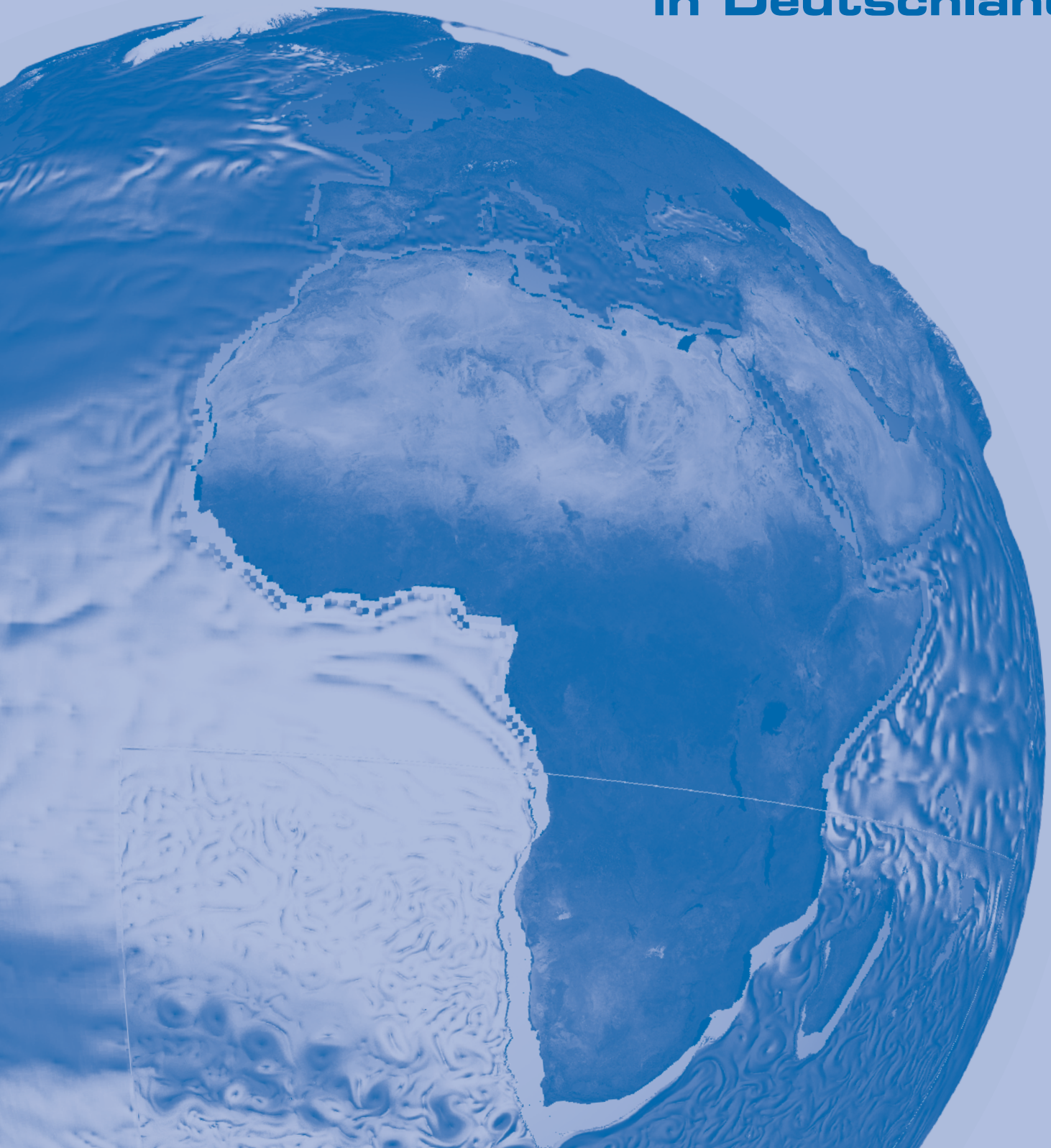
inSiDE

inSiDE • Vol. 9 No.2 • Autumn 2011

Innovatives Supercomputing in Deutschland



Editorial

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

Welcome to this new issue of InSiDE the journal on innovative Supercomputing in Germany published by the Gauss Centre for Supercomputing. As we are approaching our tenth year our journal is also growing in volume. This issue provides again a wealth of articles covering topics of High Performance Computing across the board.

GCS is about to finish the first project phase for Petaflop computing in Germany. In this issue we present the third GCS Petaflop system - a 3 Petaflop IBM cluster called SuperMUC which will be installed at the Leibniz Supercomputing Center (LRZ) at Munich in 2012. As soon as installed, the system will be made available to European users via the European HPC initiative PRACE.

PRACE has grown over the last months. Have a look at the news from PRACE presented here, which also cover the new implementation projects for PRACE. While operation and usage of European HPC systems successfully takes off, implementation projects prepare the next chapter in this European success story.

The application section again shows a wide variety of fields - proof that the thirst for supercomputing cycles is by far not satiated and will keep growing in the future. Aircraft wakes are one crucial problem to be highlighted here. New levels of performance allow resolving vortices in much more detail and provide the ability to simulate aircraft wake vortex evolution both in ground proximity and at cruise altitude.

At the other end of our spectrum, we present a comment to the problem of "how we got to be here". Read it yourself what supercomputing can do to help us answer this question.

While we lay the foundation for future successes through hardware and infrastructure, Germany and Europe are busily preparing the future through research and development projects. As we set out for Exascale systems, two projects describe how we might possibly get there. Here we describe two of the three Exascale system projects funded by the European Commission: the Collaborative Research into Exascale Systemware, Tools and Applications project (CRESTA) and the Dynamical Exascale Entry Platform project (DEEP). Further articles deal with a variety of projects that are at the core of German progress in High Performance Computing. We would like to draw your attention to a description of a European-Russian collaboration in this field which shows how Europe sets out to collaborate internationally to advance.

It's all about GPUs. Jülich Supercomputing Centre has set up a special GPU cluster. Learn more about this machine in our systems section.

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

Enjoy reading!

Contents

Industry inside

High Performance Computing is widely claimed to have an important impact on the economy of a country. HPC is considered to be a key technology in order to stay competitive (see for example [1]). However, a look at the fastest systems in the worlds reveals that still 92% of the top 100 systems are not industrial systems. This discrepancy is striking. One would expect industry to use HPC as much as possible in order to improve its competitiveness. In this article we describe the model success-fully deployed at HLRS achieving HPC integration in the simulation processes in industry on large scale computing facilities already for many years. Furthermore we present new approaches of how to extend the reach of HPC both in type of company (SME) and in type of usage.

Roadblocks

So what are the roadblocks for industrial HPC making it obviously so difficult to use HPC in industry? First, there are a cost issues. HPC is expensive. When we look at the top 100 systems in the world we have to consider investment costs in the order of 5 Mio. Euro and additional fixed operational costs in similar range over an operation period of 3 years. Furthermore the increased power demand and corresponding importance for an infrastructure with a high level of power efficiency lead to substantially raised investment costs for the infrastructure. For a top 100 system one would estimate this to be in the order of another 5 to 10 Mio. Euro at least every 5 years. All in all HPC requires a relatively high level of financial effort potentially delivering benefit in the long term but does not guarantee immediate return of investment. Additionally costs for HPC investment are fixed costs. They can hardly be reduced in time of crisis.

So the obvious alternative is outsourcing of HPC. However public HPC systems operate in a well-defined environment - which typically is exactly the opposite of what industry requires. The focus of a public HPC center is on openness (we can safely ignore the special cases of classified systems as these - by nature - are not available to industry). The service provided is best effort as typically the services are provided for free or based on a research grant. For industry openness is not a problem per se but very often simulation data and software are sensitive. So, ways have to be found to guarantee data security and safety. Best effort was acceptable 10 to 15 years ago when simulation in industry was an accompanying activity. Now, as simulation is one tool to develop new products and experiments are often made mainly to support results of extensive simulations, simulation is a core part of what industry calls production. And as much as industry is worried about stopping physical production lines, it also starts to worry about its virtual production lines - as can be found in High Performance Computing simulation.

Public HPC for Industry

High Performance Computing at the University of Stuttgart has a long tradition. At least since 25 years the computing center has operated systems together with industry. Common investments were made back in 1986. Hence, when the High Performance Computing Center Stuttgart / Höchstleistungsrechenzentrum Stuttgart (HLRS) was founded in 1996 as the first German national supercomputing center an industrial strategy was an important part of the concept. In the last 15 years this concept has been extended. Geographically this was achieved by setting up very close relations with the Karlsruhe Institute of Technology (KIT) and its Steinbuch Center for Computing (SCC). Organization wise this is amended by continuously working with our industrial partners further developing the necessary tools and organizational frameworks not only providing HPC cycles to industry but beyond that shaping the processes and methods required for industrial HPC adoption. In 2011 this has been moved to a new level introducing a couple of dramatic changes to the existing co-operation models. Changes, that are extremely important for HLRS but that can also serve as a model for a general concept of industrial HPC usage world-wide.

An industrial Solution

Industrial use of HPC systems comes in different flavors that require different approaches to meet their individual needs. There is no one size fits all approach for this. In the following we look at three key concepts that have drawn some attention in the last years:

- Simulation production in industry
- Improvement of industrial in-house methods
- Precompetitive research.

Simulation Production

Simulation as described above is part of every-day design and development process in many companies already. As such simulation is part of process chains and has to be integrated into overall processes. HPC needs to be made available on a reliable and secure basis. No efforts for research are required.

HLRS set up a solution together with industrial partners already in 1995. The Höchstleistungsrechner für Wissenschaft und Wirtschaft Betriebsgesellschaft mit beschränkter Haftung (hww) was established together with Daimler and Porsche. Over time its ownership moved from Daimler to T-Systems which is currently holding 40% of hww. Porsche is still holding 10%. Over time HLRS invited further universities to join them in hww such that today 50% of the shares of hww are evenly split between three universities and the State of Baden-Württemberg.

With HPC becoming more and more part of the production process the operational model of hww had to be changed too. Over the last 3 years the advisory board of hww has worked out a new business concept. According to this new concept High Performance Computing is no provided as a commodity via hww. hww serves as a platform and can provide access to various HPC systems on demand in a cloud-like way. The success of moving to such a service oriented approach can easily be seen from the usage of HPC systems at HLRS. Over the last years usage has increased from roughly 2 million core hours in 2007 to an expected new record of about 18 Mio core hours in 2011. Current usage numbers and discussions with industry indicate that the growth will continue over the coming two years.

The success of hww is based on a number of factors:

- Clear technical concept for operation including security measures conforming to industrial standards and operational procedures agreed upon between industry and pubic providers.

- Clear financial concept making sure that both public requirements and economic necessities are combined to find an optimum economic solution.

- Clear legal procedures that ensure that all legal regulations are met and that tax issues are resolved in such as way as to eliminate the pending risks of taxation for the public sector.

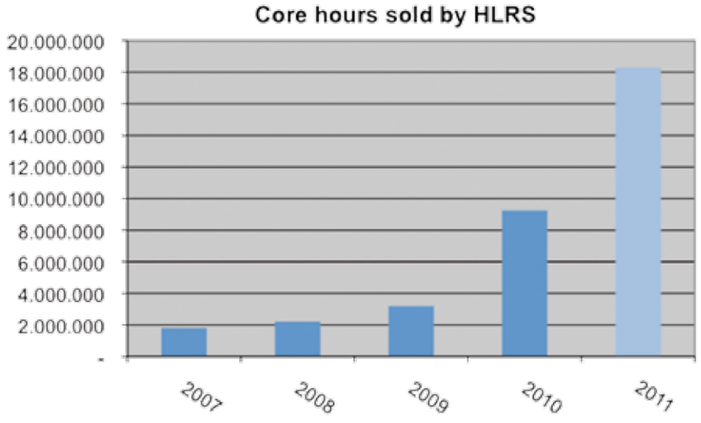
Even though sharing public resources with industry imposes new rules and additional work on the public side it also comes with a number of benefits.

- Clarification of financial aspects of HPC. A comparison of public HPC services with private cloud offerings is easily possible and can be done any time.
- Increased level of operational security which is also beneficial for the public users.
- Stabilization of financial planning as industrial usage gives political backing for HPC investment.

Improving in-house Methods

Improvement of in-house methods is often discussed in the debate about private use of public HPC resources. The argument goes that industry needs to have access to HPC in order to work with research to improve its processes.

In Germany there is a long tradition for industrial - university co-operation that has created a framework that is unique worldwide. In the fields of engineering, chemistry, electronics and many other applied sciences research departments of universities and research organizations typically work together with industry in research projects. Very often such collaborations have been established 50 or more years ago and are handed over from director to director over time. As part of such well-established frameworks public and private research are well connected. Currently we estimate that about one third of all projects running on the systems of HLRS have some relation with a public industrial research project. The very positive thing about such projects is that the results achieved are made publicly available. Only if the research is done privately and results are kept confidential industry has to go through the hww model as described above.



Another advantage of this model is that many master theses and PhD theses are done in collaboration between public research departments and industry. Through this process a lot of know how is created on both sides that benefits both sides. Young researchers get a first look into industrial processes during their thesis. Indus-try gets to work with the most advanced methods and can easily recruit well trained staff through such co-operations.

Pre-competitive Research

When it comes to international competitiveness a national approach is important. But, on the other hand, competition is not only an international issue but also a national one. So, one has to find a way to target general research topics by getting the best researcher to focus their mind on them since that may define the competitiveness of a national economy in the decades to come. And at the same time one has to find a way to keep industrial companies close to such research in order to allow an easy and quick transition of research results into the industrial simulation process. HLRS and KIT have decided to set up "Solution Centers" in order to meet both requirements. Today the follwogin are active:

- Automotive Simulation Center Stuttgart (ASCS)
- Energy Solution Center Karlsruhe (ENSOC).

Both centers bring together researchers and industry. Both centers have integrated HW and SW-vendors to make sure that whatever new methods are worked out by researchers can immediately be integrated into existing commercial software and optimized for existing hardware fast

Beyond large Scale Industries – SME-to-HPC

Usage of HPC requires a high level of expertise. Furthermore it comes with some costs. Finally there is a trust issue concerning industrial data. That's why SMEs typically are not at the forefront of public HPC usage. All three issues are extremely difficult to handle. Expertise for HPC is typically not part of the core business of smaller companies. The available budget for HPC usage is tight and even small initial investments may be prohibitively high. Finally, knowledge and data are very often the key and only assets of a small company.

HLRS and SCC are therefore specifically taregtting SMEs in the State of Baden-Württemberg. In Baden-Württemberg a lot of innovation is delivered by these small companies. Any support for them may create a substantial local return of investment

Support for SMEs relies on trust. The focus had to be on local initiatives. Trust furthermore requires a lot of personal involvement. Especially at the beginning it is important to have people at hand that listen and provide solutions if problems occur. HLRS and SCC decided to set up a new company. Following the acronym of hww (providing pure cloud like HPC compute services) the new company was named HVV (Höchstleistungsrechner und Verteiltes Rechnen Verbund). One of the key activities of HVV will be to identify SMEs in Baden-Württemberg that may benefit from HPC, to educate them on the benefits of HPC, to propose projects and to support them in integrating HPC into their own production schemes. HVV will have to work closely with HLRS and SCC to be successful. However, we expect to see a tremendous impact on HPC usage in SMEs over the next 5 years if the concept works out.

Conclusion

Support for industry in HPC is a difficult task for any public center. Howvwer, experience shows that one can be very sucessful and competetive focusing on the right issues. These where resolved successfully in the State of Baden-Württemberg and the findings can be useful beyond. Going beyond pure cycle provisioning speed of innovation is addressed by solution centers. Based on HVV we address the SME-sphere. Both activities will require more efforts in the coming years and are expected to deliver a high return of investment for society and economy. Both approaches should be extended beyond the State of Baden-Württemberg.

References

[1] **Joseph, Earl C. et al.**
A Strategic Agenda for European Leadership in Supercomputing: HPC 2020 - IDC Final Report of the HPC Study for the DG Information Society of the European Commission, http://www.hpcuserforum.com/EU/downloads/SR03S10.15.2010.pdf, last accessed August 14, 2011

• Michael Resch,

University of Stuttgart, HLRS

New HPC Systems at Leibniz Supercomputing Centre

During August 2011, a number of new computing systems have been taken into user operation at LRZ, increasing the aggregate installed peak performance by more than 100 TFlop/s. The following subsections give a short description of each system. In a final subsection, an outlook to the near future is given.

A water-cooled Cluster delivered by MEGWare

As a significant extension to LRZ's parallel tier-2 computing facilities, MEGWare has installed 178 nodes based on dual-socket AMD Magny-Cours nodes with 8 cores per socket; the systems' interconnect is also based on Infiniband QDR technology. Its total peak performance is 22 TFlop/s. This system will also serve as a research platform for energy efficiency and cooling technologies, since a large part of it is water-cooled on the system board level with a high inlet temperature (>38 °C), and a smaller part of it will be air-cooled with deployment of modern absorption cooling technology. On the software side, the SLURM scheduler will be deployed to process user jobs, and Par-Tec's (www.par-tec.com) Parastation MPI supports execution of large scale distributed memory parallelism.

Figure 1:
Dr. Herbert Huber, LRZ, shows the water-cooled MEGWare cluster.

Large shared Memory Systems from SGI

In order to fulfill some user's requirements for large-scale shared memory computing on the tier-2 level as well as to enable a low level of administrative overhead on LRZ's part, two large shared memory partitions based on the SGI Ultraviolet architecture have been installed. The system comprises a total of 2,080 Intel Westmere cores, and each of the partitions provides a shared memory area of 3.3 TBytes. The cache coherency throughout such a partition is guaranteed by SGI's proprietary NUMAlink 5 Interconnect and additions to the standard Intel chipset on the mainboards; as a trade-off, stronger ccNUMA effects for access to remote memory must be accounted for by the programmer. The peak performance of the system is 20 TFlop/s. Similar to the MEGWare cluster, jobs will be scheduled via SLURM; and SGI's ProPack software stack, which includes a highly tuned MPI implementation, is available in the Ultraviolets' programming environment.

Migration System in Preparation for IBM's SuperMUC

In order to give users of the tier-1 computing facilities the opportunity to port and tune their codes for the Petaflop-class SuperMUC (see below), IBM has delivered an island of 205 Intel Westmere-EX based nodes, each containing four 10-core sockets and 256 GBytes of memory, with a peak performance of 78 TFlop/s. The high-performance interconnect has the topology of a non-blocking fat tree and is based on QDR (quad data rate) Mellanox Infiniband technology. On the June 2011 Top 500 list, the system achieved rank 166, thereby overtaking LRZ's previous flagship system, the Altix 4700, which achieved rank 198. A fraction of the cycles will also be provided to European level (tier-0) scientific endeavours. IBM also provides a full top-level HPC software stack with LoadLeveler plus PE (parallel environment) as batch scheduler and a tuned and PE-integrated MPI implementation. The lower-level software stack, which includes compilers, optimized libraries, and processor-specific tuning tools, is provided by Intel, and this is used on all HPC systems at LRZ.

Targeting Petaflop Class Performance in 2012

After more than 5 years of continuous operation, the Altix 4700 based tier-1 system HLRB II will be replaced by its successor, named SuperMUC, which will boost the peak performance installed at LRZ beyond 3 Petaflop/s and allow LRZ to make a significant contribution to the tier-0 level of

Architectural Feature	SuperMUC (IBM)	HLRB II (SGI)
Number of cores	> 110,000	9,728
Peak performance (TFlop/s)	> 2,900	62
Main memory (TByte)	> 300	39
Memory bandwidth per core (GB/s)	6.4	4.3
Bisection bandwidth of system (GB/s)	35,600	820
Scratch disk capacity (TByte)	10,000	600
Aggregate bandwidth to scratch disk (GB/s)	200	40
Electrical power consumption (MW)	> 3	1

European supercomputing. This system, which will be delivered by IBM and whose initial operation is planned for the second quarter of 2012, was explicitly designed to be as "standardized" as possible: It potentially allows running programs developed and built on a desktop system without any porting effort. The enormous performance results from having more than 6,900 compute nodes containing two Intel SandyBridge-EP sockets with 8 cores each. Compared to the migration system described above, the Infiniband interconnect will be updated to a newer technology release (FDR) which will provide a higher communication bandwidth. The large scale architecture of the system is described by a hierarchy dubbed "islands": within an island containing some thousand cores, a fully non-blocking fat tree interconnect is established, while connectivity between islands is pruned by a factor of 4 in bandwidth, while still having fat tree topology. Eighteen of these islands make up the complete system. For such a system, high-capacity and high-throughput disk storage is also needed: 10 PBytes of disk space operated with IBM's General Parallel File System (GPFS) will be available on SuperMUC.

Great effort is being invested to achieve as energy-efficient operation and usage of the system as possible: Firstly, high-temperature water cooling on the system board level will be used to reduce the total energy consumption by up to 30%. Secondly, Intel's processor technology allows running the CPUs at different frequencies, thereby controlling energy consumption on the system level; in conjunction with monitoring and scheduling software techniques, user's jobs can be tuned by the system to run at the lowest frequency possible without a significant impact on job performance. The following table compares the features of SuperMUC with those of its predecessor, the HLRB II.

Once SuperMUC is fully operational, the migration system will be integrated as a nineteenth "fat node" island, providing large shared memory nodes to applications which need this kind of facility.

• Reinhold Bader

Leibniz Supercomputing Centre (LRZ)

Figure 2:
Dr. Reinhold Bader, LRZ, in front of the SGI Ultraviolet system.

News from PRACE: HPC Access Grants and Implementation Phase Projects

The Partnership for Advanced Computing in Europe (PRACE) is offering supercomputing resources on the highest level (tier-0) to European researchers.

Jülich Supercomputing Centre (JSC), as one of the three members of the Gauss Centre for Supercomputing (GCS), is dedicating a 35% share of its Blue Gene/P system JUGENE.

Project Access is intended for individual researchers and research groups including multinational research groups and has a one year duration. Calls for proposals for project access are issued twice yearly (May and November).

The 2nd call for proposals, this time for computing time on JUGENE and Curie, the system hosted by CEA, Paris, France, closed January 11, 2011. Ten research projects have been awarded about 360 Million compute core hours on JUGENE. Three of those research projects are from the UK, two are from Germany, and one each from Denmark, Ireland, Spain, Sweden, and Switzerland. Three research projects are from the field of Astrophysics, two each from Fundamental Physics from Medicine and Life Sciences, and one each from Chemistry and Materials, Earth Sciences and Environment, and from Engineering and Energy. More details, also on the projects granted access to Curie, can be found via the PRACE web pages **www.prace-ri.eu/PRACE-2nd-Regular-Call**

Evaluation for the 3rd call for proposals that closed June 22, 2011 is still under way, as of this writing. Details on the calls can be found on **www.prace-ri.eu/Calls-for-Proposals**

The EC-funded PRACE implementation phase projects met in Barcelona from September 14 to 16, 2011. PRACE-1IP, focussing on the operation of the tier-0 infrastructure, application support and future HPC technologies, planned the work for its second year. PRACE-2IP which additionally focusses on the coordination and integration of national tier-1 HPC systems, building on the work of the DEISA projects, held its kick-off meeting. More than 150 participants from the 21 member countries attended the meetings.

• Walter Nadler

Jülich Supercomputing Centre

JUGENE unravels Structures in Phase Change Materials

Applications

Phase change (PC) materials remain leading candidates for future computer Random Access Memory (RAM) and rewritable storage devices (CD-RW, DVD-RW, DVD-RAM, and Blu-ray Disc, BD). The battle to replace the de facto standard of optical recording materials, the Digital Versatile Disk (DVD), was decided in 2008 in favour of Blu-ray Disc, and the recording media of all BD products involve PC materials. Many readers will know that laser light is involved, but not all know how the information is recorded, read, and erased.

Information is stored in these devices in the form of microscopic bits (each less than 100 nanometers in size) in a thin layer of a polycrystalline alloy containing several elements. The bits can have a disordered (amorphous) or an ordered (crystalline) structure.

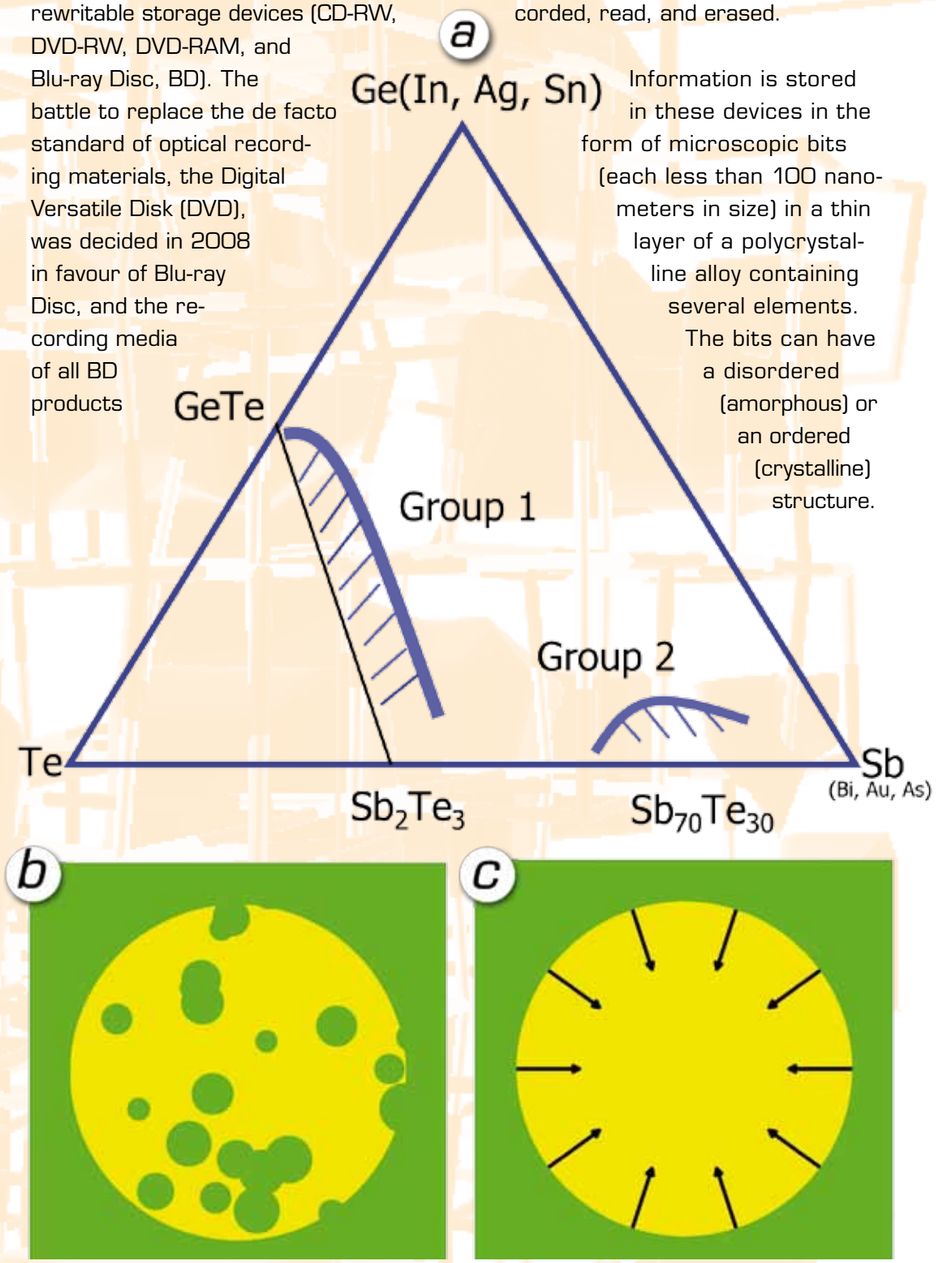


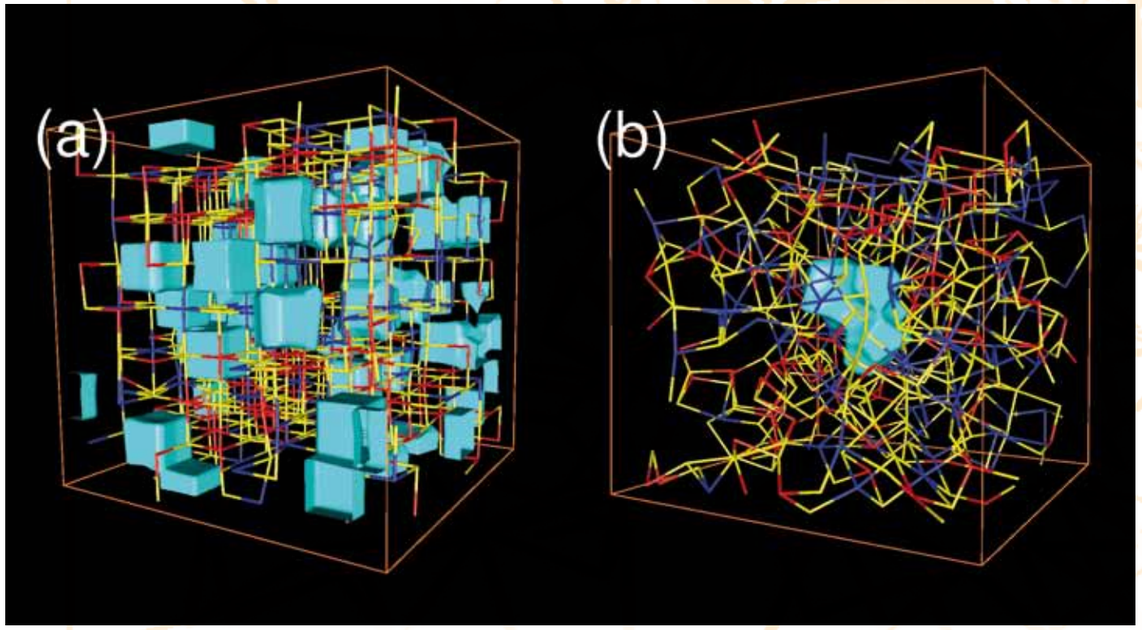
Figure 1:
Phase diagram of PC materials and crystallization patterns.
(a) The most commonly used materials for optical recording are in Groups 1 and 2,
(b) Nucleation-dominated growth recrystallization (as in GST),
(c) Growth-dominated recrystallization (as in AIST).

The transition between the two phases is not only extremely rapid (some tens of nanoseconds), but it is reversible. Amorphous bits are formed by quenching after a localized and short (~ 1 ns) pulse to above the melting temperature. Longer laser heating (of the order of 10 ns) to above the glass transition but below the melting point returns the bit to a metastable crystalline form. The state can be identified by monitoring the optical or electrical properties.

The physical requirements of PC materials, particularly the rapid crystallization, are satisfied by relatively few materials. The focus for some years now has been on alloys of three or four elements, many of which contain germanium (Ge), antimony (Sb) and tellurium (Te), and are referred to as "GST alloys" (Fig. 1a). Alloys along the line connecting GeTe and Sb_2Te_3 (Group 1) are common PC materials, particularly $Ge_2Sb_2Te_5$, which is used in DVD-RAM, and alloys near the GeTe end, which are favoured in BD applications. Also shown in Figure 1 are alloys near Sb (70%) and Te (30%) (Group 2). With small amounts of silver (Ag) and indium (In), "AIST" alloys are in widespread use in DVD-RW devices. Although both alloy families contain antimony and tellurium and appear to have much in common, the phase change mechanisms are very different. In GST materials, the amorphous bit crystallizes via nucleation, i.e. small crystallites formed in the interior grow rapidly until they cover the whole bit (Fig. 1b). The phase change in AIST alloys proceeds from the outside of the bit, where it adjoins the crystalline surroundings, towards its interior (Fig. 1c).

Materials in both groups have superior rewrite speeds and are stable at room temperature for long periods, indispensable characteristics of PC memories.

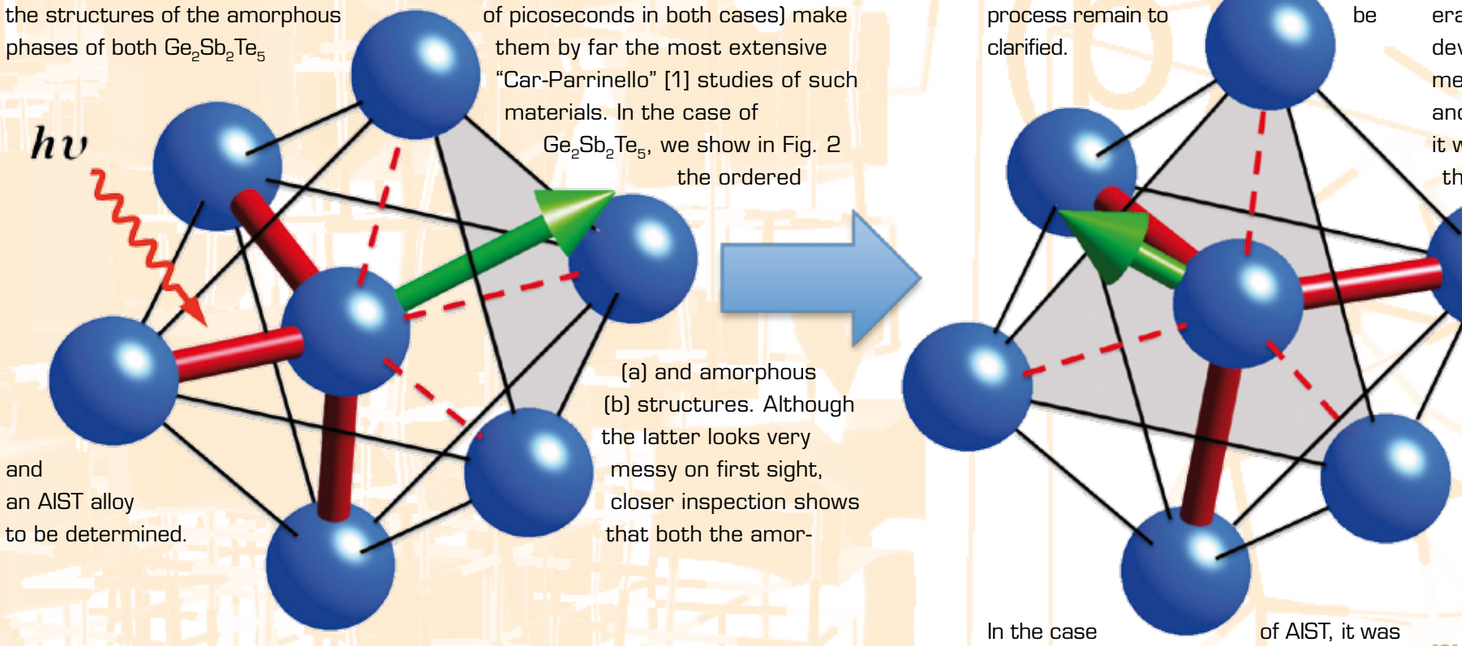
Figure 2: The simulation of 460 atoms and 52 vacancies in (a) crystalline GST and (b) amorphous GST. Red: Ge, blue: Sb, yellow: Te. Cavities are shown as light blue isosurfaces. A single large cavity is one of numerous cavities in a-GST.



Until recently, however, there was little understanding of the crucial - and rate-limiting - crystallization process and no convincing explanation for its astonishing rapidity. The combination of extensive simulations on the Jülich supercomputer JUGENE with new experimental data and x-ray spectra from the Japanese synchrotron SPring-8, the world's most powerful x-ray source, has enabled the structures of the amorphous phases of both $Ge_2Sb_2Te_5$ and an AIST alloy to be determined.

It has been possible in both cases to develop models that explain the rapid phase change.

The calculations involved density functional simulations of the cooling process from a high-temperature liquid down to room temperature, and the number of atoms in the unit cell (460 in GST, 640 in AIST) and the total simulation times (hundreds of picoseconds in both cases) make them by far the most extensive "Car-Parrinello" [1] studies of such materials. In the case of $Ge_2Sb_2Te_5$, we show in Fig. 2 the ordered (a) and amorphous (b) structures. Although the latter looks very messy on first sight, closer inspection shows that both the amorphous and crystalline phases contain the same structural units, "ABAB" rings. These four-membered rings contain two germanium or antimony atoms (A) and two tellurium atoms (B) and can rearrange in the available empty space without breaking many atomic bonds. This observation is consistent with the very rapid crystallization of the amorphous phase, but the details of the process remain to be clarified.



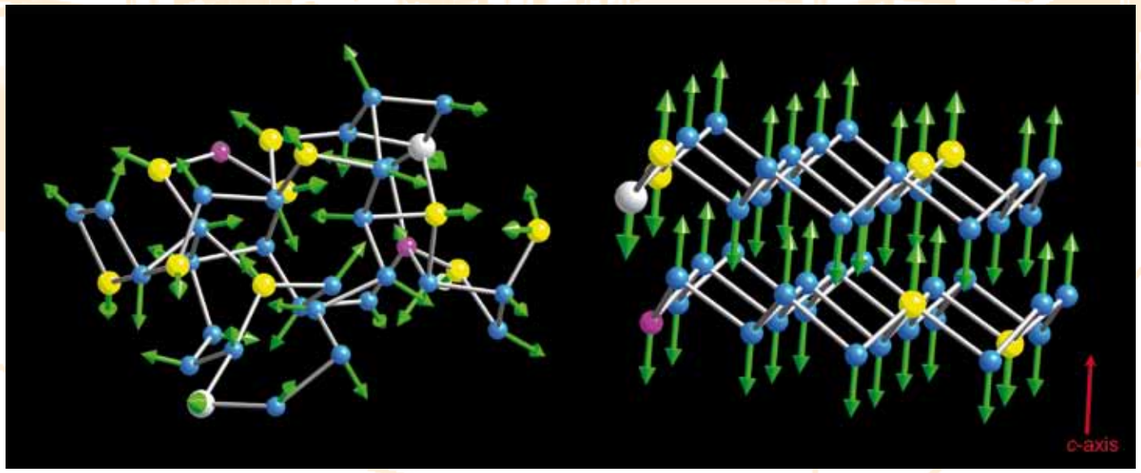


Figure 3: Model of crystallization of an AIST alloy. Upper left: a laser pulse (hv arrow) causes motion of the central antimony atom (left), which then exchanges its bonds to two neighbours. Upper right: The green vector sum of the three short red bonds changes. Below: A sequence of such processes leads from the amorphous (left) to the crystalline form (right).

In the case of AIST, it was possible to determine a structure of the amorphous phase that reproduced the x-ray scattering data of our Japanese colleagues [3].
In conjunction with the known structure of the crystalline phase, this has again allowed us to propose a model for the crystallization process (Fig. 3). In this "bond exchange model", the local environment in the amorphous bit is changed by small movements of an antimony atom (see Fig. 3, above) that result in the exchange of a short and a long bond. A sequence or avalanche of many such steps results in reorientation (crystallization), without requiring large atomic motions or the empty regions that were present in GST, but are absent here. The antimony atoms, stimulated by the laser pulse, have simply exchanged the strengths of the bonds to two neighbours.

The deeper theoretical understanding of the processes involved in writing and erasing memory devices should aid the development of phase change storage media with longer life, larger capacity, and/or shorter access times. Naturally it would be highly desirable to simulate the actual crystallization process and see whether these models are more than just plausible explanations. The time scale involved makes such calculations extremely challenging, but we are performing them for GST at the moment and are very optimistic about the prospects.

References

[1] **Car, R., Parrinello, M.**
Unified approach for molecular dynamics and density functional theory, Phys. Rev. Lett. 55, 2471 (1985)

[2] **Akola, J., Jones, R. O., Kohara, S., Kimura, S., Kobayashi, K., Takata, M., Matsunaga, T., Kojima, R., Yamada, N.**
Experimentally constrained density-functional calculations of the amorphous structure of the prototypical phase-change material $Ge_2Sb_2Te_5$, Phys. Rev. B 80, 020201(R) (2009)

[3] **Matsunaga, T., Akola, J., Kohara, S., Honma, T., Kobayashi, K., Ikenaga, E., Jones, R. O., Yamada, N., Takata, M., Kojima, R.**
From local structure to nanosecond recrystallization dynamics in AgInSbTe phase change materials, Nature Materials 10, 129 (2011)

• Jaakko Akola[1]
• Robert O. Jones[2]

[1] Tampere University of Technology, Finland

[2] Forschungszentrum Jülich, PGI-1, Jülich

Autumn 2011 • Vol. 9 No. 2 • inSiDE 15

The Effect of Dust Cooling in the Fragmentation of Star-forming Clouds for the Transition from POP. III to POP. II

How are the Stars formed in the early Universe?

The first burst of stars in the universe (Population III) are thought to give rise to massive stars, with characteristic masses of few solar masses ($M_\odot$)(Clark et al., 2011, a), in contrast with present-day star formation (Population I/II), which tends to yield stars with less then 1 $M_\odot$ (Kroupa, 2002). The most widely accepted cause for the transition from Pop. III to Pop. II is metal enrichment of the interstellar medium by the previous generations of stars (1). This suggests that there might be a critical metallicity Z_{crit} at which the mode of star formation changes.

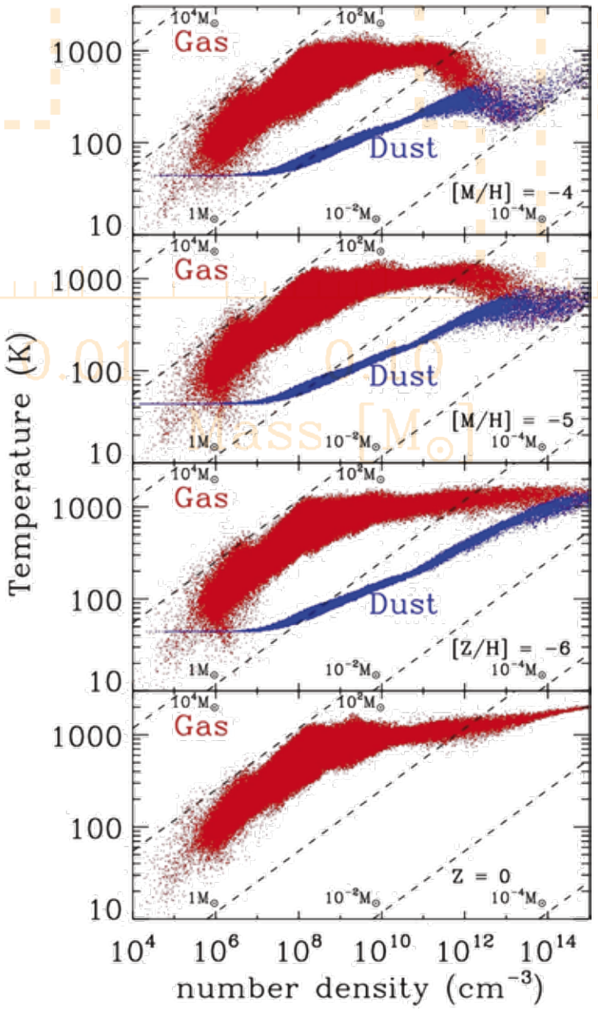


Figure 1: Results of our simulations, showing the dependence of gas and dust temperatures on gas density for metallicities 10^{-4}, 10^{-4}, and 10^{-6} and zero times the solar value. In red, we show the gas temperature, and in blue the dust temperature for the turbulent and rotating cloud. The dashed lines show constant Jeans mass values.

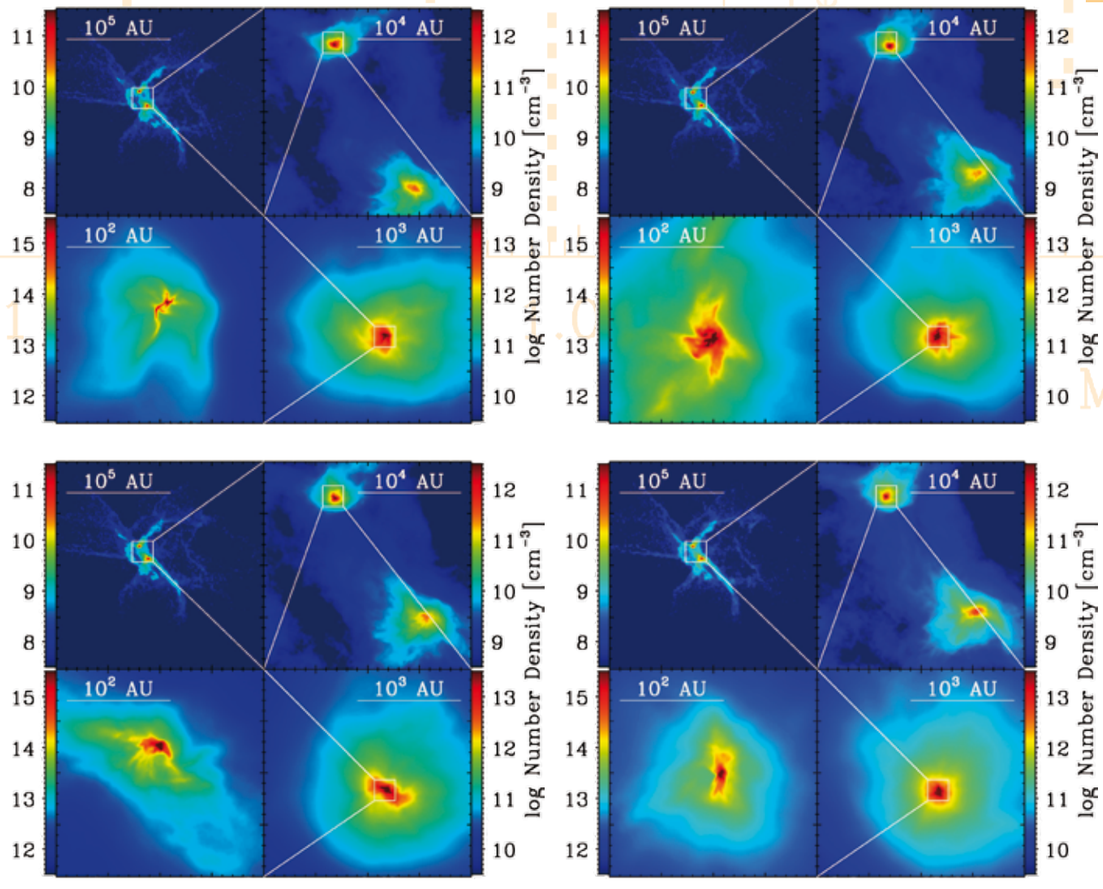
Determining Z_{crit} is important because it influences the current theory of when and how the universe was reionized at around 12 billion years ago (Gunn & Peterson, 1965), and also what kind of stars we expect to observe today (Komiya, 2011). In order to determine Z_{crit}, many physical processes should be taken into account for the star formation process, such as: hydrodynamic evolution of the collapsing gas, heating and cooling mechanisms, and chemical reactions. In the early universe, the most important cooling process at the point where the gas undergo fragmentation, and forms stars, is dust cooling (Omukai et al., 2005).

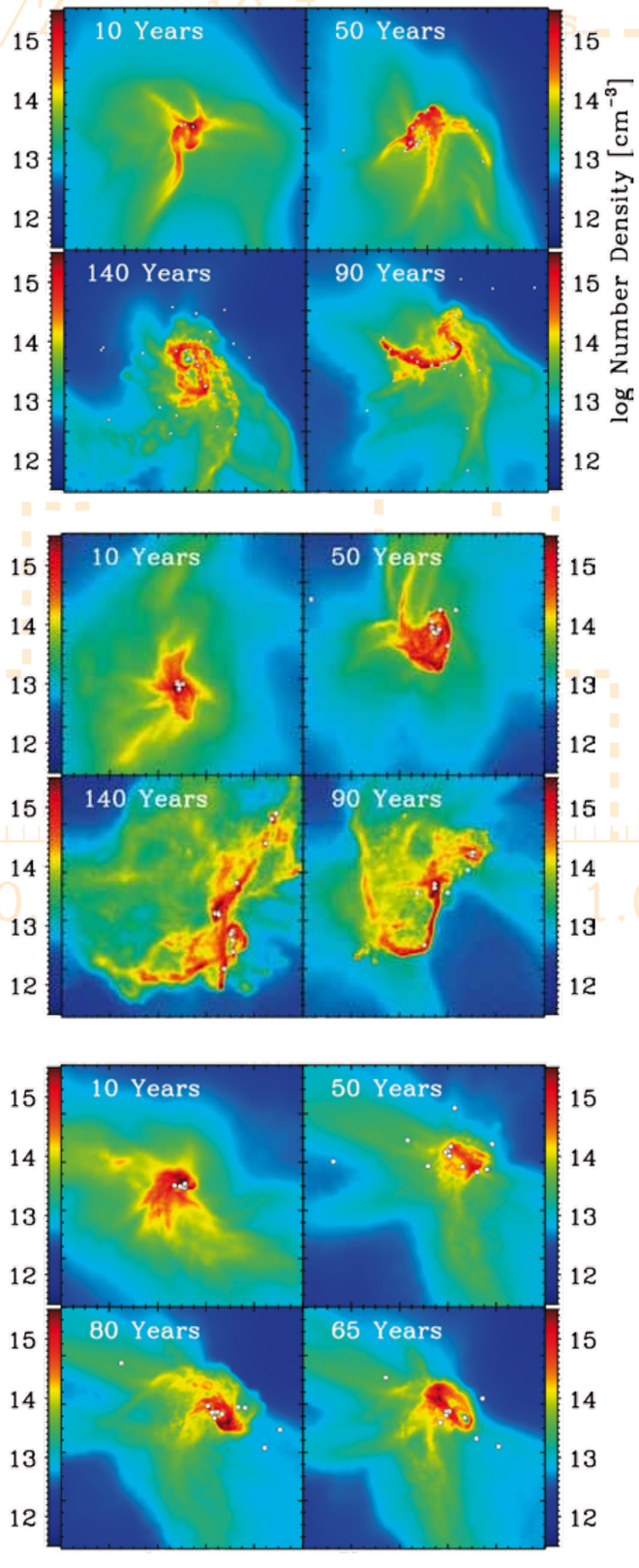
The conditions for fragmenting the gas into small protostar like objects are predicted to occur in high density gas, where the distances between the fragments can be very small (Omukai et al., 2005; Schneider & Omukai, 2010). In this regime, interactions between fragments will be common, and analytic models of fragmentation are unable to predict the mass distribution of the fragments. A full 3D treatment, following the fragments, is needed.

3D Treatment, following the Fragments, is needed.

In this work, we calculate the evolution of star-forming clouds for the metallicities (10^{-4}, 10^{-5}, $10^{-6}Z_\odot$, and 0), and observe the effect on the IMF (2). By comparing the IMF for different metallicities with the metal-free case, we determine the Z_{crit} for the transition from Pop. III to Pop. II. With this results, we also determine how different star formation properties vary in very low metallicity clouds.

Figure 2: Number density maps for a slice through the high density region for $Z=10^{-4}Z_\odot$ (top), $10^{-5}Z_\odot$, $10^{-6}Z_\odot$, and Z=0 (bottom). The image shows a sequence of zooms in the density structure in the gas immediately before the formation of the first protostar.





Simulations

We model the collapse of a low-metallicity gas cloud using a modified version of the Gadget 2 (Springel, 2005) smoothed particle hydrodynamics (SPH) code. To enable us to continue our simulation beyond the formation of the first very high density protostellar core, we use a sink particle approach (Bate et al., 1995; Jappsen et al., 2005), in the same way as in Dopcke et al. (2011).

Sink particles are created once the SPH particles are bound, collapsing, and within an accretion radius, hacc, which is taken to be 1.0 AU. The threshold number density for sink particle creation is 5.0×10^{15}cm^{-3}. At the threshold density, the Jeans length, or characteristic fragmentation length, at the minimum temperature reached by the gas is approximately 1AU, while at higher densities the gas becomes optically thick and begins to heat up. Further fragmentation on scales smaller than the sink particle scale is therefore unlikely to occur. For further discussion see Clark et al. (2011a).

Figure 3: Number density map showing a slice in the densest clump, and the sink formation time evolution, for the 40 million particles simulation, and Z = $10^{-4}Z_{\odot}$(top), $10^{-5}Z_{\odot}$(middle), $10^{-6}Z_{\odot}$(bottom). The box is 100AU x 100AU and the time is measured from the formation of the first sink particle.

We follow the thermodynamical evolution of the gas up to very high densities of order 10^{17}cm^{-3}, where the Jeans mass is $\approx 10^{-2}M_{\odot}$, and so we need a high resolution simulation to study the fragmentation behaviour.

Collisions between gas particles and dust grains can transfer energy from the gas to the dust (if the gas temperature T is greater than the dust temperature T_{gr}), or from the dust to the gas (if $T_{gr} > T$). The details of the dust cooling treatment can be seen in Dopcke et al. (2011).

We performed four simulations, varying the metallicity in $\log(Z/Z_{\odot}) = -4, -5, -6$ and the metal-free case. The simulations had 40 million SPH particles. We used these simulations to model the collapse of an initially uniform gas cloud with a number density of 10^5 particles/cm^3, temperature of 300 K and the cloud mass of 1000 M. We included small amounts of turbulent and rotational energy, with $E_{turb}/|E_{grav}| = 0.1$ and $\beta = E_{rot}/|E_{grav}| = 0.02$, where E_{grav} is the gravitational potential energy, E_{turb} is the turbulent kinetic energy, and E_{rot} is the rotational energy. The mass resolution (taken to be 100 times the SPH particle mass) was 2.5×10^{-3}M.

Computational Challenge

We follow the thermodynamical evolution of the gas up to very high densities of order 10^{17}cm^{-3}, where the Jeans mass is $\approx 10^{-2}M_{\odot}$, and so we need a high resolution simulation to study the fragmentation behaviour. The transport of angular momentum to smaller scales during the collapse leads to the formation of a dense disc-like structure, supported by rotation which then fragments into several objects. To access the hydrodynamic properties of a 1000 $M_{\odot}$ gas cloud, we need to use a minimum of 10 million SPH particles.

Each of the 40 million SPH particles simulations described above took approximately 150 k CPU hours. By using 64 CPUs, in HPC-GPU Cluster Kolob (University of Heidelberg), it took two months per simulation. This was considerably improved by using HLRB II at the Leibniz Supercomputing Centre (LRZ), and now this time was reduced to less then 15 days. In order to test more variables and cover the parameter space, we plan to run 16 more simulations, resulting in 2.6 Mio. CPU-h. For resolution studies and post-processing, there is an additional overall cost of approximately 0.3 Mio. CPU-h. In total, this project requires approximately 2.9 Mio. CPU-h.

The problem of star formation in the early universe remains not completely solved mainly because of the complexity and variety of the physical processes involved, but also because the computers we have available now do not have the power to calculate the physical processes from the big to the small scales, which can not be consistently separated in different calculations. In the ideal case, we would like to calculate the evolution of a gas cloud that fall in the first dark matter haloes.

This calculation would involve resolving the evolution from 10^6 $M_{\odot}$ to $10^{-2}M_{\odot}$, being necessary 30 Mio. CPU-h per calculation, this is within reach of the next generation of high performance computers.

Scientific Results

In Figure 1, we compare the evolution of the dust and gas temperatures in the simulations. The dust temperature, shown in the lower part of the panels, varies from the CMB temperature in the low density region to the gas temperature at much higher densities.

The transport of angular momentum to smaller scales during the collapse leads to the formation of a dense disc-like structure, supported by rotation which then fragments into several objects. Figure 2 shows the density structure in the gas immediately before the formation of the first protostar. The top panel shows a density slice on a scale comparable to the size of the initial gas distribution. The structure is very filamentary and there are two main dense clumps in the centre. If we zoom in on one of the clumps, we see that its internal structure is also filamentary. Once the conditions for sink particle creation are met, they start to form in the highest density regions (Figure 3).

Figure 4 shows the mass distribution of sink particles when we stop the calculation. We typically find masses below $1 M_\odot$, with somewhat smaller values in the $10^{-4} Z_\odot$ case compared to the $10^{-5} Z_\odot$ case. This suggests that the transition from high-mass primordial stars to Population II stars with mass function similar to that at the present day occurs early in the metal evolution history of the universe, at metallicities $Z_{crit} < 10^{-5} Z_\odot$.

This suggests that the transition from high-mass primordial stars to Population II stars with mass function similar to that at the present day occurs early in the metal evolution history of the universe, at metallicities $Z_{crit} < 10^{-5} Z_\odot$.

Acknowledgments

The present work is supported by the Baden-Württemberg Stiftung in the program Internationale Spitzenforschung via contract P-LS-SPII/18, the German Bundesministerium für Bildung und Forschung via the ASTRONET project STAR FORMAT (grant 05A09VHA), a Frontier grant of Heidelberg University sponsored by the German Excellence Initiative, the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD). All computations described here were performed at the Leibniz Supercomputing Centre, National Supercomputer HLRB II), Jülich Supercomputing Centre - Jülich Research on Petaflop Architectures, and on the HPC-GPU Cluster Kolob (University of Heidelberg).

References

[1] In astronomy, metallicity is the proportion of matter made up of chemical elements heavier than helium for a certain object

[2] IMF is the Initial Mass Function or the distribution of mass (the histogram of stellar masses) for new-born stars

[3] **Bate, M. R., Bonnell, I. A., Price, N. M.**
1995, MNRAS, 277, 362

[4] **Clark, P. C., Glover, S. C. O., Klessen, R. S.**
2008, ApJ, 672, 757

[5] **Clark, P. C., Glover, S. C. O., Klessen, R. S., Bromm, V.**
2011a, ApJ, 727, 110

[6] **Clark, P. C., Glover, S. C. O., Smith, R. J., Greif, T. H., Klessen, R. S., Bromm, V.**
2011, Science, 331, 1040

[7] **Dopcke, G., Glover, S. C. O., Clark, P. C., Klessen, R. S.**
2011, ApJ, 729, L3

[8] **Gunn, J. E., Peterson, B. A.**
1965, ApJ, 142, 1633

[9] **Jappsen, A., Klessen, R. S., Larson, R. B., Li, Y., Mac Low, M.**
2005, A&A, 435, 611

[10] **Komiya, Y.**
2011, ApJ, 736, 73

[11] **Kroupa, P.**
2002, Science, 295, 82

[12] **Omukai, K., Hosokawa, T., Yoshida, N.**
2010, ApJ, 722, 1793

[13] **Omukai, K., Tsuribe, T., Schneider, R., Ferrara, A.**
2005, ApJ, 626, 627

[14] **Schneider, R., Ferrara, A., Natarajan, P., Omukai, K.**
2002, ApJ, 571, 30

[15] **Schneider, R., Omukai, K.**
2010, MNRAS, 402, 429

[16] **Schneider, R., Omukai, K., Inoue, A. K., Ferrara, A.**
2006, MNRAS, 369, 1437

[17] **Springel, V.**
2005, MNRAS, 364, 1105

[18] **Tsuribe, T., Omukai, K.**
2006, ApJ, 642, L61

• Gustavo Dopcke
• Simon C. O. Glover
• Paul C. Clark
• Ralf S. Klessen

Institut für Theoretische Astrophysik
Universität Heidelberg

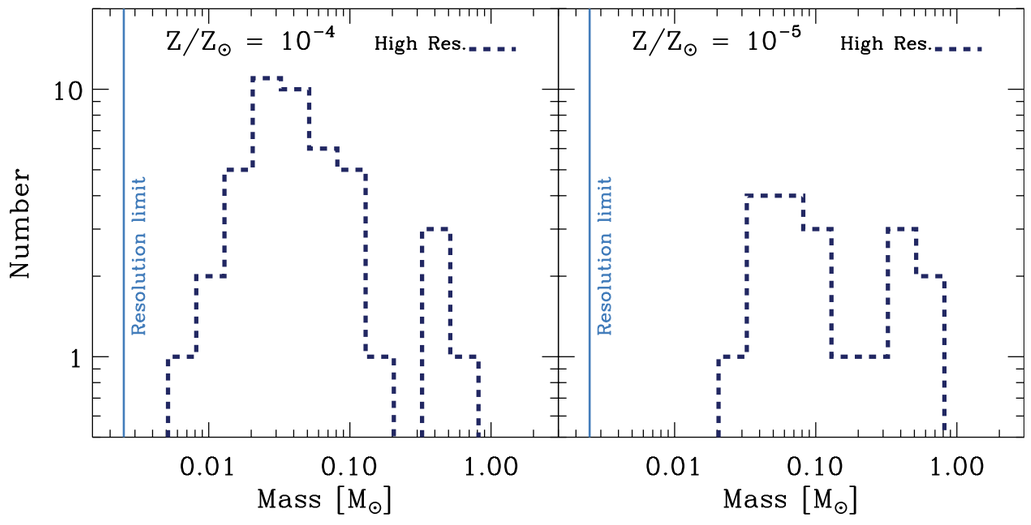


Figure 4: Sink particle mass function at the end of the simulations.High and low resolution results and corresponding resolution limits are shown. To resolve the fragmentation, the mass resolution should be smaller than the Jeans mass at the point in the temperature-density diagram where dust and gas couple and the compressional heating starts to dominate over the dust cooling. At the time shown, around 5 $M_\odot$ of gas had been accreted by the sink particles in each simulation.

Theoretical Treatment of magnetic Molecules

Applications

Molecular Magnetism is a field of condensed matter physics where the properties of magnetic molecules are investigated [1]. Although it has been known for more than 50 years that molecules can possess magnetic moments, the whole field got an enormous impetus when scientists in the late 1980s discovered that a single molecule can show a magnetic hysteresis loop.

The underlying (classical) picture is this: The spins of the magnetic ions contained in molecules such as Mn_{12}-acetate (see Fig. 1) form a resulting molecular magnetic moment. At low temperatures this moment is rather large (S=10 for Mn_{12}-acetate). Its motion is governed by an anisotropy barrier which means that there are two directions (up and down) in which the magnetic moment likes to point and that it would need to overcome an energy barrier in between. This stabilizes the up and down orientations and leads to the magnetic hysteresis. Since this is exactly the behavior of a magnetic entity storing a single bit, magnetic molecules appeared to be potential ultimately small next-generation storage devices. But since they are so small, quantum effects are important. In essence the latter result in the fact that the magnetic bistability vanishes at temperatures exceeding one or two Kelvins. Therefore, since the discovery of Mn_{12}-acetate chemists try to synthesize molecules with better barriers and physicists together with chemists try to rationalize the magnetic behavior in order to provide guidance for future synthesis.

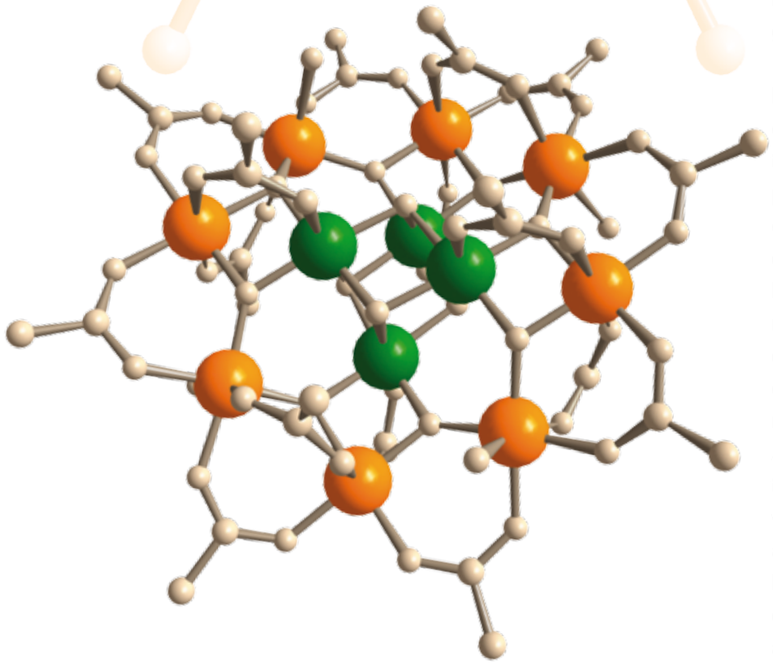


Figure 1: Sketch of the magnetic molecule Mn12-acetate. The manganese ions are highlighted in color.

Nowadays the range of research questions has broadened. It includes the investigation of magnetocaloric properties of such substances, i.e. the question whether and how efficient one could cool or heat with these materials by just sweeping the field. Efforts of maintaining and manipulating quantum coherent states are being made in order to construct constituents which might be parts of quantum computers. A last but not least focus is on effects that are driven by magnetic frustration, which labels a situation in which interacting magnetic moments are “not happy” since they cannot align parallel or antiparallel due to competing interactions.

In any case it is the task of theory to model such systems with the aim to understand the magnetic behavior. From the point of view of many-body theory a molecule is already a giant object since it consists of very many nuclei and electrons, therefore complete quantum solutions are not even within reach. But for many observables it is quite sufficient to consider the magnetic moments of the unpaired electrons together with their mutual interactions only. Such models are mathematically represented by spin Hamiltonians of which the Heisenberg Hamiltonian is one of the most famous. It has a rather simple form:

$$H = -2\sum_{k<l} J_{kl}\vec{s}_k \bullet \vec{s}_l$$

The Hamilton operator represents the energy that stems from the interactions of pairs (k,l) of spins, where J_{kl} represents the strength of the interaction. There are only two basic interactions: the ferromagnetic interaction which tends to align the spin vectors in a parallel fashion and the antiferromagnetic interaction which tends to align the spin vectors in an antiparallel fashion. Although this classical description sounds very simple the quantum problem is numerically involved since in quantum mechanics the Hamiltonian is represented by a matrix in Hilbert space whose eigenvalues we need to compute. The dimension of this vector space grows exponentially with the number of spins, i.e. $d=(2s+1)^N$ for N equal spins. This means that for e.g. 10 iron spins of s=5/2 the dimension is 60,466,176. Such a matrix, when stored completely, would need about 30 Petabyte of RAM; the determination of its eigenvalues is virtually impossible.

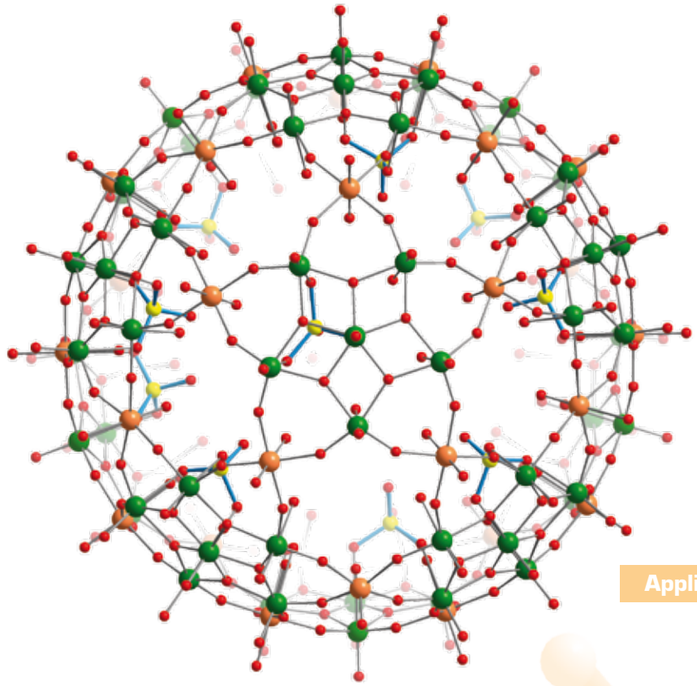


Figure 2: Sketch of the magnetic molecule $\{W_{72}V_{30}\}$; the vanadium ions are displayed as orange spheres [6].

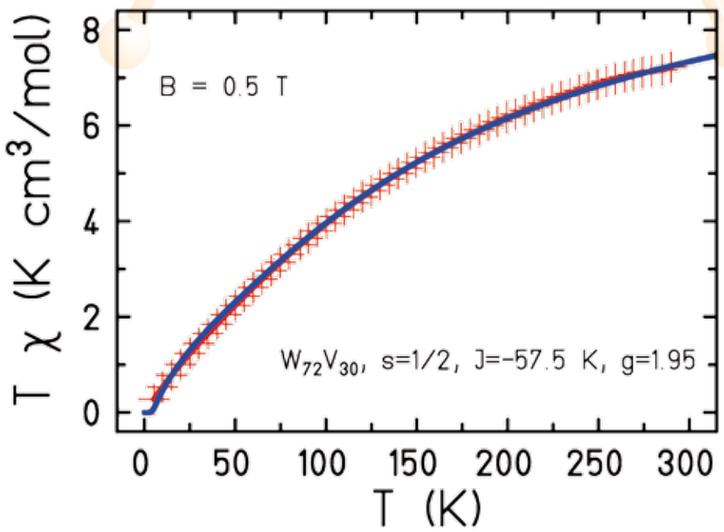
Applications

There are two approaches to ease this problem: use of symmetries and Krylow space methods. Since this article is going to focus on Krylow space methods, the first approach shall be only shortly explained. Using symmetries is one of the major paradigms in theoretical physics (and quantum chemistry). Besides its mathematical beauty it provides additional insight into physical properties as well as massive help for numerical solutions. The procedure is the following: If an operator such as the Heisenberg spin Hamiltonian above is quantum mechanically represented in Hilbert space it is given by a matrix with as many rows and columns as the dimension of the space. If a symmetry can be used then the Hilbert space can be decomposed into mutually orthogonal subspaces and the matrix needs to be evaluated only in each subspace.

Let's consider a mirror symmetry for instance. Then the Hilbert space can be split up into approximate halfs, and in each the matrix has only half of the linear dimension and needs only a quarter of RAM to be stored. The matrices can be treated one after another, which is the numerical gain. It is clear that one aims at the use of as many as possible symmetries [2].

A second approach to the eigenvalue problem is given by Krylow space methods [3] which by the way can be combined with the use of symmetries. The original idea of such methods is to obtain only extremal eigenvalues as for instance the lowest eigenvalue of a Hamiltonian, which is the ground state energy, and a few above the lowest which suffice to explain the low-lying excitation spectrum. These methods work by constructing a special vector space (Krylow space) that is only a tiny subspace of the full Hilbert space, but in this special subspace extremal eigenvalues converge rapidly against the true extremal eigenvalues. The Krylow space is generated by applying the Hamiltonian and its powers up to typically about 100 to a (randomly chosen) trial vector. This is a rather cheap operation compared to matrix diagonalisation, and it can be well parallelized [4].

Figure 3: Magnetic susceptibility of $\{W_{72}V_{30}\}$; experimental data are given by crosses, the fit is given by the blue curve.



It turns out that Krylow space methods can as well be used in order to accurately approximate quantum statistical observables, i.e. in order to evaluate the partition function with very high accuracy [5]. To this end not only the extremal eigenvalues in the respective Krylow spaces are used but all eigenvalues and eigenvectors which in essence then provide a coarse grained density of states that allows to approximate the thermodynamic functions. This enables us to model magnetic molecules of unprecedented size, i.e. with Hilbert space dimensions of up to 10^{10} [7]. A very recent example is given by the molecule $\{W_{72}V_{30}\}$ in which the 30 magnetic vanadium ions (s=1/2) have to be modeled in a Hilbert space of dimension 1,073,741,824. Despite this dimension, its magnetic properties as e.g. the susceptibility can be evaluated within a few days e.g. on 510 cores of the SGI Altix at the LRZ in Garching. This opens the prospect that recently synthesized magnetic molecules can be modeled with high accuracy in a reasonable time. We believe that the whole field of molecular magnetism will benefit from these new methods as well as their realizations on modern supercomputers.

The success of Krylow space methods partly rests on the fact that they can be very efficiently parallelized using openMP directives. Many of the necessary operations are simple matrix vector products. The performance can be even further enhanced if matrix elements do not need to be stored but can be evaluated when needed, i.e. "on the fly" [7]. We tested the scaling on various architectures, it seems to be very close to perfect. Beside tests on the SGI Altix at LRZ we run tests on ScaleMP vSMP architectures that consists of Infiniband connected commodity x86-based clusters at Bielefeld University and at RWTH Aachen, see also [8]. These clusters are turned into virtual SMP machines by means of the ScaleMP vSMP software.

The author and his group would like to thank the LRZ Garching for very valuable computing time and support. Thanks are also to Ana Maria Todea and Joris van Slageren for providing the high-resolution pictures of the magnetic molecules {W72V30} and Mn12-acetate.

References

[1] **www.molmag.de**
This is a gate to molecular magnetism. It shows research highlights, provides tutorials and lists scientists who are active in the field.

[2] **Schnalle, R., Schnack, J.**
Calculating the energy spectra of magnetic molecules: application of real- and spin-space symmetries, Int. Rev. Phys. Chem. 29 (2010) 403

[3] **Lanczos, C.**
An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Stand. 45 (1950) 255-282

[4] **Weiße, A., Wellein, G., Alvermann, A., Fehske, H.**
The kernel polynomial method, Rev. Mod. Phys. 78 (2006) 275-306

[5] **Schnack, J., Hage, P., Schmidt, H.-J.**
Efficient implementation of the Lanczos method for magnetic systems, J. Comput. Phys. 227 (2008) 4512

[6] **Jaklic, J., Prelovsek, P.**
Lanczos method for the calculation of finite-temperature quantities in correlated systems, Phys. Rev. B 49 (1994) 5065-5068

[6] **Todea, A. M., Merca, A., Bögge, H., Glaser, T., Engelhardt, L., Prozorov, R., Luban, M., Müller, A.**
Polyoxotungstates now also with pentagonal units: supramolecular chemistry and tuning of magnetic exchange in {(M)M5}12V30 Keplerates (M = Mo, W), Chem. Commun. (2009) 3351

[7] **Schnack, J., Wendland, O.**
Properties of highly frustrated magnetic molecules studied by the finite-temperature Lanczos method, Eur. Phys. J. B 78 (2010) 535

[8] **Schmidl, D., Terboven, C., Wolf, A., Mey, D., Bischof, C.**
How to Scale Nested OpenMP Applications on the ScaleMP vSMP Architecture, 2010 IEEE International Conference on Cluster Computing (CLUSTER), 2010, pp. 29-37

• Jürgen Schnack

Universität Bielefeld, Fakultät für Physik

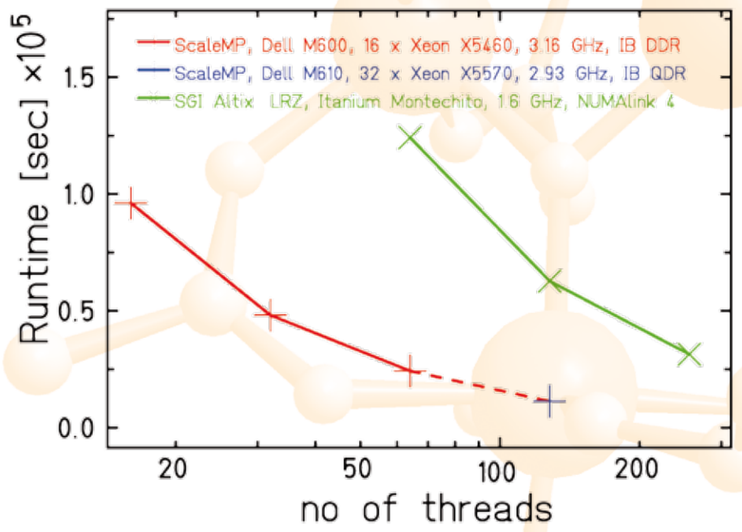


Figure 4: Runtime of a typical Lanczos run on various architectures. The scaling is very close to perfect. The higher speed of the x86-based clusters reflects its later production date (2010 vs. 2006 for the Itanium II).

The gravity Field - An important Parameter for Earth Observation

As the mass density in the Earth's interior is varying and the shape of the Earth is non-uniform, gravity is differing at each observation point. This means, that the force of attraction acting on a test mass varies between the poles and the equator, mountains and valleys, or ocean and land areas. The challenge is to find a global representation of the gravity field of the Earth, as its knowledge is of primary interest for many applications and scientific disciplines in Earth sciences.

For instance, the gravity field is directly linked to the physical shape of the Earth. This physical shape does not correspond to a sphere or an ellipsoid, but to the so-called geoid, an irregular surface which is illustrated in Figure 1. Over the oceans the geoid coincides approximately with the mean sea surface assuming that only gravity and no other external forces are acting on it. Over the continents it can be regarded as a continuation of the mean sea surface below the topography. The gravity potential is constant on the geoid, which means that there is no water flowing between points located on the geoid. Therefore, it is appropriate to use the geoid as general reference and zero level for height systems. The deviation of the geoid from a rotational ellipsoid, which is used as best approximating geometric reference body for the Earth, is between -100 and 80 m. This quantity is called geoid height.

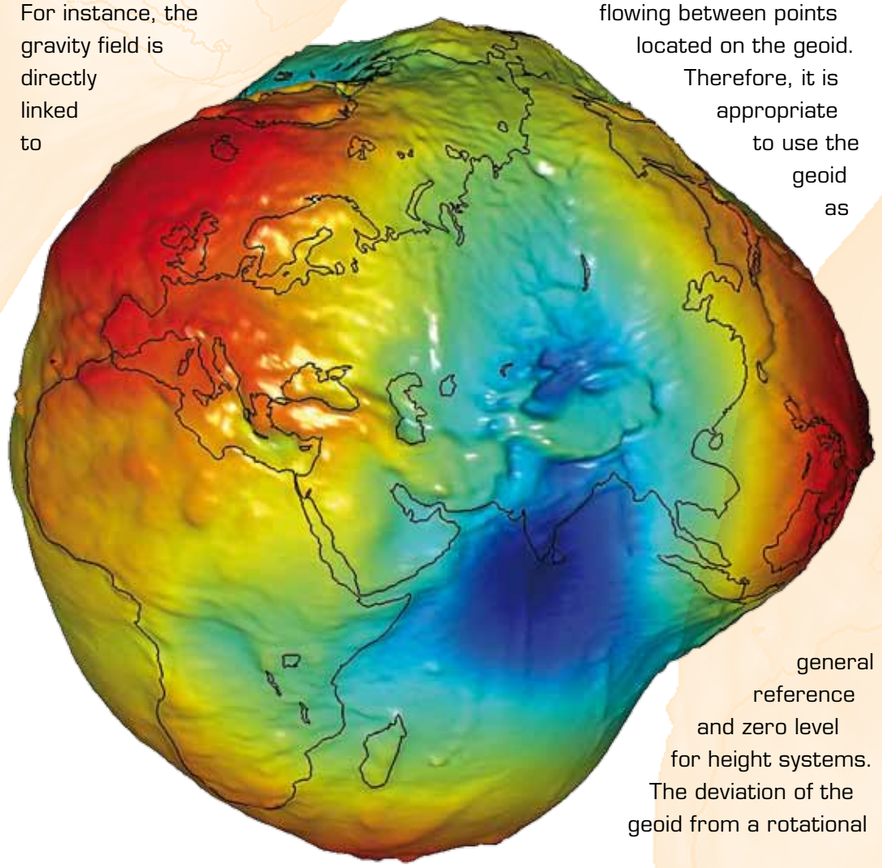


Figure 1: The geoid, the physical shape of the earth

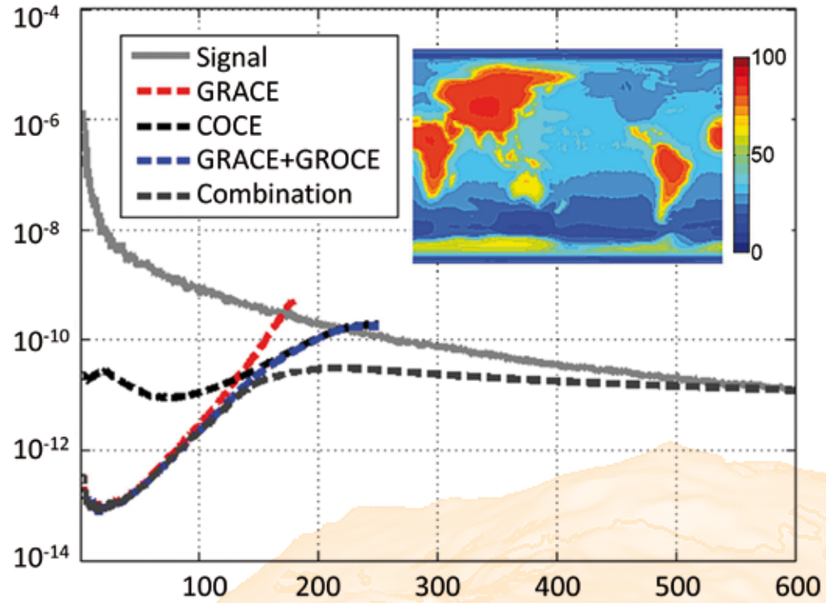


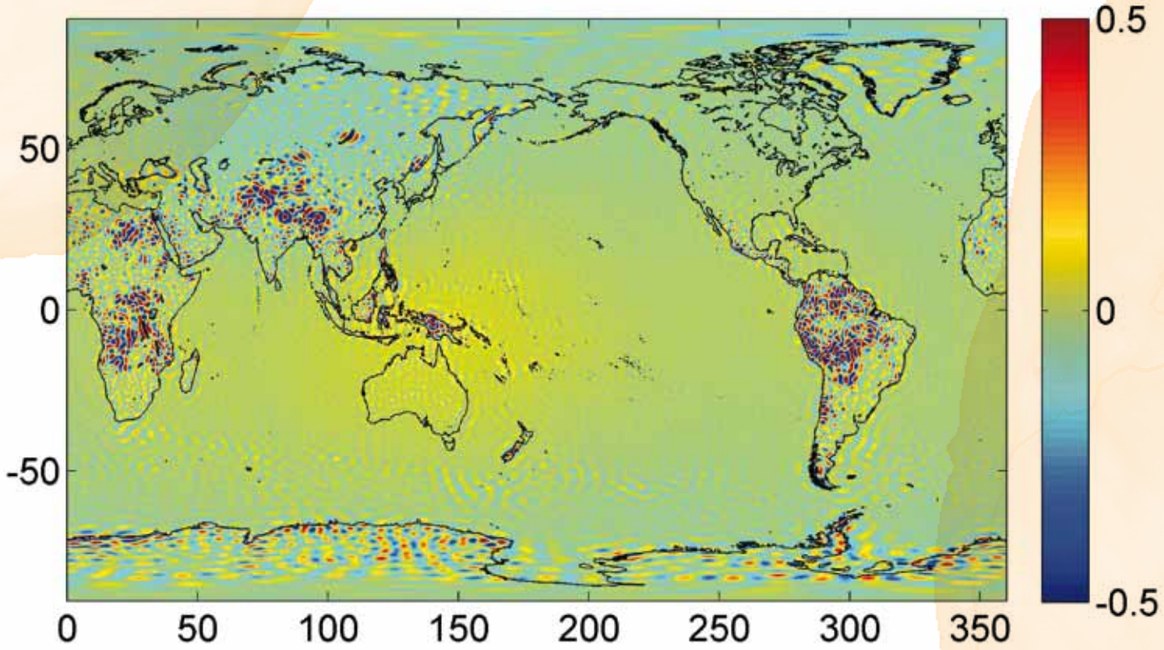
Figure 2: Degree median of the gravity field signal and the errors of the different observation methods and percental quota of satellite data to the combination solution at spherical harmonic degree 170

Also in oceanography the geoid serves as important reference surface for modelling of sea level change and ocean currents. The deviation of the ocean surface from the geoid is called dynamic ocean topography, and is one of the main drivers of ocean currents [1]. As warm and cold water is transported by these currents, the precise knowledge of the geoid is of great importance for studying climate change. Beside its importance as reference surfaces the gravity field provides insight into the Earth's interior. Geophysicists gain information about the distribution of masses and density inside the Earth. The Earth's gravity field is also sensitive to redistribution of masses due to continental water flows (hydrology), melting of continental ice sheets, or even large earthquakes. The loss of ice masses in Greenland and Antarctica as well as hydrological events like rainy seasons in the Amazon can be detected by observing the gravity field from space. Due to its importance for different scientific disciplines, the gravity field has to be determined as accurately as possible.

Since spherical harmonics, a generalization of the concept of Fourier series for the sphere, are global base functions, they are well suited to describe the Earth's gravity field. Gravity field observations on ground, from airplanes or from satellites are used to determine the spherical harmonic coefficients by solving an inverse problem. There exists a multitude of methods to observe the gravity field. First, terrestrial or airborne gravity accelerations are observed with absolute or relative gravimeters. The gravity at a certain point can be determined through free fall experiments (absolute gravimeter) or by instruments measuring the tension of springs (relative gravimeter). The advantage of this method is that the full gravity signal content is

observed, because the observations are taken on the Earth surface or close to it. However, each measurement campaign is restricted to a local area, whereas a global coverage of observations is essential for describing the global field. To achieve global coverage, the data of multiple campaigns must be collected, which is rather challenging, because in some Asian, African or South American areas only few and low-quality data sets exist. Often the access to these data is restricted. The result is an inconsistent data set with inhomogeneous coverage, which is not well suited for determining the long wavelengths of the gravity field. The second possibility to obtain gravity field observations is satellite altimetry. Altimetry is a method, which primarily determines the height of the ocean above the mean Earth ellipsoid. By subtracting the dynamic ocean topography (see above), geoid heights can be computed ocean-wide. With a further post-processing also altimetric gravity can be computed. The advantage of this procedure is that the signal content is not restricted and the dataset is consistent. Altimetric and terrestrial gravity measurements are generally combined to a more or less global set of observations. As a third method, satellites can be used to observe the Earth's gravity field. The attraction of the Earth acting on a satellite causes orbit perturbations. While in the past orbits of arbitrary satellites were analyzed, nowadays data from dedicated gravity field satellite missions exist. The first mission was CHAMP (launched in 2000), where the satellite orbit was tracked by GPS. The second mission GRACE (launched in 2002) consists of two satellites following each other on the same orbital track, and observes distance changes between them. These distance changes are mainly caused by different gravitational attractions acting on the two satellites (separation distance about 220 km). The third and most recent mission GOCE (launched in 2009) is equipped with a space gradiometer, which enables the observation of gravity gradients along a baseline of 50 cm with very high performance. In general, data from these satellite missions are consistent, almost global (depending on the chosen orbit) and of very good and homogeneous quality in the low to medium frequency part (depending on the mission concept). However, the signal content of satellite observations is restricted due to the fact, that according to Newton's law the gravity signal attenuates with increasing satellite height.

Figure 3: Improvement of the geoid due to the GOCE mission [m]



A bundle of FORTRAN90 and Matlab programs handles the pre-processing of different data sets with more than 200 million observations in case of terrestrial/altimetric observations, resulting in a single data file which is input to the main processing chain. The way from these observations to the gravity field coefficients is computationally challenging. As spherical harmonic coefficients are obtained by least-squares adjustment, the corresponding normal equation system can be very large. Up to now it was standard to estimate coefficients based on full normal equations up to spherical harmonic degree 360, which corresponds to a matrix size of 126 GByte. The coefficients of higher degrees were generally obtained by using block-diagonal techniques. At LRZ's HLRB II an approach was implemented dealing with full normal equations up to significantly higher degrees. This is useful in order to preserve the full variance-covariance information of the normal equation systems. Calculations on the HLRB II were performed up to spherical harmonic degree 600, corresponding to a normal equation size of 972 GByte [2]. The main work is done by two FORTRAN90 programs using BLACS, MPI and SCALAPACK. The first one is organized in a one-dimensional process grid, where each process assembles a certain part of the normal equation matrix (for all observations). Each of these parts handles a certain group of spherical harmonic coefficients belonging together. The second program is organized in a two-dimensional grid. The normal equation is block-cyclic distributed, and the program solves for the spherical harmonic coefficients and inverts the normal equation system for computing error estimates of the coefficients. Further routines were implemented, which handle observations of the GOCE gradiometer. The processing chain is different compared to the case of terrestrial data. Six gravity gradient observations and one 3D-position are measured per second since the start of the mission in 2009. The pre-processing requires data editing and filtering in order to cope with correlated noise. Solving of the normal equations for GOCE data is not as demanding as for terrestrial data, because, due to the satellite height, the signal is damped and the normal equation systems are restricted to degree and order 250 (corresponding to about 30 GByte normal equation file size).

The best gravity field solution is obtained by combination of all possible data sources. Figure 2 shows the gravity field signal per spherical harmonic degree (grey line) and the corresponding errors of different gravity field solutions. It can be identified, that the GRACE mission delivers the best result in the very low degrees, whereas GOCE is performing better in the low to medium degrees. Both missions together provide an excellent satellite-only gravity field solution [3]. Furthermore it is shown, that only by

including terrestrial/altimeter data a large spherical harmonic expansion can be obtained. The percental contribution of the satellite data to the combination solution is displayed in the upper right of the figure. The quota depends on the relative weights for the observation groups. In areas of poor terrestrial data quality the weight for the satellite data is higher.

where the quality of terrestrial data is poor, such as South America, Africa and the Himalaya region.

But improved gravity solution based on GOCE result also in large improvements of dynamic ocean topography estimates, because they represent a more accurate geoid solution and thus

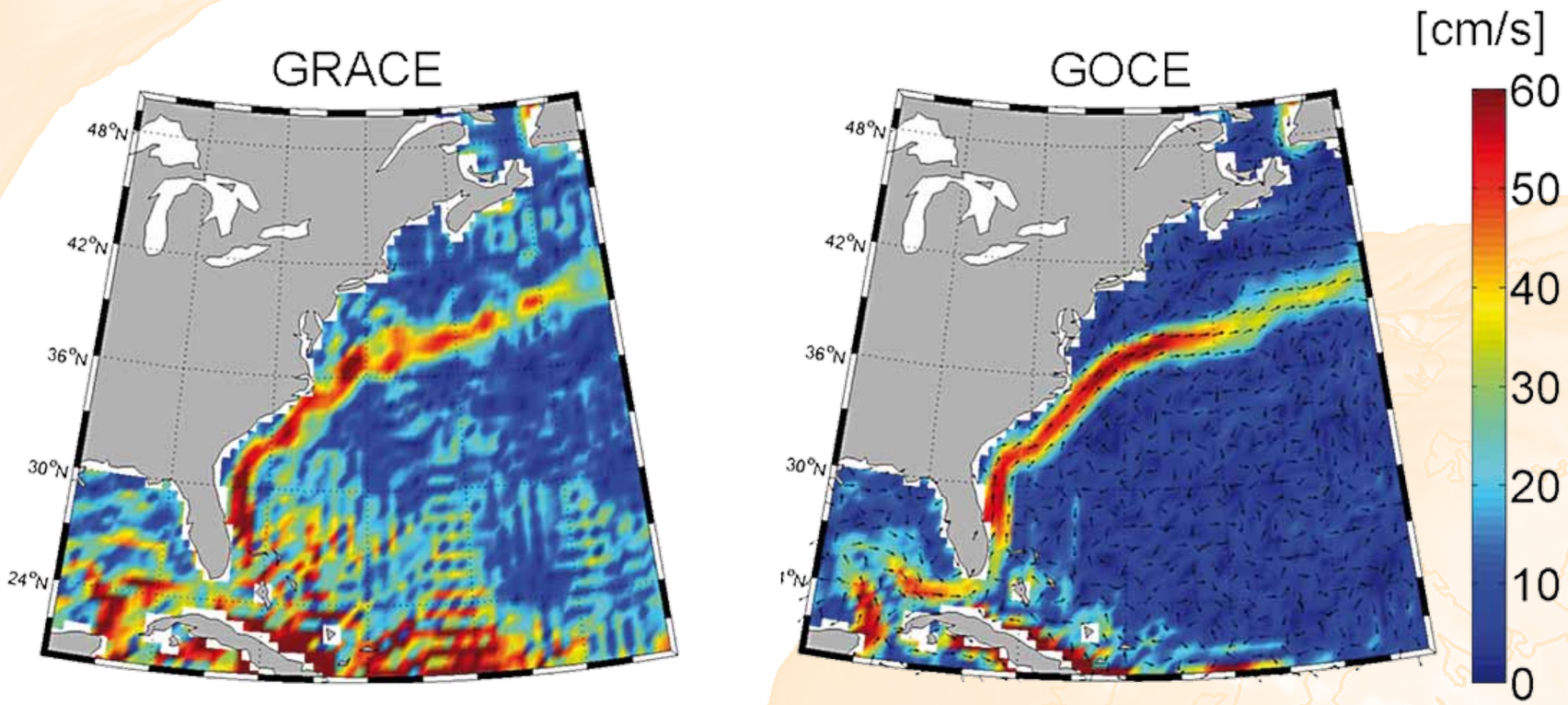


Figure 4:
Improvement in modelling of the Gulf current due to the GOCE mission

There is also a large impact of GOCE to combined gravity field solutions (due to the fact that the GOCE mission is the newest data source with the highest spatial resolution). Figure 3 illustrates the improvements of the gravity field due to the inclusion of GOCE [4]. Shown is the difference between geoid heights of a global gravity field with and without GOCE data. New and so far unknown gravity field signal detected by GOCE can be especially seen in areas

a more accurate reference surface. This was proven by an oceanographic application. The Gulf current was modelled with and without GOCE data (Fig. 4). By including GOCE data the noise is reduced significantly, and more detailed structures of the ocean current become visible.

These results show that the gravity field is an important parameter for various Earth related disciplines.

As millions of data have to be analyzed and hundred of thousands of parameters have to be estimated, high performance computer systems are an essential component to achieve such global gravity field models.

Acknowledgement

The authors gratefully acknowledge computational resources granted by LRZ. Special thanks for all the help and patience concerning our questions and incidents.

References

[1] **Bingham, R. J., Knudsen, P., Anderson, O., Pail, R.**
An initial estimate of the North Atlantic steady-state geostrophic circulation from GOCE, Vol. 38, American Geophysical Union, DOI: 10.1029/2010GL045633, 2011

[2] **Fecher, T., Pail, R., Gruber, T.**
Global gravity field determination from terrestrial data, Amercian Geophysical Union Fall Meeting, San Francisco, 2010

[3] **Pail, R., Goiginger, H., Schuh, W. D., Höck, E., Brockmann, J. M., Fecher, T., Gruber, T., Mayer-Gürr, T., Kusche, J., Jäggi, A., Rieser, D.**
Combined satellite gravity field model GOCO01S derived from GOCE and GRACE, Geophysical Research Letters, Vol 37, American Geophysical Union, DOI: 10.1029/2010GL044906, 2010

[4] **Pail, R., Bruinsma, S., Migliaccio, F., Förste, C., Goiginger, H., Schuh, W. D., Höck, E., Reguzzoni, M., Brockmann, J. M., Abrikosov, O., Veicherts, M., Fecher, T., Mayrhofer, R., Krasbutter, I., Sansò, F., Tscherning, C. C.**
First goce gravity field models derived by three different approaches, Journal of Geodesy, DOI: 10.1007/s00190-011-0467-x

• Thomas Fecher
• Thomas Gruber
• Roland Pail

Institut für Astronomische und Physikalische Geodäsie, TU München

Aircraft Wake Vortex Evolution in Ground Proximity and at Cruise Altitude

As an unavoidable consequence of lift aircraft generate a pair of counter-rotating and long-lived wake vortices that pose a potential risk to following aircraft. The therefore prescribed aircraft separations contribute significantly to capacity restrictions of large airports. But also during cruise severe encounters of wake vortices have been reported. Wake vortex behaviour is largely controlled by the prevailing meteorological conditions and the interaction with the ground. The most important meteorological parameters are wind, wind shear, turbulence, and temperature stratification.

The Deutsche Zentrum für Luft- und Raumfahrt (DLR) develops wake vortex advisory systems for airports [1,2] and en route which aim at optimizing the air traffic with respect to the measured and predicted wake vortex behaviour. As part of such systems simple probabilistic wake vortex prediction models are required that predict wake vortex behaviour accurately, robust, and fast [3]. Highly resolving Large Eddy Simulations (LES) conducted on the HLRB II supercomputer provide valuable insights in the physics of wake vortex behaviour under various environmental conditions. These LES contribute indispensable guidance for the development of the simple wake vortex models.

Figure 1: Wake vortex evolution with crosswind in ground proximity. Green vorticity structures representing the crosswind turbulence are being wrapped around the wake vortices (brown vorticity tubes).

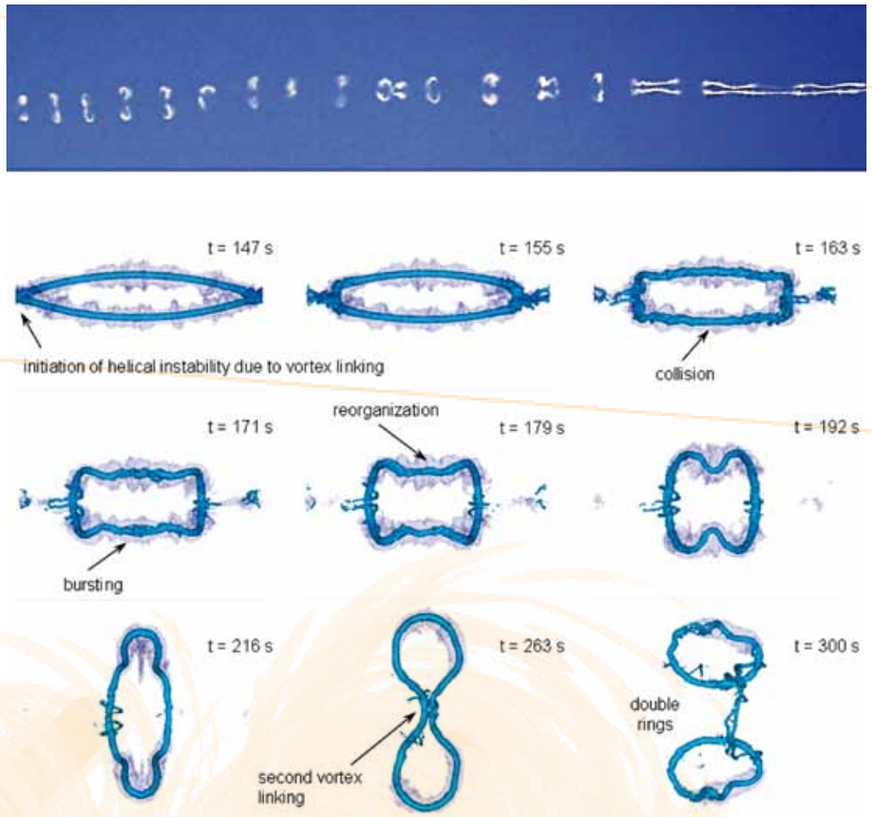


Figure 2: Above: Photo of various stages of vortex rings; flight direction from left to right (photo Sven Lüke, November 16, 2006, 8:53, http://www.4elements-earth.de). Below: LES of vortex ring formation in neutrally stratified and weakly turbulent environment.

A particular risk prevails during final approach, where the vortices can not descend below the flight path, but tend to rebound due to the interaction with the ground [4]. Moreover, the possibilities of the pilot to counteract the imposed rolling moment are restricted due to the low height of the aircraft above ground. Figure 1 shows a snapshot of the interaction of the primary wake vortices (brown) with the turbulent structures generated by the crosswind at the ground surface (green). Several phenomena are visualized in Figure 1: The turbulent streaks generated by the crosswind at the ground surface are wrapped around the primary wake vortices. During this process the streaks are intensified by vortex stretching and establish the so-called secondary vorticity structures. At the same time relatively strong vorticity sheets (brown) are generated at the ground surface. About 20 seconds later these vorticity sheets detach from the ground and start rotating around the primary vortices. Under unfavourable crosswind conditions the rebounding wake vortices may hover over the runway directly in the flight corridor of a landing aircraft.

At cruise altitude winds may be strong but the prevailing turbulence typically is very weak. At these altitudes stable temperature stratification (inversion) typically limits wake vortex lifetimes to a maximum of about 3 mins. Figure 2 depicts wake vortex evolution in a neutrally stratified weakly turbulent environment where long-lived wake vortex rings may form. The vortex rings feature intriguing phenomena both visible in the photograph (above) and in the LES (below): After linking of the vortex pair an elongated vortex ring is established and helical instabilities propagate along the vortices. Later bone-shaped vortex rings transform into the shape of an "8" followed by a double ring phase [5].

Investigations of wake vortex evolution at cruise altitude are not only relevant for passenger safety and comfort but possibly have also significant relevance for the contribution of aviation to global warming. Ice crystals from the exhaust jets may develop contrails or even induced cirrus clouds (contrail cirrus) that modify the radiation budget of the atmosphere [6].

The LES are conducted using MGLET, which is a Finite Volume solver for the incompressible Navier-Stokes equations [7]. The numerical experiments have been conducted typically in a Cartesian grid with 1 m resolution and domain sizes of 400 × 512 × 512 m^3. Larger grid sizes and higher resolutions also with adapted grid spacing have been employed in order to better resolve the vortex cores and the surface boundary layer or to realize large turbulent length scales of the ambient flow. For the simulations up to 1,024 processors on the HLRB II have been used.

- Frank Holzäpfel
- Takashi Misaka
- Anton Stephan

DLR,
Institut für Physik
der Atmosphäre,
Oberpfaffenhofen

References

[1] **Holzäpfel, F., Gerz, T., Frech, M., Tafferner, A., Köpp, F., Smalikho, I., Rahm, S., Hahn, K.-U., Schwarz, C.**
THE WAKE VORTEX PREDICTION AND MONITORING SYSTEM WSVBS - PART I: DESIGN, Air Traffic Control Quarterly, 17, 2009

[2] **Gerz, T., Holzäpfel, F., Gerling, W., Scharnweber, A., Frech, M., Kober, K., Dengler, K., Rahm, S.**
THE WAKE VORTEX PREDICTION AND MONITORING SYSTEM WSVBS PART II: PERFORMANCE AND ATC INTEGRATION AT FRANKFURT AIRPORT, Air Traffic Control Quarterly, 17, 2009

[3] **Holzäpfel, F.**
Probabilistic Two-Phase Wake Vortex Decay and Transport Model, J. Aircraft, 40, 2003

[4] **Holzäpfel, F., Steen, M.**
Aircraft Wake-Vortex Evolution in Ground Proximity: Analysis and Parameterization, AIAA J., 45, 2007

[5] **Misaka, T., Holzäpfel, F., Gerz, T., Manhart, M., Schwertfirm, F.**
Vortex bursting, tracer transport, and decay mechanisms of a counter-rotating vortex pair, subm. to Physics of Fluids, 2011

[6] **Schumann, U., Graf, K., Mannstein, H.**
Potential to reduce the climate impact of aviation by flight level changes, AIAA Paper 2011-3376, 2011

[7] **Manhart, M.**
A zonal grid algorithm for DNS of turbulent boundary layers, Comput. Fluids 33, 2004

CLUES
on how we got to be here

Among mankind's oldest and most fundamental questions are "Where do we come from?" and "How did our earth come into being?". From the dawn of civilization our ancestors have contemplated these mysteries with often beautiful creation myths. Today we know that our earth is a small planet in the solar system, and our sun is one of the few hundreds of billions of stars in the Milky Way and the Milky Way is one of a countless number of galaxies in the universe. Our Milky Way is a spiral galaxy which together with another spiral galaxy of approximately the same mass (the Andromeda galaxy or "M31") forms the main members of the so-called Local Group of galaxies. The third largest member of this group is a somewhat smaller spiral galaxy (called the Triangulum Galaxy or "M33"). In total the Local Group comprises more than 30 galaxies which are distributed in a sphere of about 10 million light-years across centred between Andromeda and Milky Way. At a distance of around 50 million light years from the Milky Way, sits a huge cluster of galaxies (the Virgo cluster) which comprises more than 1,000 member galaxies. This cluster forms the heart of the Local Supercluster which is the most massive structure in the Local Universe. The Local Universe is a tiny fraction of the observable universe, the radius of which is larger than 13 billion light years; yet the special environment around the Milky Way is also the best-known part of the whole universe.

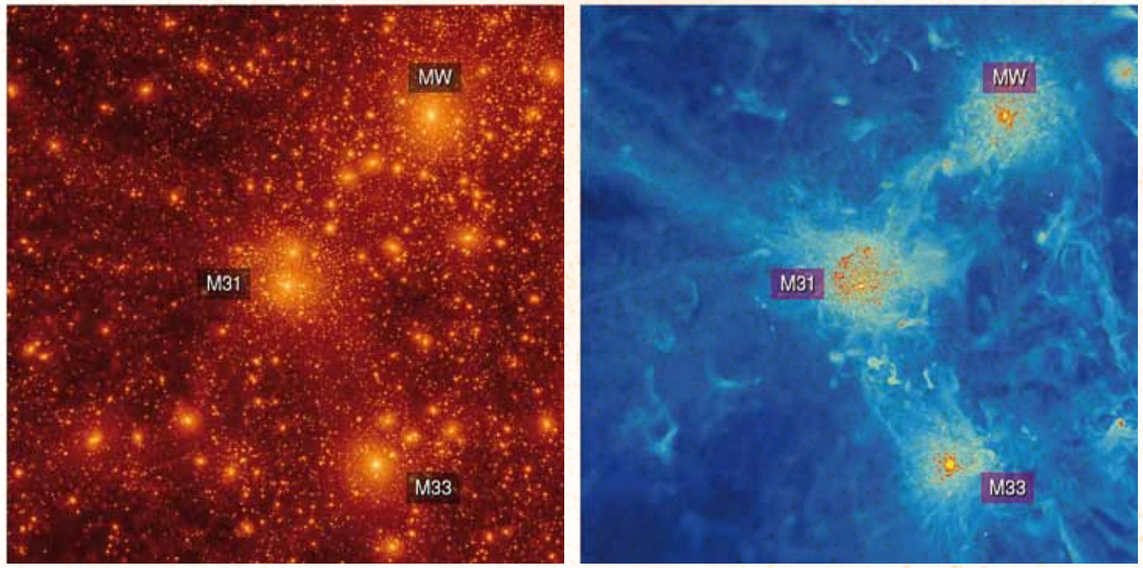


Figure 1: Dark Matter (left) and gas (right) distribution in the Local Group

At present the cosmological parameters can be determined to a precision of a few percent. We are living in a spatially flat Friedmann universe whose mass-energy content is dominated by the mysterious Dark Energy, the yet undiscovered Cold Dark Matter and baryons (atoms). The convergence to this cosmological model sets the framework for studying the formation of galaxies. The standard picture of structure formation suggests that gravitational instability drives the dark matter to cluster hierarchically in a bottom-up fashion. The gaseous baryons settle into to the Dark Matter halos (namely, bound virial Dark Matter structures) and via gas-dynamical dissipative processes cool, fragment and form stars. The non-linear nature of the gravitational dynamics and the gas-astrophysical processes make the problem of structure formation virtually intractable analytically, and therefore the field relies heavily on numerical simulations. This sets the stage for one of the most interesting and difficult problems of modern astrophysics, namely the formation of galaxies in the ΛCDM cosmology.

Over the last decade we have introduced a new approach to cosmological simulations. Rather than modelling a representative cosmological volume or simulating the formation of a randomly chosen high resolution object, our approach uses observations of the nearby universe as constraints imposed on the initial conditions of the simulations. The resulting so-called "constrained simulations" successfully reproduce our local cosmic neighbourhood embedded within the cosmic web, where 'local' means the zone within tens of million light years around us. The main goal of the CLUES project (Constrained Local UniversE Simulations - www.clues-project.org) is to study the formation of the Milky Way and its nearby galaxies. To this end we perform numerical simulations at different supercomputers (LRZ Munich, NIC Jülich, BSC Barcelona).

For our simulations we have used the GADGET-3 code developed by Volker Springel. The code follows the gravitational clustering of the Dark Matter and the gas and solves the equations of gas dynamics by means of the Smoothed Particle Hydrodynamics method. Compton cooling due to the interactions with the photons of the cosmic background radiation, radiative cooling due to atomic recombination and photoionization by the UV photons of a homogenous background from galactic sources have been implemented. The code also models the star formation in the multiphase interstellar medium.

Over the last years we have performed several constrained simulations of the Local Universe using different techniques and different resolution. As an example of a very high-resolution simulation Fig. 1 shows the Dark Matter (left) and gas (right) distribution in the Local Group. The size of the figure corresponds to approximately 6 million light years. The gas is more smoothly distributed than the Dark Matter and there seems to be gas bridges between the galaxies.

Now we zoom into the figure and show in Fig. 2 the stellar distribution in the centre of the galaxy, in the side and top view the stellar disk can be clearly seen. The size of the figure is approximately 120,000 light years. The composite colour images of the stellar distribution have been generated by mapping the UBV bands to the RGB colour channels. The galaxy had a relatively quiet mass accretion history. It grew by smooth accretion and minor mergers in the past 10 billion years so that a disk has been formed. As in nature old stars in this plot look red and young ones blue.

Our method of the reconstruction of initial conditions assumes linear physics, but obviously the Local Group itself is a highly non-linear object. Therefore, the initial conditions contain an unavoidable random component which leads to a scatter in the reconstructed properties of the Local Group even if the large-scale environment of the simulated Local Group resembles the observed one, namely the Local Supercluster. Our simulations have shown that the quiet formation history is typical for a galaxy at the environment of the Local Group. Galactic satellites tend to enter the galaxy along the filament pointing toward Virgo. Interestingly, the two main members of the Local Group, the Milky Way and the Andromeda, have exchanged some of their satellites along the cosmic history.

Figure 2: Top and side view of a stellar disk

In the coming years positions and radial velocities of many more galaxies in the Local Volume and beyond will be available. Together with our improved reconstruction methods the improved input data will allow a more precise reconstruction of the evolution of the Local Universe. The next generation of Petaflop supercomputers will allow us to increase substantially the volume and the mass and spatial resolution of the simulated Local Universe.

- Stefan Gottlöber[1]
- Yehuda Hoffman[2]

[1] Stefan Gottlöber
Leibniz-Institut für Astrophysik Potsdam (AIP)

[2] Yehuda Hoffman
Racah Institute of Physics, HU Jerusalem
Gustavo Yepes
UAM Madrid

Autumn 2011 • Vol. 9 No. 2 • inSiDE 37

Numerical Relativity Simulations of binary black Hole Spacetimes

Introduction

Albert Einstein's theory of General Relativity has fascinated scientists for more than a century. Since exact solutions of the theory can only be found in a few, very special cases, numerical simulations of the Einstein field equations, given as ten coupled nonlinear partial differential equations, were performed for the first time in 1964 [7]. However, due to the lack of computer power and memory at the time, and because of the complexity of the equations, the first modern simulations of a spacetime containing two colliding black holes in axisymmetry had to wait until 1993 [1]. Fully 3D simulations of head on-collisions [8] and non-axisymmetric mergers [3] became available in 1995-1999. Another decade later, it eventually became possible to simulate the inspiral of two black holes on quasi-circular orbits followed by their merger [9,6,2].

Inspiralling binaries of compact objects are among the strongest sources of gravitational waves, which is why a lot of effort goes into the numerical modeling of such systems. Gravitational waves are "ripples of spacetime" which, once they are detected, will allow us to look through clouds of gas, deep into the universe and even back to the big bang. This will open a new window into our universe and will greatly enhance our astrophysical understanding thereof.

The theoretical effort is accompanied by a growing network of gravitational wave detectors all over the world. As an example, the interferometric detector GEO600, a German/British collaboration, located near Hannover, with its 600 m arm length can be seen in Fig. 1.

Physical Background

From theoretical astrophysical investigations we know that a very likely and strong gravitational wave source are two black holes that orbit each other, thereby losing energy through the emission of gravitational waves, which causes their orbits to shrink and circularize. The black holes come closer and closer to each other and eventually merge, which leads to a very strong burst of gravitational radiation. The strong field regime, where the black holes are very close to each other, can only be treated numerically. In practice this means that numerical simulations typically start with an initial binary separation of 10 M, measured in units of the total mass of the system (using geometric units in which the speed of light and Newton's constant of gravity are set equal to one, c=G=1). For two black holes with 10 times the mass of the sun, this corresponds to a distance of approximately 300 km, while the black hole horizons have a diameter of 15 km. For a system with two supermassive black holes like Sagittarius A*, the black hole in the center of the Milky Way with 4.3×10^6 solar masses, a separation of D=10 M corresponds to approximately 1.3×10^8 km. For the numerical simulation, the black holes are given initial momenta which lead to an inspiral on quasi-cirluar orbits. Furthermore, the mass ratio can be specified, as well as the initial individual spins. The goal is to span the parameter space with as many sample points as needed for the establishment of a satisfactory gravitational wave catalogue, which is required for gravitational wave detection. Since there are at least eight parameters to be varied (the mass ratio, three spin components for each black hole, plus the initial separation), a large number of simulations has to be performed.

Figure 1: Gravitational wave detector Geo600 situated near Hannover. The tubes, 600 m in length, run in covered trenches at the edge of the field upwards and to the right.

Computational Code

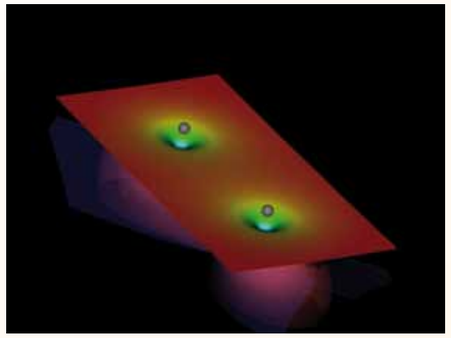
For our numerical evolutions we use the BAM code [5,4], which is designed to solve partial differential equations on Cartesian meshes, in particular a coupled system of (strongly hyperbolic) evolution equations and elliptic equations. Discretization in space is implemented by finite differences. Typical simulations use 6th order stencils, but in principal, up to 12th order stencils are available. Integration in time is performed by a method of lines using 4th order Runge-Kutta.

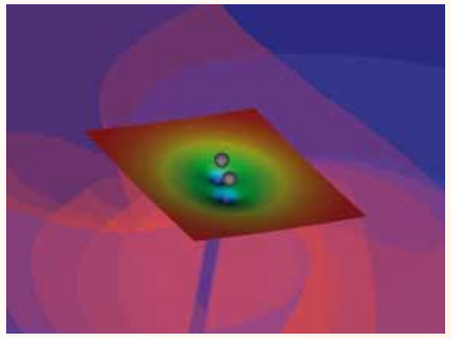
During a black hole binary simulation, different length scales have to be resolved. Close to the black holes, small but highly resolved rectangular boxes are used, whereas the outer regions are covered by very large boxes with lower resolution. In a typical run, the computational domain may be decomposed into 9 to 12 refinement levels with resolution changing by factors of two from roughly 2^{-7}M near the "centers" of the black holes (called punctures) up to 2^{+5}M at the outer boundary. The AMR is based on the standard Berger-Oliger algorithm, which relies on interpolation in time and space to exchange data between the different levels. At the refinement boundaries we use additional buffer zones

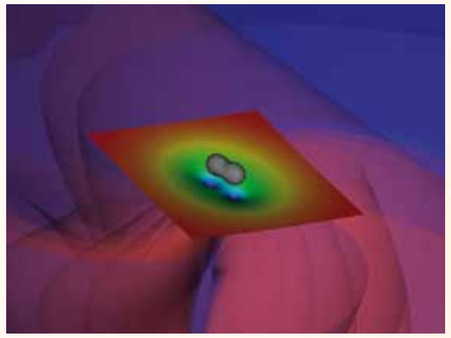
The numerical simulation of black hole binaries poses the question of how to handle the singularity inside a black hole. In the initial data, this is done using the puncture method where the singular part is split off in such a way that the numerical quantities remain finite all over the grid. The Baumgarte-Shapiro-Shibata-Nakamura system of evolution equations is then used to evolve the data. This system of equations constitutes a version of the Einstein equations which is suitable for numerical evolution. In particular, to evolve the 4D spacetime, it is split into space and time, known as 3+1 decomposition. Appropriate treatment of the center of the black holes is necessary, which is mostly accomplished by a clever choice of the numerically evolved quantities.

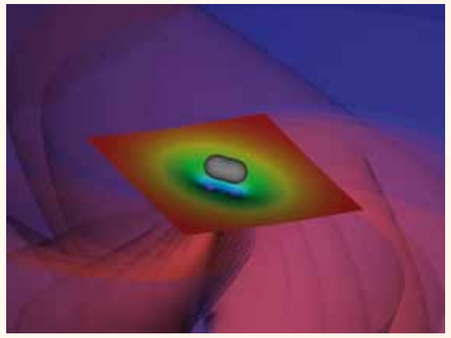
Computational Ressources

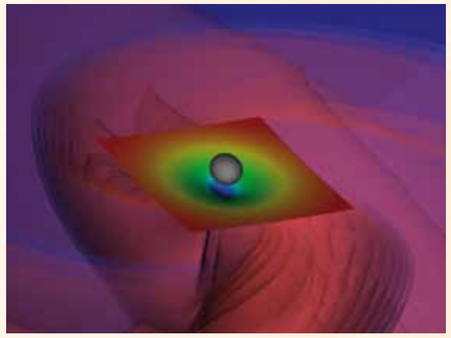
The exact computational requirements of a particular simulation depend on whether the system under study obeys simplifying symmetries. Such symmetries are available for special choices of the initial parameters. The initial configuration also determines the length of the computation. For example, for a typical production simulation that we perform for the international Numerical Relativity Analytical Relativity collaboration, wall-clock run times of up to several weeks have been required.











In a convergence series, one simulation requires about 10 refinement levels with about 100^3 points each. The evolution of the Einstein equations necessitates approximately 120 grid functions (including storage on previous timesteps) of double precision data type. This results in up to 200 GB for a single run.

The BAM code is parallelized using the Message Parsing Interface (MPI). The grid points are distributed among processors, adding buffer points to each processor for the finite difference stencils. Therefore, the number of processors which can be used efficiently depends on the size of the numerical grid. For the example mentioned, the code can efficiently use up to 256 processors since the parallelization of the time-dependent AMR structures is non-trivial. With the current goal being long evolutions, the simulations are limited by run time rather than by the available memory.

Computing Gravitational Waves

Finally, we look at the computation of gravitational waves for a system of equal mass black holes. Waves propagating on a fixed background is a linear concept, and therefore in general relativity pure waves can only be defined far away from any sources. Typically, we put extraction spheres at a certain distance from the center of mass of the binary and compute gravitational wave modes with respect to a spin-weighted spherical harmonic basis. The dominant mode is the quadrupole mode, which for equal mass binaries without spin contributes about 98% to the total energy radiated away from the system.

The upper panels of Fig. 2 shows the initial setup of two equal mass black holes with spins perpendicular to the orbital plane and with magnitudes $S_i/M_i^2 = 0.125$. The gray "spheroids" represent the black holes' horizons in 3 dimensions, which can be interpreted as a measure of the size of the holes.

The plane included in this figure shows the values of the lapse function, which corresponds to the flow of proper time at a point in space, in the part of the orbital plane near the sources. In the red region, proper time and coordinate time are nearly the same (the lapse is nearly one), whereas the proper time does not advance in the regions which are colored in blue (there, the lapse tends to zero). In the lower half of the plot, below the lapse, we can also see the radiation included in the initial data. The radiation emitted into the upper half of the plot is the mirror image of the lower half, which we exclude for the sake of better visibility. Also, the plot only shows one of two wave polarizations.

The consecutive panels in Fig. 2 display the evolution of the black holes and of the waves in time. When the black holes are close enough they merge, and a single horizon containing the until then well-separated individual horizons appears. A single distorted black hole is produced, which rings down to its final state, a stationary, spinning black hole. In the successive panels of Fig. 2, the propagation of the gravitational waves can be followed, where the amplitude changes from high to low when the colors change from red to blue.

The gravitational wave signal in terms of the quadrupole mode can be found in Fig. 3. We plot the real and imaginary parts which, except for a shift in the phase, agree with each other. The signal reaches the extraction sphere after about 100 M and its frequency increases constantly, which is the typical chirp seen for binary black hole inspirals. At approximately 300 M, where the signal

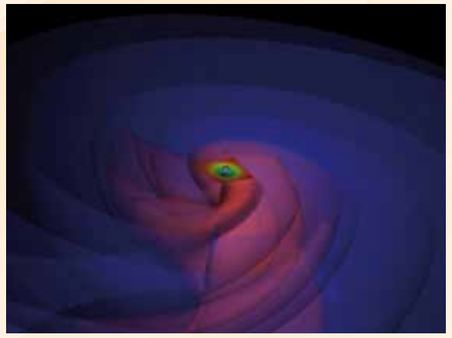


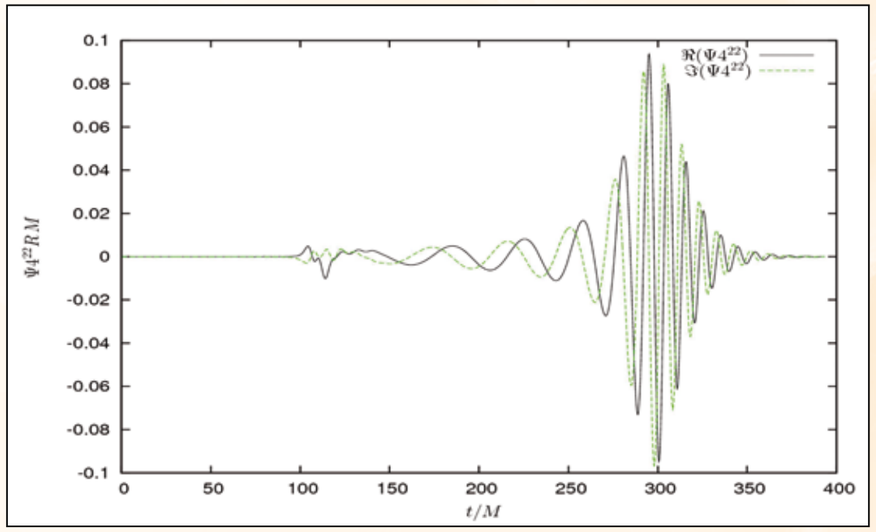
Figure 2: Gravitational wave emission of an equal mass black hole binary. The black holes are represented by their apparent horizon, the gray spheroids. The lapse function in the orbital plane is plotted as a surface in the middle of the figures, and the real part of the radiation into the z<0 half of space is shown. Red indicates high values, blue low ones. In the last plot, we view the system from large distance.

is strongest, the merger takes place. After that, we see the quasi-normal modes of the final black hole, which rings down and eventually stops emitting energy. These kind of signals are expected to be eventually found in gravitational wave detectors like GEO600.

Figure 4 shows the evolution of a slice of the event horizon of two black holes with unequal masses in time. It can be seen nicely how the two holes orbit each other, come closer and eventually merge. Thereby, a new, distorted hole is formed and we clearly see that the new horizon is larger than the original ones. In the newly formed horizon, we also see some ringing until a stationary state is achieved.

Since the breakthroughs in 2005, numerical relativity simulations of two black holes have begun to explore important portions of the parameter space. There are still parameter ranges such as large differences in the black hole masses (mass ratios in the range of 10:1 to 100:1) or high spins that have been barely touched. Since the computation of gravitational wave templates requires high accuracy, significantly higher than routinely achieved today in expensive, month-long simulations, and since this accuracy is required for a large, high-dimensional parameter space, the field of numerical relativity for compact binaries will remain highly active in the foreseeable future.

Figure 3: Real and imaginary part of the quadrupole mode of the gravitational radiation emitted by an equal mass binary with spin. The signal shows the typical increase of frequency and amplitude up to the merger at approximately t=300 M and the quasi-normal ringing of the final black hole after the merger.



Our work is embedded in the Sonderforschungsbereich/Transregio 7 “Gravitational Wave Astronomy”, a collaboration between the Friedrich-Schiller University Jena, the Max-Planck-Institute for Gravitational Physics Golm/Hannover (Albert-Einstein-Institute), the University Hannover, the Eberhard-Karls University Tübingen and the Max-Planck Institute for Astrophysics in Garching. The SFB/TR7 is funded by the German Research Foundation. We also receive funding of the Research Training Group 1523/1 “Quantum and Gravitational Fields” in Jena, as well as the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt, DLR).

Figure 4: Event horizons of two black holes with mass ratio 2:1, scliced in the orbital plane ((x.y)-plane) and evolving in time t. The horizon of the smaller black hole has the smaller diameter, starting in the lower right corner of the plot. The holes merge and form a single event horizon which is larger than the individual horizons and which reaches its final state after some time of ringing down.

References

[1] **Anninos, P., Hobill, D., Seidel, E., Smarr, L., Suen, W.-M.**
Collision of two black holes, Phys. Rev. Lett., 71:2851–2854, 1993, gr-qc/9309016

[2] **Baker, J. G., Centrella, J., Choi, D.-I., Koppitz, M., van Meter, J.**
Gravitational wave extraction from an inspiraling configuration of merging black holes, Phys. Rev. Lett., 96:111102, 2006, gr-qc/0511103

[3] **Brügmann, B.**
Binary black hole mergers in 3d numerical relativity, Int. J. Mod. Phys., 8, 1999, gr-qc/9708035

[4] **Brügmann, B., González, J. A., Hannam, M., Husa, S., Sperhake, U., Tichy, W.**
Calibration of moving puncture simulations, Phys. Rev., D77:024027, 2008, gr-qc/0610128

[5] **Brügmann, B., Tichy, W., Jansen, N.**
Numerical simulation of orbiting black holes, Phys. Rev. Lett., 92:211101, 2004, gr-qc/0312112

[6] **Campanelli, M., Lousto, C. O., Marronetti, P., Zlochower, Y.**
Accurate evolutions of orbiting black-hole binaries without excision, Phys. Rev. Lett., 96:111101, 2006, gr-qc/0511048

[7] **Hahn, S. G., Lindquist, R. W.**
The two body problem in geometrodynamics, Ann. Phys., 29:304–331, 1964

[8] **Matzner, R. A., Seidel, E.,Shapiro, S. L., Smarr, L., Suen, W.-M., Teukolsky, S., Winicour, J.**
Geometry of a black hole collision, Science, 270:941–947, November 10, 1995

[9] **Pretorius, F.**
Evolution of binary black hole spacetimes, Phys. Rev. Lett., 95:121101, 2005, gr-qc/0507014

• Doreen Müller
• Bernd Brügmann

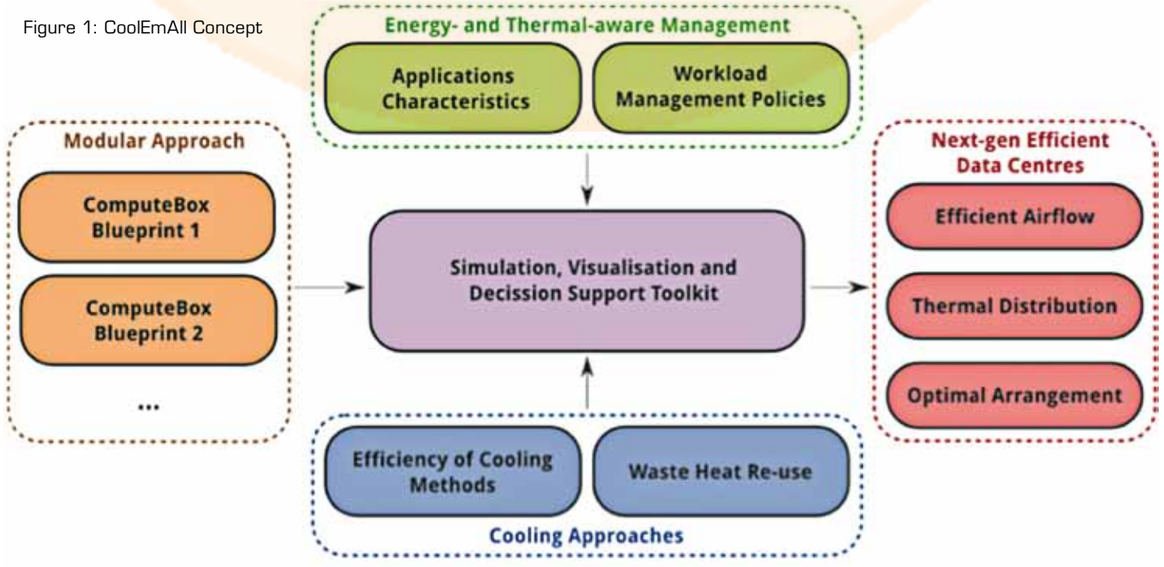
Friedrich-Schiller-Universität Jena

CoolEmAll - Platform for Optimizing the Design, Operation and Cooling of modular configurable IT Infrastructures

IT infrastructures are responsible for at least 2% of the global energy consumption making it equal to the demand of the aviation industry. Furthermore, in many current data centres the IT equipment uses only about half of the total energy for computing, whilst most of the remaining energy is required for cooling and air movement. This often results in poor Power Usage Effectiveness (PUE) values and significant CO^2 emissions. For this reason issues related to cooling, heat transfer, and IT infrastructure location are gaining more attention and are carefully studied during planning and operation of data centres.

In this context, the construction of data centres by using modular building blocks is gaining more and more attention, in particular as a potential antipole to specialized facilities. This modular approach is becoming increasingly popular due to flexibility of design, lower costs and shorter building times. This modular approach can refer to a variety of approaches - one of the most popular are data centres housed in standard shipping containers. In addition, this modular approach can also refer to pre-configured units which are joined together to build-up large computing facilities, with e.g., hundreds of square meters in size.

Figure 1: CoolEmAll Concept



However, while specialised facilities are established in current environments, there is a significant need to analyse the energy efficiency aspects of such a modular approach in order to allow for a comparison of these approaches. In particular, a deep insight into the total energy consumption of both, large data centres and smaller facilities, enforce additional research to determine how efficient this approach is. An important aspect when considering the energy efficiency of modular data centres is the cooling technique. The use of approaches such as “free air cooling” where external air is used to cool systems rather than electrical chillers can help to improve efficiency and achieve PUE ratings close to the ideal of 1.0.

The cooling and heat transfer processes are not the only important aspects influencing the energy efficiency of data centres. Actual power usage and effectiveness of energy saving methods heavily depends on the types of IT applications and workload properties. However, to take full advantage of these methods,

(i) application power usage and performance must be monitored in a fine-grained manner, and

(ii) parameters and metrics characterising both, applications and resources, must be precisely defined.

Consequently, there is a large amount of parameters impacting the energy efficiency of IT infrastructures. All these parameters should be taken into account during the design and configuration of data centres. Issues such as types and parameters of applications, workload and resource management policy based scheduling, hardware configuration, metrics defining efficiency of building blocks, and energy re-used by facilities connected to IT infrastructures are all crucial to understand and improve the energy efficiency of data centres as a whole. To carefully study these issues, simulation, visualization, and decision supporting tools are needed, supporting the optimization of the design and operation of new energy-efficient modular IT infrastructures and facilities.

To address the aforementioned IT energy efficiency issues, the main goal of CoolEmAll is to provide advanced simulation, visualization and decision support tools along with blueprints of computing building blocks for modular data centre environments. Once developed, these tools and blueprints are going to allow to minimize the energy consumption, and consequently the CO^2 emissions of the entire IT infrastructure, taking into account the corresponding facilities as well. This will be achieved by:

1. the design of diverse types of modular computing building blocks (ComputeBox Blueprints), which are going to be well defined by energy efficiency metrics

2. the development of a simulation, visualization and decision support toolkit (SVD Toolkit) that will enable the analysis and optimization of IT infrastructures assembled with these building blocks.

Therefore, these modular computing modules as well as the toolkit are going to take into account three aspects reflecting the major impact on actual energy consumption, namely the cooling model, the according application properties and workloads, as well as workload and resource management policies.

To this end, the energy efficiency of computing building blocks will be precisely defined by a set of novel metrics ex-

pressing relations between the energy efficiency and essential factors listed above. In addition to common static approaches, the CoolEmAll platform will enable studies of dynamic states of IT infrastructures based on changing workloads, management policies, cooling method, and ambient temperature. The main concept of the project is presented in Figure 1.

Therefore, CoolEmAll is going to address the following technical objectives:

1. The development of a simulation, visualization, and decision support toolkit (SVD Toolkit), allowing for analysing and designing modular IT infrastructures and facilities with resource-efficient cooling. This platform will support IT infrastructure designers, decision makers and administrators in the process of planning new infrastructures or improving existing ones, like exemplary shown in Figure 2. The intended modular approach to build and model IT infrastructures and facilities allows for many extension possibilities and high level of customization. CoolEmAll will develop a flexible simulation platform integrating models of applications, workload scheduling policies, hardware characteristics, cooling and air and thermal flows using Computational Fluid Dynamics (CFD) simulation tools. The flexibility of these models, based on model parameter settings, will ensure flexibility of the entire CoolEmAll SVD platform, allowing capturing required model settings and simulate these models for a wide range of applications, workload scheduling policies, IT-Infrastructure and hardware characteristics. Advanced visualization tools and user interfaces will allow users to easily analyse various options and optimize energy efficiency of planned IT infrastructures and facilities, as shown in Figure 3.

2. The provisioning of blueprints of computing modules and a basic prototype. The basic version of this module will enable tests and real-life experiments providing realistic behavioural information for the simulation models, allowing capturing thermal and energy efficiency behaviour on node and rack level, and will also enable the verification of these models. This prototype will include fine-grained monitoring capabilities, allowing for a detailed inspection of the entire environment. Based on this evaluation, a refined and optimized prototype will be designed for diverse scenarios including various hardware densities, cooling methods, workloads and requirements (Figure 4).

3. The definition and evaluation of thermal- and energy-aware workload scheduling and resource management policies. The proposed policies will include intelligent workload scheduling and resource management (e.g. dynamic switching off nodes, lowering frequency and voltage to avoid excessive heat generation). The corresponding decisions will be based on fine-grained hardware and application monitoring. The selection and setup of the corresponding hardware will depend on applications types, workload requirements, cooling method, and ambient temperatures. In order to reflect the evaluation of the CoolEmAll approach within a realistic environment, two major types of workload will be considered: data centre cloud workloads using virtualization and HPC applications. The proposed policies will be applied in simulations to study their impact on energy-efficiency in diverse configurations and in large scale.

4. The analysis and parameterization of applications and workloads. The CoolEmAll simulations as well as workload management techniques will take into account specific workload and application characteristics. To this end, CoolEmAll will prepare benchmarks and classification of applications and workloads. This knowledge about application properties will be used to (i) simulate their impact on thermal issues and energy efficiency and (ii) to propose thermal aware management policies.

5. The definition of specific energy efficiency metrics. Precise definitions of metrics expressing trade-offs between energy and performance will be defined. These metrics will go beyond existing ones (e.g. those defined in the Code of Conduct on Data Centers Energy Efficiency). With this respect, CoolEmAll is going to take into account metrics

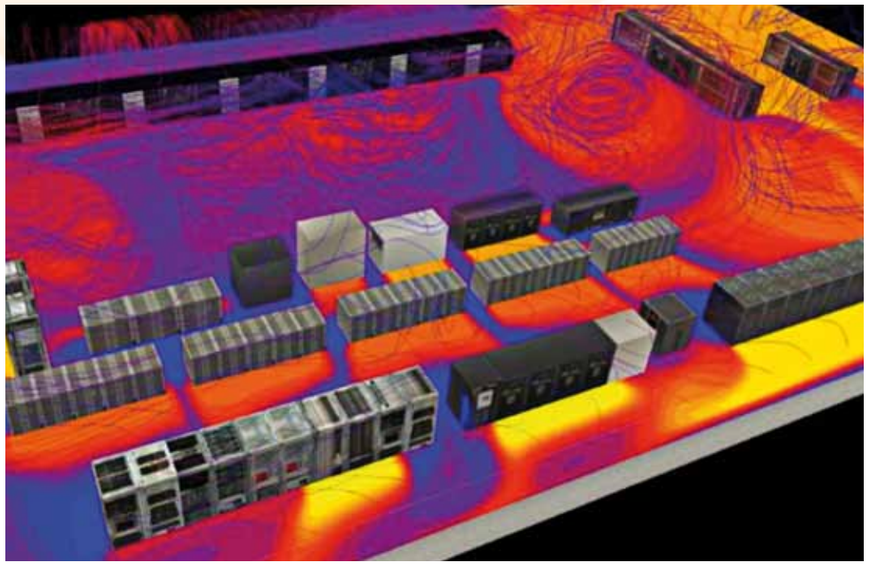


Figure 2: Airflow visualisation in a data centre

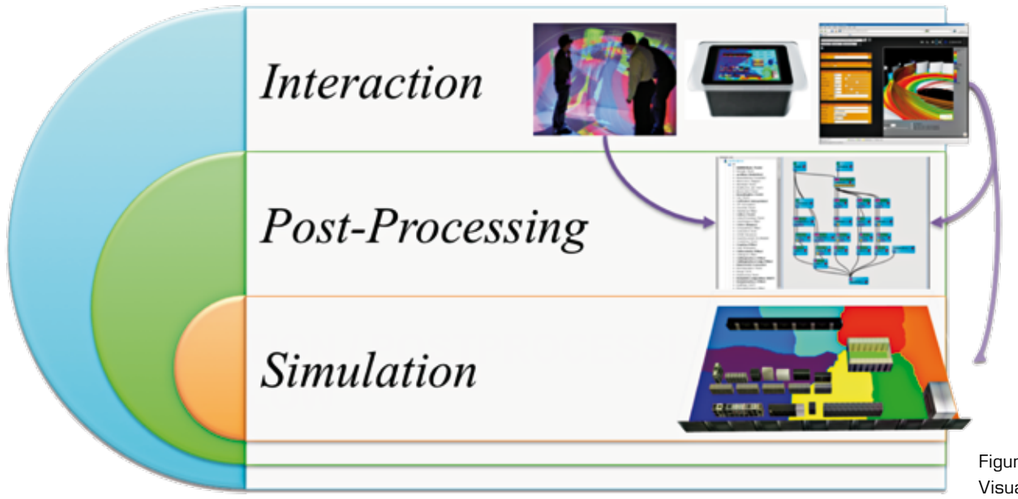


Figure 3: Visualisation Workflow

defined by other projects as well, extend them or propose additional metrics expressing classes of efficiency including relation between energy efficiency and application characteristics, workload properties, ambient temperatures, required heat re-use efficiency, etc. In particular, the metrics defined and evaluated within the GAMES project [1] are going to be taken into account.

6. Verification of simulation tools and their application for specific scenarios. The verification of the proposed methods and software will be performed by tests in real environments using a basic prototype module, real applications, as well as enhanced monitoring systems based on sensors. CoolEmAll will also perform coupled simulations for several diverse settings including large scale IT infrastructures such as whole data centres. These simulations will use collected traces (e.g. from the EU project – GAMES or partners) to plan, design and operate new and existing IT infrastructures and facilities. In this way, the final blueprints of the computing modules will be evaluated and optimized in specific settings.

Simulation and visualization technologies are essential pillars of the CoolEmAll concept, as it allows to leverage and explore new data centre arrangements and solutions. Therefore, the open source CFD package called OpenFoam [3] and the COVISE software [4], developed by HLRS, will be the integration and enhanced to provide a real-time CFD modelling capability. The resulting package enables integration of collected operational data into a simulation to achieve optimal energy-efficient and thermal-aware design of data centres consisting of modular computing units. The first prototypes of these modular computing units are going to be developed based on the existing experience of Christmann with their RECS (Resource Efficient Computing System [2]).

The use of these modular compute units is entirely in line with CoolEmAll's research into impact of cooling solutions on the energy efficiency of IT infrastructures – in particular leveraging outside air ventilation for cooling without artificial equipment as well as reusing waste heat generated by computation. The modular compute units developed as part of CoolEmAll can be used in a symbiosis deployment scenario, where computing facilities benefit other buildings that surround them, resulting in improved overall energy-efficiency of an urban area. Another result of the CoolEmAll research and simulations of the modular compute unit design and deployments could directly benefit the field of high-density server racks by modelling the airflow around them and help in finding solutions for the cooling problem of dense HPC data centres. Both modular high-density data centres and symbiosis deployments directly benefit from the SVD Toolkit (simulation, visualization and decision support), which is one of the major outcomes of the CoolEmAll project.

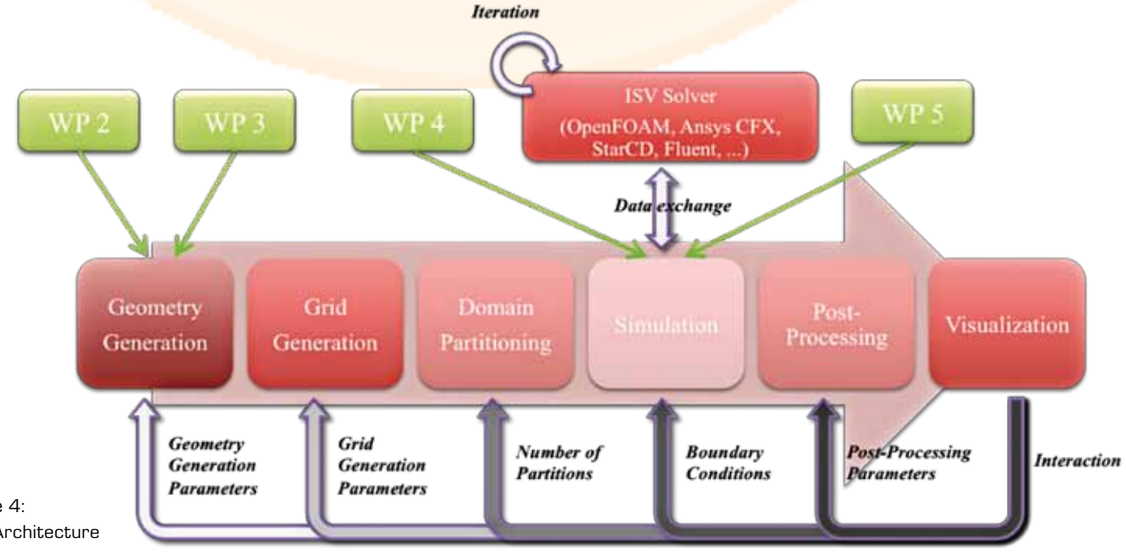


Figure 4:
SVD Architecture

Software, especially applications play significant roles both in terms of the performance and energy-efficiency of computations. Therefore, the CoolEmAll project aims to enhance existing and to develop and standardize novel fine-grained energy and thermal-aware application and hardware metrics. These will take into account both, the energy budget and thermal air flow impacts of an application running on a particular hardware. Due to the granularity of these metrics, the existing monitoring platforms are insufficient due to excessive bandwidth and processing power requirements, thus CoolEmAll will develop new monitoring solutions to face the problem for next generation green IT infrastructure.

Given these metrics and provided a fine-grained monitoring of both hardware and software, CoolEmAll is going to advance the field of cluster scheduling with new workload management algorithms and policies leveraging application characteristics and energy and thermal measurements. Proper scheduling of workloads, taking advantage of colder machines or specific hardware best suited for a given job, can significantly influence the airflow inside a data centre, reducing or even eliminating the need for artificial cooling. The impact of the workload management based on the SVD Toolkit-designed data centres emphasises the CoolEmAll's goal of a holistic approach to next generation green data centres.

HLRS contributes to CoolEmAll by acting as the technical manager of this project as well as simulation and visualization expert. In particular, HLRS is going to coordinate the design and realization of the SVD-platform and contribute with its broad experience in system-monitoring and management expertise. Finally, HLRS is going to contribute with its knowledge in the definition and analysis of energy-efficiency related metrics.

Participants

- Instytut Chemii Bioorganicznej PAN, PL
- High Performance Computing Centre Stuttgart (HLRS), D
- Universite Paul Sabatier, F
- Christmann Informationstechnik + Medien GmbH & Co. KG, D
- The 451 Group LTD, UK
- Institut de Rercerca en Energia de Catalunya, E
- Atos Origin, E

• Alexander Kipp
• Uwe Wössner

University of Stuttgart, HLRS

References

[1] http://www.green-datacenters.eu
[2] http://shared.christmann.info/download/project-recs.pdf
[3] http://www.openfoam.com/
[4] http://www.hlrs.de/organization/av/vis/covise

Dynamical Exascale Entry Platform: the DEEP Project

DEEP is one of the three Exascale projects funded in 2011 by the EU 7th framework program. The DEEP project will last three years starting in October 2011 and is developed in a collaboration constituted by 16 partners from 8 different countries, coordinated by the Research Centre Jülich. DEEP aims at developing a prototype Exascale-enabling supercomputing platform consisting of two parts: a Cluster part based on multi-core-chips with InfiniBand interconnect, and a Booster part based on Intel many-core MIC processors connected through a Terabit EXTOLL network. This Cluster-Booster Architecture (Fig. 1) will serve as proof-of-concept for a next-generation 100 PFlop/s PRACE production system. Furthermore, the innovative hot water cooling concept of DEEP has the potential to improve the power efficiency of HPC systems. A novel open source system software stack (Fig. 2) will be developed for cluster management and resource allocation, based on the ParaStation cluster management software from ParTec. ParaStationMPI will be extended with communication functions to connect the Cluster and Booster via InfiniBand and the interconnection architecture EXTOLL, developed by the University of Heidelberg.

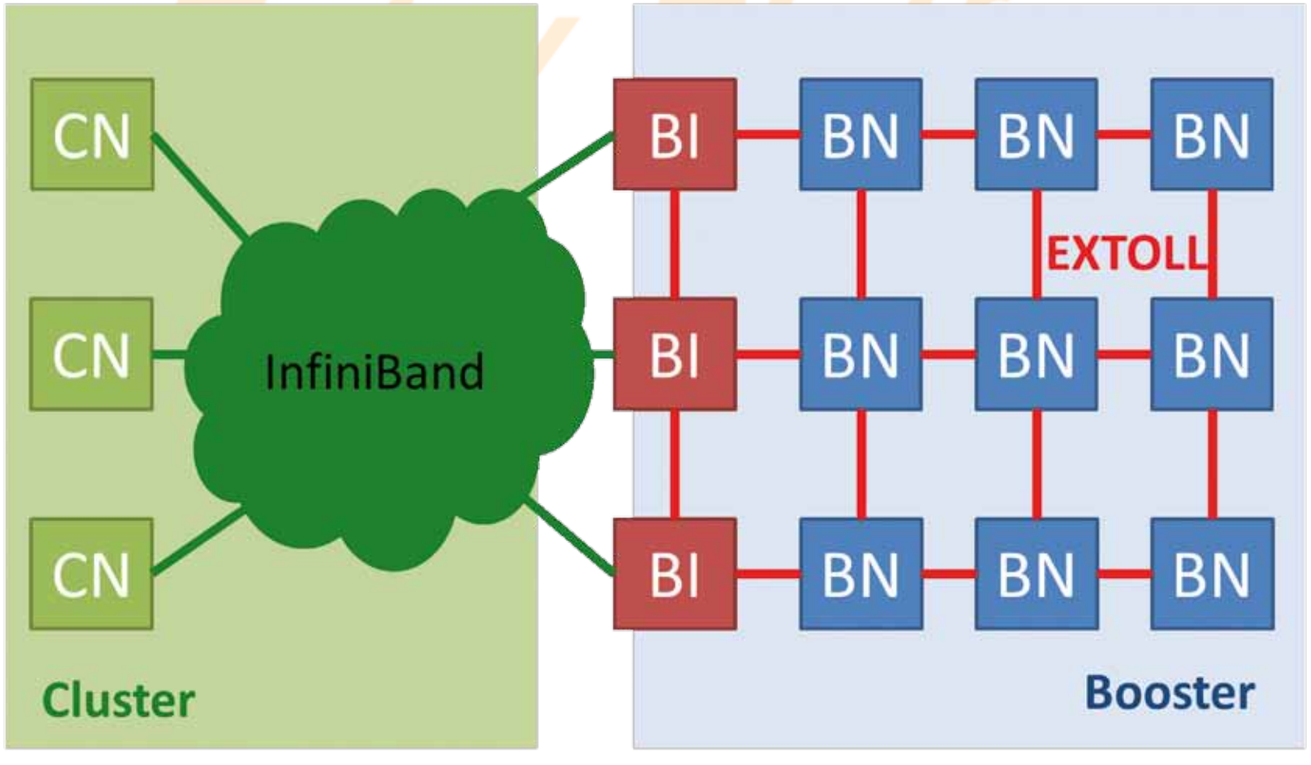


Figure 1:
Sketch of the DEEP Architecture (CN=Cluster Node, BN=Booster Node, BI=Booster Interface)

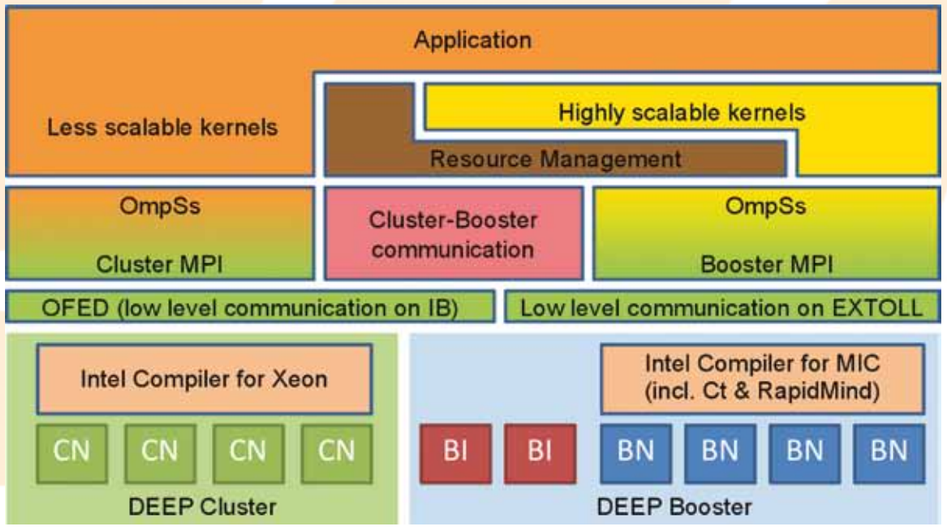


Figure 2:
Software layer for programming environment

In addition, the programming environment OmpSs (OpenMP Superscalar) of the StarSs family from Barcelona Supercomputing Centre will be ported to the DEEP System. Adapting mathematical libraries and performance analysis tools such as Scalasca, provided by the Jülich Supercomputing Centre, will complete the platform to program applications, enabling unprecedented scalability on millions of cores. Representative HPC application codes from Health and Biology, Climatology, Seismic Imaging, Industrial Design, Space Weather, and Superconductivity will be optimized on DEEP. The scalability of the DEEP hardware-software concept shall be demonstrated with respect to the generic multi-scale, adaptive grid, and long-range force parallelization models underlying the application codes. Altogether, the DEEP concept allows an extrapolation to millions of cores for future systems, with the potential to achieve Exascale between 2018 and 2020.

• Estela Suarez
• Norbert Eicker
• Wolfgang Gürich

Jülich Supercomputing Centre

Autumn 2011 • Vol. 9 No. 2 • inSiDE

UNICORE in XSEDE: Towards a large-scale scientific Environment based on Open Standards

- Morris Riedel
- Bastian Demuth

Jülich Supercomputing Centre

Evolution from TeraGrid

Starting in 2001, the National Science Foundation program TeraGrid has developed into one of the world's largest and most comprehensive Grid projects, offering resources and services to more than 10,000 scientists. It's successor, the Extreme Science and Engineering Discovery Environment (XSEDE, www.xsede.org), has started in July 2011 and is expected to excel the previous program in terms of service quality while lowering technological entry barriers at the same time. These and other goals are to be achieved in the project's five year grant period with an overall budget of $121 million. Among the partnership of 17 institutions, the Jülich Supercomputing Centre (JSC) is the only organization located outside the USA.

Open Standards-based Architecture

Since many scientific communities operate internationally, one key element of XSEDE is the use of open standards in order to promote interoperability with other distributed computing infrastructures such as PRACE in Europe. Figure 1 shows the extended reference architecture providing mandatory XSEDE Enterprise Services at every major XSEDE site as well as optionally available Community Provided Services. For many years, the JSC and several other XSEDE partners have been active in establishing the key standards that now define the interfaces of the XSEDE Enterprise Services. Within the Grid Interoperation Now (GIN) community group of the Open Grid Forum (OGF, www.ogf.org), such key standards as BES/JSDL for running remote computations have demonstrated their impact on scientific applications. Based on these standards, scientific workflows can be executed today across different infrastructures with no less than 8 different Grid middleware technologies.

Jülich's Role in XSEDE

The JSC not only contributes its extensive Grid know-how gained from European research projects and its experience in standard-based software engineering, but also a technology called Uniform Interface to Computing Resources (UNICORE, www.unicore.eu). Being developed by partners all over Europe, UNICORE is a Grid system that provides secure and seamless access mechanisms to a variety of different computer systems and platforms. It facilitates the remote execution of scientific applications as well as sharing software, resources and data. UNICORE is fully based on Web services and open standards in order to allow seamless interoperation with other standard compliant Grid systems such as Genesis II which is developed at the University of Virginia. Being complementary to the more lightweight Genesis II services, UNICORE meets all the security requirements of modern High Performance Computing centres and provides extensive support for their highly specialized hardware as well as their varying batch systems.

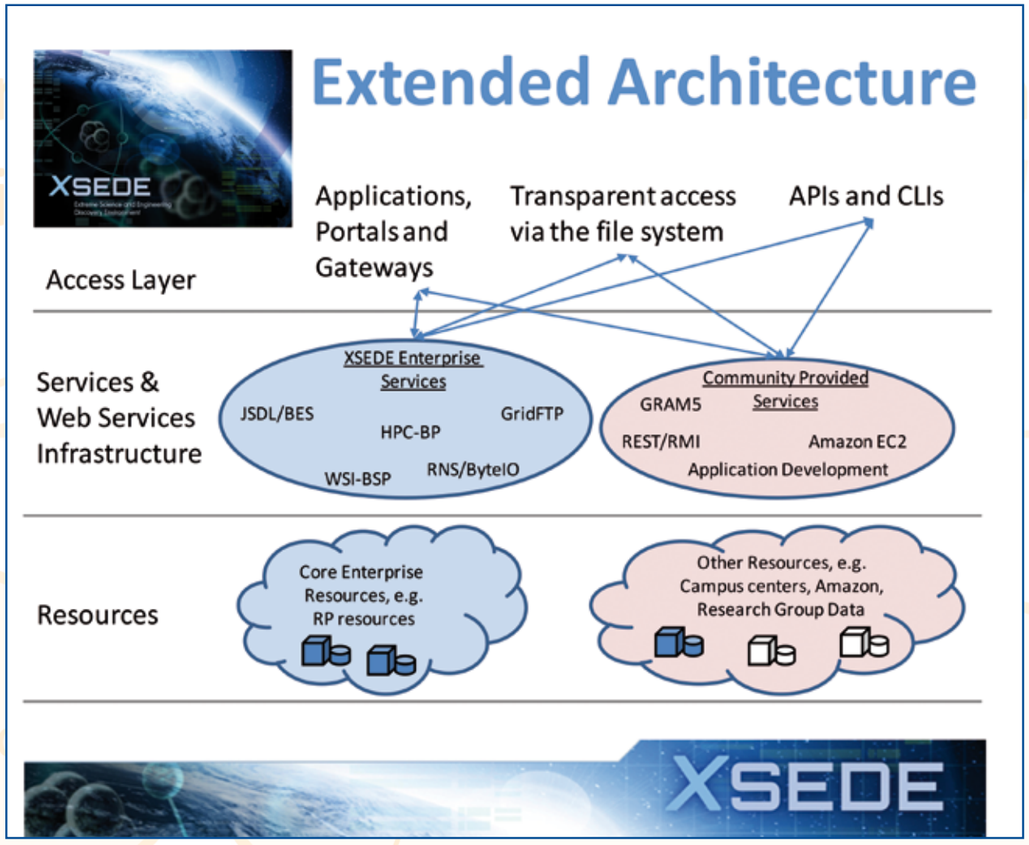


Figure 1: The current XSEDE architecture aims at providing XSEDE Enterprise Services at every major XSEDE site and optionally available Community Provided Services. The architecture will evolve over time according to end-users' needs.

Infrastructure Vision

The XSEDE architecture envisions deploying UNICORE as part of the XSEDE Enterprise Services at major US high performance centres whereas Genesis II will be used for integrating smaller computer systems such as desktop PCs in order to provide interoperability with campus Grids across the country. The resulting infrastructure is expected to cover both high performance and high throughput computing, thus enabling innovative research and discovery requiring both types of parallel computations. Moreover, collaboration between American and European scientists will be easier than ever: UNICORE will also be deployed on the European Grid Infrastructure (EGI) and is already installed on many of the systems forming the infrastructure of the European supercomputing project PRACE.

UIMA-HPC: High-Performance Knowledge Mining

UIMA-HPC will enable data mining applications to make efficient use of High Performance Computing resources. The focus will be on the bio-pharmacological area for which e.g. the PubMed database holds more than 20 million entries. Researchers in this field need to find answers to questions such as the following: For a given base structure, are there any structure variants already mentioned in literature, and if so, are there any indications of their effects? Are structure variants protected by third-party rights or are they freely available? These questions cannot be answered by sheer keyword searches.

Therefore the project aims to realize an HPC-based solution for the automated analysis of multi-modal pharmaco-chemical document databases, taking the patent-search use-case as an initial solution design driver. The combination of text and structure analysis is an innovative approach, but will be based on an existing and well-tested data analysis architecture: the Unstructured Information Management Architecture (UIMA). UIMA is a software architecture which specifies component interfaces, design patterns and development roles for creating, describing, discovering, composing and deploying multi-modal analysis capabilities. The UIMA specification is being developed by a technical committee at OASIS (Organization for the Advancment of Structured Information Standards).

The UIMA-HPC approach centres on the workflows for the automated annotation of a document corpus, the workflow comprising analysis components within the UIMA architecture.

The individual "annotation engines", such as text-mining of a document or analysis of diagrams within a document based on Optical Character Recognition (OCR), are of a computational complexity such that parallelization at the level of the heterogeneous "node" of a modern HPC system is highly appropriate, meaning parallelization for deployment on multi-core and/or GPU-accelerated processors. Handling the large quantity of documents - and the related load-balancing issues created by the diversity of computational complexity relating to individual documents - to be analyzed by independent instantiations of the annotation engines for the workflow is handled at the level of the nodes of the HPC compute system as a whole and will be realized within an adaptation of the UNICORE software system. An example workflow is shown in Figure 1 where red-coloured framed boxes denote UIMA-analysis pipelines and orange triangles split or collect data to achieve load-balancing and parallel execution of pipelines.

UIMA-HPC is a collaborative project funded in part by the German Federal Ministry for Education and Research (BMBF – Bundesministerium für Bildung und Forschung, Förderkennzeichen 01IH11012) and running for three years; the Consortium is led by FHG-SCAI and includes Forschungszentrum Jülich GmbH, scapos AG, and Taros Chemicals GmbH.

The Jülich Supercomputing Centre puts its R&D effort in the development of algorithms and tools for the distribution and collection of data as well as the calculation of the appropriate number of parallel analysis streams and monitoring.

• Mathilde Romberg

Jülich Supercomputing Centre

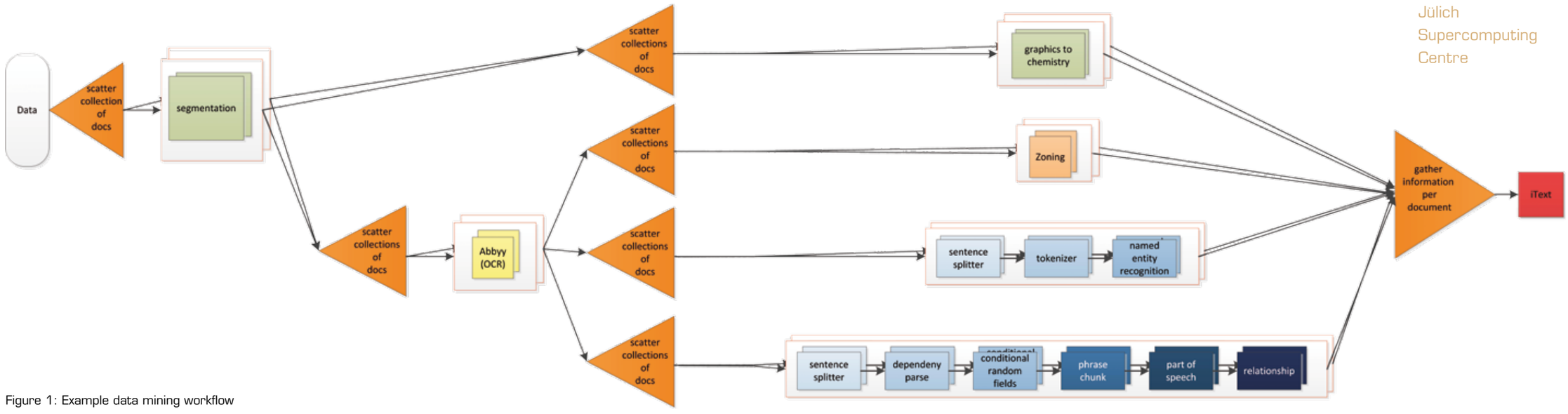
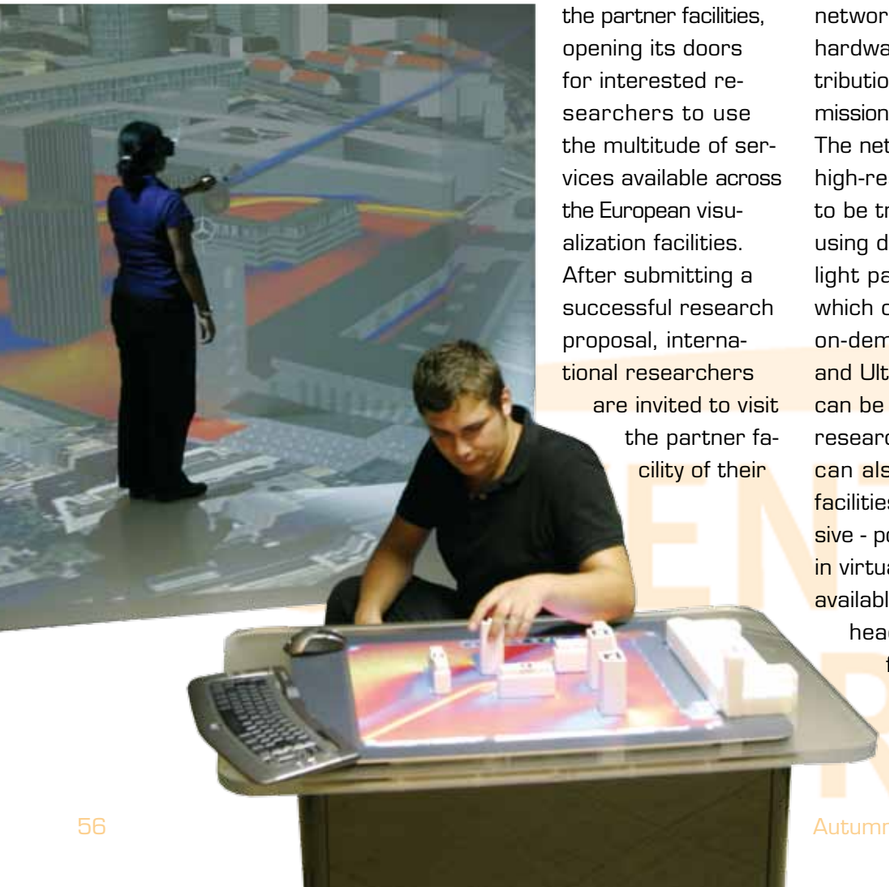


Figure 1: Example data mining workflow

VISIONAIR

VISIONAIR is establishing a European infrastructure for high level visualization facilities that is open to research communities across Europe and around the world. By integrating existing facilities to a pan-European network, it will create a world-class research infrastructure enabling to conduct cutting-edge research. Current scientific challenges such as climate evolution, environmental risks, molecular biology, health, and energy require the management of increasingly complex and voluminous information, thus calling for the development of ever more powerful visualization methods and tools. On many sites across Europe, it is infeasible to fund the necessary visualization facilities that are needed to tackle high fidelity, large screen and/or immersive visualization. VISIONAIR is targeted to fill in this gap by providing access to the partner facilities, opening its doors for interested researchers to use the multitude of services available across the European visualization facilities. After submitting a successful research proposal, international researchers are invited to visit the partner facility of their choice or what fits best to the scientific goals they have in mind to conduct their research. They are not only given access to the top visualization facilities in Europe, but are also supported in their research by funding their living and travel expenses. Researchers can choose from over 20 facilities located in 12 countries in Europe and Israel.



The project targets different fields of visualization. Scientific Visualization offers access to methods, software and hardware needed for successfully visualising scientific data, including - but not limited - to engineering, medical visualization, biology, chemistry and physics. Ultra-High-Definition facilities connected by high speed networks are targeted at users that want to create high resolution, high quality images (up to 8k) and access those by high speed networks. VISIONAIR provides the hardware and the unique network distribution services needed for transmission of the images to their end-points. The network services enable multiple high-resolution digital-media streams to be transported among global sites, using dynamically provisioned optical light paths across multiple domains, which can be used on a scheduled or on-demand basis. While Scientific and Ultra-High-Definition Visualization can be done in any environments, researchers specifically targeting VR can also apply at a multitude of facilities. Here, the focus is on immersive - possibly also haptic - experiences in virtual environments. Equipment available for researchers ranges from head mounted displays to fully fledged stereoscopic PowerWalls and CAVEs. Further specialized equipment available allows to carry out research by using Augmented Reality, a technique that allows to overlay the real environment with context dependent computer generated images. Researchers also have access to the latest developments in display technology, like holographic displays or the above mentioned 8k displays.

The project maintains an already huge database of visualization software and models that is available for all researchers for free. Thus, experts can explore the multitude of visualization packages that are already available. Software covered here ranges from processing filters, converters and readers to fully fledged modellers and visualization packages. VISIONAIR is rounded up by several research activities concerning the usability and accessibility of the facilities and their software for external researchers with a strong focus on collaboration.

VISIONAIR is a common infrastructure that grants researchers access to high level visualisation facilities and resources. Both physical access and virtual services will be provided by the infrastructure. Full access to visualization-dedicated software is organized, while physical access to high level platforms is accessible to other scientists, free of charge, based on the quality of the project submitted. Indeed, researchers from Europe and around the world are welcome to carry out their research projects using the visualization facilities provided by the infrastructure. By creating this European visualization network it will be possible to create a landmark that will have a broad visibility throughout the research communities around the world.

Within this project, the HLRS is providing access to its CAVE, PowerWalls, head mounted displays and its haptic driving simulator. Visitors will be able to interactively visualize large simulation results or even realize computational steering and interactive simulations by leveraging the power of a 40 nodes visualization cluster. The visualization software COVISE will not only allow visitors to analyse their data in Virtual Reality but they can also overlay their visualization over physical prototypes or test beds using Augmented Reality techniques. The collaborative features will allow them to analyse the simulations together with their colleagues at home or with remote scientists.

Project calls are expected to open end of 2011. Interested researchers are invited to submit a proposal at **www.infra-visionair.eu**

Partners

INPG Grenoble (F), Grenoble INP (F), University of Patras (GR), Cranfield University (UK), Universiteit Twente (NL), Universität Stuttgart (D), PSNC Posnan (PL), Université de la Mediterrané Marseille (F), CNR Genova (I), INRIA Rennes (F), KTH Stockholm (S), Technion Haifa (IL), RWTH Aachen (D), Poznan University of Technology (PL), ENSAM Aix-en-Provence (F), TU Kaiserslautern (D), University of Salford (UK), Fraunhofer IPK Berlin (D), i2cat Barcelona (ES), University of Essex Colchester (UK), MTA SZTAKI Budapest (HU), ECN Nantes (F), University College London (UK), Politecnico di Milano (I), EMIRACLE Brussels (B).

• Uwe Woessner

University of Stuttgart, HLRS, Germany

Collaborative Research into Exascale Systemware, Tools and Applications (CRESTA)

For the past thirty years, the need for ever greater supercomputer performance has driven the development of many computing technologies which have subsequently been exploited in the mass market. Delivering an exaflop (or 10^18 calculations per second) by the end of this decade is the challenge that the supercomputing community worldwide has set itself. The Collaborative Research into Exascale Systemware, Tools and Applications project (CRESTA) brings together four of Europe's leading supercomputing centres, with one of the world's major equipment vendors, two of Europe's leading programming tools providers and six application and problem owners to explore how the exaflop challenge can be met. CRESTA focuses on the use of six applications with exascale potential and uses them as co-design vehicles to develop: the development environment, algorithms and libraries, user tools, and the underpinning and cross-cutting technologies required to support the execution of applications at the exascale. The applications represented in CRESTA have been chosen as a representative sample from across the supercomputing domain including: biomolecular systems, fusion energy, the virtual physiological human, numerical weather prediction and engineering.

No one organization, be they a hardware or software vendor or service provider can deliver the necessary range of technological innovations required to enable computing at the exascale. This is recognized through the on-going work of the International Exascale Software Project [1] and, in Europe, the European Exascale Software Initiative [2]. CRESTA will actively engage with European and International collaborative activities to ensure that Europe plays its full role worldwide. Over its 36 month duration the project will deliver key, exploitable technologies that will allow the co-design applications to successfully execute on multi-petaflop systems in preparation for the first exascale systems towards the end of this decade.

Overall Concept and Objectives of the CRESTA Project

HPC systems exist to deliver results to their users from numerical simulation and modelling applications. At the centre of CRESTA are therefore six applications designed to execute well on petascale systems today and that will be expected to execute well on exascale systems tomorrow.

Co-design by Applications

Each of these applications has been carefully chosen (a) as an application that can be reasonably expected to need to run at the exascale (for reasons of problem size, time to compute etc.) and (b) to represent a key user community with a grand challenge who have the need to compute at the exascale in order to deliver their scientific or engineering results.

By understanding the current state of the limitations of the algorithms and problem sizes for each of these applications, CRESTA will be able to develop improved application performance at the petascale on then current systems (perhaps 100 petaflop/s by 2014) and define a clear roadmap for each application to get it to the exascale by the end of this decade mapped against the expected hardware designs we expect to see by the end of the decade.

Systemware for Exascale

However, application optimization and algorithmic modifications only represent part of the exascale challenge.

Systems of the scale envisaged present enormous challenges in terms of power requirements, operating system issues such as resiliency and process management, communication and programming libraries, languages, compilers, debuggers and profilers. Applications must interact with many of these aspects of the exascale systemware required to compile, link, run, debug and profile application codes.

CRESTA will therefore use the knowledge available from our target Cray platform, with estimates of what an exascale system built using these hardware technologies will look like

Table 1: CRESTA's co-design applications

Application name	Scientific grand challenge domain	Partner responsible
GROMACS	Biomolecular systems	KUNGLIGA TEKNISKA HOEGSKOLAN (KTH)
ELMFIRE	Fusion energy	ABO AKADEMI UNIVERSITY (ABO)
HemeLB	Virtual Physiological Human	UNIVERSITY COLLEGE LONDON (UCL)
IFS	Numerical Weather Prediction	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS (ECMWF)
OpenFOAM	Engineering	THE UNIVERSITY OF EDINBURGH, UNIVERSITY OF STUTTGART (HLRS, IHS), CENTRALE RECHERCHE SA (CRSA)
Nek5000	Engineering	KUNGLIGA TEKNISKA HOEGSKOLAN (KTH)

from a systems perspective, and the requirements and form of each of our applications codes, to build and explore appropriate systemware. In addition to operating system features we will look at compiler and library issues, how to debug at the exascale (with Allinea's DDT debugger), how to optimize application performance at the exascale (with TUD's Vampir tool-suite and KTH's perfminer) and how to manage the pre- and post-processing of data resulting from simulations (an often neglected area in systems design). The balance of technical work in CRESTA will be apportioned 40% on applications and 60% on systemware.

Incremental vs. disruptive Approaches

A key feature of CRESTA is its use of dual pathways to exascale solutions. Many problems in HPC hardware and software have been solved over the years using an incremental approach. Most of today's systems have been developed incrementally, growing larger and more powerful with each product release. However, we know that some issues at the exascale, particularly on the software side, will require a completely new, disruptive approach. For example, the communications overhead of a particular numerical solver may grow quickly with the number of cores. With million core systems on the horizon, the performance of such algorithms may be so poor that all speedup stops. In these cases a different method of numerical solution will be required. This may be highly disruptive to the application code but it will set it on the path to executing at the exascale.

CRESTA will therefore employ incremental and disruptive approaches to technical innovation - sometimes following both paths for a particular problem to compare and contrast the challenges associated with each approach.

Table 2: Role of applications for different research challenges

Research Challenge \ Application	GROMACS	ELMFIRE	HemeLB	IFS	OpenFOAM	Nek5000
Underpinning and cross-cutting technologies	Significant	Significant	Significant	Significant	Significant	Significant
Development environment	Significant	Essential	Essential	Essential	Significant	Significant
Algorithms and libraries	Essential	Useful	Useful	Essential	Essential	Essential
User tools (including pre and post processing)	Useful	Significant	Essential	Useful	Essential	Useful

The Co-design Process

At the heart of the proposal is the co-design process, involving guidance and feedback between the co-design applications and the systemware work packages. A high level roadmap to achieving exascale for each co-design vehicle will be developed as will a more detailed needs analysis to guide development work in the systemware WPs. They in turn will input expertise and provide solutions to the challenges. This will be a cyclical process, the needs analysis will be dynamic and updated regularly based on experiences of all the developers across WPs.

The solutions to the challenges will be different for each application and their integration with different work packages will depend on these. For each application we have characterized their interaction with each work package on a three-point scale. Tasks are either Essential, Significant or Useful based on how critical the development is to enabling the application for exascale.

Collaboration and Exploitation

No one organization, be they a hardware or software vendor or service provider can deliver the necessary range of technological innovations required to enable computing at the exascale. This is recognized through the on-going work of the International Exascale Software Project and, in Europe, the European Exascale Software Initiative. At the same time, the PRACE Research Infrastructure (PRACE-RI) [3] gives Europe, for the first time, a leadership-class HPC research infrastructure accessible by any suitably qualified European scientist with a suitable problem to solve.

Within this decade, the PRACE-RI will provide exascale computing resources for Europe's scientists. As four of the leading partners in the various PRACE projects, the supercomputing centres represented in CRESTA will ensure exploitation of the results of the CRESTA project by the PRACE-RI.

CRESTA will collaborate with EESI [4] and any subsequent project to ensure that Europe's exascale research community has a natural meeting place to share discoveries and demonstrate leadership on the world-stage. The CRESTA partners who are involved in IESP will continue this activity and seek, where appropriate to collaborate on the International stage as evidenced by the letters of support included with this proposal.

All new software developed by CRESTA will be made available as Open Source Software.

References

[1] The International Exascale Software Project, http://www.exascale.org/ [Last accessed August 23, 2011]

[2] The European Exascale Software Initiative, http://www.eesi-project.eu/ [Last accessed August 23, 2011]

[3] PRACE RI, http://www.prace-ri.eu/

[4] European Exascale Software Initiative, http://www.eesi-project.eu/

• Mark Parsons[1]
• Stefan Wesner[2]

[1] EPCC, The University of Edinburgh

[2] University of Stuttgart, HLRS, Germany

Europe and Russia to collaborate on Next-Generation Supercomputing

Faster computers do not immediately lead to better results. Incompatibility between the requirements of existing software and the capabilities of new supercomputers is a growing problem that will be addressed by a pioneering new Russian-European collaboration called APOS.

The future of High-Performance Computing means two things - massive parallelism and heterogeneity. Processor manufacturers are squeezing ever more computational cores onto their CPUs, and HPC vendors are looking to augment these many-core chips with GPU-like accelerators to deliver the next push forward for computing capacity. While such developments give scientists the potential to run ever bigger, faster, or more detailed simulations, there are significant challenges involved before today's most important application codes attain these advances.

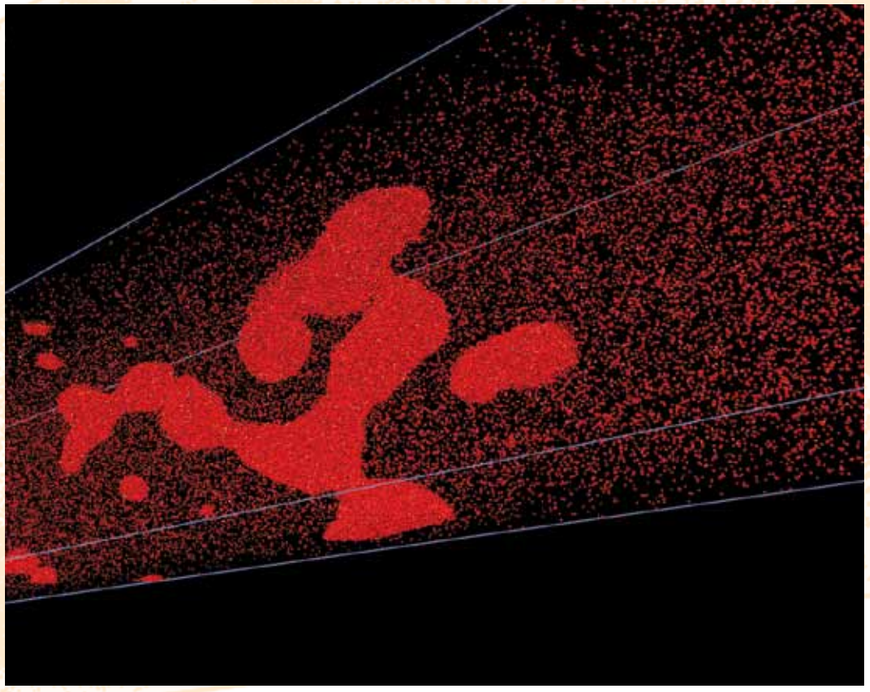


Figure 1: Simulation of a phase transition with LS1 Mardyn.

"A key consideration for exascale computers is programmability. The APOS team is working with some of today's most important HPC applications, developing the tools and techniques required to help them realize the benefits of future-generation supercomputers," said Dr. George Beckett, APOS project manager at EPCC.

Funded by the European Union's 7th Framework Programme and the Ministry of Education and Science of the Russian Federation, APOS will target popular codes from the strategically important application areas of seismic modelling, oil- and gas-reservoir simulation, computational fluid dynamics, fusion energy, and nanotechnology. It will pool expertise from computational science, High-Performance Computing and software engineering to enable next-generation science on next-generation supercomputers.

"Improving seismic image quality to find oil and gas deposits thousands of metres below the seafloor requires ever increasing computing power. Simulating oil and gas reservoir behaviour with various physics and detailed meshes also poses serious challenges in terms of programming methods that should match future architectures. We expect significant benefit from joint research with academia." Pascal Dauboin, TOTAL S.A.

APOS is a unique coupling of internationally renowned research centres in Europe and Russia, along with TOTAL - one of the world's major oil and gas groups - and CAPS - an innovative European HPC tools producer.

"Collaboration in the framework of APOS allows Russian scientists to adapt their numerical methods to modern supercomputer systems using the experience and innovative tools of European partners, and to carry out joint theoretical investigations on the basis of developed scientific software tools, to be used in practically important application areas." Boris Chetverushkin, Keldysh Institute of Applied Mathematics, Russian Academy of Sciences.

Facts and Figures

The APOS project runs for two years, from February 2011 until January 2013. The collaboration is formed from two consortia. APOS-EU is composed of University of Edinburgh, CAPS entreprise, Uniwersytet Warszawski, TOTAL S.A., and Höchstleistungsrechenzentrum Stuttgart (HLRS). APOS-RU is composed of the Keldysh Institute of Applied Mathematics (Russian Academy of Sciences), the National Research Centre "Kurchatov Institute", the Ugra Research Institute, and the Moscow Institute of Physics and Technology.

Project website www.apos-project.eu

• Colin W. Glass

APOS-EU Consortium
University of
Stuttgart, HLRS

Debugging on the next Level: TEMANEJO

Introduction

Supercomputers nowadays consist of hundreds of thousands of processing units. For example HLRS will install a 1 PFlops Cray XE6 with dual-socket 16-core AMD Interlagos nodes in over 3,500 nodes, i. e. 112,000 cores. The parallelization techniques of the programs which run on these machines, e. g. MPI and OpenMP, just to name two of the most used, have however been designed when the largest computers were 100 times smaller. Due to constant development, these techniques still work, but they have to be expanded by additional programming models in order to keep up the scalability.

A promising approach is to combine the widely used MPI with shared memory parallelization using task based parallelizm, for instance SMPSs [1]. In this programming model, pieces of code, e. g. functions, are specified as potentially parallel using special markers. Given the tasks' data dependencies, either provided by the programmer or extracted automatically by the compiler, the program's tasks are scheduled to worker threads at runtime. Programming threads already is hard, due its sharing of ressources due to possible race conditions - add to that issues of identifying and correctly stating the tasks's dependencies and performance implications makes it difficult for the programmer to come up with an efficient parallelization strategy of a given program.

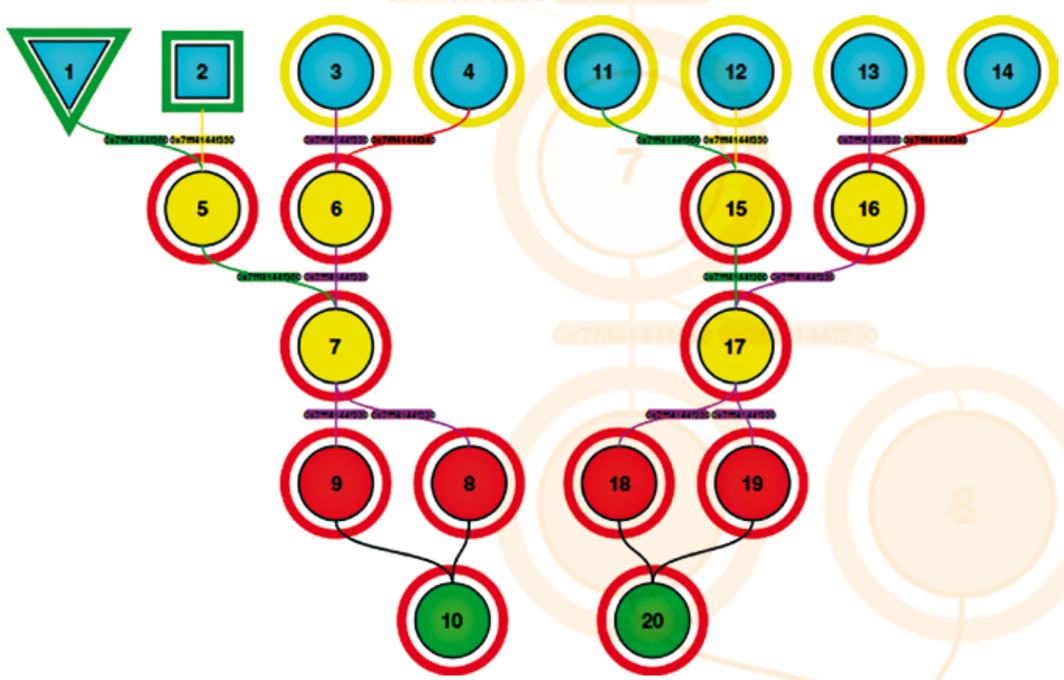


Figure 1: A simple dependency graph with ten tasks in two independent subtrees visualized by TEMANEJO. The node's interior color represent the task (the taskified function), the surrounding margin's color represent the task's state, e. g. queued (yellow) or running (green). The red margins indicate that the task has unfulfilled dependencies and can therefore not be queued yet. The node shapes (here triangle and box) denote two different worker threads for running tasks, which may be scheduled to different cores. The text labels and colours of the edges indicate the memory addresses of the dependencies.

We developed a graphical debugger [2] which is capable of displaying the relevant information in an accessible manner and giving the programmer many possibilities for interaction with the running program.

The difference between task-based programming models and other ways of parallelising applications is that it is not known a priori when a task will be executed or which core will execute it. The only information per task is the data dependencies (basically the data's memory address), and therefore which other tasks depend on it. As a result a dependency graph, to be precise a directed-acyclic graph (DAG) is created at runtime, with nodes being tasks and edges being the dependencies as seen in Figure 1.

New tasks will be added to the dependency graph while the program is running and finished tasks are removed from it. When for a given task all the dependencies are fulfilled it can be executed by any thread at any time. Therefore every run of an application can result in a different order of execution on different processing units in the system. This makes task based parallel programs extremely hard to debug with normal debuggers such as gdb. Two examples of task-based programming models are SMPSs and OMPSs. Both are developed in the TEXT project with participation of the Application, Models and Tools group at the HLRS [3].

A new Debugger

Traditional debuggers work command based on a per thread basis - to switch between concurrently executing threads, the programmer needs to issue commands and retrieve the state of the thread.

That means one steps through lines of code and switches from one thread to another. In task parallel programming, one can not see why a task can or cannot be executed. Moreover, up to now the programmer was left to her own devices imagining the dependency graph or viewing post processed graphs from program outputs.

We developed a graphical debugger for task parallel programs in order enable the programmer to visualize the dependency graph and interact

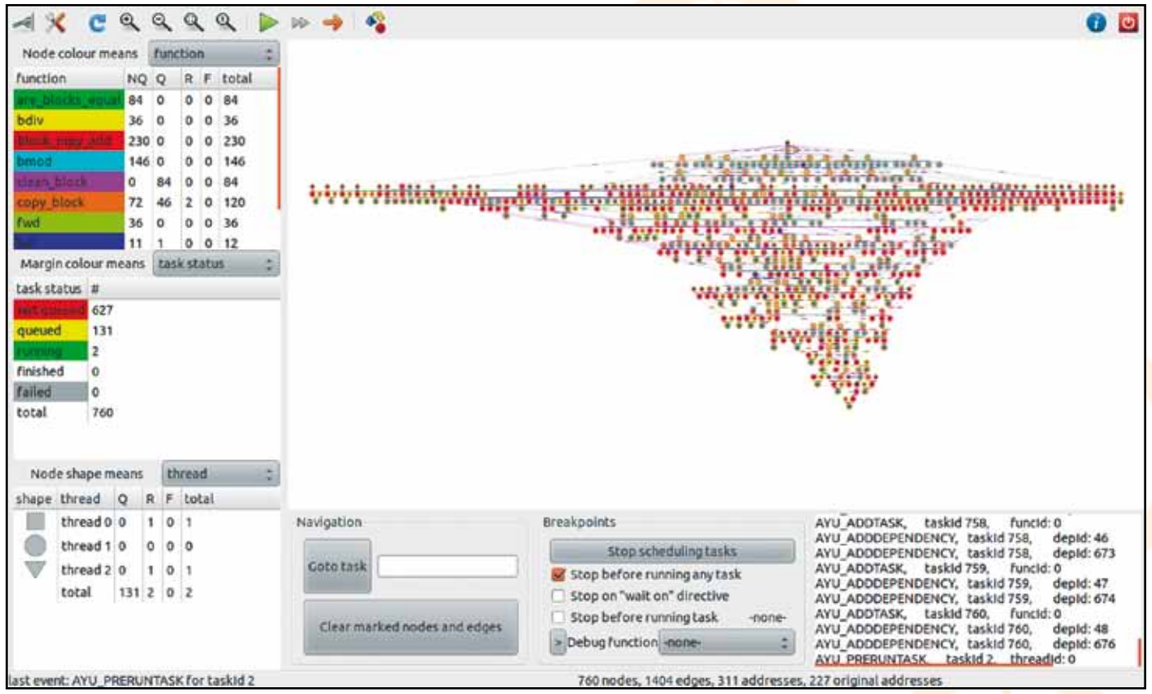
with the runtime while the program is executing (see Fig. 2).

Debugging the parallel code can be separated into two phases. Analysis of the dependency graph and controlling, i. e. changing the actual execution of the application. For the former the information about the nodes (tasks) and edges (dependencies) of the dependency graph have to be extracted and communicated to a visualization tool. For the latter a control tool (which can be the same as the visualization tool) will enable the programmer to send commands (in the following called requests) to the application.

Instead of stepping from line to line or putting breakpoints at certain lines or functions, the programmer of a task-based application wants to step through the dependency graph and inhibit or force the execution of tasks. This is supplemented by attaching a debugger like gdb to the running process on the actual compute node.

In order to analyze and control the execution of an application we instrumented the runtime developped by Barcelona Supercomputing Center (BSC) with event handlers at certain points of the tasks life cycle. The event handler is implemented as a weak reference to a library function which is dynamically linked to the application using the LD_PRELOAD environment variable.

This event handler is implemented as a library called AYUDAME (Spanish for help me) and is called at distinct moments during runtime. It performs numerous tasks, to pass the information to an external tool, like TEMANEJO via TCP connection or react to the event itself.



On the frontend side any tool may attach via the open protocol - the debugger TEMANEJO (Spanish for I handle you) performs two tasks:

1. It visualizes the dependency graph of the StarSs application giving the user the possibility to verify the correctness of the task interdependencies and to optimize the dependency structure.

2. It controls the StarSs runtime environment by sending requests to the library. Requests can be to pause the runtime under specific conditions, to block tasks, or single- and multi-step through the graph, or attaching a debugger like gdb.

For displaying the dependency graph TEMANEJO needs at least two pieces of information: Which tasks exist and what are their dependencies. The former is received when the program framework creates tasks while the latter is gathered when the dependencies to the predecessors of each newly created task are analyzed. Further information passed is the status of each task, i. e. whether they are queued, running or finished, the memory address of a dependency variable and a timestamp.

All of this information is stored in a data structure in TEMANEJO for further analysis (e. g. number of dependencies of a task, longest and shortest paths through the graph, execution duration of tasks etc.). Using the node colour, node margin colour, node shape and edge colour as configurable indicators, the programmer can access the needed information in a convinient and intuitive way, allowing reduction of information, e. g. color-coding the executing thread, showing imbalance.

Figure 2: Screenshot of TEMANEJO running a sparse matrix inversion parallelized with SMPSs.

Conclusions

With AYUDAME a scalable and extensible framework for tools for task based programming models has been developped. It has been tested with SMPSs and OMPSs, but can be used with any task based parallelization model as long as it is instrumented with calls to the AYUDAME event handler. The TEMANEJO debugger allows visualising the dependency graph, stepping through the task graph, block tasks and prioritise tasks.

Offering this functionality, TEMANEJO empowers the programmer of task parallel applications to debug and optimize the application efficiently.

References

[1] **Marjanovic, V., Labarta, J., Ayguade, E., Valero, M.**
Overlappingcommunication and computation by using a hybrid MPI/SPMSs approach. In ICS '10: Proceedings of the 24th ACM International Conference on Supercomputing, pages 5-16, New York, NY, USA, June 2-4, 2010, ACM

[2] **Brinkmann, S., Gracia, J., Niethammer, C., Keller, R.**
TEMANEJO - a debugger for task-based parallel programming models.
In Proceedings of ParCo'11, 2011, submitted for publication

[3] TEXT - Towards EXaflop applicaTions (www.project-text.eu)

• Steffen Brinkmann
• Christoph Niethammer
• José Gracia
• Rainer Keller

University of Stuttgart, HLRS

JUDGE - a new GPU Cluster in Jülich

In February 2011, the IBM iDataPlex GPU cluster JUDGE (JUelich Dedicated Gpu Experiment cluster) started production. In the first step JUDGE consists of 54 compute nodes, 2 login and service nodes, and 2 GPFS gateway nodes. Each compute node is equipped with 2 Intel Xeon X5650 (Westmere) 6 core processors running at 2.66 Ghz, 96 GB main memory, and additionally 2 NVIDIA Tesla M2050 (Fermi) GPUs with 3 GB memory. All nodes are connected with QDR Infiniband. JUDGE uses the global parallel file system GPFS from the storage cluster JUST for home file systems and a fast work file system. The nodes are connected to JUST via 2 GPFS gateway nodes that route the Infiniband to 10G Ethernet.

JUDGE is used for simulation projects in Jülich and collaborations of the Jülich Supercomputing Centre (JSC) with internal and external partners.

In the second step JUDGE was enlarged in October by 68 additional compute nodes. The new node type is the same as in configuration the first step except for the GPU type. Each of the new nodes contains 2 NVIDIA Tesla M2070 GPUs, that are equipped with 6 GB of memory each. The new enlargement is dedicated to JuBrain, a common platform for simulations on neuroscience of the Institut für Neurowissenschaften und Medizin (INM) and JSC.

The whole system consists of 122 nodes with 244 GPUs and 11.7 TB main memory providing a peak performance of 1.16 TF/s per node and about 140 TF/s in total. With the high number of GPUs and large amount of memory, JUDGE is well-suited for memory intensive applications and programs using GPU power for acceleration and allows for scaling tests of large multi-GPU applications.

Furthermore, the system is also used for educational purposes. Several GPU programming tutorial courses have been provided to support users in porting their applications to this accelerated cluster platform. In September, a tuning workshop has been organized for advanced users in collaboration with NVIDIA. Here applications from several fields, e.g. Micromagnetism simulation, PET image reconstruction, Protein Folding and Molecular Dynamics have been optimized for JUDGE.

www.fz-juelich.de/ias/jsc/judge

• Otto Büchner

Jülich Supercomputing Centre

Figure 1: IBM iDataPlex cluster "JUDGE"

Figure 2: Back side with rear door heat exchanger

Figure 3: Compute nodes with GPU trays on front side

Systems

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

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

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

Contact:
Leibniz Supercomputing Centre

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

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

Draft of LRZ's next supercomputer "SuperMUC" by IBM

Compute servers currently operated by LRZ are given in the following table

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
SGI Altix 4700 "HLRB II" Intel IA64 19 x 512-way (to be decommissioned)	9,728 Cores 39 TByte	62.3	Capability Computing	German Universities and Research Institutes, DEISA
IBM System iDataPlex "SuperMIG" (Migration system for the 3 PFLOP/s system "SuperMUC", expected for June 2012)	8,200 Cores 51 TByte	78	Capacity Computing	German Universities and Research Institutes, European Grid Projects
Linux-Cluster Intel Xeon EM64T/ AMD Opteron 2-, 4-, 8-, 16-, 32-way	4,438 Cores 9.9 TByte	36.3	Capacity Computing	Bavarian and Munich Universities, D-Grid, LCG Grid
SGI ICE Intel Nehalem 8-way	512 Cores 1.6 TByte	5.2	Capability Computing	Bavarian Universities, PRACE
SGI UV	2,080 Cores 6.1 TByte	20	Capability Computing	Bavarian Universities, PRACE
Megware IB-Cluster	2,848 Cores 2.8 TByte	22	Capability Computing, PRACE Prototype	Bavarian Universities, PRACE

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

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 co-operation industry always has acces to the most recent HPC technology.

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 HVV.

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.

Contact:
Höchstleistungsrechenzentrum
Stuttgart (HLRS)
Universität Stuttgart

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

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

View of the HLRS Cray XE6 "HERMIT"

Compute servers currently operated by HLRS

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
Cray XE6 "HERMIT" (Q4 2011)	3,552 dual socket nodes with 113,664 AMD Interlagos cores	1,045	Capability Computing	European and German Research Organizations and Industry
NEC Hybrid Architecture	12 16-way nodes SX-9 with 8 TByte main memory + 5,600 Intel Nehalem cores 9 TB memory and 64 NVIDIA Tesla S1070	146	Capability Computing	German Universities, Research Institutes and Industry, D-Grid
IBM BW-Grid	3,984 Intel Harpertown cores 8 TByte memory	45.9	Grid Computing	D-Grid Community
Cray XT5m	896 AMD Shanghai cores 1.8 TByte memory	9	Technical Computing	BW Users and Industry

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

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

JÜLICH
FORSCHUNGSZENTRUM

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

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

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

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

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

Contact:
Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich

Prof. Dr. Dr. Thomas Lippert
52425 Jülich
Germany

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

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

Compute servers currently operated by JSC

System	Size	Peak Performance (TFlop/s)	Purpose	User Community
IBM Blue Gene/P "JUGENE"	72 racks 73,728 nodes 294,912 processors PowerPC 450 144 TByte memory	1,002.6	Capability Computing	European Universities and Research Institutes, PRACE
Intel Linux CLuster "JUROPA"	2,208 SMT nodes with 2 Intel Nehalem-EP quad-core 2.93 GHz processors each 17,664 cores 52 TByte memory	207	Capacity and Capability Computing	European Universities, Research Institutes and Industry, PRACE
Intel Linux CLuster "HPC-FF"	1,080 SMT nodes with 2 Intel Nehalem-EP quad-core 2.93 GHz processors each 8,640 cores 25 TByte memory	101	Capacity and Capability Computing	EU Fusion Community
IBM Cell System "QPACE"	1,024 PowerXCell 8i processors 4 TByte memory	100	Capability Computing	QCD Applications SFB TR55, PRACE
Intel GPU Cluster "JUDGE"	122 nodes with 2 Intel Westmere 6-core 2.66 GHz processors each 244 graphic processors (NVIDIA Fermi) 11.7 TByte memory	140	Capacity and Capability Computing	selected HGF Projects

Growing Science at HLRS – Beyond Bare Metal

The importance of a supercomputing center is typically measured in speed of its supercomputer. As of today the key number to be achieved is 1 Petaflop/s and beyond. However, the system itself is by far not enough. It is only a tool to develop solutions. Hence systems have to be set in a working, stable, and well organized environment. First of all, they need an excellent infrastructure to host them. Second, they need support by the hosting center that helps users to get performance from the system Third, it takes a lot of system related research to develop new methods and harvest the full potential of a supercomputer. Finally, in order to bridge the gap between pure research and real world application of the systems a concept is required to integrate supercomputing into the real world of research, development and production.

HLRS has installed a Cray XE6 system recently. The peak performance is in the range of 1 PF/s and we expect to see sustained performance in the order of 20 - 25% of the peak performance [1]. Beyond the pure hardware HLRS has made further steps to grow its HPC-ecosystem.

Infrastructure

Power supply and cooling are one of the most pressing issues these days. Therefore HLRS has built new power cooling facilities. They provides additional 4 MW of power which will bring power supply of HLRS to a total of 5 MW. At the same time it will provide a very efficient water cooling infrastructure relying fully on free cooling up to 18 degrees outside temperature. Energy of the Cray XE6 will be used in the new research building of HLRS currently under construction.

With an increase in number of research projects and funding (see Fig. 1) also the number of scientists has grown (see Fig. 2). In order to meet the requirements of HLRS the planning for a new research building was started in 2007 and the construction work started in June 2011. The building is an extension of the existing HLRS headquarter and will provide additional 1,950 m² of office space allowing HLRS to bring all employees together in a single environment. Furthermore the new building will integrate a five-sided Virtual Reality environment. On one hand this will allow for more realistic visualization of simulation results. On the other hand the new CAVE will have a direct physical link to the supercomputers allowing new more interactive use models such as simulation steering. The building will be finished in summer 2012 and will then host HLRS and a number of collaborative staff in various projects.

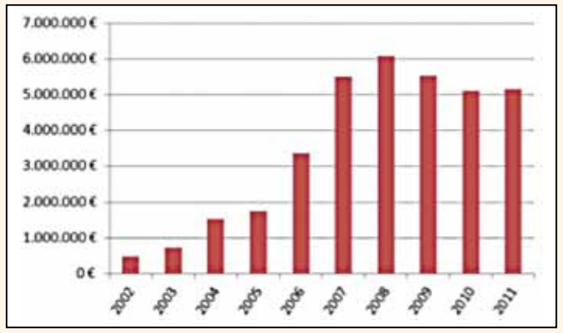


Figure 1: Project funding of HLRS over the last 10 years

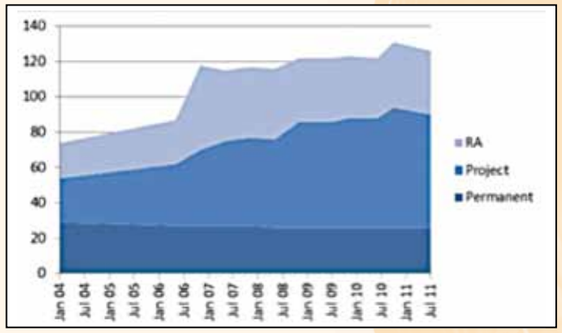


Figure 2: Number of HLRS employees over the last 7 years (RA is student research assistants

Research

Over the last years HLRS has increased the scope of its research tremendously. Four major levels of research can be identified which are merged by HLRS into a coherent program of research for High Performance Computing.

State level: At the state level HLRS has agreed with the State Ministry of Science on a major initiative to develop scalable software for High Performance Computing. The initiative will make available 30 Mio Euro for the purchase and development of highly scalable software.

Federal level: For more than two years HLRS is participating in the HPC software initiative of the German Federal Ministry of Science. HLRS is working on a variety of application driven research projects that aim at transforming flops into solutions. At the same time HLRS is taking part in a German Cluster of Excellence funded by the German Research Society (DFG) called "SimTech" [2]. SimTech has a focus on basic methods in simulation technology and is the only such cluster of excellence in Germany. Within SimTech the Director of HLRS Michael Resch serves as a principal investigator for High Performance Computing.

European level: At the European level HLRS has a long history of project research and collaboration. The most recent HPC project (CRESTA) is described in this issue of inside. CRESTA is aiming at supporting the development of Exascale systems in Europe. However, project research at HLRS goes beyond Exascale. Many of our projects aim at making HPC more productive or bringing HPC closer to the industrial environment to which HLRS is connected.

Industrial level: The collaboration of HLRS with industry is described in this issue of inSiDE elsewhere. In research we have started a Cray Center of Development recently. Over the next five years researchers from Cray and from HLRS will work on scalable applications and on tools to harvest the potential of future HPC systems.

Summary

Beyond the installation of hardware and its participation in organizational collaborations like GCS and infrastructure initiatives like PRACE [3] HLRS is heavily involved in research. Over the last two years and in the next year to come HLRS has substantially improved its infrastructure and has at the same time set up a framework for research that integrates various scopes and levels of funding to develop better solutions for its share-holders.

• Michael Resch,

University of Stuttgart, HLRS

References

[1] **Resch, M.** New HLRS System HERMIT, inSiDE, Vol. 9, No. 1, Spring 2011

[2] Simtech Excellence Cluster, www.simtech.uni-stuttgart.de

[3] PRACE, www.prace-ri.eu

Autumn 2011 • Vol. 9 No. 2 • inSiDE

Foundation of the European Open File System (EOFS) Society

On December 16, 2010 the European Open File System (EOFS) Society was founded in Munich. The initiative to co-operate on this topic originated from JSC. As a legal form the Societas Cooperativa Europaea (SCE, European Cooperative Society) was selected.

The purpose of this non-profit organization is to

- promote the establishment and adoption of an open source parallel file system
- sustain and enhance its quality, capabilities and functionality
- ensure that the specific requirements of European organizations, institutions and companies are considered
- initiate projects or to collaborate with existing projects at regional, national, European and international level in order to support Research & Development activities concerning Open File Systems
- ensure that engagement and activities with other organizations will not directly or indirectly interfere with the intellectual property or other contractual and legal obligations of its members
- facilitate the extension of business operations to non-members.

The following 14 organizations are founding members: Forschungszentrum Jülich, Bull GmbH, CEA/DAM, Data Direct Networks, Universities of Paderborn and Zürich, GSI Helmholtzzentrum für Schwerionenforschung GmbH, credativ GmbH, T-Platforms, HPCFS, Mellanox, Whamcloud, Leibniz Rechenzentrum (LRZ) and ParTec GmbH.

At the founding meeting several members expressed their satisfaction about the fact that it is the first time a Europe wide consortium has engaged in the open-source development of high-speed file systems. It was agreed that such an important piece of file system software should not be made proprietary with no absolute guarantee of future access. On the contrary, the organization should establish an open source code base, prioritize the development efforts, and improve the functionality and stability of the underlying parallel file system.

Furthermore, the intention was expressed to build a bridge to similar development efforts in the United States where in the meantime two organizations were founded around Open File Systems: HPCFS and OpenSFS.

During the Lustre™ Users Group (LUG) meeting - held in Orlando, Florida in April 2011 - EOFS was the driving force in bringing these two US organizations together. Later all members of HPCFS became members of OpenSFS and OpenSFS changed bylaws and contributorship agreement as required by HPCFS.

The next logical step was to sign a Memorandum of Understanding between EOFS and OpenSFS at ISC'11 in Hamburg, because they recognized that to succeed in this endeavour it will require collaboration and cooperation to meet their mutual goals, including

- sustaining and further developing the Lustre™ open source parallel file system
- promoting an open-source development model for the Lustre™ file system
- promoting a forum for collaboration that is vendor-neutral and
- protecting end users of the Lustre™ file system in the spirit of open source.

The two organizations wanted to establish a mechanism by which they can peer with one another in order to facilitate collaboration among working groups of each organization and to establish a single canonical Lustre™ source repository and tracking system for community development.

All this work can be considered as a big success to form a single community with global participation. It guarantees the continued future of the Lustre™ file system supported by a large number of strong supporters.

As an immediate next step EOFS organized the first European Lustre™ Workshop in Paris in September 2011 where administrators and developers from Europe and worldwide came together and exchanged their experiences, developments, tools and good practices to run this parallel filesystem.

More information is available at the **EOFS website: www.eofs.org**

EOFS
European Open File System

• Klaus Wolkersdorfer

Jülich Supercomputing Centre

JSC Guest Student Programme on Scientific Computing

As one of the leading HPC centres in Europe, the Jülich Supercomputing Centre hosts regular support activities and educational programmes in the field of Scientific Computing. One of the main priorities of these events is to introduce young academics to HPC and its applications in scientific research.

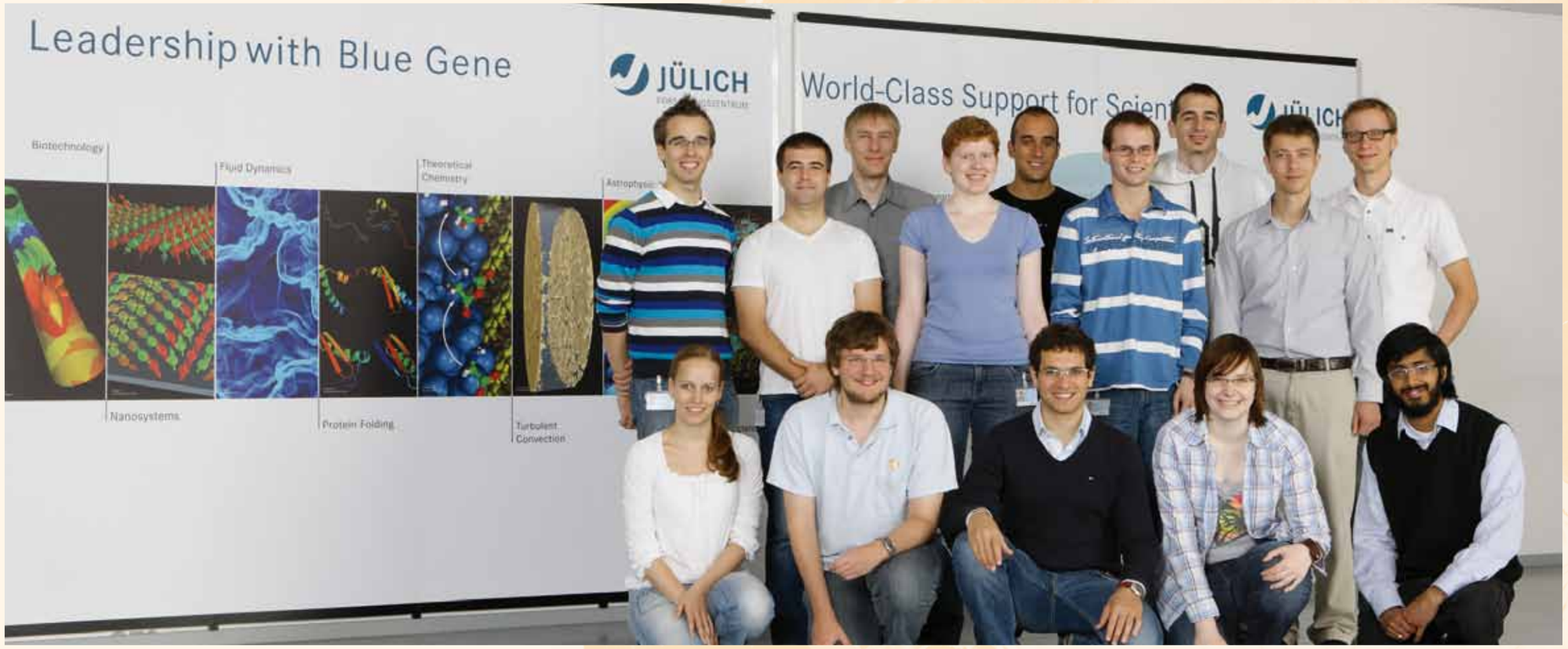
During the last eleven years, JSC has offered a very successful Guest Student Programme on Scientific Computing. Each Summer since the year 2000, between eight and eleven international students have had the opportunity to tackle exciting and challenging scientific projects in the field of HPC under the supervision of scientists from JSC, the NIC Research Group, and other institutes at Forschungszentrum Jülich. In some cases, the programme impact was such, that a number of students returned to JSC as Master or PhD candidates in order to pursue a career in HPC. Furthermore, many JSC cooperations began as a guest student project.

In 2011, the 12th JSC Guest Student Programme took place from August 01 to October 07. This year, it was co-sponsored by the Centre Européen de Calcul Atomique et Moléculaire (CECAM) and organized in cooperation with the German Research School for Simulation Sciences (GRS). Out of an again significantly increased number of excellent international applicants, 12 students from Germany, Italy, Croatia, Israel, and India joined scientists from JSC, NIC, and GRS for ten weeks. While their original scientific areas ranged from Physics, Mathematics, Chemistry, and Computer Science, their projects covered simulation, visualization, data mining, and algorithm development as well as hardware porting studies.

Besides the intensive use of JSC's reliable workhorses JUGENE and JUROPA, the newly installed GPU cluster JUDGE also played an important role this time: three projects were dedicated to GPU-based algorithms. Therefore, the traditional courses on distributed and shared memory parallel programming and performance optimization were complemented by a workshop on exploiting the capabilities of graphics processing units.

The participants had the opportunity to present and discuss their work with other students and scientists during a concluding colloquium. Additionally, more detailed reports have been prepared and were compiled into a printed JSC publication, which is also available online.

Next year's JSC Guest Student Programme will start on August 06, 2012. It is open to students from the natural sciences, engineering, computer science and mathematics after their Bachelor and before reaching their Master's degree. The application deadline has been set to April 30, 2012. Additional information as well as the reports of previous years can be found on **www.fz-juelich.de/ias/jsc/gsp**



• Mathias Winkel

Jülich Supercomputing Centre

Autumn 2011 • Vol. 9 No. 2 • inSiDE

Activities

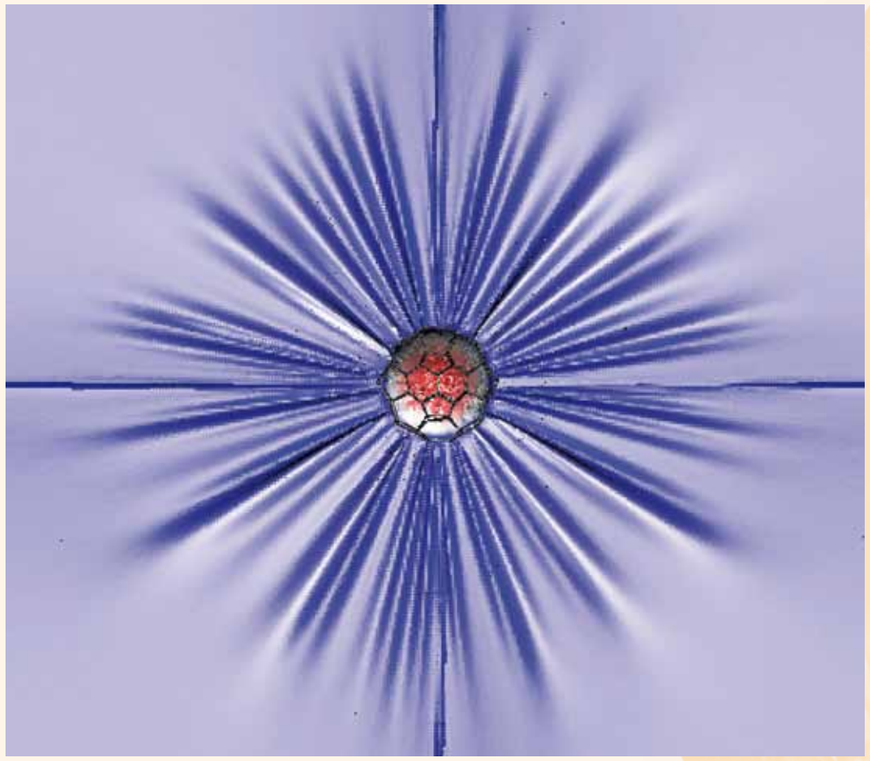
New NIC Research Group at Jülich Supercomputing Centre

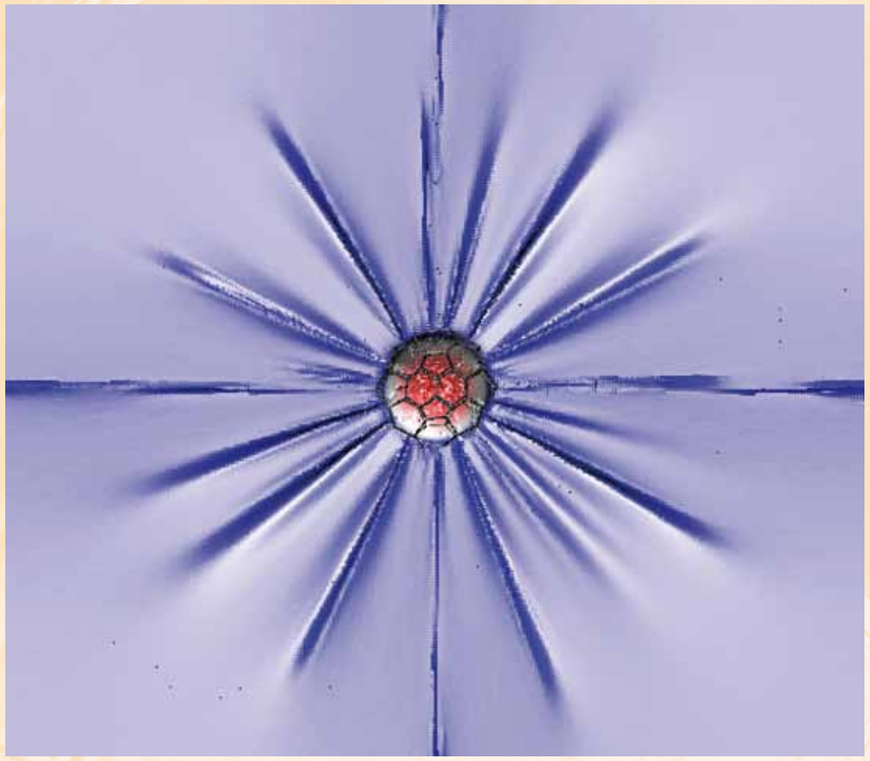
• Martin Müser

Jülich Supercomputing Centre

The partners in the John von Neumann-Institute for Computing (NIC), Jülich Research Centre among them, promote supercomputer-aided research in science and engineering by supporting research groups in selected fields. In April 2011 the NIC research group "Computational Materials Physics" started to work at the Jülich Supercomputing Centre. Its research activities involve atomistic and multiscale simulations of materials in non-equilibrium, in particular in the field of tribology, the analysis of materials under high pressures, and the development of classical force fields, in which the electrostatics of molecular systems is properly described. The group depends heavily on large-scale computational resources and contributes to the development of massively parallel simulation packages. It is lead by Prof. Martin Müser from Saarland University.

It is the main objective of the tribological section of the new NIC research group to develop and to parameterize new constitutive equations that describe interfacial dynamics of solids on macroscopic scales. This goal is achieved, in parts, by unraveling the molecular or meso-scale mechanisms that lead to friction between two rubbing bodies. One example is the study of dissipative forces in externally driven elastic manifolds that wrinkle in response to the presence of a counterbody (see figure). The NIC research group also investigates, if it is possible to alter materials used in optoelectronic memory applications so that they can be switched reversibly between their conducting, crystalline and insulating, amorphous phases through the application of stress. This question is addressed by density functional simulations of materials under compressive pressure and tensile load. A positive result could contribute to the development of fast and energy efficient memory. Lastly, our main objective of accurately describing the electrostatics of molecular systems is to make possible the simulation of redox reactions in terms of semi-empirical force fields. Goals in this line of research are the simulation of the atomistic dynamics in Galvanic elements with and without electrical load as well as the analysis of the molecular mechanisms leading to rubbing-induced electricity.





CECAM – Jülich Summer School on Fast Methods for Long-Range Interactions in Complex Systems

• Godehard Sutmann

Jülich Supercomputing Centre

After the successful tutorial on "Methods for Coulomb Solvers" in 2010, the Jülich-CECAM node organized a Summer School on "Fast Methods for Long-Range Interactions in Complex Systems", which is one of various CECAM events organized by the Jülich node (cmp. www2.fz-juelich.de/jsc/cecam). The school took place from September 12 to 16, 2011 at JSC and consisted of morning lectures and afternoon practical sessions.

About 30 participants from various countries (France, Germany, Greece, Italy, Slovenia, Spain, Sweden, UK, USA) came to Jülich to learn about modern algorithms which efficiently solve the Coulomb problem and reduce the numerical complexity from $O(N^2)$ to $O(N \log(N))$ or $O(N)$. Lecturers from Universities of Bielefeld, Chemnitz, Mainz, Stuttgart, Wuppertal and the Forschungszentrum Jülich presented state-of-the-art methods, algorithms and implementations of various approaches to tackle the long-range interactions in many-particle systems.

The motivation for organizing this Summer School arose from the BMBF funded network project ScaFaCoS (Scalable Fast Coulomb Solver), which aims at developing a scalable library for various fast methods solving the long-range interactions between particles in complex systems. The spectrum of presentations ranged from simple cut-off methods to Fourier-based methods (P3M), hierarchical tree methods, multigrid techniques, the fast multipole method (FMM) and a grid-based Maxwell solver. In addition an introduction into parallel sorting methods was given, which are required by various hierarchical implementations. For each method, emphasis was given to the theoretical foundation, the error control of approximations and parallelization algorithms.

To get participants acquainted with parallel computing, the first day included a special introduction to MPI followed by a hands-on programming session. Practical afternoon sessions focused on an introduction to the parallel interface of the ScaFaCoS library. This was included into various test programs, thereby integrating methods like multigrid, tree-methods or the fast multipole method. Participants were also encouraged to integrate the library into their own codes.

During a poster session, participants were encouraged to present their own research in short oral presentations and to discuss methods and results at the poster.

NIC Symposium 2012

• Walter Nadler

Jülich Supercomputing Centre

The John von Neumann Institute for Computing (NIC), formed by the contract partners Forschungszentrum Jülich, Deutsches Elektronen-Synchrotron DESY and GSI Helmholtzzentrum für Schwerionenforschung, will hold its sixth Symposium at Forschungszentrum Jülich on February 7 to 8, 2012. Every two years the NIC organizes this symposium to report on the activities and results obtained by the NIC projects and research groups in the last two years. It will provide an overview on a broad spectrum of computational science, with projects from Astrophysics, Chemistry, Elementary Particle Physics, Materials Science, Condensed Matter Physics, Polymers, Earth and Environment, Computer Science, and further topical research areas. To accompany the conference, an extended proceedings volume will be published.

Winter School on Hierarchical Methods for Dynamics in Complex Molecular Systems

• Johannes Grotendorst

Jülich Supercomputing Centre

Jülich will continue its successful series of winter schools. From March 05 to 09, 2012, more than 20 renowned scientists will present lectures on the topic "Hierarchical Methods for Dynamics in Complex Molecular Systems" at the rotunda of the Jülich Supercomputing Centre. This winter school will be supported by the Centre Européen de Calcul Atomique et Moléculaire (CECAM).

The school has a daily stratification pattern starting with dynamics within the realm of Materials Science with a focus on slow processes which nevertheless requires most detailed input at the level of electronic structure and interatomic potentials. In Biomolecular Science one challenge is the concurrent handling of an electronic structure based description of a "hot spot" within an enzyme with a computationally efficient treatment of the protein environment in terms of parameterized interactions. Accelerated sampling is a key issue whenever both slow and fast motion is relevant such as metadynamics, force probe molecular dynamics or nonequilibrium dynamics using fluctuation theorems. Finally, getting rid of atoms and molecules but still keeping a particle perspective is achieved by coarse-graining procedures. In Soft Matter and Life Science, the dynamics is often governed by the hydrodynamics of the solvent. A particular challenge is here to bridge the large length- and time-scale gap between the small solvent molecules and the embedded macromolecules or macromolecular assemblies (polymers, colloids, vesicles, cells). Therefore, several mesoscale simulation approaches have been developed recently, which rely on a strong simplification of the microscopic dynamics with a simultaneous implementation of conservation laws on mass, momentum and energy. Here, Lattice Boltzmann, Dissipative Particle Dynamics and Multi-Particle Collision Dynamics are most prominent.

Last but not least most efficient implementations on current-day hardware is a must, which requires facing parallel computing issues or using GPUs when designing de novo software and porting well-established numerical codes or numerical methods like the multigrid method onto new architectures.

This winter school is suited for highly motivated PhD students and PostDocs. Applications for participation can be sent in until end of December 2011. Based on the required application documents about 50 participants will be selected by the organizers. Details about the school and the application process can be found at

www.fz-juelich.de/ias/jsc/events/wshd

Creating a comfortable Climate for SuperMUC

• Ludger Palm

Leibniz Supercomputing Centre

Since our report in the last issue of inSiDE on the topping out ceremony for the new LRZ buildings a lot of progress has been made. After finalization of the construction, the technical infrastructure for water and air conditioning was installed. Every floor of the computer building contains parts of the tubing and hydraulic engineering for the hot water cooling that will enable SuperMUC to be presumably one of the most energy efficient supercomputers of the world in 2012.

When the LRZ was founded in 1962, 436 m^2 floor space were available in the Richard-Wagner-Straße 18 in the center of Munich. After moving to the building Barer Straße 21 in 1970, this increased to 3,228 m^2. 2006 LRZ moved to the new buildings on the Garching research campus north of Munich where 13,619 m^2 can be used. After the extension of the buildings LRZ now has more than 22,000 m^2 for staff and equipment.

The buildings and the air conditioning will be ready for operation in September 2011, after this the installation procedure for SuperMUC, the new 3 Petaflops system at LRZ, will begin.

Figure 1: A view from the roof of the computer building into Bavaria's white-blue sky.

HLRS Scientific Tutorials and Workshop Report and Outlook

HLRS has installed a first Cray XE6 system with AMD Interlagos processors a performance of 1 PFlop/s. We strongly encourage you to port your applications to the new architecture as early as possible. To support such effort we invite current and future users to participate in special **Cray XE6 Optimization Workshops**. With these courses, we will give all necessary information to move applications from the current NEC SX-9, the Nehalem cluster, or other systems to the upcoming Petaflop system. The Cray XE6 will provide our users with a new level of performance. To harvest this potential will require all our efforts. We are looking forward to working with our users on these opportunities. These 1 to 3-day courses in cooperation with Cray were 2011 at Febuary 02 - 04 and April 19, and November 02 - 04. Next courses in 2012 will be scheduled soon.

Additionally, the Virtual Institute for High-Productivity Supercomputing (VI-HPS) offered a workshop for tuning and code optimization specifically on our Cray XE6 pre-installation system. This **7th VI-HPS Tuning Workshop** took place at HLRS on March 28-30.

One of the flagships of our courses is the week on **Iterative Solvers and Parallelization**. Prof. A. Meister and Prof. B. Fischer teach basics and details on Krylov Subspace Methods. Lecturers from HLRS give lessons on distributed memory parallelization with the Message Passing Interface (MPI) and shared memory multi-threading with OpenMP. This course will be presented twice, in March 2012 at HLRS in Stuttgart and September 2012 at LRZ.

Another highlight is the **Introduction to Computational Fluid Dynamics**. This course was initiated at HLRS by Dr.-Ing. Sabine Roller. She is now a professor at the German Research School at RWTH Aachen, and with this, the course was held the first time there. It is again scheduled 2012, March 26-30 in Stuttgart and Sep. 17-21 in Aachen. The emphasis is placed on explicit finite volume methods for the compressible Euler equations. Moreover outlooks on implicit methods, the extension to the Navier-Stokes equations and turbulence modeling are given. Additional topics are classical numerical methods for the solution of the incompressible Navier-Stokes equations, aeroacoustics and high order numerical methods for the solution of systems of partial differential equations.

PRACE Summer Schools and SC Tutorials in 2011

Hybrid Message Passing + Shared Memory Programming. Rolf Rabenseifner, Pekka Manninen, et al.; lecture with practical, at PRACE Summer School: Taking the Most Out of Supercomputers, CSC, Espoo, Finland, Aug. 29 - Sep. 01, 2011.

Hybrid MPI and OpenMP Parallel Programming. Gabriele Jost, et al.; at DEISA/PRACE Spring School: Tools and Techniques for Extreme Scalability, EPCC, Edinburgh, Scotland (GB), March 29-31, 2011.

Rolf Rabenseifner, Georg Hager, Gabriele Jost: **Hybrid MPI and OpenMP Parallel Programming**. Half-day Tutorial No. S-01 at Super Computing 2011, SC11, Seattle, WA, USA, November 12-18, 2011.

Alice E. Koniges, Katherine Yelick, Rolf Rabenseifner, Reinhold Bader, David Eder: **Introduction to PGAS (UPC and CAF) and Hybrid for Multicore Programming**. Full-day Tutorial No. M-08 at Super Computing 2011, SC11, Seattle, WA, USA, November 12-18, 2011.

Our general course on parallelization, the **Parallel Programming Workshop**, Oct. 08-12, 2012 at HLRS, will have 3 parts: The first two days of this course are dedicated to parallelization with the Message Passing Interface (MPI). Shared memory multithreading is taught on the third day, and in the last two days, advanced topics are discussed. As in all courses, hands-on sessions (in C and Fortran) will allow users to immediately test and understand the parallelization methods. The course language is English.

Several three and four day-courses on **MPI & OpenMP** will be presented at different locations all over the year.

We also continue our series of **Fortran for Scientific Computing** in March and July 2012, mainly visited by PhD students from Stuttgart and other universities in Germany to learn not only the basics of programming, but also to get an insight on the principles of developing high-performance applications with Fortran.

With **Unified Parallel C (UPC) and Co-Array Fortran (CAF)** in July and December 2012, the participants will get an introduction of Partitioned Global Address Space (PGAS) languages.

Our series on **"GPU Programming using CUDA"** will also continue in July and December 2012.

In cooperation with Dr. Georg Hager from the RRZE in Erlangen and Dr. Gabriele Jost from TACC, the HLRS also continues with contributions on hybrid MPI & OpenMP programming at **PRACE summer schools and tutorials** at conferences; see the box on the left page.

In the table, you can find the whole HLRS series of training courses in 2012. They are organized at HLRS and also at several other HPC institutions: LRZ Garching, NIC/ZAM (FZ Jülich), ZIH (TU Dresden), TUHH (Hamburg-Harburg), and GRS/RWTH (Aachen).

2012 – Workshop Announcements
Scientific Conferences and Workshops at HLRS
14th Teraflop Workshop (December 05-07, 2011)
11th HLRS/hww Workshop on Scalable Global Parallel File Systems (April 2012)
6th ZIH+HLRS Parallel Tools Workshop (date and location not yet fixed)
High Performance Computing in Science and Engineering - The 15th Results and Review Workshop of the HPC Center Stuttgart (October 2012)
IDC International HPC User Forum (October 2012)
Parallel Programming Workshops: Training in Parallel Programming and CFD
Cray XE6 Optimization Workshops (HLRS, will be scheduled soon)
Parallel Programming and Parallel Tools (TU Dresden, ZIH, February 13-16)
Iterative Linear Solvers and Parallelization (HLRS, March 12-16)
Introduction to Computational Fluid Dynamics (HLRS, March 26-30)
GPU Programming using CUDA (HLRS, July 02-04, and December 10-12)
Unified Parallel C (UPC) and Co-Array Fortran (CAF) (HLRS, July 05-06, and Dec. 13-14)
Parallel Programming with MPI & OpenMP (TU Hamburg-Harburg, July 30 - August 01)
Iterative Linear Solvers and Parallelization (LRZ, Garching, September 10-14)
Introduction to Computational Fluid Dynamics (GRS / RWTH Aachen, September 17-21)
Message Passing Interface (MPI) for Beginners (HLRS, October 08-09
Shared Memory Parallelization with OpenMP (HLRS, October 10)
Advanced Topics in Parallel Programming (HLRS, October 11-12)
Parallel Programming with MPI & OpenMP (FZ Jülich, JSC, November 26-28)
Training in Programming Languages at HLRS
Fortran for Scientific Computing (March 19-23, and July 09-13)
URLs: www.hlrs.de/events/
www.hlrs.de/organization/sos/par/services/training/course-list/
https://fs.hlrs.de/projects/par/events/2012/parallel_prog_2012/

High-Performance Computing Courses and Tutorials

GPGPU Programming

Date and Location
October 10 - 12, 2011
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

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

Please Note: this is a first draft of the programme which may change before the workshop:
Day 1: Introduction to NVIDIA CUDA, CUDA Parallel Programming and hands-on (LRZ)
Day 2: HMPP (LRZ), Advanced CUDA (nVidia)
Day 3: PGI Accelerator compilers, GPU-accelerated Software (LRZ)

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

Web page
www.lrz.de/services/compute/courses

Introduction to the Usage of High Performance Systems, Remote Visualization and Grid Facilities at LRZ

Date and Location
October 20, 2011
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
The National High-End Systems provide nearly ten thousand cores and the Linux cluster systems at LRZ another ten thousand cores to scientists and students. Furthermore, powerful remote visualization facilities are available. Based on hands-on examples, an easy-to-follow introduction to basic Linux usage, specific information on the hard- and software of the LRZ cluster systems, the visualization systems and usage of the grid middleware (Globus Toolkit) is given. Grid certificates can be provided if needed.

Prerequisites
To obtain a grid certificate participants are required to show a valid ID card or passport.

Web page
www.lrz.de/services/compute/courses

Introduction to Molecular Modeling on Supercomputers

Date and Location
October 24 - 26, 2011
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

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

Day 1: Advanced feature of the graphical user interface Maestro (Schrodinger)
Day 2: Remote visualization services, VMD/NAMD (LRZ)
Day 3: Gromacs and Dalton (ScalaLife)

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

Web page
www.lrz.de/services/compute/courses

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

Date and Location
November 24 - 25, 2011
May 21 - 22, 2012

JSC, Research Centre Jülich

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

Web page
www.fz-juelich.de/ias/jsc/events/sc

Parallel Programming with MPI, OpenMP and PETSc

Date and Location
November 28 - 30, 2011
JSC, Research Centre Jülich

Contents
The focus is on programming models MPI, OpenMP, and PETSc. Hands-on sessions (in C and Fortran) will allow users to immediately test and understand the basic constructs of the Message Passing Interface (MPI) and the shared memory directives of OpenMP. This course is organized by JSC in collaboration with HLRS. Presented by Dr. Rolf Rabenseifner, HLRS

Web page
www.fz-juelich.de/ias/jsc/events/mpi

Scientific 3D Animation with Blender

Date and Location
December 01 - 02, 2012
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
The two-day course gives an introduction to the visualization of scientific simulation data using the open source 3D animation package Blender (www.blender.org). The participants will learn how to generate impressive and professional looking still images and animations of the data obtained from their own scientific projects. In the hands-on part of the workshop, the participants will learn all the relevant steps to create stunning 3D animations. Based on a real-world example project using a protein molecule, all steps necessary to produce a 3D animation will be discussed: importing the data, cleaning-up the geometry, assigning materials, illuminating the scene, setting key frames, rendering and post-production. At the end of the workshop, each participant will leave with his own 3D animation of an enzyme.

Prerequisites
Participants should have some basic knowledge in programming. Knowledge of a 3D animation software are helpful but not necessary.

Web page
www.lrz.de/services/compute/courses

GPU Programming

Date and Location
December 05 - 07, 2011
JSC, Research Centre Jülich

Contents
Many-core programming is a very dynamic research area. Many scientific applications have been ported to GPU architectures during the past years. We will give an introduction to CUDA, OpenCL, and multi-GPU programming using examples of increasing complexity. After introducing the basics the focus will be on optimization and tuning of scientific applications.

Web page
www.fz-juelich.de/ias/jsc/events/gpu

Introduction to SuperMUC

Date and Location
December 05 - 09, 2011
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
The course gives an introduction into the usage of the forthcoming Peta-

High-Performance Computing Courses and Tutorials

Flops Supercomputer SuperMUC at LRZ. Focus is on compilers and tools, batch system, monitoring, MPI, OpenMP, system architecture and performance analysis and optimization. Hands-on sessions featuring example applications are presented and will allow users to immediately test and understand the system tools. The course targets new users of the HPC system who will learn how to program and use the system efficiently.

Day 1: Introduction to IBM Software Stack
Day 2: Introduction to Intel Software Stack
Day 3: How to use SuperMIG (LRZ)

Prerequisites
Fortran/C, HPC Programming language, MPI, openMP.

Web page
www.lrz.de/services/compute/courses

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

Dates and Locations
December 08 - 09, 2011
July 05 - 06, 2012

Stuttgart, HLRS

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

Web page
www.hlrs.de/events/

GPU Programming using CUDA

Dates and Locations
December 12 - 14, 2011
July 02 - 04, 2012

Stuttgart, HLRS

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

Web page
www.hlrs.de/events/

Programming with Fortran

Date and Location
February 06 - 10, 2012
09:00 - 18:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
This course is targeted at scientists with little or no knowledge of the Fortran programming language, but need it for participation in projects using a Fortran code base, for development of their own codes, and for getting acquainted with additional tools like debugger and syntax checker as well as handling of compilers and libraries. The language is for the most part treated at the level of the Fortran 95 standard; features from Fortran 2003 are limited to improvements on the elementary level. Advanced Fortran features like object-oriented programming or co-arrays will be covered in a follow-on course in autumn.

To consolidate the lecture material, each day's approximately 4 hours of lecture are complemented by 3 hours of hands-on sessions.

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

Web page
www.lrz.de/services/compute/courses

Parallel Programming with MPI, OpenMP, and Tools

Date and Location
February 13 - 16, 2012
Dresden, ZIH

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

Web page
www.hlrs.de/events/

Parallel Programming of High Performance Systems

Date and Location
March 05 - 09, 2012
09:00 - 18:00

RRZE Building,
University Campus Erlangen,
Martensstr. 1

LRZ Building,
University Campus Garching,
near Munich. (via video conference if there is sufficient interest)

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

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

Day 1:
- Introduction to High Performance Computing
- HPC systems at LRZ and in Germany
- Technical aspects of software engineering: development process, usage of libraries, memory management, tools for code documentation, development environments
- Tools: Using Secure Shell, User interfaces, Source code versioning systems and Make

Day 2:
- Basic parallel programming with MPI and OpenMP
- Processor Architectures: Register, Cache, Locality, Performance metrics

Day 3:
- Principles of code optimization (unrolling, blocking, dependencies, C++ issues, bandwidth issues, performance projections)
- Using performance libraries
- Advanced MPI programming

Day 4:
- Parallel algorithms: data parallelism, domain decomposition, task parallelism, master-worker, granularity, load balancing, scalability models
- Advanced OpenMP programming
- Parallel Architectures: multi-core, multi-socket, ccNUMA, cache coherence and affinity, tools for handling memory affinity

Day 5:
- Architecture-specific optimization strategies: compiler switches, avoiding cache thrashing, exploiting SIMD capabilities
- MPI exercises

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

Web page
www.lrz.de/services/compute/courses

Iterative Linear Solvers and Parallelization

Date and Location
March 12 - 16, 2012
Stuttgart, HLRS

September 10 - 14, 2012
Garching, LRZ

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

High-Performance Computing Courses and Tutorials

and the shared memory directives of OpenMP. This course is organized by Uni. Kassel, HLRS, and IAG.

Web page
www.hlrs.de/events/

Advanced Topics in High Performance Computing

Date and Location
March 19 - 22, 2012
09:00 - 18:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
In this add-on course to the parallel programming course special topics are treated in more depth, in particular performance analysis, I/O and PGAS concepts. It is provided in collaboration of Erlangen Regional Computing Centre (RRZE) and LRZ. Each day is comprised of approximately 5 hours of lectures and 2 hours of hands-on sessions.

Day 1
- Intel Tracing Tools: MPI tracing and correctness checking
- Intel Threading tools for OpenMP correctness checking and profiling
- Profiling on SGI Altix systems: histx and lipfpm
- Introduction to Scalasca

Day 2:
- Parallel application performance analysis with Scalasca
- Parallel input/output with MPI-IO

Day 3:
- I/O tuning on high performance file systems
- Portability of binary files, big/little endian issues
- Using I/O libraries (pNetCDF, HDF5)

Day 4:
Introduction to the PGAS languages Co-Array Fortran and UPC: Partitioned global address space languages have emerged as an alternative to other parallel programming models, promising a shorter development cycle due to improved programmability while keeping the performance level on par with MPI. This course introduces the parallel facilities integrated into the Fortran language (co-arrays) and the C language (unified parallel C), respectively. A hands-on session allows to experiment with the new concepts, using prototype implementations on LRZ's HPC systems.

Prerequisites
Good MPI and OpenMP knowledge as presented in the course "Parallel programming of High Performance Systems" (see above).

Web page
www.lrz.de/services/compute/courses

Fortran for Scientific Computing

Date and Location
March 19 - 23, 2012
July 09 - 13, 2012

Stuttgart, HLRS

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

Web page
www.hlrs.de/events/

Introduction to SuperMUC

Date and Location
March 22 - 23, 2012
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

Contents
The course gives an introduction into the usage of the forthcoming Peta-Flops Supercomputer SuperMUC at LRZ. Focus is on compilers and tools, batch system, monitoring, MPI, OpenMP, system architecture and performance analysis and optimization. Hands-on sessions featuring example applications are presented and will allow users to immediately test and understand the system tools. The course targets new users of the HPC system who will learn how to program and use the system efficiently.

Day 1:
Introduction to IBM Software Stack

Day 2:
Introduction to Intel Software Stack

Day 3:
How to use SuperMIG (LRZ)

Prerequisites
Fortran/C, HPC Programming language, MPI, openMP.

Web page
www.lrz.de/services/compute/courses

Introduction to Computational Fluids Dynamics

Date and Location
March 26 - 30, 2012
Stuttgart, HLRS

Contents
Numerical methods to solve the equations of Fluid Dynamics are presented. The main focus is on explicit Finite Volume schemes for the compressible Euler equations. Hands-on sessions will manifest the content of the lectures. Participants will learn to implement the algorithms, but also to apply existing software and to interpret the solutions correctly. Methods and problems of parallelization are discussed. This course is based on a lecture and practical awarded with the "Landeslehrpreis Baden-Württemberg 2003" and organized by HLRS, IAG, and University of Kassel.

Web page
www.hlrs.de/events/

Parallel I/O and Portable Data Formats

Date and Location
March 28 - 30, 2012
JSC, Research Centre Jülich

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

Web page
www.fz-juelich.de/ias/jsc/events/parallelio

GPGPU Programming

Date and Location
April 16 - 18, 2012
10:00 - 17:00

LRZ Building,
University Campus Garching,
near Munich, Boltzmannstr. 1

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

Please Note: this is a first draft of the programme which may change before the Workshop:
Day 1: Introduction to NVIDIA CUDA, CUDA Parallel Programming and hands-on (LRZ)
Day 2: HMPP (LRZ), Advanced CUDA (nVidia)
Day 3: PGI Accelerator compilers, GPU-accelerated Software (LRZ)

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

Web page
www.lrz.de/services/compute/courses

Guest Student Programme: Education in Scientific Computing

Date and Location
August 06 - October 12, 2012
JSC, Research Centre Jülich

Contents
Guest Student Programme “Scientific Computing” to support education and training in the fields of supercomputing. Application deadline is April 30, 2012. Students of Computational Sciences, Computer Science and Mathematics can work 10 weeks in close collaboration with a local scientific host on a subject in their field.

Web page
www.fz-juelich.de/ias/jsc/events/gsp

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

Publishers

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

Editor & Design

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

Authors

Jaakko Akola
jaakko.akola@tut.fi
Reinhold Bader
Reinhold.Bader@lrz.de
Steffen Brinkman
brinkmann@hlrs.de
Bernd Brügmann
bernd.bruegmann@uni-jena.de
Paul C. Clark
pcc@ita.uni-heidelberg.de
Bastian Demuth
b.demuth@fz-juelich.de
Gustavo Dopcke
gustavotche@gmail.com
Norbert Eicker
n.eicker@fz-juelich.de
Thomas Fecher
fecher@tum.de
Colin W. Glass
glass@hlrs.de
Simon C. O. Glover
glover@uni-heidelberg.de
Stefan Gottlöber
sgottloeber@aip.de
José Gracia
gracia@hlrs.de
Johannes Grotendorst
j.grotendorst@fz-juelich.de
Thomas Gruber
Thomas.Gruber@bv.tu-muenchen.de
Wolfgang Gürich
w.guerich@fz-juelich.de
Yehuda Hoffman
hoffman@huji.ac.il
Frank Holzäpfel
Frank.Holzaepfel@dlr.de
Robert O. Jones
r.jones@fz-juelich.de
Rainer Keller
keller@hlrs.de
Alexander Kipp
alexander.kipp@hlrs.de
Ralf S. Klessen
klessen@uni-heidelberg.de
Takashi Misaka
Takashi.Misaka@dlr.de
Martin Müser
m.mueser@fz-juelich.de
Walter Nadler
w.nadler@fz-juelich.de
Christoph Niethammer
niethammer@hlrs.de
Roland Pail
pail@bv.tu-muenchen.de
Ludger Palm
ludger.palm@lrz.de
Mark Parsons
m.parsons@epcc.ed.ac.uk
Rolf Rabenseifner
rabenseifner@hlrs.de
Michael Resch
resch@hlrs.de
Morris Riedel
m.riedel@fz-juelich.de
Mathilde Romberg
m.romberg@fz-juelich.de
Jürgen Schnack
jschnack@uni-bielefeld.de
Anton Stephan
Anton.Stephan@dlr.de
Estela Suarez
e.suarez@fz-juelich.de
Godehard Sutmann
g.sutmann@fz-juelich.de
Stefan Wesner
wesner@hlrs.de
Mathias Winkel
m.winkel@fz-juelich.de
Klaus Wolkersdorfer
k.wolkersdorfer@fz-juelich.de
Uwe Wössner
woessner@hlrs.de
Gustavo Yepes
gustavo.yepes@uam.es

If you would like to receive inSiDE regularly, send an email with your postal address to klank@hlrs.de
Web page: http://inside.hlrs.de/