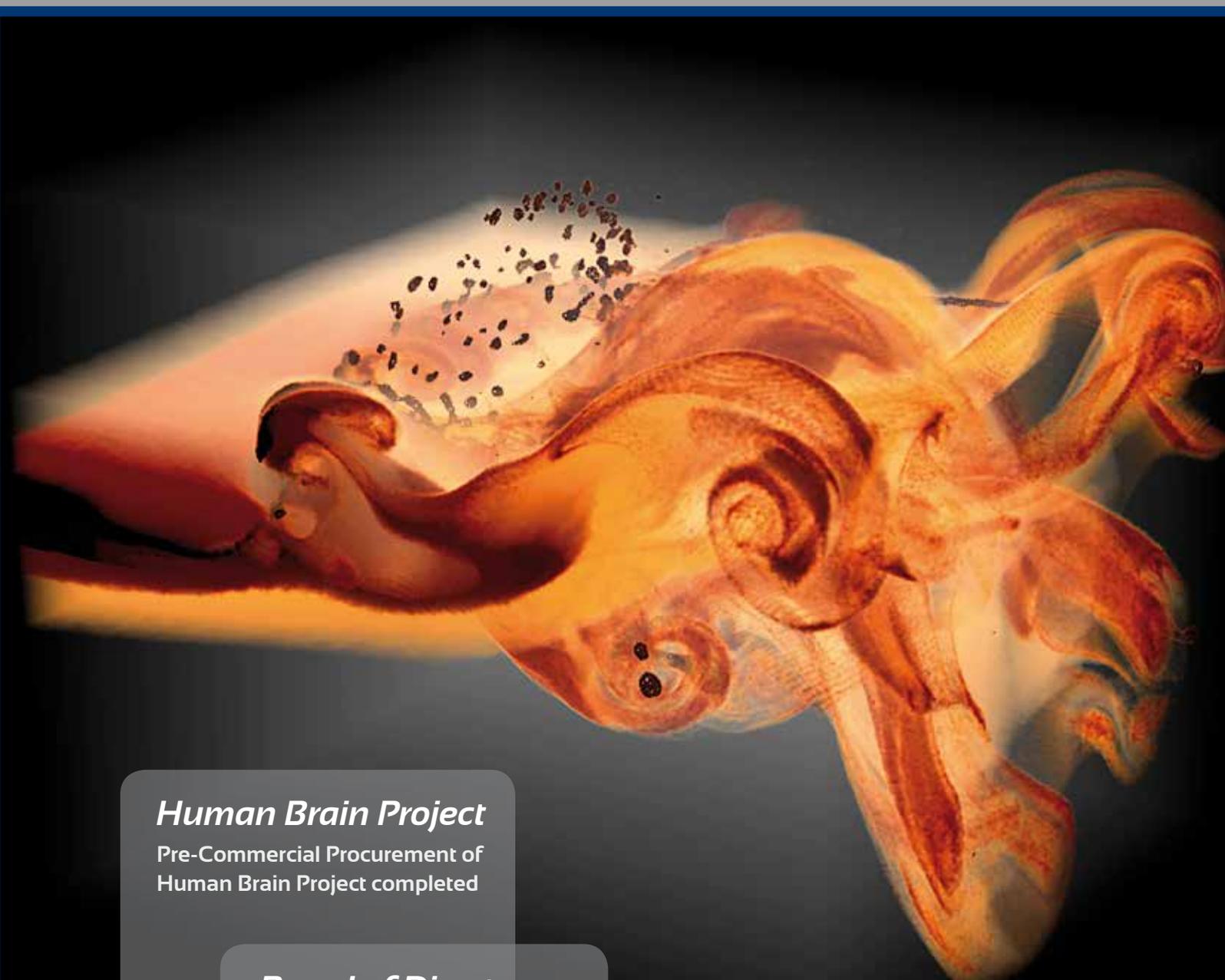


Smoothed Particle Hydrodynamics *for Numerical Predictions of Primary Atomization*



Human Brain Project

Pre-Commercial Procurement of
Human Brain Project completed

Board of Directors

New Chairman of the Board of
Directors at LRZ

Imprint



Prof. Dr. D. Kranzlmüller
Publisher of InSiDE,
Director of LRZ



Prof. Dr. Dr. Th. Lippert
Publisher of InSiDE,
Director of JSC



**Prof. Dr.-Ing. Dr. h.c. mult.
Prof. E. h. M. M. Resch**
Publisher & Editor in chief of InSiDE,
Director of HLRS

© HLRS 2017

Editor

F. Rainer Klank, HLRS
klank@hls.de

Design

Zimmermann Visuelle Kommunikation
www.zimmermann-online.info

Printing

DCC Document Competence Center
Siegmar Kästl e.K., Ostfildern

Cover picture

Samuel Braun
Institut für Thermische Strömungsmaschinen, Karlsruher Institut für Technologie

If you would like to receive InSiDE regularly, please send an e-mail with your postal address to F. Rainer Klank: klank@hls.de

Visit InSiDE online at: <http://inside.hls.de>

InSiDE

InSiDE is published two times a year by the
GAUSS Center for Supercomputing (HLRS, LRZ, JSC)

GCS
Gauss Centre for Supercomputing

Editorial

Welcome to this new issue of InSiDE, the journal on Innovative Supercomputing in Germany published by the Gauss Centre for Supercomputing (GCS). In this issue we present the new Director of LRZ at Garching Prof. Dieter Kranzlmüller. Prof. Kranzlmüller took over from Prof. Bode and also became a member of the GCS board. Prof. Michael Resch took over as chairman of the GCS Board of Directors. He follows Prof. Thomas Lippert. With these changes we celebrate 10 years of GCS and are ready for the coming challenges in HPC in Germany and Europe.

We present the use of GCS supercomputers over the last months through a variety of applications. This variety shows that GCS systems are able to cover a wide spectrum and basically make supercomputing available for all fields that rely on compute performance to break new barriers and achieve new results. Inside, we present exciting results for smoothed particle hydrodynamics for numerical predictions of primary atomization.

The project section, on the other hand, shows how much GCS and its centers are enmeshed in the European and German scientific community when it comes to research in the field of high-performance computing. We find a variety of project reports that range from basic research to applied research to industrial development. This issue especially covers the Human Brain Project and the results of its Pre-Commercial Procurement.

As usual, this issue includes information about recent events related to supercomputing in Germany over the last months and gives an outlook of future workshops in the field. Readers are invited to participate in these workshops which are now part of the PRACE Advanced Training Center and hence open to all European researchers.

Prof. Dieter Kranzlmüller
Prof. Thomas Lippert
Prof. Michael Resch

News Events

- 6 GCS: Delivering 10 Years of Integrated HPC Excellence for Germany
- 13 New Chairman of the Board of Directors at LRZ
- 14 Changes in the GCS Board of Directors
- 16 GCS Continues as Hosting Member of PRACE 2
- 17 The Jülich Supercomputing Centre Launches the New SimLab Quantum Materials
- 20 HLRS Presents Visualization Applications at CeBIT 2017
- 24 HLRS Opens New Building as Training Center
- 26 Lord Mayor of Stuttgart Fritz Kuhn visits HLRS
- 27 Pre-Commercial Procurement of Human Brain Project Completed
- 29 5th PRACE Implementation Phase Project Started
- 32 Summer of Simulation 2016
- 37 Supercomputing in Scientific & Industrial Problems (SSIP) 2017
- 39 The 16th Annual HLRS/hww Workshop
- 41 The 64th IDC HPC User Forum
- 43 8th Blue Gene Extreme Scaling Workshop
- 45 The PRACE 2017 Winterschool
- 46 2nd EasyBuild User Meeting: A Growing Community Gathers at JSC
- 48 Intel MIC Programming & High Performance Computing in Atmosphere Modelling and Air Related Environmental Hazards @ IT4Innovations, Czech Republic
- 51 Student Cluster Competition 2016
- 53 New Literature from HLRS

Applications

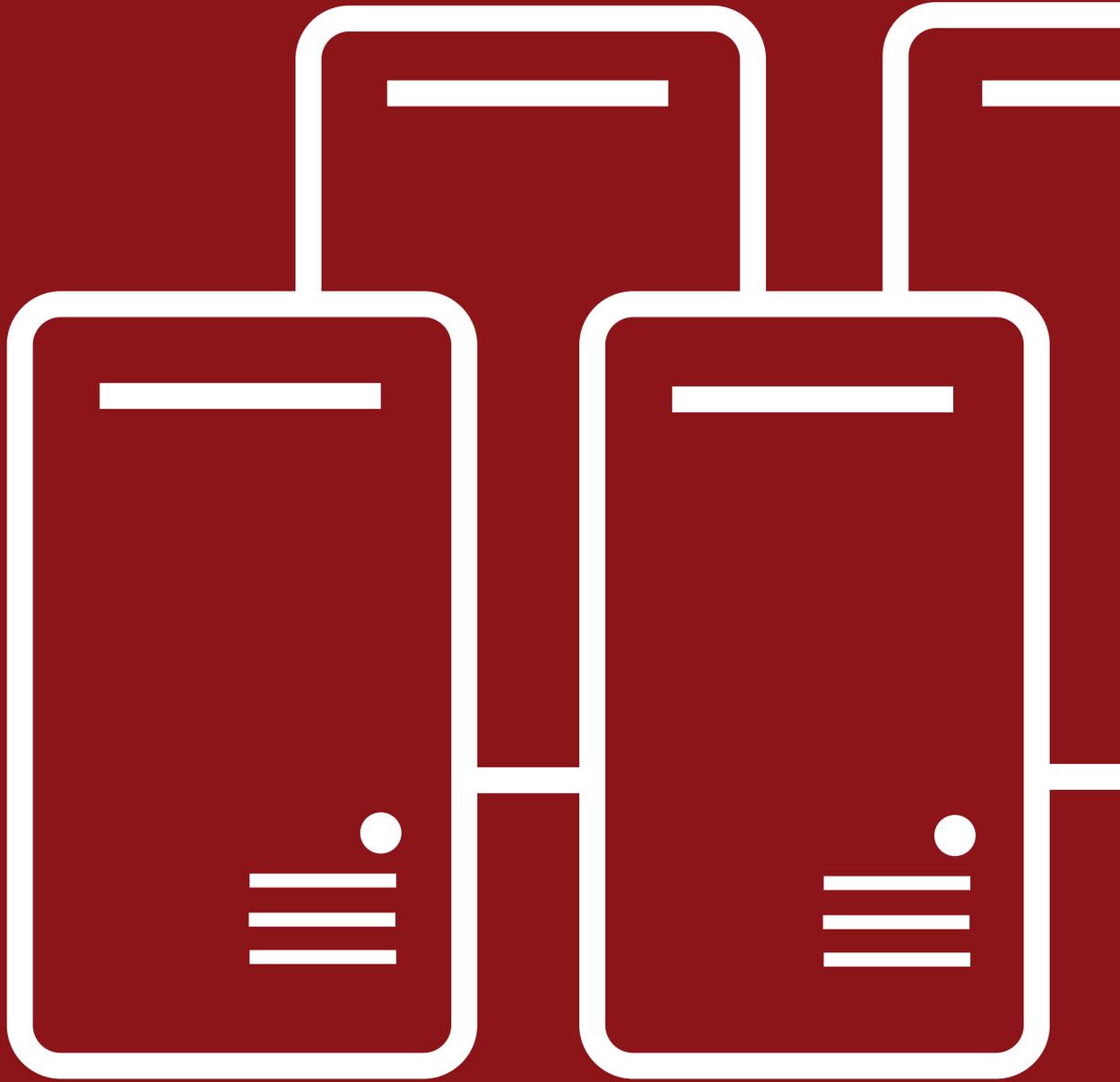
- 56 Smoothed Particle Hydrodynamics for Numerical Predictions of Primary Atomization
- 61 Large Scale I/O of an Open-Source Earthquake Simulation Package
- 65 The Next Generation of Hydrodynamical Simulations of Galaxy Formation
- 72 First-Principles Design of Novel Materials for Spintronics Applications

- 75 Applications of External Users on GCS Computers
Stability Confinement of the Brain Connectome
- 82 Lipid Transport by the ABC Transporter MDR3
- 87 Dark Matter Properties from a Supercomputer
- 92 Towards Exascale Computing with the Model for Prediction Across Scales
- 98 Rapid and Accurate Calculation of Ligand-Protein Binding Free Energies
- 102 The Cosmological Web Portal has gone Online at LRZ
- 108 EXAHD: An Exa-Scalable Approach for Higher-Dimensional Problems in Plasma Physics and Beyond
- 126 Next Generation of QPACE Operational
- 128 EXCESS Project Presents Energy-Aware Software Stack
- 131 Bringing eScience to the Cloud with PaaSage
- 134 The Mont-Blanc Project: Second Phase successfully finished
- 142 ArCTIC – Adsorption Chiller Technology for IT Cooling

Centers Systems Trainings

Projects

- 116 MoeWE – Modular Training Programme on High Performance Computing for Professionals
- 120 Helmholtz Data Federation
- 123 CATALYST: The Next Generation of Big Data Analytics at HLRS
- 148 LSC Leibniz Supercomputing Centre
- 150 HLRS Höchstleistungsrechenzentrum Stuttgart
- 152 JSC Jülich Supercomputing Centre
- 154 Many-Core Cluster “CoolMUC3” at LRZ
- 156 DGX-1 and Teramem
- 158 HPC Courses and Tutorials





News/Events

In this section you will find general activities of GCS in the fields of HPC and education.

GCS: Delivering 10 Years of Integrated HPC Excellence for Germany



In 2007, leadership from the three leading supercomputing centres—the High-Performance Computing Center Stuttgart (HLRS), the Jülich Supercomputing Centre (JSC), and the Leibniz Computing Centre of the Bavarian Academy of Sciences in Garching near Munich (LRZ)—agreed to combine their facilities and expertise, and, supported by the German Federal Ministry of Education and Research, and the science ministries of the states of Baden-Württemberg, Bavaria, and North Rhine-Westphalia, joined forces and founded the non-profit Gauss Centre for Supercomputing (GCS).

Over 10 years, the three major German supercomputing centres have been able to learn how to collaborate and integrate their vast intellectual and technological resources into a unified German supercomputing strategy.

This integration allowed the German High Performance Computing (HPC) community to deliver multiple generations of Europe’s most powerful supercomputers, classified as Tier-0 systems, and allowed GCS to offer scientists and engineers in Germany and Europe a diverse range of computing architectures and domain expertise to address some of the world’s most challenging problems related to public health, fundamental research, climate change, urban planning, new materials, scientific engineering, and green energy, among other key research areas.

Bringing three large institutions together to share expertise and coordinate research

activities required a herculean effort, and was not without risk. Working together, the directors of the three centres and their managing academic institutions helped integrate their respective manpower and computational resources to complement one another, offering users across a variety of scientific disciplines more tailored resources that would best work for their needs.

Indeed, GCS was named after one of Germany’s most famous problem-solvers, Carl Friedrich Gauss, a mathematician heralded as the father of numerical problem-solving and one who used practical solutions to answer long-standing questions.

Funding and Collaboration

GCS is jointly funded by the German Federal Ministry of Education and Research, the three states housing the three supercomputing centres, and three centres’ managing institutions: The University of Stuttgart, the Bavarian Academy of Sciences, and Forschungszentrum Jülich.

The collective power of these three centres makes Germany Europe’s HPC leader, and one of the strongest global players in HPC. The GCS structure enables the three respective centres to upgrade their machines in a “round-robin” fashion. This “life-cycle management” for the centres’ machines ensures that each centre can install a new machine every 3–4 years and GCS will constantly have a machine at the forefront of hardware and software technology. This type of close collaboration between



The 3 GCS centres (from left, clockwise): High Performance Computing Center Stuttgart (HLRS), Leibniz Supercomputing Centre in Garching/Munich (LRZ), Jülich Supercomputing Centre (JSC).

three national German HPC centres is unique in relation to national HPC strategies around the world.

Other than integrating the procurement process at the three centres, GCS also organized

integrated access methods, allowing researchers to submit a single proposal. These proposals then go through a rigorous peer review process to assess their scientific merit, then researchers get matched with the centre best-suited to their needs.

The centres work together on training courses by sharing best practices, dividing the task of training users on general HPC themes, and offering specialty training on centre-specific topics. In addition, the centres agree to work together on their outreach activities to better articulate the value that supercomputers bring to society for both government stakeholders and the public. In addition to three large-scale supercomputers, the centres offer researchers a variety of smaller clusters, data storage and analytics facilities, and sophisticated visualization resources to meet researchers' needs. Between the collaborative spirit of sharing knowledge and training activities, and the wide array of computational resources, GCS has consistently offered German and European researchers leading-edge systems, facilities, and support for the last decade.

Centre Specialization

While all three centres collaborate with researchers from across the scientific spectrum, each one does have certain specialties.

JSC is renowned for fundamental research, physics, and neuroscience, and has primarily built ultra-high-end machines with specialized architectures to support these fields. In fact, JSC's JUGENE machine—which, in 2007, appeared at number 2 on the biannual Top500 list of the world's fastest supercomputers—held Germany's highest-ever position on the list.

LRZ strongly supports geoscience, life sciences, and astrophysics research, and has been

a leader in energy efficiency for machines. LRZ consistently has one of the most energy efficient supercomputers in the world, and is capable of using far higher water temperatures to cool its SuperMUC machine—using water up to 40 degrees Celsius—than the 16 degrees common in other large-scale machines.

Due to its being situated in one of Germany's industrial hotbeds, HLRS naturally has a strong focus on engineering and industrial applications. While the centre regularly hosts one of the top 20 fastest machines in the world, it focuses on bridging the gap between cutting-edge computing technology and the commercial and industrial worlds. In addition to academic partners, HLRS has official research partnerships with automaker Porsche and German IT giant T-Systems, and holds the world record for the fastest commercial application, helping ANSYS scale its Fluent code to more than 172,000 compute cores.

The three centres' diverse capabilities and specializations have led to breakthroughs across science and engineering spectrum.

Research Highlights

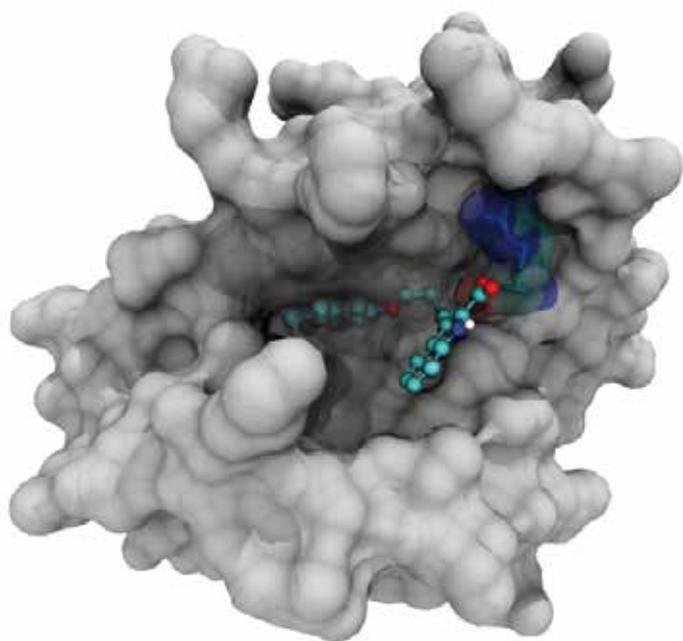
During a 2016 "extreme scaling workshop," a Pan-European research team led by University College London Professor Peter Coveney used LRZ's SuperMUC system to make a breakthrough in personalized medicine.

Many drug companies have turned to simulation to save time and money in developing



next-generation medications. Researchers can quickly simulate thousands (or millions) of combinations of molecules and human proteins, while simultaneously calculating the likelihood the two will bind together.

The Coveney team ran simulations on the SuperMUC system for 37 hours straight—using nearly all of SuperMUC’s 250,000 compute cores—combing through many combinations of proteins binding with common breast cancer drugs in the process. The team’s research helped create a roadmap for how simulation can quickly advance medicines from trials to market,



Using the SuperMUC system, the Coveney team was able to study drug efficiency based on a person’s genome. This image shows the MCL1 protein binding to the L5 ligand, shown in colour. Understanding how proteins and ligands bind together helps researchers determine a drug’s effectiveness. Image credit: Peter Coveney, University College London.

anticipate a drug’s effectiveness for individual patients, or anticipate possible side-effects.

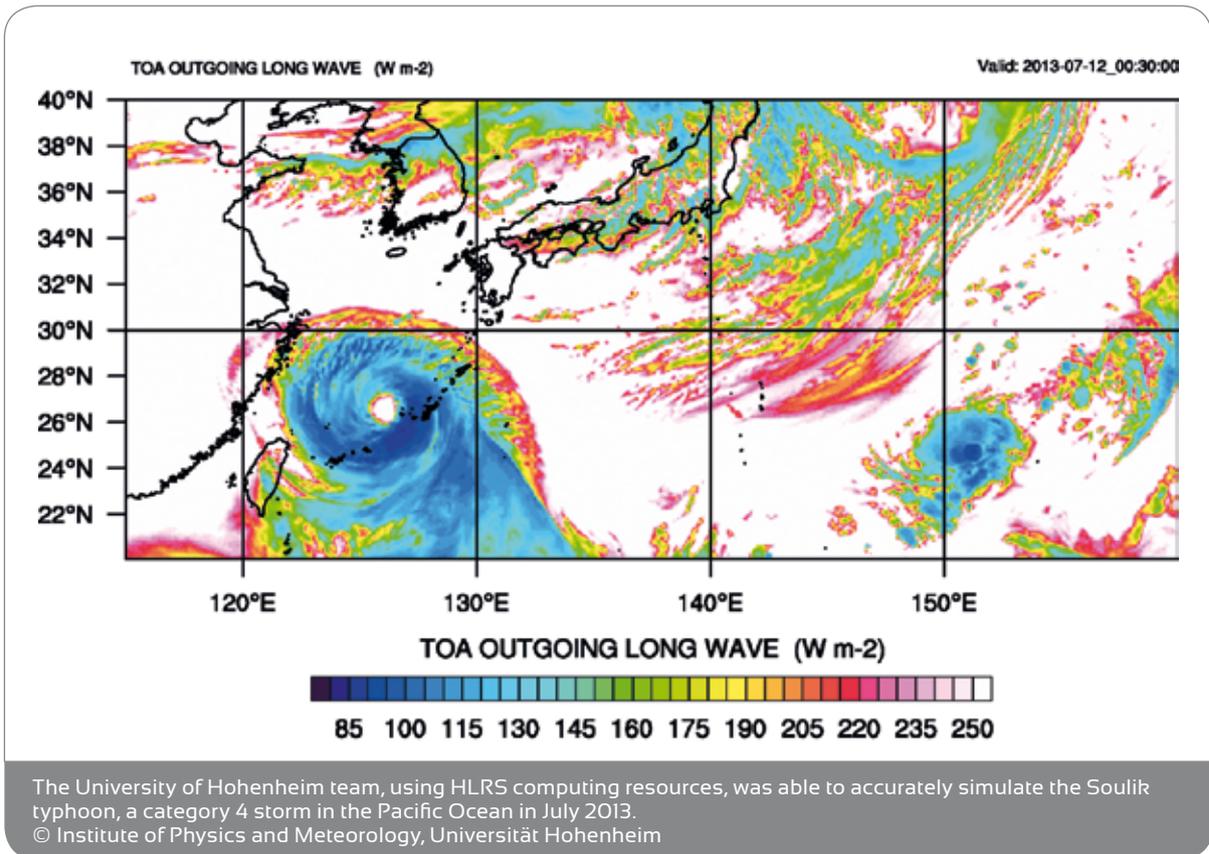
In 2015, researchers at the University of Hohenheim used HLRS’ Hornet system to run climate simulations of the Earth’s northern hemisphere at extremely precise scales.

In order for simulation to be able to accurately predict or accurately simulate past weather events, climate researchers must divide their respective maps into a very fine grid—each cube must be no larger than 20 square kilometres for researchers to be able to simulate the small-scale structures that contribute to weather events.

The Hohenheim team, led by Professor Volker Wulfmeyer, more than half of Hornet capability and accurately simulated the Soulik typhoon—a category 4 storm (meaning that wind speeds exceed 58 metres per second) that developed in the Pacific Ocean during the summer of 2013.

Researchers studying elementary particle physics—one of the most difficult science domains for experimental observation—have increasingly turned to supercomputers to make breakthrough discoveries.

In 2015, a University of Wuppertal team led by Professor Zoltán Fodor used JSC’s JUQUEEN system for a breakthrough within quantum chromodynamics (QCD). The field of QCD is focused on understanding foundational building blocks of our universe—the

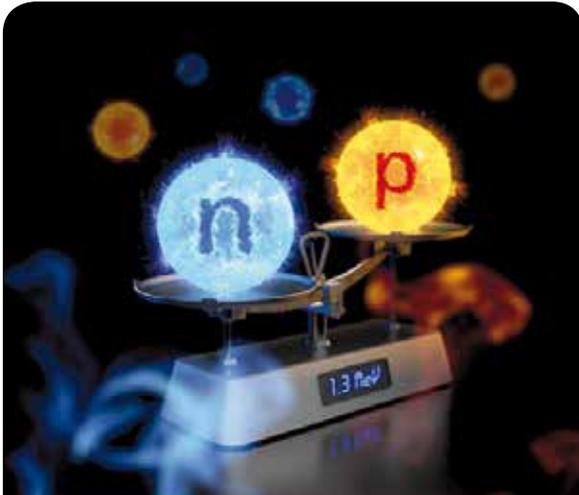


constituent particles of protons and neutrons, called quarks, gluons. These simulations requires precision—a proton is only .14 percent lighter than a neutron, but even the slightest difference in mass would have had major consequences for how galaxies, and in turn, our world, formed.

The Fodor team’s simulations, published in *Science*, helped validate the strong force—one of the four fundamental interactions in the universe, and the force that holds protons and neutrons together to form atomic nuclei. Shortly after publishing its work, the team was lauded

by Massachusetts Institute of Technology Professor Frank Wilczek—a 2004 Nobel Prize winner—in a *Nature* “News and Views” article.

In the article, Wilczek contextualizes the team’s breakthrough. “Nuclear physics, and many major aspects of the physical world as we know it, hinges on the 0.14% difference in mass between neutrons and protons. Theoretically, that mass difference ought to be a calculable consequence of the quantum theory of the strong nuclear force (quantum chromodynamics; QCD) and the electromagnetic force (quantum electrodynamics; QED). But



Using JSC's JUQUEEN, University of Wuppertal researchers were able to do an ab initio calculation of the proton-neutron mass difference, a major step forward for the field of quantum chromodynamics. Copyright: Bergische Universität Wuppertal

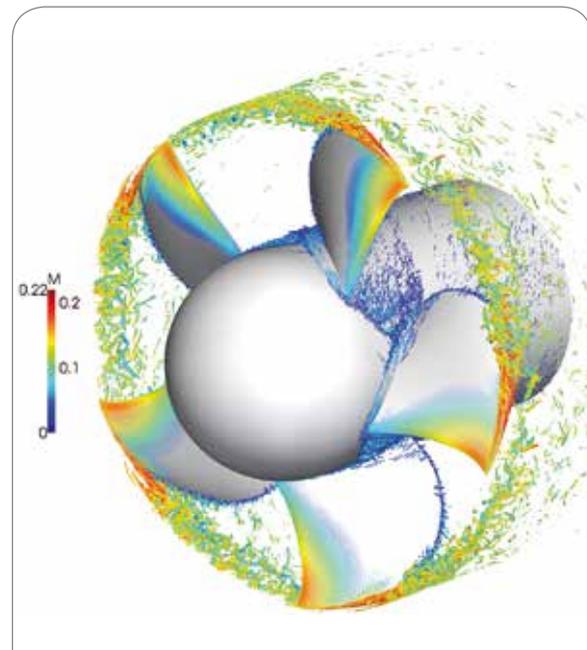
efficient aircraft engines. The team is one of the largest users of supercomputing time at HLRS, and has benefited from significant amounts of computing time at JSC as well, allowing the team to get the “best of both worlds” by having access to two diverse computing architectures.

Through its various allocations, the team has been able to successfully simulate the fluid dynamics occurring on increasingly complex engine designs, beginning with helicopter engine dynamics, the simulating the radial fan blade on a typical aircraft, then moving to more advanced space launchers, also called “Chevron nozzles.”

the required calculations are technically difficult and have long hovered out of reach. In a paper published in *Science*, [the team] report breakthrough progress on this problem,” he said in the *Nature* piece.

Some researchers, such as those led by Professors Wolfgang Schröder and Matthias Meinke from the Institute of Aerodynamics at RWTH Aachen University, benefit from the diversity and capabilities of GCS HPC resources.

The team focuses on understanding turbulence, one of the major unsolved problems in fluid dynamics. The team does not study turbulence in a purely theoretical way, though—it works closely with industry and experimentalists to understand turbulence in the context of making quieter, safer and, more energy



Using supercomputers, RWTH Aachen University researchers are able to simulate the vortical structures happening as an aircraft or helicopter blade spins. © Institute of Aerodynamics, RWTH Aachen University

Looking Ahead

Now in its 10th year, GCS has already begun implementing its strategy for the next decade of supercomputing excellence, or the “smart exascale” decade. The three GCS centres have all received funding for the next two rounds of supercomputers, and plan to not only further increase computing power, but also make their respective next-generation machines increasingly energy efficient. Faster supercomputers do not mean much without the ability to use them properly, though. To that end, GCS is further investing in training users and prospective users to make sure they spend less time porting codes or moving data and more time focused on their scientific research.

Leadership at the three GCS centres wants to continue delivering new solutions for their respective users, particularly in relation to disruptive technologies in the general HPC landscape. New computing architectures challenge traditional methods of doing simulation, and researchers need to be able to efficiently port their codes to make good use of a variety of compute architectures. Based on their decade of collaboration, the three GCS centres are excited about rising to this challenge by continuing to

offer the variety of systems and integrated support structure built during the last 10 years.

By delivering on its promises, GCS and its funding agencies have secured the finances to ensure that GCS will continue to serve as Germany’s HPC leader and one of the world’s most powerful and innovative research institutions. In the next decade, GCS centres will continue to enable breakthrough science that benefits German stakeholders, the European scientific community, industry, and society in general. GCS wants to continue collaboration between other German and European HPC centres—such as its continued membership in the Partnership for Advanced Computing in Europe, or PRACE—and contribute to the best practices that help make Europe one of the strongest HPC regions in the world.

Computing technology will continue to rapidly evolve, though, challenging those working in HPC space. GCS staff will continue to work in the spirit of the organization’s namesake, Carl Friedrich Gauss, who was quoted, “It is not knowledge, but the act of learning, not possession, but the act of getting there, which grants the greatest enjoyment.”

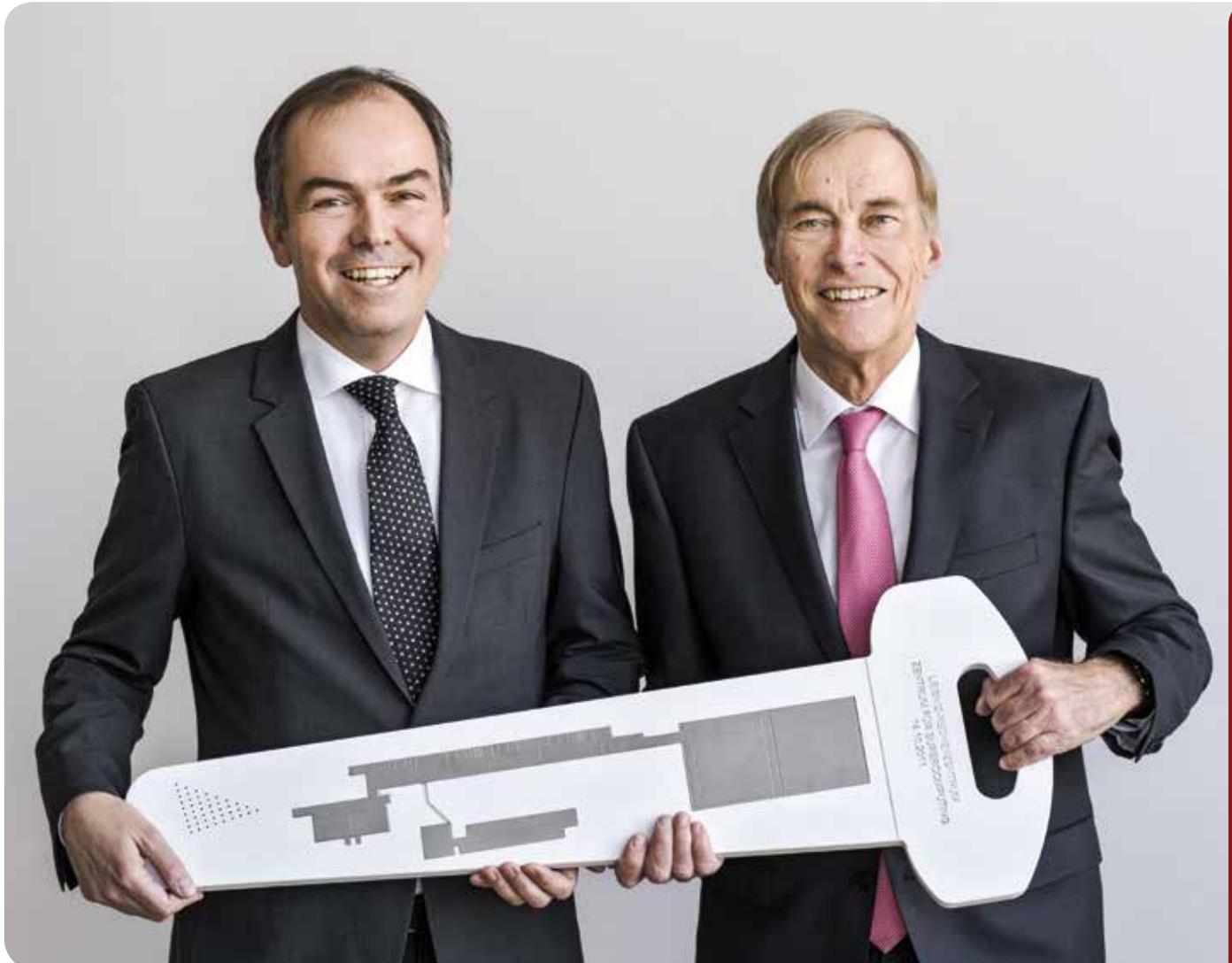


Written by Eric Gedenk

Eric Gedenk is a science writer for the Gauss Centre for Supercomputing. He is based at the High Performance Computing Center Stuttgart.

Contact: Eric Gedenk, gedenk@hlrs.de
eric.gedenk@gauss-centre.eu

New Chairman of the Board of Directors at LRZ



Starting April 1st, 2017, Prof. Dr. Dieter Kranzlmüller (left) began his new role as Chairman of the Board of Directors at Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities. He succeeds to Prof. Dr. Dr. h.c. Arndt Bode (right) who was Chairman of the Board since October 1st, 2008, and will stay one of LRZ's

directors in the future. Dieter Kranzlmüller has been one of LRZ's Directors for nearly ten years. He is full professor of computer science at Ludwig-Maximilians-Universität Munich. Just as Arndt Bode in 2008, Dieter Kranzlmüller starts as the Chairman with the procurement of the next generation super-computer at LRZ, SuperMUC-NG.

Changes in the GCS Board of Directors

Prof. Dr. Michael M. Resch, director of the High Performance Computing Center Stuttgart (HLRS), has been named the new chairman of the Gauss Centre for Supercomputing (GCS) Board of Directors as of April 7, 2017.

Resch succeeds Prof. Dr. Dr. Thomas Lippert, director of the Institute for Advanced Simulation at Forschungszentrum Jülich and head of the Jülich Supercomputing Centre who served from 2015–2017. Members of the Board of Directors serve for two years.

Resch, who takes over as the three GCS centres HLRS, JSC, and LRZ prepare to acquire their respective next-generation supercomputers, is enthusiastic about leading the next chapter of leading-edge German HPC. “As GCS enters its next phase of funding, I am excited to take the lead and, together with my two colleagues from Jülich and Munich, shape the future of supercomputing in Germany,” he said.

The three member GCS Board is completed by the directors of the other two GCS centres who



At an April 7, 2017, meeting, the GCS Board of Directors voted in Michael M. Resch as its chairman for the next two years.
© Boris Lehner for HLRS



act as BOD vice chairmen. For the new term, Prof. Dr. Dr. Thomas Lippert serves as the JSC representative and Prof. Dr. Dieter Kranzlmüller represents the LRZ. Kranzlmüller is a newly elected member of the GCS Board. He succeeds Prof. Dr. Heinz-Gerd Hegering, who - as one of the initiators and founding fathers of GCS - had been active on the GCS BOD from the very beginning and recently stepped down.

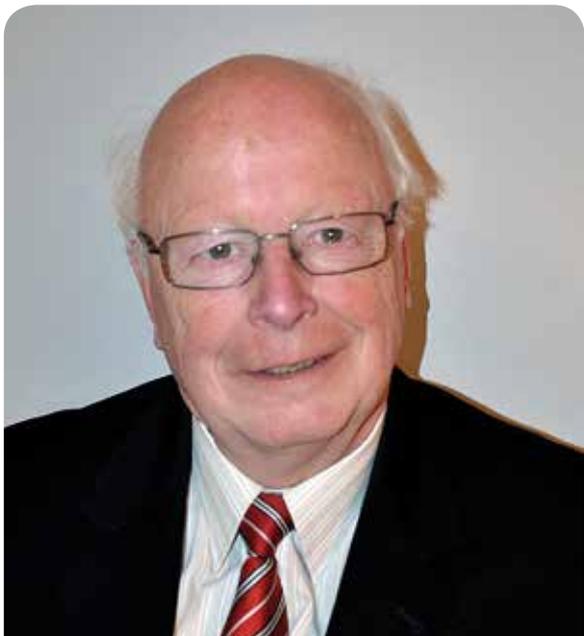
GCS management and the association members would like to whole-heartedly express their deep appreciation to Heinz-Gerd Hegering for his exemplary dedication and commitment during his 10 years of service on the GCS Board and wish him all the best for his well-deserved retirement.



Vice-Chairman of the GCS Board of Directors:
Prof. Dr. Dieter Kranzlmüller (LRZ)



Vice-Chairman of the GCS Board of Directors:
Prof. Dr. Dr. Thomas Lippert (JSC)



Prof. Dr. Heinz-Gerd Hegering (LRZ) has stepped down from the GCS Board as he enters his well-deserved retirement.

GCS Continues as Hosting Member of PRACE 2

The Gauss Centre for Supercomputing (GCS) extends its role as a hosting member of the Partnership for Advanced Computing Europe (PRACE) into the European programme's 2nd phase, PRACE 2, which runs from 2017 to 2020. At the 25th PRACE council meeting, held in Amsterdam, The Netherlands, it was agreed that the HPC systems of GCS members HLRS (High Performance Computing Center Stuttgart), JSC

hosting members (BSC representing Spain, CIN-ECA representing Italy, CSCS representing Switzerland, and GENCI representing France), the PRACE programme provides a federated, architecturally diverse supercomputing that allows for computing allocations that are competitive with comparable programmes in the USA and in Asia. PRACE 2 plans to deliver 75 million node hours per year for European research projects.



JUQUEEN – Jülich

(Jülich Supercomputing Centre), and LRZ (Leibniz Supercomputing Centre Garching/Munich) will continue to be available for large-scale European research activities of great importance. In doing so, GCS will maintain its leading role in European HPC space and will significantly contribute to boosting scientific and industrial advancements.

With providing computing time on its three HPC systems Hazel Hen (HLRS), JUQUEEN (JSC), and SuperMUC (LRZ), the GCS contributes the lion's share of computing resources for the European HPC programme. Together with the other four



SuperMUC – Garching

Any industrial or academic researcher who resides in one of the 24 PRACE member countries is allowed to apply for computing time through PRACE 2. Resource allocations are granted based on a single, thorough peer review process exclusively based on scientific excellence of the highest standard. The principal investigators of accepted research projects will be able to use computing resources of the highest level for a predefined period of time. Additionally, they will receive support from the respective centres' outstanding support teams to make the best possible use of these world-class HPC resources.



Hazel Hen – Stuttgart

The Jülich Supercomputing Centre Launches the New SimLab Quantum Materials



The SimLab Quantum Materials (SLQM) was officially established on January 1st 2017 by the Jülich Supercomputing Centre, following the sun-setting of the JARA-HPC SimLab “ab initio methods in Chemistry and Physics”. Strategically positioned at the intersection between HPC and materials science, two of the pillars of the Strategic Priority “information” of the Forschungszentrum Jülich, SLQM naturally bridges the gap between material simulations based on quantum mechanical methods and high-performance computing.

From the onset of the first theoretical breakthrough in the early ‘60s, passing through the first computational codes based on Density Functional Theory (DFT) to arrive to modern simulations involving hundreds if not thousands of atoms, materials science (MS) has evolved in a mature field producing 20,000 publications per year. This success was achieved by the continuous and painstaking effort that code developers placed in increasing the functionalities of their codes and, above all, in adapting their simulation codes to newer and more powerful computing platforms. This effort is still present today but, as simulation software gets bigger and more complex, is becoming increasingly difficult to maintain scalability, performance, and portability all at the same time. The evolution of computing architectures on the road to exascale is making the task of executing large scale and complex material

simulations even harder. Due to the limited concurrency of algorithms currently used in MS, there is a huge burden that weighs heavily on the capability of executing simulations on both state-of-the-art and emerging massively parallel platforms. Enabling materials science codes to overcome such burden crucially depends on the design, selection and optimization of modern high-performance scalable algorithms. At present, this is the most critical challenge in the advancement of scientific computing in materials science.

Research Activities

The new SimLab Quantum Materials [1] addresses the challenge exposed above with a research program centered on three specific activities. First, SLQM focuses on the development and maintenance of new modern numerical libraries tailored to specific linear algebra tasks emerging from materials science simulation software. For instance, a central task in DFT, as well as in more advanced methods dealing with excitation Hamiltonians, is the solution of an algebraic eigenvalue problem. Tackling the solution of large eigenproblems with traditional black-box libraries is not any more a feasible solution when the number of atoms (or the number of excited spectra) counts

in the tens of thousands. To this end, SLQM develops and maintain new modern iterative eigensolvers [2] which are compute bound, portable and expose an increased level of parallelism.

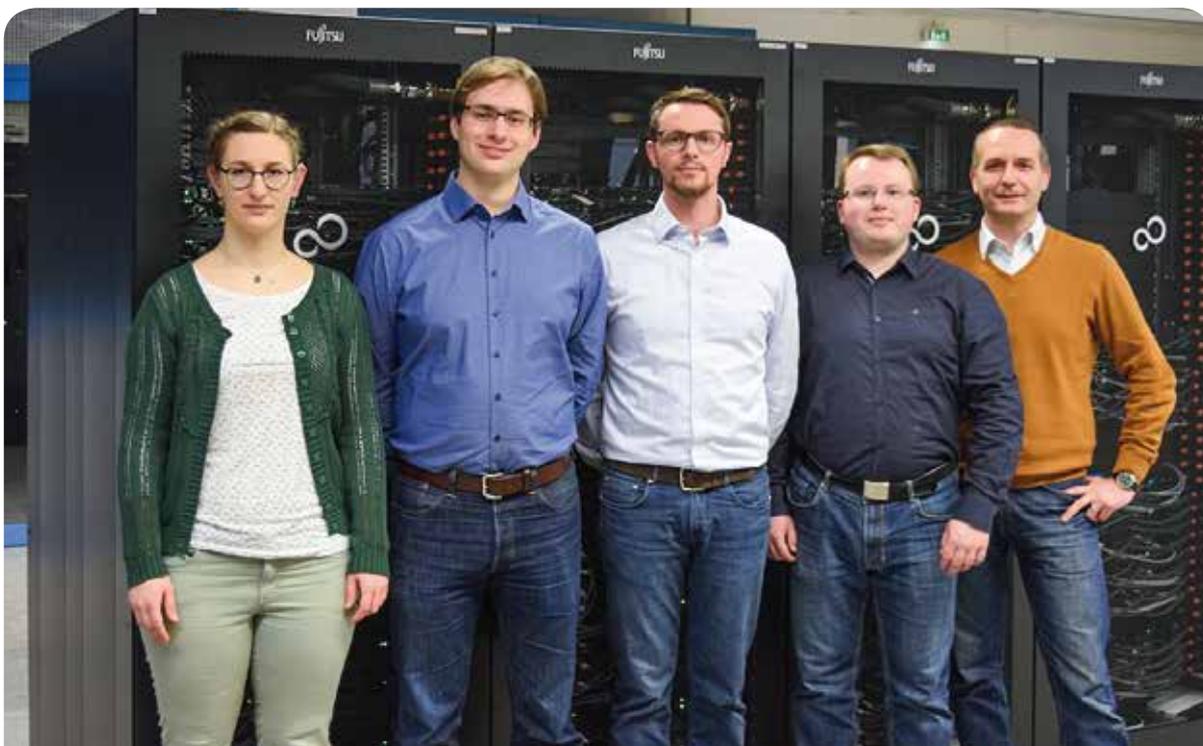


A second area of activity deals with the design and implementation of high-performance algorithms targeting specific combination of operations extracted from simulation codes in the broad area of quantum materials. An example of recent work in this sense is the development of a performance portable algorithm, termed HSDLA, for the initialization of Hamiltonian and Overlap matrices in DFT methods based on the LAPW basis set [3]. Last but not the least, the Lab carries on research on the development of new mathematical models and computational paradigms aimed at improving performance, accuracy and scalability of simulations within a well-defined methodological framework. In this context an exemplary case is provided by the

development of an adaptive preconditioner that would dramatically improve the convergence of Self-Consistent Field loops which appear almost ubiquitously in solving the Schrödinger equation.

Collaboration and Outreach

Inheriting the ongoing collaborations of the SimLab “ab initio”, SLQM has an extensive network of contacts both at the national and international level. Such collaborations orbit around research activities on specific projects as well as more systematic cooperations on a number of cross-disciplinary topics. For instance, the Lab maintains a strong presence within the Aachen Institute for Computational Science



SimLab Quantum Materials team



and Engineering at the RWTH Aachen university and in particular with the High Performance and Automatic Computation group. At the international level, SLQM has active collaborations with groups working on algorithm and method development in the US and Japan.

The SimLab also has a keen commitment to outreach and education. Members of the Lab actively participate to international conferences and workshops. The Lab also organizes events and minisymposia that reach out to different materials science communities as well as to the more general Computational Science and Engineering crowd. SLQM is actively engaged to form a new generation of interdisciplinary scientists that will constitute the backbone of future communities engaged in computational materials science.

Opening Workshop

On the 4th of April 2017, the SimLab Quantum Materials held a one-day event at the Rotunda Hall of the Jülich Supercomputing Centre [4]. The workshop represented the official opening introducing SLQM to the scientific community. This event aimed at bringing together scientists from the Forschungszentrum Jülich (FZJ),

RWTH Aachen University as well as international research institutes who were interested in participating in the research and support activities of the SimLab. By sharing their work and scientific expertise the participants were given the chance to contribute to the definition of a platform allowing for the mutual collaboration and prioritization of the Lab's activities. Each talk dealt with diverse and emerging topics in computational materials science including correlated electrons materials, topological magnetism, materials for energy, and high-performance computing, just to name a few. Each contribution was followed by constructive discussions, which often spilled over into the break between sessions. The modern and interdisciplinary nature of the topics attracted a very lively audience, comprised of about 50 scientists from all over Germany.

References

- [1] <http://www.fz-juelich.de/ias/jsc/slqm>
- [2] **M. Berljafa, D. Wortmann, and E. Di Napoli:**
An Optimized and Scalable Eigensolver for Sequences of Eigenvalue Problems. *Conc. Comp.: Pract. Exper.* 27, pp. 905-922, 2015.
- [3] **E. Di Napoli, E. Peise, M. Hrywniak and P. Bientinesi:**
High-performance generation of the Hamiltonian and Overlap matrices in FLAPW methods. *Comp. Phys. Comm.* 211, pp. 61-72, 2017
- [4] <http://indico-jsc.fz-juelich.de/event/42/>



Written by **Edoardo Angelo Di Napoli**

(SimLab Quantum Materials, Jülich Supercomputing Centre (JSC), Germany)

Contact: Edoardo Angelo Di Napoli e.di.napoli@fz-juelich.de

HLRS Presents Visualization Applications at CeBIT 2017

In March 2017, digital innovators converged on Hannover for CeBIT 2017, the world's largest trade show for information technology. Among the participants was the High Performance Computing Center Stuttgart, which presented several recent applications of its powerful visualization and simulation technologies.

The HLRS Visualization Department supports engineers and scientists by providing tools for the visual analysis of data produced by high performance computers. Using approaches based in virtual reality, augmented reality, and simulation, the Visualization Department

provides tools for representing and intuitively interacting with large, complicated data sets in ways that make them easier to interpret and use. Its presentation at CeBIT revealed some ways in which HLRS is using visualization to address complicated problems in engineering and forensics.

Visualizing Traffic Accidents and Elevator Design

Prominently displayed in the middle of the HLRS booth was a motorcycle and model driver that had been scanned into a computer using a three-dimensional laser scanner. On a screen the exhibitors demonstrated how high



Augmented Reality visualization of airflow around a race bike.



Federal Minister of Education and Research Prof. Dr. Johanna Wanka visiting the HLRS booth.

performance computing can simulate airflow around the vehicle in motion. By combining such models with visualization strategies, it becomes easier for engineers to optimize aerodynamic properties during the design process.

The HLRS team also presented an augmented reality application that they are developing for computer aided crash analysis. Called VIS-DRAL (Virtual Investigation, Simulation and Documentation of Road Accidents via Laser-scanning), the approach uses 3D laser scanners to scan the vehicles that were involved in

an accident in great detail, capturing the shape of the damaged vehicles after the collision. Scanners can also capture other important factors in a crash, such as the condition of the road surface and the locations of other objects in the surrounding environment. These data are then imported into the computer and superimposed upon one another, making it possible to simulate models of the crash. This enables the HLRS investigators to reconstruct precisely where the two vehicles must have collided and to infer other information such as how fast they must have been traveling.

“Right now it can be very difficult for an accident investigator or insurance inspector to determine exactly what happened,” says Uwe Wössner, head of the Visualization Department. “We think that simulation could provide a more reliable way to reconstruct these kinds of events.”

In an additional presentation, HLRS described its collaboration with Thyssen-Krupp on the development of MULTI, a new elevator concept that dispenses with cables and makes it possible for cars to move horizontally as well as vertically through a building. Although such a project offers exciting new opportunities for moving people more efficiently, it also presents unique challenges, as the motions of multiple cars must be coordinated and optimized like trains in a transit network.

HLRS contributed to the project in several ways. One involved designing an immersive virtual reality environment that simulated the experience of traveling in the elevator. The visualization included multiple layers of detail within the building, making it possible for users to “erase” walls and closely inspect the elevator’s machinery. HLRS also simulated airflow around two cars passing one another inside an elevator shaft, which can cause vibrations that affect passengers’ comfort. These studies enabled Thyssen-Krupp engineers to identify problems in their initial designs and to develop solutions that would address them.

Federal Minister Visits HLRS Booth

Another highlight of HLRS’s participation at CeBIT was a visit from Prof. Dr. Johanna Wanka, Germany’s Federal Minister for Education and Research. At the HLRS booth she spoke at great length with Dr. Wössner and enjoyed a demonstration of a virtual smoke wand that illustrates airflow around the motorcycle when in motion. Dr. Wanka expressed great interest in the HLRS’s activities, particularly in the wide variety of ways in which high performance computing and data visualization can be used to address societal and engineering challenges. HLRS was one of four exhibitors at CeBIT from the University of Stuttgart. Also participating were the startups Blickshift Analytics, DICEHub, and MeSchup, all of which originated in the University’s research community. The exhibitors from Stuttgart joined more than 500 exhibitors from the state of Baden-Württemberg. Such a strong presence highlighted the importance of new digital technologies in the research and industrial development of southwestern Germany.

Other key themes represented at CeBIT included data analytics, the Internet of Things, drones and unmanned systems, and virtual and augmented reality, among many others. HLRS is looking forward to participating in CeBIT 2018, which is scheduled to take place once again in Hannover on June 11-15, 2018.



3D scanning of a race bike and the booth and surrounding.

HLRS Opens New Building as Training Center

HLRS is Europe's largest training institution for high-performance computing (HPC). In 2016 more than 1,000 trainees took part in our courses and approximately 400 attendees participated in our workshops and seminars. These numbers are likely to grow in the coming years as HLRS implements new training concepts addressing customers from industry and from small and medium-sized engineering companies.

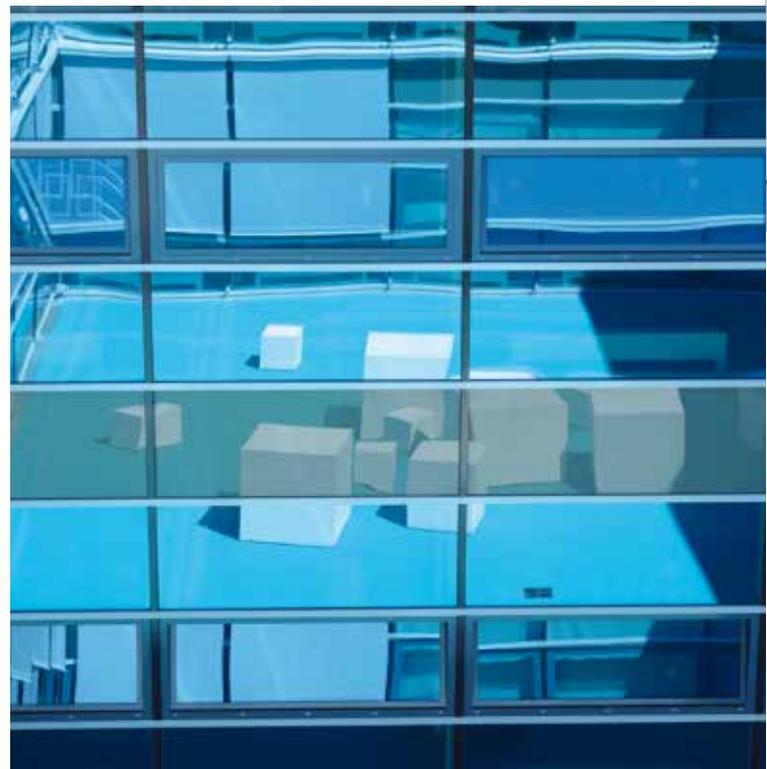
To accommodate these increased demands, HLRS recently opened a new HPC training center as an annex to our existing building. The facility, whose construction costs were supported by a nearly 6 million Euro investment from the University of Stuttgart, started operation on March 7, 2017.

The centerpiece of the approximately 2,000 square meter building is a 254 square meter classroom. State-of-the-art IT and multimedia equipment accommodates simultaneous hands-on programming training for more than 60 attendees. When configured for conferences and lectures, the facility can hold more than 150 participants. The new building also includes much needed office space as well as several smaller meeting rooms. Heating comes from waste heat generated by the HLRS supercomputer, located just next door.

An open space outside the classroom offers a pleasant environment for informal conversation while catering can be provided onsite. An interior courtyard also features an installation by artist Harald F. Müller based on the symbolist

poet Stéphane Mallarmé's concept of the "throw of the dice". Müller also designed the building's color scheme, which harmonizes with that of the previously existing building.

"As a scientific research institution, it is HLRS's duty to foster science and research. However, we also have the mission to transfer know-how from science to industry. This is why we engage in new initiatives that include training programs tailored for nonscientific users, specifically for engineers in the greater Stuttgart region who need to add simulation to their skillsets. The new facility will improve our ability to achieve this goal," emphasizes HLRS Director Michael Resch.





Written by Norbert Conrad
High Performance Computing Center Stuttgart (HLRS)

Contact: conrad@hlrs.de

Lord Mayor of Stuttgart Fritz Kuhn visits HLRS



From right to left: Head of HLRS Operations Thomas Beisel, Lord Mayor of Stuttgart Fritz Kuhn, Director of HLRS Michael Resch, Director of Economic Development of Stuttgart Ines Aufrecht, Mayor for Administration, Culture, and Legal Affairs Fabian Mayer

On February 17, 2017 Lord Mayor Fritz Kuhn visited HLRS for the first time to get a view of the center and to discuss topics of common interest. Lord Mayor Kuhn was accompanied by Mayor Fabian Mayer and Ines Aufrecht, Director of Economic Development.

The group was given an introduction to HLRS and toured to the computer room, where they viewed Europe's fastest productive supercomputer and got an impression of the effort it takes to operate such systems. The tour was completed with a visit to the HLRS virtual reality lab, where scientists from HLRS showed samples of simulations to support urban planning

processes as part of the Reallabor project funded by the state of Baden-Württemberg.

In the discussion, Lord Mayor Kuhn showed a keen interest both in the potential of HPC technology for urban planning and for the regional economy. Prof. Resch emphasized the openness of HLRS to collaborate more intensely with the city of Stuttgart on projects both in science and industry but also in contemporary arts. Both sides agreed to continue the discussion focusing on the issue of network connectivity for small and medium sized enterprises but also on the potential of collaboration in cultural affairs.

Pre-Commercial Procurement of the Human Brain Project Completed



One of the key goals of the Human Brain Project (HBP) [hbp] is the creation of a research infrastructure for neuroscience, which also comprises high-performance computing (HPC) and high-performance data analytics (HPDA) systems. This area of research comes with new and challenging requirements, and thus the HBP also needs to work on enabling new technologies and architectures.

In 2014, the HBP therefore launched a pre-commercial procurement (PCP) of research and development services, which was successfully completed in January 2017. During a wrap-up workshop at JSC in March, the contractors of the last phase of the PCP, Cray and a consortium of IBM and NVIDIA, and researchers from the HBP discussed the resulting solutions and their evaluation.

One of the technical goals of this PCP was to facilitate new approaches to the integration of dense memory technologies. New solutions should enable global access to distributed dense memory. IBM expanded previous work on a Distributed Storage Access (DSA) layer to enable access to distributed dense memory within a global address space through an RDMA-type interface. On a pilot system, which was deployed at JSC in late summer 2016, it could be demonstrated that this interface allows to achieve both, extremely high bandwidth as well as a high IOPS rate. The competitor Cray proposed a solution based on a commodity technology, namely Ceph. So far, this technology has hardly been used for HPC,

and its potential for software-defined storage architectures based on dense memory has not yet been exploited.

For many of the use cases of the HBP, visualization plays a key role and was thus chosen as a second focus of the PCP. The goal was to enable scalable visualization capabilities that are tightly integrated into large-scale HPC and HPDA systems. A big challenge was with the NVIDIA team to enable the complex visualization software stacks of the HBP on GPU-accelerated POWER servers. The new OpenPOWER HPC architectures do rely heavily on the compute performance provided by GPUs and thus are particular suitable for large-scale visualization.

The third technical focus area was dynamic resource management. The HBP sees this as a requirement to improve the utilization of future HPC architectures and to support the HBP's complex workflows. In this context, new features allowing to change the resources available to a running job have been added to resource managers, and mechanisms for resizing job sizes have been designed and implemented.

The PCP is a quite new instrument in Europe for working with commercial operators and promoting the development of innovative solutions. Within the HBP, its use could be successfully demonstrated for enabling new HPC capabilities. The solutions developed within the PCP will become available to a broader



The two pilot systems, developed by Cray (left) and IBM-NVIDIA (right), are designed to meet the specific requirements of neuroscience applications. Both systems are connected to an high-performance storage system from IBM/Lenovo (middle).

community and can already be exploited by HBP scientists through the two pilot systems, JURON and JULIA, which have been deployed at JSC as part of this project.

References

[hbp] Human Brain Project:

<https://www.humanbrainproject.eu>

[hbp-infra] B. Orth et al.:

"Towards an European infrastructure for brain research." Inside, Autumn 2016.

[dsa] B. Metzler et al.:

"Direct Storage-class Memory Access".
https://www.openfabrics.org/images/eventpresos/workshops2014/DevWorkshop/presos/Monday/pdf/O9_DirectStorage.pdf

Written by Dirk Pleiter and Boris Orth

(Jülich Supercomputing Centre (JSC), Germany)

Contact: Dirk Pleiter, d.pleiter@fz-juelich.de, Boris Orth, b.orth@fz-juelich.de

5th PRACE Implementation Phase Project Started



In response to the European Commission's EINFRA-11-2016 Call within the new European framework programme Horizon 2020 PRACE partners from 25 countries submitted a successful proposal and started the 5th Implementation Phase project (PRACE-5IP) on 1 January 2017. The project will assist the transition from the initial five-year period (2010-2015) of the Research Infrastructure established by the Partnership for Advanced Computing in Europe (PRACE) to PRACE 2 and support its implementation.

Key objectives of PRACE-5IP are:

Provision of Tier-0 service based on excellence and innovation.

With this project, it is proposed that the now regular PRACE Access Calls would continue and constitute the bulk of the audited allocation of resources. A rigorous and proven peer-review model which meets or exceeds the standards of comparable processes in participating countries, and indeed globally, will be further enhanced. Tailored services will be developed to facilitate collaboration with groups with specific requirements e.g. ESFRI projects, Horizon 2020 programmes and industry. Enabling efficient use of the Tier-0 systems via targeted user and application support is a key objective of this project. In addition to the continuation of the application enabling, documentation and benchmarking of the previous projects, PRACE already has a programme underway to provide computational resources and, more importantly at this point, human support and requirements assessment, to the recently established FET Centres of Excellence (CoEs).

Support a functional European HPC Eco-system.

The primary service provision at the Tier-0 level is dependent on a well-functioning Tier-1 service at national level. This is exemplified in the domain of training where PRACE currently operates six PRACE Advanced Training Centres (PATCs). In this project PRACE will develop a "PRACE Training Centre" (PTC) brand, i.e. a standardised quality controlled syllabus appropriate to Tier-1. This will not only provide a service to the wider community but will also raise the profile of PRACE and prompt Tier-1 users to look beyond their borders in due course. In parallel with this, code development is needed. This is a naturally cross-tier activity with initial developments made at smaller scales. PRACE will continue such work with programmes such as Preparatory Access and expert technical effort on selected codes. PRACE has also a functional programme" at Tier-1 level which operates the DECI Calls. This optional programme will be retained.

A sustainable governance and business model.

PRACE undertakes a periodic, independent auditing exercise to formally account for delivery of commitments made. In concert with the development of the PRACE model, the audit process must develop, too. Criteria will be made more specific and reporting more comprehensive, reflecting the maturity of the Research Infrastructure. Despite these steps the legal and taxation complexity of a multi-national funded entity like PRACE will remain and models beyond those currently employed will be examined.



PRACE collaborators at the PRACE-5IP Kick-off.

Work with European Technology Platform for HPC (ETP4HPC) and Centres of Excellence in HPC.

PRACE's complementary working relationship with ETP4HPC is exemplified by its commonly executed Horizon 2020 project EXDCI. Furthermore, PRACE and ETP4HPC have eight common partners. Within the collaboration with ETP4HPC specific technical reports will be produced addressing emerging user requirements along with Exascale performance and usability testing. In the context of the Centres of Excellence (CoE's), today PRACE is establishing a strong working relationship with these important activities and in this project will put in place a programme of initiatives that address domain specific challenges facing individual centres.

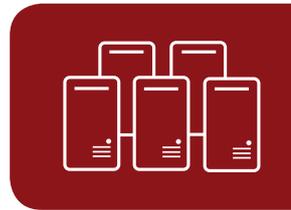
Provide tailored Training and skills development programmes.

Additionally to the operation of the PRACE Advanced Training centres (PATCs) and new PRACE Training Centres (PTCs), PRACE-5IP will produce material which can either be used directly by end-users or to support intermediaries such as teaching staff. In particular, the

project will support Massively Open Online Courses (MOOCs) or blended on-site taught courses with some MOOC elements. The effective implementation of the PTC programme will enable the PRACE brand to be associated with a wider number of courses with an increased number of contact points into academia in partner countries while amortising the cost of the generation of high quality material over a larger number of consumers.

HPC system interoperation.

PRACE will study the use of Containers as a means of exchanging configurations and packaged applications, enhancing user mobility. In addition, OpenStack's applicability to general HPC workloads will be examined. Together these methods could form a lingua franca for interoperating at the technical level with other classes of computing infrastructure. At the non-technical level, as PRACE access is based on the peer review requirements of Member States, equivalence relationships will be studied to identify where resource exchange is both advantageous to end users and within the bounds of conditions imposed at the national level.



PRACE-5IP is again coordinated and managed by by Forschungszentrum Juelich. It has a budget of nearly 15.9 Mio € including an EC contribution of 15 Mio €. The duration will be 28 months.

Over 250 researchers collaborate in PRACE from 49 organisations¹ in 25 countries. The PRACE-5IP Kick-Off meeting took place at the Greek Research and Technology Network S.A., Athens from 1-2 February 2017 with over 130 researchers attending the event.

Synopsis of the PRACE projects

The European Commission supported the creation and implementation of PRACE through six projects with a total EC funding of 97 Mio €. The partners co-funded the projects with more than 33 Mio € in addition to the commitment of 400 Mio € for the initial period by the hosting members to procure and operate Tier-0 systems and the in-kind contribution of Tier-I resources on systems at presently twenty partner sites. The following table gives an overview of the PRACE projects.

¹ This includes 26 Beneficiaries and 23 linked Third Parties from associated universities or centres.

Project Id Grant Number	Number Partners	Budget Mio €	EC Funds Mio €	Duration	Status
PRACE-PP RI-211528	16	18.9	10.0	01.01.2008 – 30.06.2010	Completed
PRACE-IIP RI-261557	21	28.5	20.0	01.07.2010 – 30.06.2012	Completed
PRACE-2IP RI-283493	22	25.4	18.0	01.09.2011 – 31.08.2013	Completed
PRACE-3IP RI-312763	26	26.8	19.0	01.07.2012 – 30.06.2014	Completed PCP ² to end 31.12.2017
PRACE-4IP EINFRA-653838	26	16.5	15.0	01.02.2015 – 30.04.2017	Ongoing
PRACE-5IP EINFRA-730913	26	15.9	15.0	01.01.2017 – 30.04.2019	Started
Total		132.0	97.0		

² The Pre Commercial Procurement Pilot (PCP) is part of the PRACE-3IP project, however with a duration of 66 month.

Written by Florian Berberich

Jülich Supercomputing Centre (JSC), Germany

Contact: Florian Berberich, f.berberich@fz-juelich.de

Summer of Simulation 2016

Current multi and many core architectures pose a challenge to simulation codes with respect to scalability and efficiency. Particularly life and material-science simulations often cannot just increase their system sizes because underlying algorithms scale unfavorably and insights from larger systems are limited. Still, computational demands are high due to the required abundant sampling of phase space or molecular structures and more accurate physical descriptions.

In May 2016, the biolab@lrz of the LRZ applications support initiated the "Summer of Simulation" (SoS16) to help young scientists in these fields to tackle their problems on current supercomputers. Masters and PhD students employing molecular dynamics or quantum chemical simulations were called to submit a short one-page proposal describing their project. The aim was to port these projects on SuperMUC, find an optimally scaling setup, and run their applications during the summer-semester break.

The SoS16 started with a kickoff meeting in July, where the eight participants from the Ruhr-University Bochum (RUB), the University of Bonn (UB), the Friedrich-Alexander University Erlangen-Nurnberg (FAU), and the Technical University Munich (TUM), respectively, presented their projects and were assigned one of the four tutors from the LRZ BioLab. The projects spanned a broad range from highly accurate surface chemistry over enzymatic reactions to simulations of nanoporous gold with a few hundred millions of particles.

In the following five weeks, the students had to get their codes and simulations running on SuperMUC and optimize the setup. Here, each project had an initial budget of one million CPUh for preparatory simulations and to demonstrate the scalability of their project. With the guidance of their tutors, the students prepared follow-up proposals to apply for up to nine million additional CPUh. After a speedy review process, a total of 50 MCPUh had been granted and were available until 15th of October for the simulations.

At the closing workshop end of October each student presented the progress and results made over summer and handed in a final report. Currently, most projects are evaluating their data for scientific publications. Furthermore, follow-up proposals for CPU time based on their experiences obtained during SoS16 are under way.

Three of the projects involved the density functional theory code CPMD (www.cpmc.org) for first-principles molecular dynamics to explore chemical reactions on metal oxide surfaces. For these many electron systems ultra-soft Vanderbilt pseudopotentials are applied, for which comparatively small plane-wave basis sets suffice and drastically reduce the overall computational costs. On the downside, this limits the number of MPI processes that can be used in the simulation because CPMD is parallelized by distributing planes of the real space mesh that supports the plane-wave basis and the electron density. During the SoS16 and with the help of the LRZ BioLab the hybrid OpenMP/MPI parallelization of the Vanderbilt code within CPMD

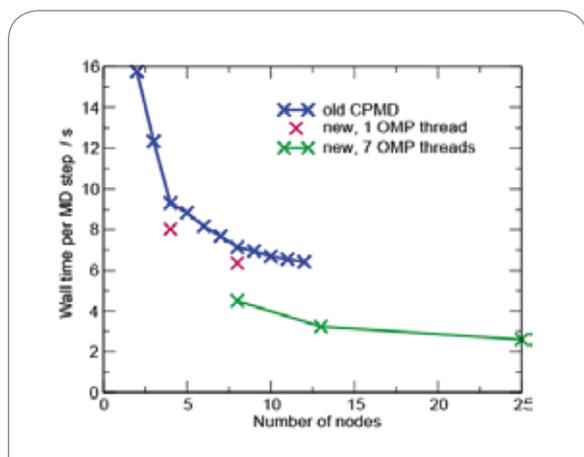


Fig. 1: Improved scaling of CPMD on SuperMUC Phase 2.

was revised and permitted an optimal setup of 4 MPI processes times 7 OpenMP threads per node of SuperMUC Phase 2. Together with the tuning of the Intel MPI library, a notable speed up was achieved as is shown in Figure 1.

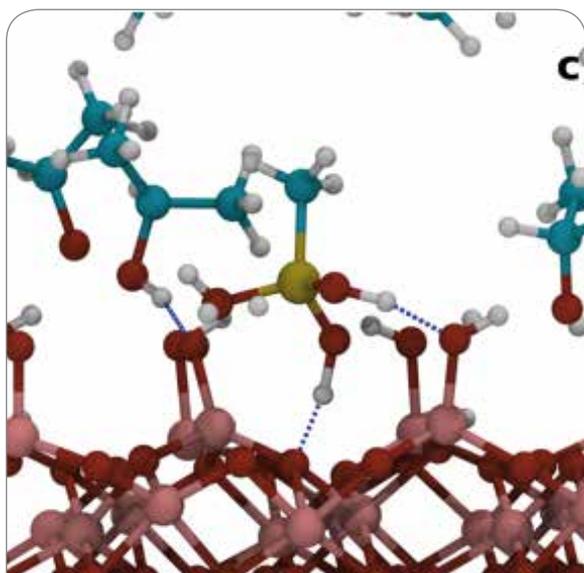


Fig. 2: Methylsilanetriol condensing on an aluminum oxide surface.

Paul Schwarz (group Prof. B. Meyer, FAU) used this CPMD version to study the reaction pathways of the condensation of methylsilanetriol on aluminumoxide (see Figure 2). Hannah Schlott from the same group developed a simulation protocol to generate amorphous surface structures of zinc oxide and other oxides by an annealing procedure. The generated structures will serve to study catalytic reactions on such unordered interfaces.

The third project by Niclas Siemer (group Prof. D. Marx, RUB) dug even deeper into the CPMD code. For the simulation of gold clusters on titanium dioxide surfaces he had to employ a so called Hubbard U correction to account for the poor description of the localized electrons in the d-shells of the titanium atoms by standard DFT functionals. The initial implementation of the U correction in CPMD, however, was developed for small metal clusters only, and never tested for larger systems. Correspondingly, it showed

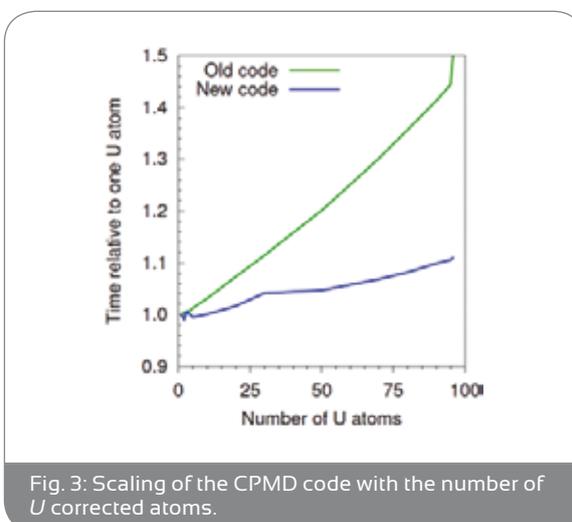


Fig. 3: Scaling of the CPMD code with the number of U corrected atoms.

a rather poor scaling behavior. During the first stage of the project he revised thread parallelized, and optimized the U correction code at the LRZ. Figure 3 shows that the originally more than 40% overhead for about a hundred U corrected atoms is reduced to well below 10%. Furthermore, the overall time per integration step on 20 nodes was reduced from 55 sec to below 11 sec, yielding a speedup of more than 5. With these improvements computing the dissociation energy of O_2 on the TiO_2/Au surface was feasible during the SoSI6 period.

Two other quantum chemistry projects focused on small metal containing clusters. Sascha Thinius (group Prof. T. Bredow, UB) computed the structure of chalcopyrite nano particles in order to improve the description of leaching processes of such copper ores. The particular focus of the project was selecting the optimal program and setup because even small models of such nanoparticles easily contain a few thousand atoms, for which the computation of a wave function is very demanding. Here, the GPAW program was shown to scale up to 900 cores and to outperform VASP by up to a factor of twenty.

To treat even larger metal oxide systems at high accuracy Martin Paleico (group PD J. Behler, RUB) is developing an interaction potential that is based on a neural network (NNP) for zinc oxide/copper systems. ZnO Cu is used as catalyst in the industrial synthesis of methanol and a better understanding of this reaction by simulation could help to improve its efficiency. To

reach the required accuracy of a few meV/atom the neural network has to be trained with energies for a large set ($\sim 10^5$) of different ZnO/Cu structures covering up to a few hundred atoms each. Generating these data was carried out on SuperMUC with the Vienna Ab-initio Simulation Package (www.vasp.at) using 20 nodes per structure. The resulting accuracy of the NNP is shown in figure 4 for the ZnO zincblende and wurtzite crystal structures.

Two projects originated from the biophysics/biochemistry community. Florian Kandzia (group Prof. Zacharias, TUM) analyzed the domain dynamics of the heat-shock protein Hsp90. This protein is an essential molecular chaperone that facilitates the folding of many other proteins and plays a vital role in the assembly of functional complexes. In order to bind to a variety of proteins its homologues adopt very different 3-dimensional conformations ranging from tightly closed to loosely connected dimers.

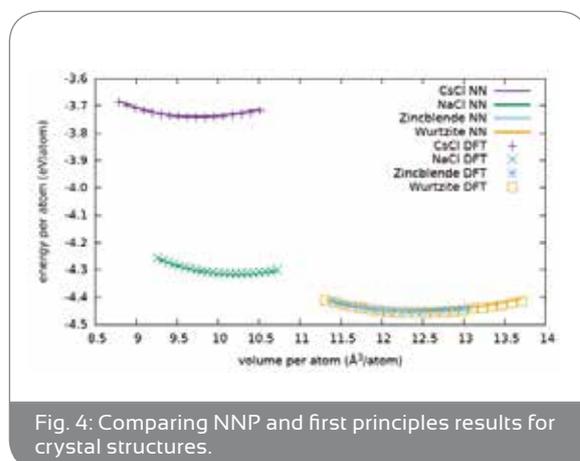


Fig. 4: Comparing NNP and first principles results for crystal structures.



Based on a pre-equilibrated structure, 560 initial conformations for a 2D umbrella sampling simulation scheme along a center-of-mass distance and center-of-mass torsion collective variable (see figure 5) were generated by preparatory simulations. Structural exchange between these windows by a replica exchange algorithm enhanced the sampling and convergence during the production run with the molecular dynamics package Amber16 (ambermd.org) on SuperMUC Phase 1. Using 80 cores per replica the resulting 44800 cores for each run cover approximately one third of the supercomputer partition, which is close to the maximally allowed job size during normal operation.



Fig. 5: Collective variables used for umbrella sampling. Purple, cyan and brown molecules correspond to different conformational states.

The biochemistry project by Sophie Mader (group Prof. Kaila, TUM) employed a new method for computational enzyme design. Here, amino acids of enzymes are mutated randomly,

the properties of the mutant are calculated and a Metropolis Monte Carlo procedure decides if the mutation is accepted or declined based on its catalytic activity. The target system to be enhanced by this QM/MM Monte Carlo method was the computationally designed enzyme “CE6” for the Diels-Alder reaction (see Figure 6). Mutations are performed by the tool VMD, then structures are relaxed using the molecular dynamics package NAMD (www.ks.uiuc.edu/Research/), and properties are calculated with the quantum chemical program Turbomole (www.turbomole.com).

The challenge of this project was to adapt the python-scripted workflow, which steers the different applications, to be used on SuperMUC Phase 1. NAMD ran efficiently on two nodes, whereas the quantum chemistry application

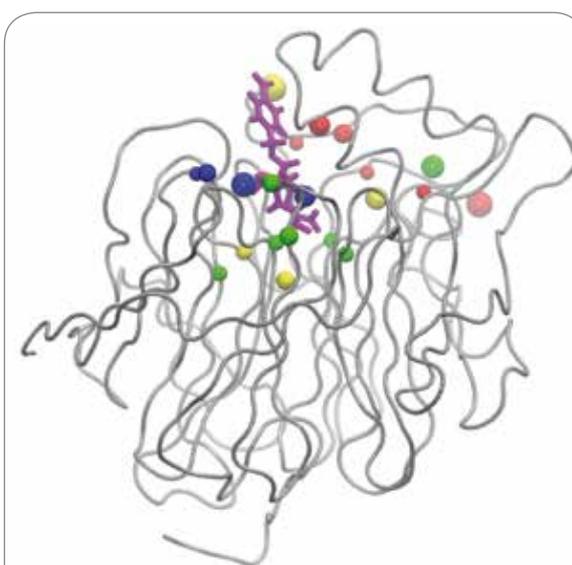


Fig. 6: Mutated residues around the active center of CE6.

Turbomole did not scale beyond node boundaries. Here, however, the two independent Turbomole calculations per mutation step were distributed on the two nodes. A further massive parallelism was achieved by running 200 Monte Carlo paths simultaneously, which totals to a job size of 6400 cores per run.

The largest systems were investigated by Zhuo-cheng Xie (group Prof. Bitzek, FAU) exploring the material properties of nanoporous gold. Nanoporous metals are a popular field in materials science due to their potential technological applications in actuation, catalysis and sensing. Based on experimental structures of nanoporous gold, a model system containing 450 million gold atoms was constructed and its material properties were investigated. Simulations with the LAMMPS molecular dynamics program were run on SuperMUC Phase 2 using up to four islands (2048 nodes). Figure 7 shows a snapshot of a simulation compressing the gold cluster in z-direction.

All projects were carried out by curious, industrious and eager students and it was a great pleasure for the tutors to work with them. Moreover, the close contact with the different projects showed hurdles and pitfalls, whose fixing improved the usability of SuperMUC in general. Since almost all groups are currently working

on follow up projects, the SoS program helps to promote life science applications on SuperMUC. Currently, the biolab@lrz is preparing a new SoS call for 2017.

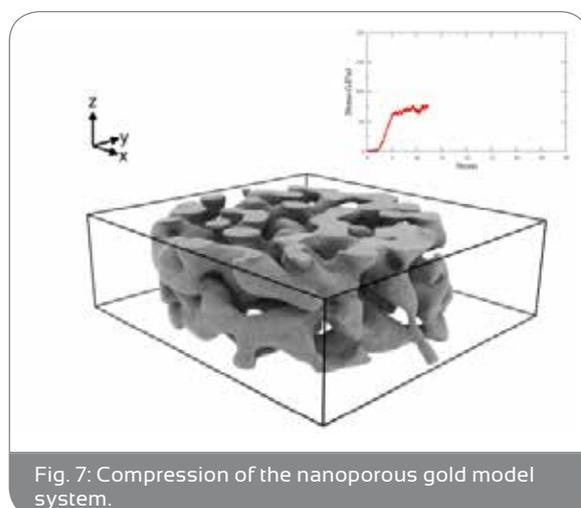


Fig. 7: Compression of the nanoporous gold model system.

Kind support by the SuperMUC steering committee, Prof. Wellein in particular, is gratefully acknowledged.

Written by Gerald Mathias
Leibniz Supercomputing Centre (LRZ)

Contact: Gerald.Mathias@lrz.de

Supercomputing in Scientific & Industrial Problems (SSIP) 2017



2nd German-Russian Conference took place at HLRS March 27 until March 29.

For the second time scientists from Russia and Germany met to discuss supercomputing in a both scientific and industrial context. The conference was organized by HLRS and the Keldysh Institute of the Russian Academy of Science. About 70 scientists met at HLRS to focus on all issues of supercomputing ranging from mathematical problems to issues of programming of large scale systems and to real world applications of simulation on supercomputers.

After a warm welcome by Prof. Michael Resch from HLRS and Prof. Boris Chetverushkin from the Russian Academy of Science the first day was devoted to real world applications. On the second and the third day mathematical and computer science issues were addressed but also industrial usage of HPC in a variety of fields. As part of the workshop a tour through the computer room of HLRS with its fastest European system gave participants a view of the hardware aspects of high performance computing. This was complemented by a visit to the HLRS virtual reality environment allowing



Dr. Orlov, Peter the Great St. Petersburg Polytechnic University.



Prof. Dr. Chetveruschkin, Keldysh Institute of Applied Mathematics, Moscow

participating scientists to see how simulation results are turned into visual representations enabling insights from numbers.

The second German Russian workshop helped to deepen the collaboration of scientists from both countries. The organizers agreed to plan for a next workshop to be held again in Russia in 2018. All participants are looking forward to meet again next year.

Written by Christopher Williams

High Performance Computing Center Stuttgart (HLRS)

Contact: Natalia Currle-Linde, linde@hlrs.de



Dr. Titarev, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

The 16th Annual HLRS/hww Workshop



The 16th annual HLRS/hww Workshop on Scalable Global Parallel File Systems “Memory Class Storage” took place from April 3rd to April 5th at HLRS in Stuttgart. Once again about 65 participants met to discuss recent issues and new developments in the area of parallel file systems and the required infrastructure and devices around.

Prof. Michael Resch, Director HLRS, opened the workshop with a greeting address in the Rühle Saal of the recently opened HLRS training centre on Monday afternoon.

In the first presentation, Felix Hupfeld, Quobyte, gave an introduction and a technical deep

dive into the newly developed Quobyte file system. Coming from the object storage world but adopting features from high performance file systems like ExtremeFS this new storage system wants to bring together both worlds. Peter Braam, well-known in the community as inventor of the Lustre file system, presented his new work on campaign storage. Representing the new start-up with a similar name, he showed his ideas and plans for a new storage paradigm for HPC centres. Accelerating storage has been the main topic of Robert Haas of IBM. In his presentation, he covered several topics in this area: from a report on Spectrum Scale (formerly known as GPFS) on the upcoming Coral system over the optimization of the data analytics stack



to the development of cognitive storage where machine learning is used to choose the best storage class for a data object.

For the main focus of the workshop, Torben Kling-Petersen of Seagate discussed the issues of burst buffers and introduced a development for flash acceleration especially for mixed workloads. Further discussions on Memory Class Storage have been provided by Jan Heichler, DDN, and Erich Focht, NEC, both talking about its opportunities and potential differences to current approaches.

Johann Lombardi, Intel, introduced the current developments around DOAS, the new storage paradigm for the exascale era. Johann introduced the storage model and showed the different available features. The newest information around BeeGFS was provided by Franz-Josef Pfreundt. In his talk, he showed different new features and different set-ups of the file system in the field. In addition, he showed how it can support distributed deep learning. This was followed by a presentation on a HSM solution for Lustre and BeeGFS, which has been the topic of Ulrich Lechner, GrauData.

Covering the network part of the workshop, Jerry Lotto, Mellanox, gave an update on high performance networking technology and he showed current developments which will specifically support storage requirements. Covering

sustainability aspects, Klaus Grobe covered the energy efficiency challenge for ICT and provided insight into the solution approaches at Adva.

Furthermore, Lenovo, Dell EMC and NEC storage experts gave insight into their solutions for high performance storage systems.

In the academic-research-oriented session, Eugen Betke of DKRZ presented their work on an MPI-IO in memory driver for the Kove XPD. A new solution for I/O optimization has been introduced by Xuan Wang, HLRS. In his work, he uses a machine learning approach to transparently set file system and I/O parameters to increase I/O performance and such to optimize application run time.

HLRS appreciates the great interest it has once again received from the participants of this workshop and gratefully acknowledges the encouragement and support of the sponsors who have made this event possible.



The 64th IDC HPC User Forum



The 64th IDC HPC User Forum took place on February 28th/March 1st 2017 at HLRS premises in Stuttgart. The agenda covered a variety of valuable talks, from political, technical and business viewpoints. With an impressive number of over 105 registrations, a broad and diverse audience of HPC stakeholders shared information and viewpoints about the latest developments in HPC and its communities.

After an official opening by Peter Castellaz, from the Ministry of Science, Research and Arts of Baden-Wuerttemberg, Gustav Kalbe, Head of the High Performance Computing & Quantum Technologies Unit at the European Commission (EC), provided insights into the HPC and Quantum Strategy of Europe, with a certain focus on the Exascale aims and activities now and planned in the future.

This was supported by a presentation of Michael Malms from the European Technology Platform for High Performance Computing (ETP4HPC) who gave a comprehensive overview of the mission of the ETP, the setup of partners and the strategic research agenda, which is provided as input for the European HPC strategy.

In the following afternoon sessions, updates on the Gauss Centre for Supercomputing (GCS, presented by Dr. Claus-Axel Müller), the Leibniz Supercomputing Centre (LRZ, presented

by Professor Arndt Bode), the Jülich Supercomputing Center (JSC, presented by Professor Thomas Lippert) and HLRS (presented by Dr. Bastian Koller) were given, also providing insight into the planned next activities with regard to new technology.

After a short break, the focus was directed on the end-users and application perspective, when Dr. Bastian Koller presented the plans for a European Center of Excellence in Engineer-



ing and Professor Mark Parsons from the Edinburgh Parallel Computing Centre presented the Fortissimo EU initiative for supporting Small and Medium Enterprises with HPC:

The first day concluded with a presentation from Jürgen Kohler, from Daimler, presenting how digitalization is used in context of the



Mercedes-Benz Car Research and Development.

In the morning of the second day, IDC presented their recent market updates and research findings before Professor Michael Resch, Director of HLRS, gave his perspective on the HPC Community.

Andreas Wierse from the SICOS GmbH then took the token to elaborate how HPC can be provided to SMEs. Following a short coffee break, Peter Hopton from Iceotope presented the activities and results of the EU ExaNeSt

project, developing solutions for Interconnection Networks, Storage, and Cooling.

The final presentations of the 64th HPC User Forum were provided by Arno Kolster from PayPal about "HPC for Advanced Business Analytics", followed by the meeting wrap up by Earl Joseph and Steve Conway.



Written by Basti Koller

High Performance Computing Center Stuttgart (HLRS)

Contact: koller@hlrs.de

8th Blue Gene Extreme Scaling Workshop



From 23 to 25 January 2017, Jülich Supercomputing Centre (JSC) organized its eighth IBM Blue Gene Extreme Scaling Workshop. The entire 28-rack JUQUEEN Blue Gene/Q was reserved for over 50 hours to allow six selected code teams to investigate and improve the scalability of their applications. Ultimately, all six codes managed to run using the full complement of 458,752 cores (most with over 1.8 million threads), qualifying two new codes for the High-Q Club [1,2].

The MPAS-A multi-scale non-hydrostatic atmospheric model (from KIT & NCAR) and “pe” physics engine rigid body simulation framework (from FAU Erlangen-Nürnberg) were both able to display good strong scalability and thereby qualify for High-Q Club membership. Both exclusively used MPI parallelization, with the latter demonstrating strong and weak scalability to over 1.8 million processes in total. Available compute node memory and the lack of support for nested OpenMP parallel regions limited MPAS-A to a single MPI task per core, however, substantial code improvements in the two years since its first workshop participation (particularly the use of SIONlib for massively-parallel file reading and writing), combined with a larger 2km cell-size global mesh (147 million grid cells), were key success factors.

ParFlow (developed by the University of Bonn and FZJ-IGB) is an integrated hydrology model simulating saturated and variably saturated subsurface flow in heterogeneous porous media that had recently demonstrated how

improvements to its solver, coupling with the p4est parallel mesh manager, allowed it to scale to the 458,752 cores of JUQUEEN to qualify for High-Q Club membership. During the workshop the focus was investigating the performance of the writing of output files with its SILO library, which was requiring prohibitive amounts of time for larger numbers of MPI processes.

KKRnano is a Korrington-Kohn-Rostocker DFT/Green’s function simulation of quantum nano-materials from FZJ-IAS which is being extended to support a million atoms. While solver components were found to perform acceptably, Fortran direct access file I/O impeded overall scalability. The performance of the latest version of the CPMD (Car-Parinello Molecular Dynamics) code with a large 1500-atom organic-inorganic hybrid perovskite system was also investigated by a team from RWTH-GHI and FZJ-IAS/INM.

The final code was a prototype multi-compartment neuronal network simulator designed for massively-parallel and heterogeneous architectures, NestMC (JSC SimLab Neuroscience). Implementations using 64 OpenMP or C++ threads per MPI process were compared and weak-scaling limitations identified.

Detailed workshop reports provided by each code-team, and additional comparative analysis to the 28 High-Q Club member codes, are available in a technical report [3]. The participants greatly appreciated the opportunity to have dedicated access to the entire JUQUEEN

system over the three day period to investigate their applications' performance and scalability. The workshop was immediately followed by the second "Big Blue Gene Week" dedicated to exploiting JUQUEEN for capability computing jobs, including extreme-scale atmospheric science, materials science and neuroscience simulations by High-Q member codes and prior scaling workshop participants.

References

- [1] **The High-Q Club:**
<http://www.fz-juelich.de/ias/jsc/high-q-club>
- [2] **Brömmel, D., Frings, W., Wylie, B.J.N.:**
Extreme-scaling applications en route to exascale, Proc. Exascale Applications and Software Conference 2016, ACM, 2016
<http://dx.doi.org/10.1145/2938615.2938616>
- [3] **Brömmel, D., Frings, W., Wylie, B.J.N. (eds.):**
JUQUEEN Extreme Scaling Workshop 2017, Tech. Report FZJ-JSC-IB-2017-01
<https://juser.fz-juelich.de/record/828084>



The PRACE 2017 Winter School “Fuelling Scientific Discovery with HPC Infrastructure” [1] took place at Tel Aviv University, Tel Aviv, Israel from February 6-9, 2017. The school was organized by the Inter-University Computation Center (IUCC) and attended by 58 participants from Israel, Spain, Italy, Turkey and Austria. Four lecturers from Israel and three from the Jülich Supercomputing Centre provided expert knowledge on a range of topics.

The emphasis of the school program was to provide a thorough introduction and overview of how to use PRACE resources and to raise awareness of the value of the resources available to local and regional users.



The Porter School of Environmental Studies Building, Tel Aviv Israel, a LEED [2] Platinum certified building, hosted the PRACE 2017 Winter School.

The school started out on Monday morning with a general introduction to parallel computing including getting started guides to MPI and OpenMP by Guy Tel-Zur. In the afternoon, the participants learned why they should pay attention to parallel

IO and how to take advantage of it in their own programs from Sebastian Lührs.

A welcome dinner at the ‘Milk & Honey Whisky Distillery’ [3], Israel’s first whisky distillery, on Monday evening, gave the international participants, lecturers, and the school’s program organizers the chance to get to know each other and learn about whisky distilling and tasting. It included a a tour of the distillery and a tasting of some of the spirits produced.

The session on parallel IO continued on Tuesday morning and was followed by a introduction to parallel performance analysis and tuning by Bernd Mohr. On Wednesday the focus moved to many-core architectures such as GPUs and Xeon Phi Knights Landing. Both of these sessions were taught by Jan Meinke. The Thursday morning session approached GPUs from an algorithmic point of view presented by Dennis Rapaport. The school concluded with two parallel sessions on LAMMPS (Dan Mordehai) and multiprocessing in Python (Mordechai Butrashvily). Each session consisted of a mix of lectures and hands-on exercises for which the participants were granted access to computing resources in Tel Aviv and Jülich.

References

- [1] **PRACE 2017 Winter School “Fuelling Scientific Discovery with HPC Infrastructure”:** <https://events.prace-ri.eu/event/549/>
- [2] **LEED:** <http://leed.usgbc.org/leed.html>
- [3] **Milk & Honey Whisky Distillery:** <http://mh-distillery.com>

Written by Jan H. Meinke

(Jülich Supercomputing Centre (JSC), Germany)

Contact: Jan Meinke, j.meinke@fz-juelich.de

2nd EasyBuild User Meeting: A Growing Community Gathers at JSC

Installing software on supercomputers is a significant burden for user support groups in many research centres. In recent times, open source projects have been created to help alleviate this. One such project, EasyBuild, enables the installation and presentation of a coherent stack of scientific software in an automated, reproducible way. It has grown into a thriving community and currently supports over 1000 software packages. JSC has embraced the project and become a core part of its community. As a result, JSC has hosted the 2nd EasyBuild User Meeting in February. In this event participants of 21 different institutions shared ideas as well as development and implementation strategies during the 3 days event, in a very successful meeting.

As supercomputing becomes more and more ubiquitous in a growing variety of research fields, the community of users expands and becomes more diverse. The direct consequence of this fact is the larger amount of software requested by HPC users, and larger variability in requirements between communities. Maintaining a software stack in 2017 is significantly more difficult than it was 10 years ago. Noticing this situation, Ghent University developed EasyBuild, a package manager for scientific software.

EasyBuild was developed with HPC centers in mind. It provides support for over 1000 software packages, and ensures that the compilation and module generation are done in a reproducible



Fig. 1: On Friday just the participants involved in EasyBuild development stayed to share ideas during the Hackathon.



manner. At JSC, EasyBuild became a core part of their strategy to maintain software on their clusters. As the first early adopter among large research centers, JSC also became a core part of the EasyBuild community.

With this in mind, it was natural for JSC to host the 2nd EasyBuild User Meeting. The event had over 35 attendees -with 14 different nationalities- from 21 different international institutions. During the meeting 10 different presentations -including 2 remote- were broadcast live over the internet. Among these presentations the participants could see how CSCS uses EasyBuild in their production Cray system, based on the work presented in [1], and how JSC manages its whole software infrastructure, based on the work presented in [2].

The 3 day event included one and a half days for a "hackathon". The target of these sessions was clearly focused on developing new features that can benefit the wider community. Members proposed ideas in a roundtable discussion, and created teams of collaborators with common interests.

The meeting itself was a very successful event with a forward looking perspective, including adding increased support for site customisations, new file formats to allow deeper collaboration on specific software packages and support

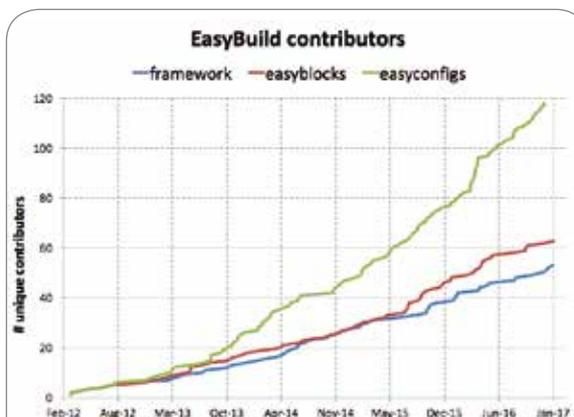


Fig. 2: The growth and engagement of the community is obvious when looking at the number of unique collaborators.

for new software packages and streamlined workflows. All participants were exposed to ideas that can benefit their institutions and the users of their systems once these ideas -both existing and to-be-developed- get deployed.

Institutions involved:

EMBL, CSCS, Compute Canada, Free University Brussels, Ghent University, IDRIS, Illumina, JSC, RWTH, University of Hanover, New York University Abu Dhabi, Ottawa Hospital Research Institute, STFC, TACC, The Francis Crick Institute, Universite Catholique de Louvain, University of Liege, University of Birmingham, University of Michigan, University of Muenster, University of Namur

References

- [1] Forai, P.; Hoste, K.; Peretti-Pezzi, G.; Bode, B.: Making Scientific Software Installation Reproducible On Cray Systems Using EasyBuild, Cray User Group Meeting 2016, London, England, 2016
- [2] Alvarez, D.; O'Cais, A.; Geimer, M.; Hoste, K.: Scientific Software Management in Real Life: Deployment of EasyBuild on a Large Scale System, 3rd International Workshop on HPC User Support Tools (HUST), Salt Lake City, UT, USA, pp. 31-40, 2016

Written by Damian Alvarez and Alan O'Cais
(Jülich Supercomputing Centre (JSC), Germany)

Contact: Damian Alvarez, d.alvarez@fz-juelich.de, Alan O'Cais, a.ocais@fz-juelich.de

Intel MIC Programming & High Performance Computing in Atmosphere Modelling and Air Related Environmental Hazards @ IT4Innovations, Czech Republic

For the third time the Czech-Bavarian Competence Centre for Supercomputing Applications, CzeBaCCA, has organised a technical Intel Many Integrated Core (MIC) programming workshop combined with a scientific workshop about HPC simulations in the field of geo- and environmental sciences. The Czech-Bavarian Competence Centre was established in 2016 by the Leibniz Supercomputing Centre (LRZ), the Department of Informatics of the Technical University of Munich (TUM) and the IT4Innovations National Supercomputing Centre of the Czech Republic to foster the Czech-German collaboration in high performance computing. One of the main objectives of the Competence Centre is to trigger new collaborations between Germans and Czechs via scientific workshops and many-core architecture specific technical trainings.

The successful series of workshops was initiated in February 2016 with an introductory MIC programming workshop and a symposium on “SeisMIC – Seismic Simulation on Current and Future Supercomputers” at IT4Innovations (see inSiDE Vol. 14 No. 1 p. 76ff, 2016). The series was continued with an extended MIC workshop concentrating on simulations on the Intel Xeon Phi based SuperMIC cluster at LRZ combined with a scientific workshop on “High Performance Computing for Water Related Hazards” in June 2016 at LRZ (see inSiDE Vol. 14 No. 2 p. 25ff, 2016).

The third edition of this workshop series took place on 7 – 9 February 2017 at IT4Innovations, again, and combined a 2-day Intel Xeon Phi



Fig. 1: Participants of the “Intel MIC programming workshop” at IT4Innovations, February 7 – 8, 2017. The main lecturers in the first row (Dr. B. Jansik (IT4Innovations), Dr. V. Weinberg, Dr. M. Allalen (LRZ), f.l.t.r.) are holding Intel Xeon Phi coprocessors.



Fig. 2: Participants of the symposium “High performance computing in atmosphere modelling and air related environmental hazards” at IT4Innovations, February 9, 2017.

programming workshop with a one-day scientific workshop. The MIC programming workshop focused on the Salomon supercomputer at IT4Innovations, currently one of the largest Intel Xeon Phi based systems in Europe. The workshop was offered within the PRACE Advanced Training Centre (PATC) curriculum of LRZ and attracted around 25 participants from Austria, Germany, Poland, Slovakia and the Czech Republic. The topics of the first one and a half days reached from an introduction about the Salomon supercomputer at IT4Innovations and the Intel MIC architecture, various Intel Xeon Phi programming models (native mode, off-loading, MKL, OpenMP, MPI etc.) to advanced topics about vectorisation and performance

optimisation, interleaved with many hands-on sessions on the Salomon system at IT4Innovations. During a public half-day plenary session five invited speakers from TU Czestochowa, Brno University of Technology and IT4Innovations talked about MIC experience and best practice recommendations using Intel Xeon Phi based systems. The speakers discussed the acceleration of scientific applications like the ESPRESO domain decomposition library, the Blender Cycles Render Engine or the k-Wave toolbox.

The Intel MIC programming workshop was followed by a one-day symposium on “High performance computing in atmosphere modelling



IT4Innovations
national
supercomputing
center



and air related environmental hazards". Nine presenters addressed recent topics in atmospheric modelling, simulation of air pollution, spread of hazardous substances, weather and climate prediction, and global navigation satellite systems. A special focus was on demands and desired features of (future) simulation software, parallelisation for current and novel HPC platforms, as well as establishing scalable simulation workflows on supercomputing environments. Furthermore, partnership initiatives of LRZ in environmental computing were presented. The next series of workshops will take place at LRZ on 26 – 30 June 2017 and will combine a 3-day Intel MIC Programming workshop concentrating on the 2nd generation of the Intel Xeon Phi processor named "Knights Landing" with a symposium on "HPC for natural hazard assessment and disaster mitigation".

Acknowledgements

The Czech-Bavarian Competence Centre for Supercomputing Applications is funded by the Federal Ministry of Education and Research. The Intel MIC programming workshop was financially also supported by the PRACE-4IP project funded by the European Commission's Horizon 2020 research and innovation programme (2014-2020) under grant agreement 653838.

References

- [1] <https://www.lrz.de/forschung/projekte/forschung-hpc/CzeBaCCA/>
- [2] <http://www.training.prace-ri.eu/>
- [3] <https://www.lrz.de/services/compute/courses/>
- [4] <http://training.it4i.cz/kurzy-it4innovations/>

Written by Michael Bader,
(Technical University of Munich (TUM), Germany)

Momme Allalen, Arndt Bode, Anton Frank, Megi Sharikadze, Volker Weinberg,
(Leibniz Supercomputing Centre (LRZ), Germany)

Ondřej Jakl, Branislav Jansík, Martin Palkovic, Vít Vondrák
(IT4Innovations, Czech Republic)

Contact: Volker Weinberg, Volker.Weinberg@lrz.de



Congratulations to Teams PhiClub and segFAUlt for "A Job Very Well Done"!

The 2016 edition of the Supercomputing Conference (SC) saw the hitherto largest number of student teams participating in the Student Cluster Competition (SCC), which for the 10th time was included in this annually recurring international High Performance Computing (HPC) event. In total, 14 teams had been accepted for this year's multi-disciplinary HPC challenge for undergraduate students, and among them—the only European participants—were two teams from Germany:

- Team **PhiClub** of the Technical Universität München (TUM), and
- Team **segFAUlt** representing the Friedrich-Alexander Universität Erlangen-Nürnberg (FAU)

Although none of the German representatives made it onto the podium, the captains



of both team PhiClub and team segFAUlt asserted that participating in this international challenge and mastering „stress situations which could easily be compared to real working scenarios“ has been an awesome experience to be remembered by all team members alike.

Evidence of their ability to cope with extreme pressure provided e. g. outstanding results achieved by both teams in two particular competition components:

a) Scientific Reproducibility, a brand new component added to the SCC challenge: Here, students were requested to reproduce a research paper from SC15 rather than focus on throughput of prescribed data sets. Team segFAUlt secured 100 of 100 possible points and team PhiClub achieved 95/100.

b) The Interview: The young students from TU München, most of them being only in their 3. semester, excelled in deeply impressing the



Team PhiClub of TUM (left) and team segFAUlt of FAU (right) hard at work in their SCC booths ...

critical HPC jury with their demonstration of overall HPC knowledge. This earned them the highest possible score in this competition component, and team segFAULT came in a no-less respectable third.

China's University of Science and Technology received top honors in this year's SCC, winning both categories: best performance of the LINPACK benchmark application and best overall team performance. The 11 teams, completing the round of this year's SCC participants, represented the following universities:

- Huazhong University of Science and Technology, China
- Nanyang Technological University, Singapore
- National Tsing Hua University, Taiwan
- Northeastern University, Auburn University, United States
- Peking University, China
- San Diego State University, United States
- Boston University, UMass Boston, MIT, United States
- Universidad EAFIT, Colombia
- University of Illinois Urbana-Champaign, United States
- University of Texas at Austin, Texas State University, United States
- University of Utah, United States

The Gauss Centre for Supercomputing has proudly acted as financial sponsor of both team PhiClub and team segFAULT for SCC 2016. The dedication and commitment of the students shown in the course of the preparation for the competition and throughout the challenge are much appreciated and deeply honoured. Congratulations to all team members and their mentors for a job very well done!



Written by Regina Weigand
(GCS Public Relations).

Contact: Regina Weigand, r.weigand@gauss-centre.eu

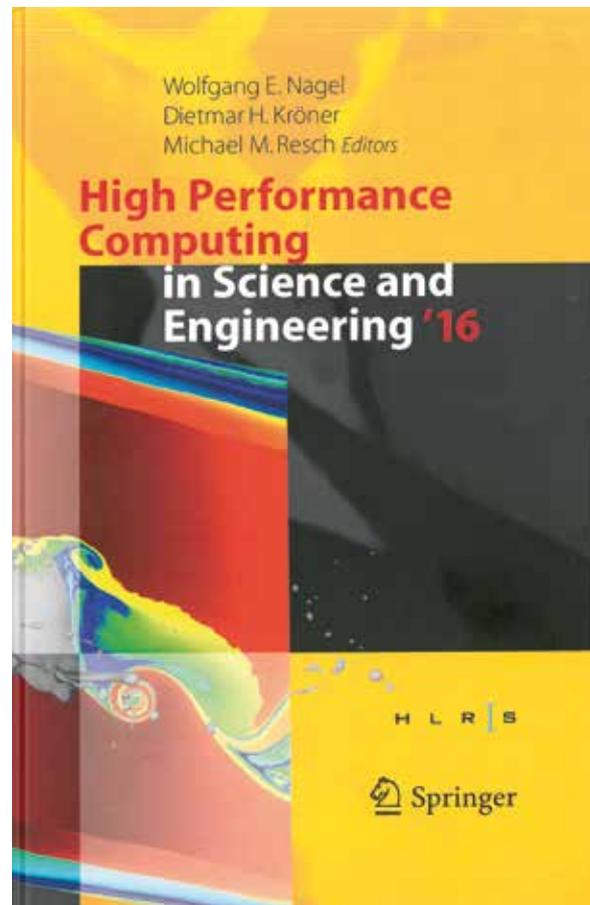


Scientists of the HLRS, Albert-Ludwig-Universität Freiburg, and the Technische Universität Dresden published a new book: High Performance Computing in Science and Engineering '16. The book covers all fields of computational science and engineering ranging from computational fluid dynamics to computational physics and from chemistry to computer science with a special emphasis on industrially relevant applications. The basis for the report were the latest advances in this field discussed at the 19th Results and Review Workshop of the HLRS in October 2016 in Stuttgart, Germany. Presenting findings of one of Europe's leading systems, this volume treats a wide variety of applications that deliver a high level of sustained performance.

The conference proceeding covers the main methods in High Performance Computing. Its outstanding results in achieving the best performance for production codes are of particular interest for both scientists and engineers.

Transactions of the High Performance Computing Center, Stuttgart (HLRS) 2016

Herausgeber: Nagel, Wolfgang E., Kröner, Dietmar H., Resch, Michael M. (Eds.)



Book: <http://www.springer.com/de/book/9783319246314>

eBook: <http://www.springer.com/de/book/9783319470658#otherversion=9783319470665>





Applications

In this section you will find the most worth mentioning applications of external users on GCS computers.

Smoothed Particle Hydrodynamics for Numerical Predictions of Primary Atomization



Fig. 1: Sequence of a bag breakup event. The time increment between two consecutive images is 75 microseconds.

The design goal for future propulsion systems of civil aircrafts is to minimize the environmental impact as well as to optimize the economic benefit. At least for long range flights, liquid fuels will continue to be the preferred and most appropriate energy carrier. Minimizing harmful emissions and, at the same time, maximizing the efficiency of an aircraft engine may be thermodynamically contradictory. The key technology which enables both, low emissions and high efficiencies, is a well controlled combustion process. Controlling combustion mainly comprises a well defined fuel placement inside the combustion chamber. However, the simulation of fuel atomization was not feasible due to the enormous computational requirements of this multi-scale problem and the accuracy required for a correct handling of multiphase flows. Presently, sophisticated combustion simulations still rely on fuel droplet starting conditions, which are a rather rough estimate of the real spray properties. Furthermore, the physical effects of air assisted atomization are not understood in detail.

The present numerical simulation is the first of its kind in which the qualitative and quantitative

features of an engine-relevant air-assisted atomizer were successfully predicted. The vast experimental data-basis for validation and comparison has been provided by Gepperth (2016) [1].

Numerical method

Due to severe shortcomings of commonly used Eulerian simulation techniques for multiphase flows, we developed the parallel flow solver *super_sph*, which is based on the Smoothed Particle Hydrodynamics method. Due to its Lagrangian description of the flow, mass is strictly conserved. Interface diffusion, which is a big shortcoming of grid-based methods when it comes to the treatment of violent multiphase flows, is not an issue. Due to the self organization of the numerical discretization points (referred to as particles), the shape and connectivity of liquid structures is not affected by their position within the computational domain nor by any grid-related preferential directions.

Being a relatively new numerical method outside of astrophysical applications, the experience with SPH for common multiphase simulations was limited due to a lack of functionality.

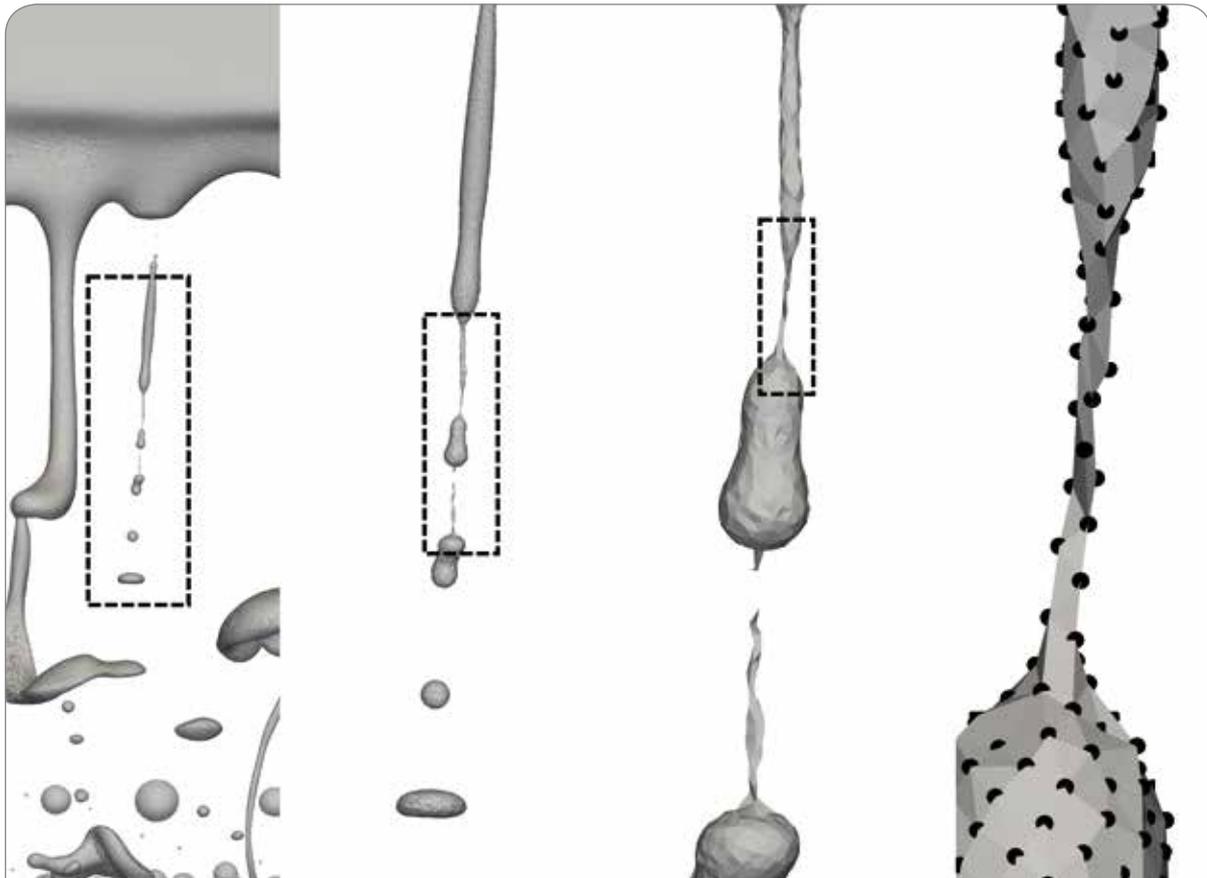


Fig. 2: Close-ups of a Rayleigh breakup event. The height of the left-most image is 3.8 mm, the height of the right-most image is 90 microns. The self organizing behavior of the discretization points (black dots) allows to discretize quasi-1D liquid structures.

In the past 4 years, the ITS (Institut für Thermische Strömungsmaschinen) has successfully addressed the missing capabilities. Arbitrary periodic boundary conditions, translational and rotational, as well as robust inlet and outlet boundary conditions have been developed [2]. Furthermore, an accurate wall-wetting behavior has been implemented by slightly adapting the commonly used surface tension model [3]. Density ratios of 1000 can be handled without any numerical stability issues, which is

indispensable for imposing physically correct fluid properties.

Beside the physically (sufficiently) correct modeling of the involved fluid flows and the code functionality, performance is a crucial feature when it comes to the simulation of atomization phenomena. This is due to the fact, that the length scales to be resolved cover at least 4 orders of magnitude in time and space. The diameter of the smallest droplets is (supposed

to be) in the range of 1 micron, the nozzle length scale is in the order of centimeters. It is worth to mention, that the smallest length scales to be resolved are not defined by turbulent eddies of the gaseous flow, but by the size and curvature of the small liquid structures. The simulation to be presented constitutes a DNS (Direct Numerical Simulation), turbulence is handled directly without the need of modeling it.

Computational aspects

The code framework `super_sph` is parallelized via MPI and a proper domain decomposition scheme. Due to the weakly compressible explicit formulation, a very low inter-processor communication footprint can be achieved. This leads to an excellent parallel performance. Strong scalability tests have been performed over 3 orders of magnitude. Even with domain sizes reaching less than 1500 particles per core, the efficiency remains above 0.6 compared to the baseline performance. In order to enhance the serial code performance, emphasis has been put on cache efficiency. This is mainly due to the fact, that vectorization is quite challenging regarding the arbitrary connectivity of the discretization points. An efficient cache usage, optimized data structures and access patterns allow to remedy the low level of vectorization. The absolute code performance is measured in terms of particle-iterations per CPU-second. On the ForHLR II cluster, we obtain more than 144000 particle iterations per CPU-second using 2560 compute cores and more than 123000 particle iterations per CPU-second using 10000 cores. Typical values of grid based multiphase methods

(Level Set, Volume of Fluid) are in the range of 4000 to 7000 cell iterations per CPU-second for comparable flow configurations. As surface tension effects limit the admissible time step, incompressible grid based methods and the weakly compressible SPH method are limited to the same time stepping if the spatial resolution is identical. Using SPH, the time required for solving large-scale multiphase problems can be reduced by a factor of 20 to 50 compared to established grid-based Finite Volume methods.

Results

A 3D simulation of a small section of a planar prefilming air-blast atomizer has been set up, for which a vast experimental database exists. In the experiment, an airfoil shaped prefilmer is exposed to an air stream. At the upper side of the prefilmer surface, a liquid film is fed through small drill holes. High aerodynamic forces push the liquid film to the trailing edge of the prefilmer. Here, the liquid accumulates and forms flapping ligaments, which finally detach from the prefilmer lip. The numerical domain covers the region in the vicinity of the trailing edge. At the inlet, the air velocity is prescribed by a piecewise defined profile with a maximum velocity of 50m/s. The domain size is approximately 6x6.23x4mm. An inter-particle spacing of 5 microns yields 1.2 billion particles. The domain has been decomposed and distributed onto 2560 processor cores. Within 60 days on the compute cluster ForHLR I, 1.1 million time steps could be completed, which corresponds to 14.6 milliseconds of physical time. About 1100 time steps have been dumped to the file system. A



Fig. 3: Development and deflection of the gaseous boundary layer. Volume rendering of the mapped particle IDs, 360 million voxels are used for visualization.

data size of 62GB per time step yields an overall required disk space of nearly 70TB.

For post-processing, the Lagrangian nature of the SPH method allows to filter the data by e.g. the fluid type. Hence, for the present multiphase simulation, it is possible to reduce the data size by approximately 99% if only the liquid phase is of interest. The resulting small datasets, consisting of only 10 to 20 million particles per time step, can be handled, visualized and further processed on usual desktop computers. In order to visualize the surface of the liquid structures, an interface reconstruction scheme based on the alpha-shape algorithm is applied. This further reduces the required computational resources and allows for an interactive analysis of the data.

Droplet statistics are extracted by applying a Connected Component Labeling technique.

Regarding the qualitative results of the simulation, all phenomena were predicted as observed in the experiment. The liquid fuel accumulates at the trailing edge of the prefilmer, forms flapping ligaments and finally is atomized. There exist two main breakup mechanisms. The first one is the disintegration of thread-like filaments, which can be identified as Rayleigh breakup. The second one is the formation and blow-up of bag shaped structures which finally burst. During the burst, many very small droplets are generated by the disintegration of the liquid skin of the bubble. Within the simulated period of time,

two main bag breakup events could be detected. In Figure 1 a sequence of such a bag breakup is depicted, where the time increment between two consecutive images is 150 microseconds. The gaseous phase is not depicted. The image section covers the region downstream the atomizer trailing edge. Like in the experimental observations, the different breakup modes coexist, each one generating different droplet size spectra and trajectories.

A more detailed example of a Rayleigh breakup is visualized in Figure 2, where different close-ups are depicted. The highest zoom-level clearly reveals the self-organizing behavior of the numerical discretization points. Shortly before the breakup, a quasi-1D filament is formed and

connects the droplets which are about to pinch off. The breakup of these thin filaments eventually creates very small satellite droplets.

For the qualitative analysis of the gaseous phase, SPH offers an easy way to derive Lagrangian Coherent Structures (LCS) without further post-processing steps. By simply assigning every particle a unique ID, vortices, residence times or the degree of mixing can be easily visualized and accessed. In Figure 3, the gaseous phase stemming from the inflow boundary layer is depicted. The system of vortices appearing downstream of the prefilmer lip is clearly visible. It is mainly influenced by the liquid structures, which are still attached to the trailing edge or which are about to disintegrate. The vortices itself generate strong velocity gradients, which interact with the fuel droplets and eventually trigger further breakup events. For the visualization of the gaseous phase in Figure 3, particle data has been mapped onto a Cartesian grid, where the grid spacing corresponds to 1.5 times the mean particle distance. The resulting grid consist of 360 million cells, and can be handled and visualized on a desktop computer with only 16GB of RAM.

Regarding the quantitative results of the spray analysis, the characteristic droplet diameters DV10, DV50 and DV90 and the mean diameter

D32 coincide with the experimentally obtained data with a maximum deviation of 10%. However, due to the very limited amount of simulated breakup events and due to the high numerical costs, a statistically converged comparison with experiments will not be possible in the near future.

Acknowledgements

This work was performed on the computational resource ForHLR I and (partly) ForHLR II funded by the Ministry of Science, Research and the Arts Baden-Württemberg and DFG ("Deutsche Forschungsgemeinschaft"). The authors would especially like to thank the helpful and patient support offered by the Steinbuch Centre for Computing at KIT and the excellent performance and visualization related workshops at HLRS.

References

- [1] **Gepperth, S., Koch, R., and Bauer, H.-J.:**
„Analysis and comparison of primary droplet characteristics in the near field of a prefilming airblast atomizer“. In: ASME Turbo Expo 2013: Turbine Technical Conference and Exposition. American Society of Mechanical Engineers, 2013.
- [2] **Braun, S., Wieth, L., Koch, R., and Bauer, H.-J.:**
„A Framework for Permeable Boundary Conditions in SPH: Inlet, Outlet, Periodicity“. In: 10th International SPHERIC Workshop, 2015.
- [3] **Wieth, L., Braun, S., Koch, R., and Bauer, H.-J.:**
„Modeling of liquid-wall interaction using the Smoothed Particle Hydrodynamics (SPH) method“. In: ILASS Europe, 26th Annual Conference on Liquid Atomization and Spray Systems, 2014.

Written by **Samuel Braun, Rainer Koch, Hans-Jörg Bauer**

Institut für Thermische Strömungsmaschinen, Karlsruher Institut für Technologie

Contact: Samuel Braun, samuel.braun@kit.edu



Large Scale I/O of an Open-Source Earthquake Simulation Package

Many metropolitan areas are affected by active seismic zones. Large earthquake events in these areas often claim numerous victims and lead large economic loss. Unfortunately, reliable earthquake predictions are out of scope with the present state of knowledge and technology. Therefore, it is crucial to construct the infrastructure and buildings in such a way that they can withstand the ground shaking of earthquakes. Simulating the high resonance frequencies of buildings, essential for civil engineering, is one of the biggest challenges in seismology. It requires knowledge about the Earth's interior, accurate physical models and computational resources in the peta- and exascale range.

SeisSol is an open-source earthquake simulation package. The code focuses on local wave propagation in detailed models and complicated fault zones. The strong coupling of the seismic waves and dynamic rupture simulations is designed to capture mutual influences (e.g. seismic waves triggering new rupture processes). To resolve the complex geometries (topology, fault systems, etc., see Figure 1), SeisSol is based on fully unstructured tetrahedral meshes. For a high-order discretization in space and time, SeisSol combines the Discontinuous Galerkin method with Arbitrary high-order DERivative (ADER-DG) time stepping. The kernels are implemented as small matrix-matrix multiplications. With Intel's LIBXSMM [1] as backend, highly optimized kernel code is generated for several Intel architectures, including Intel Knights Corner and Knights Landing. On both architectures, SeisSol's kernels achieve more

than 40% of the peak performance in node-level benchmarks. Since the communication scheme is completely local, the high peak performance rates can be maintained in large scale simulations.

SeisSol is actively developed in a collaboration between Ludwig-Maximilians-University (group of Alice-Agnes Gabriel) and Technical University of Munich (group of Michael Bader). Recent improvements to the code base include a clustered local time-stepping [2] and new physical models (attenuation [3] and plasticity [4]). To simulate the high frequencies and to access to the full potential of SeisSol, we are also tuning the complete workflow to support meshes with 100 million to 1 billion cells and 10^{11} to 10^{12} degrees of freedom.

Large Scale I/O

As part of its large scale workflow, SeisSol uses netCDF and XDMF/HDF5 to read unstructured meshes from disk and write visualization data [5]. The netCDF-based mesh format is customized to include pre-computed communication structures for efficient parallel initialization. The XDMF/HDF5 output automatically aggregates data from different MPI processes to create larger I/O buffers. The format is compatible to ParaView and VisIT and data can be visualized without further post-processing.

To complete the workflow, we implemented a checkpoint-restart mechanism for SeisSol and overlapped I/O operations with computation. For writing checkpoints, we use overwrite+flush

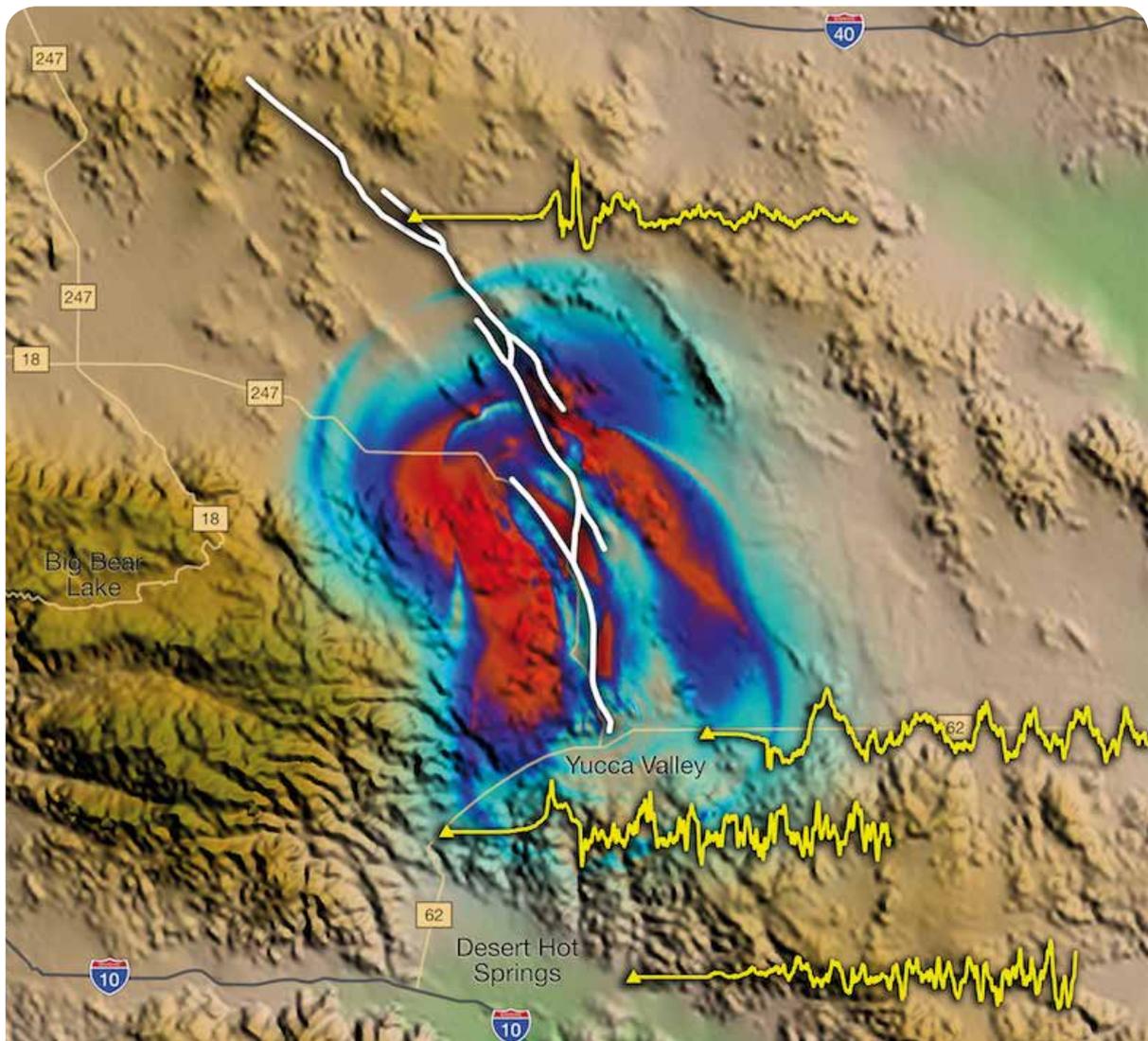


Fig. 1: The complex fault system of the 1992 Landers earthquake contains several branches which lead to an unpredictable rupture process. Figure obtained from [6].

(instead of open+write+close) to reduce expensive metadata operations in parallel file systems. For flexibility, SeisSol's checkpoints can be written using different I/O libraries (POSIX I/O, MPI-IO, HDF5, or SIONlib). On Hazel Hen, we had the unique opportunity to evaluate the

checkpoint implementation on a newly installed Lustre file systems before it was generally available. In addition, our results were not influenced by concurrent applications of other users. For large checkpoints all I/O libraries could utilize 72-84% of the available peak bandwidth of the



Lustre file system. For smaller checkpoints, HDF5 suffered from the additional overhead (compared to MPI-IO) and SIONlib from the missing flush operation (Figure 2).

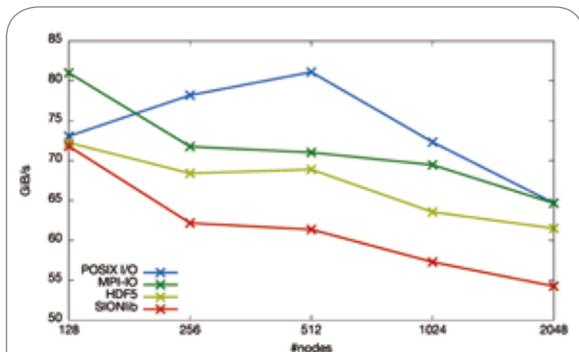


Fig. 2: Strong scaling of the checkpoint implementation depending on the I/O library. The theoretical peak bandwidth of the file system is 104 GB/s.

To overlap the output (checkpoints and visualization output) with the computation in SeisSol, we integrated our novel asynchronous I/O library. The library supports I/O threads running on the compute nodes as well as staging nodes which are excluded from the computation. The library is responsible for moving data buffers from the compute cores to the I/O threads (or nodes) and provides functions to start the I/O task and wait for completion. The implementation of the actual I/O routine is left to the application developer allowing them to choose the best I/O library for their purpose. In SeisSol, asynchronous I/O works with the XDMF/HDF5 output (Figure 3) and for checkpoints.

Combining the local time-stepping with the asynchronous I/O library allowed us to simulate the 1992 Landers earthquake in less than

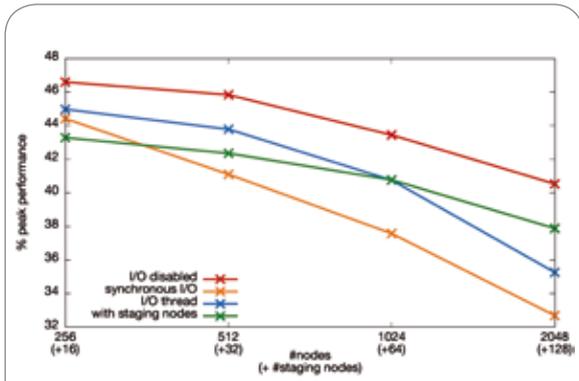


Fig. 3: Comparison of the different asynchronous I/O modes for the XDMF/HDF5 wave field output on SuperMUC.

3 h on 2112 nodes (with a peak performance of 2 PFlops) on Hazel Hen. The 1 TB of output data written during the simulation consists of 81 snapshots each containing 191 million tetrahedra (Figure 4). The same simulation without local time-stepping but also without the large scale output took more than 7 h on the whole SuperMUC Phase 1 system (with a peak performance of 3.2 PFlops) [6].

Outlook

Most parts of SeisSol's I/O implementation were originally designed to work with SeisSol only. However, we have identified several other applications that have similar demands. The XDMF/HDF5 output was recently extended to support adaptive mesh refinement and will be integrated into the ExaHype and the Terra Neo project. ESPRESO, a parallel FETI solver, will use the asynchronous I/O library to reduce the time spend in the VTK output.

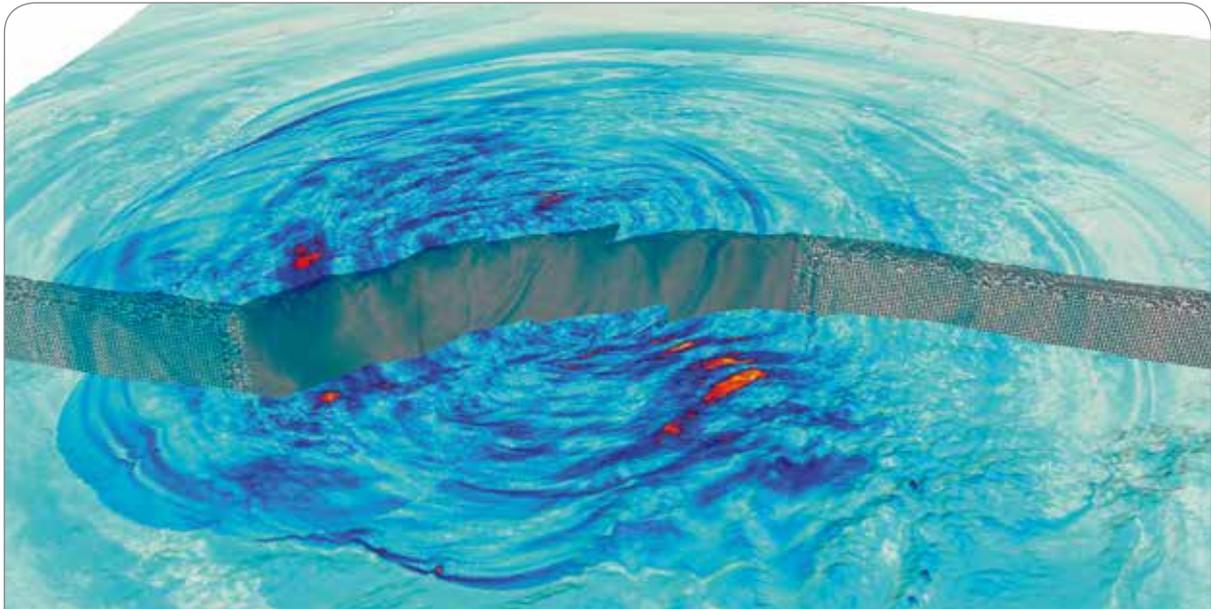


Fig. 4: Visualization of the 1992 Landers earthquake computed on 191 million tetrahedra (9.6×10^8 degrees of freedom).

References

- [1] **A. Heinecke, G. Henry, M. Hutchinson, and H. Pabst, LIBXSMM:**
Accelerating Small Matrix Multiplications by Runtime Code Generation. The International Conference for High Performance Computing, Networking, Storage and Analysis, 2016.
- [2] **A. Breuer:**
High Performance Earthquake Simulations, 2015. PhD Thesis.
- [3] **C. Uphoff, and M. Bader:**
Generating high performance matrix kernels for earthquake simulations with viscoelastic attenuation. The 2016 International Conference on High Performance Computing & Simulation, 2016.
- [4] **S. Wollherr, and A.-A. Gabriel:**
Dynamic rupture with off-fault plasticity on complex fault geometries using a Discontinuous Galerkin method: Implementation, verification and application to the Landers fault system. Submitted.
- [5] **S. Rettenberger, and M. Bader:**
Optimizing Large Scale I/O for Petascale Seismic Simulations on Unstructured Meshes, IEEE International Conference on Cluster Computing, 2015.
- [6] **A. Heinecke, A. Breuer, S. Rettenberger, M. Bader, A.-A. Gabriel, C. Pelties, A. Bode, W. Barth, X.-K. Liao, K. Vaidyanathan, M. Smelyanskiy, and P. Dubey:**
Petascale High Order Dynamic Rupture Earthquake Simulations on Heterogeneous Supercomputers, The International Conference for High Performance Computing, Networking, Storage and Analysis, 2014. Gordon Bell Finalist.-

Links

<http://www.seissol.org>
<https://github.com/SeisSol/SeisSol>

Written by **Sebastian Rettenberger, Michael Bader**

Technical University of Munich, Department of Informatics, Chair of Scientific Computing

Contact: Sebastian Rettenberger, rettenbs@in.tum.de



The Next Generation of Hydrodynamical Simulations of Galaxy Formation

Scientific background

Galaxies are comprised of up to several hundred billion stars and display a variety of shapes and sizes. Their formation involves a complicated blend of astrophysics, including gravitational, hydrodynamical and radiative processes, as well as dynamics in the enigmatic „dark sector“ of the Universe, which is composed of dark matter and dark energy. Dark matter is thought to consist of a yet unidentified elementary particle, making up about 85% of all matter, whereas dark energy opposes gravity and has induced an accelerated expansion of the Universe in the recent past. Because the governing equations are too complicated to be solved analytically, numerical simulations have become a primary tool in theoretical astrophysics to study cosmic structure formation. Such calculations connect the comparatively simple initial state left behind by the Big Bang some 13.6 billion years ago with the complex, evolved state of the Universe today. They provide detailed predictions for testing the cosmological paradigm and promise transformative advances in our understanding of galaxy formation.

One prominent example of such a simulation is our Illustris simulation from 2014 [1, 2]. It tracked the small-scale evolution of gas and stars within a representative portion of the Universe, using more than 6 billion hydrodynamical cells and an equally large number of dark matter particles. Illustris yielded for the first time a reasonable morphological mix of thousands of well-resolved elliptical and spiral galaxies. The simulation reproduced the observed distribution of

galaxies in clusters and the characteristics of hydrogen on large scales, and at the same time matched the metal and hydrogen content of galaxies on small scales. Indeed, the virtual universe created by Illustris resembles the real one so closely that it can be adopted as a powerful laboratory to further explore and characterize galaxy formation physics. This is underscored by the nearly 100 publications that have been written using the simulation thus far.

However, the Illustris simulation also showed some tensions between its predictions and observations of the real Universe, calling for both, improvements in the physical model as well as in the numerical accuracy and size of the simulations used to represent the cosmos. For example, one important physical ingredient that was missing are magnetic fields. In fact, our Universe is permeated with magnetic fields – they are found on Earth, in and around the Sun, as well as in our home galaxy, the Milky Way. Often the magnetic fields are quite weak, for example, the magnetic field on Earth is not strong enough to decisively influence the weather on our planet. In galaxies like the Milky Way, the field is however so strong that its pressure on the interstellar gas in the galactic disc is of the same size as the thermal pressure. This suggests that magnetic fields could play an important role in regulating star formation, but their origin still remains mysterious.

Another challenge that had become clear in our past work is that the regulation of star formation in massive galaxies through the energy output of growing supermassive black holes was not

adequately described by Illustris. Even though the adopted model was so strong and violent that it caused an excessive depletion of the baryon content of galaxy groups and low-mass galaxy clusters, it proved insufficient to reduce the star formation in the central galaxies in these systems to the required degree, causing these galaxies to become too massive. In essence, this showed a serious failure of the underlying theoretical model. Fixing it requires replacing it with something considerably different.

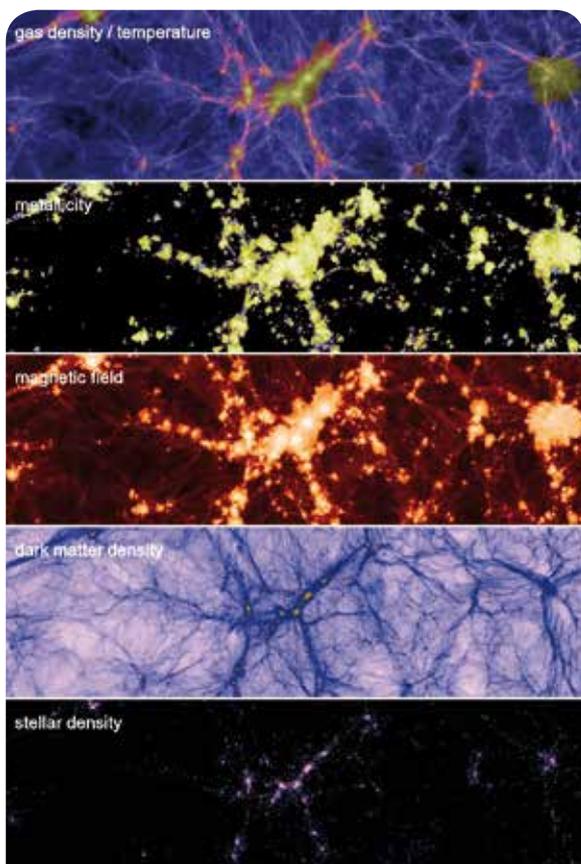


Fig. 1: Thin projections through one of our Illustris TNG simulations, showing different physical quantities as labelled.

We thus felt compelled to work on a new generation of simulations with the goal of advancing the state-of-art on all of these fronts by using a new comprehensive model for galaxy formation physics and updated numerical treatments. At the same time, we aimed for higher numerical resolution, larger volume covered and hence better statistics, as well as an improved accuracy in our hydrodynamical solvers. Using Hazel-Hen and a GCS grant for computer time, we were able to succeed on many of these aims, and to produce a novel, scientifically very interesting set of simulation models, which we now call “The Next Generation Illustris Simulations” (IllustrisTNG).

New physics modelling

For IllustrisTNG, we developed a new kinetic feedback model for AGN driven winds [3], motivated by recent theoretical suggestions that conjecture advection dominated inflow-outflow solutions for the accretion flows onto black holes in the low accretion rate regime. In terms of energetic feedback, we distinguish between a quasar mode for high accretion rates where the feedback is purely thermal, and a kinetic mode for low accretion rate states where the feedback is purely kinetic. The distinction between the two feedback modes is based on the Eddington ratio of the black hole accretion. In the kinetic feedback state, strong quenching of cooling flows and star formation in the host halo is possible, such that the corresponding galaxy can quickly redden.

Another important change we made relates to the modelling of galactic winds and outflows



[4], which now scale differently with the Hubble rate, and also take metallicity-dependent cooling losses better into account. The net effect of this is a stronger suppression of star formation in small galaxies, yielding an improved faint-end of the galaxy luminosity function.

Importantly, we have also added magnetic fields to our simulations, using a new implementation of ideal magnetohydrodynamics in our AREPO code [5, 6]. This opens up a rich new area of predictions that are still poorly explored, given that the body of cosmological magneto-hydrodynamic simulations is still very small. In particular, it allows us to study the strength of magnetic field amplification through structure formation as a function of halo mass and galaxy type.

For IllustrisTNG, we have also improved our modelling of chemical enrichment, both by using updated yield tables that account for the most recent results of stellar evolution calculations, and by making the tracking of different chemical elements (H, He, C, N, O, Ne, Mg, Si, Fe) more accurate and informative. For example, we developed a special chemical tagging method that separately accounts for metals produced by asymptotic giant branch stars, type-II supernovae, and type-Ia supernovae. This has not been done before in such hydrodynamical simulations.

Finally, we also developed a novel hierarchical timestepping scheme in our AREPO code that solves this in a mathematically clean fashion. This is done by recursively splitting the

Hamiltonian describing the dynamics into a slow and a fast system, with the fast system being treated through sub-cycling. An important feature of this time integration scheme is that the split-off fast system is self-contained, i.e. its evolution does not rely on any residual coupling with the slow part. This means that poorly populated short timesteps can be computed without touching any parts of the system living on longer timesteps, making these steps very fast so that they not slow down the main calculation significantly.

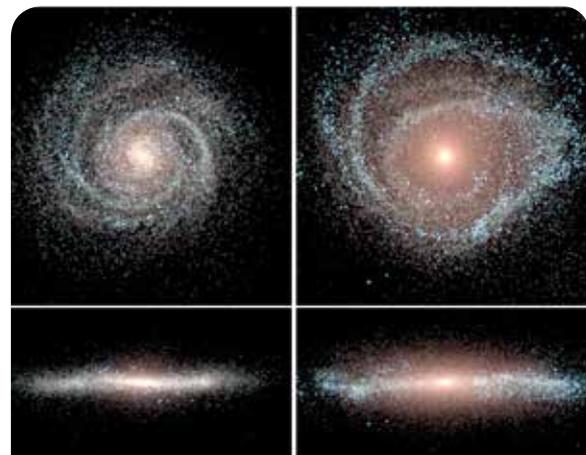


Fig. 2: Two disk galaxies in halos of mass $8.3 \times 10^{11} M_{\odot}$ (left) and $2.0 \times 10^{12} M_{\odot}$ (right).

Computational challenge

The AREPO code [5] we developed for cosmological hydrodynamics uses a finite-volume approach on a three-dimensional, fully dynamic Voronoi tessellation. The moving mesh is particularly well suited to the high dynamic range in space and time posed by the galaxy formation problem. The very low advection errors of AREPO are very helpful for the highly

supersonic flows occurring in cosmology and for treating subsonic turbulence within the gas of virialized halos. These properties make it superior to smoothed particle hydrodynamics and adaptive mesh refinement codes that use a stationary Cartesian mesh. AREPO also follows the dynamics of dark matter with high accuracy, as required to compute cosmic structure growth far into the non-linear regime.

The simulations carried out in the IllustrisTNG project represent a significant challenge not only in terms of size and spatial dynamic range, but also in terms of the dynamic range in timescales. In particular, the strong kinetic feedback by black holes, which couples to the densest gas in galaxies, induces very small timesteps for a small fraction of the mass. Over the course of 13 billion years of cosmic evolution that we cover, we needed to do up to 10^7 timesteps in total. This would be completely infeasible with time integration schemes that employ global timesteps, but even for the new individual timestepping we have used in AREPO, this represents a formidable problem. It could only be tackled by making the computation of sparsely populated timesteps extremely fast so that they do not dominate the total CPU time budget.

In addition to the challenging dynamic range in timescales, we also aim for a larger number of resolution elements, and a larger simulation volume than realized previously. This is necessary to study the regime of galaxy clusters better (which are rare and can only be found in a sufficiently large volume), and to allow a sampling

of the massive end of the galaxy and black hole mass functions. The primary science runs of IllustrisTNG consist of two large full-physics calculations (and a third one targeting dwarf galaxies is underway in follow up work), each significantly more advanced and also larger than the older Illustris simulation. This is complemented with matching dark matter only simulations, as well as a series of lower resolution calculations to assess numerical convergence. The calculations include magneto-hydrodynamics and adopted the newest cosmological models as determined by the Planck Satellite.

We have used between 10752 and 24000 cores on Hazel-Hen, benefitting in full from the large memory, high communication bandwidth, high floating point performance and high I/O bandwidth of this machine. The peak memory consumption of our largest run is about 95 TB RAM, and each of our simulation time slices weighs in with several TB. In fact, we have already transferred more than 300 TB of final production data to the Heidelberg Institute of Theoretical Studies, in part by using fast gridftp services offered by HLRS.

First results and outlook

In Figure 1, we illustrate the large-scale distribution of different physical quantities in one of our IllustrisTNG simulations. From top to bottom, we show projections of the gas density field, the mean mass-weighted metallicity, the mean magnetic field strength, the dark matter density, and the stellar density. The displayed regions are about 350 million lightyears across from left

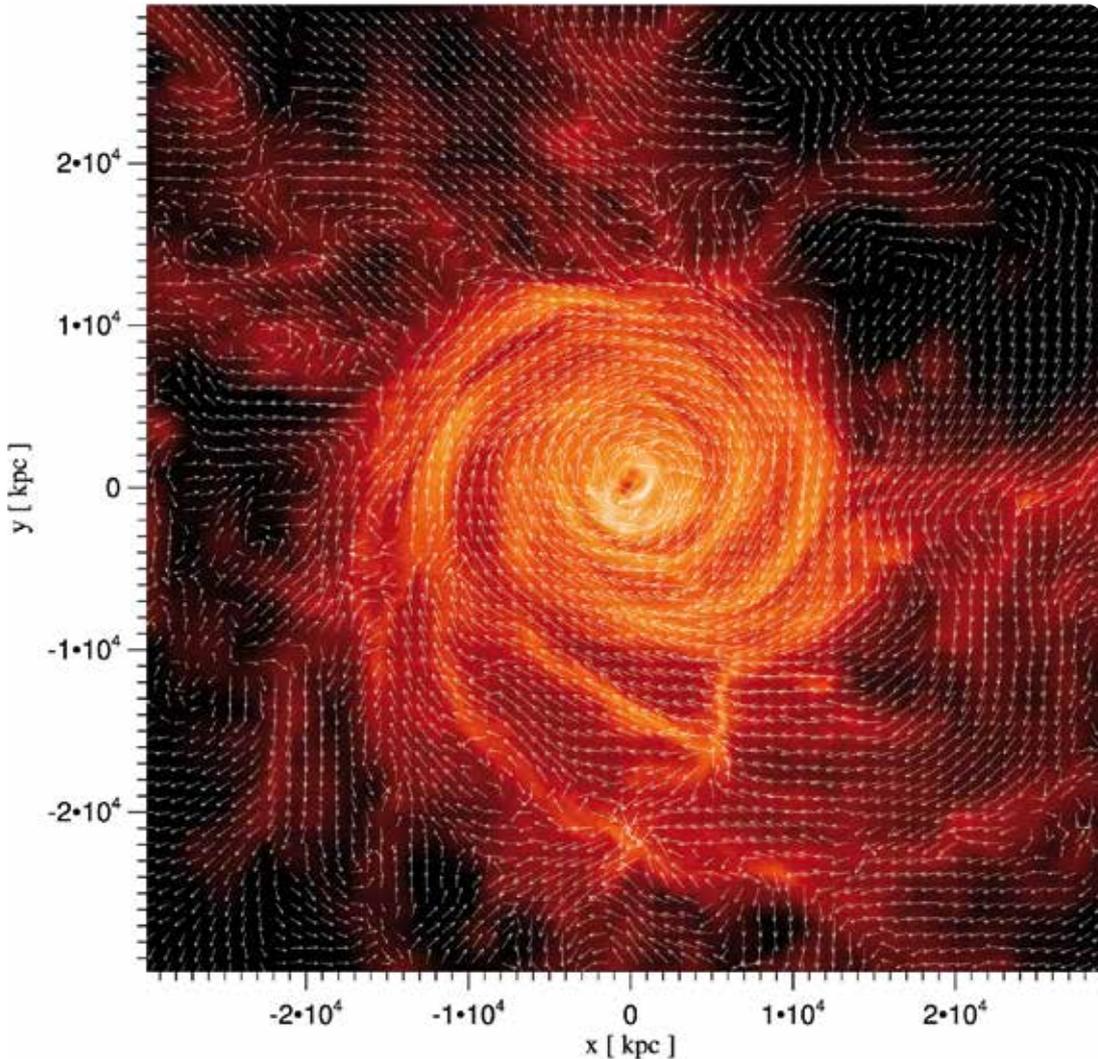


Fig. 3: Magnetic field direction (vectors) in the disk galaxy displayed on the left in Fig. 2, with the gas density structure shown in the background (color scale).

there is clear evidence for very strong outflows in them, causing widespread heating as they impinge on the gas in the intergalactic medium.

The rightmost panel in Figure 1 displays the stellar mass density. Clearly, on the scales shown in this image, the individual galaxies appear as very small dots, illustrating that the stellar component fills only a tiny fraction of the volume. However, our simulations have enough resolution and dynamic range to actually

to right. On large scales, the dark matter and the diffuse gas trace out the so-called cosmic web that emerges through gravitational instability. The color in the gas distribution encodes the mass-weighted temperature across the slice. The largest halos are filled with hot plasma, and

resolve the internal structure of these galaxies in remarkable detail. This is shown in Figure 2, which zooms in on two disk galaxies formed in our simulations. The one on the right hand panel is in a more massive halo and has a more massive black hole. This in fact has made it start to

transition into the quenched regime, which here begins by a reduced star formation in the center as a result of kinetic black hole feedback. The outskirts of the galaxy still support some level of star formation, causing blue spiral arms.

In Figure 3, we plot the magnetic vector field of this galaxy, overlaid on a rendering of the gas density in the background. We see that the field is ordered in the plane of the disk, where it has been amplified by shearing motions to sizable strength. Interestingly, there are multiple field reversals and a complicated topology of the field surrounding the disk. The realistic field topologies predicted here should be very useful for studying the propagation of cosmic rays in the Milky Way. Already now we can say that our calculations demonstrate that an extremely tiny magnetic field left behind by the Big Bang is sufficient to explain the orders of magnitude larger field strengths observed today. In fact, the field strengths we measure in our galaxies agree quite well with observational constraints.

In Figure 4, we show a break down of the total metal content in the gas phase of IllustrisTNG at different times as a function of gas density. The histograms are normalized to the total metal content in the gas at the corresponding epoch, so that the distributions inform about the question at which gas densities the majority of the metals can be found. Most of the metals are actually stored at gas densities that correspond to the circumgalactic medium, whereas only a smaller fraction is contained in the star-forming interstellar medium, and relatively little in the

low-density intergalactic medium. These distributions are shaped by the galactic winds in the simulation, and determining them observationally will provide powerful constraints on our theoretical models.

The scientific exploitation of the IllustrisTNG simulations has just begun. We expect that they will significantly expand the scientific possibilities and predictive power of hydrodynamical simulations of galaxy formation, thereby forming the ideal basis for the comparison with real data. Obtaining this valuable data has only been possible thanks to the power of the Hazel-Hen supercomputer, and mining the rich scientific results this data has in store will keep us and

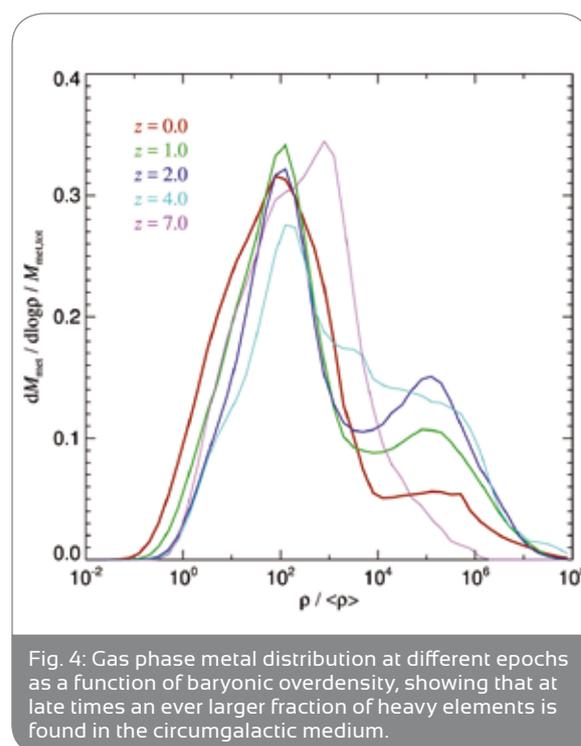


Fig. 4: Gas phase metal distribution at different epochs as a function of baryonic overdensity, showing that at late times an ever larger fraction of heavy elements is found in the circumgalactic medium.



many of our colleagues in the field busy for years to come.

Acknowledgements

We gratefully acknowledge support by the High Performance Computing Center in Stuttgart, and computer time through project GCS-ILLU on Hazel-Hen. We also acknowledge financial support through subproject EXAMAG of the Priority Programme 1648 SPPEXA of the German Science Foundation, and through the European Research Council through ERC-StG grant EXAGAL-308037, and we would like to thank the Klaus Tschira Foundation.

References

- [1] **Vogelsberger M., Genel S., Springel V., Torrey P., Sijacki D., Xu D., Snyder G., Bird S., Nelson D., Hernquist L.:** Properties of galaxies reproduced by a hydrodynamic simulation, 2014, *Nature*, 509, 177
- [2] **Genel S., Vogelsberger M., Springel V., Sijacki D., Nelson D., Snyder G., Rodriguez-Gomez V., Torrey P., Hernquist L.:** Introducing the Illustris project: the evolution of galaxy populations across cosmic time, 2014, *Monthly Notices of the Royal Astronomical Society*, 445, 175
- [3] **Weinberger R., Springel V., Hernquist L., Pillepich A., Marinacci F., Pakmor R., Nelson D., Genel S., Vogelsberger M., Naiman J., Torrey P.:** Simulating galaxy formation with black hole driven thermal and kinetic feedback, 2017, *Monthly Notices of the Royal Astronomical Society*, 465, 3291
- [4] **Pillepich A., Springel V., Nelson D., Genel S., Naiman J., Pakmor R., Hernquist L., Torrey P., Vogelsberger M., Weinberger R., Marinacci F.:** Simulating Galaxy Formation with the IllustrisTNG Model, 2017, *Monthly Notices of the Royal Astronomical Society*, submitted, arXiv:1703.02970
- [5] **Springel V.:** E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh, 2014, *Astrophysical Journal Letters*, 401, 791
- [6] **Pakmor R., Marinacci F., Springel V.:** Magnetic Fields in Cosmological Simulations of Disk Galaxies, 2017, *Astrophysical Journal Letters*, 783, L20

Written by Volker Springel, Rainer Weinberger, Rüdiger Pakmor

Heidelberg Institute for Theoretical Studies and Heidelberg University, Germany

Annalisa Pillepich

Max-Planck Institute for Astronomy, Heidelberg, Germany

Dylan Nelson

Max-Planck-Institute for Astrophysics, Garching, Germany

Mark Vogelsberger, Federico Marinacci, Paul Torrey

Massachusetts Institute of Technology, Cambridge, USA

Lars Hernquist, Jill Naiman

Harvard University, Cambridge, USA

Shy Genel

Center for Computational Astrophysics, Flatiron Institute, New York, USA

Contact: Volker Springel, volker.springel@h-its.org

First-Principles Design of Novel Materials for Spintronics Applications

Motivation

Spintronics, or spin transport electronics, is a field of technology that exploits the electron spin, possibly in addition to the electron charge, to achieve the next generation of electronic devices. As magnetic effects occur at smaller energy scales than the electronic ones, spintronic devices promise low-power dissipation devices. Furthermore, as magnetic interactions are enhanced at the nanoscale, spin-based devices can also allow one to achieve device miniaturization.

Despite many active and passive spin-devices have been achieved, the quest for optimal materials for spintronics applications is still open. For instance, the Datta spin transistor [1] is based on the concept that it is possible to manipulate the spin of the electrons in the channel material via an external electric field, the ability to act on the spin being proportional to the Spin Orbit Coupling (SOC) of the material. For this, materials with simultaneously high SOC and long spin coherence length are sought for. However, these requirements are often in conflict: typical semiconductors present high SOC but short spin coherence length.

On the contrary, carbon nanomaterials have a remarkably long spin diffusion length (two order of magnitudes larger than in inorganic semiconductors) but extremely small SOC. Graphene presents an additional difficulty: it has a vanishing electronic band gap, which makes impossible to switch off a graphene-based transistor.

Computations

We propose to overcome both difficulties, weakness of the SO interaction and lack of a band-gap, by placing graphene on a magnetic semiconducting substrate. As representative of this family, we choose a hexagonal Mn-based magnetoelectric: BaMnO_3 . We have used first-principles techniques as implemented in the SIESTA code, to compute the structural, electronic and magnetic properties of the graphene/ BaMnO_3 interface [2]. The use of high-performance machines has been critical to address this material system, due to the large size (large number of atoms) needed to model various graphene/ BaMnO_3 slabs. Calculations have been performed on several machines: JUROPA/JURECA at FZJuelich (Germany) and LINDGREN at KTH (Sweden), as the work lasted three years.

Results

Using this approach, we have shown that the spin polarization is induced in the pristine carbon network [Fig. 1], exclusively due to the strong interaction between the Carbon π and the Mn d states (proximity interaction). Analysis of the electronic band structure shows that the effect is general and valid for any RMnO_3 (with R a rare-earth material) compound. The resulting hybrid system is half-metallic: majority carriers present an electronic band gap, while minority carriers have no gap. Hence, the graphene- BaMnO_3 can act as an injector of 100% spin polarized carriers. Since BaMnO_3 is not only magnetic but also insulating, we suggest using a thin layer of such a material to simultaneously achieve spin injection and high resistance contacts. This approach



has the advantage of combining in one material – the magnetic insulator – the two main features of the most efficient injection scheme known to date: ferromagnetic contacts followed by a tunnel barrier of insulating material deposited on graphene [3]. A possible device exploiting these results is a spin-FET in which spin injection is obtained by graphene-BaMnO₃ in its ground state. Then, the FET channel can be made by

graphene on an insulating material, such as hexagonal Boron Nitride (hBN), which preserves the long intrinsic spin coherence length of carbon. A high SOC substrate would enable control of the spin in the channel. The continuous graphene with a modulated substrate minimizes problems with contact resistances and interface mismatch in the transport direction.

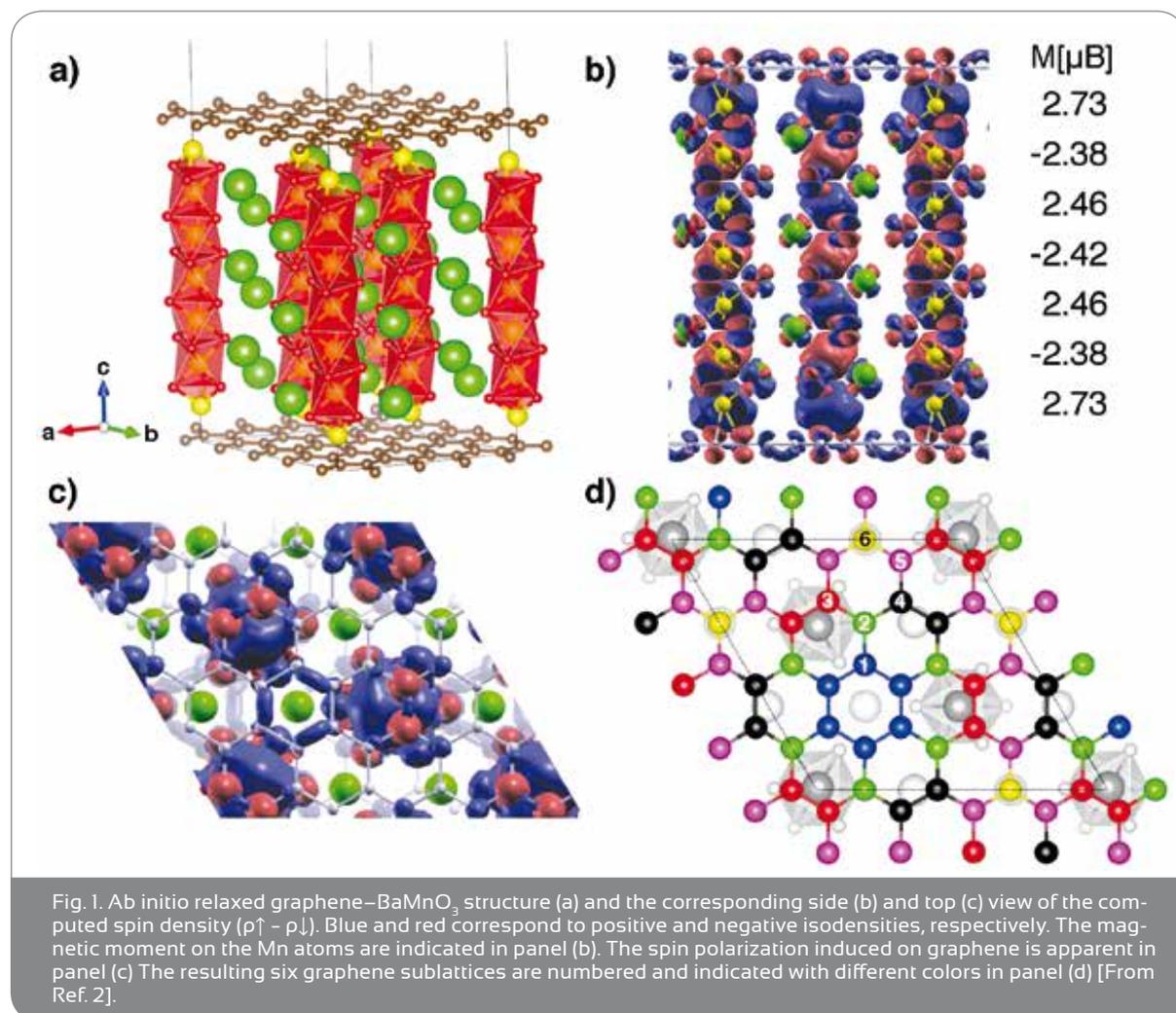


Fig. 1. Ab initio relaxed graphene-BaMnO₃ structure (a) and the corresponding side (b) and top (c) view of the computed spin density ($\rho\uparrow - \rho\downarrow$). Blue and red correspond to positive and negative isodensities, respectively. The magnetic moment on the Mn atoms are indicated in panel (b). The spin polarization induced on graphene is apparent in panel (c) The resulting six graphene sublattices are numbered and indicated with different colors in panel (d) [From Ref. 2].

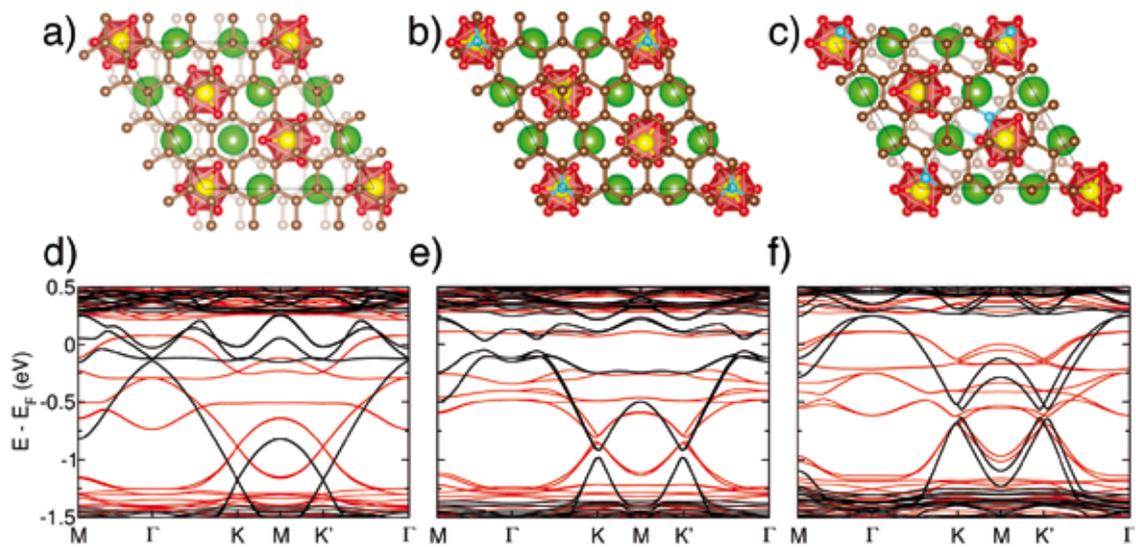


Fig. 2: Structural and electronic properties of the graphene-BaMnO₃ slab. Ball-and-stick model (top) and electronic band structure (bottom) of graphene-BaMnO₃ in its ground state (left), and of the relaxed structure with one (middle) and two (right) substitutional Boron atoms. The position of the Dirac cones is shifted towards the Fermi energy (E_F) with increasing doping. Red and black solid lines indicate, respectively, majority and minority spin channels. The Fermi energy of the hybrid system is taken as a reference. Atom colors: Mn yellow, Ba green, O red, C gold, B light blue. [From Ref. 2].

Going further, our simulations show that the high-mobility region, characteristic of graphene electronic band structure, are preserved in the hybrid system. High-mobility is achieved when the energy has a linear dependence on the momentum, resulting in the so-called Dirac cones. The latter are found in the band structure and the splitting between majority and minority cones is quite large (~300 meV), but they occur at quite low energy. We address this issue by showing that doping of graphene with acceptors can be used to tune the Dirac cones, moving them into the experimentally accessible energy range [Fig. 2]. The velocity of the two types of carriers is quite different, so spin dependent transport is expected in this hybrid material.

Acknowledgements

Z.Z. acknowledges EC support under the Marie-Curie fellowship (PIEF-Ga-2011-300036) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) grant ZA 780/3-1. Z.Z. also acknowledges computational resources from the PRACE-3IP project (FP7 RI-312763) resource Lindgren based in Sweden at KTH, and the JARA-HPC projects jara0088, JIAS16, JHPC39.

References

- [1] **Datta, S and Das, B.:**
Appl. Phys. Lett. 56 665 (1990)
- [2] **Zanolli, Z.:**
Graphene-multiferroic interfaces for spintronic applications, Scientific Reports, 6, 31346 (2016)
- [3] **W. Han, R. K. Kawakami, M. Gmitra, J. Fabian:**
Nat. Nanotechnol. 9, 794 (2014) and references therein

Written by Dr. Zeila Zanolli

Physics Department, RWTH Aachen University and European Theoretical Spectroscopy Facility (ETSF)

Contact: Zeila Zanolli, zanolli@physik.rwth-aachen.de

Applications of External Users on GCS Computers

Stability Confinement of the Brain Connectome



What is the influence of specific connections between areas in the mammalian cortex on its dynamical state? A novel theoretical framework tackles this question by combining techniques from statistical physics with large-scale numerical simulations on a supercomputer.

Theoretical neuroscience aims to understand the working principles of the brain. The cortex is the outermost structure of the brain of humans and other mammals, and is responsible for sensory processing and higher cognitive functions. It is subdivided into cortical areas comprising up to hundreds of millions of neurons each. The neurons form a network, where each neuron receives signals from on average 10000 other neurons, either located in the same area or in remote parts of the brain. The corresponding wiring pattern is described by the 'connectome'.

A central topic in systems neuroscience is the link between the connectome and the activity of the neurons. For this purpose, we develop multi-scale network models, combining realistic cell densities on the order of 10^4 – 10^5 cells per mm^3 with experimental data on brain connectivity. These models represent the neural network as a system of coupled differential equations with a dimension proportional to the number of neurons, typically solved on a supercomputer to obtain the network dynamics.

The first sanity check compares the model dynamics with basic physiological observations. The anatomical data typically come with

large uncertainties, so that the activity does not automatically match the observations. This lack of knowledge calls for a parameter exploration within the uncertainties of the connectome. The high dimensionality of the networks, however, prohibits testing each parameter combination in simulation (Figure 1, black arrows). We therefore describe here an analytical method which circumvents this problem (Figure 1, red arrows) and finds a realization of the model parameters within the uncertainties that yields plausible network activity.

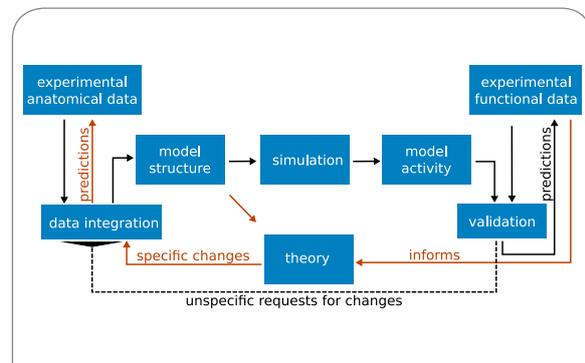


Fig. 1: The integrative loop between experiment, model, and theory. Black arrows represent the classical forward modeling approach: experimental anatomical data are integrated into a model structure, which gives rise to the activity via simulation. Red arrows symbolize the new approach: informed by functional data, an analytical theory of the model identifies critical connections to be modified, refining the integration of data into a consistent model structure and generating predictions for anatomical studies. Figure taken from [1].

A full account of the work summarized here has recently appeared as Schuecker et al. (2017), "Fundamental Activity Constraints Lead to Specific Interpretations of the Connectome", PLOS CB, 13(2):1–25.

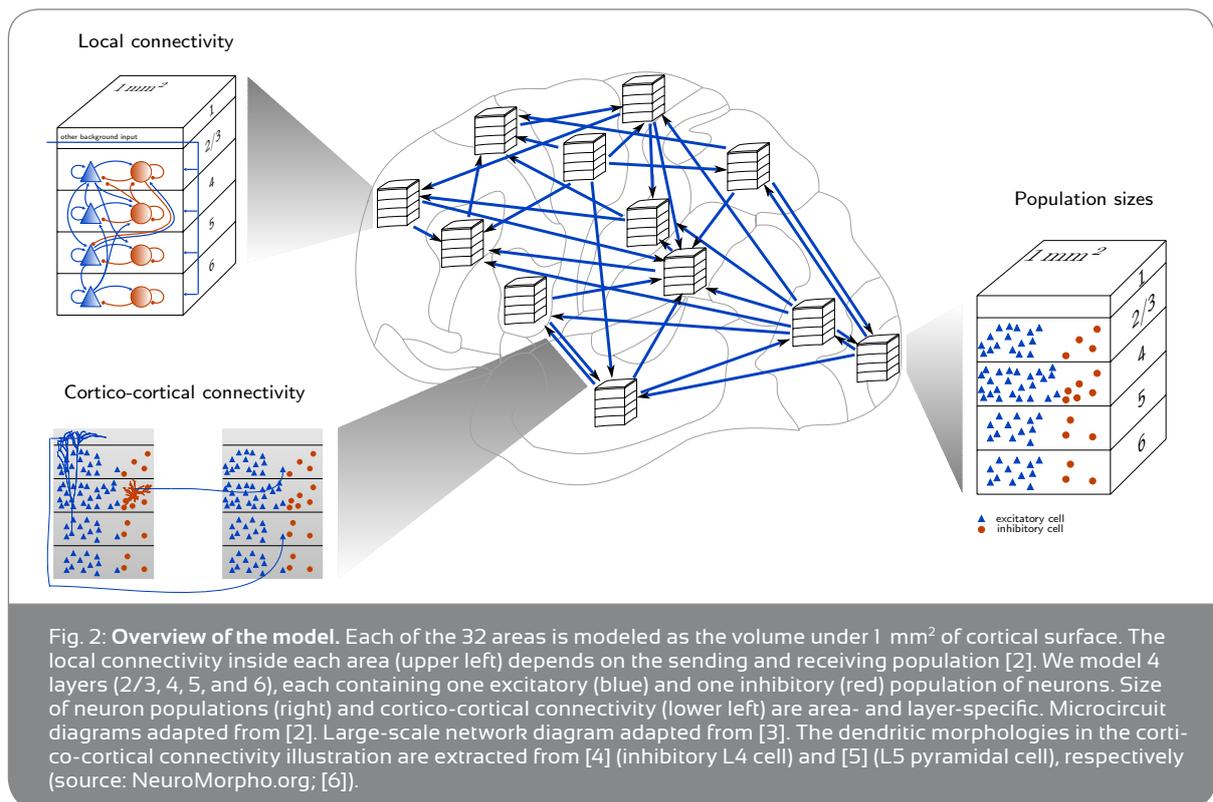


Fig. 2: **Overview of the model.** Each of the 32 areas is modeled as the volume under 1 mm² of cortical surface. The local connectivity inside each area (upper left) depends on the sending and receiving population [2]. We model 4 layers (2/3, 4, 5, and 6), each containing one excitatory (blue) and one inhibitory (red) population of neurons. Size of neuron populations (right) and cortico-cortical connectivity (lower left) are area- and layer-specific. Microcircuit diagrams adapted from [2]. Large-scale network diagram adapted from [3]. The dendritic morphologies in the cortico-cortical connectivity illustration are extracted from [4] (inhibitory L4 cell) and [5] (L5 pyramidal cell), respectively (source: NeuroMorpho.org; [6]).

Multi-scale model of visual cortex

We study a multi-scale network model of all vision-related areas of macaque monkey cortex (Figure 2). The model represents each cortical area as the network under 1 mm² of cortical surface with realistic cell densities. The model fulfills two basic constraints of cortical networks: each cell forms connections with on average 10,000 other neurons and the probability for two neurons in the same 1 mm² to be connected is roughly 10%. Typically a model area comprises four layers that each contain two different populations of neurons: one with excitatory cells, which increase the activity of the connected neurons, and one with inhibitory

cells, which decrease target cell activity. The population sizes are derived from a large collection of experimental studies describing laminar thicknesses and cell densities, architectural categories of areas [7], and proportions of excitatory and inhibitory neurons in each layer [8]. Neurons are connected to each other in a random fashion with a connection probability that depends on the respective populations to which they belong. These connection probabilities are derived from a collection of experimental data from different sources [9, 10, 11]. Overall, the network contains approximately 4 million neurons interconnected with $4 \cdot 10^{10}$ synapses. Inputs to the model from subcortical and non-visual



cortical areas are replaced by random input, whose strength is a global free parameter.

Inevitable gaps and uncertainties in the anatomical data are further constrained by electrophysiological measurements of neuronal activity. Neurons communicate through action potentials or spikes; sharp electrical signals of a few milliseconds' width. The number of spikes sent per second, the firing rate, is measured to lie roughly between 0.05 spikes / s and 30 spikes / s, providing bounds for the simulated activity.

Distributed parallel simulation at cellular resolution

Numerical simulations represent each neuron by an equivalent circuit that mimics the dynamics of the constituents of the biological cell, mathematically described by a system of differential equations. In this study we use the leaky integrate-and-fire (LIF) model neuron. The resistance of a parallel RC circuit models the voltage-gated ion channels of the cell membrane, and the capacitance mimics the membrane's ability to store electrically charged ions. Whenever the membrane potential, modeled as the voltage across the capacitance, exceeds a fixed threshold, the neuron sends a spike to all its connected partner cells. The entire neuronal network is thus a system of coupled differential equations with directed, delayed, and pulsed interaction.

Our simulations are carried out with the neuronal network simulation code NEST [12] an

open-source software that is developed in an international collaboration. NEST distributes the neurons evenly over the thousands of compute nodes of a supercomputer so that the equations of the neurons are solved in parallel. However, each compute node is a full computer like a PC and today often houses one or more processors that can solve tens of equations in parallel without slowing down. Therefore, NEST simultaneously exploits parallel computing on two levels: inside the compute nodes (using OpenMP) and across compute nodes (using MPI).

The sub-threshold dynamics of single cells is usually simple and fast to solve, in particular in the case of the LIF neuron model with linear and thus exactly solvable differential equations [13]. However, each neuron sends its spikes to 10,000 other cells, so that the bulk of the computational load lies in the communication of spikes between neurons, which are usually placed on different compute processes. It is therefore beneficial to use as few compute nodes of a high-performance system as possible, maximally filling the working memory of each compute node. The simulations presented here were carried out on the JUQUEEN supercomputer at Forschungszentrum Jülich and utilized 65,536 threads on 1024 compute nodes. The network instantiation takes ~ 5 minutes and simulating 1 s of biological time amounts to ~ 12 minutes wall-clock time, depending on the activity level in the neuronal network.

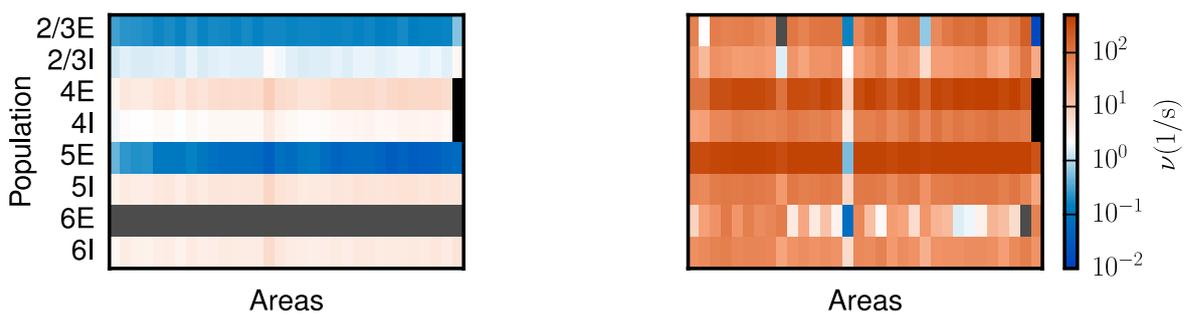


Fig. 3: Bistability of the multi-area network model. Left: State with low population-averaged firing rates in the network model at low external drive. Right: For increased external drive, the network enters the high-activity state. The color bar refers to both panels. Black indicates that a certain neuron population does not occur in the respective layer, gray indicates extremely low rates ($< 10^{-2}$ Hz).

Stability in a high-dimensional network model

Simulations of the model (Figure 3A) reveal that, though realistic levels of activity can be achieved for populations in layers 2/3 and 4, the excitatory populations in layers 5 and 6 of the majority of areas show vanishingly low or zero activity in contrast to empirical data. To elevate these firing rates, we increase the external drive to the corresponding populations. Already a perturbation of a few percent, however, causes the network to enter a state with unrealistically high rates (Figure 3B). Thus, with the chosen parameters the activity of the model does not match the measured experimental data. The high dimensionality of the model, moreover, prohibits a systematic parameter search that would counter the bistability.

Mixed analytical - simulation approach

How can we adjust the dynamics of the model to experimental data when the variables of the individual neurons numbering millions prohibit

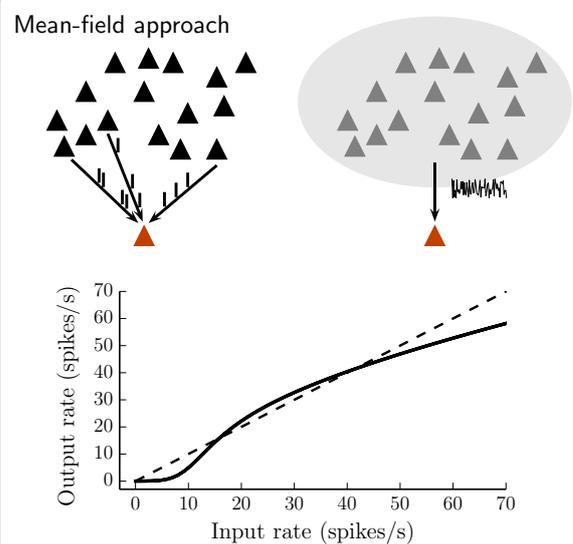


Fig. 4: Analytical reduction of network activity. Top: Visualization of the mean-field approach: In the full network simulation, each neuron receives a series of spikes from the neurons projecting to it. If the spiking in the network is sufficiently asynchronous and irregular, the input can be approximated by a randomly fluctuating current. Bottom: Collective input-output relationship $\Phi(\nu)$ for the example network consisting of one excitatory population.



a systematic search for parameters? We use mean-field theory, a technique known from statistical physics, to analytically describe the population-averaged spiking rates in the model, thereby reducing the dimensionality to 254, the number of populations of the model. The main idea of the mean-field reduction is to approximate the input from a large number of sources by an effective Gaussian process (Figure 4). The approach yields a self-consistency equation for the stationary firing rates $\mathbf{v} = \Phi(\mathbf{v})$ that describe the stationary network state, where v_i is the rate of population i and Φ is a non-linear transfer function mapping the input of the neurons to their output rate.

We apply the mean-field approach to render the multi-area model consistent with experimentally measured activity. Let us illustrate the method on a simplified network of only one excitatory population. The fixed points are given by the intersections of the transfer function with the identity line (Figure 4). Due to the non-linearity of Φ combined with the positive excitatory feedback, two stable attractors emerge with an unstable one in between, a bistable situation similar to that in the multi-area model (Figure 3). Figure 5A visualizes the situation by a corresponding energy landscape.

Our goal is to increase the activity of the low-activity (LA) state to a realistic range. However, this decreases its global stability, the distance to the unstable fixed point, as the latter shifts in the opposite direction (Figure 5B). Eventually, fluctuations drive the system to the unrealistic

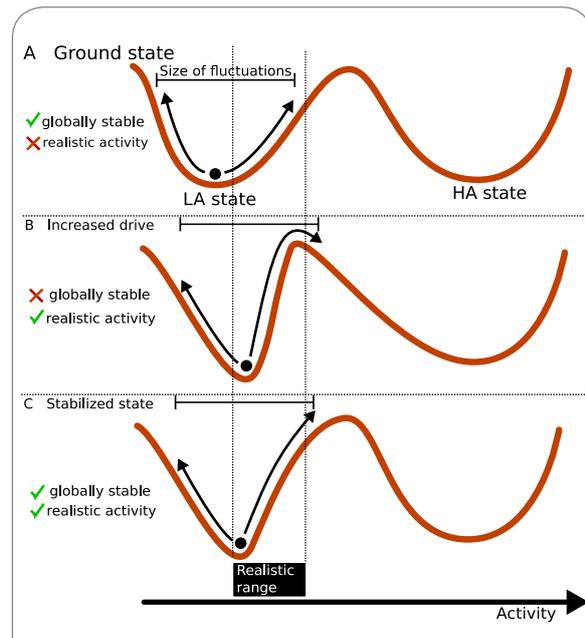


Fig. 5: Illustration of stabilization method. A The network activity fluctuates around the low-activity (LA) attractor. It is globally stable: fluctuations do not drive the network to the high-activity (HA) state. B Increasing the drive moves the LA state to the realistic regime, but reduces the global stability; stochastic fluctuations may kick the system to the HA state. C Targeted modifications to the connectivity move the unstable fixed point back to its original position, thus ensuring global stability.

high-activity (HA) state. The mean-field theory exposes the dependence of the fixed points on the model's structural connectivity, allowing us to adapt specific connections to shift back the unstable fixed point. These adaptations lead to a realistic, globally stable LA state (Figure 5C).

The set of small, specific structural modifications reveals critical roles for layer 5 and a loop between two frontal areas (Figure 6A). The model with the adapted connectivity achieves a dynamical state with realistically low but non-zero activity in all populations (Figure 6B).

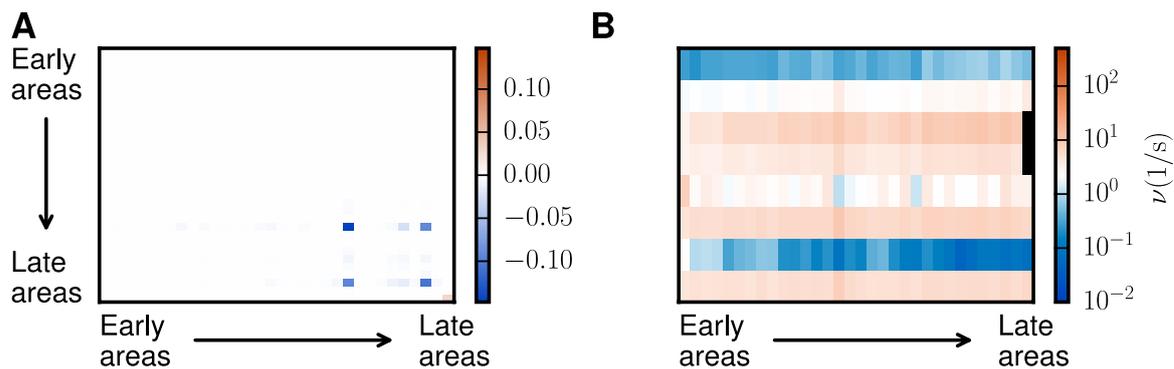


Fig. 6: Network after stabilization. A Relative changes in the connection density between all pairs of areas. Predominantly, the connectivity between frontal areas FEF and 46 is reduced. B Population-averaged firing rates obtained from a network simulation using the modified connectivity.

Ongoing research and outlook

The presented analytical method solves the data integration challenge for the brain-scale network model of macaque visual cortex. It exploits the constraints on the activity, being on average low and sufficiently stable. In future, the resulting consistent model may serve as a research platform to address open questions on the emergence of brain-wide low-frequency oscillations and the distinct frequency bands of bottom-up and top-down signaling.

So far, the method is restricted to stationary firing rates. In future studies, also higher-order statistical measures of activity can be used as constraints. Functional magnetic resonance imaging,

for example, provides information on the functional connectivity between areas as a second-order measure. When combined with analytical predictions of functional connectivity, our method may shed light on the anatomical connection patterns underlying inter-area communication.

The multi-scale spiking network model currently represents each area as the network corresponding to a patch of 1 mm² surface area. In cortex, however, the areas vary in size by orders of magnitude, cover up to 1400 mm², and contain up to hundreds of millions of neurons. In future, we aim to study the impact of scaling the network to more realistic relative sizes of up to 800 million neurons in total.

References

- [1] Schuecker, J., Schmidt, M., van Albada, S. J., Diesmann, M., & Helias, M. (2017): Fundamental activity constraints lead to specific interpretations of the connectome. *PLoS Comput. Biol.* 13 (2), 1–25.
- [2] Potjans, T. C., & Diesmann, M. (2014): The cell-type specific cortical microcircuit: Relating structure and activity in a full-scale spiking network model. *Cereb. Cortex* 24 (3), 785–806.
- [3] Kunkel, S., Potjans, T. C., Morrison, A., & Diesmann, M. (2009): Simulating macroscale brain circuits with microscale resolution. In *Proceedings of the 2nd INCF Congress of Neuroinformatics*. doi:10.3389/conf.neuro.11.2009.08.044.
- [4] Stepanyants, A., Hirsch, J. A., Martinez, L. M., Kisvarday, Z. F., Ferencsik, A. S., & Chklovskii, D. B. (2008): Local potential connectivity in cat primary visual cortex. *Cereb. Cortex* 18 (1), 13–28.



- [5] **Mainen, Z. F., & Sejnowski, T. J. (1996):**
Influence of dendritic structure on firing pattern in model neocortical neurons. *Nature* 382, 363–366.
- [6] **Ascoli, G. A., Donohue, D. E., & Halavi, M. (2007):**
Neuromorpho.org: a central resource for neuronal morphologies. *J. Neurosci.* 27 (35), 9247–9251.
- [7] **Hilgetag, C. C., Medalla, M., Beul, S. F., & Barbas, H. (2016):**
The primate connectome in context: Principles of connections of the cortical visual system. *NeuroImage*.
- [8] **Binzegger, T., Douglas, R. J., & Martin, K. A. C. (2004):**
A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* 39 (24), 8441–8453.
- [9] **Bakker, R., Thomas, W., & Diesmann, M. (2012):**
CoCoMac 2.0 and the future of tract-tracing databases. *Front. Neuroinformatics* 6 (30).
- [10] **Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., & Kennedy, H. (2014a):**
Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *J. Compar. Neurol.* 522(1), 225–259.
- [11] **Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M. A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marini, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D. C., & Kennedy, H. (2014b):**
A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* 24 (1), 17–36.
- [12] **Bos, H., Morrison, A., Peyser, A., Hahne, J., Helias, M., Kunkel, S., Ippen, T., Eppler, J. M., Schmidt, M., Seeholzer, A., Djurfeldt, M., Diaz, S., Moren, J., Deepu, R., Stocco, T., Deger, M., Michler, F., & Plesser, H. E. (2015):**
Nest 2.10.0.
- [13] **Plesser, H. E., & Diesmann, M. (2009):**
Simplicity and efficiency of integrate-and-fire neuron models. *Neural Comput.* 21, 353–359.

Written by Maximilian Schmidt^{*1,2}, Jannis Schuecker^{*1}, Sacha J. van Albada¹,
Markus Diesmann^{1,3,4}, Moritz Helias^{1,4}

¹Institute of Neuroscience and Medicine (INM-6), Institute for Advanced Simulation (IAS-6) and JARA BRAIN Institute I Jülich Research Centre, Jülich, Germany

²Laboratory for Neural Circuit Theory, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

³Department of Psychiatry, Psychotherapy and Psychosomatics, Medical Faculty, RWTH Aachen University, Aachen, Germany

⁴Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany

*These authors contributed equally

Contact: Maximilian Schmidt, maximilian.schmidt@riken.jp



Maximilian Schmidt

After obtaining his Diploma in Physics at Heidelberg University, Dr. Maximilian Schmidt joined the INM-6 at Forschungszentrum Jülich to work on large-scale simulations of neuronal networks in the cortex. He received his PhD in Physics in 2016 from RWTH Aachen University. After continuing as postdoc in the group of Prof. Markus Diesmann at INM-6, he took up a postdoc position at RIKEN Brain Science Institute in Tokyo, Japan.



Jannis Schuecker

Jannis Schuecker obtained his Master of Science in Physics at Bielefeld University. Since 2013 he is a doctoral researcher at INM-6 at Forschungszentrum Jülich focusing on mean-field theory for complex neuronal networks.

Lipid Transport by the ABC Transporter MDR3

ABC transporters are a large family of transmembrane proteins that mediate the translocation of a wide range of substrates across the membrane bilayer. Despite recent structural advances, the exact mechanism by which ABC transporters exert their functions at the molecular level has remained elusive so far, and several studies suggest a surprising mechanistic diversity within the members of the ABC transporter family. Using molecular dynamics simulations and free energy calculations, we investigated a novel substrate translocation pathway for the highly specialized ABC transporter multidrug resistance protein 3 (MDR3).

The human ATP-binding cassette (ABC) transporter multidrug resistance protein 3 (MDR3, ABCB4) plays a vital role in bile formation. It is primarily expressed in the canalicular membrane of hepatocytes, where it translocates phosphatidylcholine across the lipid bilayer and thereby promotes phospholipid secretion into bile. Phospholipids are an essential component of biliary micelles, which represent the preferred transport vehicle for the detergent bile acids. High concentrations of free bile acids would cause a loss of the membrane integrity of biliary epithelial cells. Accordingly, mutations in the ABCB4 gene are associated with Progressive Familial Intrahepatic Cholestasis type 3 (PFIC-3), a rare hereditary disease that ultimately results in liver failure.

Despite its high sequence similarity of 86% to the well-known drug efflux pump, multidrug resistance protein 1 (MDR1, P-glycoprotein),

the substrate spectrum of both transporters is fundamentally distinct: While MDR1 transports a wide range of structurally unrelated hydrophobic compounds, MDR3 is a floppase specific for lipids with a phosphatidylcholine head group [1]. Since most of the amino acids involved in substrate binding in MDR1 are identical in MDR3 [2], the lipid specificity of MDR3 must arise from a different region of the protein.

A conspicuous arrangement of non-conserved, hydrophilic amino acids in transmembrane helix 1 (TMH1) of MDR3 suggests an alternative pathway for substrate translocation that does not involve the “classical” translocation pathway via the central cavity of the transporter. Instead, we hypothesize that lipid translocation occurs along the surface of TMH1, where it is facilitated by the interaction between the phosphatidylcholine head group and the side chains of these hydrophilic residues. Transport of lipids and lipid-like molecules that partially or solely occurs along designated surface cavities has already been described for ABC transporters [3] and other lipid transporters [4].

With this project, we aim to obtain a better understanding of the molecular mechanisms by which MDR3 translocates phospholipids. Atomic-level insights into the transport process of this highly specialized ABC transporter are not yet available and could be a major step on the way towards a complete understanding of the inner workings of ABC transporters.



Methods

Molecular dynamics (MD) simulations at the classical mechanical level are at present the most appropriate way to explore the dynamics and energetics of complex biological molecules. In MD simulations, Newton's equations of motion are solved by numerical integration. They are used to estimate the equilibrium properties of biomolecular systems and to describe the quality and timescales of biomolecular processes controlled by conformational changes.

To assess whether the proposed pathway along TMH1 of MDR3 is a more efficient alternative to a spontaneous, i.e. unassisted, phospholipid flip-flop, we performed MD simulations with the Amber software suite [5]. In particular, we employed steered molecular dynamics (sMD) simulations and umbrella potential-restrained MD simulations ("umbrella sampling") to calculate if the free energy barrier of spontaneous phospholipid flip-flop decreases when moving the lipid along a pathway lined by the hydrophilic residues in TMH1 of MDR3.

Phospholipid flip-flop is accompanied by a highly unfavorable transfer of the charged lipid head group into, and then out of, the hydrophobic core of the lipid bilayer. Consequently, the free energy landscape associated with this process is steep and requires extensive and biased sampling to be accurately reproduced by MD simulations. The simulation system used to study MDR3-mediated phosphatidylcholine flip-flop (depicted in Figure 1) contains ~240,000 atoms and requires at least 2 μ s of

total simulation time divided into 100 umbrella windows of 20 ns length to yield converged free energy profiles. Until now, this project consumed approximately $4 \cdot 10^6$ core hours on the general-purpose cluster JURECA.

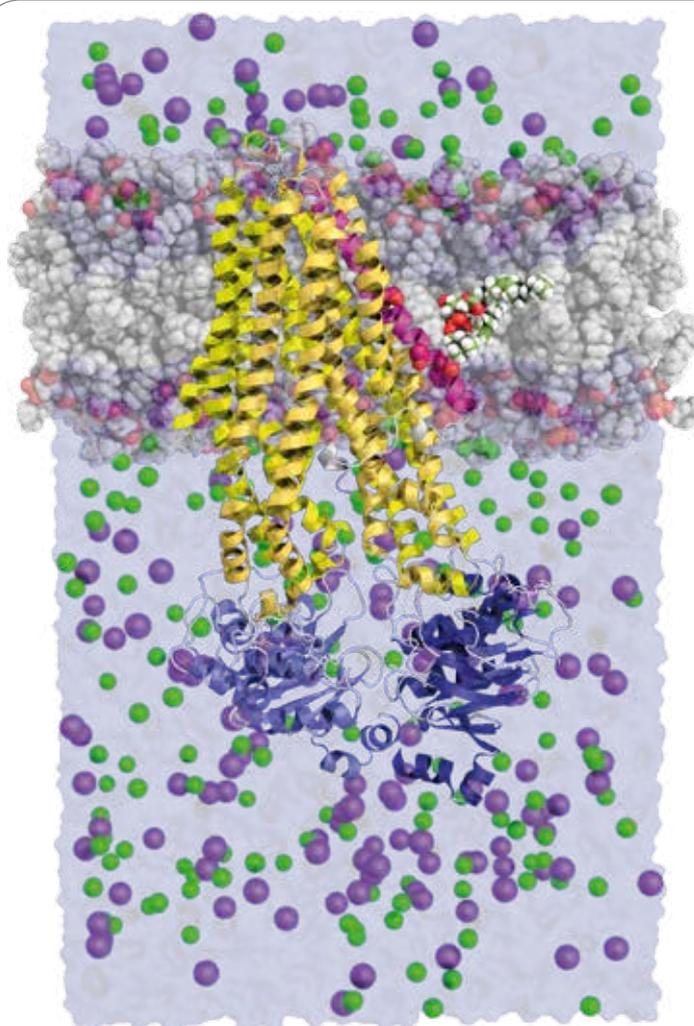
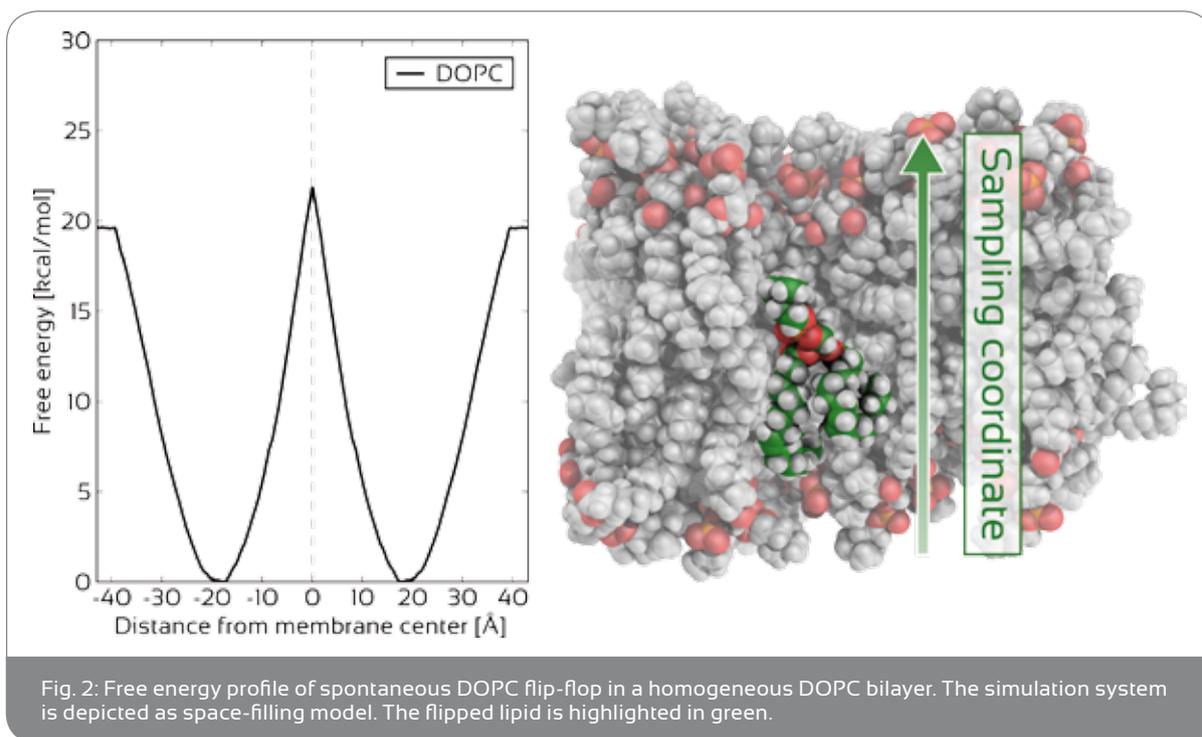


Fig. 1: Simulation system containing MDR3 (yellow and blue) embedded in a 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) bilayer (grey). NaCl (purple and green spheres) at a concentration of 154 mM mimics conditions in the intra- and extracellular space. The DOPC molecule being flipped through the bilayer is depicted in green.



Results

Free energy profile of spontaneous flip-flop. As a reference for spontaneous phospholipid flip-flop, we calculated the free energy profile for the passage of a single 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) through a homogeneous DOPC bilayer. Figure 2 depicts the obtained profile, which displays an energy barrier of 20.8 kcal mol⁻¹ and agrees both qualitatively and quantitatively with calculated profiles for related lipid species [6].

Free energy profile of MDR3-assisted flip-flop.

The free energy profile obtained for DOPC flip-flop along TMH1 of MDR3 is remarkably different from the reference profile. As shown in Figure 3, the barrier height is decreased by

7.0 kcal mol⁻¹, rendering MDR3-mediated flip-flop approximately five orders of magnitude faster than spontaneous flip-flop. According to our hypothesis, this effect should primarily be attributable to salt bridge and hydrogen bond interactions between the DOPC head group and the side chains of the hydrophilic amino acids in TMH1. Indeed, the free energy profile of MDR3-mediated flip-flop starts to diverge from the reference profile once the phosphatidylcholine head group reaches a “buffer zone” in which interactions with the hydrophilic serine and threonine residues are possible. A similar profile calculated along TMH7 of MDR3, which is the structural equivalent of TMH1 in the second pseudohalf of the transporter but does not show hydrophilic amino acids exposed to the

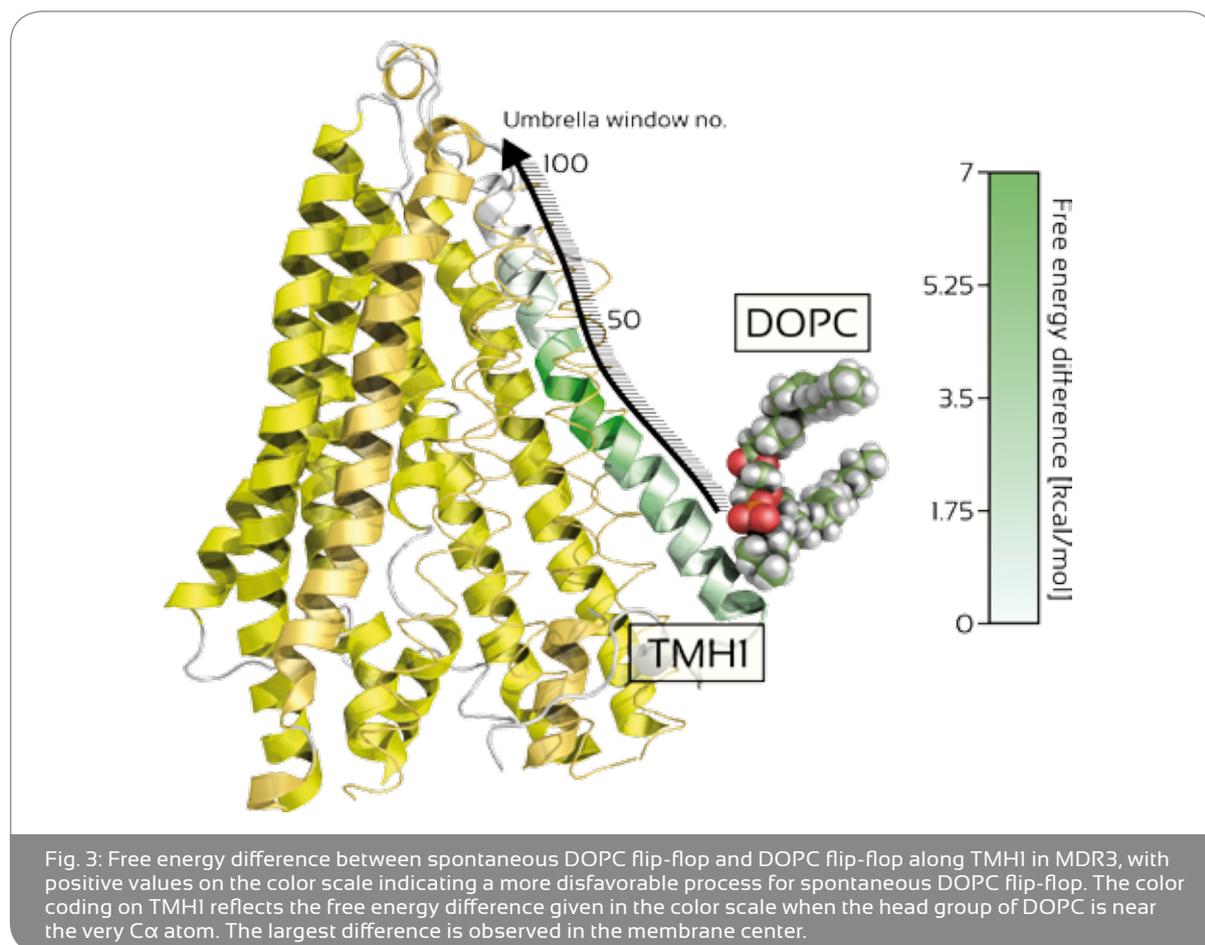


Fig. 3: Free energy difference between spontaneous DOPC flip-flop and DOPC flip-flop along TMH1 in MDR3, with positive values on the color scale indicating a more unfavorable process for spontaneous DOPC flip-flop. The color coding on TMH1 reflects the free energy difference given in the color scale when the head group of DOPC is near the very C α atom. The largest difference is observed in the membrane center.

membrane, does not display any reduction in barrier height. Thus, the obtained profile along TMH1 strongly supports our hypothesis of an alternative substrate translocation pathway for the lipid floppase MDR3.

Comparison to experimental data. While experimentally determined rates of spontaneous phosphatidylcholine flip-flop amount to 0.04 h^{-1} [7], the ATPase activity of MDR3 has been determined as 828 h^{-1} [8]. Assuming that one molecule of DOPC is translocated during each

ATPase cycle, MDR3-mediated DOPC transport is thus sped up by a factor of $\sim 2 \times 10^4$, relating to a lowering of the energy barrier by about 6 kcal mol^{-1} at 300 K . This value is in excellent agreement with our computed barrier height decrease of $7.0 \text{ kcal mol}^{-1}$.

Ongoing research / outlook. In order to rule out the possibility of a “classical”, cavity-mediated phospholipid translocation in MDR3, additional $3 \mu\text{s}$ of unbiased MD simulations on different MDR3 systems were performed. Using

configurations from these MD simulations as templates, we aim to construct free energy profiles for cavity-mediated DOPC transport in a similar way as presented for TMHI-mediated transport. The results obtained from our simulations will furthermore serve to guide mutation experiments for biochemical validation of our predictions.

References

- [1] van Helvoort, A., Smith, A. J., Sprong, H., Fritzsche, I., Schinkel, A. H., Borst, P., van Meer, G.: MDR1 P-glycoprotein is a lipid translocase of broad specificity, while MDR3 P-glycoprotein specifically translocates phosphatidylcholine. *Cell* (87), 507-517, 1996.
- [2] Kluth, M.: Molecular in vitro analysis of the human ABC transporter MDR3, Doctoral Thesis, Düsseldorf, 2014.
- [3] Perez, C., Gerber, S., Boilevin, J., Bucher, M., Darbre, T., Aebi, M., Reymond, J. L., Locher, K. P.: Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* 433-438, 2015.
- [4] Brunner, J. D., Lim, N. K., Schenck, S., Duerst, A., Dutzler, R.: X-ray structure of a calcium-activated TMEM16 lipid scramblase. *Nature* (516), 207-212, 2014.
- [5] Case, D. A., Betz, R. M., Botello-Smith, W., Cerutti, D. S., Cheatham, T. E., Darden, T. A., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., Lin, C., Luchko, T., Luo, R., Madej, B., Mermelstein, D., Merz, K. M., Monard, G., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Sagui, C., Simmerling, C. L., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., Xiao, L., York, D. M., Kollman, P. A.: AMBER 2016. University of California, San Francisco, 2016.
- [6] Tieleman, D. P., Marrink, S. J.: Lipids out of equilibrium: energetics of desorption and pore mediated flip-flop. *Journal of the American Chemical Society* (128), 12462-12467, 2006.
- [7] Kornberg, R. D., McConnell, H. M.: Inside-outside transitions of phospholipids in vesicle membranes. *Biochemistry* (10), 1111-1120, 1971.
- [8] Kluth, M., Stindt, J., Dröge, C., Linnemann, D., Kubitz, R., Schmitt, L.: A mutation within the extended X loop abolished substrate-induced ATPase activity of the human liver ATP-binding cassette (ABC) transporter MDR3. *Journal of Biological Chemistry* (290), 4896-4907, 2015.

Written by Michele Bonus, Isabelle Eichler, Holger Gohlke

Department of Mathematics and Natural Sciences, Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

Michele Bonus graduated in pharmacy from the Heinrich Heine University Düsseldorf in 2012 and has been working in the research group of Holger Gohlke since 2013. Beyond standard MD simulations, he works with the AMBER implementations of accelerated MD (aMD), targeted MD (tMD), and adaptively biased MD (ABMD), as well as techniques for the calculation of free energies from MD simulations.

Isabelle Eichler graduated in pharmacy from the Heinrich Heine University Düsseldorf in 2016 and is currently a trainee pharmacist in the research group of Holger Gohlke.

Holger Gohlke is Professor of Pharmaceutical and Medicinal Chemistry and Managing Director of the Institute of Pharmaceutical and Medicinal Chemistry at the Heinrich Heine University, Düsseldorf. After appointments as Assistant Professor in Frankfurt and Professor in Kiel, he moved to Düsseldorf in 2009. As of 2017, he also heads the NIC research group "Computational Biophysical Chemistry" at the Forschungszentrum Jülich. His current research focuses on the understanding, prediction, and modulation of interactions involving biological macromolecules from a computational perspective.

Contact: Michele Bonus, michele.bonus@hhu.de
Isabelle Eichler, isabelle.eichler@hhu.de
Holger Gohlke, gohlke@hhu.de

Dark Matter Properties from a Supercomputer



During the last decades, cosmology has delivered a grand wealth of groundbreaking discoveries. The understanding of the Universe at very large and very small scales advanced tremendously, both by experimental work, e.g. by the measurement of the accelerated expansion of the current Universe [1, 2] and by theoretical work, e.g. by the determination of the nature of the QCD phase transition in the early Universe [3]. Very often theoretical advances drive experimental ones, like the prediction of gravitational waves finally lead to “first light” for a gravitational wave observatory [4]. It also works the other way round just as frequently. For example high precision measurements revealed [5], that the usual type of matter accounts only for about 5% of the total energy content of the Universe, with a factor 5 more attributed to Dark Matter and the rest to Dark Energy. In turn, theory has proposed several candidates for Dark Matter to be confirmed or ruled out by experiment.

A theoretically well motivated Dark Matter candidate is the axion particle. It explains a peculiar feature of the strong interaction: it is surprisingly symmetric under the P-transformation, the transformation which exchanges left and right. To see, what is surprising about this, let us go back in time to the last century.

In the first half of the 20th century most physicist thought, that all fundamental processes

are symmetric under the P-transformation: left behaves the same as right. There is no way to distinguish the two from each other. A revolution came in the late 50’s, when P-violation was found in an experiment, which was studying the weak interaction: a certain weak process happened many more times than its mirror image process. It turned out, that Nature does actually make a difference between left and right. The result was shocking for most physicists at that time. While many of them were still recovering, Pauli, one of the most brilliant among his contemporaries, immediately recognized, that the real problem, that has to be explained is:

I am not so shocked about the fact, that GOD is left handed, but much more that as a left-handed he appears to be symmetric in his strong actions. [...] Why is the strong interaction left-right symmetric?¹

Later Pauli’s remark was laid on solid theoretical foundations. The theory of the strong interaction was developed, it is called Quantum-chromodynamics (QCD). It describes how quarks and gluons, the constituents of protons and neutrons, interact. It is possible to introduce a parameter into QCD, which violates the P-symmetry, it is usually called θ [6]. A-priori the value of θ can be any number, like many other fundamental parameters in particle physics. However experimentally it is found to be consistent with zero, with an extremely

¹ “Ich bin nicht so sehr durch die Tatsache erschüttert, dass der HERR die linke Hand vorzieht, als vielmehr durch die Tatsache, dass er als Linkshänder weiterhin symmetrisch erscheint, wenn er sich kräftig ausdrückt. Kurzum, das eigentliche Problem scheint jetzt in der Frage zu liegen: Warum sind starke Wechselwirkungen linksrechts symmetrisch?” See eg. in Martin Gardner: The Ambidextrous Universe, 1967.

good precision. The current bound on θ is $\theta \leq 10^{-10}$ [7], thus QCD is P-symmetric to a very good approximation. Such a fine-tuning begs for an explanation.

A nice explanation was given by Peccei and Quinn [8]: instead of considering θ as a fixed parameter, they suggested to treat it as a dynamical field, whose value can change with space and time. Now, if the potential of this field is such that it has a minimum at $\theta = 0$, then the θ field evolves in time and relaxes to this minimal value, and effectively explains, why θ is small. Theorists have come up with all sorts of possible potentials, that have their minima at $\theta = 0$. The immediate consequence is, the existence of a very weakly interacting particle: the axion [9, 10]. Whether this is the correct explanation, and whether axions exist is not known. There are several experiments around the world looking for them, none of them successful yet.

It was realized soon after the proposal, that since axions couple weakly to ordinary matter, they are perfect candidates for Dark Matter. Even if they interact weakly, they could be produced in sufficient amount during the Big-Bang. Assuming that axions are the only source of Dark Matter, one can calculate their mass. For this we have to know how the axion potential looks like today and how it looked like during the Big-Bang.

In this project we calculated the axion potential from QCD for the whole history of our Universe.

From this we gave an estimate on the axion's mass [11]. This can help to design future experiments looking for these particles.

Objectives, challenges and methods

Our goal was to calculate the axion potential for the whole history of our Universe. At early times the Universe was much hotter than now, and with expanding it cooled down rapidly to reach its current temperature today (0.235 meV). So what we needed is the temperature dependence of the axion potential. For the simplest axion model, the potential only receives contributions from QCD. To compute this, the equations of QCD had to be solved. As QCD is a highly non-linear theory and its coupling constant is not particularly small, a non-perturbative technique is required to work out its properties, for which we used the lattice discretization of QCD.

In performing these computations we faced two challenges. The first one is an algorithmic issue. Determination of the axion potential using a standard lattice QCD algorithm is analogous to the following simple problem: one has to determine the ratio of red and blue balls in a black bag, by randomly picking balls from the bag. See Figure 1/a for an illustration. The ratio is essentially the curvature of the axion potential. For small temperatures there are similar number of red and blue balls in the bag, by picking a few hundred one can give a very good estimate on the ratio. However as the temperature is increased, the ratio drops rapidly and one needs more and more random picks, which costs more and more CPU time. To calculate the potential

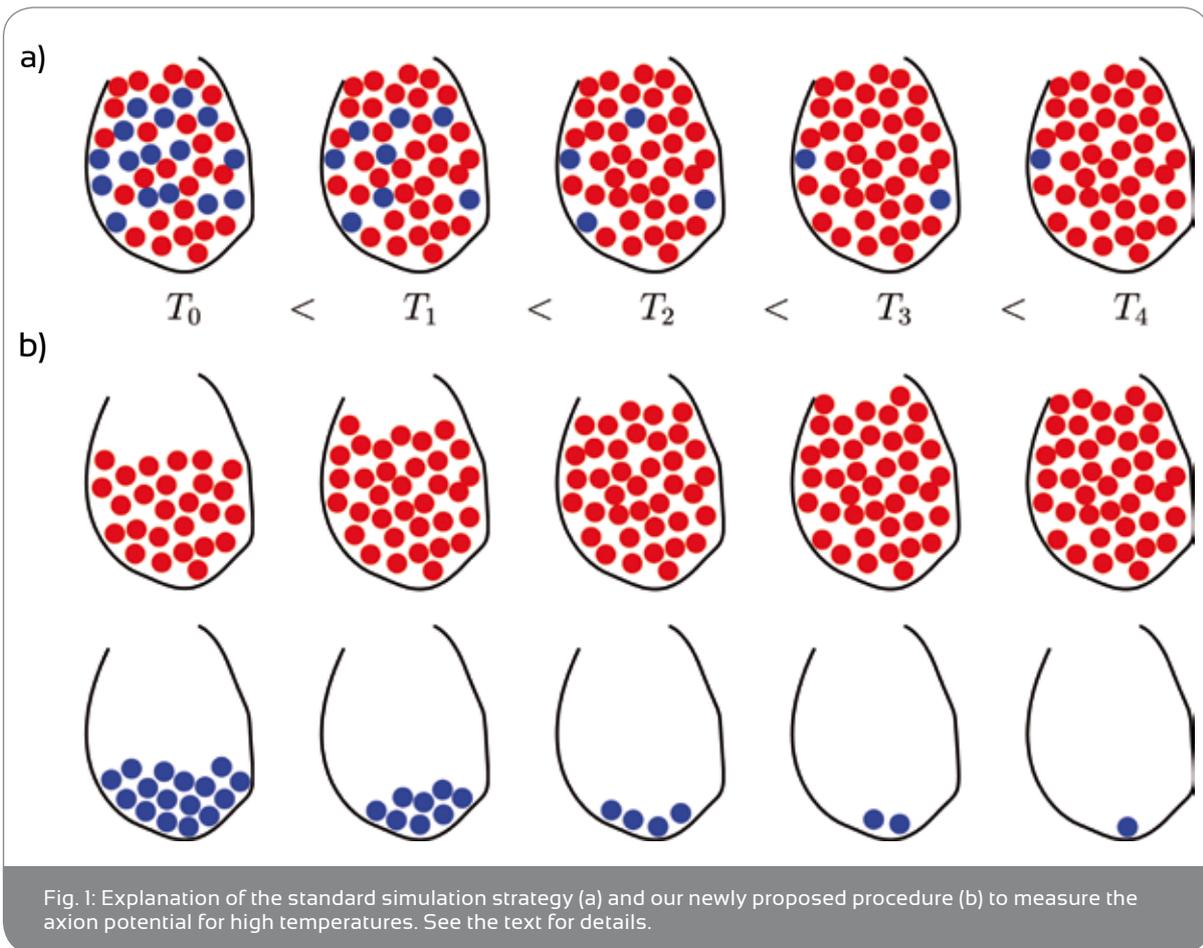


Fig. 1: Explanation of the standard simulation strategy (a) and our newly proposed procedure (b) to measure the axion potential for high temperatures. See the text for details.

using this standard approach in the interesting temperature region (~ 2 GeV) one needs about 10^{10} years of computational time even on a supercomputer.

We came up with an alternative procedure, see Figure 1/b. First at a relatively small temperature, let us call it T_0 , where the standard approach was still feasible, we measured the ratio in the usual way. Then for higher temperatures we separated the balls into two bags, and carried out simulations with either only red or

only blue balls. Then we have measured, how the number of blue/red balls were changing as the temperature was increased. Using these temperature differences plus the starting value of the ratio at T_0 we could then calculate the ratio at higher temperatures.

Another challenge was related to the large discretization artefacts in the axion potential. These artefacts are on the 10% level in typical lattice QCD simulations, and can be get rid of by performing the so-called continuum extrapolation

procedure. Here one takes lattices with smaller and smaller lattice spacing and performs an extrapolation to the continuum limit. The axion potential turned out to have much larger errors, and the continuum extrapolation to be much more difficult than usual. Here we also designed a new procedure, to get rid of the large discretization artefacts and demonstrated in several cases the effectiveness of the procedure.

Beside the axion potential one also needs equations governing the expansion of the Universe to determine the mass of the axion. The expansion is governed by the visible matter content, for which the thermodynamical properties, pressure and energy density, had to be determined. In our paper we also calculated the QCD component of these. Here the challenge was to go up to a sufficiently high temperature, where the non-perturbative lattice QCD results could be connected to perturbation theory. We were able to reach a temperature of 1 GeV, after which we smoothly connected to known thermodynamics results from perturbation theory.

Results and outlook

In Figure 2 and Figure 3 we show the two main results of our work [11]. The first is the temperature dependence of the curvature of the axion potential. This is the first determination of this quantity from first principles in a range of temperatures relevant for axion cosmology with control over all errors. We extended significantly the reach of previous lattice determinations, and managed to give a result where all systematic errors were estimated. From the second plot one

can read off the energy density and pressure of our Universe in a temperature range of five orders of magnitude.

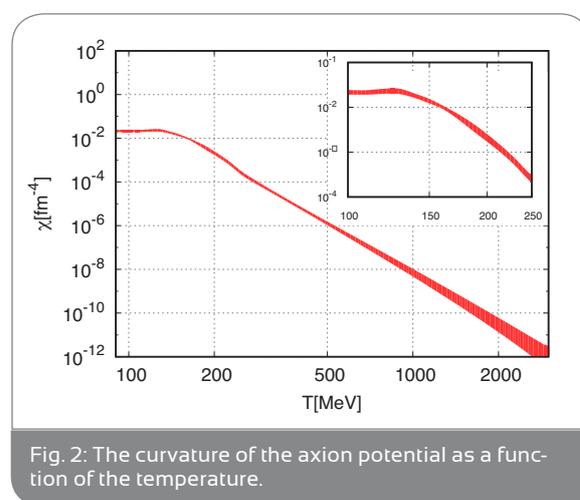


Fig. 2: The curvature of the axion potential as a function of the temperature.

Combining these two results one obtains the mass of the axion: $m_A = 28(1)\mu\text{eV}$. This number assumes, that all Dark Matter is made of axions and also assumes the simplest cosmological production scenario. It is important to mention, that there exist more complicated axion production scenarios and also more complicated axion like particle models. The current best estimation of these scenarios increases the axion mass together with its uncertainty considerably: $50\mu\text{eV} \leq m_A \leq 1500\mu\text{eV}$. It is an important, though very non-trivial task to decrease the size of these uncertainties in the future.

The resulting value for the mass is an important hint for experimentalists how to design experiments looking for axion particles in the near future. If the experimental search succeeded, the axion would be the first confirmed



constituent of Dark Matter and an evidence for physics of an unknown world.

Acknowledgements

We thank our colleagues, Sz. Borsanyi, J. Gunther, S. Katz, T. Kawanai, T. Kovacs, A. Pasztor, A. Ringwald, J. Redondo, for a fruitful collaboration. We also thank M. Dierigl, M. Giordano, S. Krieg, D. Nogradi and B. Toth for useful discussions. This project was funded by the DFG grant SFB/TR55, and by OTKA under grant OTKA- K113034. The work of J.R. is supported by the Ramon y Cajal Fellowship 2012-10597 and FPA2015- 65745-P (MINECO/FEDER). The computations were performed on JUQUEEN at Forschungszentrum Jülich (FZJ), on SuperMUC at Leibniz Supercomputing Centre in München, on Hazel Hen at the High Performance Computing Center in Stuttgart, on QPACE in Wuppertal and on GPU clusters in Wuppertal and Budapest.

References

- [1] **A. G. Riess et al. [Supernova Search Team]:**
"Observational evidence from supernovae for an accelerating Universe and a cosmological constant," *Astron. J.* 116 (1998) 1009
- [2] **S. Perlmutter et al. [Supernova Cosmology Project Collaboration]:**
"Measurements of Omega and Lambda from 42 high redshift supernovae," *Astrophys. J.* 517 (1999) 565
- [3] **Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, K. K. Szabo:**
"The Order of the quantum chromodynamics transition predicted by the standard model of particle physics," *Nature* 443 (2006) 675
- [4] **B. P. Abbott et al. [LIGO Scientific and Virgo Collaborations]:**
"Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.* 116 (2016) no.6, 061102
- [5] **P. A. R. Ade et al. [Planck Collaboration]:**
"Planck 2013 results. XVI. Cosmological parameters," *Astron. Astrophys.* 571 (2014) A16
- [6] **G. 't Hooft:**
"Symmetry Breaking Through Bell-Jackiw Anomalies," *Phys. Rev. Lett.* 37 (1976) 8.
- [7] **J. M. Pendlebury et al.:**
"Revised experimental upper limit on the electric dipole moment of the neutron," *Phys. Rev. D* 92 (2015) no.9, 092003
- [8] **R. D. Peccei and H. R. Quinn:**
"CP Conservation in the Presence of Instantons," *Phys. Rev. Lett.* 38 (1977) 1440.
- [9] **S. Weinberg:**
"A New Light Boson?," *Phys. Rev. Lett.* 40 (1978) 223.
- [10] **F. Wilczek:**
"Problem of Strong p and t Invariance in the Presence of Instantons," *Phys. Rev. Lett.* 40 (1978) 279.
- [11] **S. Borsanyi et al.:**
"Calculation of the axion mass based on high-temperature lattice quantum chromodynamics," *Nature* 539 (2016) no.7627, 69

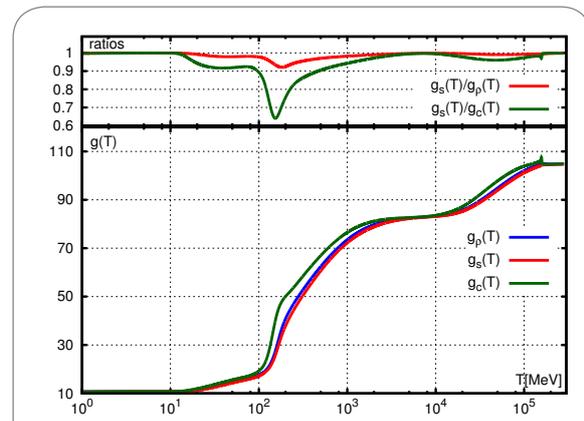


Fig. 3: Thermodynamical properties (energy, entropy density and speed of sound) of our Universe as the function of the temperature.

Written by Zoltan Fodor, Simon Mages and Kalman Szabo

(University of Wuppertal and Jülich Supercomputing Centre (JSC), Germany)

Contact: Kalman Szabo, k.szabo@fz-juelich.de

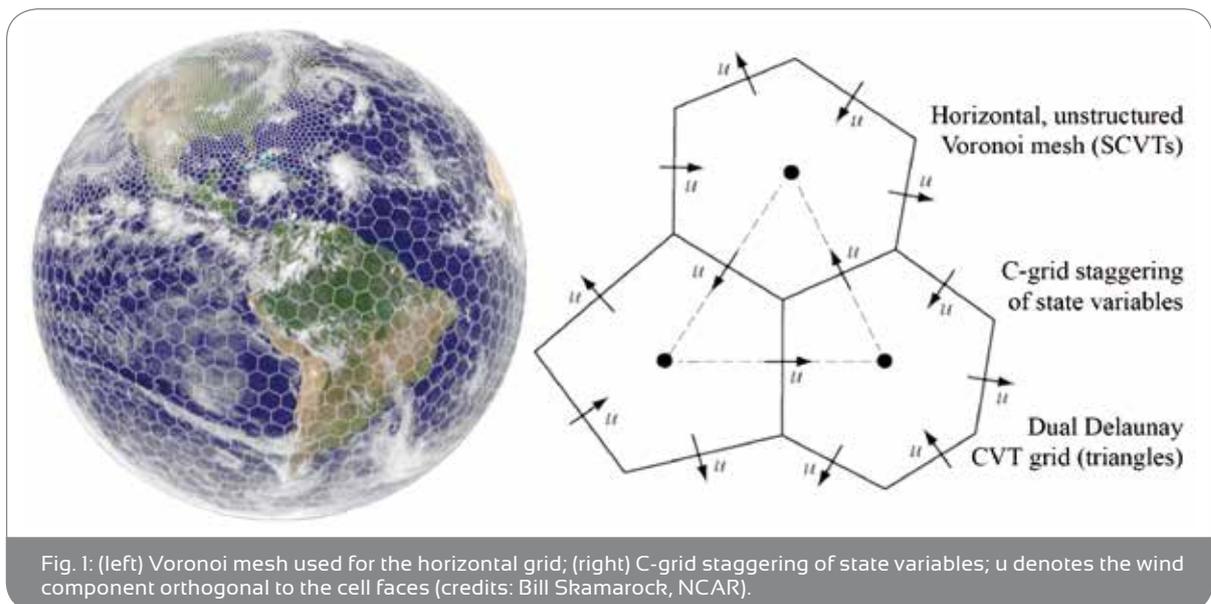
Towards Exascale Computing with the Model for Prediction Across Scales

The Model for Prediction Across Scales is a suite of earth system modelling components designed with massively parallel applications on current and next-generation HPC platforms in mind. In this contribution, we present steps towards exascale computing that allow global atmospheric simulations with a convection-permitting grid spacing that requires tens of millions of grid columns.

The weather- and climate-modelling community is seeing a shift in paradigm from limited area models towards novel approaches involving global, complex and irregular meshes. A promising and prominent example therefore is the Model for Prediction Across Scales (MPAS, Skamarock et al. 2012). MPAS is a novel set of earth system simulation components and consists of an atmospheric core, an ocean core, a land-ice core and a sea-ice core. Its distinct

features are the use of unstructured Voronoi meshes and C-grid discretisation (see Fig. 1) to address shortcomings of global models on regular grids and the use of limited area models nested in a forcing data set, with respect to parallel scalability, numerical accuracy and physical consistency. The unstructured Voronoi meshes employed by all MPAS cores allow variable mesh refinement across the globe with smooth transitions between areas of different resolution, making it possible to conduct global modelling experiments at reasonable computational costs.

Yet, with exascale computing projected for the end of this decade and in light of the fact that energy requirements and physical limitations will imply the use of accelerators and the scaling out to orders of magnitudes larger numbers of cores than today, it is paramount to prepare





modern codes like MPAS for this future. In two extreme scaling experiments on FZJ (Research Centre Juelich) JUQUEEN in 2015 and on LRZ SuperMUC in 2016, the atmospheric core, MPAS-A, has been tested on up to nearly half a million cores for a global 3km mesh with more than 65 million grid columns (Heinzeller et al. 2016). These experiments highlighted several bottlenecks, such as the disk I/O and the initial bootstrapping process during the model setup, that need to be addressed to fully exploit the capabilities of exascale computing systems. The dynamical solver of MPAS-A, on the other hand, maintained a very high parallel efficiency up to 24,576 nodes on JUQUEEN. For larger numbers of nodes, the parallel efficiency dropped quickly. This is related to the number of owned cells versus the number of halo cells per MPI task, for which data need to be exchanged at every time step between adjacent patches (Fig. 2).

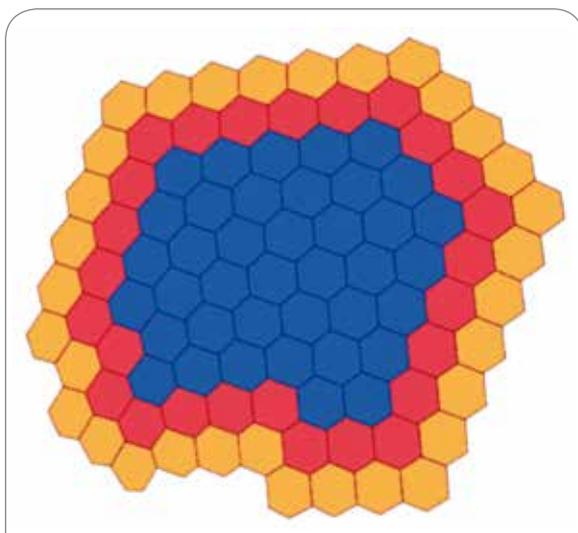


Fig. 2: Block of cells owned by an MPI task and two layers of halo cells.

Methods

The I/O in the current release, version 5.0, of MPAS is facilitated through the Parallel I/O library (PIO, <https://github.com/NCAR/ParallelIO>), a wrapper around the commonly used netCDF4/HDF5 and parallel-netCDF libraries. The parallelism in MPAS is based on the dual grid of the Voronoi mesh, the Delaunay triangulation, which connects the cell centres. In an offline pre-processing step, the METIS graph partitioning software is used to divide the globe into separate patches for each MPI task (Fig. 3). By using an efficient hybrid MPI+OpenMP parallelisation, it is possible to reduce the number of MPI tasks and thus to improve the ratio of owned cells to halo cells.

Implementation of alternative I/O layer. We implemented an alternative I/O layer based on the SIONlib library (<http://www.fz-juelich.de/ias/jsc/EN/Expertise/Support>), designed for massively parallel applications. Our implementation is designed for maximum flexibility and can be used together with the existing I/O capabilities. When SIONlib is chosen, each MPI task writes its share of the data into a separate block of the SIONlib output files. For good parallel write performance, it was paramount to align the size of the block owned by an MPI task with the file system block size.

The SIONlib implementation writes the data on disk in a scrambled format: each task writes the data for its owned and halo cells, using its own local index arrays. The standard I/O formats in MPAS use the PIO library to re-arrange all data

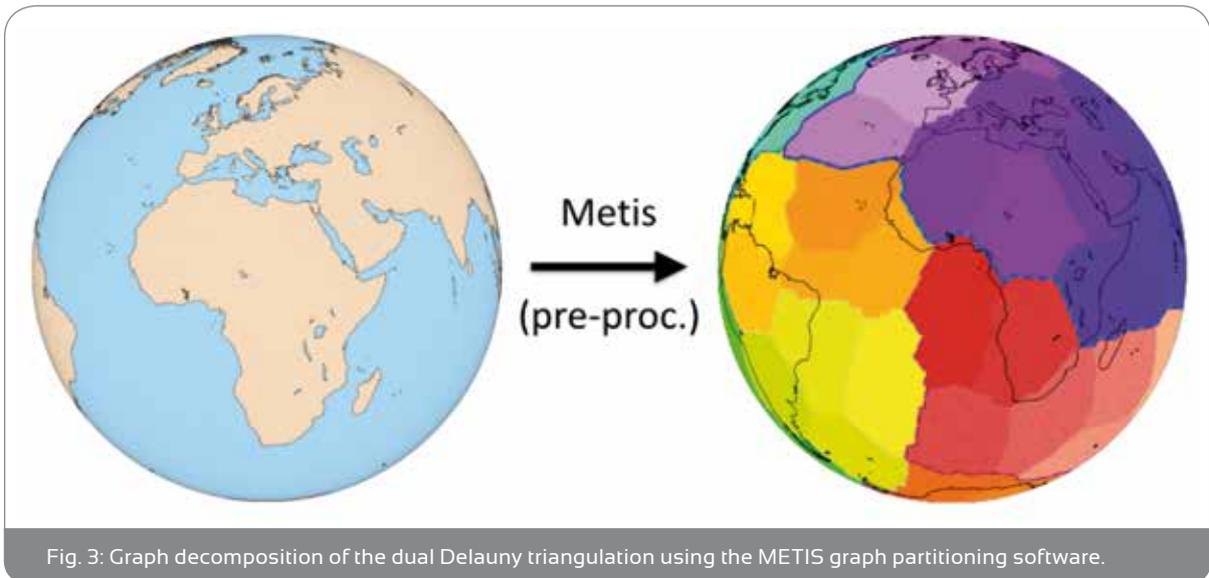


Fig. 3: Graph decomposition of the dual Delaunay triangulation using the METIS graph partitioning software.

into a global list of cells, vertices and edges, independent of the number of MPI tasks. For this reason, reading data in SIONlib I/O format requires the same number of MPI tasks and the same graph decomposition that were used for writing the data. While this may be seen as a limitation, it offers huge potential to speed up other parts of a typical MPAS model run (see next section).

Speedup of model initialisation. The model initialisation contains two steps in which most of the time is spent: the reading of initial conditions from either a “cold start” file or from a restart file from a previous run, and the bootstrapping process. The new SIONlib I/O layer can be used to reduce the times required not only for the reading of the initial conditions, but also for the bootstrapping process. As the data are written to disk in a task- and decomposition-dependent order, they already contain a large share

of the information that is otherwise calculated during the bootstrapping process. This alternative bootstrapping method was implemented in the code, and further possibilities to reduce the model initialisation times were identified.

Improvement of hybrid parallelisation. The hybrid parallelisation in the atmospheric core is implemented for the dynamical solver, which is responsible for the time integration of geophysical fluid flow equations, and does not affect the model initialisation or file I/O. Since most of the time in realistic applications is spent in the dynamical solver, small improvements made to the hybrid code can have large effects.

In the original version of the hybrid MPAS-A code, parallel sections are created repeatedly around individual steps of the time integration routine that benefit from a shared-memory parallelisation within each MPI task. For larger



numbers of threads, which are required for effective use of many-core systems such as the Intel Xeon Phi Knights Landing (KNL), this adds significant overhead to the model execution. The solution developed here solves this problem by creating the threads for the entire lifetime of one call to the time integration routine. The different sub-steps of the time integration are wrapped in OpenMP parallel-do loops or in master-only sections as needed.

Results

Parallel disk I/O. The new I/O layer was primarily designed for massively parallel applications of the code. For serial runs or small-scale applications of MPAS, the read and write performance with SIONlib is comparable to netCDF. For medium- to large-scale applications, the performance gain using SIONlib is significant and encourages its use for internal data (i.e., data used only by MPAS) and possible also for external data (i.e., data used by other applications such as visualisation tools) in combination with a post-processing and conversion tool.

For large- to extreme-scale applications, the speedup using SIONlib can reach a factor of 10 compared to parallel-netCDF (and even a factor of 60 compared to netCDF4) for read and write operations and justifies its usage for both internal and external data. This way, a model integration on several thousand cores can proceed quickly by dumping its output to disk in SIONlib format, instead of being held up by using a much slower netCDF format.

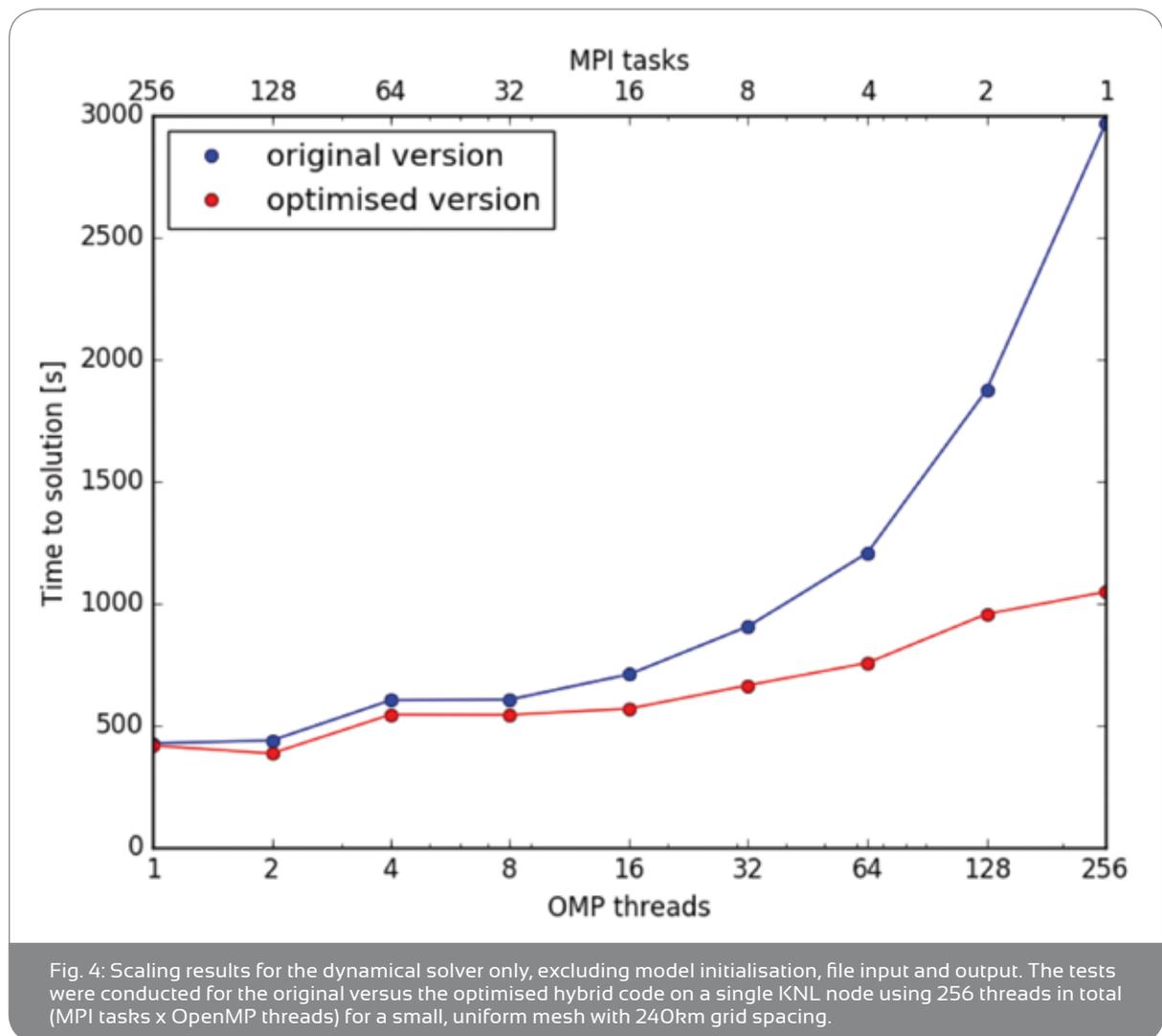
Model initialisation. Bootstrapping directly from SIONlib files reduces the model initialisation costs significantly. This becomes apparent from Table 1 for the largest parallel configurations tested on SuperMUC. On this system, the inter-island communication is slower than the intra-island communication (one island consists of 512 nodes). Bootstrapping from netCDF files takes more than 600s on 2048 nodes and can reach up to one hour on 4096 nodes (not shown). Using SIONlib instead, the initial bootstrapping times remain nearly constant and are a factor of 5 to 6 smaller for 2048 nodes on SuperMUC.

Mesh (#cells)	Nodes	P.P.N.	I/O in	I/O out	Bootstrapping	Read	Write
2km (147,456,002)	1536	16 x 1	netCDF4	netCDF4	582	701	6142
2km (147,456,002)	1536	16 x 1	pnetCDF	pnetCDF	568	353	1086
2km (147,456,002)	1536	16 x 1	SIONlib	SIONlib	110	18	62
2km (147,456,002)	2048	16 x 1	pnetCDF	pnetCDF	619	301	1081
2km (147,456,002)	2048	16 x 1	SIONlib	SIONlib	111	14	95

Table 1: Timing results for an extreme-scaling experiment on LRZ SuperMUC, focusing on different I/O formats and times required (in seconds) for the initial bootstrapping process, stream input (read) and stream output (write). P.P.N. denotes the number of MPI x OpenMP processes per node.

Hybrid parallelisation. The differences in the hybrid version of the atmospheric solver suggest improvements predominantly for large numbers of OpenMP threads on modern many-core systems such as the KNL. The significance of the relatively small change to the hybrid implementation of the dynamical solver are demonstrated by a KNL single-node tests on the 240km mesh

(Fig. 4). From 256 x 1 to 16 x 16 threads (MPI x OpenMP), the differences between the two versions of the code are small. The performance of the solver decreases steeply for larger numbers of OpenMP tasks in the original version of the code, but only slowly in the new version. In the extreme case of 1 x 256 tasks on a single KNL, the new hybrid code is three times as fast as the





old version. While the dynamical solver alone is slower the more OpenMP threads and the less MPI tasks are used, the remaining parts of the model run (initialisation, file I/O) are faster (not shown). The overall performance of the new hybrid code is thus highly competitive with other hybrid implementations used in atmospheric modelling.

Conclusions

The addition of a new SIONlib I/O layer, designed primarily for massively parallel applications, and the optimisation of the hybrid version of the dynamical solver led to demonstrated performance improvements from small to extreme scales on current mainstream architectures and next-generation many-core systems. As a side effect, the model setup costs, stemming from the initial bootstrapping process, are reduced by either of the two methods. While the SIONlib I/O layer allows to bypass the non-scaling parts of the bootstrapping process and, at the same time, lowers the memory footprint, a more efficient hybrid parallelisation allows to reduce the number of MPI tasks and this way speeds up the model setup. The two development paths taken in this project are complementary and represent one step towards reaching exascale capabilities with the Model for Prediction Across Scales. The work presented in this contribution

was done within a KONWIHR IV project using computational resources provided by LRZ on SuperMUC.

References

- [1] **Heinzeller D., Duda M.G., Kunstmann H.:**
Towards convection-resolving, global atmospheric simulations with the Model for Prediction Across Scales (MPAS) v3.1: an extreme scaling experiment, *Geosci. Model Dev.*, 9, 77-110, 2016; doi: 10.5194/gmd-9-77-2016
- [2] **Skamarock, W.C., Klemp, J.B., Duda, M.G., Fowler, L., Park, S.-H., Ringler, T.D.:**
A Multi-scale Nonhydrostatic Atmospheric Model Using Centroidal Voronoi Tessellations and C-Grid Staggering, *Monthly Weather Review*, 240, 3090-3105, 2012; doi:10.1175/MWR-D-11-00215.1

Written by Dominikus Heinzeller^{1,2}, Michael G. Duda³

¹Augsburg University, Institute of Geography, Augsburg, Germany

²Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Garmisch-Partenkirchen, Germany

³National Center for Atmospheric Research, Mesoscale and Microscale Meteorology Laboratory, Boulder, CO, USA

Contact: Dominikus Heinzeller, heinzeller@kit.edu

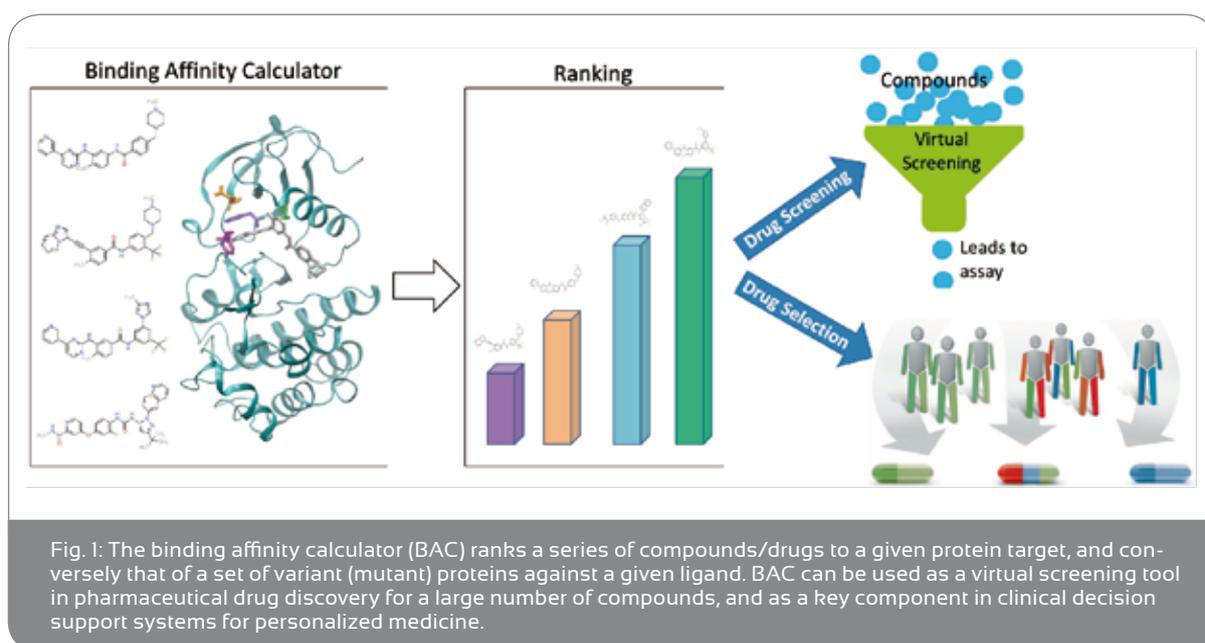
Rapid and Accurate Calculation of Ligand-Protein Binding Free Energies

Most drugs work by binding to specific proteins and blocking their physiological functions. The binding affinity of a drug to its target protein is hence a central quantity in pharmaceutical drug discovery and clinical drug selection (Fig. 1). For successful uptake in drug design and discovery, reliable predictions of binding affinities need to be made on time scales which influence experimental programs. For applications in personalized medicine, the selection of suitable drugs needs to be made within a few hours to influence clinical decision making. Therefore, speed is of the essence if we wish to use free energy based calculation methods in these areas. Our work is on developing an automatic workflow which ensures that the binding affinity results are accurate and reproducible, and can be delivered rapidly.

Computations

The binding affinity calculations would be very lengthy, tedious, and error-prone to perform manually. They consist of a large number of steps, including model building, production MD and data analytics performed on the resulting trajectory files. To perform modeling and calculation with optimal efficiency, we have developed the Binding Affinity Calculator (BAC) [1], a highly automated molecular simulation based free energy calculation workflow tool (Fig. 1). Its execution is much faster and more error-proof when performed in an automated fashion. A user-friendly version of BAC, namely uf-BAC, has been developed to extend its accessibility to nontechnical users.

Two approaches are included in BAC for the binding free energy calculations of ligands





to proteins. One is ESMACS (enhanced sampling of molecular dynamics with approximation of continuum solvent) [2]; the other is TIES (thermodynamic integration with enhanced sampling) [3]. The underlying computational method is based on classical molecular dynamics (MD). In MD simulations, macroscopic properties corresponding to experimental observables are defined in terms of ensemble averages. Free energy is such a measurement. ESMACS and TIES use ensemble averaging and the recognition of the Gaussian random process (GRP) properties computed from MD trajectories. On multi-core machines such as SuperMUC, ensemble simulations play into our hands because, in the time it takes to perform one such calculation, all of the members of an ensemble can be computed. The method is therefore fast, with free energies being determined within around 8 hours.

We have found that an ensemble consisting of ca 25 replicas for an ESMACS study, and an ensemble of a minimum of 5 replicas for a TIES study are required per free energy calculation in order to guarantee reproducibility of predictions. Our approaches have now been standardized; ESMACS and TIES have been applied by us to over 20 different sets of compounds and protein targets, of which many have been performed using the substantial allocation of cycles on the GCS (Gauss Centre for Supercomputing) Supercomputer SuperMUC at Leibniz Supercomputing Centre.

In an unprecedented project, we ran a Giant Workflow on Phases 1 and 2 of SuperMUC, from which more than 60 free energy calculations were performed in 37 hours between 11-13 June 2016 prior to a system maintenance. This was the first time that both phases of SuperMUC were jointly allocated exclusively for one project. The accumulated compute power of phase 1 and 2 amounts to 5.71 PFlop/s Linpack, which would have ranked it in 9th position in the top 500 Supercomputer list (June 2016). In contrast to a monolithic application, which would span the whole machine via MPI, the big challenge was to keep the system busy by the multiple threads of the workflow. This required constant monitoring of the job's progress and immediate fixing of problems during the time of the run.

We not only attained all our planned objectives for the Giant run but achieved even more than anticipated, thanks to the exceptional performance of the computer. The GCS issued a press release, and distinguished science writer and journalist Dr. Roger Highfield wrote a blog post about the experience [4].

Results

With the resource allocation in the current project, we have applied ESMACS and TIES to study a wide range of proteins which have diverse functions in the human body and are important targets for pharmaceutical drug design and discovery, and for clinical therapies. We have made very important progress in our research this year. We have been able to produce rapid, reliable, accurate and precise predictions of binding

free energies using both ESMACS and TIES. Studies of many of the molecular systems have been completed and results published [3, 5, 6], others are either at the post-processing stage or at earlier stages where more simulations and calculations are required. Our predictions from ensemble simulations, some of them performed blindly, are in good agreement with experimental findings, including those released to us by leading pharmaceutical companies worldwide after our computational predictions were made [3, 5, 6]. Our findings have demonstrated that this approach is able to deliver an accurate ranking of ligand binding affinities quickly and reproducibly.

We have recently reported the performance of the TIES approach when applied to a diverse set of protein targets and ligands [3]. The results (Fig. 2) are in very good agreement with experimental data (90% of calculations agree to within 1 kcal/mol), while the method is reproducible by construction. Statistical uncertainties of the order of 0.5 kcal/mol or less are achieved.

In direct collaborations with two pharmaceutical companies, our approaches were tested in a realistic pharmaceutical setting [5, 6]. The calculations were performed, initially blind, to investigate the ability of our methods to reproduce the experimentally measured trends. Good

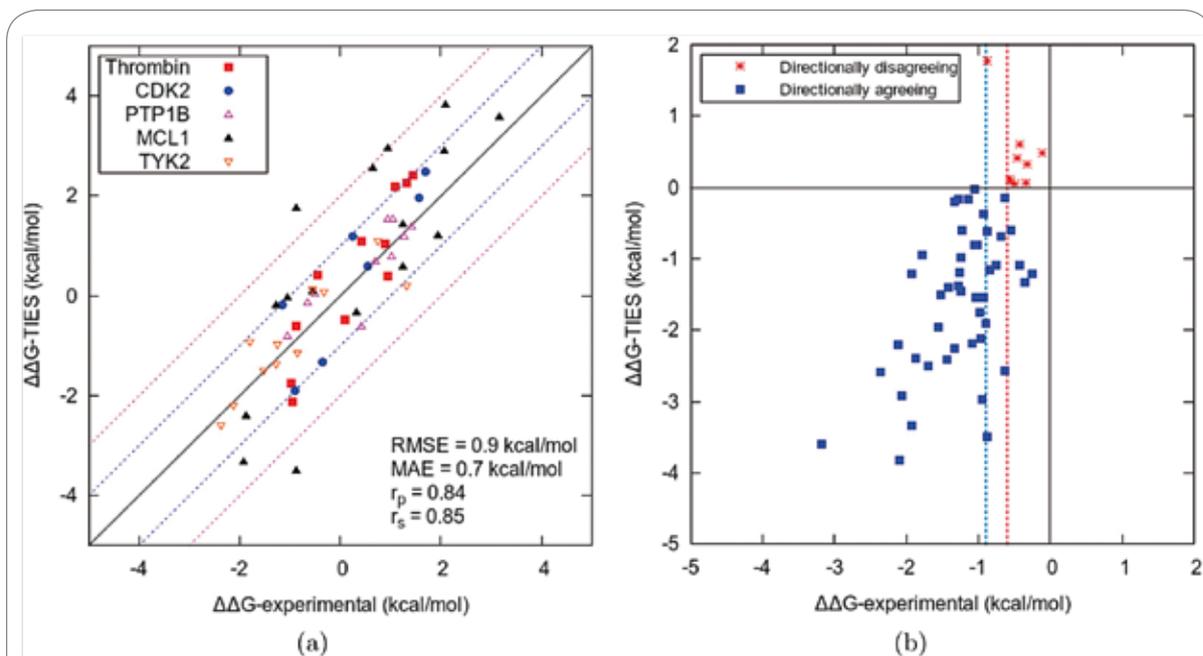


Fig. 2: (a) Good correlation has been obtained between TIES-predicted relative binding affinities and experimental data, and (b) the 84% of correctness has been achieved for directional agreement, meaning the computational prediction and experimental observation have the same sign.

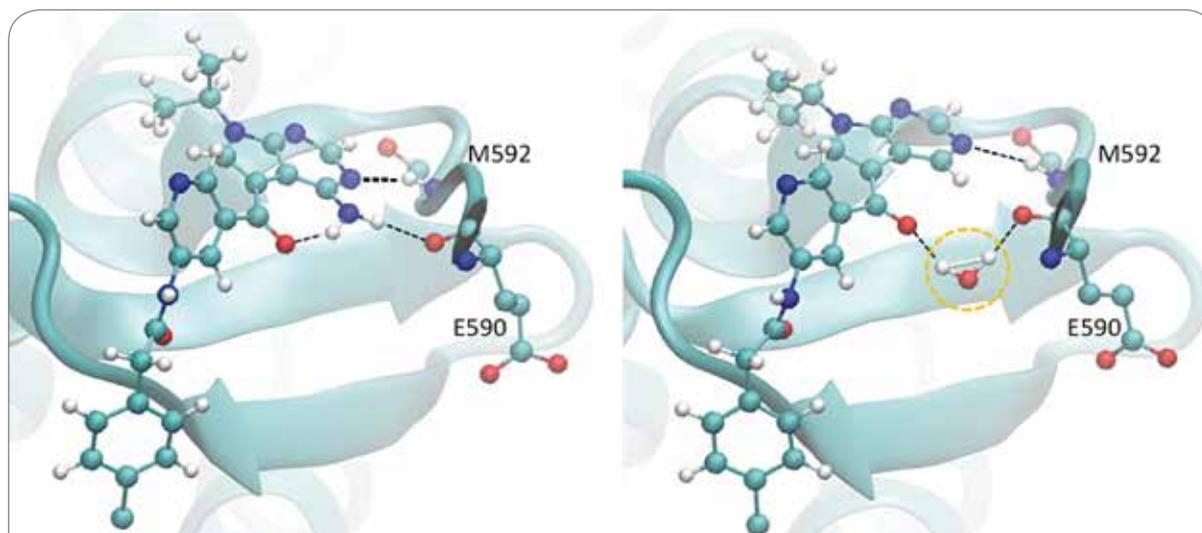


Fig. 3: Suggestions of modification obtained from ESMACS and TIES study. Ligands are bound to a target called TrkA, in which key residues GLU590 and MET592 are highlighted. The deletion of the -NH₂ group has only a minimal effect on the binding affinities, but could potentially lead to an enhanced kinome selectivity profile. See Details in [6].

correlations were obtained from both of the methods. Energetic and dynamic information at the atomistic level are forthcoming from the simulations, which cannot be obtained from experiments. Such information not only explains the experimental observations, but sheds light on how to make modifications in the laboratory to improve the ligand binding and/or ligand selectivity (Fig. 3).

Acknowledgements

We acknowledge the Leibniz Supercomputing Centre for providing access to SuperMUC (<https://www.lrz.de/services/compute/>) and the very able assistance of its scientific support staff.

References

- [1] S. Sadiq, D. Wright, S. Watson, S. Zasada, I. Stoica, P. Coveney. 2008: *J. Chem. Inf. Model.* 48, 1909–19.
- [2] S. Wan, B. Knapp, D. W. Wright, C. M. Deane, P. V. Coveney: *J. Chem. Theory Comput.* 2015, 11, 3346–3356.
- [3] A. Bhati, S. Wan, D. Wright, P. Coveney: *J. Chem. Theory Comput.*, 2017, 13, 210–222
- [4] http://www.gauss-centre.eu/SharedDocs/Pressemitteilungen/GAUSS-CENTRE/EN/2016-03_SuperMUC_Pers_Med.html <https://blog.sciencemuseum.org.uk/supercomputer-bid-to-create-the-first-truly-personalised-medicine/>
- [5] S. Wan, A. Bhati, S. Zasada, I. Wall, D. Green, P. Bamborough, P. Coveney: *J. Chem. Theory Comput.*, 2017, 13, 784–795
- [6] S. Wan, A. Bhati, S. Skerratt, K. Omoto, V. Shanmugasundaram, S. Begal, P. Coveney: *J. Chem. Inf. Model.*, 2017, 57, 897–909

Written by Peter V. Coveney and Shunzhou Wan

University College London, London, UK

Contact: p.v.coveney@ucl.ac.uk, shunzhou.wan@ucl.ac.uk

The Cosmological Web Portal has gone Online at LRZ

Advanced computer simulations nowadays follow the evolution of galaxies and galaxy clusters in unprecedented high precision, producing hundreds of terabytes of complex, scientific data. Exploiting these data-sets is a challenging task. The physics at play, namely the large-scale gravitational instability coupled to complex galaxy formation physics, is highly non-linear and some aspects are still poorly understood. To capture the full complexity, such simulations need to incorporate a variety of physical processes in the calculations, including those that are considered particularly important for the development of the visible universe: first, the condensation of matter into stars, second, their further evolution when the surrounding matter is heated by stellar winds and supernova explosions and enriched with chemical elements, and third, the feedback of super-massive black holes that eject massive amounts of energy into the universe.

To understand the imprint of the dynamical structures and environment of galaxy clusters and groups onto their observables will be essential to interpret the data from current and upcoming astronomical surveys and instruments like PLANCK, South Pole Telescope (SPT), Dark Energy Survey (DES), Euclid, eROSITA, Athena and many more. Direct comparison of true observables from such simulations with observations are essential to constrain the formation history of cosmological structures and to identify the involved physical processes shaping their appearance. It is therefore important to make such simulation data available to a large astrophysical community and to allow scientists

to perform analysis tools via standard interfaces. Therefore, the LRZ now hosts the Cosmological Web Portal for accessing and sharing the output of large, cosmological, hydro-dynamical simulations with a broad scientific community, allowing in its current stage to perform virtual observations of simulations from the Magnetium Pathfinder Project.

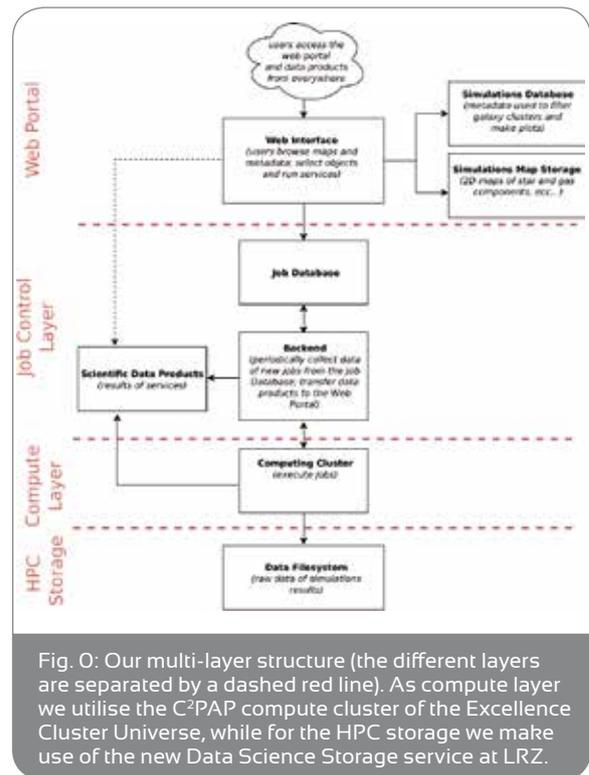


Fig. 0: Our multi-layer structure (the different layers are separated by a dashed red line). As compute layer we utilise the C²PAP compute cluster of the Excellence Cluster Universe, while for the HPC storage we make use of the new Data Science Storage service at LRZ.

The Web Portal

The web portal is based on a multi-layer structure as illustrated in figure-0. Between those layers, data and processes flow over the web portal with its web interface, several databases, the backend within the job control layer, the compute cluster (where the analysis tools are actually



executed) and the storage system (where the raw simulation data are stored). The need for a separation between the web interface and the backend arises from both the necessity of users to run personalized jobs on raw data, managed by a job scheduler of the compute cluster and the protection of the data from unauthorized access. As compute layer, currently the C²PAP compute cluster, operated by the Excellence cluster Origin and Structure of the Universe (www.universe-cluster.de) is used while for the HPC

storage we use the new Data Science Storage service at LRZ. All other processes are virtualised using the LRZ machines. Almost all parts of the implementation is based on common packages and publicly available libraries, except the core of the backend, which is a customized component tailored for the data flows, job requests and specific needs of the used scientific data analysis software. All services are based on standard post-processing tools as used in scientific analysis of such simulations.



Fig. 1: This Snapshot of the visual exploration front end shows the stellar component as basic background (Magneticum/Box2b/hr simulation; visualization based on SPLITCH, see [4]). The layer-spy option allows to dynamically overlap a movable lens with different other visualisation. All objects chosen in the "Restrict" menu are marked with green circles. The blue circle marks the currently chosen cluster, the pop-up info window shows its basic properties. The overlay "pie chart" in the lower right corner shows examples of different visualizations which can be chosen.

Exploring simulations

The visual frontend allows to explore the cosmological structures within the simulation, based on panning through and zooming into high resolution, 256 megapixel size images which are available for numerous outputs of the different simulations.

Generally, two different components can be visualized. Either the density of the stellar component colour coded by the mean age of the stellar population or the diffuse baryonic medium. For the later, either the density (via its X-ray emission), the pressure (via its Compton Y parameter, visible in the millimetre wavelength regime) or the non-thermal structures (turbulence and shocks, visible as fluctuations in the thermal pressure) within the Inter Cluster Medium (ICM) can be chosen. Figure-1 shows an example.

The layer-spy option can be used to show a second, smaller, visualization within a lens which can be moved freely over the whole image. Here, any combinations of visualizations can be used and moving the layer spy over the visible objects reveals immediately and intuitively the connection between the different components. The layer spy also can be set to the same visualization but for the previous or next output in time. This then gives a direct impression of the evolution and dynamics of the objects.

All objects which are selected via the restrict dialogue are marked by a green circle. The user can select an object, which then appears as marked with a blue circle. An additional pop-up

shows then some further information about the cluster.

Select Clusters

Selected 2995 over 10083 clusters.

M_{500c} M_{500} [KeV]
 T_{500c} [erg/s]
 L_x
 L_y
 M_{gas}
 M_{star}
 V_{500c}
 M_{star}/M_{gas}

Definition: All cluster properties are defined within 500 times the critical density

Further restrictions based on sub-structures/member galaxies:

Compact groups Merging clusters

M_{500} M_{500} zero include also galaxies with zero gas mass
 M_{gas} [km/s]
 d_v [km/s]
 v_t [km/s]
 v_s [km/s]
 $dist$ [kpc/h]

Definition: At least one satellite have to fulfill all restrictions of this section

Download CSV table - View table

Close

Fig. 2: The "Restrict" form allows users to select objects with complex properties from the simulation. The upper part allows users to restrict on global properties (like total mass, temperature, etc.) while the lower panel allows to add further constrains on the properties of substructures. By this users can search either for compact groups or merging clusters. Merging clusters as in the shown example, can be defined by the presence of substructure with complex properties, like stellar mass or gas mass content in combination with its velocity (radial, tangential or total) and position within the cluster.

The restrict dialogue as shown in figure-2 allows to select objects according to a list of detailed constrains on global, as well as on internal properties. This is realized via performing complex queries on the meta-data of the galaxy clusters and groups. The user sets up the query interactively by using sliders to define the ranges



of interest for different global properties, like mass or temperature but also by gas and star fraction and even by some dynamical state indicators like centre shift or stellar mass fraction between central galaxy and satellite galaxies (shown in the upper part of figure-2). Furthermore, additional constrains can be placed on the internal structure by either the „Compact Groups“ or the „Merging Cluster“ menu. The first one allows to define, how many galaxies above a given mass limit are to be located within a given distance to the centre. The second allows to restrict to clusters which

have sub-structures with given gas and stellar mass, moving with selected velocities within a defined distance to the centre. Here, the sign of the radial velocity allows to define in-falling or outwards moving substructures. The help function, which is switched on/off by pressing „h“, reveals additional information and further details on any red bordered element as soon as the mouse cursor hovers over it.

The services

There are two classes of services available in the web interface. The first class of services are all designed to complete the interactive exploration of the simulation and to explore and select interesting objects. They are based on the frontend of the portal. The second class of services do involve the backend part and allow the user to obtain information based on unprocessed simulation data.

Exploration services

The CLUSTERFIND service builds on top of the restrict dialogue. It additionally shows the resulting list of objects in form of a table and allows to produce histograms or scatter plots from any combination of result table columns (indexed by the column name). The data points in the scatter plot can be coloured and labelled with table column names. An example result of CLUSTERFIND is shown in figure-3. The produced table can then be exported as CSV-table and individual clusters can be selected by clicking on the table entry or the data points in the plot. This, for example, easily allows to select prominent outliers.

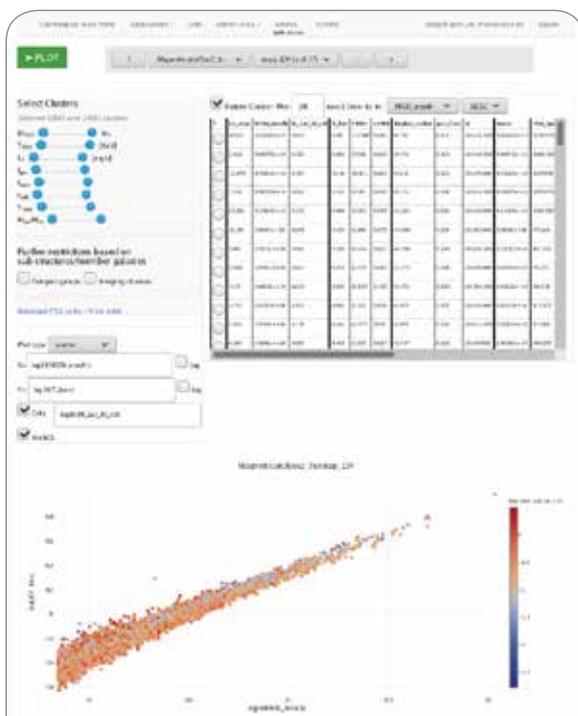


Fig. 3: The Cluster Find service, shows tables listing the clusters of the search result (right upper part). The selection of objects works identical to the “Restrict” form (shown in figure-2). The interactive plot tool allows to produce histograms or scatter plots from the requested data table.

The CLUSTERINSPECT service works similar to the CLUSTERFIND, except that, once a cluster is

selected, the generated table displays the properties of all member galaxies of the cluster. The interactive plotting tool allows then to visualize any galaxy property from the table in the same way as described above.

Post-processing services

The SIMCUT service allows users to directly obtain the unprocessed simulation data for the selected object. These data are returned in the original simulation output format. Therefore, the user may analyse the data in the same way as he would do for his own simulations.

The SMAC service allows the user to obtain 2D maps produced by the map-creation program SMAC. It allows to integrate various physical quantities along the line of sight. Currently, the service allows to produce column densities for the gas component or for the entire matter, the mass-weighted temperature bolometric X-ray surface brightness and the thermal or kinetic Sunyaev-Zel'dovich (SZ) maps. The maps are returned in standard FITS files.

The PHOX service allows to perform synthetic X-ray observations of the ICM and Active Galactic Nuclei (AGN) component of selected galaxy clusters (see [3] for details). Here the user can choose among current and future X-ray instruments to make use of the actual specifications of these instruments. He obtains an idealized list of X-ray photons which in ideal case would be obtained by such instrument. To produce even more realistic results, the user can additionally request a detailed instrument simulation based

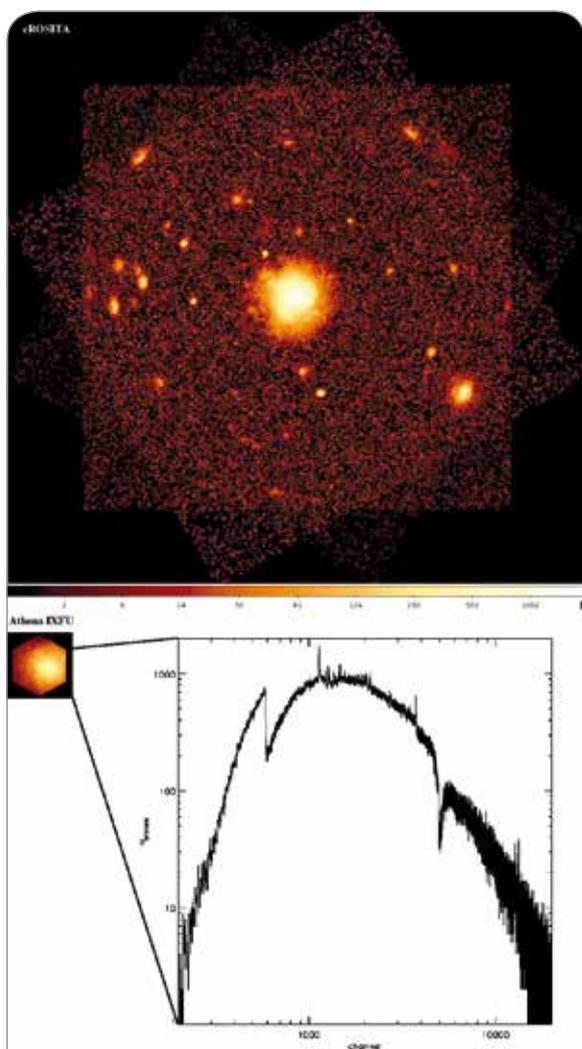


Fig. 4: Results for a virtual X-ray observation of a galaxy cluster using the eROSITA instrument (upper panel). eROSITA has a very large field of view and is supposed to be launched this year. The image is an overlay of the individual eROSITA telescopes (each chip has a different rotation). The lower part show the expectation for the XIFU chip on the future Athena instrument. Athena is expected to have an extremely high spectral resolution, as illustrated by the shown spectrum.



on special software available for the individual X-ray satellite missions. With this, the results are based on full instrument simulation, including energy dependent efficiencies of the detectors and the geometry of the instrument, as can be seen in the examples of figure-4. The result is then obtained in form of a so called event file in the FITS format, which is identical (beside some keywords in the headers) to what would be obtained from a real observation.

Conclusions

The virtual observations obtained via the web portal can be used to explore the theoretical performance of future X-ray satellites like eROSITA or Athena and will help to perform a realistic exploration of the potential of such experiments and thereby shed light on their ability to detect real galaxy clusters and groups across cosmic time. It also allows to easily explore how well global and internal properties can generally be inferred from current and future X-ray missions. Here, thanks to the large, underlying cosmological simulations, users can select systems across a wide range of mass and dynamical states. This allows to test the assumptions of hydrostatic equilibrium and spherical symmetry, typically made when interpreting real X-ray observations.

References

- [1] Ragagnin, A., Dolag, K., Biffi, V., Cadolle Bel, M., Hammer, N.J., Krukau, A., Petkova, M., Steinborn, D., 2016:
A web portal for hydrodynamical, cosmological simulations, *Astronomy and Computing*, 10.1016/j.ascom.2017.05.001 (in publication, see also arXiv:1612.06380).
- [2] Dolag, K., Hansen, F.K., Roncarelli, M., Moscardini, L., 2005:
The imprints of local superclusters on the Sunyaev-Zel'dovich signals and their detectability with Planck. *MNRAS* 363, 29–39.
- [3] Biffi, V., Dolag, K., Böhringer, H., Lemson, G., 2012:
Observing simulated galaxy clusters with PHOX: a novel X-ray photon simulator. *MNRAS* 420, 3545–3556. doi:10.1111/j.1365-2966.2011.20278.x, arXiv:1112.0314.
- [4] Dolag, K., Reinecke, M., Gheller, C., Imboden, S., 2008:
Splotch: visualizing cosmological simulations. *New Journal of Physics* 10, 125006. doi:10.1088/1367-2630/10/12/125006, arXiv:0807.1742.w

Written by Anatonio Ragagnin^{1,2}, Klaus Dolag^{3,4}, Nicolay Hammer¹, Alexey Krukau^{1,2} and the Cosmological Web Portal Team

¹ Leibniz-Rechenzentrum (LRZ), Boltzmannstr. 1, 85748 Garching bei München, Germany

² Excellence Cluster Universe, Boltzmannstr. 2, 85748 Garching bei München, Germany

³ Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany

⁴ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching bei München, Germany

Contact: Nicolay.Hammer@lrz.de

EXAHD: An Exa-Scalable Approach for Higher-Dimensional Problems in Plasma Physics and Beyond

With the construction of ITER, the world's largest plasma fusion reactor expected to begin operation as early as 2025, scientists will be one step closer to proving the feasibility of plasma fusion as an alternative source of clean energy. Numerical simulations are one of the main driving forces behind this enterprise, but new computational techniques are needed in order to achieve the levels of resolution required to gain further insights into the underlying physics.



Fig. 1: The interior of the tokamak (torus-shaped) fusion reactor at the Max Planck Institute for Plasma Physics in Garching, Germany. (Image: IPP; Volker Rohde)

A future powered by plasma?

Germany has long been on the forefront of plasma fusion research, one of the most promising fields in the road to sustainable, carbon-free energy. With virtually no nuclear waste or risk of large-scale accidents and with readily available fuel material (mostly deuterium and tritium), the process of fusion in highly magnetized and extremely hot plasma could represent the future of clean energy for generations to come. In particular, two projects drive the search for the optimal plasma configuration in Germany: the ASDEX Upgrade project in Garching, using a

reactor with a tokamak geometry (Fig. 1) and the recently inaugurated Wendelstein 7-X reactor in Greifswald, which was built using the alternative stellarator geometry (Fig. 2), both commissioned to the Max Planck Institute for Plasma Physics. Since the conception of the two projects, numerical simulations have played a crucial role in trying to better understand both the turbulent properties of the confined plasma and the geometrical configuration that a reactor should have in order to optimize energy production. It is in this context that the project EXAHD, one of 16 projects within the German Priority Programme SPPEXA: Software for Exascale Computing, aims to facilitate the endeavor of plasma fusion research.

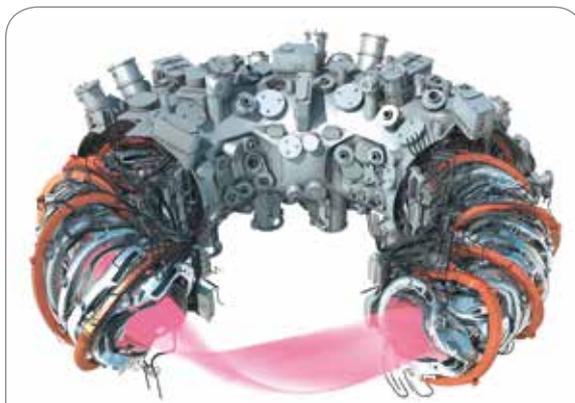


Fig. 2: Schematic of the Wendelstein 7-X fusion reactor in Greifswald, the world's largest reactor with a stellarator geometry. (Image: IPP)

One of the main challenges in trying to extract energy from confined magnetized plasma is the appearance of anomalous heat transport, which leads to a detrimental dissipation of energy. This anomalous transport is usually caused by microturbulence arising from the

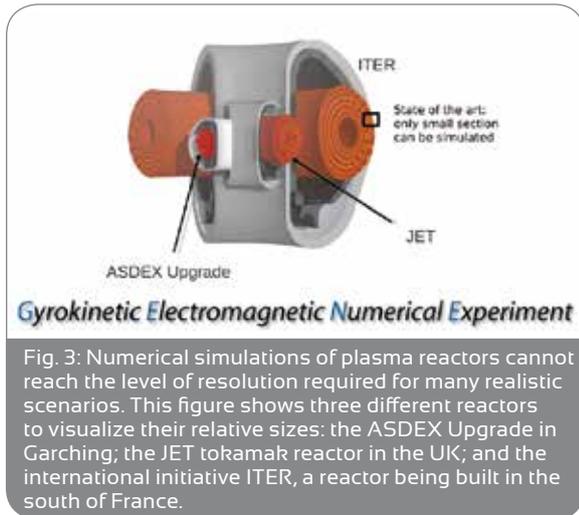
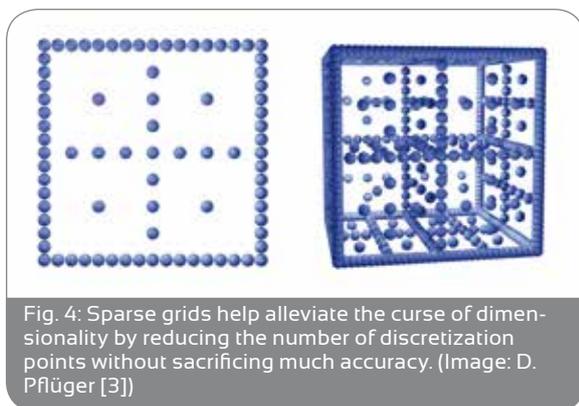


Fig. 3: Numerical simulations of plasma reactors cannot reach the level of resolution required for many realistic scenarios. This figure shows three different reactors to visualize their relative sizes: the ASDEX Upgrade in Garching; the JET tokamak reactor in the UK; and the international initiative ITER, a reactor being built in the south of France.

strong temperature and density gradients in the plasma. Although numerical simulations can help us understand this phenomenon in more detail, the level of resolution required to capture the relevant turbulent effects is large. The code GENE, for example, one of the most efficient plasma microturbulence solvers in the physics community, uses five-dimensional spatial grids, corresponding to the three position coordinates x, y, z and two velocity coordinates v, μ of a plasma particle [1]. For a typical simulation scenario, the 5D grid will have roughly $128 \times 64 \times$



$512 \times 64 \times 16$ grid points – 232 in total – requiring about 2 terabytes of data only to store the function values [2]. Other interesting simulation scenarios require many more grid points (Fig. 3), but they cannot be carried out with the resources of current supercomputers. This slows down progress in fusion research.

The curse of dimensionality

The exponential increase in the number of discretization points as a function of the dimension is not a new problem. In fact, there is a vast number of numerical techniques available to deal with high-dimensional problems. Sparse grids are one such method – indeed, one of the best-established. A sparse grid is a computational grid with considerably fewer points than usual Cartesian full grids. They are constructed by asking which grid points (and thus basis functions in a hierarchical function space) give us the most information about our function a priori, and which ones we can get rid off while maintaining a good numerical accuracy. Examples of sparse grids in 2 and 3 dimensions can be seen in Fig. 4. The reduction in the number of grid points allows one to reach higher resolutions, which is what physicists need. But there is a price to pay: discretizing a problem on a sparse grid is not an easy task – in fact, it is quite cumbersome, and for a complex code like GENE, it is unrealistic. It would mean rewriting the code almost from scratch!

But there is one way to circumvent these difficulties and to obtain the benefits of sparse grids without rediscrctizing the spatial domain.

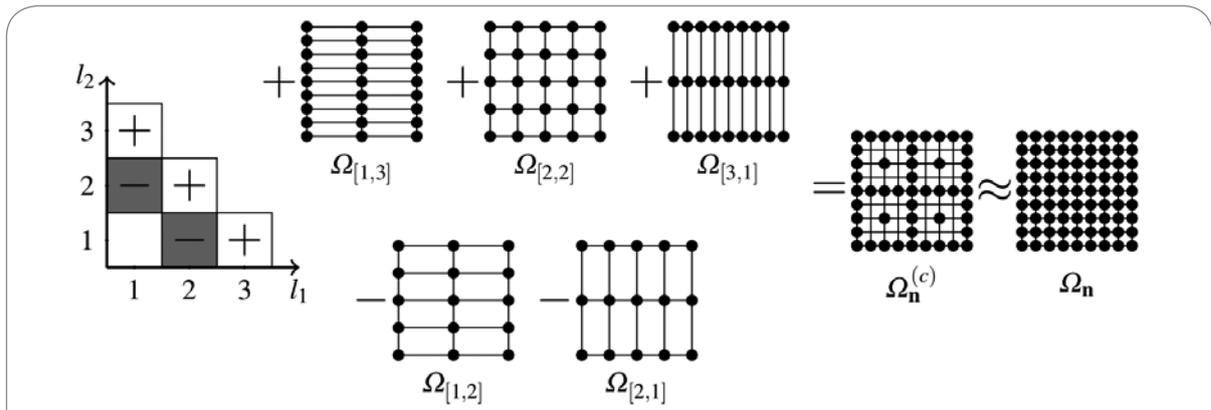


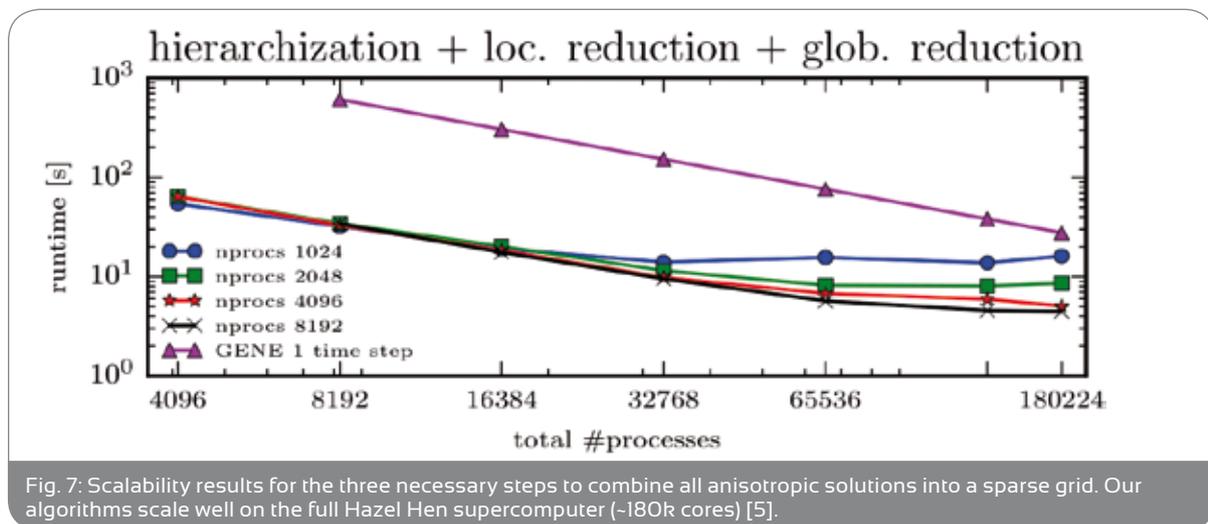
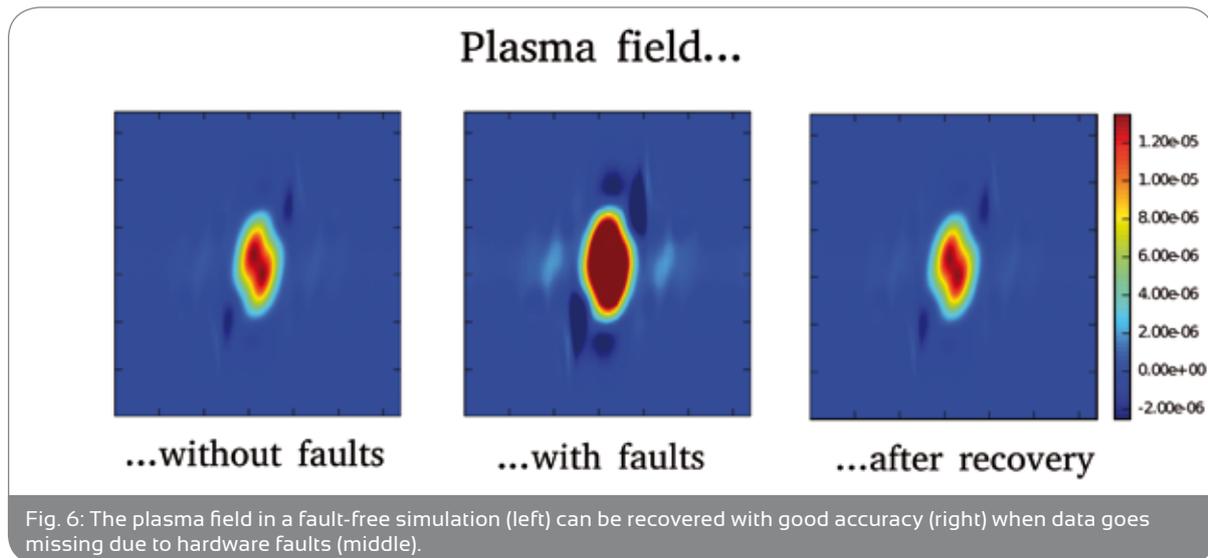
Fig. 5: The Combination Technique builds an approximation of a sparse grid by combining full anisotropic grids (here, five of them in two dimensions, with coefficients +1 or -1).

The idea consists in approximating the solution on a sparse grid by computing the solution on many coarse anisotropic full grids of different resolutions, and to combine the different solutions with certain weights to recover the sparse grid structure. It is called the Combination Technique, and it is illustrated in Fig. 5 for a small 2D example. This is the main idea behind EXAHD: instead of solving a high-dimensional PDE on one single full grid of high resolution (which might be infeasible), we solve the same PDE on these various anisotropic grids, by simply calling our existing code with different discretization resolutions. Afterwards we combine the various solutions together, thus obtaining an approximation of the full grid solution [4]. The Combination Technique can therefore be thought of as an extrapolation method. And this approach has an additional advantage: it offers a second level of parallelism. The different coarse solutions can be computed independently of each other, and they can be combined only at the end, or every certain number of timesteps. The solution on each coarse grid can in turn be solved with the

underlying parallel algorithm. GENE, for example, uses a very efficient domain decomposition in all five spatial dimensions.

New algorithms for new HPC systems

One of the main goals of the EXAHD project is to develop scalable and efficient algorithms to run the Combination Technique on distributed systems. To make the best use of the computational resources, we use a manager-worker scheme, whereby all the available processes are divided into groups, each group usually encompassing plenty of nodes [5]. A master process assigns each group a batch of tasks – a set of anisotropic grids where the PDE has to be solved - using a load balancing scheme [6]. The groups then work independently of each other, solving its set of tasks, and once done, the manager triggers a signal to combine all the solutions. This combination step requires global, but reduced communication, and it can be performed very efficiently and it is not expensive compared to the cost of solving the PDE on the different grids.



The Combination Technique also has several properties that make it tolerant to various types of system faults. For instance, if some of the groups fail to compute the tasks assigned to them due to a hardware defect, it is still possible to find a good combination of the successfully computed solutions, by simply adapting the combination weights (see Fig. 6). This approach, though lossy,

has the advantage of not relying on checkpoint/restart, and scales very well [7]. Similarly, if any of the individual solutions are affected by silent data corruption (usually errors in the floating point data that do not trigger error signals), one can use the information from other solutions to detect wrong results. This can be done using tools from robust regression and outlier detection [8].

The different components of our algorithm have been extensively tested on the supercomputer Hazel Hen – the fastest in Germany (see Fig. 7). The good scalability properties of the Combination Technique make us confident that this approach will allow us to push the resolution barriers in plasma physics and beyond.

EXAHD is a collaborative project between the University of Stuttgart (Pflüger), the Technical University of Munich (Bungartz), the University of Bonn (Griebel), and the Max Planck Institute for Plasma Physics (Dannert), with international collaborations with Prof. Frank Jenko from the UCLA and Prof. Markus Hegland from the ANU.

References

- [1] **Jenko, F., et al.:**
Electron temperature gradient driven turbulence, *Physics of Plasmas*, pp. 1904-1910, AIP Publishing, 2000.
- [2] **Mohr, B., Frings, W.:**
Jülich Blue Gene/P extreme scaling workshop 2009, FZ Jülich Technical Report, <http://juser.fz-juelich.de/record/8924/files/ib-2010-02.ps.gz>, pp. 1-4, FZ Jülich, 2010.
- [3] **Pflüger, D.:**
Spatially adaptive sparse grids for high-dimensional problems, PhD Thesis, p. 14, Verlag Dr. Hut, 2010.
- [4] **Pflüger, D., et al.:**
EXAHD: An exa-scalable two-level sparse grid approach for higher-dimensional problems in plasma physics and beyond, *European Conference on Parallel Processing*, pp. 565-576, Springer International Publishing, 2014.
- [5] **Heene, M., Pflüger, D.:**
Scalable algorithms for the solution of higher-dimensional PDEs, *Software for Exascale Computing-SPPEXA 2013-2015*, pp. 165-186, Springer International Publishing, 2016.
- [6] **Heene, M., Kowitz, C., Pflüger, D.:**
Load balancing for massively parallel computations with the sparse grid combination technique, *PARCO*, pp. 574-583, 2013.
- [7] **Heene, M., Parra Hinojosa, A., Bungartz, H.-J., Pflüger, D.:**
A massively-parallel, fault-tolerant solver for time-dependent PDEs in high dimensions, *Euro-Par 2016*, accepted, 2016.
- [8] **Parra Hinojosa, A., Harding, B., Hegland, M., and Bungartz, H.-J.:**
Handling silent data corruption with the sparse grid combination technique, *Software for Exascale Computing-SPPEXA 2013-2015*, pp. 187-208, Springer International Publishing, 2016.

Written by Alfredo Parra Hinojosa

Technical University of Munich, Department of Informatics, Chair of Scientific Computing

Contact: Alfredo Parra Hinojosa, hinojosa@in.tum.de



Alfredo Parra Hinojosa is a doctoral researcher at the Technical University of Munich, Chair of Scientific Computing. He currently works on fault-tolerant algorithms to solve high-dimensional problems. He is a researcher in the project "EXAHD", which is part of the initiative "SPPEXA (Software for Exascale Computing)" funded by the German Research Foundation (DFG). Mr. Parra Hinojosa's research interests also include low-rank tensor approximations, fractional calculus and random differential equations.







Projects

In this section you will find information about current internal (research) projects of GCS.

MoeWE – Modular Training Programme on High Performance Computing for Professionals

The MoeWE project team, High Performance Computing Center (HLRS), Ulm University, University of Freiburg and SICOS BW, are currently developing a training programme on high performance computing (HPC) tailored to professionals. The concept will transfer practical-oriented knowledge to industry and specialized IT-companies while meeting rising demands in an exciting area of computer science and application.

Although HLRS already offers several HPC trainings to over 1,000 participants per year, they often address research-related topics. Spotting the need in the market, the MoeWE team have set ambitious goals to transfer and utilize current “supercomputing” theoretical-based knowledge for real-world problems.

Rising demand for IT-experts in HPC, simulation and parallel programming

To our knowledge, an in-depth investigation does not exist exploring the exact demands and expert profiles in HPC in Germany as well as in Austria and Switzerland. However, according to regular exchanges with industry executives, expert interviews, and market analysis, there is a high demand for progressive knowledge in IT areas like simulation, modelling and programming. The MoeWE team aims to meet this need and to establish a new field of innovation transfer.

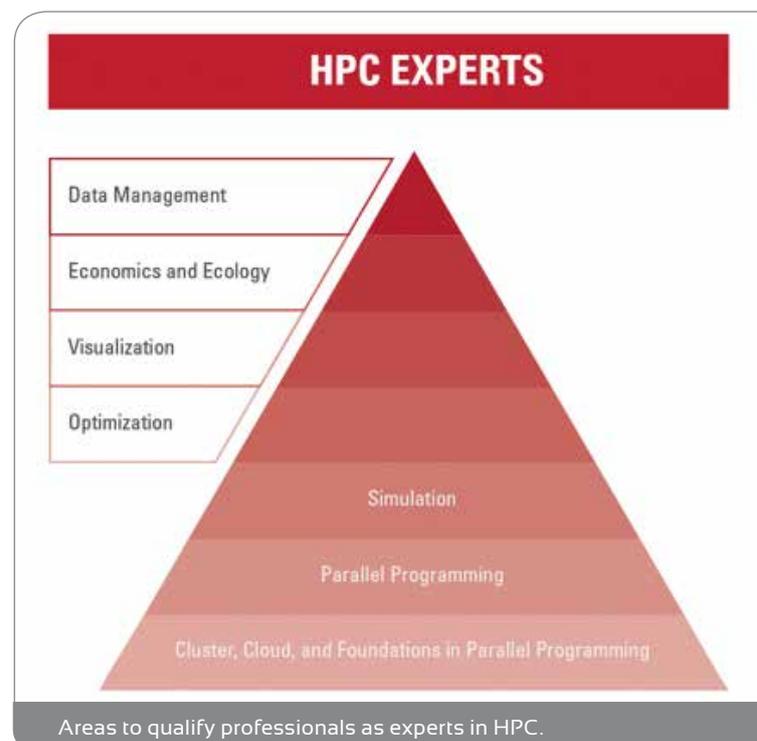
A new modular programme will train IT-professionals to understand and apply parallel and distributed computing as well as set up and operate super computers in their companies.

MoeWE – a project funded by the European Social Fund and the Federal State of Baden-Wuerttemberg

As part of MoeWe, the aforementioned project partners address this topic with a new, innovative training programme.

The programme covers the following areas to qualify professionals as experts in HPC:

- Introduction to IT
- Parallel programming (IT-languages, tools, libraries and scheduling)
- Simulation (visualization, optimization, techniques and models of simulation)
- Cluster, cloud and high-performance-computing





- Business administration, ecology and economy
- Data management

The first courses will take place in the second half of 2018, and will be free-of-charge for all participants.

Training approach: online and face-to-face

MoeWE offers modular and flexible courses to become a HPC expert so that professionals are provided development opportunities and can meet long-term needs beside their work and personal life.

The training approach is based on a blended learning concept which combines E-learning and face-to-face learning. This approach takes advantage of both learning forms. Therefore, participants are flexible to study when and where it best suits their work and personal schedules.

An academic qualification is not required. Participants can enroll in selected modules or the entire program based on their background and needs.

Practitioners from industry can take advantage of training programme opportunities

The target audience is professionals and executives from small and medium-sized businesses. Additionally, our training programme is attractive for specialized IT companies interested in HPC expertise.

Our “blended learning” approach is particularly suitable for practitioners due to mix of face-to-face sessions as well as online learning and hands-on exercises with HPC software and hardware. Moreover, the online learning part will consist of video conferences, screen demonstrations of software and computing, online-scripts and online-videos. Because of this, participants can tailor the training program to their personal situation in a flexible and self-defined pace.

First module’s outline: “Introduction to High Performance Computing”

In 2018, the first module will be conducted as a fundamental orientation of the training programme. The module will have a 100 hours learning investment. The group of HPC-expertise trainers will chair this module over a period of about 10 weeks. The topics to be covered will be an introduction to approaches and the structure of fast computers, and the principal thinking in parallel programming. The initial group of participants will aid in achieving valuable learning experiences and exchange of ideas.

The module will begin with an initial face-to-face session, serving as the module orientation (kick-off) as well as closing face-to-face session at the end of the module. Online video conferences will take place every three weeks during evening hours. Online self-learning periods fill in the rest of the learning experience. In the self-learning periods, professionals are able to learn at their own pace and can tailor the workload to best fit their needs.

Future plans

The MoeWE project team is currently analysing the market, developing the didactic concept and selecting the content of the first module, which meets industry needs.

Advisory board and expert meetings will be established to guarantee the training programme continually meets future industry demands. A full set of 7 to 9 modules should be

established by the end of 2019. The maximum number of participants has yet to be determined, but the MoeWE team expects 15 to 20 professionals trained per module.

Project details

Funding Agency: The European Social Fund (ESF) and the Ministry of Science, Research and the Arts of the state of Baden-Württemberg.
Runtime: 1.7.2016 – 31.12.2020



GEFÖRDERT VOM MINISTERIUM FÜR WIRTSCHAFT,
ARBEIT UND WOHNUNGSBAU BADEN-WÜRTTEMBERG
AUS MITTELN DES EUROPÄISCHEN SOZIALFONDS



Further details

Interested in joining the MoeWE training program? Please send an e-mail (contact details below) and you will be provided with further information.

Department

Communication & Industrial Trainings



Contact

Jutta Oexle, oexle@hirs.de

Link

<https://www.hirs.de/moewe>

Written by Ludger Benighaus¹, Jutta Oexle¹, Hanna Skubski¹, Christopher A. Williams²

¹High Performance Computing Center Stuttgart (HLRS)

²Ulm University

Contact: Jutta Oexle, oexle@hirs.de



Blended Learning: mix of face-to-face sessions as well as online learning

Helmholtz Data Federation

Large-scale experiments and simulations in science generate an increasing amount of data. The transformation of data and information to findings and knowledge, however, also needs a new quality of storage and analysis capability. The Helmholtz Association now takes an architectural role in the permanent, secure, and usable storage of data. For managing big data in science, it has established the Helmholtz Data Federation (HDF). Within the next five years, about EUR 49.5 million will be invested into multi-disciplinary data centers and modern data management. The HDF will establish a data federation comprising three elements: innovative software technologies, excellent user support and leading-edge storage and analysis hardware.

The HDF as a national research data infrastructure constitutes the long-term federation of powerful, multi-disciplinary data centers. Combining these federated data storages with the existing expertise and knowledge of the six partners in research data management and user support provides a unique research infrastructure, which will promote and foster the transformation of data into knowledge and thereby support excellent science in Germany and beyond.

The federation is built on efficient software methods and tools of distributed data management and secure network links among each other and within Helmholtz, to university partners and further research organizations in Germany and internationally via DFN. The HDF represents the nucleus of a national research data infrastructure across science organizations,

which is open to users in the whole German science community. International connections will make it compatible with the future European Open Science Cloud (EOSC).

The data centers at Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Deutsches Elektronen-Synchrotron (DESY), GSI Helmholtz Centre of Heavy Ion Research, German Cancer Research Centre, Forschungszentrum Jülich, and Karlsruhe Institute of Technology with strong topical profiles (Figure 1) are enhanced with leading-edge storage and analysis resources and technologies. This will ensure that the ever increasing volume of valuable research data in various scientific disciplines is stored and archived, long term access is guaranteed, data ownership is preserved and new perspectives can arise for intra- and interdisciplinary transformation of data into knowledge with high relevance for science, industry and society.

Federating resources and knowledge is a common principle in science and well adopted in various research domains. The guiding principle behind the HDF is the open federation of leading-edge storage and analysis hardware through innovative software as well as excellent user support for the preservation of data and metadata itself, their integrity, provenance, moral and legal ownership as well as their original access rights. Initially this federation is build up across several Helmholtz Centers connecting to a majority of science disciplines in Helmholtz. The HDF can be used by scientists from



Fig. 1: Centres and Use Cases.

Helmholtz, universities and other research organizations and institutes across Germany ultimately leading to a nation-wide, federated infrastructure for research data of the entire German science system. Conceptually additional data and computing centers from Helmholtz, universities and other research organizations in Germany can be added in an efficient, secure and transparent way. The federation is established through structural elements of innovative data management software, security and identity mechanisms, broadband network connections as well as a competence network of human experts. Besides the federation and sharing of research data, the federation and sharing of knowledge, expertise and software among the HDF partners and with the scientific communities using the HDF will provide unprecedented advancements beyond state of the art. The federated approach of the HDF will foster existing and new scientific collaborations inside and across scientific domains and communities. New collaborations will arise between biology,

life science and photon science for the understanding of biological structures and processes or between energy, climate and marine research to optimize energy systems of the future based on renewable energies. The HDF will allow for mutual use of data by cross-linking and annotating. The development and deployment of methods to enhance sharing and re-use by applying common standards will bring science one step closer to universal data access, where researchers from different disciplines will have the chance not only to search but also to find answers from data collected in other scientific domains. For example, information from metagenomics (an approach to reveal the full diversity of life in each specific sampling location) can be combined with environmental parameters (physical, chemical, and other biological) in these locations. This will enable a much more detailed understanding of ecosystems, which in turn will help to predict changes in biological productivity (e.g. fisheries, agriculture) under conditions of climate change.

Written by Daniel Mallmann

Jülich Supercomputing Centre (JSC), Germany

and Prof. Dr. Achim Streit

Karlsruhe Institute of Technology, Steinbuch Centre for Computing

Contact: Daniel Mallmann, d.mallmann@fz-juelich.de

Prof. Dr. Achim Streit, achim.streit@kit.edu

CATALYST: The Next Generation of Big Data Analytics at HLRS



We are living in the data era. With the rise of technologies, including sensor networks and cyber-physical systems, we witness not just the growing volume of data, but also in the increase in speed and requirement to analyze data effectively in near real-time. Here, a new research field called data-intensive science arose in order to tackle today's and future challenges of Big Data. At the High Performance Computing Center Stuttgart (HLRS), customers tend to execute more data-intensive applications, processing and producing more data than ever

before. With today's computing power, automotive companies tend to execute several hundred crash simulations; up to 1 Petabyte of resulting data can thus be generated within a single day. Since it no longer becomes feasible that data is processed and analyzed manually by domain experts, HLRS and Cray have launched the CATALYST project to advance the field of data-intensive computing by converging HPC and Big Data to allow a seamless workflow between compute-intensive simulations and data-intensive analytics.



Fig. 1: Prof. Dr. Michael M. Resch (HLRS) and Nurcan Rasig (Cray) at the handover of the Cray Urika-GX. Photo: Boris Lehner.



Urika-GX

Hazel Hen, the current HPC flag-ship system of HLRS, is extended with specific data analytics hardware designed by Cray—a Urika-GX (cf. Figure 1). In the first phase of the project, the Urika-GX system is operated in two individual configurations: a larger configuration (48 nodes) for production, and a smaller configuration (16 nodes) for development and testing. The new data analytics system supports innovative Big Data technologies such as Hadoop and Spark, both of which boost data analytics in engineering. Moreover, the special-purpose Cray Graph Engine enhances the analysis of semantic data, which is commonly present in biology and chemistry.

Project goals

CATALYST investigates the hardware of the Urika-GX and its usefulness with a particular focus on applications from the engineering domain. Since the majority of today's data analytics algorithms are oriented towards text processing (e.g., business analytics) and graph analysis (e.g., social network studies), we are further in need to evaluate existing algorithms with respect to their applicability for engineering. Thus, CATALYST will examine future concepts for both hardware and software. CATALYST will pursue multiple case studies from divergent domains throughout the next years. In order to support end users best, CATALYST foresees to incorporate the Urika-GX with the existing HPC system at HLRS. This ambitious goal requires to tackle various operative challenges including security aspects, fast and

secure data transfer, and accounting, to name but a few [1].

Case study

Our first case study was conducted in collaboration with Cray. In the past, we have randomly observed performance variations of our Cray XC40 system, which is composed of more than 7,000 compute nodes. Performance variability on HPC platforms is a critical issue with serious implications on the users: irregular runtimes prevent users from correctly assessing performance and from efficiently planning allocated machine time. The hundreds of applications, which are sharing thousands of resources concurrently, escalate the complexity of identifying the causes of runtime variations. Thus, monitoring today's IT infrastructures has actually become a big data challenge on its own.

Novel analytics tools—like the ones installed on the Urika-GX—enable exploring new ways to use data for identifying and understanding performance variability. In this context, we have developed a Spark-based tool for analyzing system logs with the goal of identifying applications that show high variability (victims), and applications potentially causing the variability (aggressors). Understanding the nature of both types of applications is crucial to developing a solution to these issues (cf. Figure 2). The analysis was conducted over a span of two weeks. We identified 472 victims, and 2,892 potential aggressors. Seven of those potential aggressors were running on more than 1,000 nodes and three of them were found repeatedly. On a larger dataset



(data aggregated over three months), the same analysis found 3,215 victims, 67,908 aggressors, and 17 of them using more than 1,000 compute nodes. We will use the new insights to further improve our operations.

Outlook

CATALYST will be evaluated in cooperation with partners from both industry and academia new possible applications of Big Data. Furthermore, we will continue to advance the integration of the Big Data system into our existing HPC infrastructure to guarantee a seamless workflow between simulations and analytics.

Acknowledgement

The project is funded by the State of Baden-Württemberg, Ministry of Science, Research and the Arts Baden-Württemberg. Cray Inc. is partner in the project. Daimler AG is an associative partner.

References

[1] D. Hoppe, M. Gienger, T. Bönisch, O. Shcherbakov and D. Moise:

“Towards Seamless Integration of Data Analytics into Existing HPC Infrastructures,” in Cray User Group, Redmond, WA, USA, 2017.

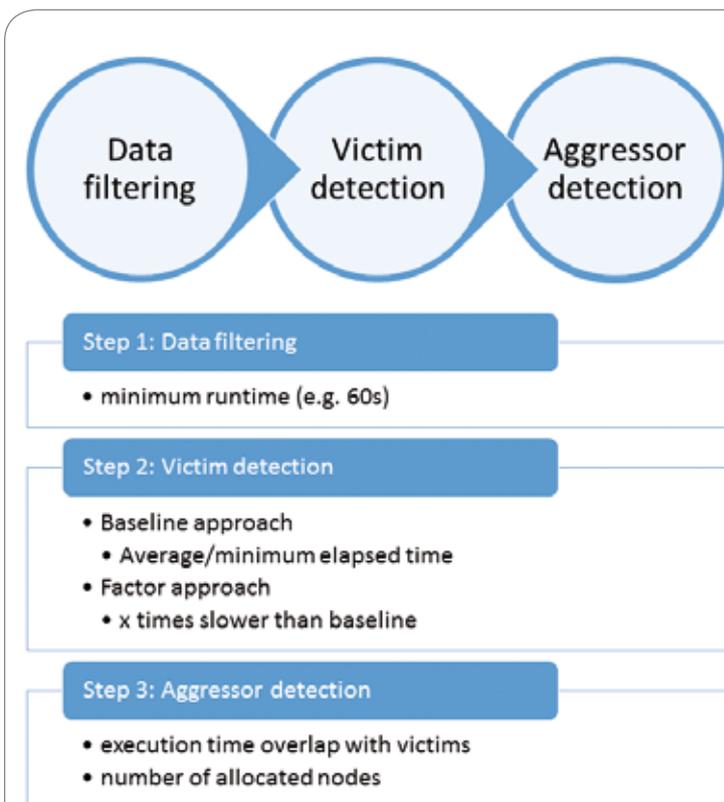


Fig. 2: Three steps of the analysis workflow (data filtering, victim detection, and aggressor detection) including configuration parameters.

Written by Thomas Bönisch¹, Michael Gienger¹, Dennis Hoppe¹, Bastian Koller¹, Oleksandr Shcherbakov¹, Diana Moise²

¹High Performance Computing Center Stuttgart (HLRS)

²Cray, Inc.

Contact: Michael Gienger, gienger@hlrs.de

Next Generation of QPACE Operational

The “QCD Parallel Computing Engine” (QPACE) project started in 2007 by the universities of Regensburg and Wuppertal in the framework of the transregional Collaborative Research Centers SFB/TRR 55. Meanwhile several generations of QPACE supercomputers have been put into operation. Most recently, QPACE 3 has been installed at Jülich Supercomputing Centre.

The goal of the QPACE project is on the one hand the development of novel application optimized supercomputer architectures and on the other hand the creation of competitive research infrastructure for simulating the theory for strong interactions, namely quantum chromodynamics, on a lattice. The strategy of the project is to integrate extremely fast processors in a particularly dense way. Unlike previous generations, QPACE 3 was not the result of a joint development project but of an open tendering process, which resulted in a contract awarded to Fujitsu.

QPACE 1 [QPACE1] and QPACE 2 [QPACE2] were based on the IBM PowerXCell 8i and the first generation of Intel Xeon Phi, respectively. For QPACE 3 the second generation Xeon Phi is used (codename Knights Landing). While the selected Xeon Phi 7210 with its 64 cores is not the fastest available processor of its kind, it was expected to be the most power efficient choice. All generations of QPACE had been designed for power efficiency and were ranked top in the Green500 list [green500]. Thanks to strong support from Fujitsu in Augsburg (Germany) it had been possible to have QPACE 3 listed at rank #5 on the November 2016 list.

All these generations do also have in common the use of direct liquid cooling. Initially, this approach was selected to minimize costs by maximizing density. Meanwhile other benefits of liquid cooling have become more important. As direct liquid cooling allows to significantly increase temperature of the liquid that leaves the data centre, full year free cooling becomes an option, which again helps to reduce power consumption. The Fujitsu CS600 servers use technology from the Danish company Asetek to enable outlet liquid temperatures beyond 40°C and mounting of servers both from front and back of the racks. All 352 compute nodes, i.e. almost 1 PFlop/s compute performance, could thus be accommodated in just 4 racks only.

Until summer 2017, the size of the QPACE 3 installation will be significantly increased. In parallel, work on the next generation has started. QPACE 4 is being developed with Cray and will again exploit highly advanced processor technology, but now based on the ARM architecture.

References

[QPACE1] G. Goldrian et al.:

“QPACE: Quantum Chromodynamics Parallel Computing on the Cell Broadband Engine”. In: *Computing in Science Engineering* 10.6 (Nov. 2008), pp. 46–54. issn: 1521-9615. doi: 10.1109/MCSE.2008.153.

[QPACE2] P. Arts et al.:

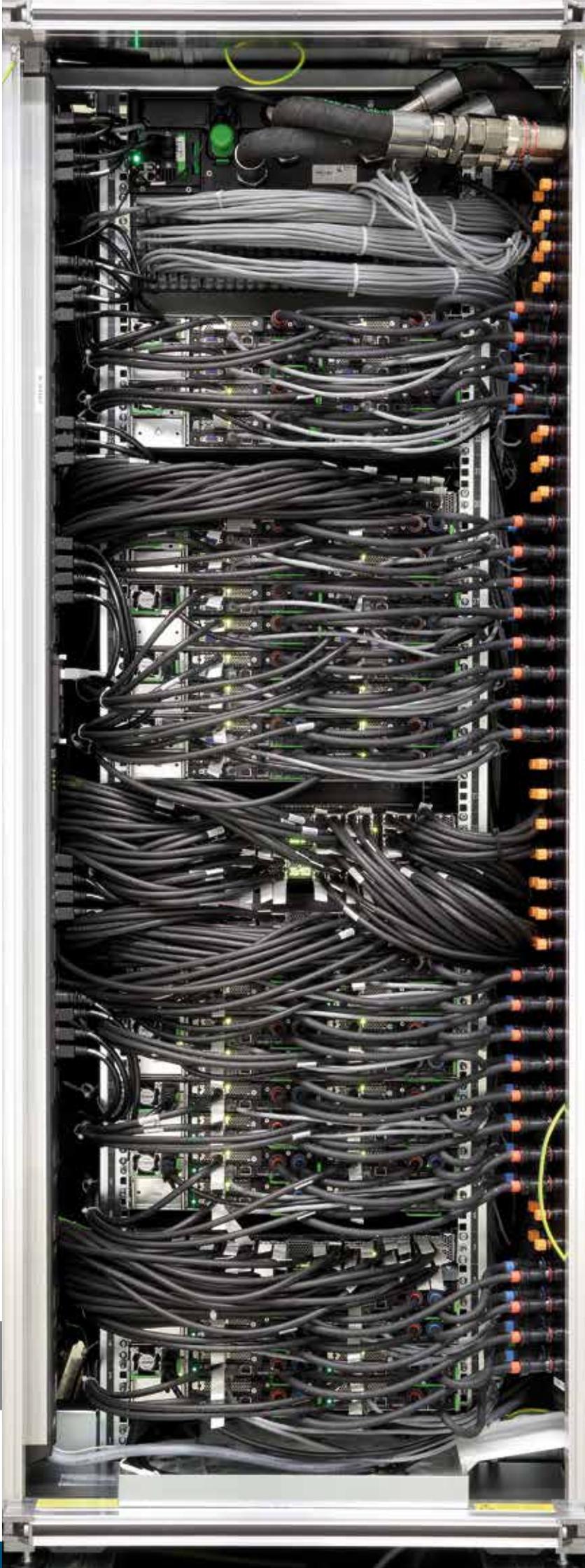
“QPACE 2 and Domain Decomposition on the Intel Xeon Phi”. In: *PoS LAT2014* (2014), p. 001. arXiv:1502.04025 [hep-lat].

[green500]: <https://www.top500.org/green500/>

Written by Dirk Pleiter

Jülich Supercomputing Centre (JSC), Germany

Contact: Dirk Pleiter, d.pleiter@fz-juelich.de



Single QPACE 3 rack with a water manifold at the right side, which is connected to a heat exchanger in the top of the rack.

EXCESS Project Presents Energy-Aware Software Stack

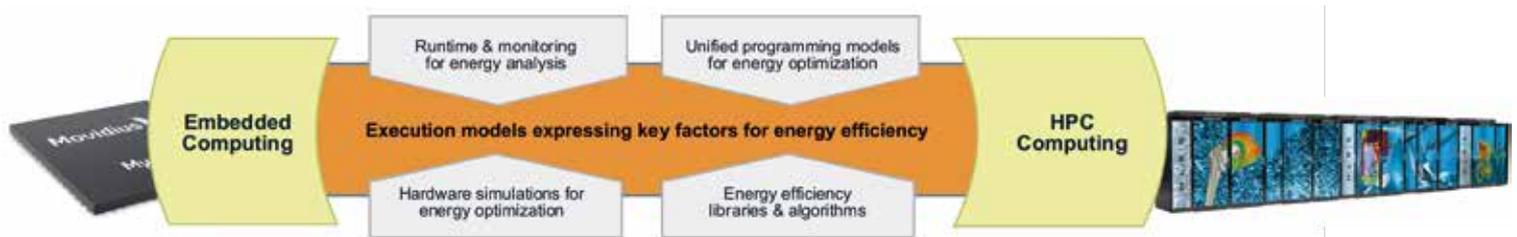


Fig. 1: Illustration of the iterative hardware/software co-design process implemented in EXCESS.

Reducing energy is a leading design constraint of current and future HPC systems. Aside from investing into energy-efficient hardware, optimizing applications is also key to substantially reduce the energy consumption of HPC clusters. The EXCESS (Execution Models for Energy-Efficient Computing Systems) project started three years ago with the fundamental idea of implementing a holistic approach to realize energy-saving solutions for both high performance and embedded systems (cf. Figure 1). The consortium was composed of partners bringing in HPC expertise (HLRS), embedded systems (Movidius), and energy efficient computing (LIU, UiT). The project was led by CHALMERS; it finished successfully in August 2016.

EXCESS' Energy-Aware Software Stack

The project demonstrates newly developed energy-aware solutions on a specially established testbed. The testbed at HLRS integrates both high performance as well as embedded systems including x86 multicore CPU-based servers (plus GPUs) and Movidius Myriad2 boards for ultra-low power processing; the

testbed is complemented by an external power measurement system. We highlight some key components developed in EXCESS to allow for energy-aware development and application execution.

ATOM monitoring framework

Software developers are usually in the dark when it comes to quantifying energy consumption of their applications. HPC clusters rarely provide capabilities to monitor energy consumption at a fine-granular level. EXCESS developed ATOM—a light-weight near-real time monitoring framework—to lower the hurdle of energy-aware development [1]. ATOM enables users to monitor applications at run-time with ease. In contrast to existing frameworks, ATOM profiles applications at high resolution, focuses on energy measurements, and supports a heterogeneous infrastructure. Further, ATOM allows software developers to optimize their applications at run-time by requesting instant monitoring feedback. ATOM is already widely used within the project, and across multiple European projects such as DreamCloud and PHANTOM.

EXCESS implemented “EXCESS HPC Services” (cf. Figure 2) as an extension to the standard HPC resource manager PBS Torque to compute individual tasks of an HPC application in a distributed manner on both HPC and embedded resources. “EXCESS HPC Services” is composed of the following main components: a single producer (the leg implant simulation generates new tasks), multiple worker (tasks are processed either by HPC resources or integrated Myriad2 boards), and a task queue (responsible for assigning tasks to workers). Figure depicts the overall architecture. EXCESS could demonstrate the successful porting of an HPC-based application to such a distributed infrastructure, which allows for saving energy when tasks are delegated to embedded devices.

Acknowledgements

EXCESS is funded by the EU 7th Framework Programme (FP7/2013-2016) under grant agreement number 611183. Visit www.excess-project.eu for more information about EXCESS. The entire EXCESS software stack is available as open source at www.github.com/excess-project.

References

- [1] **D. Hoppe, Y. Sandoval and M. Gienger:**
„ATOM: A Near-Real Time Monitoring Framework for HPC and Embedded Systems,” in PODC, San Sébastian, 2015.
- [2] **F. Pi, D. Hoppe, M. Gienger and D. Khabi:**
„Energy-aware Scheduling with StarPU,” in EXCESS Workshop, Gothenburg, 2016.
- [3] **R. Schneider:**
„Identification of anisotropic elastic material properties by direct mechanical simulations: estimation of process chain resource requirements,” High Performance Computing on Vector Systems, 2010.

Written by Bastian Koller, Uwe Küster, Dmitry Khabi, Dennis Hoppe, Fangli Pi and Michael Gienger

HLRS Höchstleistungsrechenzentrum Stuttgart

Contact: Bastian Koller, koller@hlrs.de
Uwe Küster, kuester@hlrs.de
Dmitry Khabi, khabi@hlrs.de
Dennis Hoppe, hoppe@hlrs.de
Fangli Pi, pi@hlrs.de
Michael Gienger, gienger@hlrs.de

Bringing eScience to the Cloud with PaaSage

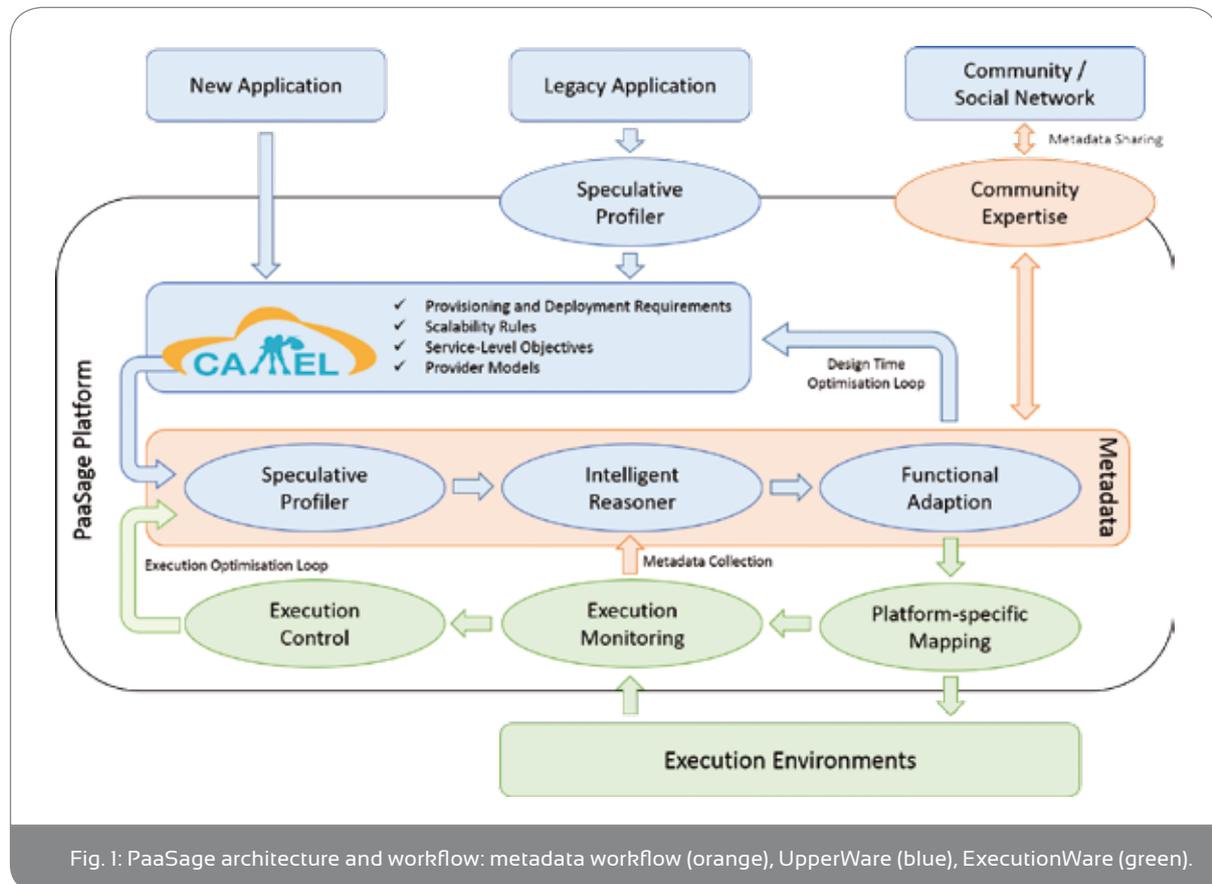


Migrating applications to multiple cloud platforms is a challenging task. PaaSage—a model-based cross-cloud deployment platform—has proven, however, that the migration process is feasible. Migration is convenient enough to be performed by end users having limited knowledge. PaaSage overcomes current migration challenges by avoiding a vendor lock-in, accounting for the heterogeneity of cloud platforms, and abstracting non-standardized APIs and architectures. The project ended after four years in September 2016 with an excellent notice awarded by the European Commission.

The consortium of 18 partners was composed of modelling experts, infrastructure providers, software specialists, and seven use case providers coming from both academia and industry.

Architecture

How could PaaSage implement a holistic framework for cross-cloud deployment? PaaSage’s architecture is geared towards its “develop once, deploy everywhere” paradigm. Users are encouraged to represent (existing) applications using a newly developed cloud modelling-language named CAMEL—Cloud Application



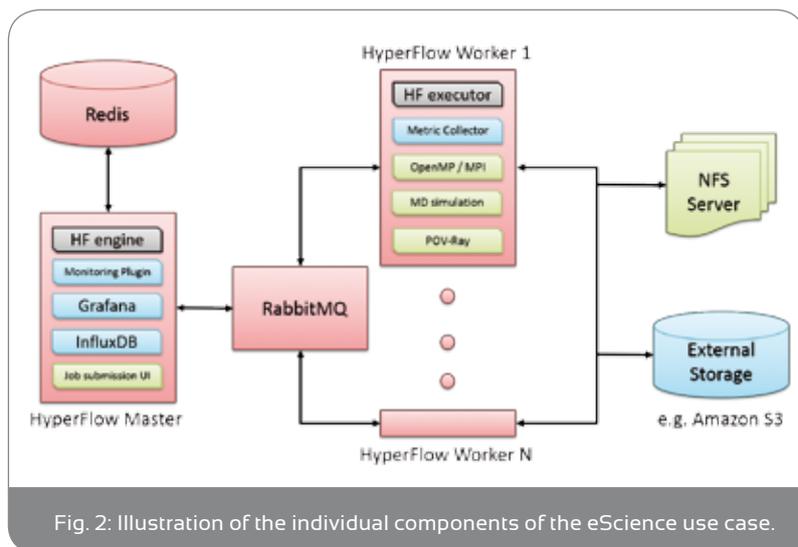


Fig. 2: Illustration of the individual components of the eScience use case.

the ExecutionWare is to monitor, re-configure, and optimize running applications. Monitoring data is passed continuously to the UpperWare. If the UpperWare should find a better deployment solution at run-time, a re-deployment can automatically be triggered. Figure 1 illustrates the architecture.

Seven success stories

A use case, fueled by a cooperation of HLRS and AGH,

Modeling and Execution Language [1]. PaaSage supports new users with comprehensive documentation and training materials. PaaSage also provides an Eclipse-based editor for creating application models. CAMEL models include not only required components of an application, but also various user requirements such as 1) a set of preferred cloud providers for deployment, 2) hardware requirements, 3) auto-scaling options, and 4) optimization criteria (e.g., response time below a given threshold). The CAMEL model is then passed to the so-called UpperWare, where application and user requirements are mapped against a metadata database to identify potential cloud providers that satisfy all requirements. The UpperWare returns an initial feasible deployment solution, which is passed to the next component—ExecutionWare. The ExecutionWare provides a unified interface to multiple cloud providers, and thus can handle platform-specific mappings and different cloud provider architectures and APIs. The purpose of

achieved to bring HPC-based applications into the cloud [2]. HPC is usually first choice when it comes to executing eScience applications. However, the trend is towards a hybrid HPC/cloud model, where high performance resources are combined with the advantages of the cloud: flexibility, high availability, and disaster recovery to name but a few. Applications that benefit from such a hybrid model are, in particular, eScience applications. eScience applications usually include pre- and post-processing steps, which are not compute-intensive. These tasks can be moved to the cloud, whereas compute-intensive tasks continue to run on HPC infrastructure. We demonstrated in PaaSage that we could model a representative molecular dynamics (MD) simulation workflow with CAMEL once, and then deploy individual tasks—managed by PaaSage and HyperFlow—on different cloud infrastructures. What is HyperFlow? HyperFlow is a workflow engine that enables users to execute scientific workflows on available resources [3].



PaaSage, on the other hand, then provisions these resources (e.g., virtual machines) and deploy the actual application. The solution did not touch the existing source code of the MD simulation, and required to set up a MPI cluster in the cloud. Figure 2 depicts the individual components of the model and their communication.

Additional case studies demonstrated that PaaSage can serve many-faceted requirements across different domains: flight scheduling (LSY), financial applications (IBSCY), scientific data farming experiments (AGH), CAE applications (asc(s), after sales (be.wan), and the public sector (EVRY).

Impact

PaaSage impressively demonstrated that its platform can overcome today's challenges when migrating to the cloud: no vendor lock-in, unified API to deploy on multiple cloud infrastructures, automatic scaling features, and optimization goals. The extensive set of demonstrators proved the wide applicability of PaaSage to diverse use cases. PaaSage is available as open source, and it will be further developed in a European project called Melodic (Multi-cloud Execution-ware for Large-scale Optimised Data-Intensive Computing).

Acknowledgements

PaaSage is funded by the EU 7th Framework Programme (FP7/2013-2016) under grant agreement number 317715. Visit www.paasage.eu for more information about PaaSage. CAMEL models for use cases are available in PaaSage's social network: socialnetwork.paasage.eu.

References

- [1] **A. Rossini:**
„Cloud Application Modelling and Execution Language (CAMEL) and the PaaSage Workflow,” in 4th European Conference on Service-Oriented and Cloud Computing (ESOCC ,15), Taormina, Italy, 2015.
- [2] **M. Malawski, B. Balis, K. Figiela, M. Pawlik, M. Bubak, D. Krol, R. Slota, M. Orzechowski, J. Kitowski and D. Hoppe:**
„Molecular Dynamics with HyperFlow and Scalarm on the PaaSage Platform,” in 5th European Conference on Service-Oriented and Cloud Computing (ESOCC ,16), Vienna, Austria, 2016.
- [3] **B. Balis, K. Figiela, M. Malawski, M. Pawlik and M. Bubak:**
„A Lightweight Approach for Deployment of Scientific Workflows in Cloud Infrastructures,” in International Conference on Parallel Processing and Applied Mathematics (PPAM ,15), Lublin, Poland, 2015.

Written by Bastian Koller, Michael Gienger and Dennis Hoppe

Höchstleistungsrechenzentrum Stuttgart

Contact: Bastian Koller, koller@hls.de
Michael Gienger, gienger@hls.de
Dennis Hoppe, hoppe@hls.de

The Mont-Blanc Project: Second Phase successfully finished

Until recently, the design and implementation of the majority of the HPC systems have not primarily focused on energy efficiency. However, it is unanimously accepted within the scientific and industrial community that, as system computing capabilities approach the exascale, future HPC centres will severely being constrained by their massive power consumption. The aim of the EU-funded Mont-Blanc project has been to contribute in addressing these significant challenges by designing a new type of computer architecture capable of setting future, global HPC standards in terms of energy efficiency.

During the first phase of the project (October 2011 to June 2015), the focus was on the development and evaluation of an HPC prototype using energy-efficient embedded ARM technology available at the time, as well as porting a set of representative applications to this system. The now successfully completed second phase (October 2013 to January 2017) concentrated on an initial design of the Mont-Blanc Exascale architecture by exploring different alternatives for the compute node as well as advancing the system software stack. In particular, it enabled further development of the OmpSs parallel programming model to automatically exploit multiple cluster nodes, transparent application check-pointing and software-based fault tolerance, and added support for ARMv8 64-bit processors. In addition, the project allowed the development of parallel programming tools for OmpSs and for the adaptation of system software such as job schedulers to the ARM ecosystem. Due to the shifted focus, several new

partners joined the Mont-Blanc consortium for the second phase of the project, in particular Allinea, Inria, University of Bristol, and HLRS from the University of Stuttgart. Like the first phase, the second phase of the project was also coordinated by the Barcelona Supercomputing Center (BSC).

An important aspect of Mont-Blanc is hardware prototyping, in particular for the evaluation of the various software components developed in the project. In this regard, the consortium has designed and built a prototype based on 1080 ARMv7 nodes that has been installed at BSC. The interested reader can find more information on the architectural details of the Mont-Blanc system in the InSiDE magazine issue of Autumn 2015 or at the website <http://www.montblanc-project.eu>. Additionally, in the second phase of the project, researchers from the consortium focused their efforts on the evaluation of small ARMv8-based cluster systems, in order to assess their performance/energy trade-off and to keep track of the evolution of ARM platforms in the HPC domain (with special focus on ARM 64-bit solutions). Examples of these systems include but are not limited to a 16-node Nvidia Jetson TX1 cluster, a 3-node Applied Micro X-Gene 2 cluster and a 4-node Cavium ThunderX cluster. More details on the hardware configuration and system software of these small clusters is available at the project main website.

All three Gauss centres, i.e., HLRS, JSC, and LRZ, have been involved in the second phase of



the Mont-Blanc project. Their overall contributions, detailed in the following sections, mainly focused on:

- Energy-aware scheduling algorithms (LRZ),
- Performance and debugging tools (JSC and HLRS),
- Porting of a scientific application for evaluating the project's developments (HLRS).

Energy-aware scheduling

In line with its vision, LRZ had the primary role of advocating an energy-efficient operation of the system by conducting research activities in Work Package 4 on runtime systems. During the first phase of the project, LRZ developers implemented a fine-grained monitoring tool that, among other system parameters, is capable of retrieving the power consumption of the platform at the granularity of a computing node. In the second phase of the project, LRZ researchers successfully collaborated with Bull/Atos for the deployment and test of three experimental scheduling algorithms on the Mont-Blanc prototype that introduce energy-aware features in the resource and job scheduling management system. Specifically, the Power Adaptive Scheduling (PAS) algorithm, the Energetic Fairshare Scheduling (EFS) algorithm and the Energy Cap Scheduling algorithm have been developed. While the first two algorithms are included in SLURM since version 15.08, the integration of the last one is still work in progress.

The PAS algorithm [1] allows for dynamic adjustment of the instantaneous power consumption

of a cluster in order to stay below a pre-defined tolerable power consumption, that is, a "power cap". This is possible by reducing the number of usable resources of the system and/or operating them at lower power. In this way, the scheduler is capable of running jobs only if their consumed power does not contribute in exceeding the defined cap. The original version of the algorithm overestimates the power consumed by compute nodes by using the (theoretical) maximum power consumption for each CPU frequency. The close collaboration between LRZ and Bull/Atos resulted in an improvement of the algorithm by developing a "Power Plugin" that enables the scheduler to acquire the real power consumption of nodes. This yields a more precise estimation of the total cluster power consumption with better assessment of eventual power cap violations, consequently improving the resource utilization of the system. Figure 1 illustrates an example of the advantages offered by this feature with or without the activation of the "Power Plugin".

The EFS algorithm [2] is a modified version of the more common fairshare algorithm where CPU hours are considered as the main resource for accounting users. In EFS, instead, a counter further accumulates the energy consumed by jobs associated with a user and aligns it with the shares of each user account. Energy-efficient users will then be favoured with shorter waiting times in the scheduling queue compared to less energy-efficient ones. In this way, EFS attempts to provide incentives to users for optimizing their codes in order to save more energy.

CDF of wait time between two power management approaches for the Light-ESP benchmark with the CoMD application of 100 jobs on 40 nodes (2 CPUs/node)

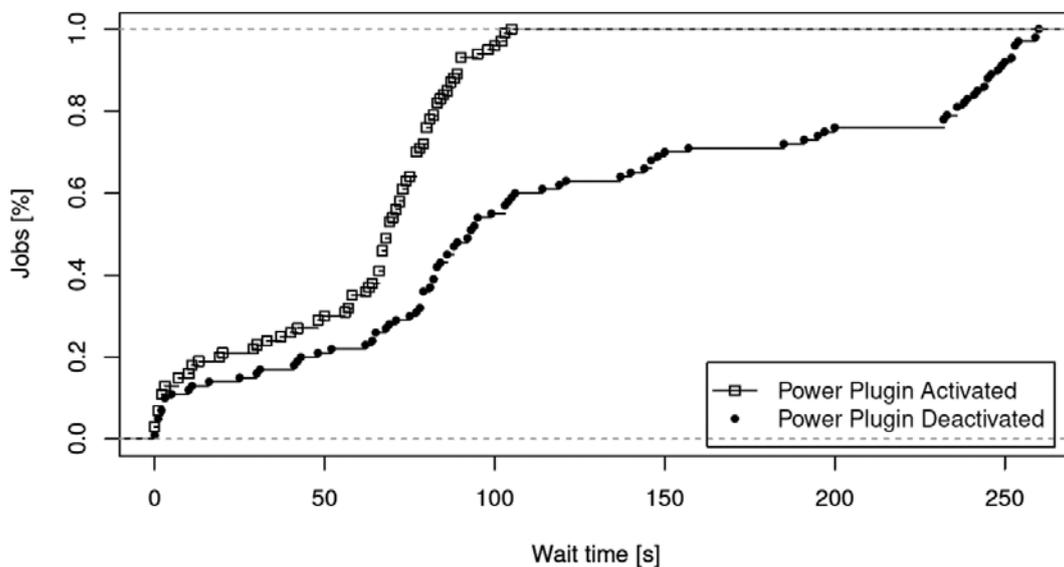


Fig. 1: Cumulative Distribution Function of job waiting time with and without the “Power Plugin”. Note that when the plugin is activated, jobs wait less time in the queue, improving the system utilization. This is due to a better estimation of the power consumed by jobs, based on real values retrieved from the Mont-Blanc power monitoring tool, allowing them to start if the power cap is satisfied. When the plugin is deactivated, power consumed by jobs is always estimated as the maximum possible for the selected CPU frequency, increasing the probability of potential power cap violations and increasing the average job waiting time.

Testing of the EFS scheduler on the Mont-Blanc prototype has been possible through the realization of an additional custom “Energy Plugin” in SLURM, which provides the energy consumed by a user’s jobs over their execution time. Results successfully demonstrated the effectiveness of the algorithm.

Finally, the Energy Cap Scheduling algorithm [3] implements a mechanism to operate a cluster under energy budget constraints by extending the PAS algorithm concept with energy consumption. In contrast to the PAS algorithm, the

Energy Cap algorithm covers more realistic use-case scenarios, where system operators need to establish and maintain an energy budget mainly over a certain period, due to rising electrical energy costs and often due to contractual agreements with providers.

In addition to these activities, LRZ researchers further developed a proof-of-concept for automatic fingerprinting of scientific applications [4], allowing for an optimal selection of the CPU frequency for running user’s jobs and consequently contributing to a more energy-efficient



operation of the system. Initial tests show promising results, motivating further research efforts towards this direction.

Performance analysis and debugging tools

The majority of the efforts of HLRS and JSC were in Work Package 5 on development tools, where both partners extended and improved their debugging and performance analysis tools, respectively.

Besides leading this work package, JSC focused on improving the instrumentation and measurement infrastructure Score-P, the Scalasca Trace Tools for automated analysis of event traces, and the performance report explorer Cube including the underlying libraries. As a first step, the aforementioned tool suites were ported to the 64-bit ARMv8 architecture. Moreover, the Cube software design was reworked to allow for a better extensibility through a newly developed plugin architecture and API [5].

While this plugin architecture is now used by all default Cube-internal views (e.g., system tree, box plot, and topology view), it also allows the development of additional views without touching the main Cube code base. For example, the Cube plugin interface has been used by BSC to develop two new visualizations for the results of their Folding tool [6], which combines coarse-grained sampling measurements with phase instrumentation from iterative applications to statistically improve the accuracy of the measurement results for each phase.

Furthermore, JSC actively participated in the OpenMP tools working group on the definition of the OpenMP tools interface (OMPT), which was voted into the “OpenMP Version 5.0 Preview 1” technical report [7] by the OpenMP Architecture Review Board. OMPT will enable tools to reliably work across different implementations of the OpenMP API. In particular, the Mont-Blanc partners BSC and JSC contributed a proposal to track tasks and their dependencies, BSC implemented draft versions of the OMPT API in their Nanos++ OmpSs runtime, and JSC developed a Score-P prototype based on the OMPT interface, which was subsequently validated with both the Nanos++ runtime as well as an extended version of the LLVM OpenMP runtime developed by the OpenMP tools working group.

With respect to the Score-P instrumentation and measurement infrastructure, JSC enhanced the prototypical support for the OmpSs programming model already developed during the first phase of the Mont-Blanc project. Most notably, the ability to support OmpSs@Cluster mode, where tasks are automatically distributed across a set of worker nodes by the OmpSs runtime, has been evaluated and an initial prototype has been implemented. In addition, Score-P has been enhanced to measure OmpSs tasks offloaded to GPU accelerators. The latter also motivated improvements regarding an integrated analysis and presentation of hybrid applications using multiple programming models—such as MPI, OpenMP or OmpSs, and CUDA or OpenCL—in combination.

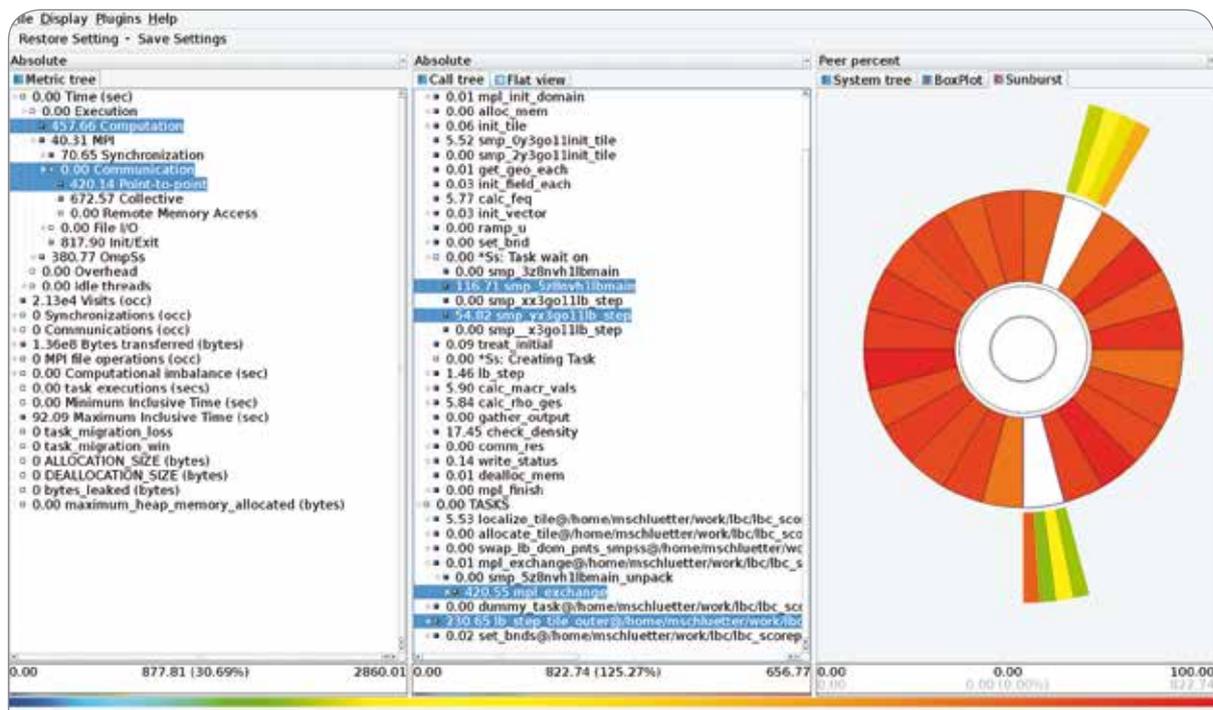


Fig. 2: Screenshot of the Cube performance report browser showing a profile of the OmpSs version of the LBC code measured on the Cavium ThunderX mini-cluster. The left pane shows the various performance metrics collected, the middle pane shows the call tree of the main application (top) as well as the asynchronously executed tasks (bottom), and the right pane shows the “Sunburst” diagram as an alternative visualisation for the system tree, implemented as a plugin using the newly defined Cube plugin API.

HLRS, on the other hand, concentrated on development of Temanejo [8] (Figure 3), a graphical debugger for task-based programming models. The foremost purpose of Temanejo is to display the task-dependency graph of applications, and to allow simple interaction with the runtime system in order to control various aspects of the parallel execution of an application.

Today’s parallel programming models offer high-level concepts, such as task and data-dependencies, to the developers of parallel applications. Traditional debuggers, on the other

side, aim to support a large variety of programming languages and models, and thus need to use the lowest common denominator, such as system calls and POSIX threads. Developers thus can only bridge this semantic gap by having specialised skills and know-how on both sides. Model-centric debugging closes the semantic gap between high-level programming model and the debugging by simply using the same high-level concepts and semantics. In particular, model-centric debuggers represent the state of an application in terms of the programming model and support interaction using its semantics.

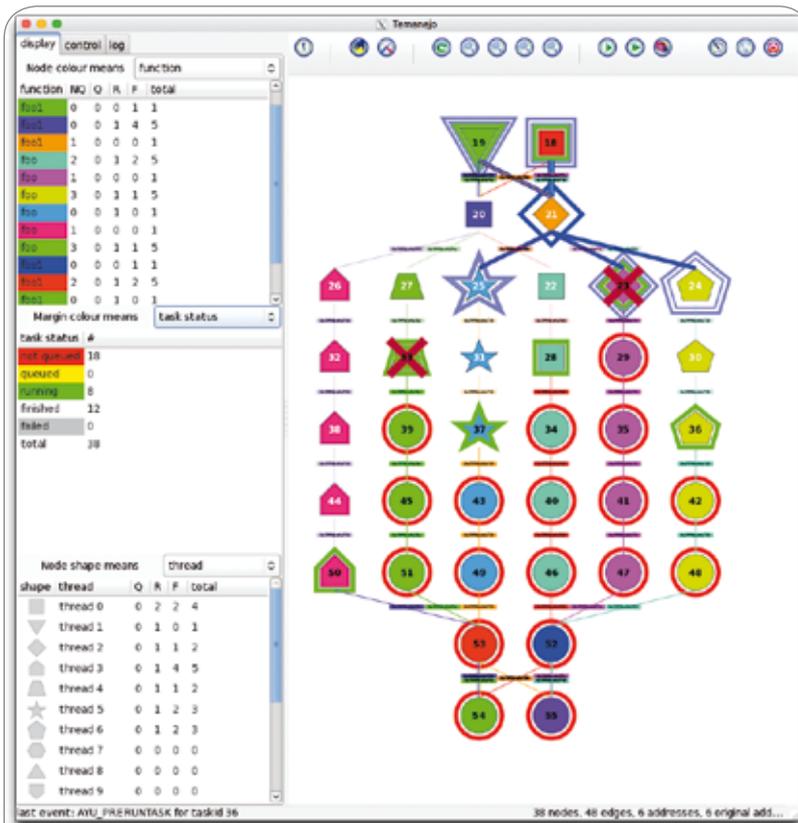


Fig. 3: Screenshot of Temanejo showing a basic dependency graph. The node color represents the function names, the margin color represents the task states and shape represents the different threads.

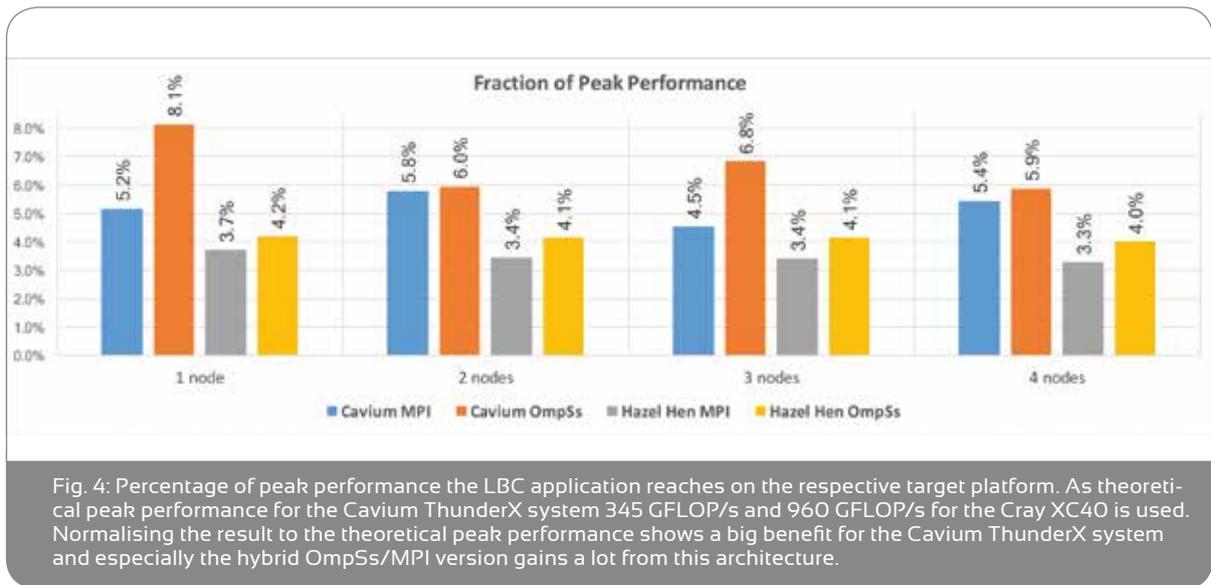
Temanejo was redesigned to support multi-process debugging with multiple attached programming models. Therefore, HLRS had to develop a completely new backend library, called Ayudame. This library is now capable of interfacing with different programming models like OmpSs, StarPU, OpenMP and also MPI. In fact, Ayudame may use the aforementioned monitoring interface OMPT to intercept events from OpenMP and OmpSs runtimes. The native event system of OmpSs continues to be supported. In order to control the programming model's behaviour,

HLRS developed the Tasking Control Interface (TCA) [9] as an extension of the OMPT interface to allow interoperability.

Application porting for evaluation

As part of Work Package 3 on applications, HLRS took an existing version of the LBC code and parallelised it using OmpSs. During the parallelisation process, HLRS developed three different code versions. The first version is a very basic implementation and is based on the fork&join model. This version needs to synchronise all tasks before the communication with the neighbouring processes can be started. The second version hides the communication inside the computation. Therefore, only the tasks necessary for

communication need to be synchronised before the communication with the neighbouring processes can be started. After the communication, the synchronisation with the remaining tasks is necessary. The third version is similar to version two, but in addition the OmpSs programming model is aware of the underlying MPI communication inside tasks. This feature allows the programming model to interrupt the communication task and execute computation tasks while waiting for the MPI communication to finish.



All of the above-described versions were implemented with the help of Temenajo and the tools available in the Mont-Blanc consortium. A performance analysis done by JSC helped us to detect and solve a performance-critical issue. Figure 3 shows a comparison of LBC on two different platforms, the Hazel Hen (Cray XC40) located at HLR5 and the ARM-based Cavium ThunderX mini-cluster. On both platforms, the hybrid OmpSs+MPI implementation shows better performance than a pure MPI version.

Impact

For the three Gauss centres, the second phase of the Mont-Blanc project has been an excellent environment to continue established research lines and develop methods and software packages to a state which is ready for production environments. For instance, the energy-aware scheduling algorithms have been pushed upstream to SLURM since version 15.08 and

more solutions and improvements of the existing mechanisms and on application fingerprinting are expected to become available in future releases. Similarly, the extensions made to the performance analysis and debugging tools Cube, Scalasca, Score-P, and Temenajo are either already available or will be released soon. In general, the Gauss centres are committed to release all software developed within Mont-Blanc funding under an open-source license. The end of the second phase of Mont-Blanc establishes an important milestone, which marks the availability of necessary system middleware/software for research and development in the next phase of the project. Most importantly however, the standardisation of the OMPT interface has a significant impact on the HPC community as a whole. It greatly simplifies the development of new tools for OpenMP via a portable interface across different runtimes, including the Nanos++ OmpSs runtime.



Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under the Mont-Blanc project (<http://www.montblanc-project.eu>), grant agreement n° 288777 and n° 610402.

References

- [1] **Y. Georgiou et al.:**
"Adaptive Resource and Job Management for Limited Power Consumption," IEEE International Parallel and Distributed Processing Symposium Workshop (IPDPSW) 2015, 25-29 May 2015, Hyderabad, India.
- [2] **Y. Georgiou et al.:**
"A Scheduler-Level Incentive Mechanism for Energy Efficiency in HPC," 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2015, 4-7 May 2015, Shenzhen, China.
- [3] **P. F. Dutot et al.:**
"Towards Energy Budget Control in HPC," 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid) 2017, 14-17 May 2017, Madrid, Spain (accepted for publication).
- [4] **The Mont Blanc Consortium:**
"Deliverable D4.9: Final Evaluation of the Compiler and the Runtime System," the Mont-Blanc Project, January 2017.
- [5] **P. Saviankou, M. Knobloch, A. Visser, B. Mohr:**
"Cube v4: From Performance Report Explorer to Performance Analysis Tool", Procedia Computer Science, 51:1343-1352, June 2015.
- [6] **H. Servat, G. Llort, J. Gimenez, K. A. Huck, J. Labarta:**
"Unveiling Internal Evolution of Parallel Application Computation Phases", ICPP 2011: 155-164.
- [7] **OpenMP Architecture Review Board:**
"OpenMP Technical Report 4: Version 5.0 Preview", Nov 2016.
- [8] **S. Brinkmann, J. Gracia, C. Niethammer, R. Keller:**
"TEMANEJO - a debugger for task based parallel programming models," Oct 2012.
- [9] **M. Nachtmann, J. Gracia:**
"Enabling Model-Centric Debugging for Task-Based Programming Models – a Tasking Control Interface," in "Tools for High Performance Computing 2015, Proceedings of the 9th International Workshop on Parallel Tools for High Performance Computing", Sep 2016.

Written by Daniele Tafani¹, Marc Schlütter², Markus Geimer², Bernd Mohr²,
Mathias Nachtmann³ and José Gracia²

¹Leibniz Supercomputing Centre

²Jülich Supercomputing Centre

³Höchstleistungsrechenzentrum Stuttgart

Contact: José Gracia, gracia@hlsr.de

ArCTIC – Adsorption Chiller Technology for IT Cooling

Towards a mechanical chiller free data centre: Adsorption refrigeration at LRZ

LRZ deploys high temperature direct liquid cooled (HT-DLC) HPC systems to utilize chiller-less cooling and reduce the energy spent on cooling. HT-DLC also facilitates heat re-use, e.g. for producing still needed cold water via adsorption refrigeration. This article describes the CoolMUC-2 system at LRZ, the first production level installation of a Top500 HPC system with adsorption technology.

While most data centres operate their IT systems with air cooling around 20°C, direct-liquid cooling is becoming the de-facto standard in HPC. Due to the superior thermal properties of water over air, DLC supports increasing compute densities while also improving Power Usage Effectiveness (PUE). This improved effectiveness is mainly due to the fact that even water at 40°C (and higher) is still sufficient to keep the hot computer components such as CPUs or memory chips within safe operating conditions. The return water from the computer system can then be easily cooled in most climate zones year-round chiller-free, avoiding the need for energy hungry mechanical chillers. At LRZ's location in southern Germany, inlet temperatures of 30°C can be sustained without chillers year-round. The annually averaged Coefficient of Performance (COP) for free-cooling setups is typically around 20, meaning that 1kW of electrical energy has to be spent to remove 20kW of heat. In contrast, even for very efficient mechanical chillers, COPs above 4 are rarely observed. Hence, chiller-free cooling is around five times more efficient than using mechanical chillers.

Another benefit of using DLC is the fact that the heat energy from the HPC system has been captured in water and can be readily re-used in further processes, e.g. for heating office buildings or de-icing walkways in winter. In summer, it can be used as heat source for a thermally driven refrigeration process, such as adsorption refrigeration: the latest generation of adsorption chillers is capable of utilising warm water at temperature levels as low as 50°C to drive an adsorption process that generates cold water at 20°C. This cold water can then be used to cool components that cannot be cooled with HT-DLC, for example storage systems, network switches, and tape archives.

In 2015, LRZ has installed a pilot HPC system to demonstrate and assess the use of adsorption chillers in data centres: the CoolMUC-2 system. It is a Linux compute cluster comprising 6 racks with a total of 384 Lenovo NeXtScale nx360M5 WCT nodes. Each node is equipped with two Intel Xeon E5-2697 v3 14-core Haswell CPUs, 64GB of main memory, and a Mellanox ConnectX-3 FDR Infiniband HCA. The CPUs are direct liquid cooled, whereas the memory modules and the Infiniband adapter are cooled via heat pipes connected to liquid cooled cold rails. The cluster operates at inlet water temperatures of 45°C, which yields outlet water temperatures of around 50°C. The return cooling loop connects to 6 SorTech eCoo 2.0 Adsorption Chillers with a total nominal cooling capacity of 60kW. The cold generated by the adsorption chillers is used to cool the rear door heat exchangers of the SuperMUC Phase2 storage system with an



average power consumption (and hence heat dissipation) of around 50kW. This installation won the 2016 Data Center Dynamics (DCD) EMEA Energy Improver's Award.

The SorTech eCoo 2.0 chillers operate according to the principle of solid matter sorption, also referred to as adsorption. An adsorption chiller requires three water loops. The high temperature (HT) circuit transports the driving heat to the adsorption chiller. The low temperature (LT) circuit transports the generated cold water to the consumer. And lastly, the medium temperature (MT) loop removes the heat generated during the adsorption process via a cooling tower.

The whole system has been extensively instrumented with sensors to monitor temperatures and flow rates in each individual cooling loop as well as the power consumption of all components. This allows for a detailed analysis of the efficiency of the adsorption chiller setup:

In typical day-to-day operations, the average power consumption of CoolMUC-2 is 120kW. Around 72% of the heat generated by the compute nodes is captured in the hot-water cooling loop, whereas the remaining 28% dissipate into ambient air and are removed by the computer room air handler (CRAH) units. Hence, approximately 90kW of heat are available to drive the adsorption refrigeration process. In 2016 the six adsorption chillers generated an average of around 45kW of cold water at 21°C from the heat provided by the CoolMUC-2 cluster.

Using adsorption chillers in a data centre is particularly interesting as they only require little electrical energy to generate cold: electricity is basically only used to drive the pumps in the three cooling loops and to run the fans of the cooling tower. In 2016 this power consumption averaged at 7.2kW. So with only 7.2kW of electricity, 135kW of heat (90kW from the HT loop and 45kW from the adsorption process generating the same amount of cold in the LT loop) have been removed. This translates to a combined COP of 19.3 which is on par with free-cooling.

Yet, looking at these isolated numbers does not show the complete picture. Although the adsorption chillers themselves use only little electrical energy, the higher water temperatures that are required to drive the adsorption process cause higher electricity consumption in other parts of the system. On the one hand, they cause increased leakage currents in the CMOS components of the compute nodes. Experiments with the CoolMUC-2 system showed that the power draw increases by 1.3% when going from 30°C inlet temperature to 40°C and by 1.8% when going to 45°C. This amounts to an additional power draw of about 2.1kW that causes the COP to drop to 14.5 which is still good.

On the other hand, the higher water inlet temperatures also have an impact on the heat capture rate, i.e., the fraction of heat captured in the hot-water cooling loop. As the temperatures level in the cooling loop increase, so does the temperature difference to ambient air (25°C at LRZ). The higher the temperature difference,

the more heat is transferred to air instead of the hot-water cooling loop (2nd Law of Thermodynamics). In the case of CoolMUC-2, the heat capture rate drops from 87.4% at 30°C to 79.5% at 40°C, and further to 72.2% at 45°C. The remaining heat has to be removed from the computer room by the inefficient CRAH units. At 45°C inlet temperatures, 33.4kW of heat have to be removed via air. With the existing infrastructure at LRZ, this translates to an electrical power draw of 7.5kW. At 30°C, only 14.8kW of heat would have to be removed via air using 3.3kW of electrical power.

Taking these numbers into account in the COP calculations, the total COP for the whole system drops to 11.3 at 45°C. Although this is significantly worse than the initially calculated COP of 19.3, the installation of the adsorption chillers still improved the efficiency of the cooling: running CoolMUC-2 without adsorption chillers at 30°C would result in a total COP of 8. This is mainly due the storage racks that need to be cooled via the existing mechanical chiller-supported cooling loop.

Using adsorption chillers in a data centre is feasible, reduces the energy spent on cooling, and can make use of abundantly available waste heat even in summer. However, it induces additional but hidden electrical energy consumption.

Yet, things should improve in the future: next-generation adsorption chillers should allow for lower water temperatures, which would reduce leakage currents and heat dissipation into air. In addition, as HT-DLC becomes more and more mainstream in HPC, the heat recovery rate should improve with better node design. Heat capture rates of 95% seem feasible with insulated and fan-less racks and LRZ is in fact currently procuring a new cluster system with a guaranteed heat capture rate of 97% at 40°C inlet temperatures.

References

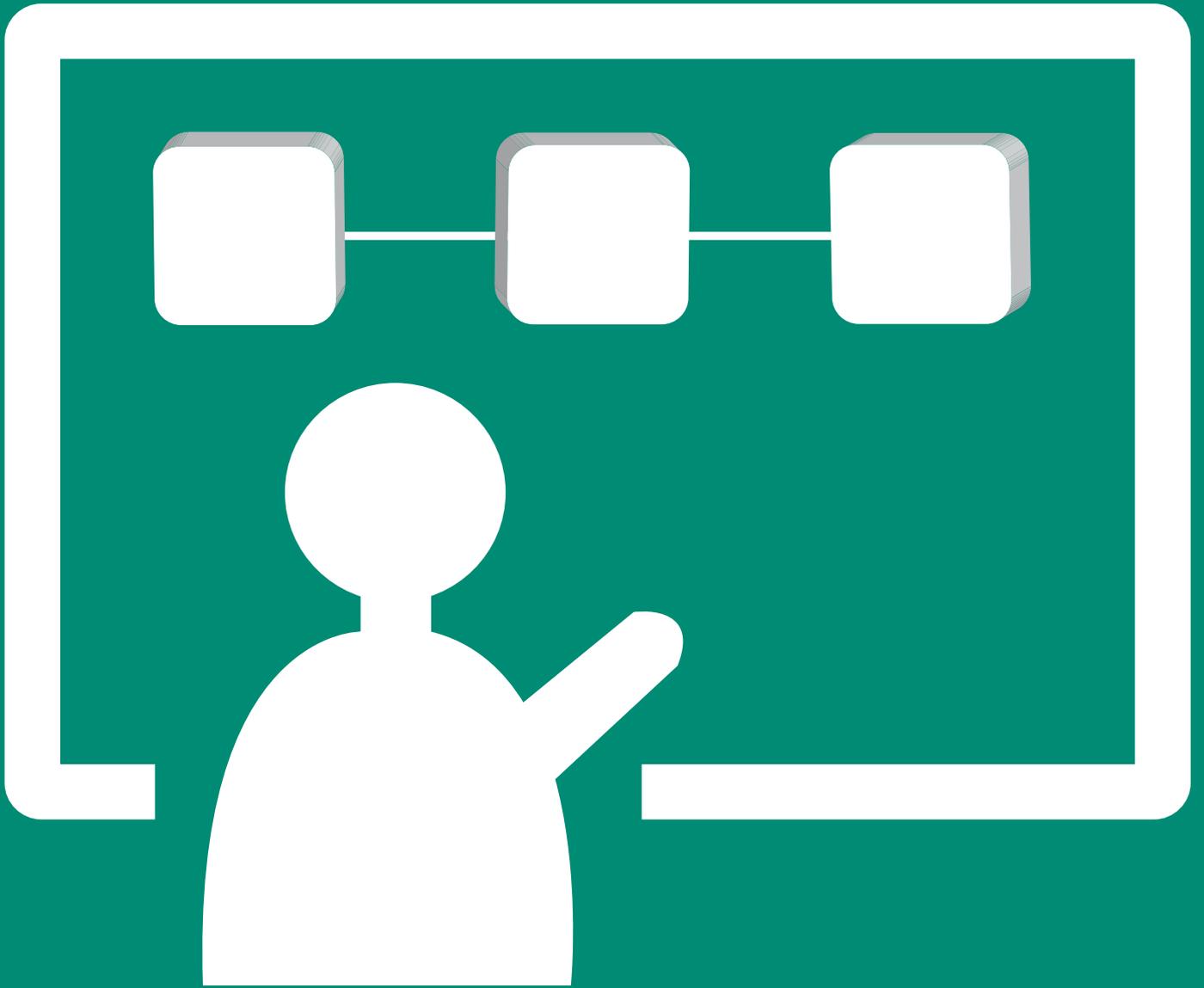
- [1] **Wilde, T., Ott, M., Auweter, A., Meijer, I., Ruch, P., Hilger, M., Kühnert, S., Huber, H.:**
CoolMUC-2: A Supercomputing Cluster with Heat Recovery for Adsorption Cooling, Proceedings of the 33rd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), Institute of Electrical and Electronics Engineers (IEEE), 2017
- [2] **Ott, M., Wilde, T., Huber, H.:**
ROI and TCO Analysis of the First Production Level Installation of Adsorption Chillers in a Data Center, Proceedings of the 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), Institute of Electrical and Electronics Engineers (IEEE), 2017

Written by Michael Ott, Torsten Wilde, Herbert Huber.

Leibniz Supercomputing Centre of the Bavarian Academy of Sciences and Humanities.

Contact: {ott|wilde|huber}@lrz.de.





Centers / Systems

ems / Trainings

In this section you will find an overview about the upcoming training program and information about the members of GCS.



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

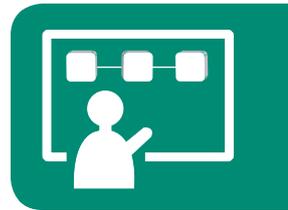
- Giving general IT services to more than 100,000 university customers in Munich and for the Bavarian Academy of Sciences
- Running and managing the powerful communication infrastructure of the Munich Scientific Network (MWN)

- Acting as a competence centre for data communication networks
- Being a centre for large-scale archiving and backup, and by
- Providing High Performance Computing resources, training and support on the local, regional, national and international level.

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



Picture of the Petascale system SuperMUC at the Leibniz Supercomputing Centre.



Leibniz Supercomputing Centre
 Prof. Kranzlmüller
 Boltzmannstraße 1, 85748 Garching near Munich
 Germany
 Phone +49-89-358-31-80 00
 kranzlmue@lrz.de
 www.lrz.de

Compute servers currently operated by LRZ

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
"SuperMUC Phase 1" IBM System x	IBM iDataPlex 9216 nodes, 147456 cores, 288 TByte, FDR10	3,185	Capability Computing	
	205 nodes, 8,200 cores Westmere EX 52 TByte, QDR	78	Capability Computing	
	32 accelerated nodes Knights Corner 76 GByte, FDR14	100	Prototype System	
"SuperMUC Phase 2" Lenovo Nextscale	3072 nodes, 86016 cores, Haswell EP 197 TByte, FDR 14 IB	3,580	Capability computing	German universities and research institutes, PRACE (Tier-0 System)
"CoolMUC2" Lenovo Nextscale	252 nodes, 7,056 cores Haswell EP 16.1 TByte, FDR 14 IB	270	Capability computing	Bavarian Universities (Tier-2)
"CoolMUC3" Megware Slide SX	148 nodes, 9472 cores, Knights Landing, 14.2 TByte, Omnipath	383	Capability Computing	Bavarian Universities (Tier-2)
Compute Cloud Linux-Cluster	200 nodes, 2700 cores	18	Capability Computing	Bavarian Universities, LCG Grid

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

HLRS

First German National Center

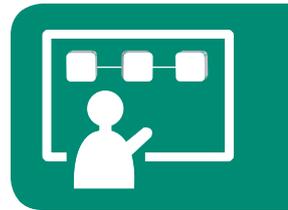
Based on a long tradition in supercomputing at University of Stuttgart, HLRS (Höchstleistungsrechenzentrum Stuttgart) was founded in 1995 as the first German Federal Centre for High Performance Computing. HLRS serves researchers at universities and research laboratories in Europe and Germany and their external and industrial partners with high-end computing power for engineering and scientific applications.

Service for Industry

Service provisioning for industry 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 cooperation, industry always has access to the most recent HPC technology.



View of the HLRS Cray XC40 "Hazel Hen"



**Hochleistungsrechenzentrum Stuttgart (HLRS),
Universität Stuttgart**

Prof. Dr.-Ing. Dr. h.c. Dr. h.c. Prof. E.h. Michael M. Resch
Nobelstraße 19, 70569 Stuttgart, Germany
Phone +49-711-685-8 72 69
resch@hls.de
www.hls.de

Bundling Competencies

In order to bundle service resources in the state of Baden-Württemberg HLRS has teamed up with the Steinbuch Center for Computing of the Karlsruhe Institute of Technology. This collaboration has been implemented in the non-profit organization SICOS BW GmbH.

World Class Research

As one of the largest research centers for HPC, HLRS takes a leading role in research. Participation in the German national initiative of excellence makes HLRS an outstanding place in the field.

Compute servers currently operated by HLRS

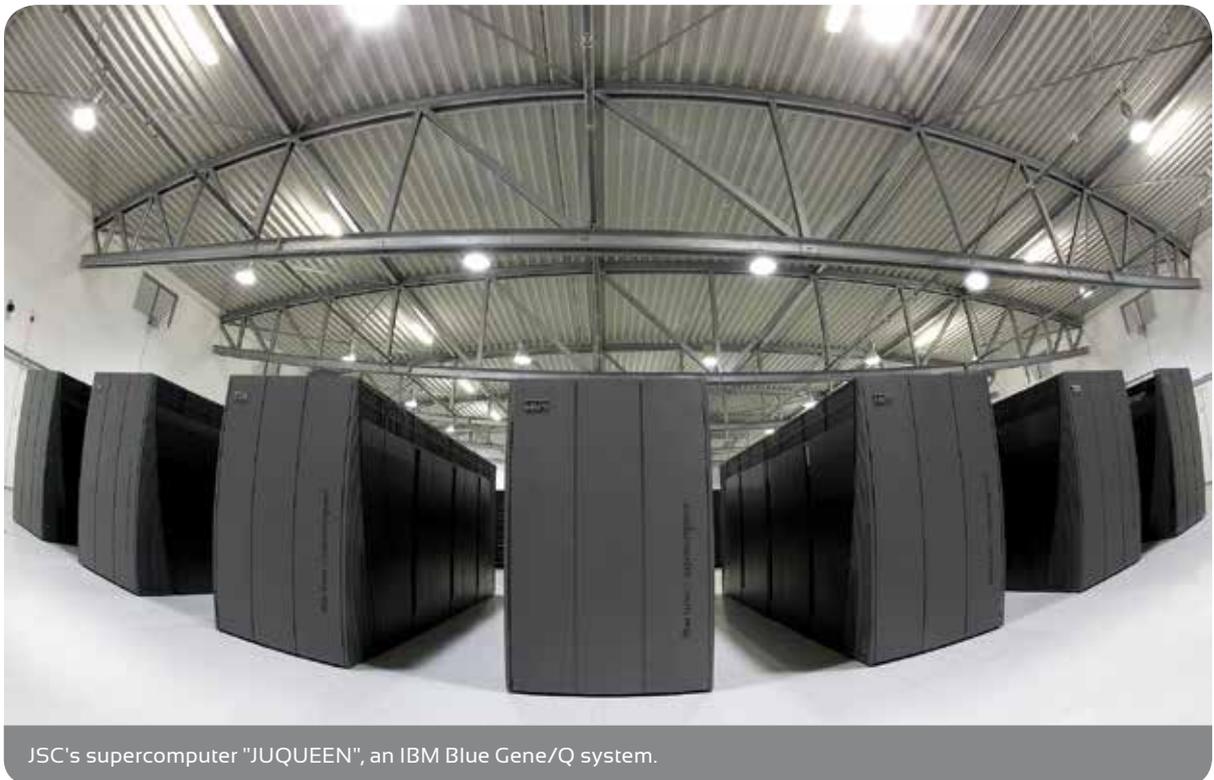
System	Size	Peak Performance (TFlop/s)	Purpose	User Community
Cray XC40 "Hazel Hen"	7,712 nodes 185,088 cores 1 PB memory	7,420	Capability Computing	European (PRACE) and German Research Organizations and Industry
NEC Cluster (Laki, Laki2) heterogenous computing platform of 2 independent clusters	911 nodes 9,988 cores 23 TB memory	170	Laki: 120,5 TFlops Laki2: 47,2 TFlops	German Universities, Research Institutes and Industry
NEC SX-ACE	64 nodes 256 cores 4 TB memory	16	Vector Computing	German Universities, Research Institutes and Industry

A detailed description can be found on HLRS' web pages: www.hls.de/systems

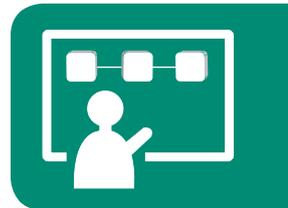


The Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich is committed to enabling scientists and engineers explore some of the most complex grand challenges facing science and society. Our research is performed through collaborative infrastructures, exploiting extreme-scale supercomputing, and federated data services.

Provision of supercomputer resources of the highest performance class for projects in science, research and industry in the fields of modeling and computer simulation including their methods. The selection of the projects is performed by international peer-review procedures implemented by the John von Neumann Institute for Computing (NIC), GCS, and PRACE.



JSC's supercomputer "JUQUEEN", an IBM Blue Gene/Q system.



Supercomputer-oriented research and development in selected fields of physics and other natural sciences by research groups and in technology, e.g. by doing co-design together with leading HPC companies.

Higher education for master and doctoral students in close cooperation with neighbouring universities.

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

Jülich Supercomputing Centre (JSC)
 Forschungszentrum Jülich
 Prof. Dr. Dr. Thomas Lippert
 Wilhelm-Johnen-Straße, 52425 Jülich, Germany
 Phone +49 -24 61-61-64 02
 th.lippert@fz-juelich.de
 www.fz-juelich.de/jsc

Compute servers currently operated by JSC

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM Blue Gene/Q "JUQUEEN"	28 racks 28,672 nodes 458,752 processors IBM PowerPC® A2 448 Tbyte mainmemory	5,872	Capability Computing	European Universities and Research Institutes, PRACE
Intel Linux Cluster "JURECA"	1,884 SMT nodes with 2 Intel Haswell 12-core 2.5 GHz processors each 150 graphics processors (NVIDIA K80) 281 TByte memory	2,245	Capacity and Capability Computing	European Universities, Research Institutes and Industry
Intel Xeon Phi Cluster "QPACE"	352 Intel Xeon Phi 7210 17 TByte memory	937	Capability Computing	QCD Applications SFB TR55

Many-Core Cluster "CoolMUC3" at LRZ

Upon conclusion of a European procurement process, Leibniz Supercomputing Centre (LRZ) has signed a contract with MEGWare (<https://www.megware.com>) for delivery and installation of an HPC cluster based on Intel's many-core architecture. Innovations are not limited to the node architecture, but also extend to interconnect and cooling infrastructure.

LRZ's procurement targets for the new cluster system were:

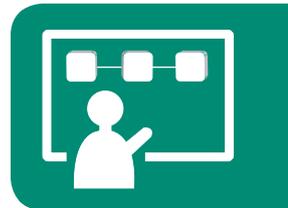
- To supply a system to its users that is suited for processing highly vectorizable and thread-parallel applications,
- provides good scalability across node boundaries for strong scaling, and
- deploys state-of-the-art high-temperature water-cooling technologies for a system operation that avoids heat transfer to the ambient air of the computer room.

Furthermore, the system features an extensible design, enabling seamless addition of further compute nodes of various architectures. Its baseline installation will consist of 148 computational many-core Intel "Knight's Landing" nodes (Xeon Phi 7210-F hosts) connected to each other via an Intel Omnipath high performance network in fat tree topology. A standard Intel Xeon login node will be available for development work and job submission. CoolMUC3 will be comprised of three water-cooled racks, using an inlet temperature of at least 40 °C, and one rack for the still air cooled components (e.g. management servers) that use less than 3% of the systems total power budget. A very high fraction of

waste heat deposition into water is achieved by deployment of liquid-cooled power supplies and thermal isolation of the racks that suppresses radiative losses. Also, the Omnipath switches will be delivered as water-cooled implementations and therefore do not require any fans.

Because of the low-core frequency as well as the small per-core memory of its nodes, the system is not suited for serial throughput load, even though the instruction set permits execution of legacy binaries. For best performance, it is likely that a significant optimization effort for existing parallel applications must be undertaken. To make efficient use of the memory and exploit all levels of parallelism in the architecture, typically a hybrid approach (e.g. using both MPI and OpenMP) is considered a best practice. Restructuring of data layouts will often be required in order to achieve cache locality, a prerequisite for effectively using the broader vector units. For codes that require use of the distributed memory paradigm with small message sizes, the integration of the Omnipath network interface on the chip set of the computational node can bring a significant performance advantage over a PCI-attached network card.

LRZ has acquired know-how throughout the past three years in optimizing for many-core systems by collaborating with Intel. This collaboration included tuning codes for optimal execution on the previous-generation "Knight's Corner" accelerator cards used in the SuperMIC prototype system; guidance on how to do such optimization will be documented on the LRZ web



server, and can be supplied on a case-by-case basis by the LRZ application support staff members. The Intel development environment (“Intel Parallel Studio XE”) that includes compilers, performance libraries, an MPI implementation and additional tuning, tracing and diagnostic tools,

assists programmers in establishing good performance for applications. Courses on programming many-core systems as well as using the Intel toolset are regularly scheduled within the LRZ course program.

(See <https://www.lrz.de/services/compute/courses>)

Overview of CoolMUC3 characteristics

Hardware	
Number of nodes	148
Cores per node	64
Hyperthreads per core	4
Core nominal frequency	1.3 GHz
Memory (DDR4) per node	96 GB (Bandwidth 80.8 GB/s)
High Bandwidth Memory per node	16 GB (Bandwidth 460 GB/s)
Bandwidth to interconnect per node	25 GB/s (2 Links)
Number of Omnipath switches (IOOSWE48)	10 + 4 (48 Ports each)
Bisection bandwidth of interconnect	1.6 TB/s
Latency of interconnect	2.3 µs
Peak performance of system	394 TFlop/s
Infrastructure	
Electric power of fully loaded system	62 kVA
Percentage of waste heat to warm water	97%
Inlet temperature range for water cooling	30 – 50 °C
Temperature difference between outlet and inlet	4 – 6 °C
Software (OS and development environment)	
Operating system	SLES12 SP2 Linux
MPI	Intel MPI 2017, alternatively OpenMPI
Compilers	Intel icc, icpc, ifort 2017
Performance libraries	MKL, TBB, IPP, DAAL
Tools for performance and correctness analysis	Intel Cluster Tools

The performance numbers in the above table are theoretical and cannot be reached by any real-world application. For the actually observable memory bandwidth of the high bandwidth memory, the STREAM benchmark will yield approximately 450 GB/s per node, and the commitment for the LINPACK performance of the complete system is 236 TFlop/s.

DGX-1 and Teramem New Special Systems at LRZ: Machine Learning and Big Data

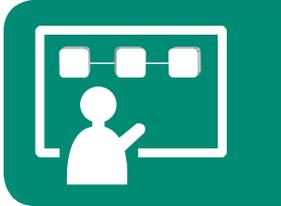
Beginning of 2017, LRZ will provide two new special systems which are dedicated to applications in Big Data and Machine Learning.

The Big Data System consists of a 4-way HP DL 580 Gen9 system containing 96 cores and a total memory of 6.1 TB. Genome analysis, in-memory databases and post processing of large HPC simulations are its major application targets. The system is operated as part of the LRZ Linux Cluster, thus, all Linux Cluster Software is automatically available on the system which consists of Big Data Applications like a 64-bit version of R that can allocate arrays up to the full size of the RAM (6 TB). For R users, it is furthermore possible to use RStudio in the browsers to run jobs on the new system. Users can use the system either interactively or via scheduling batch jobs using SLURM.

The Machine Learning System DGX-1 is a "Super-computer in a box" with a single precision peak performance of 80 TFlop/s. It contains eight high-end GPGPUs from Nvidia (P100) with 16 GB RAM and 28.672 CUDA-compute units which are connected to each other by a NVLink Interconnect and a host x86 compatible system with 40 cores. Users can reserve the whole DGX-1 exclusively and run complex machine-learning tasks, which are available via Docker images. A set of preinstalled images covers deep learning toolkits such as TensorFlow, Theano, CNTK, Torch, Digits and Caffe.

Both systems are available for all scientists of Bavarian universities free of charge.



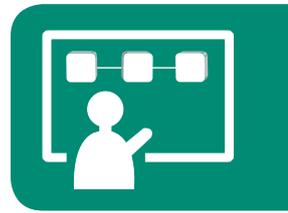


Timetable

High Performance Computing Courses and Tutorials

Course Title	Location	Date
Introduction to hybrid programming in HPC	Stuttgart	Jun 12, 2017
High-performance Computing with Python (PATC course)	Jülich	Jun 12–13, 2017
High-performance scientific computing in C++	Jülich	Jun 20–21, 2017
Intel MIC / Knights Landing Programming Workshop (PATC course)	Garching	Jun 26–28, 2017
Introduction to OpenGL	Jülich	Jun 27, 2017
Usage of VTK for scientific-technical visualization	Jülich	Jun 27, 2017
Introduction to ParaView for visualization of scientific data	Jülich	Jun 28, 2017
HPC for natural hazard assessment and disaster migration	Garching	Jun 28–30, 2017
Introduction to Parallel In-Situ Visualization with VisIt	Jülich	Jun 29, 2017
Introduction to UPC and Co-Array Fortran (PATC course)	Stuttgart	Jun 29–30, 2017
Efficient Parallel Programming with GASPI (PATC course)	Stuttgart	Jul 3–4, 2017
Using R at LRZ	Garching	Jul 4, 2017
Introduction to Cluster Filesystems	Stuttgart	Jul 6, 2017
NEC SX-ACE – Vectorization and Optimization	Stuttgart	Jul 11–12, 2017
Introduction to parallel programming with MPI and OpenMP	Jülich	Aug 15–18, 2017
CFD with OpenFOAM®	Stuttgart	Aug 28–Sep 1, 2017
Iterative Linear Solvers and Parallelization	Garching	Sep 4–8, 2017
Advanced Fortran Topics (PATC course, tbc.)	Garching	Sep 11–15, 2017
Introduction to Computational Fluid Dynamics	Stuttgart	Sep 11–15, 2017
Introduction to the Cray XC40 HPC System at HLRS	Stuttgart	Sep 19, 2017
Cray XC40 Optimization and Scaling Workshop	Stuttgart	Sep 20–22, 2017
High Performance Computing in Science and Engineering	Stuttgart	Oct 5–6, 2017
Porting code from Matlab to Python	Jülich	Oct 9–10, 2017
Introduction to GPU programming using OpenACC	Jülich	Oct 16–17, 2017
Parallel Programming Workshop (MPI, OpenMP and advanced topics) (PATC course, tbc.)	Stuttgart	Oct 16–20, 2017
Scientific Visualization	Stuttgart	Nov 6–7, 2017
Software Development in Science	Jülich	Nov 20, 2017
Advanced C++, Focus on Software Engineering	Stuttgart	Nov 20–23, 2017
Vectorisation and Portable Programming Using OpenCL	Jülich	Nov 21–22, 2017
Introduction to the Programming and Usage of the Supercomputer Resources at Jülich	Jülich	Nov 23–24, 2017
Advanced Parallel Programming with MPI and OpenMP	Jülich	Nov 27–29, 2017
Node-level performance engineering (PATC course, tbc.)	Garching	Nov 30–Dec 1, 2017
Fortran for Scientific Computing	Stuttgart	Dec 4–8, 2017

Visit InSiDE Online for Details



You will also find the detailed course list online at
<http://inside.hlr.de/courses.html>
or by scanning the QR-Code with your mobile device.

Complete and updated list of all GCS courses:

http://www.gauss-centre.eu/gauss-centre/EN/Training/Training/_node.html

Further training courses at members' sites:

<http://www.fz-juelich.de/ias/jsc/courses>

<http://www.lrz.de/services/compute/courses/>

<http://www.hlr.de/de/training/>

Get the Latest News about Innovative High Performance Computing in Germany.

The design goal for future propulsion systems of civil aircrafts is to minimize the environmental impact as well as to optimize the economic benefit. At least for long range flights, liquid fuels will continue to be the preferred and most appropriate energy carrier. Minimizing harmful emissions and, at the same time, maximizing the efficiency of an aircraft engine may be thermodynamically contradictory. The key technology which enables both, low emissions and high efficiencies, is a well controlled combustion process. Controlling combustion mainly comprises a well defined fuel placement inside the combustion chamber. However, the simulation of fuel atomization was not feasible due to the enormous computational requirements of this multi-scale problem and the accuracy required for a correct handling of multiphase flows. Presently, sophisticated combustion simulations still rely on fuel droplet starting conditions, which are a rather rough estimate of the real spray properties. Furthermore, the physical effects of air assisted atomization are not understood in detail.