics study of native oxide growth on Si(001), Physical Review Letters 95, 196101, 2005

[3] **Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A., Joannopoulos, J. D.**
Iterative minimization techniques for ab-initio total-energy calculations: molecular dynamics and conjugate gradients, Reviews of Modern Physics 64, 1045, 1992

[4] **Zimmermann, J., Colombi Ciacchi L.**
DFT modelling of oxygen adsorption on CoCr surfaces, in: High-Performance Computing in Science and Engineering, W. Nagel et al., Eds., Springer Verlag, Berlin, pp. 173-186, 2008

[5] **Dianat, A., Zimmermann, J., Seriani, N., Bobeth, M., Pompe, W., Colombi Ciacchi, L.**
Ab initio study of element segregation and oxygen adsorption on PtPd and CoCr binary alloy surfaces, Surface Science 602, 876, 2008

[6] **Milosev, I., Strehblow, H.-H.**
The composition of the surface passive film formed on CoCrMo alloy in simulated physiological solution, Electrochimica Acta 48, 2767, 2003

• Janina Zimmermann
• Lucio Colombi Ciacchi

University of Karlsruhe, izbs

Fraunhofer IWM Freiburg

Autumn 2008 • Vol. 6 No. 2 • inSiDE 19

Interactive Simulations on the Grid

It is every scientists dream: To have interactive access to high performance computing from the comfort of one's own office. Interactive simulations on the Grid promise to deliver exactly that: Complex computations are performed on the Grid, a dedicated remote visualization system creates powerful graphics in real time and the final rendered images are displayed on the client's run of the mill desktop PC or even Laptop. The scientist can directly see the output of his most complex simulation and even interact with it in real time. User input is directed back to the Grid and the simulation is steered in the direction that yields the most insight to the scientific question at hand.

Computing on the Grid is an easy and fast way to perform complex numerical tasks. The user simply submits a job, Grid middleware runs the code as a batch job on a suitable machine, and delivers the final result back to the user. To make sense of the vast simulation data, powerful visualization hard- and software is needed (lots of RAM, powerful graphics cards and expensive software licenses). All of this is provided through a dedicated remote visualization server. In contrast to a conventional internet connection*, the processing and rendering of the data sets are performed on the remote visualization server, generating fully rendered images which are then compressed into jpeg format and transported to the client, sending a continuous stream of screenshots over the internet. It's like watching a personalized movie that is rendered as you watch it, with a plot that evolves exactly the way you want it to.

* It has been possible for many years to redirect the graphical output of a program running on one machine to the display connected to another computer (this "network transparency" is one of the basic features of the X window system used on all Unix-like operating systems). However, rendering in this case is performed on the client's machine. Sending complete data sets generates tremendous network traffic and requires high-end graphics workstations on the client side.

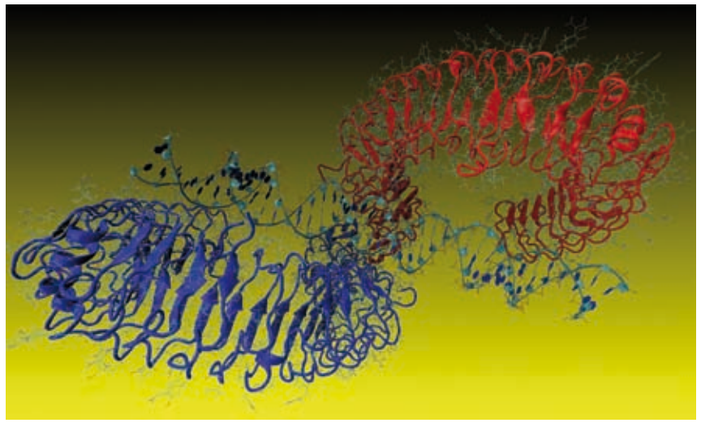


Figure 1: Two toll-like receptor molecules (coloured in red and blue, respectively) bound to viral double stranded RNA (see text)

Real-world Example

To demonstrate the power of interactive simulations on the Grid, in the following we present a real-world example from biological physics, describing a computer experiment that tries to elucidate the inner workings of the human immune system, especially how it detects the presence of a virus. Here is a brief summary of the scientific background: One of the defining features of certain viruses is that their genetic code is stored in double stranded RNA (dsRNA). Living cells, however, store their genetic code in double stranded DNA (dsDNA, the DNA double helix) and only produce single stranded RNA in the process of gene transcription. Therefore, historically one of the oldest detection mechanisms to reveal the presence of a virus is to look for the presence of dsRNA. This is done by specialized molecular machines, the so called toll-like receptors (see Figure 1).

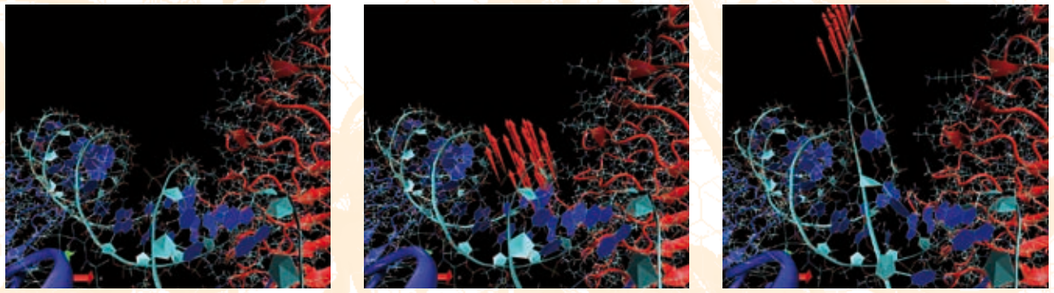


Figure 2: VMD screenshots of viral dsRNA bound to the toll-like receptor. In the center, the dsRNA can be seen (the light and dark blue disks represent the nucleobases, the backbone is shown as a light blue tube). On the left and right, parts of the two toll-like receptor molecules can be seen, coloured in blue and red, respectively. Left: The starting configuration of viral dsRNA and receptor (solvated in water – the water molecules are not drawn). Middle: In an AFM-like computer experiment, a force has been applied to a nucleobase in the centre of the picture (red arrows) Right: The molecular dynamics simulation on the Grid calculates the correct response of the total system (receptor molecules plus viral dsRNA and surrounding water – ca. 300,000 atoms), one strand of the dsRNA is pulled away.

For a computer experiment, we need to know the exact molecular structure of the biomolecules involved, which, for large biomolecules that can consist of several hundred thousand or even millions of atoms, is a very challenging task in itself. A very recent publication presented the molecular structure of a toll-like receptor bound to a piece of dsRNA [1]. This data set is especially interesting since it captured the instant when the immune system detected the presence of viral dsRNA. It is still unknown how exactly this detection mechanism works, i.e. where exactly the receptor molecule binds to the dsRNA. An answer to this scientific question may come from interactive molecular dynamics that allow the user to perform the analogy of single-molecule AFM* experiments in real time in the computer. The complex biomolecule (receptor plus viral dsRNA in water) is simulated on the Grid and visualized on the remote visualization system. Simply by clicking and dragging on a specific nucleobase of the viral dsRNA, the user can apply a force to a selected part of the viral dsRNA (see Figures 2 and 3). The molecular dynamics simulation (NAMD) on the Grid now calculates the response of the combined system to this force. The user can watch if and how the applied force is able to tear the viral dsRNA away from the receptor. The position where the viral dsRNA last loses contact to the receptor molecule will be a good candidate for the position where the recognition site of the viral dsRNA is located in the receptor molecule. Traditionally, this kind of computer experiment has already been possible by writing a numerical force vector into the input file of the molecular dynamics simulation and running the simulation as a batch job. Using the described approach of interactive molecular dynamics, however, allows the scientist to decide on-the-fly whether the currently performed computer experiment is useful or even

* Atomic force microscopy (AFM) utilizes an extremely sharp needle and allows experiments on single biomolecules. The biomolecule is firmly attached at one end with the AFM tip gently pulling at the other end. By measuring force vs. distance it is possible to determine when specific bonds are opened, and when the biomolecule unfolds and eventually breaks.

valid. For example, was the force applied at the right position? If not, either nothing happens, or the biomolecule could be torn apart, rendering the simulation useless in both cases and wasting valuable CPU time in the case of non-interactive batch runs.

The Components

- **Uniform Interface to Computing Resources (UNICORE):** a Grid computing technology that provides seamless, secure, and intuitive access to distributed Grid resources such as supercomputers or cluster systems. The UNICORE client has a powerful graphical user interface that covers the most common usage scenarios, such as application execution and multi-step, multi-site workflows [2].

- **Grid Security Infrastructure OpenSSH (GSI-OpenSSH):** a part of the Globus Toolkit that provides a modified version of OpenSSH [3]. GSI-OpenSSH adds support for X.509 proxy certificate authentication and delegation, providing a single sign-on remote login and file transfer service.

- **Nanoscale Molecular Dynamics (NAMD):** a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems. NAMD scales to thousands of processors [4,5].

- **Visual Molecular Dynamics (VMD):** a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics. VMD has a plug-in for interactive molecular dynamics that provides seamless integration of VMD with NAMD and allows computational steering of the molecular dynamics simulation [4,5].

- **VirtualGL:** an open source package that redirects openGL commands to the server's graphics cards, compresses final rendered images into jpeg format and sends them over the internet to the virtualGL client (vglclient) that listens on port 4242 (or port 4243 in case of SSL encryption) [6].

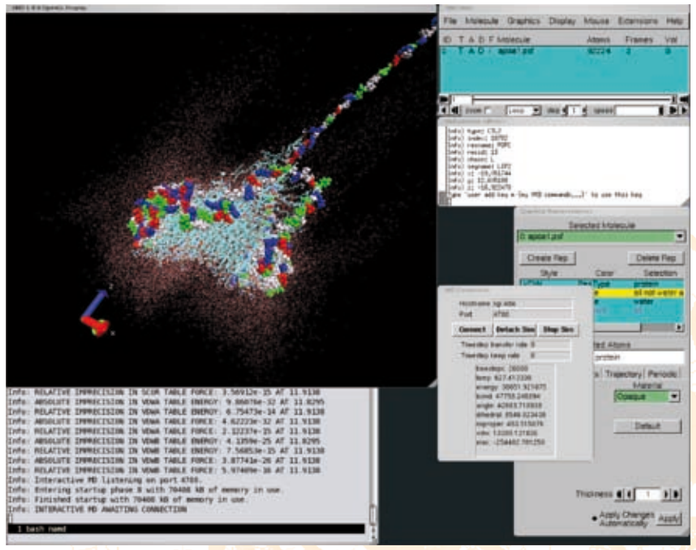


Figure 3: Interactive simulations on the Grid at work (screenshot). The molecular dynamics simulation NAMD runs in the background. In the foreground, VMD is used to visualize the current configuration of the molecule in real time. The user can interact with the simulation with the mouse (computational steering).

How it's done

The four main elements of the simulation infrastructure are depicted in Figure 4: submission host, visualization host, gateway, and the execution host on the Grid. The scientist sends a NAMD job script to the Grid using the graphical user interface of the UNICORE client. UNICORE automatically generates an appropriate job description file that is submitted to a batch scheduler system on the execution host where the simulation is run (in our case 510 cores of the sgi Altix 4700 at the LRZ). Output generated by NAMD is sent to the visualization host (here: the remote visualization server at the LRZ). Various security policies, such as firewalls, typically prevent a direct communication channel between the execution host and the visualization host. Therefore, an SSH tunnel (port forwarding) has to be established, for which GSI-OpenSSH offers a great advantage since authentication on all machines is done using the X.509 certificate of the user. The tunnel can pass through several intermediary hosts and connects a port on the gateway machine with a port on the execution host allowing direct communication. The output of the molecular dynamics simulation is sent to the remote visualization server, where VMD reads the new co-ordinates after each time step and displays the current configuration of the molecule. Fully rendered images are compressed and sent to the client using virtualGL. Mouse events generated by the user are sent back to the supercomputer on the Grid, which reacts accordingly (computational steering).

In our example, for every time step of the MD simulation (300,000 atoms, 510 cores, calculating ~20 time steps per second), circa 0.4 GByte/s of data are generated and have to be transferred between the Grid and the remote visualization system. The client machine, on the other hand, only receives a stream of jpeg images with roughly 1 MByte/s.

Conclusion

Interactive simulations on the Grid allow scientists from anywhere in the world to perform the most complex numerical simulations on the Grid, to directly see a live visualization of their data sets through remote visualization and to have the extra advantage of a means to directly interact with the simulation (computational steering). The scientist does not have to worry about buying, installing and/or upgrading expensive hardware or software, but can focus on the task at hand: finding an answer to the scientific question.

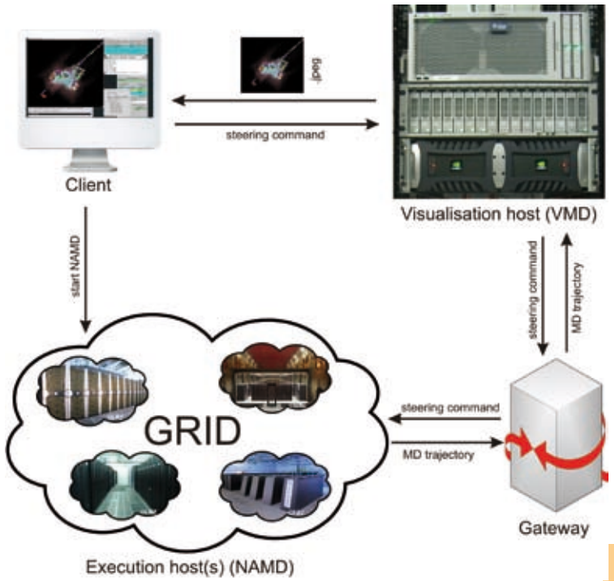


Figure 4: The four main elements of the simulation infrastructure: submission host, visualization host, gateway and the execution host on the Grid. The molecular dynamics simulation NAMD is run on the Grid execution host and sends its output (MD trajectory) to the visualization host, where VMD reads the new co-ordinates after each time step and displays the current configuration of the molecule. Fully rendered images are compressed and sent to the client using virtualGL. Mouse events generated by the user are sent back to the supercomputer on the Grid, which reacts accordingly (computational steering).

References

[1] **Liu, L., Botos, I., Wang, Y., Leonard, J.N., Shiloach, J., Segal, D.M., Davies, D.R.** Structural basis of toll-like receptor 3 signaling with double-stranded RNA, Science, v320, pp. 379-81, 2008

[2] **For more information on UNICORE, visit http://www.unicore.eu/**

[3] **For more information on the Globus toolkit, visit http://www.globus.org/**

[4] **Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L.,Schulten, K.** Scalable molecular dynamics with NAMD, J. Comp. Chem., 26(16), 1781-1802, 2005

[5] **For more information on NAMD and VMD, visit http://www.ks.uiuc.edu/**

[6] **For more information on virtualGL, visit http://www.virtualgl.org/**

- Ferdinand Jamitzky
- Ilya Saverchenko
- Helmut Satzger

Leibniz Supercomputing Centre

Supersonic Wall-bounded Turbulent Flows

- Somnath Ghosh
- Christoph Schaupp
- Rainer Friedrich

Technische Universität München, Lehrstuhl für Aerodynamik

Supersonic turbulent Flow in axisymmetric Nozzles and Diffusers

Direct numerical simulations of compressible turbulent channel flows have only recently led to proper understanding of compressibility effects in wall-bounded flows at supersonic speeds. Changes in density from the wall to the core region are responsible for changes in the turbulence structure (Reynolds stress anisotropy) in such flows. This opens up new challenges for turbulence modelling.

Expansion/compression of a supersonic wall-bounded flow, however, gives rise to more complicated effects on the turbulence structure which need to be investigated in detail in order to improve turbulence models. With this aim, we study supersonic flow through nozzles and diffusers with circular cross-sections by means of direct and large-eddy simulations. The incoming flow is a supersonic fully-developed pipe flow. The pipe flow and nozzle/diffuser flow simulations run simultaneously and the coupling is done by use of MPI routines. The DNS code reaches 400 Mflops/core while using 64 cores for the nozzle flow simulations. The number of grid points used for these simulations is 256*128*91 in the streamwise, circumferential and radial directions for both pipe and nozzle flow domains.

The expansion/compression effect in the nozzle/diffuser is kept weak when compared to the mean shear rate, but the effects on the turbulence structure are found to be dramatic. Mean expansion in the nozzle decreases the turbulence intensity drastically (Figure 1). Analysis of Reynolds stress transport equations shows that, as a dominant effect, the decrease of production and pressure-strain terms leads to the decrease in Reynolds stresses. Similar but opposite effects are found in the diffuser simulations. Here, the turbulence intensities are enhanced due to mean compression (Figure 2). The reasons are the same as in the nozzle, i.e. increase in production and pressure-strain terms. The DNS of the diffuser flow typically uses 256*256*140 points for the pipe flow domain and 384*256*140 for the diffuser flow domain and runs on 128 cores achieving 300 Mflops/core. The physics of the diffuser flow is more complicated than that of the nozzle flow and we hope to gain more insight from the DNS data.

Large-eddy simulations of the nozzle flow have also been carried out and the results are in good agreement with those from the DNS.

The outcome of this work is a comprehensive database of compressible wall-bounded flows undergoing acceleration and deceleration, which can be used to improve turbulence models.

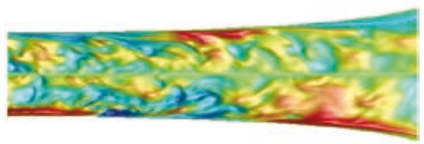


Figure 1: Axial velocity fluctuations in the nozzle (DNS)

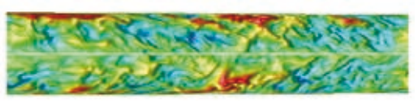


Figure 2: Axial velocity fluctuations in the diffuser (DNS)

Large Eddy Simulation of a reacting hydrogen Jet in supersonic Cross-flow

One concept for the propulsion of next-generation space transportation vehicles – reusable with increased payload compared to present-day designs – relies on the scramjet engine, which has been discussed for many years. The idea is to allow combustion of fuel using as long as possible the oxygen present in the atmosphere during the ascent from the earth's surface to orbit, rather than carrying all oxygen necessary on the vehicle itself. This concept appears straightforward, but difficulties arise at high flight Mach numbers. Compression of incoming air flow is achieved through a ramp system gradually decelerating the supersonic/hypersonic flow. Subsequently, heat is added via exothermic chemical reaction. Then, the flow is expanded in a nozzle to generate thrust. For this concept to work efficiently at high Mach numbers, the flow through the combustion chamber should stay in the supersonic regime. In spite of extensive research in the past, the injection of fuel, efficient mixing of fuel and air stream with low losses, and chemical reaction within a short combustion chamber are still great challenges in the design of such a propulsion device. One reason for this is that knowledge about combustion processes in turbulent supersonic flows is still incomplete. This study is intended to contribute to the understanding of the flow phenomena in question. As experimental evidence is difficult to obtain under the flow conditions considered, accurate numerical simulations may play an important role. Here, the method of Large Eddy Simulation (LES) is applied to the injection of a hydrogen jet into the supersonic turbulent flow in a plane channel. This flow situation, depicted in Figure 3, is representative of many aspects of fuel injection into a Scramjet combustion chamber, though it is geometrically simplified in order to allow the use of numerical methods of high accuracy on cartesian meshes. The flow phenomena expected, specially shock systems developing due to the interaction of the jet with the supersonic cross-flow, necessitate high resolution and thus a grid comprising a large number of points (17.9 million for the case shown). Additionally, time-accurate simulations of wall-bounded flows are very expensive due to small timesteps. Therefore, computations are performed in parallel on a number of CPUs of the high-performance computing system HLRB II at LRZ. Figure 4 shows the local Mach number in a cut through the domain centerline. The bulk Mach number of the oncoming channel flow is Ma = 2 and the bulk Reynolds number is Re = 5000. Important features of the flow, known from experimental investigations, are found in LES (separation upstream and downstream of the slot, bow and separation shocks, barrel shock structure and Mach disk).

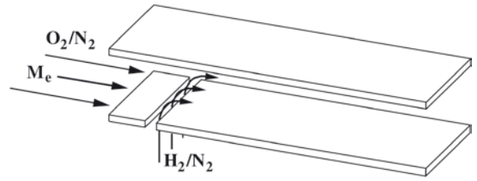


Figure 3: The flow situation considered: injection of a jet through a spanwise slot into plane supersonic channel flow. Reactants are diluted by nitrogen.

References

[1] **Ghosh, S., Sesterhenn, J., Friedrich, R.** Large-Eddy simulation of supersonic turbulent Flow in axisymmetric Nozzles and Diffusers, International Journal of Heat and Fluid Flow, 29, pp. 579-590, 2008

[2] **Schaupp, Ch., Friedrich, R.** Large-Eddy simulation of plane jet Injection into supersonic turbulent Cross-flow in: High-Performance Computing in Science and Engineering, Garching/Munich 2007, Springer-Verlag, 2008, to appear

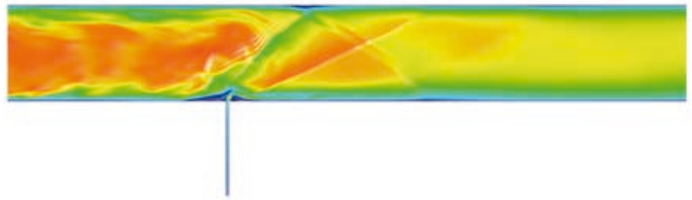


Figure 4: Plot of local Mach number M = |u|/c. The bulk Mach number of the incoming air flow is Ma = 2. Here, the entire computational domain is shown. All boundaries are placed as far from the injection orifice as possible to reduce the influence of approximate boundary conditions on the results.

On Supercomputing in Handwritten Whiteboard Note Recognition

Meetings play an important role in any organizational structure: the Intel Corporation schedules around 3 million meeting hours and another 56 thousand hours each year for training their employees on how to hold a meeting efficiently. High effort is taken to investigate how computers can be used to make meetings more efficient and analyze them automatically (see e.g. [1]), or making expensive meetings available in fully digital form. Therefore, special locations like the IDIAP smart meeting room (see [2]) use cameras, microphones, and a sampling device for recording notes written on a whiteboard. Automatic recognition of such recorded notes is a relatively new and a challenging task in pattern recognition (see [3]) and can be seen as a milestone towards analyzing meetings automatically.

Why using Supercomputers for handwriting Recognition?

Although represented in a similar form as on handheld-devices (so called personal digital assistants, PDAs), which already offer limited capabilities of handwriting recognition, the problem of whiteboard note recognition requires much more sophisticated algorithms, [3]. First, in contrast to handwriting recognition on a PDA, the recognizer cannot be trained on the specific handwriting of a single user: being an instrument to communicate ideas among the participants mutually, writer independent script recognition is required. Second, there is a need to recognize whole text lines rather than isolated words.

Besides, in a whiteboard scenario the writer stands rather than sits and moves the whole arm and body during writing introducing a high level of distortions to the handwriting. In order to test the writer independence of a recognizer, millions of characters have to be validated – a task, which can only be solved in parallel within reasonable time.

Recognition

For recognition of whiteboard notes, up to 24 features are extracted from all sample points of the recorded data, e. g. the pen pressure, describing whether the pen touches the whiteboard, the writing direction and the curvature of the pen trajectory. A complete list and short description can be found in [4]. After the feature extraction, each handwritten character can be described by the corresponding feature sequence. Due to varying writing speeds and writing styles, the same character may consist of a different number of sample points and different feature values. In order to compensate for these dynamic sequences of variable length, it is convenient to represent the character rather by the statistical distribution of its features than by the features themselves. This is realized by so-called Hidden-Markov-Models (HMMs, see [5]) which consist of two statistical processes: one describing the temporal structure of the handwritten character and another for representing the statistics of the observed features. The first process can be described as a Markov-chain, the latter e.g. by mixtures of Gaussian distributions. The estimation of each HMM's parameters, namely the transition probabilities of the underlying Markov-chain and the Gaussian mixtures, is a computational expensive task: starting with just one mixture, the transition probabilities are adjusted from training observations. Afterwards, the number of mixtures is increased by one and the transition probabilities are refined. This continues until a maximum of 32 mixtures is reached.

Recognition is performed by presenting an unknown script pattern to all HMMs and selecting the HMM which yields the highest observation probability for the unknown sample. A maximum-likelihood segmentation and recognition of connected characters forming words and whole text lines can be realized by interconnecting several HMMs. To prevent the models from over-fitting the training data, after each training iteration a validation is necessary. Due to the computationally expensive training procedure and the quite costly recognition process, the validation step after each iteration is usually omitted (see [6]), which compromises the quality of the models. By using the parallel computational capabilities of the HLRB II, a combined training and validation procedure can be realized.

Another goal of our efforts is to use the HLRB II to find a rigorously reduced feature set, achieving an acceptable, writer independent recognition performance and to further use these findings to improve computationally less expensive recognition algorithms (e.g. by replacing the Gaussian mixture distribution of the HMM observation by a discrete distribution) running on low cost hardware. Thus running these algorithms on the HLRB II can be seen as an important step towards robust whiteboard recognition on a standard computer.

Parallelization

In order to provide well-trained models that benefit from validating their parameter after each training iteration the parallelization, which is displayed in Figure 1, is used. Each of the three writer-independent subsets (one for training, one for validation, and one for a final, independent test), containing 5,500,000 training, 3,700,000 validation and 2,900,000 test observation frames, respectively, is partitioned according to the number N of CPUs used for parallelization and forms one instance each for training, validation and testing. For the training, where the model-parameters of the previous iteration are required, so-called

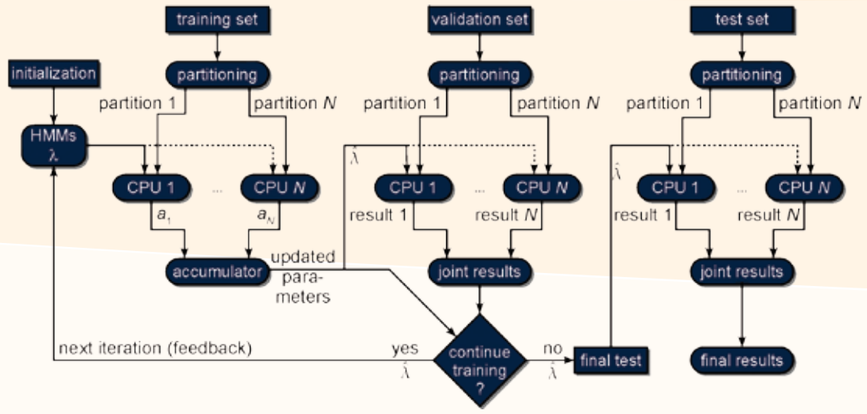


Figure 1: Parallelization scheme for HMM-based recognizer as used in on-line handwriting recognition. The training set is divided into N partitions (test instances) which individually train character models, deriving so-called accumulators (left). In a similar way, the validation (middle) and test procedure (right) can be parallelized.

accumulators "a_i" are the preliminary optimization results. After each training instance has finished calculating its accumulator, all accumulators are combined to a new model. After training, the model parameters are validated. Thereby, each validation instance produces a result, which is again combined to a score for the current model parameters. Depending on the system's design parameters (e.g. the maximum number of iterations or the number of Gaussians), the model parameters are either fed back to the training system providing the new initialization of another training iteration, the number of Gaussian mixtures is increased, or evaluated on the independent test set.

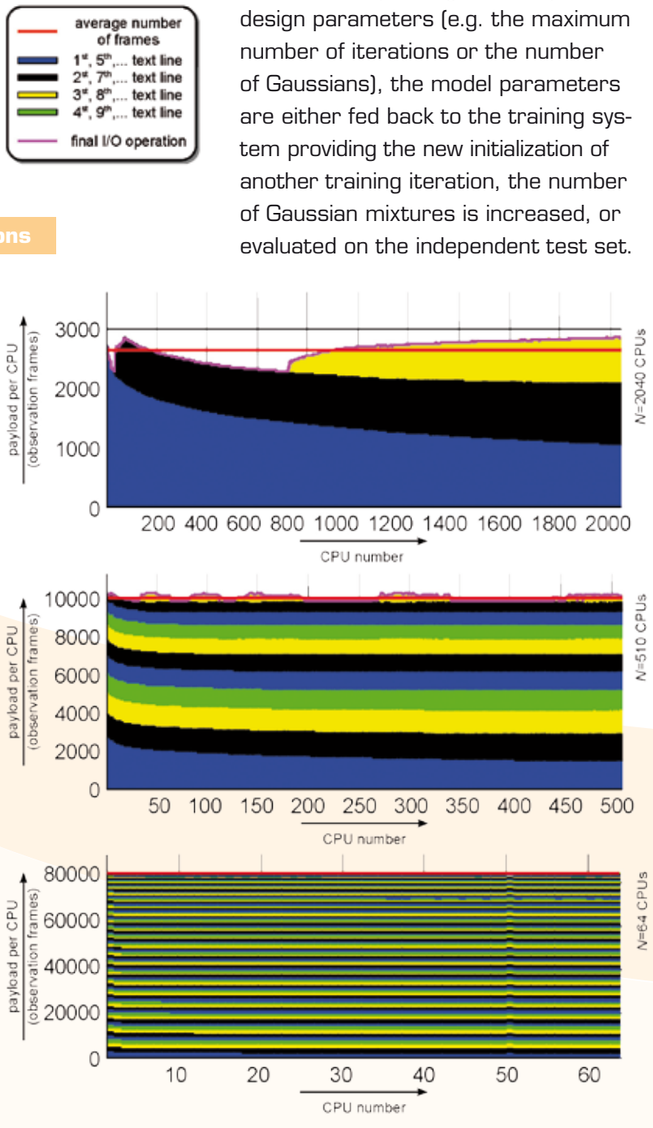


Figure 2: Average (red line) and individual payload per CPU represented by the number of observation frames per CPU and iteration for varying number N of CPUs

A crucial factor influencing the computational performance of the system is the balancing of the training, validation, and test instances. This is provided by a frame-wise partitioning of the data. In order to show the scalability of our parallelization scheme, the number of observation frames processed by each CPU per iteration (the "payload") is shown Figure 2, in case of using either N=64, N=510, or N=2040 CPUs in parallel. Additionally, the average number of observation frames (the average payload) is given. For N=64 and N=510, the CPUs' individual computational payload is well balanced with little deviation from the average payload. In a realistic run, a parallelization using N=64 … 256 CPUs performs adequately efficient. All parameters, results, and data are kept in memory for providing the necessary I/O-operation at the end of each iteration at a high speed.

Preliminary Results

One important task is to lower the computational requirements of a HMM-based handwritten whiteboard note recognizer. Instead of modeling the observation probabilities of the HMMs by mixtures of Gaussians, the features can be quantized by a vector quantizer [4]. Their actual occurrence can then be used as observation probability, leading to discrete HMMs. When performing quantization, it turns out that the pressure information looses its significance (see [4]) as illustrated in Figure 3: the red line denotes the character level accuracy of a_b=62.6% for a discrete HMM system, using all 24 features for quantization. The green line indicates the accuracy of a_r=62.5% of a discrete system where the feature vector has been reduced by the pen's pressure information before quantization – a slight relative drop of r=0.2%.

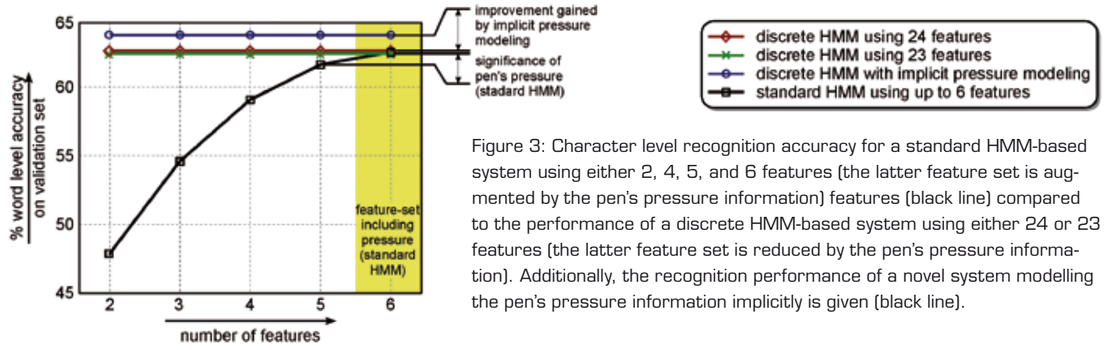


Figure 3: Character level recognition accuracy for a standard HMM-based system using either 2, 4, 5, and 6 features (the latter feature set is augmented by the pen's pressure information) features (black line) compared to the performance of a discrete HMM-based system using either 24 or 23 features (the latter feature set is reduced by the pen's pressure information). Additionally, the recognition performance of a novel system modelling the pen's pressure information implicitly is given (black line).

Additionally, Figure 3 shows the character accuracies of a continuous system using varying numbers of features, chosen by the sequential forward floating selection algorithm [7]. As can be seen, the pressure Information is one of the six most significant features amongst the selected features. These findings motivate a novel VQ-design, enabling an implicit modeling of the pen's pressure information, [4]. The result is also shown: a significant, relative improvement of r=1.8 % can be reported in case of the implicit modeling of the pressure. This shows how the results derived by supercomputing can improve low cost systems.

Outlook

The main reason for using supercomputing in the field of handwritten whiteboard note recognition is to get novel insights: they can enable an improvement of existing recognition algorithms running on devices with less performance. Taking the loss of the information of the pen's pressure, when vector quantization is performed, as an example, investigations conducted on the HLRB II led to an implicit modeling of the pressure information and thereby to a significant improvement. In future work, supercomputers can help to investigate the ability of combining several approaches for handwriting recognition, providing a well-selected set of recognizers and parameters for a more accurate recognition performance at much lower computational cost.

References

[1] **Al-Hames, M., Lenz, C., Reiter, S., Schenk, J., Rigoll, G.**
Robust Multi-Modal Group Action Recognition in Meetings from Disturbed Videos with the Asynchronous Hidden Markov Model, Proceedings of the International Conference on Image Processing, pp. 213–216, 2007

[2] **Moore, D.**
The IDIAP smart meeting room, Research Report 7, 2002

[3] **Liwicki, M., Bunke, H.**
HMM-based on-line Recognition of handwritten Whiteboard Notes, Proceedings of the IWFHR, pp. 595-599, 2006

[4] **Schenk, J., Schwärzler, S., Ruske, G., Rigoll, G.**
Novel VQ Designs for discrete HMM on-line handwritten Whiteboard Note Recognition, Pattern Recognition, LNCS 5096, pp. 234-243, 2008

[5] **Rabiner, L.R.**
A Tutorial on Hidden Markov Models and selected Applications in Speech Recognition, Proceedings of the IEEE, 77(2), pp. 257-285, 1989

[6] **Günter, S., Bunke, H.**
HMM-based handwritten Word Recognition: on the Optimization of the Number of States, Training Iterations, and Gaussian Components, Pattern Recognition, 37(10), pp. 2069-2079, 2004

[7] **Pudil, P., Novovicova, J., Kittler, J.**
Floating Search Methods in Feature Selection, Pattern Recognition Letters, 15(11), pp. 1119-1125, 1994

• Joachim Schenk

Technische Universität München, Departement for Human-Machine Communication

Autumn 2008 • Vol. 6 No. 2 • inSiDE

Prebiotic Peptide Synthesis on Blue Gene Platforms at "Iron–Sulfur–World" Conditions

The exploration of possible scenarios for prebiotic molecular synthesis, including biopolymers such as peptides, is paramount to understanding how primitive life emerged on the young Earth. Amongst a vast amount of different hypotheses, evidence accumulated that mineral surfaces in conjunction with water at extreme thermodynamic conditions might offer favorable reaction environments. In particular, an unusually precise proposition for biomolecular synthesis on iron/sulfur minerals in hot-pressurized water as found close to deep-sea hydrothermal vents has been detailed in the framework of the so-called "Iron-Sulfur-World" (ISW) scenario [1]. Certainly, understanding the interplaying fundamental issues of unusual chemical reactivity at extreme conditions, liquid state theory of solvation, and physical chemistry of mineral/water interfaces turn out to be of overriding importance here. In the intricate reaction chain from small molecules to functional proteins the formation of the peptide bond as such is, without any doubt, a key step. Although significant experimental support of the major ingredients has been accumulated by several groups, pertinent experiments lack detailed molecular insight into how small "inorganic" reactants transform into biomacromolecular products. Still, preliminary ab initio molecular dynamics (AIMD) simulations at ISW conditions were carried out only a few years back [2-4].

In the long-term project "Full in Silico Exploration of Possible Routes to Prebiotic Peptide Synthesis by Ab Initio Metadynamics" devoted to fundamental research in Chemistry the importance of high temperature and pressure ISW reaction conditions including mineral-surfaces is being assessed by AIMD techniques. The primary goal here is to provide vital molecular level understanding about the pertinent reactions in the "virtual lab" [5] which is otherwise difficult or even impossible to obtain in real laboratory experiments. Greatly expanding our initial work [2-4], our computations on the IBM Blue Gene systems JUBL and JUGENE at the John von Neumann Institute for Computing (NIC) at Forschungszentrum Jülich (FZJ) during the last two years have unveiled very important mechanistic and energetic details of peptide synthesis at ISW conditions. It is stressed that the unprecedented computational complexity of our in silico prebiotic peptide synthesis demanded an investment of up until now about five nanoseconds of AIMD simulation time in total, which was only possible due to generous access to these efficient resources at NIC. In 2008 the above-mentioned project has been elected to be the first "NIC Excellence Project of the Year".

The "Virtual Lab" Approach to Chemistry

Recent advances in both computer technologies and simulation methods, in particular Car-Parrinello AIMD [6,7] in conjunction with efficient sampling methods like the powerful "metadynamics" technique developed by Laio and Parrinello (see Ref. [8] for a review), make it possible to study truely complex chemical reaction networks in the "virtual lab" [5]. All calculations presented here were performed using Hohenberg-Kohn-Sham density functional theory in its efficient plane wave pseudopotential implementation [7] within the CPMD software package [9].

Since the beginning of the 1990ies, the CPMD code [9] has been designed by Jürg Hutter from the onset to run efficiently on all kinds of parallel platforms as explained in detail in NIC Lecture Notes [7]. However, low-latency interconnects are required to run this parallel AIMD code efficiently. To get around load balancing problems on platforms of ever growing processor numbers a second level of parallelization named the "task-grouping" of processors has been implemented into CPMD some time ago [7]. Furthermore, the hierarchical multi-level strategies [7] that combine distributed-memory and shared-memory parallelization are highly suited for ultra-dense massively parallel HPC machines such as the Blue Gene architecture [10] in particular.

However, electronic structure calculations and thus AIMD require non-trivial parallelization strategies as the character of the underlying off-lattice quantum problem is not easily suited for partitioning without making use of further approximations such as done in linear scaling algorithms. This is due to the spatially non-local character of quantum-mechanical wavefunctions. The multiple walker metadynamics algorithm [8], which we have successfully implemented, is a linear scaling algorithm in itself and thus improves the net scaling behavior of CPMD for our given system sizes at hand on a large number of processors. Figure 1 shows the performance of CPMD on the JUGENE Blue Gene/P installation at NIC employing up to one half of the whole machine with and without multiple walkers. Still, due to the inherent scaling limitations in any quantum simulation code for typical system sizes like that in the test case shown, it is much less efficient to go beyond a Blue Gene mid-plane for practical simulations; see for instance in Figure 1

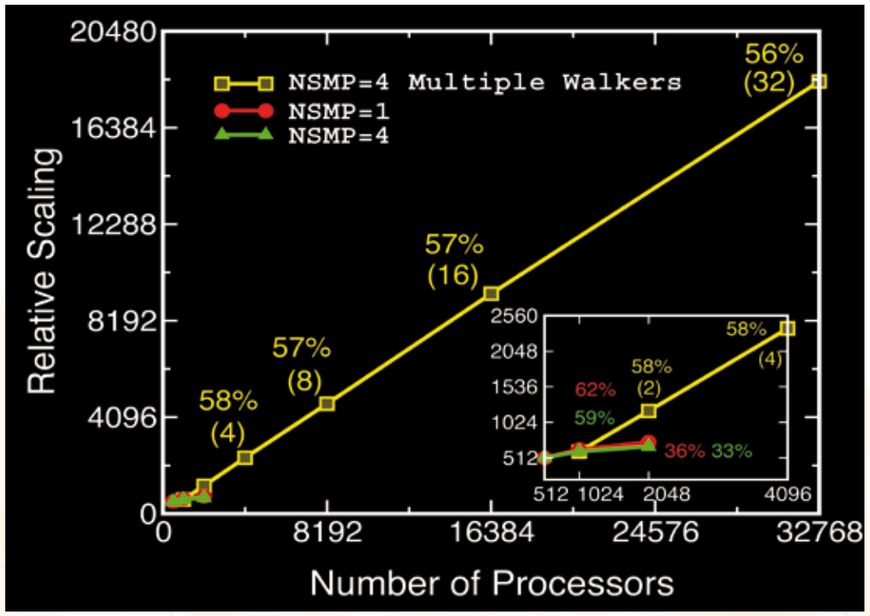


Figure 1: Relative scaling performance when using N processors with respect to 512 processors (i.e. the ratio of the computer time for one AIMD step per processor using N processors to that for 512 processors multiplied by N) for the CPMD code for one of the systems studied on JUGENE at NIC. Red and green lines are using Open MP threads 1 and 4, respectively. The yellow line is the performance using the multiple walker technique where the number of walkers, n, is reported in parenthesis. In the inset the scaling behavior going from 512 processors to 4,098 processors is magnified for clarity.

that the scaling dropped down to about 33 % when using a full mid-plane with respect to using only a quarter of a mid-plane. However, the parallelization among walker replicas is extremely efficient due to the loose coupling of the walkers such that the communication characteristics of Blue Gene systems can be exploited using many racks for a single AIMD simulation. As usage of ***n*** walkers will decrease the total length of the simulation by a factor of about ***n*** the effective scaling of a multiple walker algorithm can be estimated based on CPU time per AIMD step divided by ***n***. It is evident based on the scaling behavior shown in Figure 1 that use of all processors of JUGENE is now possible for our system of interest without the need to increase the size of the system even when certain processor groups must communicate via high-latency interconnects.

The Key Result: Peptide Synthesis Cycle

The simulations of peptide bond formation [11,12] between two glycine molecules were carried out using three vastly different reaction conditions: ambient bulk water at about 300 K and 0.1 MPa (ABW), hot-pressurized bulk water at about 500 K and 20 MPa (HPW), and hot-pressurized water at the pyrite interface (PIW). The effective free energy barriers estimated along the peptide synthesis cycle leading to diglycine are reported in Figure 2. It is evident from the mechanism depicted in Figure 2 that it is the neutral form of the amino acid glycine 2 that is required and not the zwitterionic form 1 for its reaction with the COS molecule to form thiocarbamate (see step B). The thiocarbamate 3, in turn, leads to an activated amino acid in form of its so-called Leuchs anhydride 5 that easily adds to another amino acid (or peptide) to form a peptide bond (in step E) which finally yields an elongated peptide. As apparent by the computed free energy surface Figure 3 HPW extreme conditions stabilize the neutral form 2, consistent with the lowering of the dielectric constant of HPW, whereas in ABW neutral glycine converts easily to the zwitterion 1 on an ultrafast timescale of ≈ 1 ps.

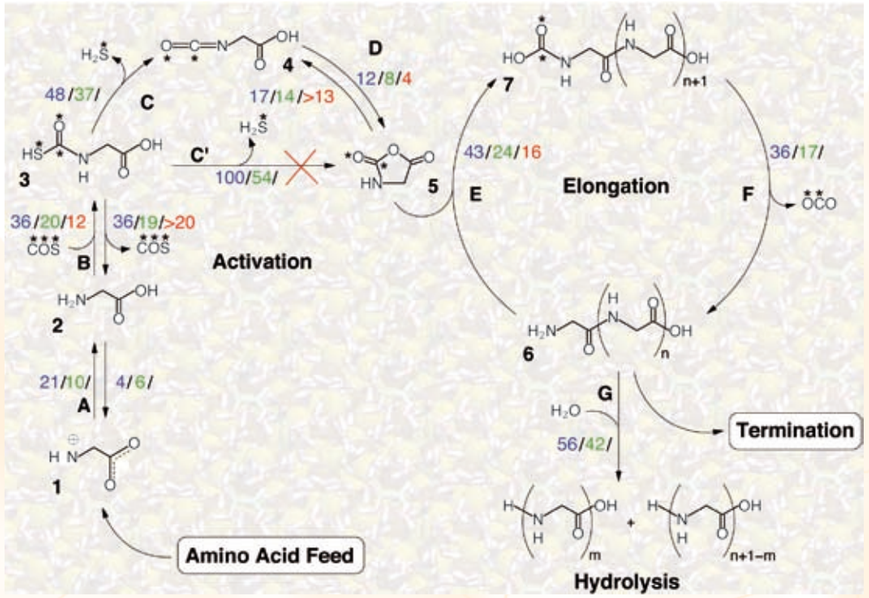


Figure 2: Full peptide synthesis cycle comprising input of an amino acid (or peptide), here glycine, and its activation followed by elongation using another amino acid (or peptide), here another glycine, as well as termination and hydrolysis as a major reverse reaction. The calculated free energy barriers (given in k_BT energy units) for individual steps of the mechanism leading to diglycine formation are color coded. Blue: ambient bulk water (ABW), green: hot-pressurized bulk water (HPW), red: hot-pressurized water at the pyrite interface (PIW). The crossed direct formation path C' is very unlikely in view of its high activation free energy compared to the indirect path via isocyanate 4. See Ref. [11] for details.

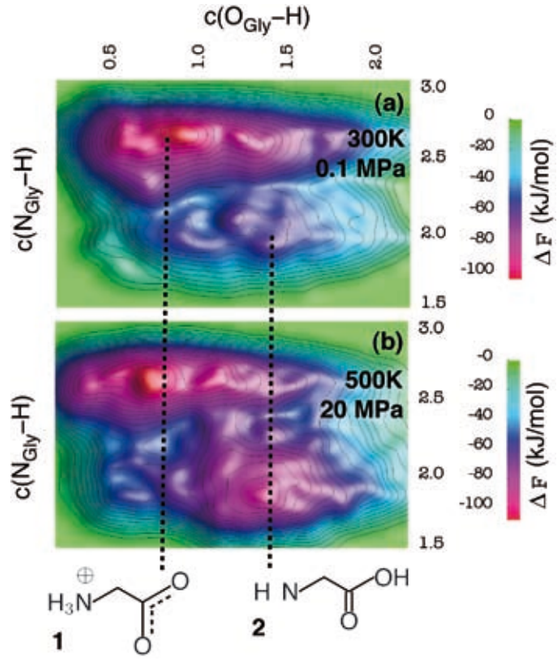


Figure 3: Free energy surface for the conversion between zwitterionic 1 and neutral form 2 of glycine (a) in ambient bulk water (ABW) and (b) in hot-pressurized bulk water (HPW); color bar shows the relative free energy ΔF in kJ/mol energy units. Metadynamics AIMD simulations were performed using two collective variables: coordination numbers of nitrogen and carboxy oxygen to all hydrogen atoms in the system, $c(N_{Gly} - H)$ and $c(O_{Gly} - H)$, respectively. See Ref. [12] for details.

In other words, such extreme thermodynamic conditions are found to increase the concentration of the neutral amino acid by shifting the equilibrium between the neutral and the charged zwitterionic forms of amino acids toward neutral form, thus favoring the formation of peptides. This is an interesting result and immediately reveals the importance of hot-pressurized conditions for this route to peptides. Moreover, it was found that these extreme HPW conditions speed up the production of peptides by accelerating individual steps of the whole peptide synthesis cycle according to the free energy barriers reported in Figure 2.

Another discovery from our simulations is the so-called isocyanate pathway leading to the formation of the activated form of amino acid, Leuchs anhydride 5, from thiocarbamate 3. Compared to a direct cyclization of the thiocarbamate 3 to form Leuchs anhydride 5, the indirect isocyanate pathway, i. e. first forming an isocyanate 4 which rapidly cyclizes to Leuchs anhydride, is very much lower in terms of free energy barriers. This result confirms earlier experimental speculations about such a route including an equilibrium between 5 and 4. Our calculations also shed light on the productivity of the cycle including the formation of isocyanate being the rate-determining step of the whole peptide cycle. Based on a simple estimate, the time scale for forming peptide bonds along this route is in the order of a few minutes at hot-pressurized conditions whereas it would be several years at ambient conditions! It is also shown by the simulations that hydrolysis of the synthesized peptide is slower than the rate-determining formation step and, therefore, a net accumulation of peptide can be expected in agreement with experimental conclusions.

Possible roles of Fe/S mineral surfaces were also investigated via simulations using an ideal pyrite surface, FeS_2(001), as the simplest model. By decreasing the entropic contribution to the free

energy barriers this surface is found to accelerate several reaction steps in the peptide synthesis cycle by lowering the corresponding free energy barriers up to a factor of two, which increases the respective reaction rates exponentially. In addition, more interesting effects like scaffolding of reactant molecules favoring the formation of transition states and thereby speeding up the reaction by several orders of magnitude have been identified. In particular, the pyrite surface is found to favor the preformation of the five-membered ring that is characteristic of Leuchs anhydride 5 as demonstrated in Figure 4.

Based on free energy calculations and careful examination of various reaction mechanisms, our studies underpin the importance of extreme conditions and mineral surfaces for peptide synthesis along a putative route [11,12]. These comprehensive simulations delineate pathways connecting the crucial activation and elongation steps through which peptides can be produced out of amino acids and COS via an indirect isocyanate/Leuchs anhydride route. Importantly, the data provide convincing evidence that all steps along the proposed synthesis cycle are clearly favored in hot-pressurized water when compared to ambient conditions providing in total a productive synthesis cycle [11]. Beyond the specific case, these findings imply that "different chemistry" must be considered when discussing putative prebiotic synthesis scenarios at extreme aqueous conditions.

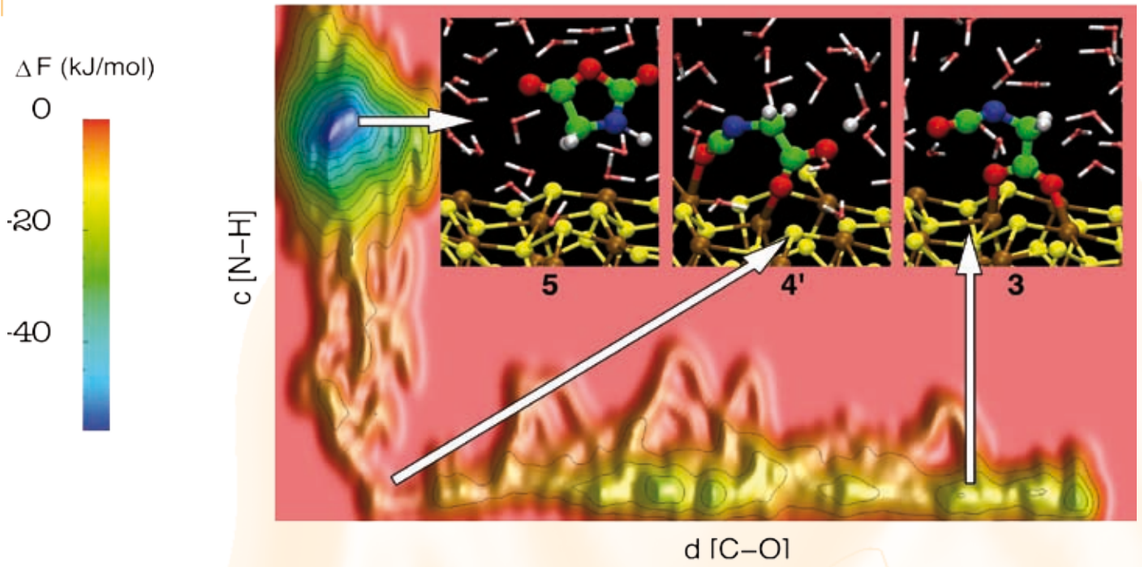


Figure 4: Mechanism of the formation of Leuchs anhydride 5 from thiocarbamate 3 in hot–pressurized water at the pyrite interface (PIW) based on the free energy surface; color bar shows the relative free energy ΔF in kJ/mol energy units. Metadynamics AIMD simulations were performed using two collective variables: distance between the carbon atom of the COS entity to one of the carboxylate oxygen atoms d[C–O] and coordination number of nitrogen to all hydrogen atoms in the system c[N-H]. Three representative real space configurations sampled from these simulations at the pyrite–water interface demonstrate scaffolding due to FeS_2(001) by preformation of the cyclic topology of Leuchs anhydride upon bidentate chemisorption. Color code: hydrogen (white), oxygen (red), carbon (green), nitrogen (blue), sulfur (yellow), iron (ocher); labeling is according to Figure 2. See Ref. [11] and upcoming publications for details.

What Next?

Although a significant step forward, the cycle in Figure 2 is based on a set of disconnected free energy calculations whereas a single, global free energy landscape is necessary to fully explore this rather complex reaction network including the roles of reverse and side reactions. This, again, will be a challenge to both algorithms to sample free energies beyond three or four dimensions and platforms to carry out such a unified, ultramassive simulation. Another key issue that still remains largely unexplored is the role of mineral surfaces since defective surfaces are known to be most active in heterogeneous catalysis whereas an ideal, non-defective pyrite has been used up to now in our "virtual lab" [5]. Including defects at the mineral/water interface, most desirably as "dynamical degrees of freedom" in an AIMD simulation, certainly adds a lot more complexity to the problem but would be another necessary step forward. However, looking back at recents developments it appears to optimists that such dreams might become true much sooner than currently expected.

Thanks

We are most grateful to Alessandro Curioni and Jürg Hutter for help in porting CPMD onto the Blue Gene/L and P platforms JUBL and JUGENE at NIC. This research is supported by Deutsche Forschungsgemeinschaft via Normalverfahren MA 1547/7.

References

[1] **Wächtershäuser, G.**
Groundworks for an Evolutionary Biochemistry: The Iron–Sulfur World, Progress in Biophysics and Molecular Biology, 58, pp. 85-201, 1992

[2] **Boehme, C., Marx, D.**
Glycine on a Wet Pyrite Surface at Extreme Conditions, The Journal of the American Chemical Society, 125, pp. 13362-13363, 2003

[3] **Pollet, R., Boehme, C., Marx, D.**
Ab Initio Simulations of Desorption and Reactivity of Glycine at a Water-Pyrite Interface at "Iron-Sulfur World" Prebiotic Conditions, Origins of Life and Evolution of Biospheres, 36, pp. 363-379, 2006

[4] **Nair, N. N., Schreiner, E., Marx, D.**
Glycine at the Pyrite-Water Interface: The Role of Surface Defects, The Journal of the American Chemical Society, 128, pp. 13815-13826, 2006

[5] **Marx, D.**
Theoretical Chemistry in the 21st Century: The "Virtual Lab", in: Proceedings of the "Idea–Finding Symposium: Frankfurt Institute for Advanced Studies", (Greiner, W.; Reinhardt, J., Eds.), EP Systema, Debrecen, pp. 139-153, 2004

[6] **Car, R., Parrinello, M.**
Unified Approach for Molecular Dynamics and Density–Functional Theory, Physical Review Letters, 55, pp. 2471-2474, 1985

[7] **Marx, D., Hutter, J.**
Ab Initio Molecular Dynamics: Theory and Implementation, in: Modern Methods and Algorithms of Quantum Chemistry, (Grotendorst, J., Ed.), John von Neumann Institute for Computing (NIC), Forschungszentrum Jülich, Germany, Vol. 3, pp. 301-449, 2000

[8] **Laio, A., Parrinello, M.**
Computing Free Energies and Accelerating Rare Events with Metadynamics, in: Computer Simulations in Condensed Matter: From Materials to Chemical Biology, Ferrario, M.; Ciccotti, G.; Binder, K., Eds., Springer-Verlag, Berlin Heidelberg, Vol. 1, pp. 315-347, 2006

[9] **CPMD**
Hutter, J. et al., Copyright: IBM Corp 1990-2008 and MPI für Festkörperforschung Stuttgart 1997-2001; see http://www.cpmd.org

[10] **Hutter, J., Curioni, A.**
Car-Parrinello Molecular Dynamics on Massively Parallel Computers, ChemPhysChem, 6, pp. 1788-1793, 2005

[11] **Schreiner, E., Nair, N. N., Marx, D.**
Influence of Extreme Thermodynamic Conditions and Pyrite Surfaces on Peptide Synthesis in Aqueous Media, The Journal of the American Chemical Society, 130, pp. 2768-2770, 2008

[12] **Nair, N. N., Schreiner, E., Marx, D.**
Peptide Synthesis in Aqueous Environments: The Role of Extreme Conditions on Amino Acid Activation, The Journal of the American Chemical Society, in press, 2008

• Nisanth N. Nair
• Eduard Schreiner
• Dominik Marx

Ruhr-Universität Bochum, Lehrstuhl für Theoretische Chemie

Autumn 2008 • Vol. 6 No. 2 • inSiDE

The Mystery of the Dark Side

During the last three decades cosmology became a very important part of astronomy. An increasing number of large ground based telescopes and very sensitive satellite-based telescopes delivered many data which lead to new insights about the properties of the present universe and its evolution. Meanwhile such parameters like the mean density or the expansion velocity are already known with high precision. According to the current paradigm the total mean energy density of the universe is equal to the critical density in Friedmann's equation and, therefore, the three-dimensional space is flat. However, baryons, i.e. the normal observed matter, contribute only about 5% to this density. The remaining 95% are made of some still unknown Dark Matter and Dark Energy, often called the Dark Side of the Universe. One of the main aims of modern cosmology is to unravel the mystery of the dark components in our Universe.

Dark Matter means a kind of particles which interact only gravitationally. These still undetected particles are a necessary ingredient without which one could not understand and explain both the formation of large scale structure and the inner structure of galaxies as measured from their rotation curves. But, according to observations, these particles only contribute about 21% of the measured total energy density of the universe. The remaining 74% are made of Dark Energy, an even more mysterious ingredient of our universe. As opposed to matter, Dark Energy is assumed to be homogeneously distributed in the universe. It has a constant energy density and a negative pressure that – according to General Relativity – accelerates the cosmic expansion. Thus, the cosmic pie consists of about 5% normal matter (which includes also the negligible contribution of radiation and neutrinos at present), 21% of Dark Matter and 74% of Dark Energy.

We observe that the universe expands with a velocity of about 72 km/s/Mpc, the Hubble constant. Thus temperature and matter density decrease with time while the dark energy stays constant. Looking backwards in time the density becomes infinite at about 13.7 billion years ago – the Big Bang. According to theory, during an inflationary stage, shortly after Big Bang, the universe expanded exponentially and quantum fluctuations became scale free classical density fluctuations. Later-on the density fluctuations in the dark matter component started to grow whereas the baryons were tightly coupled to the radiation field until the temperature decreased below a critical value and free electrons and nuclei combined into atoms. This occurred some 370,000 years after the Big Bang. At this moment the universe became transparent to radiation that is being observed today as a cosmic microwave background radiation (Nobel prize 1978 for its first detection).

Tiny fluctuations in the temperature of this radiation (Nobel prize 2006 for the detection of the fluctuations) give us insight into the properties of the early universe when the density fluctuations were small and could be described analytically by linear theory. During the following more than 13 billion years the fluctuations grew due to gravitational instability and formed the observed nonlinear structures in the universe. Since most of the particles are dark matter particles which interact only gravitationally they constitute the backbone of the large scale structure within which galaxies formed. This process of nonlinear gravitational clustering is extremely complex and can be described only by means of sophisticated numerical simulations that challenge the most powerful supercomputers in the world. These simulations take into account the expansion of the universe according to General Relativity. As we said before, the expansion depends on the properties of the Dark Energy. The matter density field within a simulated volume is represented by point particles. The interaction between these particles is described according to Newton's gravitational law. Therefore, the non-linear evolution of the initial density field is equivalent to compute the gravitational forces among a set of N-body particles. The larger the number of particles used, the better the resolution of the simulation will be. At present, high resolution simulations contain typically billions of N-body particles. Depending on the desired resolution, several 100,000 CPU hours are necessary for such simulations on the most powerful parallel computers, thanks to the development of efficient parallel algorithms. The situation becomes even more complicated if baryons are also taken into account. Unlike Dark Matter, baryons do not only interact gravitationally, but also through other fundamental interactions (electromagnetic and nuclear as well). Such gas-dynamical simulations which include also radiative cooling processes, star formation and the feedback of stars on the gas are much more complex and, therefore, numerically much more expensive. In most cases only the evolution of relatively small volumes can be simulated.

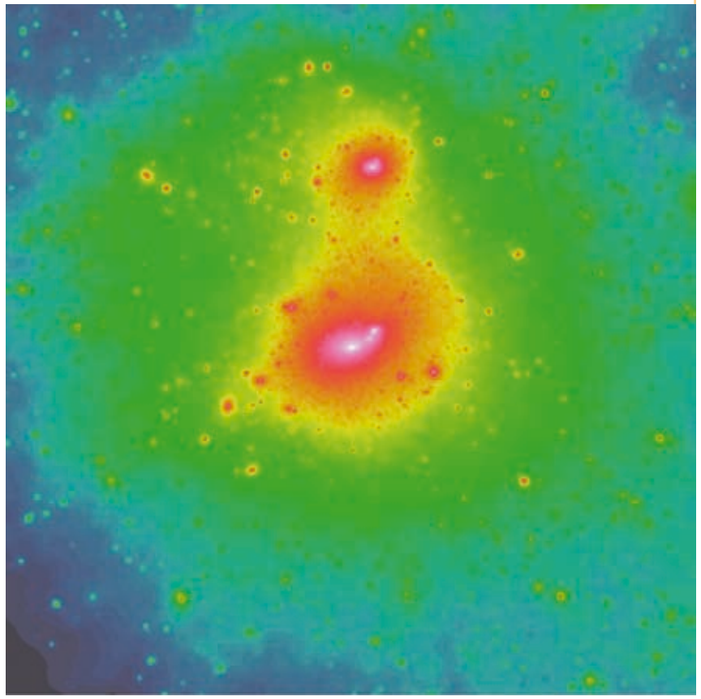


Figure 1: A Milky-way sized halo is simulated in the Cold Dark Matter cosmology. The dark matter distribution is shown color coded to the logarithm of the projected density. Besides the massive satellite, close to the center, more than hundred smaller ones are clearly visible in the figure.

Within the DEISA Extreme Computing Initiative we were granted with 2 million CPU hours during the past two years. We used this computational time on supercomputers in Barcelona, Munich and Jülich to study the problems of small scale structure formation. A substantial part of these simulations has been performed at MareNostrum in Barcelona whereas most of the analysis has been done on the shared memory machines at Jülich and Munich. Most of these simulations were done with 1024^3 particles, although the initial conditions were generated with 64 times more (4096^3) particles. This allows us to resimulate some specific objects in the simulation boxes with higher resolution (up to 64 times more than the rest of the box). The galactic object shown in the pictures is one such example.

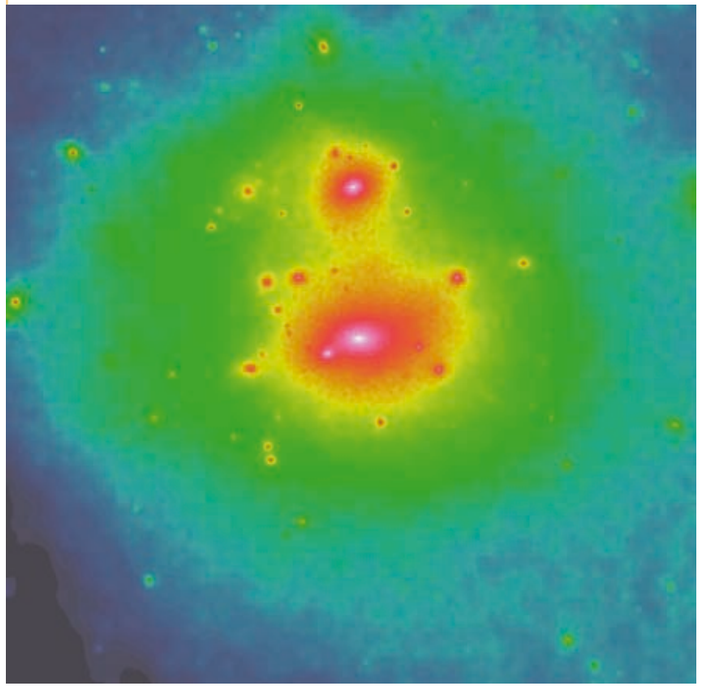


Figure 2: The same halo as shown in Figure 1 but this time simulated with Warm Dark Matter particles. Their projected distribution is shown at about the same time moment than in the previous case. As can be easily appreciated, there are much less satellites than in Figure 1.

At least three co-ordinates and three velocities must be stored for 1024^3 particles. Thus, in single precision and for dark matter only simulations at least 24 GB of data have to be stored for each time step. Since we are interested in the evolution of structures we store typically 135 time steps (every 100 mega-years). Thus terabytes of data have been generated. These data had to be transferred between the different supercomputers for permanent storage and to perform many different data analyzes. During this data transfer we have benefited from the DEISA high capacity dedicated network among the different computer centers. We used the Globus software for parallel data transfer with sustained transfer rates of several tens of Mbytes/s.

The inflationary paradigm predicts a scale free initial spectrum of density perturbations which has widely been confirmed on large scales by measurements of the cosmic microwave temperature fluctuations. Thus, massive objects like galaxy clusters with hundreds of galaxies are expected to look similar to galactic size dark matter halos. However real galaxies do not have hundreds of satellites, as it is shown in simulations, but only a few tens, in the best case. There are two possible explanations for the absence of satellites in simulations: Either the satellites do not exist at all or they are invisible because they do not contain stars. In the first case, this would be related to the nature of the dark matter particles, while in the other case, it would require a proper modeling of the baryonic physics. In order to disentangle between these two situations, we have been doing large scale N-body simulations with different dark matter candidates: a massive (more than Giga electron Volt mass) weakly interacting particle named as Cold Dark Matter (CDM) and another candidate with lighter mass (of the order of kilo electron Volt) called Warm Dark Matter (WDM).

Comparing Figure 1 and Figure 2 we can appreciate the strong difference in the number of satellites orbiting around a galactic size dark matter halo depending on the kind of dark matter particles assumed. We show the projected dark matter density around one halo color-coded according to the logarithm of the density in each pixel. Within this dark matter halo resides a Milky Way sized galaxy (in the center of the plot) with one very massive satellite in the upper part. All the other satellites have much less mass. Figure 1 corresponds to the CDM simulation, while Figure 2 is for the WDM simulation. As can be clearly seen, the number of low mass satellites in Figure 2 is considerably smaller: If we assume that the dark matter particles have a relatively light mass, i.e. Warm Dark Matter, the small scale fluctuations are washed out and during further evolution only a very few low mass satellites formed. This would explain why we observe only a few satellites within our Local Group.

Another open problem is the nature of Dark Energy. The simplest assumption is a constant vacuum energy which acts like a cosmological constant in Einstein's equations. A straightforward generalization is dark energy with a time-independent equation of state. Such a generalization leads to a slightly modified expansion history of the universe and, therefore, to a slightly different evolution of objects. Within the German astrogrid we were running a set of simulations assuming different equations of state of the Dark Energy. The number of objects above a certain mass per volume element (the mass function) changes with time and depends on the expansion speed. The tiny oscillations in the power spectrum are another characteristics. Besides a model with a cosmological constant (constant vacuum energy) we simulated also the evolution of structures in Dark Energy models with negative pressure as a function of energy density. The predicted mass functions and power spectra at different redshifts can be compared with future observations to constrain the parameters of the equation of state of Dark Energy.

• Stefan Gottlöber[1]
• Gustavo Yepes[2]

[1] Astrophysikalisches Institut Potsdam

[2] Universidad Autónoma de Madrid, Grupo de Astrofísica

Stuttgart Research Centre for Simulation Technology and Cluster of Excellence in Simulation Technology SimTech Cluster of Excellence

From Isolated Numerical Approaches to an Integrative Systems Science

In the 21st century's society, simulation technology has become completely indispensable, it dominates all areas of life. The cluster of Excellence in Simulation Technology, approved within the German national "Initiative of Excellence" is embedded in the Stuttgart Research Centre for Simulation Technology. It represents a massive platform for further developing scientific methods and techniques in all branches of modeling and simulation techniques. HLRS is taking a lead in the cluster being responsible for High Performance Computing research.

Scientific Goals

Currently, Simulation Technology is one of the most challenging research areas in contemporary engineering and natural sciences. It offers solutions for a broad variety of applications with an enormous economic and scientific impact. Recent and future developments in technology and society pose challenges which require Simulation Technology to advance to a new level. The Cluster of Excellence in Simulation Technology will strengthen the sustainable evolution of research areas that are of fundamental importance in both science and economy, as well as to conceive new and visionary applications of simulation sciences. The goal of SimTech is to further intensify the university's engagement in the field of Simulation Technology in order to maintain and expand the excellent standing in this important field in a dedicated intersectorial structure. Establishing integrative prediction methods and solutions for future challenges, the cluster concentrates on six methodical Research Areas ranging from Molecular Dynamics and Advanced Mechanics via Computational Mathematics and Systems Analysis to Data Management and Interactive Visualisation as well as to Hybrid High-Performance Computing Systems and Simulation Software Engineering. An Integrative Platform acts as a bracket of reflection and evaluation. Through transfer of scientific results to applications, the cluster will furthermore contribute to the economic benefit of our society.

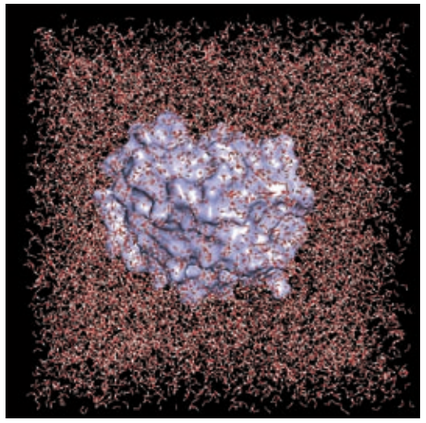


Figure 1: Simulation of a β-lactamase enzyme in water

The Role of HLRS in SimTech

The director of HLRS Prof. Michael Resch is a Principal Investigator and one of the two Coordinators responsible for the Research Area: Hybrid High Performance Computing. The focus of the work of HLRS in the cluster of excellence is on Parallel Programming Models and Tools. Models and tools to handle large scale hybrid systems are needed to harvest the potential of millions of components in future supercomputing systems hiding the complexity from the non-expert user. HLRS has therefore set the following short and medium term goals for the work in the cluster

- Development of parallel programming models capable of extending existing approaches to large scale systems and later to hybrid systems
- Development of methods and tools to support large scale parallel programming and further to hybrid systems
- Development of resource management methods to describe and manage workflows for large scale simulations on distributed and hybrid systems

Resource Management

The efficient utilization of concepts of parallelization confronts the scientific software developer with many problems when having to solve complex applications on distributed heterogeneous resources or Grid resources. Very often this is due to the fact that the subject areas are already highly complex in themselves. So far for many fields there has been no systematic approach for the development of complex time-critical Grid applications. HLRS will be devoted

- to the study of the peculiarities of the organization of distributed heterogeneous computation environments and of their components in order to solve the problem of the efficient execution of complex applications

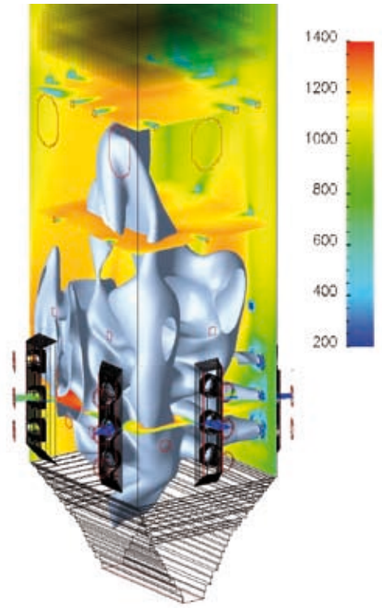


Figure 2: Temperature distribution [°C] as 3D projection with 1,300° C isosurface (grey) to visualize the fire location

- to the development of resource management methods guaranteeing a certain level of functional reliability for time-critical applications (applications that have to be done within a certain time like business simulations) and real-time simulations (applications that are part of a process and run permanently like weather forecast) to improve the quality of the handling, the design, the deployment and the execution of complex large-scale simulations

HLRS already developed a basic concept for the solution of complex applications in scientific and engineering experiments on distributed heterogeneous resources. Part of this work was done in the European research projects GridCoord and CoreGrid. The concept comprises two main components: a parallel language GriCoL for the description of complex scientific experiments and the Science Experimental Grid Laboratory (SEGL) for the control and execution of complex experiments.

• Natalia Currle-Linde
• Michael Resch

University of Stuttgart, HLRS

DECI – The DEISA Extreme Computing Initiative

DEISA (Distributed European Infrastructure for Supercomputing Applications) is a consortium of leading national Supercomputing Centres in Europe founded with the purpose of advancing computational sciences in the area of high performance computing. The consortium has designed, deployed and is currently operating a complex and heterogeneous high-performance, distributed supercomputing environment on a continental scale with an aggregate peak performance of about one PetaFlop/s. To enhance DEISA's impact on science and technology, the DEISA Extreme Computing Initiative (DECI) was launched in 2005. The DECI consists of the identification, enablement, deployment and operation of "flagship" applications in selected areas of science and technology. These leading, ground breaking applications must deal with complex, demanding, innovative simulations that benefit from the exceptional resources of the Consortium. Projects supported by DECI are chosen on the basis of innovation potential, scientific excellence and relevance criteria. Multinational proposals are especially encouraged.

Applications Enabling

For the advancement of computational sciences in the area of supercomputing, application enabling is of key importance. As a prerequisite to handle and master this challenging task, DEISA created the Applications Task Force (ATASKF), constituted as a team of leading experts in high performance and Grid computing. ATASKF deals with all the different tasks of application enabling for capability computing. Application specialists have enabled and optimized applications for usage in DEISA. Tasks included: scaling of applications, workflows, coupled applications, determination of best suited architecture(s) in DEISA, adaptation of applications to the DEISA infrastructure and architecture dependent optimizations.

Scaling of parallel programs for the efficient usage on systems with thousands of processor-cores is a very challenging task. Design, deployment and optimization of workflow applications allow for the chaining of simulation and pre- and post-processing steps, involving successive operations on large data sets performed on several platforms.

Extreme Computing Projects

DECI has supported and continues to support the most challenging supercomputing projects in Europe which require the special resources and skills available within DEISA. A European Call for Extreme Computing Proposals was published annually in spring. After receiving around 40 and 50 proposals in 2005 and 2006, respectively, more than 60 proposals from the 2007 call competed for DEISA resources in 2008, asking for more than 70 million computing hours.

By selecting the most appropriate supercomputer architectures for each project, DEISA is opening up the most powerful HPC architectures available in Europe for highly challenging projects, mitigating the rapid performance decay of a single national supercomputer within its short lifetime cycle of typically about five years, as implied by Moore's law.

Figure 1: DEISA brochures 2008, available from the German DEISA sites HLRS, JSC, LRZ and RZG

So far scientists from about 160 different universities or research institutes from 15 different European countries with collaborators from four other continents (North and South America, Asia and Australia) have benefited. The scientific projects cover major areas of science including Astro Sciences, Earth Sciences, Engineering, Life Sciences, Materials Science, Plasma Physics, and Quantum Chromodynamics.

Scientific achievements through DEISA – especially through DECI - have been reported in the recently pulbished booklet "DEISA – Advancing Science in Europe". Additionally, the DEISA Digest 2008 provides information about recent developments, targeting a more general audience. Hardcopies are available from the four German DEISA sites HLRS, JSC, LRZ, and RZG, and PDF versions of both publications can be downloaded from http://www.deisa.eu/publications/results

As one of the highlights, results from a German/British group with Principal Investigator Kurt Kremer were published as the cover story of NATURE (Nature, 447, pp. 461-464, 2007). For almost two decades, physicists have put considerable efforts into the investigation of membrane mediated interactions. Simulations in DEISA have now revealed that curvy membranes make proteins attractive. So far more than 120 scientific publications have appeared which were based on DECI activities.

In 2008, the European interest in DECI is persisting. The DECI call 2008, conducted in May and June 2008, resulted in 65 proposals requesting more than 120 Million core hours. The scientific evaluations through the national evaluation committees and the technical evaluations by ATASKF will be conducted during summer, with results and decisions expected for autumn.

To even further enhance its impact on Europe's capability computing, DEISA has now started to extend the service provisioning model from the single project support through the DECI to also supporting Virtual European Communities and European projects in key areas of science.

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

KONWIHR II: Fostering the Efficient Use of High Performance Computers

Bavarian Competence Network for Technical and Scientific High Performance Computing

Computational Science has established itself as the "third pillar" of scientific enquiry alongside theory and experiment. The discipline of Computational Science enhances scientific investigation by building and testing models of complex phenomena, yielding new information, innovation and fresh insight into the research process that is otherwise not obtainable. Many of these problems require outstanding computational resources going far beyond the capabilities of desktop PCs. Prominent and important examples in this context are the simulation of turbulent flows, the computation of quantum mechanical many-body interactions in theoretical physics, the evolution of protein dynamics in life sciences or the propagation of earthquake waves.

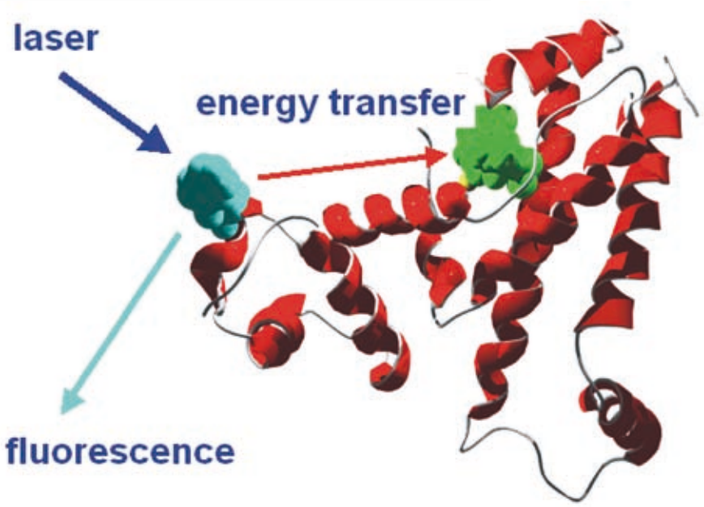


Figure 1: KONWIHR project ENZYMECH (Prof. T. Clark, Computer-Chemie-Zentrum, Universität Erlangen-Nürnberg)

ENZYMECH was concerned with the cheminformatics, modeling and simulation of biological systems. The schematic diagram above shows the principles of fluorescence resonant energy transfer (FRET), which is also known as the "spectroscopic ruler" and is used to measure distances in proteins and other biological systems. FRET was simulated for the first time using a combination of molecular mechanics and quantum mechanical calculations. The results showed that the commonly used interpretation of FRET results can lead to large errors in "measured" distances.

Fostering the use of modern supercomputers within scientific and industrial research communities to solve scientifically challenging and technically highly relevant problems has been the main focus of the Competence Network for Technical and Scientific High Performance Computing in Bavaria (KONWIHR). KONWIHR has been established by the Federal State of Bavaria in May 2000 with a total budget of about four million Euros and has supported 25 to 30 larger scientific projects and numerous short evaluation projects at Bavarian universities in the 2000-2005 timeframe. A special focus of the KONWIHR activities has been on porting, optimization and parallelization of application codes for the "Höchstleistungsrechner in Bayern I (HLRB I)" which was the first federal supercomputer at LRZ Munich and, at the time of installation, the first TFlop/s system in Europe.

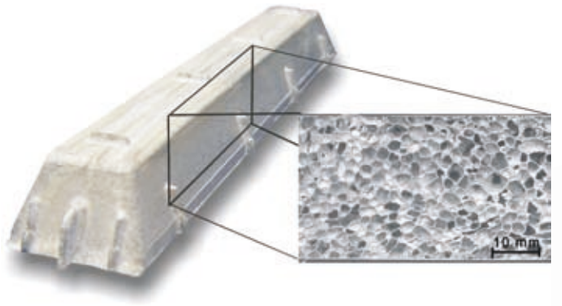
An important aspect of KONWIHR is that it has paved the way to install highly competitive research and support groups at the computing centres of Erlangen (RRZE) and Munich (LRZ). These groups provide extensive and qualified user training as well as high level user support for all challenges in high performance computing, ranging from code debugging and optimization through massive and hybrid parallelization. Of course, this work is fully integrated into the research activities of the scientific groups, requiring a broad level of scientific HPC research at the computing centres themselves. The HPC expertise of both centres also enabled a smooth transition for many Bavarian research groups from HLRB I to HLRB II which was ranked as a TOP 10 computer worldwide in 2007. Complementing the HLRB II, which aims at large scale parallel applications, both centres also acquired large computer clusters in the 2006–2008 timeframe for embarrassing or moderate parallel applications from all over Bavaria.

In 2006 and 2007, the very successful work of KONWIHR has been continued beyond its official lifetime in selected key projects through follow-on funding provided by the Federal State of Bavaria. In summary, all projects reported on substantial progress of their research through KONIWHR funding. Thematic foci of the KONWIHR activities have been multi-physics applications in engineering, molecular dynamics simulations in life sciences, parallel methods for highly correlated quantum systems and evaluation of new programming approaches as well as multi-threaded hardware architectures in computer science. A small collection of scientific areas addressed within KONWIHR is presented in Figures 1 - 3. Beyond the applied and fundamental sciences research activities, KONWIHR projects also provided substantial contributions to HPC related research as indicated by several papers presented at various SC conferences between 2002 and 2007 [1-4].

A new sustainable infrastructure for HPC in Bavaria is currently being launched through KONWIHR-II. A KONWIHR directorate and advisory board have been established (cf. Fig. 4). A first KONWIHR workshop with ten project proposals has been held at RRZE on June 23, 2008. The decision about the final project funding, which will be granted for a two year timeframe with a one year option for prolongation, is expected to be released soon by the Bavarian State Ministry of Sciences, Research and the Arts.

Figure 2: KONWIHR project FreeWIHR
(Prof. Rüde, Lehrstuhl für Systemsimulation, Dr. Körner, Lehrstuhl für Werkstoffkunde und Technologie der Metalle, Universität Erlangen-Nürnberg):

Free surface Lattice-Boltzmann simulations (lower part) allow for computation and modeling of complex technical processes such as the foaming of lightweight aluminum parts (upper part).



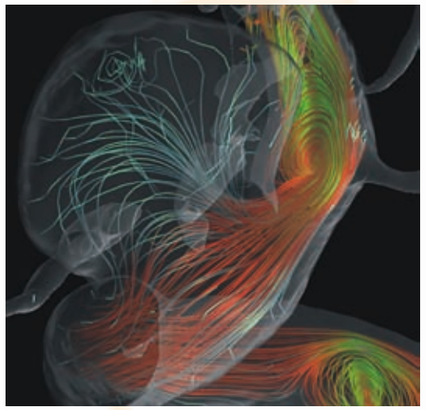


Figure 3: KONWIHR project BESTWIHR
(Prof. Durst, Dr. Beronov, Lehrstuhl für Strömungsmechanik, Universität Erlangen-Nürnberg)

Lattice-Boltzmann (LB) methods as implemented and optimized in BESTWIHR are not limited to turbulence research. In fact they are also capable of simulating flow in complex geometries such as pulsing blood flow in aneurysms. Results of a LB simulation for the flow velocity in an aneurysm and the adjacent blood vessels are shown.

KONWIHR Directorate	KONWIHR Advisory
Ministerialrat Antretter	Prof. Durst, FMP Technologies, Erlangen
Prof. Bode, TU München	Prof. Nagel, TU Dresden
Prof. Hanke, University of Würzburg	Prof. Rannacher, University of Heidelberg
Prof. Urbanek, Ohm-Hochschule Nürnberg	Prof. von der Linden, TU Graz
Dr. Wies, BMW	
Dr. Wellein, RRZE	
Prof. Zenger, BAdW	

Figure 4: Members of the KONWIHR directorate and KONWIHR advisory

The focus of KONWIHR-II (http://www.konwihr.uni-erlangen.de/) will be to further extend the services and competence of the HPC centres of excellence at LRZ and RRZE and to provide funding and know-how to leading edge as well as emerging HPC projects. The goal is to prepare user applications for the technological/disruptive changes expected in the area of Petaflop computing such as massively parallel processing (100,000+ cores) or homogeneous and heterogeneous many-core architectures. These activities will be decisive for the future success of Bavaria in the context of the installation of European-level supercomputing centres and the scientific amortization of the related large investments.

References

[1] **Stamatakis, A. P., Ludwig, T., Meier, H. M., Wolf, J.**
Accelerating parallel Maximum likelihood-based phylogenetic Tree Calculations using Subtree Equality Vectors, Proceedings of the 2002 ACM/IEEE Conference on Supercomputing, 2002, pp. 1-16, CD-ROM, ISBN 0-7803-9802-5

[2] **Pohl, T., Thürey, N., Deserno, F., Rüde, U., Lammers, P., Wellein, G., Zeiser, T.**
Performance Evaluation of parallel Large-scale Lattice Boltzmann Applications on three Supercomputing Architectures, Proceedings of the IEEE/ACM SC 2004 Conference, pp. 1-13, ISBN 0-7695-2153-3

[3] **Bergen, B., Hülsemann, F., Rüde, U.**
Is 1.7×10^{10} Unknowns the largest Finite Element System that can be solved today? Proceedings of the ACM/IEEE SC 2005 Conference, p. 5, ISBN 1-59593-061-2

[4] **Ott, M., Zola, J., Stamatakis, A., Aluru, S.**
Large-scale maximum likelihood-based phylogenetic analysis on the IBM BlueGene/L, Proceedings of the ACM/IEEE Conference on Supercomputing 2007, Article No. 4, ISBN 978-1-59593-764-3

- Arndt Bode[1]
- Gerhard Wellein[2]

[1] Technische Universität München, Lehrstuhl für Rechnertechnik und Rechnerorganisation

[2] Friedrich-Alexander-Universität Erlangen-Nürnberg, Regionales Rechenzentrum Erlangen

The QPACE Project: QCD parallel Computing on the Cell Broadband Engine

The building blocks of the atomic nucleus, protons and neutrons, are known not to be elementary since the 1950s. Rather, they are composed of even smaller building blocks: quarks and gluons. The theory describing the interactions between quarks and gluons is called Quantum Chromodynamics (QCD), which is well established both experimentally and theoretically. However, the study of some of the most important properties of the theory is only possible by numerical simulations, using a discretized formulation of the theory known as Lattice Quantum Chromodynamics (LQCD). While this formulation renders numerical simulations possible, these require a huge amount of computational resources. To carry out such calculations, highly scalable massively parallel computers providing hundreds of TFlops of computing power are required. Scientific progress in this field is limited by the availability of suitable computing resources.

The QPACE project addresses this issue, as its goal is to design and build a novel cost-efficient capability computer that is optimized for LQCD applications. This research area has a long tradition of developing such computers (see, e.g., [1,2]). Previous projects were based on system-on-a-chip designs, but due to the rising costs of custom ASICs the QPACE project is pursuing a different strategy: a powerful commercial multi-core processor is tightly coupled to a custom-designed network processor. The latter is implemented using a modern Field Programmable Gate Array (FPGA), which has several distinct advantages over a custom ASIC: shorter development time and cost, lower risk, and the possibility to modify the hardware design of the network processor even after the machine has been deployed.

The development of QPACE is a common effort of several academic institutions together with the IBM Research and Development Lab in Böblingen (Germany). The academic partners include the Universities of Regensburg and Wuppertal as well as the research labs DESY and Jülich and the Universities of Ferrara and Milano. The project is mainly funded by the Deutsche Forschungsgemeinschaft (DFG) in the framework of SFB/TR-55 and by IBM. First prototype hardware is already available, and the testing of the final hardware configuration is expected to be completed at the end of 2008. In early 2009 we plan to start the manufacturing of several large machines with an aggregate peak performance of 200 Tflops (double precision). The ambitious goal of the project is to make these machines available for research in lattice QCD by the middle of 2009.

The QPACE Architecture

The building block of QPACE is a node-card based on IBM's PowerXCell 8i processor and a Xilinx Virtex-5 FPGA (see Figure 1). The PowerXCell 8i is the second implementation of the Cell Broadband Engine Architecture [3] and is very similar to the Cell processor used in Sony's PlayStation 3. The main reason for using this enhanced Cell processor is its support for high-performance double precision operations with IEEE-compliant rounding. The Cell processor contains one PowerPC Processor Element (PPE) and 8 Synergistic Processor Elements (SPE). Each of the SPEs runs a single thread and has its own 256 kBytes on-chip memory (local store, LS) which is accessible by direct memory access (DMA) or by local load/store operations to/from 128 general-purpose 128-bit registers. An SPE in the PowerXCell 8i processor can execute two instructions per cycle, performing up to 8 single precision (SP) or 4 double precision (DP) floating point (FP) operations. Thus, the total SP or DP peak performance of all 8 SPEs of a single processor is 204.8 GFlops or 102.4 GFlops, respectively (at a clock speed of 3.2 GHz).

The processor has an on-chip memory controller supporting a memory bandwidth of 25.6 GB/s and a configurable I/O interface (Rambus FlexIO) supporting a coherent as well as a non-coherent protocol with a total bidirectional bandwidth of up to 25.6 GB/s. Internally, all units of the processor are connected to the coherent element interconnect bus (EIB) by DMA controllers.

In QPACE the I/O interface is used to interconnect the PowerXCell 8i processor with the network processor (Xilinx V5-LX110T). This is possible because of a special feature of the RocketIO transceivers in the Xilinx Virtex-5 FPGAs. We will be using 2 FlexIO links between the multi-core compute processor and the network processor, with an aggregate bandwidth of 6 Gbytes/second per direction.

The node-cards are connected in a three-dimensional torus with nearest-neighbor connections. The physical layer of the torus network links relies on commercial standards for which well-tested and cheap communication hardware is available. This allows us to move the most timing-critical logics out of the FPGA. Specifically, we are using the 10 Gbit/s transceiver PMC Sierra PM8358 (in XAUI mode). On top of this standard physical layer we have designed a lean custom protocol optimized for low latencies. Unlike in other existing Cell-based parallel machines, in QPACE it will be possible to perform communications directly from the local store (LS) of any SPE on one processor to the LS of any SPE of one of the 6 neighboring processors. The data do not have to be routed through main memory (therefore decreasing the pressure on the performance-critical memory controller) or the PowerPC processor element. Rather, the data are moved via the EIB directly to or from the I/O-interface. The tentative goal is to keep the latency for LS-to-LS copy operations on the order of 1µs.

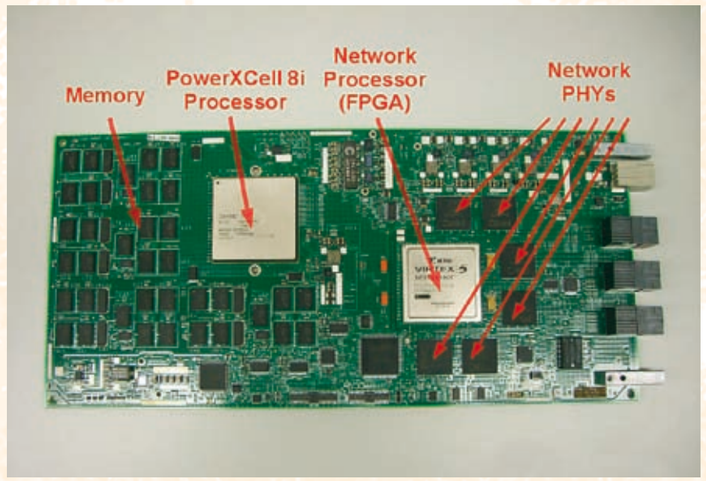


Figure 1: QPACE node-card with a PowerXCell 8i processor, 4 GBytes of main memory, an FPGA and 6 high-speed network transceivers

32 node-cards are mounted on a single backplane. One dimension of the three-dimensional torus network is completely routed within the backplane. The nodes can be arranged as 1 x 4 x 8 nodes or multiple smaller partitions. For larger partitions, several backplanes can be interconnected by cables. 8 backplanes are integrated into a single rack, hosting a total of 256 node-cards with a total peak performance of 26 TFlops (DP). A system consisting of n racks can be operated as a single partition with 2n x 16 x 8 nodes. To obtain smaller partitions without re-cabling we use a special feature of the PMC Sierra PM8358, which provides a redundant link interface. An example for how this feature can be used to partition the machine is shown in Figure 2. The properties of the physical layer of the network have been investigated in detail in a test setup (see Figure 3). In Figure 4 an example for an eye diagram is shown for a lane with maximum distance between transmitter and receiver.

The root-cards are also part of a global signal tree network. Via this network signals and interrupts can be sent by any of the node-cards to the top of the tree. There the signals are reduced, and the result is propagated to all node-cards of a given partition.On each node-card the network processor is also connected to a Gbit-Ethernet transceiver. The Ethernet ports of all node-cards can be connected to standard Ethernet switches that are integrated in the QPACE rack. Depending on the I/O-requirements the Ethernet bandwidth between a QPACE rack and a front-end system can be adjusted by changing the bandwidth of the uplinks of the switches.

Each node-card consumes up to 130 Watts. To remove the generated heat a cost-efficient liquid cooling system is being developed, which enables us to reach high packaging densities. The power consumption of a single QPACE rack is about 35 kWatts. This translates into a power efficiency of about 1.5 Watts/GFlops (DP, peak).

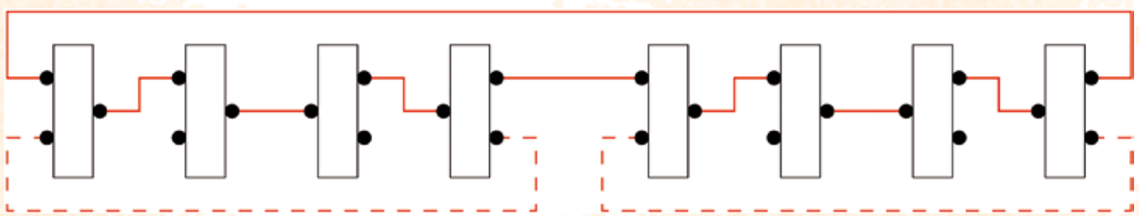


Figure 2: Using redundant links, 8 node-cards can be connected periodically as 1 x 8 or 2 x 4 node-cards.

On each backplane there are 2 root-cards which manage and control 16 node-cards each (e.g., when booting the machine). The root-card hosts a small Freescale MCF5271 Microprocessor operated using uClinux [4]. The microprocessor can be accessed via Ethernet, and from it one can connect to various devices on the node-cards via serial links (e.g., UART).

Application Software and Performance

During an early phase of this project a performance analysis has been done based on simple models which typically only take bandwidth and throughput parameters of the hardware into account [5]. The overall performance of LQCD applications strongly depends on how efficient the execution of just one basic step can be implemented, namely the product of a large but sparse matrix (the so-called Lattice Dirac Operator) and a vector (a quark field). For one particular version of this matrix we estimated for realistic parameters a theoretical efficiency of about 30%. The main restrictions come from the performance of the memory controller. In this case a sophisticated strategy had been assumed for reading data from memory and storing results back to memory, such that external memory accesses are minimized. In a real implementation of this application kernel it has been demonstrated that on a single processor an efficiency of 25% can be achieved [6].

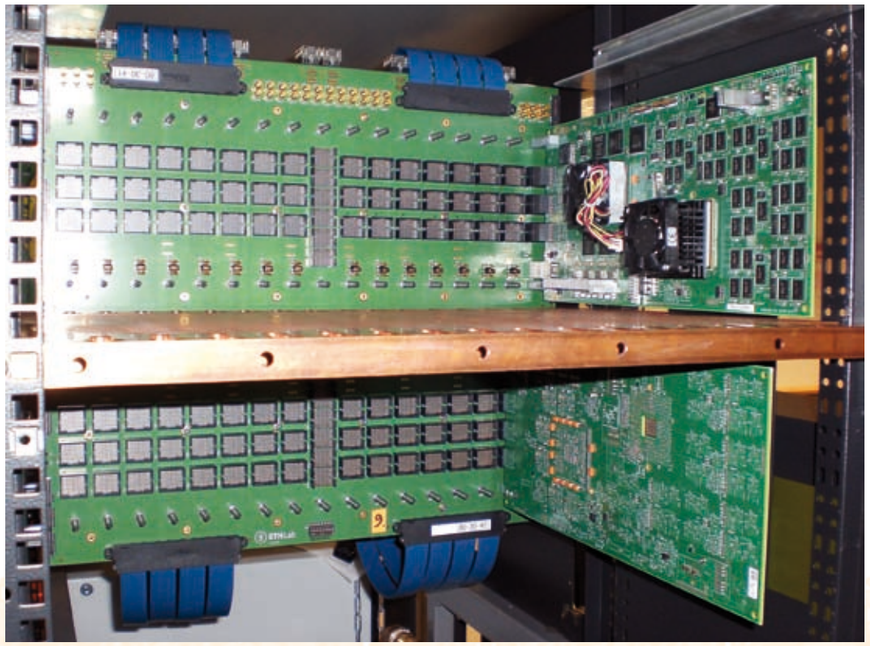


Figure 3: QPACE backplane together with 2 prototype node-cards. The copper plate in the middle is liquid cooled. In the final system the node-cards will be mounted in a housing which is directly attached to this cold plate. Also visible are the (blue) communication cables.

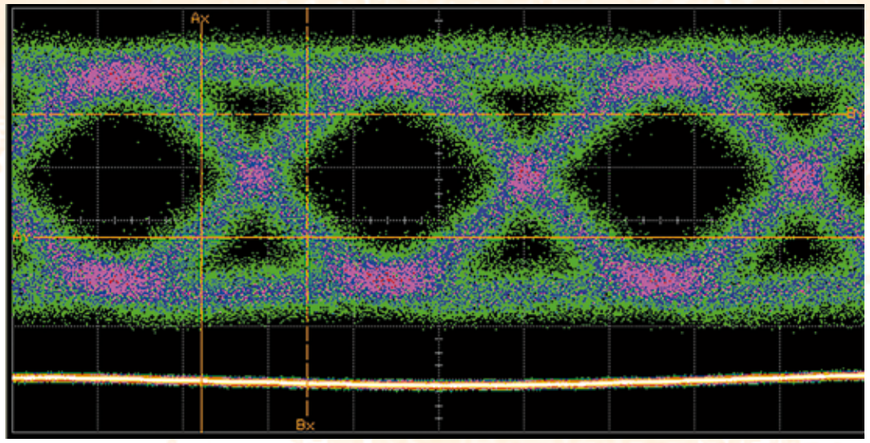


Figure 4: Eye diagram measured for a signal which is routed through about 50 cm board material and 50 cm cable. In case of QPACE this is the maximum possible distance. During this measurement the link was running at 3.125 GHz.

Efficient implementation of applications on the Cell processor is obviously more difficult compared to more standard processors. One has to carefully choose the data layout to maximize utilization of the memory interface. The overall performance of the program furthermore critically depends on the utilization of the on-chip parallelizm.

To relieve the programmer from the burden of porting efforts we apply two strategies. For a number of kernel routines which are particularly performance relevant we will provide highly optimized implementations which can be accessed through library calls. To facilitate the implementation of the remaining parts of the code we plan to port or implement software layers that hide the hardware details.

Summary

QPACE is a next-generation massively parallel computer optimized for LQCD applications. It leverages the power of modern multi-core processors by tightly coupling them within a custom high-bandwidth, low-latency network. The system is not only optimized with respect to procurement costs vs. performance but also in terms of power consumption, i.e., operating costs. The machines that will become available in 2009 will significantly increase the computing resources available for LQCD calculations in Germany.

Acknowledgements

We would like to thank all members of the QPACE development team at the academic sites and at the IBM labs in Böblingen, La Gaude and Rochester for their hard work making this endeavor possible. We also acknowledge the following companies which contribute to the project by various means: Eurotech (Italy), Knürr (Germany), Rambus (US), Xilinx (US), Zollner (Germany).

References

[1] **Belletti, F. et al.**
Computing for LQCD: apeNEXT, Computing in Science & Engineering, 8, p. 18, 2006

[2] **Boyle, P.A. et al.**
Overview of the QCDSP and QCDOC computers, IBM Journal of Research and Developement, 49, 351, 2005

[3] **IBM**
Cell Broadband Architecture, Version 1.0, 2005

[4] **UCLINUX**
http://www.uclinux.org

[5] **Goldrian, G. et al.**
QPACE: QCD parallel Computing on the Cell Broadband Engine, Computing in Science and Engineering, 2008 (accepted for publication)

[6] **Nobile, A.**
PhD thesis, Milano University, 2008

• Dirk Pleiter[1]
• Tilo Wettig[2]

[1] DESY, Zeuthen Site

[2] University of Regensburg, Departement of Physics

New Cell-based HPC Cluster System for GRS

On 1 July 2008, the German Research School for Simulation Sciences (GRS) received a high-performance computer system which is based on the latest implementation of the Cell Broadband Engine (Cell/B.E.) architecture. This IBM system with 35 compute nodes, 8 GB memory per node, InfiniBand interconnect, and a total computational capability of 7 TFlop/s was deployed in Jülich and will be operated by the Jülich Supercomputing Centre.

The compute cluster is based on IBM BladeCenter QS22 nodes, which incorporate two PowerXCell 8i processors. This type of processor contains a set of special-purpose processing cores which can be used as computational accelerators to augment the main PowerPC processor resulting in a peak double precision floating point performance of 108.8~GFlop/s. The PowerXCell 8i processor represents recent trends in processor architecture related to multi-core processors with specialized functions.

Early experiences with cell processors have already been gained in Jülich since the beginning of 2007, when the project JUICE (JUelich Initiative CEll cluster) was established. Cell-specific implementations of applications, e.g. based on the Lanczos Eigenvalue Algorithm and the multigrid method, were developed and documented in a technical report. The new cluster – called JUICEnext – will help to investigate new programming methodologies, models, and techniques. Existing compute-intensive simulation codes will be adapted to exploit the computational accelerators of the PowerXCell 8i thus gaining significant performance improvements.

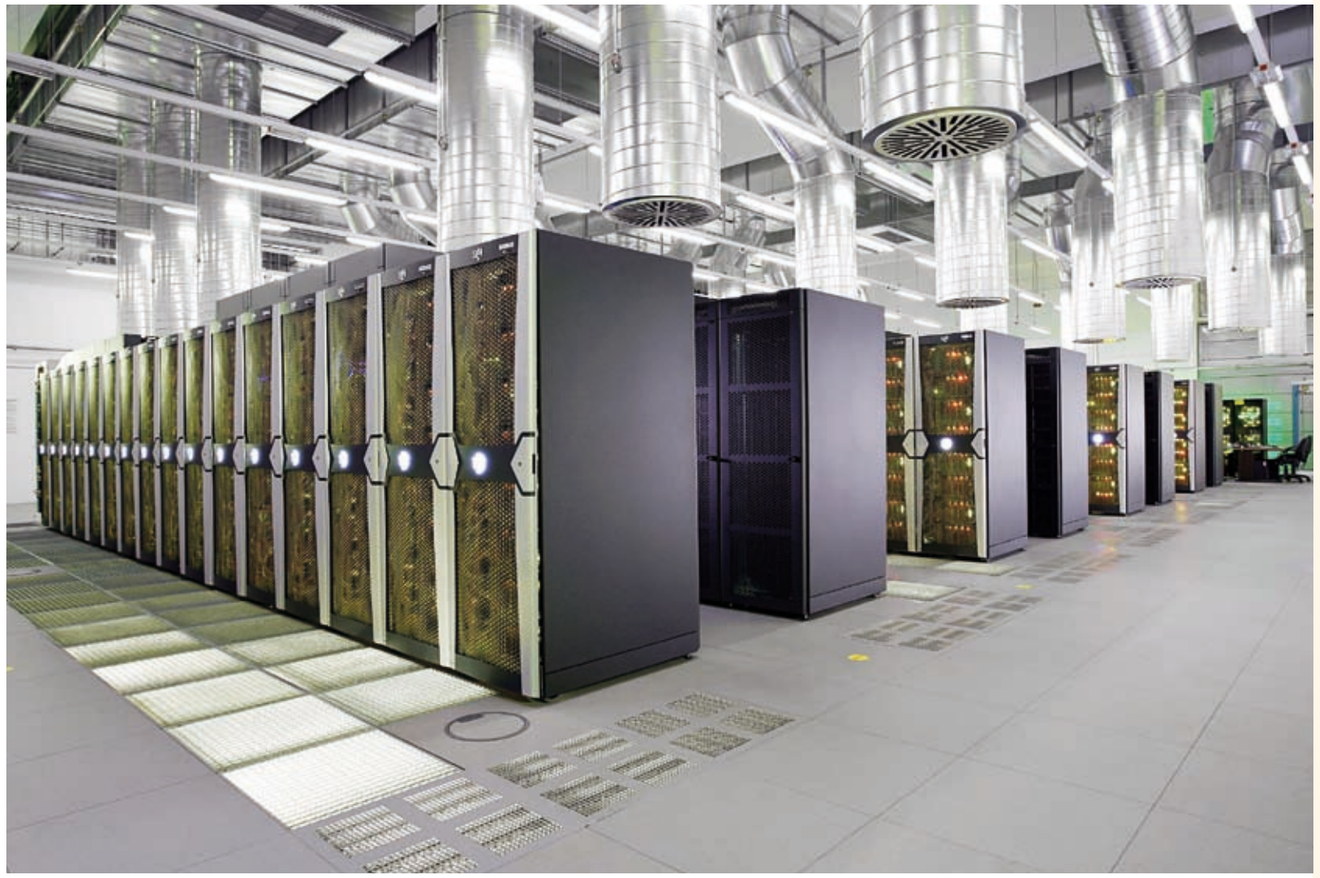
For further information see:
http://www.fz-juelich.de/jsc/juice

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences (Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, LRZ) in Garching near Munich provides national, regional and local HPC services.

Each platform described below is documented on the LRZ WWW server; please choose the appropriate link from www.lrz.de/services/compute

Contact
Leibniz-Rechenzentrum

Dr. Horst-Dieter Steinhöfer
Boltzmannstraße 1
85748 Garching/München
Germany
Phone +49-89-3 58 31-87 79



View of „Höchstleistungsrechner in Bayern HLRB II", an SGI Altix 4700
Foto: 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	62.2	Capability Computing	German Universities and Research Institutes, DEISA
Linux-Cluster SGI Altix 4700 Intel IA64 256 way	256 Cores 1 TByte	1.6	Capability Computing	Bavarian Universities
Linux-Cluster SGI Altix 3700 BX2 Intel IA64 128-way	128 Cores 0.5 TByte	0.8	Capability Computing	Bavarian Universities
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- and 16-way	1,602 cores 4.3 Tbyte	9.1	Capacity Computing	Bavarian and Munich Universities, D-Grid
Linux-Cluster Intel Xeon EM64T AMD Opteron 2-, 4-way	912 cores 1.4 TByte	6.6	Capacity Computing	LHC Computing Grid Tier 2

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. Michael M. Resch
Nobelstraße 19
70569 Stuttgart
Germany
Phone +49-711-685-872 69
resch@hlrs.de
www.hlrs.de

View of the NEC SX-8 at HLRS

Compute servers currently operated by HLRS are

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
NEC SX-8	72 8-way nodes 9.22 TByte memory	12.67	Capability Computing	German Universities, Research Institutes and Industry
TX-7	32-way node 256 GByte memory	0.19	Preprocessing for SX-8	German Universities, Research Institutes and Industry
IBM BW-Grid	498 2-way nodes 8 TByte memory	45.9	Grid Computing	D-Grid Community
NEC SX-9	8 16-way nodes 4 TByte memory	12.8	Grid Computing	D-Grid Community
Intel Nocona	205 2-way nodes 0.62 TByte memory	2.5	Capacity Computing	Research Institutes and Industry
AMD Opteron	230 2-way nodes 1,63 TByte memory	3.6	Capacity Computing	Research Institutes and Industry
Cray XD1	48 2-way nodes 0.24 TByte memory	0.54	Industrial Development	Industry and Research Institutes

A detailed description can be found on HLRS's web pages: www.hlrs.de/hw-access

The John von Neumann Institute for Computing (NIC) is a joint foundation of Forschungszentrum Jülich, Deutsches Elektronen-Synchrotron DESY, and Gesellschaft für Schwerionenforschung GSI to support supercomputer-aided scientific research and development. Its tasks are:

Provision of supercomputer capacity for projects in science, research and industry in the fields of modeling and computer simulation including their methods. The supercomputers with the required information technology infrastructure (software, data storage, networks) are operated by the Jülich Supercomputing Centre (JSC) and by the Centre for Parallel Computing at DESY in Zeuthen.

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

At present, a research group Elementary Particle Physics, headed by Zoltan Fodor with staff members in Zeuthen and Jülich exists; a successor of the group Computational Biology and Biophysics in Jülich is in preparation.

Education and training in the fields of supercomputing by symposia, workshops, school, seminars, courses, and guest programmes.

Contact
John von Neumann Institute
for Computing (NIC)
Jülich Supercomputing Centre (JSC)
Institute for Advanced Simulation (IAS)

Prof. Dr. Dr. Thomas Lippert
52425 Jülich
Germany
Phone +49-24 61-61-64 02
th.lippert@fz-juelich.de

www.fz-juelich.de/nic
www.fz-juelich.de/jsc

The IBM supercomputer "JUGENE" in Jülich (Photo: Research Centre Jülich)

Compute servers currently operated by JSC in Jülich or DESY in Zeuthen are

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM Blue Gene/P "JUGENE"	16 racks 16,384 nodes 65,536 processors PowerPC 450 32 TByte memory	222.8	Capability computing	German Universities, Research Institutes and Industry
IBM Power6 575 "JUMP"	14 SMP nodes 448 processors POWER6 1.8 TByte memory	8.4	Capability computing	German Universities, Research Institutes and Industry
IBM Cell System "JUICEnext"	35 Blades 70 PowerXCell 8i processors 280 GByte memory	7	Capability computing	German Research School
IBM BladeCentre-H "JULI"	2 racks 56 Blades 224 PowerPC 970 MP cores 224 GByte memory	2.2	Capability computing	Selected NIC Projects
AMD Linux Cluster "SoftComp"	125 compute nodes 500 AMD Opteron 2.0 GHz cores 708 GByte memory	2.5	Capability computing	EU SoftComp Community
AMD Linux Cluster "JUGGLE"	44 compute nodes 176 AMD Opteron 2.4 GHz cores 352 GByte memory	0.85	Capacity and capability computing	Selected D-Grid Projects
apeNEXT (special purpose computer)	4 racks 2,048 processors 512 GByte memory	2.5	Capability computing	Lattice gauge theory Groups at Universities and Research Institutes
APEmille (special purpose computer)	4 racks 1,024 processors 32 GByte memory	0.55	Capability computing	Lattice gauge theory Groups at Universities and Research Institutes

Continuing the Standardization Effort: Fortran 2008

In March 2008, the technical subcommittee responsible for the ISO standardization of the Fortran language [1] has released a draft [2] for a new standard, named Fortran 2008 and scheduled to be officially released in 2010 provided it passes the ISO voting procedure. Although originally envisioned to be comprised of minor changes only, with focus on fixing defects identified in the Fortran 2003 standard and only a few features added, one very large and controversial change has also been integrated: Coarrays for parallel programming. The following sections describe the two largest features proposed for Fortran 2008.

Submodules

The module concept introduced by Fortran 90 allows for modularization of code; however for implementation of very large concepts under circumvention of monstrously large modules there are disadvantages: Distribution of a design into multiple modules often requires violation of encapsulation properties; the dependencies between the modules typically lead to re-compilation cascades even when only implementations – i.e., no type definitions, global objects or interfaces – are modified. For this reason, the submodule concept was developed first as a separate technical report (TR19767), and now integrated into the Fortran 2008 draft together with corrections to defects identified in the meantime. Submodules allow separation of procedure implementations from their interface definitions ("separate module procedures"); each submodule accesses the entities of its single ancestor module or submodule by host association, which in particular allows access also to private entities defined in the ancestor. Changes which only impact the implementation require re-compilation of the submodule itself and its sibling submodules, but not of the files accessing the ancestor module by use association. Finally, it is now possible to set up module references within submodules which formerly would have led to – disallowed – circular dependencies between modules.

Coarrays

With the change of focus of processor architecture away from a continuing increase in scalar processing power toward explicit parallelism, as exemplified by the increasing number of processor cores in standard PCs, all computer languages must face the challenge of providing integrated parallel facilities in their semantics. For Fortran, the partitioned global address space (PGAS) approach was chosen, and the well-known coarray extension – the first version of which was developed by Bob Numrich and John Reid in 1997 [3]– will be available for implementation of SPMD-like parallel algorithms. Under the control of the run time environment, a user-specified number of images of a single executable run asynchronously on a computing system. In the Fortran source, the programmer has the possibility to declare objects as coarrays

```
type(mytype) :: x(ndim)[*]
```

characterized by the coindex in angular braces, thereby enabling access to data defined on one image from another image

```
! pull data from image 4 to image 3

if (this_image() == 3) then

y(1:size(x)) = x(:)[4]

end if
```

(this also demonstrates one of the new intrinsics for parallel execution control). It is the programmer's duty to guarantee the correct sequence of events in his parallel algorithm by inserting image control statements, e.g.

```
sync images(4)
```

at the right places in his code.

Compared to the library approach (e.g. MPI) the development, coding and debugging effort should be considerably reduced, because the type system is fully integrated with the coarray functionality. As opposed to most instances of shared memory programming with directives, memory locality is fully under the programmer's control, which allows to more easily avoid clogging of system interconnects especially on NUMA architectures. The one-sided communication semantics allows an optimized implementation to efficiently utilize the potential of the interconnect hardware, and allows the programmer to interleave computation and communication. It is expected that the coarray feature will be included in at least some commercial compilers before the Fortran 2008 standard is officially published; an open-source implementation is available even now [4].

Other Features

An overview of all new Fortran 2008 features is provided in an article by John Reid [5].

Notes and References

[1] The main development body is the J3 US standards committee (http://www.j3-fortran.org/); development is done in collaboration with WG5 of the International Standards Organization (http://www.nag.co.uk/sc22wg5/).

[2] http://j3-fortran.org/doc/standing/links/007.pdf

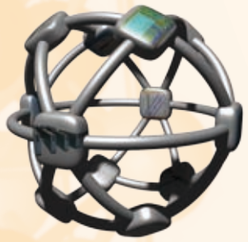
[3] See the coarray overview by John Reid, ftp://ftp.nag.co.uk/sc22wg5/N1701-N1750/N1724.pdf and references therein

[4] The G95 Fortran compiler (http://www.g95.org) partially supports the coarray syntax. A commercial add-on which supports running multiple images is available on LRZ's HPC systems. See http://www.lrz-muenchen.de/services/software/parallel/pgas/ for details.

[5] This is available via ftp://ftp.nag.co.uk/sc22wg5/N1701-N1750/N1735.pdf

Author

Dr. Reinhold Bader works in the High Performance Computing Group at the LRZ; since autumn 2007 he has been member of the German Delegation to WG5.



The 2nd Parallel Tools Workshop at HLRS

With the introduction of multi-core processors, programmers are facing the problem to learn new parallel programming paradigms, new parallelization techniques, most probably run into buglets due to concurrent execution.

The 2nd Parallel Tools Workshop held at HLRS on July 7th and 8th built on the experience of the successful workshop last year. This year, we were able to attract 85 participants, mainly from German and Austrian Universities and research institutions; but also guests and speakers from France, the Netherlands, the UK, Spain, Poland and the US.

As last year, a wide variety of programming tools, ranging from single-core, sequential-code analysis tools, as well as parallel, highly-scalable tools for HPC-environments were presented. Sessions were organized for IDEs and Libraries for Parallel programming, Sequential and Parallel Performance Analysis and a session on Parallel Debuggers.

The main aim of the workshop is to bring the actual users of the tools to the developers; having interaction and exchanging ideas between tool developers and most importantly for users getting hands-on experience with the tool being presented. Therefore we enjoyed live-demonstrations from scalable parallel debuggers such as DDT presented by Mark O'Connor, Allinea getting to know the new memory replay features of Totalview by Chris Gottbrath, Totalview Technologies. Greg Watson, IBM and Thomas Köckerbauer, GUP were demonstrating the capabilities and versatility of the Eclipse IDE with the Parallel Tools Platform (Eclipse/PTP) and the g-Eclipse workbench framework, respectively. The Open MPI project and Sun's contribution were presented by Len Wisniewski, Sun, Open MPI's internal memory debugging features were demonstrated by Shiqing Fan, HLRS. Another talk and hands-on-session demonstrated new Marmot's MPI-checking capabilities, presented by Valentin Himmler, HLRS.

The talk by Herbert Cornelius, Intel presented the software toolchain available by Intel. CPAT, Cray's Performance Analysis Tools were demonstrated by Luiz DeRose (see picture) from Cray. Innovative methods to quickly gather the loss of performance were being shown with a demo on Acumem's SlowSpotter, presented by Magnus Vesterlund and another talk with hands-on-session on Valgrind's tool Callgrind and its GUI KCachegrind by Josef Weidendorfer, TUM.

Finally a wide-range of stand-alone parallel performance analysis applications was selected. We were very glad to welcome Allen Malony, University of Oregon, giving a presentation of the TAU toolchain, Jesus Labarta, Barcelona Supercomputing Centre presenting the Paraver tool and Brian Wylie from Jülich Supercomputing Centre giving a live-demo showing the Scalasca toolset.

Of course, such a session would not be complete, if the Vampir analysis Toolset were missing. Holger Brunst from Technical University of Dresden was providing examples of very large traces analysed using VampirServer.

As such a variety of programs is overwhelming the audience, this year we provided a Springer-publication, covering the tools, showing strengths and particular features, including scientific content from projects such as ParMA.

Additionally, to allow self-studying and experimenting after the workshop, we created a Knoppix-based DVD, containing the tools, called Parallel Tools Distribution (PTD). This DVD was used in the Workshop as Live-DVD booting a common Linux, offering all the tools, precompiled based on Open MPI. Vendors kindly provided short-term licenses for commercial tools, like the Intel Compiler, Acumem's Slowspotter or DDT. Based on the experience gained with this Live-DVD, we will continue this effort in the continuation of the HPC-Europa Project.

After the success of this year's workshop, we would like to continue this event in the future, possibly including contributions on new programming paradigms and languages, and possibly with early adopters of tools for these programming paradigms.

Workshop: Introduction to Computational Fluid Dynamics

Computational Fluid Dynamics is one of the main research areas with application in the High Performance Computing sector. This is due to the fact that even modern supercomputers still may not provide enough ressources for simulations with a full resolution of all physically relevant phenomena. The course "Introduction to Computational Fluid Dynamics" has been introduced into the HLRS Parallel Programming Workshop in order to support researchers who have such CFD supercomputing applications in mind. The course gives an overview of the state of the art in numerical methods used in today's CFD tools and also adresses current challenges in flow simulations. The emphasis is placed on finite volume methods for the compressible Euler- and Navier-Stokes equations. Included topics are explicit and implicit temporal discretizations, spatial discretization on structured and unstructured grids as well as turbulence modelling. Additionally, numerical methods for the incompressible Navier-Stokes equations, aero-acoustics and high order numerical methods which are expected to be the core of future CFD methods are addressed. The last day is dedicated to the parallelization of explicit and implicit solvers.

The course is usually held in late February/early March at the University of Kassel and in late September/early October at the HLRS in Stuttgart.

The 11th Results and Review Workshop of the HLRS

The 11th Results and Review Workshop of the High Performance Computing Center Stuttgart (HLRS) was held on September 29-30 in Stuttgart.
More than 40 scientists from all over Germany presented their results which they had achieved over the last year on the supercomputers at HLRS and the Karlsruhe HPC systems. After a welcome by director Prof. Michael Resch and an introduction of the head of the steering committee of the HLRS Prof. Wolfgang Nagel lectures and poster sessions were held. The focus was on the work in the fields of computational fluid dynamics, physics and chemistry. As HLRS is preparing for its next supercomputer installation the discussion was also focused on the future requirements of these fields. Intensive discussions were continued during social events.

As every year the steering committee had to find the best three contributions of the workshop combining the quality of the paper and the lecture. These three papers were awarded the traditional golden spike award. The winners of this year are:

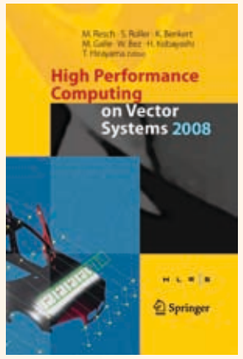
- B. Müller from the MPI for Astrophysics in Garching for the work on Super Nova simulations
- D. von Terzi from the Institute for Hydromechanics of the Karlsruhe Institute of Technology (KIT) for the work on Hybrid Techniques for Large-Eddy Simulations of Complex Turbulent Flows
- T. Zeiser from the Computing Center of the University of Erlangen for the work on vector computers in a world of commodity clusters, massively parallel systems and many-core many-threaded CPUs: recent experience based on an advanced lattice Boltzmann flow

These three papers and all contributions shown at the workshop are published in the HLRS series by Springer (see next page).

Tools for High Performance Computing

With the advent of multi-core processors, implementing parallel programming methods in application development is absolutely necessary in order to achieve good performance. Soon, 8-core and possibly 16-core processors will be available, even for desktop machines. To support application developers in the various tasks involved in this process, several different tools need to be at his or her disposal. This workshop will give the users an overview of the existing tools in the area of integrated development environments for clusters, various parallel debuggers, and new-style performance analysis tools, as well as an update on the state of the art of long-term research tools, which have advanced to an industrial level. The proceedings of the 2nd Parallel Tools Workshop guide participants by providing a technical overview to help them decide upon which tool suits the requirements for the development task at hand. Additionally, through the hands-on sessions, the workshop will enable the user to immediately deploy the tools.

Resch, M., Keller, R., Himmler, V., Krammer, B., Schulz, A. (Eds.)
Tools for High Performance Computing, Proceedings of the 2nd International Workshop on Parallel Tools for High Performance Computing, July 2008, HLRS, Stuttgart 2008, XII, 204 p., 64 illus., 49 in color, Hardcover
ISBN: 978-3-540-68561-6



High Performance Computing on Vector Systems 2008

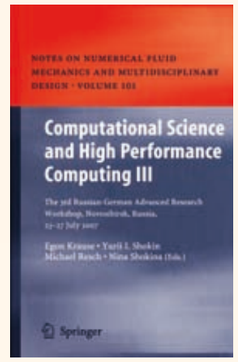
This book presents the state of the art in High Performance Computing and simulation on modern supercomputer architectures. It covers trends in hardware and software development in general and specifically the future of vector-based systems and heterogeneous architectures. The application contributions cover computational fluid dynamics, fluid-structure interaction, physics, chemistry, astrophysics, and climate research. Innovative fields like coupled multi-physics or multi-scale simulations are presented. All papers were chosen from presentations given at the 7th Teraflop Workshop, held in November 2007 at Tohoku University, Japan, the 8th Teraflop Workshop, held in April 2008 at the HLRS, Germany, and at the Japan Atomic Energy Agency in April 2008.

Resch, M.; Roller, S.; Benkert, K.; Galle, M.; Bez, W.; Kobayashi, H.; Hirayama, T. (Eds.)
High Performance Computing on Vector Systems 2008, 2009, Approx. 245 p., Hardcover
ISBN: 978-3-540-85868-3

High Performance Computing in Science and Engineering ‘08

This book presents the state-of-the-art in simulation using supercomputers. Leading researchers present results achieved on systems of the Stuttgart High Performance Computing Center (HLRS) for the year 2008. The reports cover all fields of computational science and engineering, ranging from CFD and computational physics and chemistry to computer science, with a special emphasis on industrially relevant applications. Presenting results for both vector-based and microprocessor-based systems, the book makes it possible to compare the performance levels and usability of various architectures. As the HLRS operates the largest NEC SX-8 vector system in the world, this book gives an excellent insight into the potential of such systems. The book further covers the main methods utilized 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 coloured illustrations and tables.

Nagel, W. E.; Kröner, D. B.; Resch, M. (Eds.)
High Performance Computing in Science and Engineering ‘08, Transactions of the High Performance Computing Center, Stuttgart (HLRS) 2008
Approx. 615 p., Hardcover, Due: November 26, 2008, ISBN: 978-3-540-88301-2



Computational Science and High Performance Computing III

This volume contains 18 contributions to the Third Russian-German Advanced Research Workshop on Computational Science and High Performance Computing presented in July 2007 at Novosibirsk, Russia. The workshop was organized jointly by the High Performance Computing Center Stuttgart (HLRS) and the Institute of Computational Technologies of the Siberian Branch of the Russian Academy of Sciences (ICT SB RAS) The contributions range from computer science, mathematics and high performance computing to applications in mechanical and aerospace engineering. They show a wealth of theoretical work and simulation experience with a potential of bringing together theoretical mathematical modelling and usage of High Performance Computing systems presenting the state of the art of computational technologies.

Krause, E.; Shokin, Y.I.; Resch, M.; Shokina, N. (Eds.)
Computational Science and High Performance Computing III, The 3rd Russian-German Advanced Research Workshop, Novosibirsk, Russia, 23 - 27 July 2007, Series: Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Vol. 10 2008, XVI, 258 p., 121 illus, with online files/ update, Hardcover, ISBN: 978-3-540-69008-5

Transactions of the 3rd HLRB / KONWIHR Status and Result Workshop

The book reports on leading-edge projects on the High Performance Computer in Bavaria (HLRB) and presents state-of-the-art simulations and algorithms in science and engineering. The projects originate from the scientific fields of fluid dynamics, astrophysics and cosmology, computational physics including high energy physics, computational chemistry and materials sciences, geophysics, biosciences, and computer sciences. The articles provide an excellent overview of the broad range of applications that require high performance computing for solving challenging problems. For each project the scientific background is described, along with the results achieved and methodology used. The book also describes the latest advances in high performance applications and reports on the performance and scaling numbers.

Wagner, S., Steinmetz, M., Bode, A., Brehm, M. (Eds.)
High Performance Computing in Science and Engineering, Garching/Munich 2007, Transactions of the Third Joint HLRB and KONWIHR Status and Result Workshop, Springer-Verlag Berlin Heidelberg, 2008, 1st Edition (Sept. 2008), 710 p., Hardcover, ISBN-10: 3540691812

The 5th Russian-German School on Parallel Programming and High Performance Computing

The present event continues the tradition of such Schools in Novosibirsk/Russia, and pursues the success of the previous Schools. The School was initiated first in 2004 as collaboration between the Director of the Institute of Computational Technologies, Academician Y.I. Shokin and the Director of the Stuttgart High Performance Computing Center, Professor M. Resch, and this year was joined by Professor D. Kröner from the University of Freiburg. The School aims to provide High Performance Computing- and Parallelization Techniques to young scientists at the beginning of their career.

Inspirational and highly professional lectures have been given by Harald Klimach (HLRS), Dr. Andreas Dedner (University of Freiburg) and Eugen Volk (HLRS). A presentation of HLRS and DEISA has been done by Harald Klimach. A videoconference between the HLRS and the Institute of Computational Technologies of the Siberian Branch of the Russian Academy of Sciences (ICT SB RAS) has been organized, when Uwe Küster has given a lecture on program optimization for single processor performance. Following a tradition that has already been formed, this School has been organized as two courses, allowing participation of scientists with different levels of knowledge. The basic course has included Parallel Architectures and Programming Models, MPI, OpenMP and Tools for Programming with Multiple Threads. In this year Andreas Dedner from the University of Freiburg has given lectures on programming framework DUNE (Distributed and Unified Numerics Environment).

The advanced course has covered MPI-2 features, debugging and performance analysis for parallel programs, domain decomposition techniques, and single processor optimization. For the first time of this School, Eugen Volk has given lectures and exercises about Grid-Technologies, including GT4 and UNICORE, and presented the current European Grid Project BEinGRID with the business experiment AgroGrid as an example. On the last day of the School the participants had the opportunity to attend an excursion to the newly installed HP-Cluster for the new supercomputing center of the Novosibirsk State University.

The course was attended in each week by nearly 40 students and scientists from Angarsk, Irkutsk, Kemerovo, Krasnoyarsk, Novosibirsk, Tomsk and Almaty (Kazachstan). Another highlight of the School has been a scientific session, where participants have had a possibility to present their own works. The talks have included Parallel Programming, Mathematics, Computational Fluid Dynamics, Molecular Dynamics and Bio-Informatics.

Two books of course material have been prepared by R. Rabenseifner, T. Bönisch, R. Keller, B. Krammer, H. Klimach and E. Volk. The 5th School has been supported by HLRS, Prof. Dr. Kröner of University of Freiburg, ICT SB RAS and its director Prof. Yurii Shokin, Siberian Branch of Russian Academy of Sciences, Russian Foundation for Basic Research.

Winter School

Multiscale Simulation Methods in Molecular Sciences

Jülich will continue its successful series of winter schools. From March 2 to 6, 2009, the winter school "Multiscale Simulation Methods in Molecular Sciences" will take place at the rotunda of the Jülich Supercomputing Centre. More than 20 renowned scientists will present lectures on what can be called "eclecticism in simulation". Eclecticism is a conceptual approach that does not hold rigidly to a single paradigm or set of assumptions, but instead draws upon multiple theories, styles, or ideas to gain complementary insights into a subject, or applies different theories in particular cases.

In particular three strings of themes will be covered focusing on how to deal with hard matter, soft matter, and bio matter when it is necessary to cope with disparate length and time scales. Therein aspects like coarse graining of molecular systems and solids, quantum/classical hybrid methods, embedding and multiple time step techniques, creating reactive potentials, multiscale magnetism, adaptive resolution ideas or hydrodynamic interactions will be discussed in detail.

In addition, another string of lectures will be devoted to the genuine mathematical and the generic algorithmic aspects of multiscale approaches and their parallel implementation on large, multiprocessor platforms including techniques such as multigrid and wavelet transformations. Although this is beyond what can be achieved in a very systematic fashion given the breadth of the topic, introductions into fundamental techniques such as molecular dynamics, Monte Carlo Simulation, and electronic structure (total energy) calculations in the flavour of both wavefunction-based and density functional based methods will be provided.

This winter school is suited for highly motivated PhD students and PostDocs. Applications for participation can be sent in until end of December 2008. Based on the required application documents about 50 participants will be selected by the organizers. Details about the school and the application process can be found at http://www.fz-juelich.de/conference/wsms

German Research School for Simulation Sciences launched

To address the pressing need to train the next generation of computational scientists and engineers, Forschungszentrum Jülich and RWTH Aachen University have established the German Research School for Simulation Sciences. Bringing together the specific strengths of its founders in the fields of science, engineering, and high-end computing in an unprecedented, synergistic manner, the new graduate school provides a unique environment for cutting-edge interdisciplinary research and education in the applications and methods of HPC-based computer simulation in science and engineering. Equipped with new dedicated facilities in Aachen and on the Jülich campus, privileged access to the state-of-the-art computing and visualization resources of Jülich Supercomputing Centre and the Center for Computing and Communication of RWTH Aachen University, plus five new professorships (in addition to the diverse group of existing faculty from Aachen and Jülich contributing to the school's activities), the German Research School for Simulation Sciences offers innovative interdisciplinary graduate training programs. Currently, these include a doctoral program and a non-consecutive master's course in simulation sciences. Both programs are intended for outstanding, extremely well-qualified students. Detailed information, in particular on admission, may be found at www.grs-sim.de

The German Research School for Simulation Sciences is jointly funded by the Federal Republic of Germany (BMBF), the State of North Rhine-Westphalia (MIWFT), the Helmholtz Association, Forschungszentrum Jülich, and RWTH Aachen University.

German Research School
for Simulation Sciences

Jülich Blue Gene/P Porting and Scaling Workshop 2008

From 22 to 24 April 2008, JSC organized the first Jülich Blue Gene/P Porting, Tuning, and Scaling Workshop to provide first-hand knowledge on how to port applications to the Blue Gene/P system, debug them on large numbers of nodes, and how to scale existing Blue Gene codes to the new full 16-rack system JUGENE of JSC. For this purpose, JSC provided CPU time in the order of one Blue Gene/P rack month for the 43 participants from France, Germany, Ireland, Spain, Sweden, Switzerland, the UK and USA.

Six advisors from the JSC parallel support team, as well as three experts from IBM and one debugging expert from Totalview Technologies, supported the participants with these tasks. On the first morning, IBM speakers introduced the Blue Gene system and architecture and provided practical tips and tricks on how to port and tune codes to the machine.

The other two days of the workshop were dedicated to hands-on sessions where the participants tried to port and tune their application codes, which they brought to the workshop, with the help of the JSC, IBM, and Totalview experts. Almost all the participants succeeded in scaling their code to at least one Blue Gene/P rack (4,096 cores). A few participants were able to make use of four, eight, or even the full 16 rack (65,536 cores) system. Participants brought Fortran, C, and C++ codes parallelized with MPI and hybrid MPI/OpenMP from all kinds of application areas (DFT, electronic transport and structure, CFD, neuron simulation, discrete dipole approximation, molecular simulations, convective turbulence simulation, quantum chemistry, finite element tearing, astrophysics, exact diagonalization, and others).

For example, Jens Henrik Göbbert from the Institute for Technical Combustion of the RWTH Aachen University was able to investigate the scaling behavior of the 3D fast Fourier transform on BlueGene/P which is the most time-consuming portion of his DNS code which simulates flame propagation of premixed combustion with simplified chemistry. His experiments showed a speedup of 25,850 on 32,768 cores. Developers of the RACOON framework for time-dependant mesh-adaptive computations on structured grids, headed jointly by Jürgen Dreher and Rainer Grauer of Ruhr-University Bochum studied the scalability of their codes for various problems. Weak scaling of a compressible Euler- and MHD-turbulence problem was perfectly linear up to 16,384 cores. Also, to the surprise of the team, RACOON's multigrid, which does not scale on Linux clusters, showed also good scaling on the BLueGene/P for 4,096 cores.

At the same time as the practical experiments, more detailed tutorials on the parallel debugger Totalview and the parallel scalable performance analysis tool Scalasca, developed by JSC, were presented and prompted the participants to experiment with these tools. During the workshop, Totalview succeeded in attaching to 32,768 processes in one experiment. Participants effectively used Scalasca to make performance measurements of their application on up to 8,192 cores. For more information including the presentation slides see:

http://www.fz-juelich.de/jsc/bg-ws08

High Performance Computing Courses and Tutorials

Parallel Programming with MPI, OpenMP and PETSc

Date & Location
November 26-28, 2008
JSC/NIC, Research Centre Jülich

February 16-19, 2009
ZIH, Dresden

Contents
The focus is on MPI, OpenMP, and PETSc. Hands-on sessions (C/Fortran) will allow to immediately test and understand the basic constructs of the presented programming models. Course language is English.
Lecturer: R. Rabenseifner (HLRS)

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

Introduction to modern Fortran and associated Tools

Date & Location
January 29 - February 6, 2009
LRZ Building, Garching
A video conference to RRZE will be organized if there is sufficient interest.

Contents
This course is targeted at audiences with two levels of expertise: Participants with basic programming experience in C, C++ or Fortran should attend the first part of the course, while for participants with extensive Fortran 95 knowledge it will be sufficient to take part in the lectures covering the advanced Fortran features, according to individual interest. The participants of the course have the opportunity to experiment with the lecture materials in hands-on sessions.

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

Iterative Linear Solvers and Parallelization

Date & Location
March 2-6, 2009
HLRS, Stuttgart

Contents
The focus is on iterative and parallel solvers, the parallel programming models MPI and OpenMP, and the parallel middleware PETSc. Different modern Krylov Subspace Methods (CG, GMRES, BiCGSTAB ...) as well as highly efficient preconditioning and multigrid techniques are presented in the context of real life applications. Hands-on sessions (C/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 University of Kassel, HLRS, IAG, and LRZ.

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

Parallel Programming of High Performance Systems

Date & Location
March 2-8, 2009
LRZ Building, Garching

RRZE Building, Erlangen
(Video Conference)

Contents
This course 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. Prerequisite for this course is good knowledge and practice in programming with at least one of the standard HPC languages: Fortran 95 (or higher), C or C++.

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

Introduction to Computational Fluid Dynamics

Date & Location
March 9-13, 2009
University of Kassel

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

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

Fortran for Scientific Computing

Date & Location
March 23-27, 2009
HLRS, Stuttgart

Contents
This course is dedicated for scientists and students to learn (sequential) programming of 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

NEC SX-8 Usage and Programming

Date & Location
March 30-31, 2009
HLRS, Stuttgart

Contents
The course is focused on vectorizing and parallelizing on NEC SX-8 and SX-9.

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

3rd HLRS Parallel Tools Workshop

Date & Location
July 6-7, 2009 – HLRS, Stuttgart

Contents
This workshop offers to the industrial and scientific user community, as well as the tools developers itself an in-depth workshop on the state-of-the-art of parallel programming tools, ranging from debugging tools, performance analysis and best practices in integrated developing environments for parallel platforms. Participants and tools developers itself will get the chance to see the strengths of the various tools. Therefore, this workshop is focused on persons who already know about parallel programming. Hands-on sessions will give a first touch and allow to test the features of the tools.

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

Education in Scientific Computing

Date & Location
August 3 - October 9, 2009
NIC/JSC, Research Centre Jülich

Contents
Guest Students' Programme "Scientific Computing" to support education and training in the fields of supercomputing. Application deadline is April 30, 2009.

Webpage
http://www.fz-juelich.de/jsc/gaststudenten

CECAM Tutorial Programming Parallel Computers

Date & Location
2009 (tentative)
JSC/NIC, Research Centre Jülich

Contents
This tutorial provides a thorough introduction to scientific parallel programming. It covers parallel programming of distributed and hybrid memory model architectures, as well as parallel I/O. Lectures will alternate with hands-on exercises.

Webpage
http://www.cecam.org/tutorials.html

inSiDE

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

Publishers

Prof. Dr. Heinz-Gerd Hegering, LRZ
Prof. Dr. Dr. Thomas Lippert, NIC
Prof. Dr. Michael M. Resch, HLRS

Editor

F. Rainer Klank, HLRS klank@hlrs.de

Design

Stefanie Pichlmayer pichlmayer@hlrs.de

Authors

Reinhold Bader reinhold.bader@lrz.de
Arndt Bode bode@in.tum.de
Lucio Colombi Ciacchi lucio@izbs.uni-karlsruhe.de
Natalia Currle-Linde linde@hlrs.de
Rainer Friedrich r.friedrich@lrz.tu-muenchen.de
Somnath Ghosh somnath.ghosh@aer.mw.tum.de
Stefan Gottlöber sgottloeber@aip.de
Ferdinand Jamitzky ferdinand.jamitzky@lrz.de
Harald Klimach klimach@hlrs.de
Hermann Lederer lederer@rzg.mpg.de
Ulrich Maas maas@itt.mach.uni-karlsruhe.de
Dominik Marx dominik.marx@theochem.ruhr-uni-bochum.de
Nisanth N. Nair nisanth.nair@theochem.ruhr-uni-bochum.de
Dirk Pleiter dirk.pleiter@desy.de
Michael Resch resch@hlrs.de
Sabine Roller roller@hlrs.de
Helmut Satzger helmut.satzger@lrz.de
Ilya Saverchenko ilya.saverchenko@lrz.de
Christoph Schaupp christoph.schaupp@aer.mw.tum.de
Joachim Schenk schenk@mmk.ei.tum.de
Eduard Schreiner eschrein@ks.uiuc.edu
Rainer Stauch stauch@itt.mach.uni-karlsruhe.de
Gerhard Wellein wellein@rrze.uni-erlangen.de
Tilo Wettig tilo.wettig@physik.uni-regensburg.de
Gustavo Yepes gustavo.yepes@uam.es
Janina Zimmermann janina.zimmermann@iwm.fraunhofer.de

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