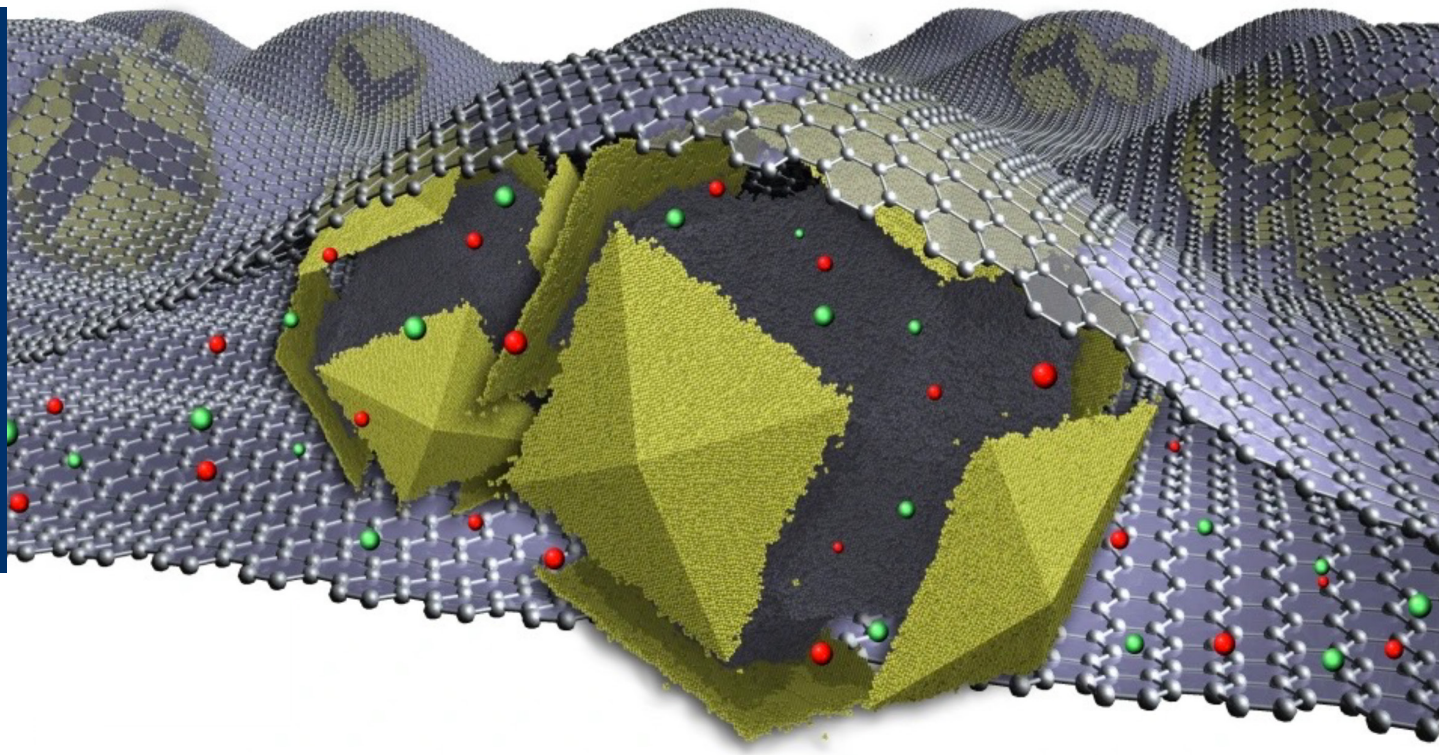
FAU

Friedrich-Alexander-Universität
Erlangen-Nürnberg

nhr.fau.de



Erlangen National High Performance Computing Center (NHR@FAU)

Annual Report 2022

NHR FAU

Annual Report 2022

Erlangen National High Performance Computing Center (NHR@FAU)
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
c/o Erlangen Regional Computing Center (RRZE)

Martensstraße 1
91058 Erlangen
hpc-support@fau.de
https://nhr.fau.de

Contents

1 NHR@FAU at a glance 9
1.1 Organization 9
1.2 Executive board, divisions and steering committee 10
1.3 Services 10

2 Scientific Steering Committee 13
2.1 Compute time applications 13
2.2 Project statistics 14
2.3 Resource usage 15
2.4 Projects active at NHR@FAU 16

3 Systems & Services 17
3.1 Compute resources 19
3.2 Storage resources (Tier-3 with dedicated extensions for NHR) 21
3.3 System availability and usage 2022 22

4 Training & Support 23
4.1 Training activities 24
4.2 NHR PerfLab Seminar series 28
4.3 Projects 29
4.4 NHR Graduate School 30
4.5 Dissemination 30

5 Software & Tools 33
5.1 Tool development 34
5.2 Services 37
5.3 Talks 38

6 Research 39
6.1 Performance modeling and performance engineering 39
6.2 Performance tools 42
6.3 Building blocks for sparse linear algebra and stencil solvers 42
6.4 Software engineering for HPC and data analytics 43
6.5 Projects 45
6.6 PhD defenses 48
6.7 Publications 48
6.8 Talks 50

7 NHR@FAU application focus 51
7.1 Atomistic Simulation Center (ASC) 51
7.2 NHR@FAU Project Support by liason scientists 52
7.3 Activities of liason scientists 55

8 Appendix 71
8.1 Publications which used NHR@FAU Resources 71
8.2 Compute Projects 76

Preface

Dear Reader,

Prof. Dr. Gerhard Wellein,
NHR@FAU Director

we highly appreciate your interest in the 2022 Annual Report of the Erlangen National High Performance Computing Center (NHR@FAU). In 2021, the foundations of our new center were laid, while in 2022, we were able to bring it to full operational capabilities on a national level. Some exemplary cornerstone developments in 2022 include: full operation of two Top500-class supercomputers, implementation of a nation-wide peer review application process, the release of our HPC portal for easy compute time project management, broadening of our application and user support, and substantial extension of our training program. Most importantly, our team helped numerous scientists to advance their research. In 2022, more than 65 research groups from all over Germany used our NHR resources. In addition, more than 140 groups from FAU and other universities accessed our local offerings for small/mid-sized compute projects.

NHR@FAU is one of nine national HPC centers at German universities (NHR centers). It continues HPC services formerly provided by RRZE to FAU's and northern Bavarian universities' researchers. As an NHR center, we provide our services and expertise to scientists from all German universities and are a member of the NHR Alliance (Verein für Nationales Hochleistungsrechnen e.V.). Within the NHR Alliance, NHR@FAU leads the joint performance engineering (PE) activities and provides special application expertise in the field of atomistic simulations.

It is a major objective of NHR@FAU to cover HPC across all involved fields, including HPC infrastructure design and operation, user support from level one to in-depth collaborations with developers and application scientists, HPC training, HPC research, and application expertise. To this end, we are organized in four divisions: *Systems & Services*, *Training & Support*, *Software & Tools*, and *Research*. The research-oriented structure of NHR@FAU is strengthened by a group of liaison scientists. Further, in

2022, a scientific steering committee has been established which has set up a review and allocation procedure for NHR compute time applications. Thus, we were able to also accept large-scale (LS) compute time projects within the first NHR-wide call for LS projects in Q2/2022. This annual report elaborates on all major activities of NHR@FAU in detail; however, I would like to briefly highlight some of them upfront.

Our *Systems & Services* division continued its efforts to install the new HPC systems *Alex* (GPGPU module) and *Fritz* (CPU module) and make them operational in the first half of 2022. After stability tests and an extensive early user phase, both systems entered the TOP500 lists of 2022. Initially with restricted configuration, *Fritz* achieved rank 323 with a LINPACK value of 2.2 PF/s in June and jumped to rank 151 with 3.6 PF/s in November, running its final full configuration. The NVIDIA A100 part of *Alex* entered the list in June with 2.9 PF/s LINPACK Performance on rank 184 and was ranked 16th on the Green500 list. It moved up to rank 174 in the TOP500 in November after a hardware upgrade by approximately 10%. In particular for *Fritz*, this performance level comes at a high power consumption, so energy-efficient cooling is mandatory. Here, our team was able to reduce the cooling overhead to as low as 10% by optimizing the parameters of the warm-water cooling infrastructure.

Our *Training & Support* division has been working on returning to on-site or hybrid formats for the training events, seminars, and the HPC Café. We are prepared to continue this mixture of events into the future. In addition, our training program has been extended in 2022 with courses on OpenMP, CUDA (C/C++ and Python variants), and GPGPU performance analysis tutorials. Also, we have implemented an application-specific five-day course on GROMACS, a molecular dynamics package used by many of our customers. As regular operation started on *Fritz* and *Alex*, we saw a steep increase in the number of support requests. Particularly in the field of molecular dynamics (MD) simulations, a broad range of support services has been established and was immediately accepted by our users: from basic advice on optimized settings of runtime parameters to methodological project support involving the expertise of our *liaison scientists*. As a consequence of our broad training activities, user support projects, and compute time projects, an editorial team has been established which is responsible for dissemination and event organization.

The *Software & Tools* division develops, maintains, and supports tools and software for HPC. The most prominent product is our popular *LIKWID* tool suite, which is continuously being adapted to new architectures and new features are being developed. Within the NHR Alliance, we also focus on the joint development of cluster-wide and job-specific monitoring solutions. Our *ClusterCockpit* framework has made it into operational use on the new NHR@FAU clusters *Fritz* and *Alex* and also other HPC centers in Germany. Furthermore are our tools a pivotal component of the project *Energy-Efficient HPC* (EE-HPC), which will be funded through the BMBF GreenHPC initiative from 2022 to 2025 and is led by NHR@FAU. Another major topic in 2022 was the implementation and introduction of the *HPC Portal*, which allows easy management and supervision of NHR projects and their assigned resources. As it immediately proved to be an easy-to-use and valuable tool, we plan to move all existing and new Tier-3 HPC projects to the portal as well.

The HPC *Research* division pools scientific activities at FAU related to the HPC focus topics of NHR@FAU. PhD students funded by third-party projects are currently pursuing research in the fields of performance modeling, hardware-efficient sparse solvers, or automatic code generation. Beyond many scientific contributions, our young

researchers achieved a best paper award at the bi-annual *International Conference on Parallel Processing and Applied Mathematics* (PPAM).

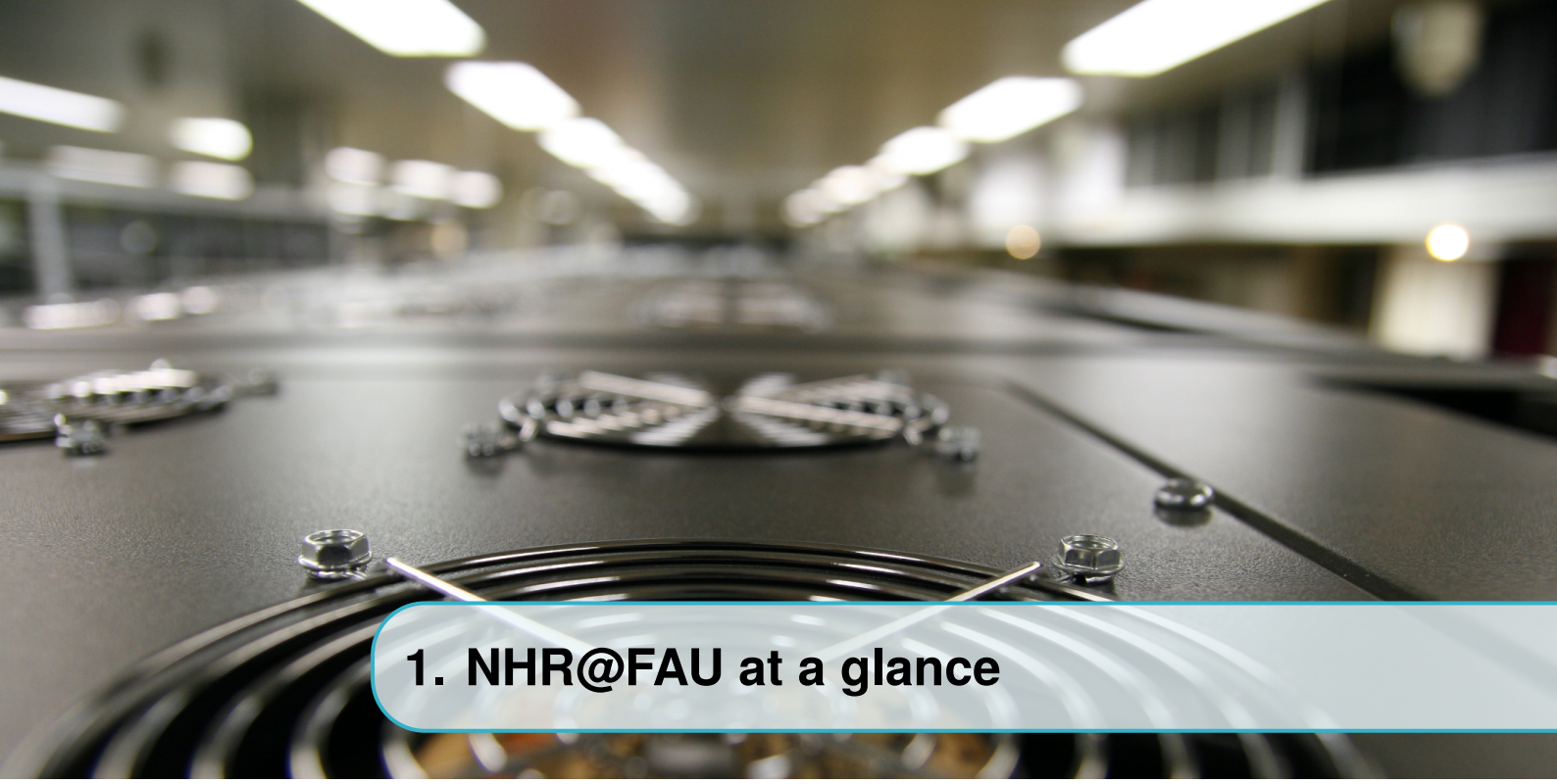
In addition, we are pleased to congratulate two former PhD students on their doctorates: Faisal Shahzad and Julian Hammer successfully defended their PhD theses in 2022. The main topic of Faisal's work on "Efficient Application-level Fault Tolerance Methods for Large Scale HPC Applications" was fault tolerance of massively parallel programs on supercomputers. He developed *CRAFT*, a C++ framework for painless and flexible checkpoint/restart and automatic failure recovery, which makes it very easy for application developers to include fault tolerance features in their code. Julian's work, titled "Design and Implementation of an Automated Performance Modeling Toolkit for Regular Loop Kernels," revolved around the automatic performance prediction of regular loop kernels. He developed and co-developed the tools *Kerncraft*, *OSACA*, *asmbench*, and *pycachesim*, all of which contributed to the overall goal of automating the tedious process of building analytic, first-principles performance models such as the Roofline Model and the ECM Model. We thank Faisal and Julian for their scientific contributions and active involvement in the HPC team over the years and wish them best of luck with new challenges.

We are also looking forward to contribute to two new projects which have been positively evaluated within the BMBF SCALEXA call and will start in 2023: *DAREXA* investigates the use of reduced precision and data compression in plasma simulation, while *StroemungsRaum* addresses the challenges of advanced solvers and parallel-in-time methods for CFD applications on Exascale machines.

Finally, on a national scope, NHR@FAU contributed to various NHR activities such as the first *NHR Graduate School Course Week* and organized the public *NHR PerfLab Seminar* and the *PE Jour fixe* along with NHR centers in Aachen, Paderborn, and Berlin. Furthermore, a well-received workshop was co-organized with NHR centers in Berlin and Paderborn in November 2022, which brought together numerous researchers in the field of atomistic simulations.

In 2022, we have leveled up to provide high-quality HPC services, user support, and adequate modern training activities to researchers all over Germany. These customer-oriented activities are embedded into an active environment of research and tool development, which reaches beyond FAU and NHR. NHR@FAU is on its way and will be an integral part of the national HPC landscape in the years to come. I would like to express my sincere appreciation to all who worked together to make this happen.

Gerhard Wellein
on behalf of the NHR@FAU Executive Board



1. NHR@FAU at a glance

1.1 Organization

High performance computing (HPC) is a key research priority at the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). The use of HPC is central to numerous research activities across all of FAU's faculties; FAU's HPC-related research and teaching/training efforts are recognized internationally, and FAU has continuously developed its HPC infrastructure.

In 2021, these efforts have enabled FAU to establish the Erlangen National High Performance Computing Center (NHR@FAU) as one out of nine national centers for HPC (NHR centers) at German universities. The NHR centers receive federal and state funding, and their services are open to all German universities. NHR@FAU closely collaborates and coordinates with the other NHR centers; they jointly operate the NHR Alliance („NHR-Verein"). Besides fostering coordination and collaboration of the NHR centers, the NHR Alliance also supports the scientific computing and HPC-related research, e.g., by running the *NHR Graduate School*.

NATIONALES HOCHLEISTUNGS RECHNEN

Complementing its national duties, NHR@FAU is also responsible for FAU's local HPC infrastructure and HPC services. These offerings, which have been provided by RRZE before 2021, are open to FAU researchers (and the regional RRZE-supported universities) only. Despite being a separate organization within FAU, the NHR@FAU closely collaborates with RRZE in many ways, most notably in the areas of infrastructure maintenance and development, system administration, networks, and Identity Management. Reflecting its national (Tier-2) and local (Tier-3) HPC service activities, NHR@FAU receives annual budgets from both the NHR program and the FAU.

NHR@FAU is also a pillar of the HPC activities in the State of Bavaria. It complements the Tier-0/1 compute offerings and support services of the Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and hosts one of the two offices of *KONWIHR*, the Competence Network for Scientific High Performance Computing in Bavaria. Within *KONWIHR*, NHR@FAU and LRZ collaborate with scientists at Bavarian universities to exploit the potential of the massive computational power available.

The organizational concept of NHR@FAU covers HPC across all involved fields, including HPC infrastructure design and operation, user support from level one to in-depth collaborations with developers and application scientists, HPC training, HPC research, and application expertise. NHR@FAU aims to be the focal point of FAU's HPC activities to provide high-quality application and user support, up-to-date training and teaching offerings, and efficient and reliable compute capabilities to our users from Erlangen and all over Germany.

1.2 Executive board, divisions and steering committee

The NHR@FAU director holds the professorship for HPC at the Department of Computer Science and was the lead principal investigator (PI) of the NHR@FAU proposal. Together with the other PIs and the technical director of RRZE, the center director forms the NHR@FAU *Executive Board* (see Figure 1.1), which meets regularly to discuss system operations, support and training activities, and budget planning.

NHR@FAU has four divisions: *Systems & Services*, *Training & Support*, *Software & Tools*, and *Research*. The research-oriented structure of the center is further supported by a group of *liaison scientists*, who establish a sustainable link between NHR services and ongoing research at FAU in key focus topics of NHR@FAU.

The scientific steering committee is responsible for the scientific peer-review based resource application and allocation process. It is supported by NHR@FAUadministrative office and closely coordinates with the executive board and systems administration as well as support activities.

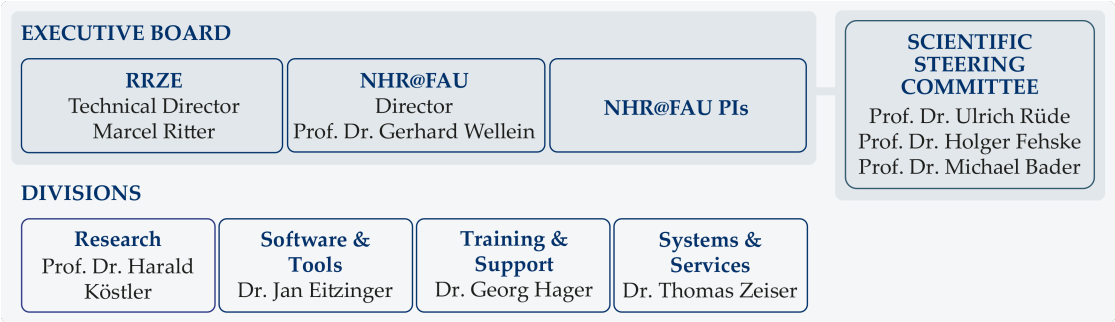


Figure 1.1: Executive Board, Steering Committee, and divisions of NHR@FAU

1.3 Services

The main application topics covered by NHR@FAU are in the areas of atomistic simulations of chemical processes, biomolecular structures, and material properties that are investigated with classical and quantum-mechanical methods; applications used in chemistry, life sciences, materials science, and beyond. FAU's unique computational chemistry, biology, and materials science community can provide comprehensive, application-specific competence and support for the full spectrum of classical and quantum-mechanical atomistic simulation methods and the most popular application codes. User consulting and expert support within NHR@FAU is provided for the full range of electronic structure techniques for static calculations in chemistry, catalysis, and materials science.

The main HPC methodology focus is on node-level performance engineering for CPUs and GPGPUs. The internationally recognized model-driven performance engineering (PE) approach is the foundation of performance-related research, user

consulting, training, performance monitoring, and all aspects of code optimization and parallelization. Research topics comprise novel performance models, PE concepts, efficient CPU/GPGPU implementation strategies, automatic generation of hardware-efficient codes, and micro-architectural analysis and benchmarking. The PE approach is applicable to a broad range of application fields. These activities are complemented by the development and maintenance of performance tools that support node-level PE.

As a core competence, NHR@FAU provides and further develops its internationally recognized expertise in node-level performance engineering for CPUs and GPGPUs, including performance analysis, analytic performance modeling, and resource-based performance monitoring. The pivotal guideline for support actions, services, and training activities is the PE approach. Based on hardware specifications and a performance analysis that contains runtime profiling and inspection of hotspot code, a hypothesis is normally formulated about a potential hardware bottleneck at the main hotspot; this hypothesis should be connected with a performance expectation or even a full-fledged performance model.

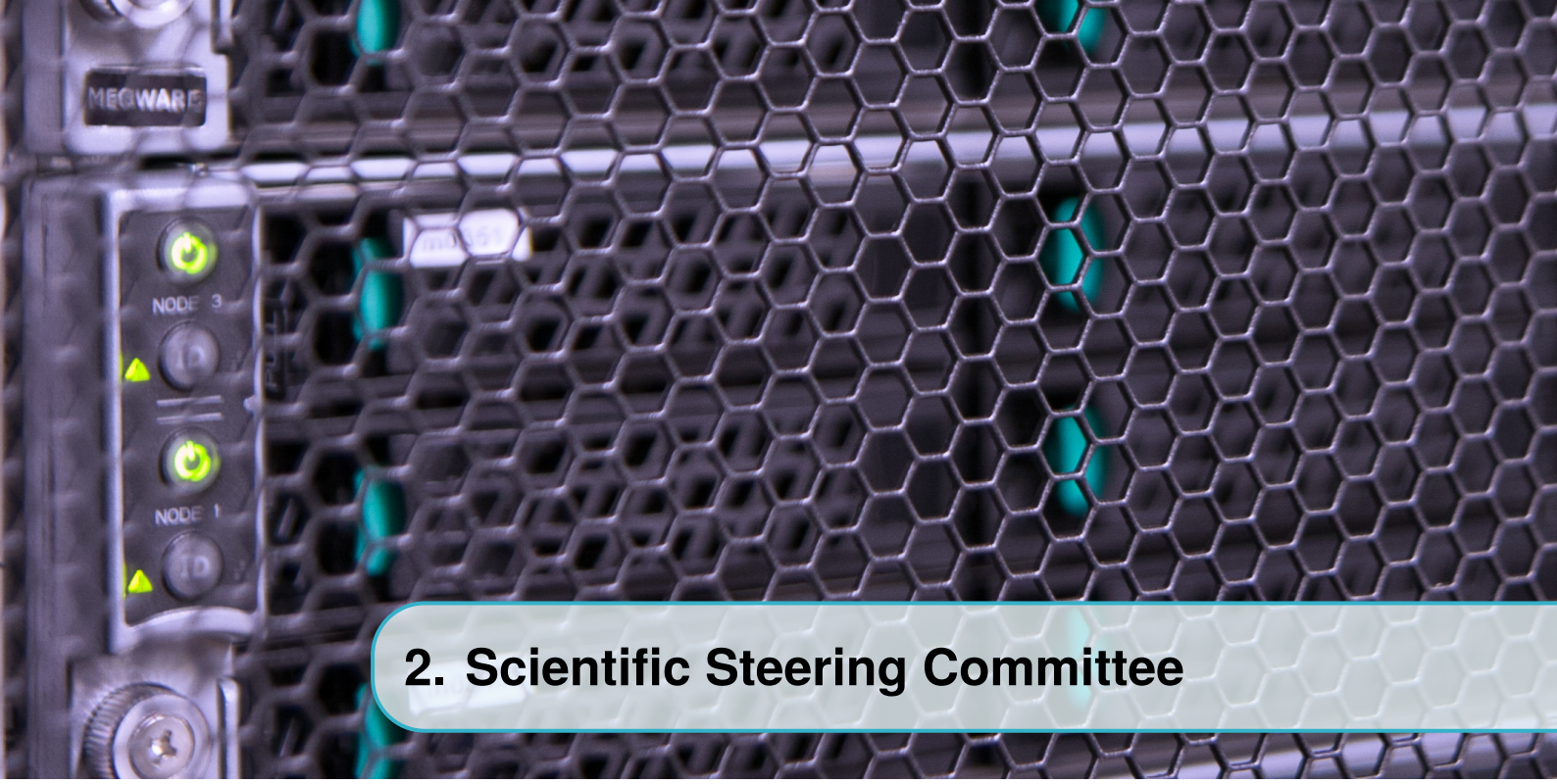
Within the regular support process, PE activities are initiated in different ways: (1) detection of performance issues via job-specific performance monitoring, (2) a user's own performance analysis, (3) a user request for more effective use of a resource allotment, or (4) a user request for performance models of their own code in order to get better insight into performance issues.

The continuous job performance monitoring of relevant resource metrics like bandwidths, flop rates, network traffic, I/O utilization, etc. that was already in place at RRZE has been adapted to NHR@FAU. In addition, the widely used *LIKWID* tool suite is instrumental in PE activities and a pivotal component in monitoring solutions. NHR@FAU continues to develop and adapt *LIKWID* to future CPU and GPGPU architectures. *LIKWID* support is provided as a service to the NHR Alliance and the entire HPC community.

Numerical methods for simulations are focused on scalable iterative solvers, where FAU already had expertise on scalable hardware-efficient sparse solvers, modern matrix-free finite element multigrid methods, and lattice-Boltzmann solvers, including work on adaptive parallel mesh refinement with advanced mesh generation, load balancing, and visualization strategies. The work at NHR@FAU stands out because it integrates the whole method stack ranging from modeling to the hardware-aware and highly optimized implementations for extreme-scale computing on heterogeneous architectures. NHR@FAU has broad experience with building sustainable HPC software and is leading in automatic code generation techniques, where efficient parallel codes are automatically generated from abstract specifications.

Algorithmic user support is offered for these numerical methods, and expert support and training is provided in two directions: (1) manual implementation or automatic generation of hardware-efficient kernels and solvers for CPUs and GPGPUs and (2) defining and implementing maintainable and flexible numerical algorithms and libraries for full-scale applications. An important goal is to foster a software ecosystem of portable and inter-operable applications and libraries for the German HPC community.

In-depth courses on code analysis and optimization are offered to staff from other centers, so-called "train the trainer". Advanced domain-specific HPC training complements the curriculum. In 2022, we started to develop a new tutorial on "Core-Level Performance Engineering" and, as a follow-up to the well-established lecture on "Iterative Solvers and Parallelization," a course on "NLPE for Iterative Solvers." Monthly webinars with crash courses on using application codes and domain-specific tools at NHR@FAU are offered.



2. Scientific Steering Committee

The scientific steering committee of NHR@FAU comprises supercomputing experts from within and outside of FAU. The committee advises the NHR@FAU management and is responsible for important decisions concerning the compute time application process. This includes setting guidelines for the allocation of resources, decisions on accounting formalities, approval of requested resources, and participation in the procurement of hardware and software.

The scientific steering committee currently consists of the chairperson Prof. Dr. Ulrich Rüde from the Department of Computer Science at FAU, Prof. Dr. Michael Bader from the School of Computation, Information, and Technology at the Technical University of Munich, and Prof. emeritus Dr. Holger Fehske from the Institute of Physics at the University of Greifswald.

Scientific steering committee: Prof. Dr. Ulrich Rüde, Prof. Dr. Holger Fehske, Prof. Dr. Michael Bader (f.l.t.r.)

2.1 Compute time applications

Scientists from German universities are able to gain access to NHR@FAU's resources by submitting a scientific application for compute time. So far, applications for compute time are handled via a local, template-based process where the application form is directly sent to NHR@FAU. A connection to the NHR-central application platform *JARDS* (Joint Application, Review and Dispatch Service) is planned for 2023.

Project applications can be submitted in different categories. We differentiate test, normal, and large-scale applications, as specified by the NHR Alliance. The NHR@FAU steering committee decides on the assignment of expert consultants and on the final allocation of approved resources for all project types. For large-scale projects, the NHR@FAU steering committee makes a recommendation, which is then decided upon quarterly by the central NHR steering committee. The scheme shown

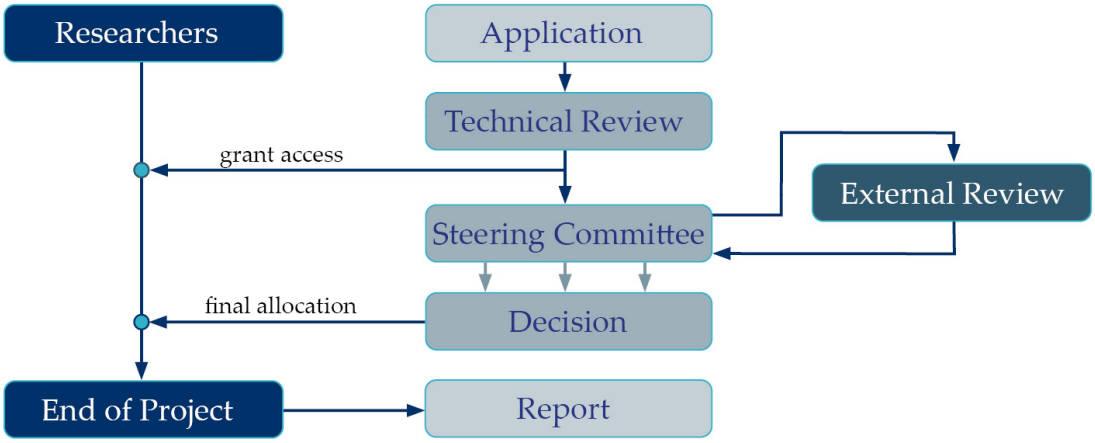


Figure 2.1: Access to HPC Resources

in Figure 2.1 summarizes the general application workflow including the tasks of the steering committee.

Applications for test and normal projects can be submitted at any time. Large-scale project applications are collected on a quarterly basis and are presented in four committee meetings to be decided upon. Test and normal projects are generally able to start using the compute resources within a fortnight and are fully approved within four to six weeks.

Besides the involvement in the application and reviewing process, the local steering committee also supports the NHR@FAU expert advisors in their various activities, such as application porting and optimization. This helps to identify problems in an early phase of the project, far before the first progress report will be available.

2.2 Project statistics

Processing of NHR applications for computing time began in the second quarter of 2022. A total of 29 normal project applications were received and approved by the local steering committee in the reporting year. Two projects were assigned to Tier-3 basic provision due to insufficient resource requirements. Within this project category, a total of 149 million CPU core-hours and 1.25 million GPGPU-hours were awarded.

Of the 35 normal and large-scale project applications approved in the reporting year, 26 came from Bavaria, which corresponds to a share of 74%. A total of 18 applications were received from FAU, as shown in Figure 2.2. The remaining applications were spread among five federal states.

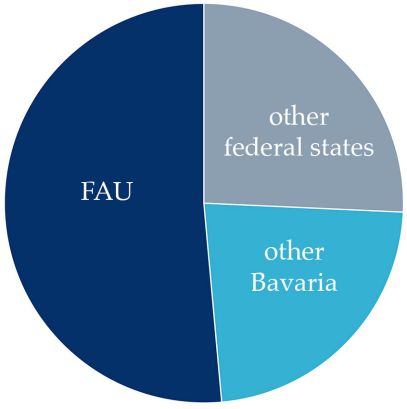


Figure 2.2: Geographical distribution of the approved NHR compute time projects in 2022.

The picture slightly changes when looking at the awarded CPU core hours and GPGPU hours as shown in Figure 2.3 and Figure 2.4, respectively. Less than 50% of the CPU resources are assigned to projects from FAU as the typical project size for non-FAU proposals is larger. These include two large-scale

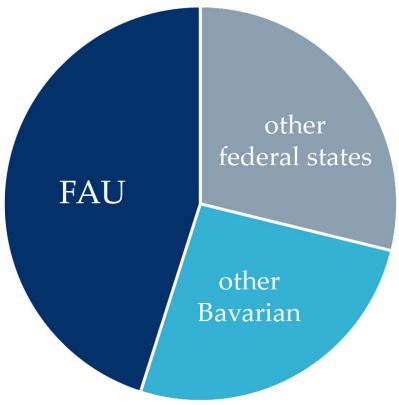


Figure 2.3: Geographical distribution of the 242 million core-hours awarded to NHR projects in 2022

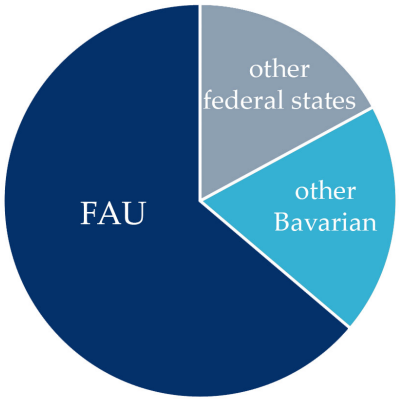


Figure 2.4: Geographical distribution of the 1.72 million GPGPU-hours awarded to NHR projects in 2022

non-FAU proposals from physics with a combined volume of 55 million core-hours. On the GPGPU side, FAU dominates the awarded resources driven by three large-scale projects (in summary 0.42 million GPGPU-hours), which are located in the fields of atomistic simulations and digital linguistics.

For testing or porting user-written applications to our NHR systems and for improving performance, researchers can apply for dedicated projects with limited resources via a simplified procedure. These test projects, which only require a technical review, usually serve to prepare a regular NHR application. In the reporting year, 21 working groups made use of this option.

2.3 Resource usage

Table 2.1 presents an overview of available resources for NHR projects at NHR@FAU and their actual utilization in 2022. When interpreting the utilization, it must be taken into account that in the first two quarters, installation measures were still being carried out on the *Fritz* CPU cluster, and extensive stability tests were conducted on both the *Fritz* and *Alex* systems. Regular NHR@FAU user operations only began in the third quarter.

Table 2.1: Overview of available and used resources for NHR projects at NHR@FAU in 2022.

	CPU cluster *Fritz* [mio h]				GPGPU cluster *Alex* [mio h]			
	compute	core-hours		%		GPGPU-hours		%
2022	nodes	available	used	usage	GPGPUs	available	used	usage
Q1	708	102	1.78	1.7	348	0.70	0.03	4.3
Q2	708	102	9.68	9.5	348	0.70	0.64	91
Q3	708	102	53.0	52	348	0.70	0.32	46
Q4	708	102	84.1	82	348	0.70	0.38	54
total		408	149			2.80	1.37	

It is not surprising that in this early phase of NHR@FAU user operations, half of the applications (24) came from FAU, ten from the rest of Bavaria, and eleven from the federal states. On the other hand, it is good to see that the awareness of the national service NHR@FAU is not concentrated in Bavaria but also reaches the other federal states of Germany.

Figure 2.5 shows the usage of NHR@FAU's systems in 2022, broken down according to the origin of the application and the research discipline indicated by the applicant, respectively. Test projects were not included.

The distribution of the research areas addressed by the applications clearly shows a focus on life sciences, physics, and chemistry. This maps perfectly to the scientific focus of NHR@FAU namely atomistic simulations and their use in these areas. Additionally, there is a significant number of projects applying machine learning techniques, which attests to the current attractiveness of AI-related topics such as image classification, segmentation, and deep learning.

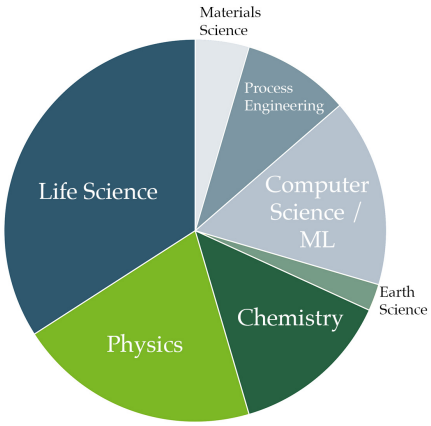
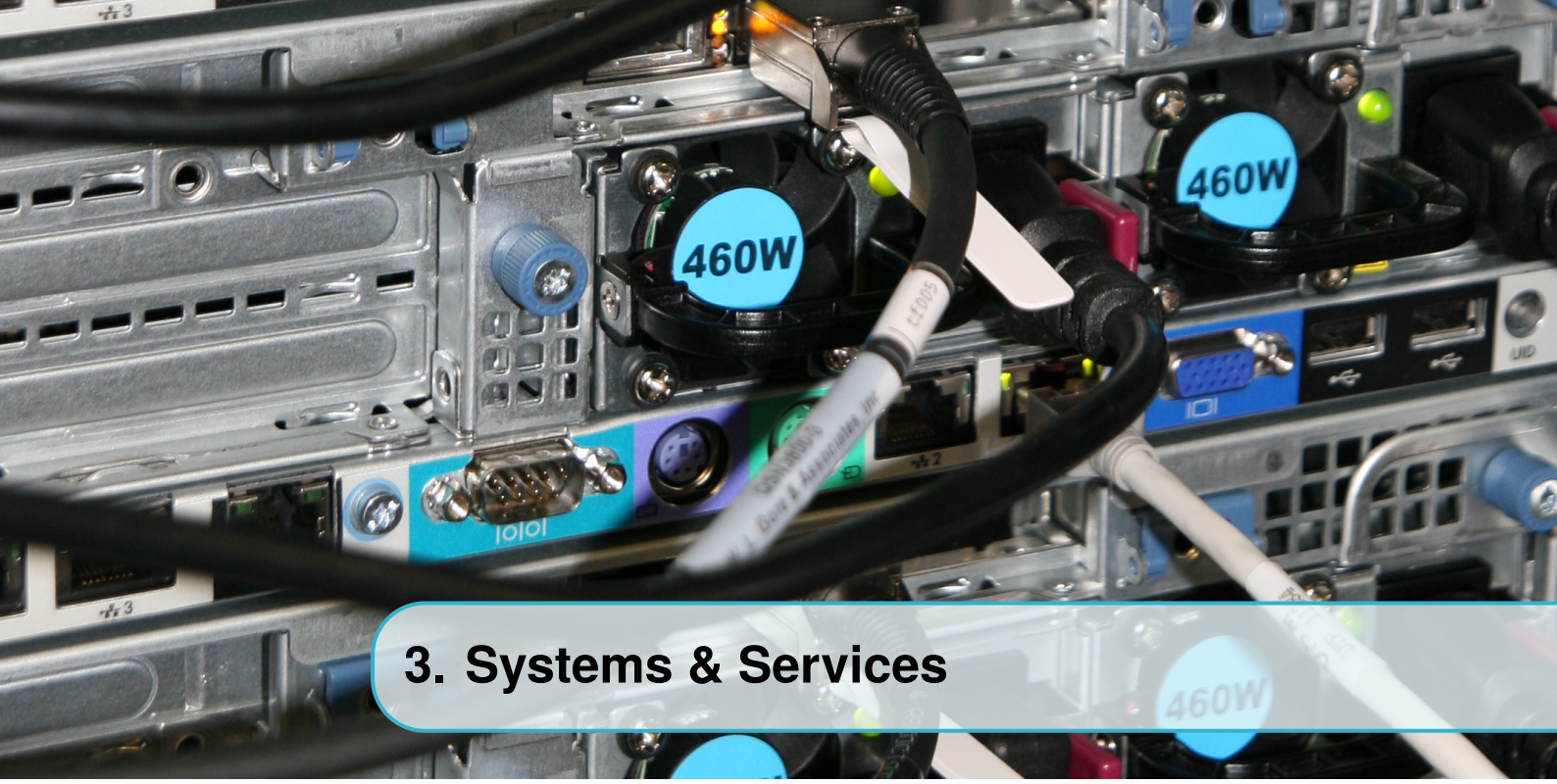


Figure 2.5: Distribution of resource usage of NHR@FAU systems among scientific disciplines in 2022.

2.4 Projects active at NHR@FAU

A list of projects that were actively using the NHR@FAU resources in 2022 is provided in the Appendix (Chapter 8).



3. Systems & Services

Division head: *Dr. Thomas Zeiser*

Dr. Thomas Zeiser

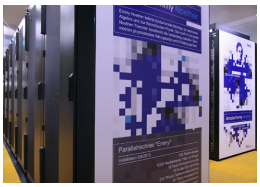
The *Systems & Services* division at NHR@FAU is responsible for operation of the HPC systems, including procurement and software installation, and the required infrastructure, including power and cooling. Many support activities are also carried out or supported by the *Systems & Services* division. The facility team within the *Systems & Services* division not only takes care of the current infrastructure but also drives the planning of a new data center building to be completed before the end of the decade.

RRZE has been operating Linux-based HPC clusters as Tier-3 compute resources for users of FAU and regional customers of RRZE since 2003. Centralization of HPC systems and services has been well accepted and the HPC team grew over time. With the inception of NHR@FAU in January 2021, the next level has been reached and all HPC activities and existing systems were technically taken over by NHR@FAU while still relying on the basic services of RRZE. For the Tier-3 operation of the HPC systems and services, dedicated funding is available. However, synergies of NHR and Tier-3 will be exploited wherever possible—including, but not limited to the joint procurement and operation of systems and services.

The year 2022 was strongly characterized by the start-up of the high-performance computers *Alex* (GPGPU cluster) and *Fritz* (parallel computer). In Erlangen, HPC customers are traditionally closely involved in the design of the HPC hardware strategy and in all steps of the procurement process including early testing of the new systems. An overview of HPC compute systems as well as the available architectures and some notable highlights are shown in Table 3.1.

Large parts of *Fritz* and *Alex* had already been delivered by the end of 2021. However, owing to the general shortage of IT (and other) components, it took until mid-2022 for all remaining pieces, in particular Infiniband network adapters and power distribution units, to be delivered. Thus, the general start-up of both systems took until mid of the year while first (test) users had already been active on both systems since the end of 2021 or early in 2022.

Table 3.1: Overview of the HPC systems of NHR@FAU. Tier-2 systems (NHR) are accessible through the NHR application process, while Tier-3 systems (FAU) are accessible by researchers from FAU without an NHR application.

System	Architecture	Resource	Access	Highlight
Fritz 	Intel Ice Lake HDR100-IB	> 70,000 cores	NHR FAU	#151: Top500 list 11/2022
Alex 	NVIDIA A40/100 partially HDR-IB	584 GPGPUs	NHR FAU	#17: Green500 list 06/2022 #174: Top500 list 11/2022
Meggie 	Intel Broadwell OmniPath100	> 14,000 cores	FAU	major workhorse for FAU/region
Emmy 	Intel IvyBridge QDR-IB	> 11,000 cores	FAU	decommissioned in Q3/2022 after nine years of operation
Woody 	in the meantime mostly Intel IceLake/ throughput	> 2,500 cores	FAU	major renewal in 2022; partially financed by users
TinyFAT 	mostly AMD Rome/ large memory	> 2,400 cores	FAU	all financed by users
TinyGPU 	diverse NVIDIA GPUs/various generations	> 150 GPGPUs	FAU	all financed by users; growing since 2009
Test cluster 	mixed		NHR FAU	≈15 CPU architectures/generations (x86_64, ARM); ≈10 different accelerators (NVIDIA, AMD, NEC)

3.1 Compute resources

3.1.1 *Fritz* parallel computer (NHR & Tier-3)—full operation since mid-2022

https://hpc.fau.de/systems-services/documentation-instructions/clusters/fritz-cluster/

Fritz (system integrator: MEGWARE) is a high-performance compute resource with high-speed interconnect, i.e., a parallel computer. It is intended for moderately-sized, multi-node parallel workloads. *Fritz* is funded by NHR and DFG INST 90/1171-1. Thus, it will not only be the main resource for NHR projects at NHR@FAU but, to a certain extent, also serve as FAU's basic Tier-3 resource for high-end demand.

The compute nodes of *Fritz* have been delivered in November/December 2021 with an additional extension in fall 2022. The configuration of *Fritz* at the end of 2022 is as follows:

- Four front-end nodes with the same CPUs as the compute nodes but 512 GB of RAM and 100 GbE connection to RRZE's network backbone.
- One visualization node with the same CPUs as the compute nodes but 1024 GB of RAM, one NVIDIA A16 GPU, 30 TB of local NVMe SSD storage, and a 100 GbE connection to RRZE's network backbone.
- 992 compute nodes with direct liquid cooling (DLC), each with two Intel Xeon Platinum 8360Y *Ice Lake* CPUs (36 cores per chip) running at a base frequency of 2.4 GHz, 54 MB Shared L3 cache per chip, and 256 GB of DDR4-RAM.
- Lustre-based parallel file system with a capacity of about 3.5 PB and an aggregated parallel I/O bandwidth of > 20 GB/s.
- Blocking fat-tree HDR100 InfiniBand with up to 100 GBit/s bandwidth per link and direction; there are islands with 64 nodes (i.e. 4,608 cores); the blocking factor between islands is 1:4.
- Operating system: AlmaLinux 8 (RHEL clone).
- Batch system: Slurm.

The LINPACK performance of 3.578 PFlop/s on 986 nodes resulted in rank 151 on the Top500 list in November 2022. The direct liquid cooling of the processors and memory of the compute nodes ensures an efficient operation of *Fritz*, significantly lowering the operating costs. The annual PUE (power usage efficiency) of the system was better than 1.1; thus, the electrical overhead for cooling was less than 10%.

3.1.2 *Alex* GPGPU cluster (NHR & Tier-3)—full operation since mid-2022

https://hpc.fau.de/systems-services/documentation-instructions/clusters/alex-cluster/

Alex (system integrator: MEGWARE) is a high-performance compute resource with NVIDIA GPGPU accelerators and partially high-speed interconnect. It is intended for single- and multi-GPGPU workloads, e.g., from molecular dynamics or machine learning. *Alex* is funded by NHR and DFG INST 90/1171-1. Thus, *Alex* serves both as NHR's project resource and as FAU's basic Tier-3 resource.

Most compute nodes of *Alex* were delivered in October/November 2021. Several extensions have been integrated during the year 2022. The configuration of *Alex* at the end of 2022 is as follows:

- Two front-end nodes, each with two AMD EPYC 7713 *Milan* CPUs (64 cores per chip) running at 2.0 GHz with 256 MB Shared L3 cache per chip, 512 GB of RAM, and 100 GbE connection to RRZE's network backbone but no GPGPUs.
- 20 GPGPU nodes, each with two AMD EPYC 7713 *Milan* CPUs (64 cores per chip) running at 2.0 GHz with 256 MB Shared L3 cache per chip, 1,024 GB of DDR4-RAM, eight NVIDIA A100 (each 40 GB HBM2 @ 1,555 GB/s; HGX board with NVLink; 9.7 TFlop/s in FP64 or 19.5 TFlop/s in FP32), two HDR200 InfiniBand HCAs, 25 GbE, and 14 TB of local NVMe SSDs.
- 15 GPGPU nodes, each with two AMD EPYC 7713 *Milan* CPUs (64 cores per chip) running at 2.0 GHz with 256 MB Shared L3 cache per chip, 2,048 GB of DDR4-RAM, eight NVIDIA A100 (each 80 GB HBM2 @ 1,555 GB/s; HGX board with NVLink; 9.7 TFlop/s in FP64 or 19.5 TFlop/s in FP32), two HDR200 InfiniBand HCAs, 25 GbE, and 14 TB on local NVMe SSDs; seven of these nodes belong to HS Coburg and one to FAU Audiolabs/FHG-IIS.
- 38 GPGPU nodes, each with two AMD EPYC 7713 *Milan* CPUs (64 cores per chip) running at 2.0 GHz with 256 MB Shared L3 cache per chip, 512 GB of DDR4-RAM, eight NVIDIA A40 (each with 48 GB DDR6 @ 696 GB/s; 37.42 TFlop/s in FP32), 25 GbE, and 7 TB on local NVMe SSDs.
- Native access to the Lustre-based parallel file system of *Fritz* through 25/100 GbE.
- A dedicated central NVMe storage.
- Operating system: AlmaLinux 8 (RHEL clone).
- Batch system: Slurm.

A LINPACK performance of 3.24 PFlop/s has been measured in October 2022 resulting in rank 174 on the November 2022 Top500 list and rank 33 on the Green500 list of November 2022.

For molecular dynamics codes like GROMACS, an NVIDIA A40 GPGPU delivers a performance very similar to a much more expensive NVIDIA A100 GPGPU. Even for many machine learning workloads, the price/performance ratio of the NVIDIA A40 GPGPUs is more than competitive, while applications requiring double precision calculations certainly rely on the NVIDIA A100. The mixture of NVIDIA A40 and A100 allows to maximize the overall cluster performance for a broad range of applications, while still serving very diverse needs.

3.1.3 *Meggie* parallel cluster (Tier-3, from 2016)

Meggie (system integrator: MEGWARE) is a system that is designed for running parallel programs using significantly more than one node. It is intended for distributed-memory (MPI) or hybrid-parallel programs with medium to high communication requirements and consists of originally 728 compute nodes, each with two Intel Xeon E5-2630v4 *Broadwell* CPUs (10 cores per chip) running at 2.2 GHz, 64 GB of RAM, and OmniPath interconnect. 64 compute nodes and the parallel file system had to be turned off during 2022 due to fatal hardware failure.

3.1.4 *Emmy* parallel cluster (Tier-3, from 2013)

Emmy (manufacturer: NEC) is a high-performance compute resource with high-speed interconnect. It was intended for distributed-memory (MPI) or hybrid-parallel programs with medium to high communication requirements and served also for high-capacity

throughput workload towards the end of its lifetime. *Emmy* consists of 560 compute nodes (each with two Xeon 2660v2 *Ivy Bridge* CPUs (10 cores per chip), running at 2.2 GHz, 64 GB of RAM, and QDR InfiniBand). *Emmy* also features a small accelerator partition with NVIDIA Tesla K20 and Tesla V100 GPUs. (The Intel Xeon Phi accelerators were shut down some years ago when Intel abandoned their software support.) **Emmy has been turned off in September 2022 due to its high power consumption in relation to its compute performance.**

3.1.5 *Woody* throughput cluster (Tier-3, multiple phases 2013–2022)

Woody is the preferred cluster for serial/single-node throughput jobs and has a long history. The nodes changed over time while the name was kept. The currently more than 200 nodes consist of (a) single-socket nodes with rather high-frequency quad-core Intel Xeon E3-12xx processors (E3-1240 v3 *Haswell*; E3-1240 v5 *Skylake*; E3-1240 v6 *Kaby Lake*) and (b) new dual-socket nodes from 2022 with Intel Icelake processors. All nodes have 8 GB of RAM per core and a local HDD/SDD. While the single-socket nodes only use 1 Gbit Ethernet, the new Icelake nodes are connected via 25 Gbit Ethernet. Both node types form a single cluster with quite similar per-core characteristics.

3.1.6 *TinyFAT* cluster (Tier-3, 2016/2020)

TinyFAT is another small special-purpose cluster. It consists of almost 50 dual-socket Intel *Broadwell* or AMD *Rome* nodes with 256 or 512 GB of main memory and local SSD storage.

3.1.7 *TinyGPU* cluster (Tier-3, multiple phases since 2009)

TinyGPU started in 2009 as a small special-purpose research test bed. Over the years, *TinyGPU* grew and no longer is that tiny. It now consists of more than 40 nodes with, in total, more than 200 NVIDIA GPGPUs of different generations (NVIDIA GeForce GTX1080, GeForce GTX1080Ti, GeForce RTX2080Ti, GeForce RTX3080, Tesla V100, and A100). Almost all nodes have been funded by different research groups across FAU; NHR@FAU only takes care of the proper housing and operation.

3.2 Storage resources (Tier-3 with dedicated extensions for NHR)

NHR@FAU operates several storage systems of different quality but also price/performance ratios serving different needs. Most file servers have 100 Gbit Ethernet connections.

- All throughput and GPGPU/large-memory nodes have local HDDs/SSDs for fast job-local storage.
- Each parallel computer has a parallel file system which is dedicated to that cluster; there are no capacity quotas enforced but high-watermark deletion is applied. The main usage is for checkpoint-restart files.
- `$WORK` consists of several capacity-optimized NFS file servers (more than 4 PB in total); there is no or only very limited backup. These Linux servers can cope rather well with many small files.
- `$VAULT` is a high-quality file system where especially larger files can be kept for longer times; `$VAULT` has a total capacity of 6 PB in a single file system and is served by an IBM Spectrum Scale cluster consisting of six servers (two NDS and

Table 3.2: Core-hours or GPU-hours used by NHR and Tier-3 on the different systems in 2022.

system	NHR	Tier-3
Fritz	149 mio core-h	68 mio core-h
Meggie	–	93 mio core-h
Emmy	–	34 mio core-h
Woody	–	8 mio core-h
TinyFAT	–	24 mio core-h
Alex	1.37 mio GPU-h	0.7 mio GPU-h
TinyGPU	–	0.7 mio GPU-h

four CES nodes; about 700 NL-SAS HDDs plus some SSDs in hardware RAIDs provide the capacity).

- `$HOME` is served by the same IBM Spectrum Scale cluster as `$VAULT`. The main difference is the frequency of snapshots (every 30 minutes vs. once per day).
- For long-term offline storage, an IBM TS4500 tape library with currently eight LTO8 tape drives and two expansion frames for up to 3,370 tapes is available. Archiving data is a manual process.

3.3 System availability and usage 2022

There were no major outages in the year 2022. Table 3.2 shows an overview of core- and GPU-hours used. User jobs were running on *Meggie* for nearly 80% of the cluster's uptime, whereas the *Woody* throughput cluster was utilized by approximately 35%. The other two clusters available for Tier-3 users, *TinyFAT* and *TinyGPU*, were allocated by users for 60% and 50% of the time, respectively. The newly procured systems for both Tier-3 and NHR users, the parallel computer *Fritz* and the GPGPU cluster *Alex*, only slowly gained popularity among users. In 2022, 40% of *Fritz* were used and roughly 45% of *Alex* ran user jobs. NHR@FAU expects the usage of its computers to rise in 2023.

4. Training & Support

Division head: *Dr. Georg Hager*

The *Training & Support* division at NHR@FAU is responsible for HPC training efforts and all support activities beyond level two. This entails training and event coordination, conducting local courses and user engagement events, organizing invited talks, participation in and organization of external training, application-specific support, performance engineering of user code, expert support for algorithms and libraries, and participation in third-party funded research projects. Due to the cross-cutting nature of many of these activities, members of the *Training & Support* division are typically operative in other divisions as well.

Dr. Georg Hager

The RRZE HPC group, out of which NHR@FAU has emerged, has a long history of teaching, training, and support in HPC. As early as 1998, scientists from FAU were educated in programming and performance optimization for the then-current vector supercomputers VPP700 at LRZ Munich and its smaller satellite system, the VPP300 at RRZE. In 2000, the installation of the federal "HLRB" system at LRZ and the concomitant *KONWIHR* funding fostered the intensification of training and user support activities. Over the years, the RRZE HPC group could build up considerable expertise in parallel programming, performance modeling, and optimization, which led to the development of signature lectures and tutorials. This process was facilitated by a constant and lively mutual exchange among research, user consulting, teaching, systems administration, and training staff. Since 2019, the monthly *HPC Café* and *HPC in a Nutshell* events have been strengthening the communication with new and experienced HPC customers. With the advent of the NHR program in 2021, the spectrum of training offers could be broadened even further. In the course of the year 2022, the training program could be extended substantially. We have also engaged liaison scientists to contribute their specialized knowledge and skills in courses about accelerator programming and the GROMACS MD package.

The *Training & Support* division is further responsible for all HPC user support that requires intense interaction with the customer, such as software configuration, code porting, parallelization, performance analysis and optimization, and the selection of appropriate algorithms and libraries. These services are not only offered to NHR users but also to third party-funded projects conducted by the NHR@FAU *Research* division and its project partners, and to *KONWIHR* projects.

4.1 Training activities

4.1.1 Courses and tutorials

In the following, we list courses and tutorials that were conducted in 2022 and were either organized by NHR@FAU, conducted by others with substantial contributions from NHR@FAU, or performed on invitation by NHR@FAU (see Table 4.1). Most of these events were conducted online via Zoom.

Node-Level Performance Engineering

This "signature" course covers performance engineering approaches on the compute node level, conveying the required knowledge to develop a thorough understanding of the interactions between software and hardware. Pipelining, SIMD, superscalarity, caches, memory interfaces, ccNUMA, etc., are covered. A cornerstone of node-level performance analysis is the Roofline model, which is introduced in due detail and applied to various examples from computational science. This tutorial is designed to be modular and can be given in various formats.

Node-Level Computer Architecture and Performance Engineering

This special edition of the *Node-Level Performance Engineering* course was presented by Markus Wittmann and Georg Hager as a contribution to the first *NHR Graduate School Course Week* in June 2022 at ZIB in Berlin.

Introduction to Hybrid Programming

A tutorial conducted by Georg Hager in close collaboration with colleagues from HLRS Stuttgart (Rolf Rabenseifner) and TU Wien (Claudia Blaas-Schenner). Most HPC systems are clusters of shared memory nodes. To use such systems efficiently both memory consumption and communication time has to be optimized. Therefore, hybrid programming may combine the distributed memory parallelization on the node interconnect (e.g., with MPI) with the shared memory parallelization inside of each node (e.g., with OpenMP or MPI-3.0 shared memory). This course covers several parallel programming models on clusters of shared-memory nodes, combining MPI with OpenMP, MPI-3.0 shared memory, or accelerators. Numerous case studies and micro-benchmarks demonstrate the performance-related aspects of hybrid programming. Hands-on sessions are included on all days.

Parallel Programming of High Performance Systems

A three-day introduction to HPC. It is a long-standing collaborative course by NHR@FAU (and previously RRZE) and LRZ Garching. This online course is targeted at students and scientists with interest in programming modern HPC hardware, specifically the large-scale parallel computing systems available in Jülich, Stuttgart and Munich but also smaller clusters in Tier-2/-3 centers and departments.

Table 4.1: Courses and tutorials in 2022; if not stated otherwise, courses were held online "(e)" denotes an invited teacher from outside FAU.

Event	Date(s)	Place	Teacher(s)
Node-Level Performance Engineering	Jun 14	NHR Course Week	G. Hager, M. Wittmann
	Jun 28–Jul 1	HLRS	G. Hager, J. Eitzinger, B. Wesarg
	Nov 13–18	SC22, Dallas, in person	G. Hager, T. Gruber, G. Wellein
	Dec 5–7	LRZ Garching	G. Hager, G. Wellein
Introduction to C++ for beginners	Oct 10–14	NHR@FAU	S. Dimitrovic (e)
Modern C++ Software Design	Oct 5–7	NHR@FAU	K. Iglberger (e)
Introduction to Hybrid Programming tutorial	Apr 5–7	TU Wien	G. Hager, R. Rabenseifner [HLRS] and C. Blaas-Schenner [TU Wien]
	Jun 22–24	LRZ Garching	
	Dec 12–14	TU Wien	
Parallel Programming of High Performance Systems (PPHPS22)	Mar 8–10	NHR@FAU	A. Afzal, M. Wittmann, G. Hager, LRZ
Roofline Modeling and Performance Engineering	Jul 24	Univ. of Lugano, in person	G. Hager
Introduction to parallel programming with OpenMP	Oct 4	NHR@FAU	M. Wittmann
Fundamentals of Accelerated Computing with CUDA C/C++	Apr 21–22	NHR@FAU	S. Kuckuk
	Nov 28	NHR@FAU	
Fundamentals of Accelerated Computing with CUDA Python	Aug 2–3	NHR@FAU	S. Kuckuk
	Sep 22–23	NHR@FAU	
Performance Analysis on GPUs with NVIDIA tools	Sep 29	NHR@FAU	D. Ernst
GPU Performance Analysis	Jun 19–24	IHPCSS	D. Ernst
Using the *LIKWID* tool suite	Jun 29	LRZ	T. Gruber
	Nov 3	Univ. of Cambridge	
GROMACS	Dec 7–11	FAU, in person	R. Böckmann

Fundamentals of Accelerated Computing with CUDA Python & Fundamentals of Accelerated Computing with CUDA C/C++

These NVIDIA "Deep Learning Institute" courses were offered for the first time in 2022 by Sebastian Kuckuk, who underwent the required training and tests to become a certified *NVIDIA DLI University Ambassador*. These courses are the foundation for any serious performance-aware accelerator programming.

GPU Performance Analysis

This lecture given by Dominik Ernst at the International HPC Summer School (IHPCSS) for several years has been converted into a half-day tutorial with hands-on exercises, which is now offered twice a year at NHR@FAU. It complements the NVIDIA DLI courses (see above) since it covers the use of NVIDIA performance tools for the analysis of GPU programs written in CUDA C/C++, OpenACC, or OpenMP.

Introduction to C++ for beginners

Held by Slobodan Dmitrovic, is a one-week course with a focus on the introduction of the essential language features and the syntax of C++. Additionally, it introduced many C++ software development principles, concepts, idioms, and best practices, which enable programmers to create professional, high-quality code from the very beginning.

Modern C++ Software Design

Held by Klaus Iglberger, it is a three-day course on software development with the C++ programming language. Its focus is on essential C++ software development principles, concepts, idioms, and best practices, which enable programmers to create professional, high-quality code. The course will give insight into the different aspects of C++ (object-oriented programming, functional programming, generic programming) and will teach guidelines to develop mature, robust, maintainable, and efficient C++ code.

Introduction to Parallel Programming with OpenMP

A new full-day course with hands-on exercises, designed and conducted by Markus Wittmann. It introduces OpenMP from the ground up for developers who have no prior experience with parallel programming and focuses on CPUs only. An advanced part covering performance issues and accelerator programming will complement the introduction; the full two-day event will be conducted twice per year from 2023 on.

Roofline Modeling and Performance Engineering

An invited tutorial with hands-on exercises conducted by Georg Hager at the *2022 CSCS-USI Summer University on Effective High-Performance Computing and Data Analytics* in Serpiano, Switzerland. Attendees were familiarized with the Roofline model in the context of sparse matrix computations and could analyze sparse linear algebra code on the *Piz Daint* supercomputer at CSCS in Lugano.

Tutorial on the GROMACS molecular dynamics package

A five-day hands-on course conducted by the group of Rainer Böckmann, Professorship of Computational Biology. This course covered an introduction into the molecular dynamics engine GROMACS, including fundamental commands and applications. Participants learned how to prepare and run simulations of biomolecular systems including membranes and proteins at an atomistic and

coarse-grained level of resolution. Post-processing and analysis of simulation trajectories were a large part of the tutorial. The course is usually embedded in the Bachelor programs of Biology and Integrated Life Sciences but had five places available for people from NHR. The course was held in person and took place in the CIP of the Biology Department.

LIKWID

The popular *LIKWID* tool suite was presented by Thomas Gruber as part of several courses at other sites in 2022. *LIKWID* (see Section 5.1.1) is in wide use at many computing centers worldwide for affinity control, performance analysis, and monitoring. It contributes significantly to the visibility and reputation of NHR@FAU within the HPC community.

4.1.2 HPC Café

The *HPC Café* was initiated in late 2019 as a new way to foster and facilitate communication between the RRZE HPC group and HPC customers. It also complements the traditional HPC support channels by providing a deliberately informal setting. Every second Tuesday of a month, all HPC customers are invited to a freestyle Q&A session, followed by a short presentation about a "focus topic" of interest. In 2022, the *HPC Café* has been conducted as a hybrid event whenever possible, with HPC customers and speakers attending in person and the option for external attendees to connect via Zoom. Talks were usually recorded and published on the FAU video portal (https://www.fau.tv/course/id/1146) and/or the NHR@FAU YouTube channel (https://youtube.com/NHRFAU). A complete list with all events and links to slides and video recordings can be found at https://hpc.fau.de/systems-services/support/hpc-cafe/. The following focus topics have been covered in 2022:

- December 13, 2022 (hybrid): *Using the Zenodo document and data repository* by Dr. Jürgen Rohrwild, FAU University Library.
- November 8, 2022 (hybrid): *NHR@FAU User Forum*. Invited top users talked about their experience with compute clusters at NHR@FAU.
- October 11, 2022 (hybrid): *Modern Fortran* by Dr. Reinhold Bader, LRZ Garching.
- September 13, 2022 (online): *News from NHR@FAU cluster configurations, resource monitoring* by Thomas Zeiser.
- July 12, 2022 (online): *Job monitoring for end users*—ClusterCockpit *and PyTorch Profiling at work* by Jan Eitzinger.
- June 14, 2022 (hybrid): *The new clusters at NHR@FAU: hardware, access, application process* by Thomas Zeiser and Gerhard Wellein.
- May 10, 2022 (online): *From C++98 to C++11/14/17* by Jan Hönig, FAU.
- April 12, 2022 (online): *SLURM: Basics, best practices, advanced usage* by Katrin Nusser.
- March 8, 2022 (online): *GROMACS—Best practices* by Anna Kahler.

- February 8, 2022 (online): HPC Café für Kontaktpersonen (Liaisons) by Georg Hager.
- January 18, 2022 (online): *Using File Systems Properly* by Markus Wittmann and Georg Hager.

4.1.3 HPC in a nutshell

Each month, one day after the *HPC Café*, a two-hour online introduction into NHR@FAU systems, access formalities, and basic usage was provided by Katrin Nusser and Markus Wittmann. This well-received format aims at reducing the entry barrier for new and inexperienced users. The content is continuously being updated to reflect recent changes in NHR@FAU systems and access rules.

4.2 NHR PerfLab Seminar series

The *NHR PerfLab Seminar* is a collaborative effort by the NHR centers at FAU, RWTH Aachen, ZIB Berlin, and the University of Paderborn to combine resources and activities around performance engineering, code analysis, HPC computer architectures, and tools. The *NHR PerfLab Seminar* is a series of public talks about those (and more) topics. Since the inception of the seminar, NHR@FAU has taken on the role of organizing the talks, speaker acquisition, and pre- and post-event dissemination. If possible, all material likes slides and video recordings were published on the NHR@FAU website and YouTube channel. A full list of past seminars with links to slides and video recordings is available at `https://hpc.fau.de/research/nhr-perflab-seminar-series/`. The following talks have been given in 2022:

- December 20, 2022 (hybrid): *Quantum Computing Today and Tomorrow* by Michael Hartmann, FAU.
- November 29, 2022 (online): *Enabling next-generation processor simulations with SimEng* by Simon McIntosh-Smith, University of Bristol.
- November 8, 2022 (online): *High-Performance Implementations for High-Order Finite-Element Discretizations of PDEs* by Martin Kronbichler, University of Augsburg.
- October 18, 2022 (hybrid): *New Features in MPI 4.0* by Rolf Rabenseifner and Tobias Haas, High Performance Computing Center Stuttgart (HLRS).
- September 6, 2022 (online): *LARC: A Case Study in Enhancing CPUs with Copious 3D-Stacked Cache* by Jens Domke, RIKEN.
- June 21, 2022 (online): *Parallel Graph Processing—a Killer App for Performance Modeling* by Ana Lucia Varbanescu, University of Twente.
- June 9, 2022 (hybrid): *Innovative Scientific Computing by Integration of (Simulation+Data+Learning) in Information Technology Center, The University of Tokyo* by Kengo Nakajima, University of Tokyo.
- May 10, 2022 (online): *Designing for Portable, Efficient and Explainable Performance* by Tze Meng Low, Carnegie Mellon University.
- April 26, 2022 (online): *The Role of Idle Waves in Modeling and Optimization of Parallel Programs* by Ayesha Afzal, NHR@FAU.
- April 12, 2022 (online): *gprofng: The Next Generation GNU Profiling Tool* by Ruud van der Pas, Oracle Corp.

- April 5, 2022 (online): *OpenMP Offloading* by Christian Terboven, RWTH Aachen University IT Center.
- February 15, 2022 (online): *Designing Next-Generation Numerical Methods with Physics-Informed Neural Networks* by Stefano Markidis, KTH Royal Institute of Technology.
- February 1, 2022 (online): *Performance Engineering for Sparse Matrix-Vector Multiplication with the Recursive Algebraic Coloring Engine* by Christie L. Alappat, NHR@FAU.

4.3 Projects

4.3.1 EoCoE-II

The EoCoE-II project (for details see Section 6.5.4) ended officially in 2021. In 2022, some loose ends were tied up and a proposal for a third project phase was submitted. Over the course of the project, significant performance improvements could be achieved for the flagship codes *Alya*, *GYSELA*, and *EURAD-IM*. Especially for *Alya*, the optimizations developed for GPUs could be applied across many parts of the matrix assembly and led to acceleration factors of up to 40. A paper about the optimization strategies applied is in preparation.

4.3.2 KONWIHR

The main objective of *KONWIHR* is to foster the efficient use of high-performance computers and to increase and broaden their impact on research. *KONWIHR* support scientists at Bavarian universities in adapting and developing their numerical applications or other data- or computation-intensive codes for modern parallel computer architectures. Funding is based on the duration of the project and can amount to a maximum of €10,000 (small project of three months) or up to €50,000 for projects with a duration of twelve months. The projects must be carried out in close cooperation with one of the two computing centers (LRZ and NHR@FAU). In particular, the funded project staff member should spend an extended period at one center.

NHR@FAU receives *KONWIHR* funding to coordinate *KONWIHR* activities in the northern part of Bavaria and to support *KONWIHR* projects in optimizing and adapting their codes and workflows for HPC systems. *KONWIHR* is currently led by Prof. Bungartz (TU Munich) and Prof. Wellein (FAU).

In 2022, *KONWIHR* granted the following project applications:

- *Vectorization of SPHinXsys, a strong-coupling, meshless multi-physics and AI-aware library*, Dr. Xiangyu Hu, TU Munich, small project.
- *Improving FAIRness of HPC research data*, Prof. Stemmer, TU Munich, small project.
- *ALPACA in Florence—first steps towards the Ponte Vecchio: Porting ALPACA to Intel's upcoming GPU architecture Ponte Vecchio*, Prof. Adams, TU Munich, large project.
- *Porting of Lattice QCD simulation software to GPUs*, Prof. Wettig, University of Regensburg, large project.
- *HPC mixed precision quantization of encoder-decoder deep neural networks*, Prof. Kist, FAU, small project.
- *Massively parallel solvers for geophysical flow problems with strong viscosity variations*, Prof. Rüde, FAU, small project.

- *Continuous Benchmarking for the GHODDESS framework*, Prof. Köstler, FAU, small project.
- *Optimization of the ALF implementation of the auxiliary-field quantum Monte Carlo algorithm: porting to GPUS and symmetry considerations*, Prof. Assaad, Uni Würzburg, large project.
- *HPC-Stützpunkt at the University of Regensburg*, Prof. Wettig, University of Regensburg, Basis project.

In addition, workshops have been organized to allow the *KONWIHR* projects to present their work and discuss with other projects and with the expert personnel from NHR@FAU and LRZ:

- October 12, 2022: *KONWIHR* workshop for projects from 2022-1.
- April 21, 2022: *KONWIHR* workshop for projects from 2021-2.

4.4 NHR Graduate School

In 2022, the first cohort of new PhD students entered the three-year program of the *NHR Graduate School*. The *NHR Graduate School* awards up to nine PhD scholarships each year to doctoral students from all over the world. During the program, each student is required to spend six months at one of the other NHR centers.

In the fall of 2021, Raviraj Mandalia from Gujarat (India) successfully applied for a PhD scholarship at NHR@FAU. On April 1, 2022, he started his position at the Chair for Theoretical Chemistry of Prof. Görling. Earlier that year, he had completed a Master of Science in physics at St. Xavier's College in India. The work planned during the PhD comprises both applied projects as well as projects in the area of method development in electronic structure theory. During his first year, Mr. Madalia carried out electronic structure calculations for novel two-dimensional materials in collaboration with experimental working groups in Germany and Brazil. First work analyzing the structure of porous graphene nanoribbons has been submitted for publication and is presently under review. In a second project, structures and energies of two-dimensional metal-organic frameworks adsorbed on graphene are investigated. In the area of method development, Mr. Mandalia is active in the development of novel density-functional methods that enable a substantially more accurate description of electronic structures than present conventional density-functional methods. Here, Mr. Mandalia provided high-level reference data for electronic energies by carrying out full configuration interaction calculations for selected atoms and small molecules and started to investigate strategies to deal with symmetries in real space as well as spin space in the construction of orbital-dependent functionals for the correlation energy of electronic structures of atoms and molecules.

4.5 Dissemination

Dissemination of activities and results was an integral part of the NHR@FAU activities in 2022. Event and course announcements, success stories from the support (see below), calls for compute time projects, talk recordings, etc., were disseminated via social media channels (YouTube, Twitter), our bi-monthly newsletter (see below), the NHR@FAU home page, and mailing lists that reach out to NHR@FAU system users, *KONWIHR* project PIs, and interested subscribers all across the NHR Alliance.

Each semester, a 90-minute online introductory talk on *Resources for High Performance Computing at FAU* was given at the FAU Graduate Center in order to inform graduate students of the opportunities for using NHR@FAU HPC clusters for their research. Finally, the concept of the NHR Alliance and the resources at NHR@FAU were detailed in various online and on-site talks at Universities:

- September 21, 2022 (online): *The National High Performance Computing Alliance and NHR@FAU: New structures and opportunities* by G. Wellein at the Ruhr University Bochum
- November 23, 2022: *The National High Performance Computing Alliance and NHR@FAU: New structures and opportunities* by G. Hager and G. Wellein at the University of Bayreuth.
- December 19, 2022: *The National High Performance Computing Alliance and NHR@FAU: New structures and opportunities* by G. Hager and G. Wellein at the University of Regensburg.

4.5.1 Success stories from the support

Frequently, customers observe low or fluctuating performance of their jobs, or the cluster monitoring indicates low resource utilization. While most of these cases can be categorized as simple accidental misuse, some require deeper investigation by support staff. The ensuing consultation occasionally results in impressive performance improvements. We document such cases regularly at https://hpc.fau.de/about-us/success-stories/.

4.5.2 NHR@FAU newsletter

The bi-monthly NHR@FAU newsletter is a publication that summarizes recent and upcoming events at NHR@FAU and the NHR association. Beyond events, every newsletter highlights special or noteworthy achievements by NHR@FAU scientists and an FAQ corner. Finally, the "newsletter spotlight" introduces an NHR@FAU scientist in some detail. In 2022, six newsletters were published. All newsletters are available at https://hpc.fau.de/nhrfau-newsletters/.

A mailing list has been set up to alert about the publication of a new issue. It is open for everyone to subscribe at: https://lists.fau.de/cgi-bin/listinfo/nhr-newsletter.

4.5.3 Social media

In order to widen its reach to potential customers and audiences, NHR@FAU leverages social media channels. Currently, we concentrate mostly on dissemination via recorded lectures, seminars, and tutorials, which are published either on the NHR@FAU YouTube channel (https://www.youtube.com/NHRFAU) or the FAU video portal (https://www.fau.tv/course/id/1146). Our recordings of parallel programming lectures are particularly well received and generate considerable feedback in the HPC community and among students.

5. Software & Tools

Division head: *Dr. Jan Eitzinger*

Dr. Jan Eitzinger

The *Software & Tools* division at NHR@FAU is responsible for the development of Open-Source software and for providing and administrating services for NHR@FAU and the NHR alliance. To this end, the division develops, maintains, and supports Open-Source software that is published on our GitHub account (`https://github.com/RRZE-HPC`): the *LIKWID* performance tools, The Open Source Architecture Code Analyzer (*OSACA*), the *Kerncraft* Loop Kernel Analysis and Performance Modeling Toolkit, *The Bandwidth Benchmark*, *MachineState*, and the *ClusterCockpit* Monitoring Framework.

The division participates in third party projects related to performance and monitoring tools, develops and maintains software for internal use at NHR@FAU such as the *HPC Portal*, and provides and administrates services (NHR Moodle, NHR@FAU *HPC Portal*, NHR@FAU job-specific monitoring for all cluster systems).

We are a member of the Virtual Institute—High Productivity Supercomputing (VI-HPS) and the SPEC Research Group. In 2022, we contributed our expertise in performance and monitoring tools and software development in the NHR central projects *Standards and interfaces for a system-wide job-specific performance monitoring* (project lead) and *Cx as a service for sustainable HPC research software engineering*.

The RRZE HPC group, out of which NHR@FAU has emerged, started to develop the *LIKWID* Performance Tool Suite already in 2009. From 2013 to 2016, it had the project lead in the BMBF *FEPA* project, which developed a system-wide job-specific monitoring infrastructure. From 2017 to 2020, the group was a partner in the BMBF Metacca project and contributed its LIKWID and Kerncraft tools. Also from 2017 to 2020, the RRZE HPC group was the project lead in the DFG *ProPE* project, where among other activities the initial development for the ClusterCockpit monitoring framework began. ClusterCockpit is now besides LIKWID the second largest Open Source project at NHR@FAU.

Since fall 2022, the BMBF project *EE-HPC* has been strengthening our activities in cluster-wide performance and energy monitoring. NHR@FAU is the project lead with the further partners HLRS Stuttgart, RWTH Aachen, Deutsches Klimarechenzentrum (DKRZ), and Hewlett Packard Enterprise (HPE). NHR@FAU organized a project kickoff meeting from October 27–28, 2022 in Erlangen.

5.1 Tool development

The tool development at NHR@FAU is well integrated with our main research topics of performance engineering and performance modeling. LIKWID and OSACA are important tools used in many of our research and support projects and enable us to adopt and research novel processor architectures at a very early stage. With Kerncraft, analytical performance models (the well-known Roofline model and the ECM model, which was developed at NHR@FAU) are accessible also for non-expert users. The ClusterCockpit monitoring stack bundles our activities in cluster-wide monitoring solutions.

A new addition to our software portfolio is *MD-Bench* (https://github.com/RRZE-HPC/MD-Bench), a performance-focused prototyping harness for state-of-the-art short-range molecular dynamics algorithms. MD-Bench is a vehicle for research on performance engineering for molecular dynamics algorithms and strengthens our expertise in this area. Its source code is simple, understandable, and extensible, and therefore well suited for benchmarking, teaching, and researching MD algorithms. A paper about MD-Bench was presented at the PPAM conference 2022 in Gdansk, Poland, where it won the Best Paper Award.

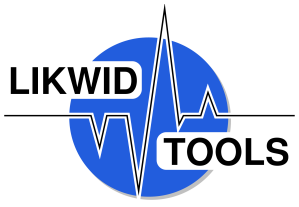
5.1.1 LIKWID performance tool suite

LIKWID (https://github.com/RRZE-HPC/likwid) is an easy-to-use yet powerful command line performance tool suite for the GNU/Linux operating system. LIKWID is maintained by Thomas Gruber.

Currently, LIKWID consists of seven core tools of which `likwid-perfctr` (counting hardware performance events), `likwid-topology` (display node topology), `likwid-pin` (control thread and process affinity), and `likwid-bench` (microbenchmarking framework) are most prominent. The tools are specific to hardware architecture and currently support x86-64, ARM, and Power processors, and NVIDIA GPUs.

LIKWID is by far our most popular open-source project, and its tools are used worldwide for research, teaching, or in production environments, e.g., at NERSC (Lawrence Berkeley National Laboratory), CSCS (Swiss National Supercomputing Center) in Lugano, the National Super Computer Center in Guangzhou, the Vienna Scientific Cluster (VSC), or the Barcelona Supercomputing Center (BSC).

Moreover, LIKWID is used by the members of the Gauss Center for Supercomputing (GCS)—LRZ Garching, HLRS Stuttgart, and JSC Jülich—and at the IT provider for the Max Planck Society (MPCDF). Some of the Tier-2/3 HPC sites in Germany using LIKWID are RWTH Aachen, TU Dresden, KIT Karlsruhe, University of Paderborn, University of Konstanz, University of Gießen, and the national research center DESY in Hamburg.



In 2022, support for the AWS Graviton 3 and the HiSilicon TaiShan v110 chips was added in LIKWID. Moreover, Apple released the M1, its first ARM chip; fter the Asahi Linux team ported Linux to the Apple M1, support for these chips was added to LIKWID as well. To ensure software functionality during development, LIKWID was integrated in the CI/CD infrastructure of NHR@FAU. Furthermore, the CI/CD infrastructure creates software packages for different Linux-based operating systems. The Julia team (mainly Carsten Bauer from NHR center PC2) created a Julia interface to *LIKWID* (https://github.com/JuliaPerf/LIKWID.jl).

5.1.2 OSACA Open Source Architecture Code Analyzer

OSACA is a tool that can analyze x86 and Arm64 assembly code and produce runtime predictions assuming steady-state execution and no cache misses.

By taking data dependencies into account, OSACA provides not only a throughput prediction as the best-case scenario but also the critical path and loop-carried dependencies for loop kernels. OSACA is maintained by Jan Laukemann and available at https://github.com/RRZE-HPC/OSACA.

A tool like OSACA is needed for analytic performance modeling, e.g., to formulate ECM or refined Roofline models. While there exist similar tools, such as the Intel Architecture Code Analyzer (IACA), LLVM's Machine Code Analyzer (LLVM-MCA), or the uops.info Code Analyzer (uiCA), they lack accuracy in prediction, are not open source, discontinued in development, or do not provide support for non-x86 architectures. OSACA, on the other hand, can also handle all modern Intel and AMD x86 architectures and several Arm processors. It is available as a Python3 module and a CLI application, but it is also integrated into the Compiler Explorer at https://godbolt.org, which allows using OSACA from a browser without any installation.

In 2022, we enhanced the supported micro-architectures by the AMD Zen 3 (co-dename *Milan*) and HiSilicon TaiShan v110 that are used in modern supercomputers around the globe. Furthermore, we created a brand new tutorial from scratch around the *OSACA* tool, tackling the important topic of core-level performance engineering, which was held for the first time in October for a selected audience of the NHR alliance. We aim to establish this tutorial at leading HPC and performance engineering conferences throughout the upcoming years.

5.1.3 Kerncraft loop kernel analysis and performance modeling toolkit

Kerncraft is a loop kernel analysis and performance modeling toolkit. It allows automatic analysis of loop kernels using the Execution Cache Memory (ECM) model and the Roofline model, together with their validation via actual benchmarks. Kerncraft provides a framework for investigating the data reuse and cache requirements by static code analysis.

By employing Intel IACA or our OSACA tool together with static source code analysis, Kerncraft can give a good overview of both in-core and memory bottlenecks and use this data to construct predictive, white-box performance models. Kerncraft was developed and maintained by Julian Hammer, who has left FAU after finishing his PhD. We currently search for funding and a new maintainer for further development of Kerncraft.

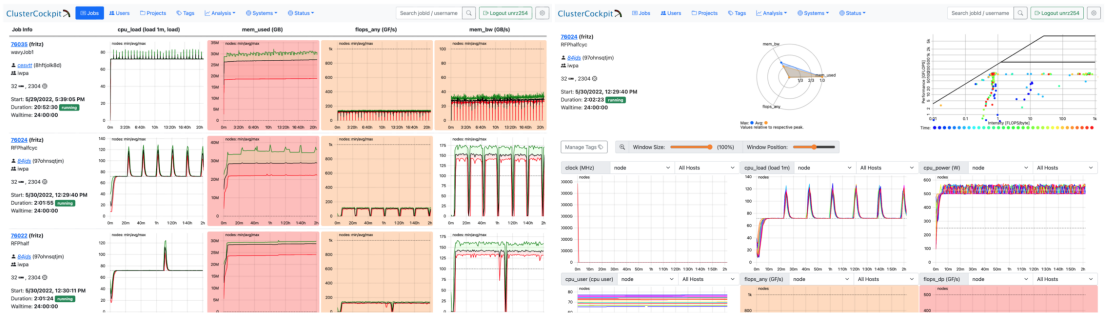


Figure 5.1: Screenshots of the *ClusterCockpit* web interface. Left: job list of running jobs; right: single-job view showing the resource utilization in a polar plot and diagnostic Roofline plot.

5.1.4 ClusterCockpit monitoring framework

ClusterCockpit is a full-stack framework for job-specific performance monitoring on HPC clusters. ClusterCockpit is available at https://github.com/ClusterCockpit. It started as a side project for a web frontend during the DFG ProPE project and has grown into a cooperative effort to provide a tailor-made software stack for job-specific performance monitoring.

ClusterCockpit currently comprises cc-backend (web API backend), cc-frontend (web-based user interface), cc-metric-collector (a node agent to measure and forward node metric data), and cc-metric-store (a simple metric time series in-memory cache). A focus of ClusterCockpit is to define standards and interfaces for an interoperable monitoring ecosystem. We document and specify generic data structure descriptions and developed a job archive specification that allows to archive and share job performance data in a portable manner.

Four members of the team (Jan Eitzinger, Thomas Gruber, Christoph Kluge, and Rafael Ravedutti) contributed to the development of ClusterCockpit in 2022, focusing on stability and ease of installation and maintenance. ClusterCockpit is in production use at NHR@FAU, PC2 Paderborn, and DKRZ Hamburg. It is funded as part of the BMBF EE-HPC project (https://eehpc.clustercockpit.org), which started in fall 2022 (see Sect. 6.5.3). Our aim is to establish ClusterCockpit as an attractive open-source offering for job-specific performance monitoring in academic HPC centers.

5.1.5 MachineState

MachineState is a Python3 module and CLI application for documenting and comparing hardware and software settings known to affect performance. It is our contribution to enabling deterministic and reproducible benchmark results on today's complex processor and node architectures. MachineState is maintained by Thomas Gruber and available at https://github.com/RRZE-HPC/MachineState.

5.1.6 The Bandwidth Benchmark

The Bandwidth Benchmark (https://github.com/RRZE-HPC/TheBandwidthBenchmark) is inspired by the famous STREAM main memory bandwidth microbenchmark by John McCalpin. It can be seen as "STREAM on steroids" and contains eight streaming kernels with varying data access patterns.

Since it is a simple, modular C99 implementation with a simple yet flexible makefile-based build system, it can also be used as a blueprint for other microbenchmarking projects. For teaching purposes, a single-file version of this benchmark is available. The project's Wiki pages (https://github.com/RRZE-HPC/TheBandwidthBenchmark/wiki) feature results for many processor architectures. The Bandwidth Benchmark is maintained by Jan Eitzinger.

5.2 Services

5.2.1 HPC Portal

Modern digital workflows are key for efficient project management, handling of user accounts, and resource access. To this end, we are developing the HPC Portal web interface, which provides principal investigators with the means for self-administration of approved projects. Access to computational NHR resources are streamlined by utilization of user-supplied public SSH keys, which allows secure, password-less authentication for these accounts.

The *HPC Portal* supports login via Single-Sign-On (SSO), allowing users to use the credentials of their respective home organization. It is in production use (https://portal.hpc.fau.de and the default user management solution for external users and teaching accounts. It is planned to transfer existing Tier-3 groups to the *HPC Portal* and make it the only supported management solution by the end of 2023. The *HPC Portal* is maintained by Christoph Kluge.

5.2.2 NHR Moodle learning platform

Since 2021, we have been hosting and administrating a Moodle server for general use within the NHR Alliance (https://moodle.nhr.fau.de/). Login is enabled through DFN-AAI Single-Sign-On authentication. The service is deployed as a docker container on a high-availability VM cluster at RRZE and uses the central RRZE SQL database and tape backup services. The NHR Moodle service is maintained by Christoph Kluge.

5.2.3 NHR@FAU *ClusterCockpit* service

For usage within the FAU, we administer a central *ClusterCockpit* monitoring server that can be reached at https://monitoring.nhr.fau.de. For this purpose, we procured a dedicated server with large main memory capacity and fast NVMe file storage.

The server is used for all cluster systems and provides job-specific performance monitoring for support personnel and HPC users. FAU HPC users can authenticate via LDAP. External NHR users can delegate their authentication and initiate a *ClusterCockpit* session with a button in the *HPC Portal*. This service is jointly maintained by Katrin Nusser from *Systems & Services* and Jan Eitzinger, Thomas Gruber, Christoph Kluge from *Software & Tools*.

5.2.4 Services for internal use

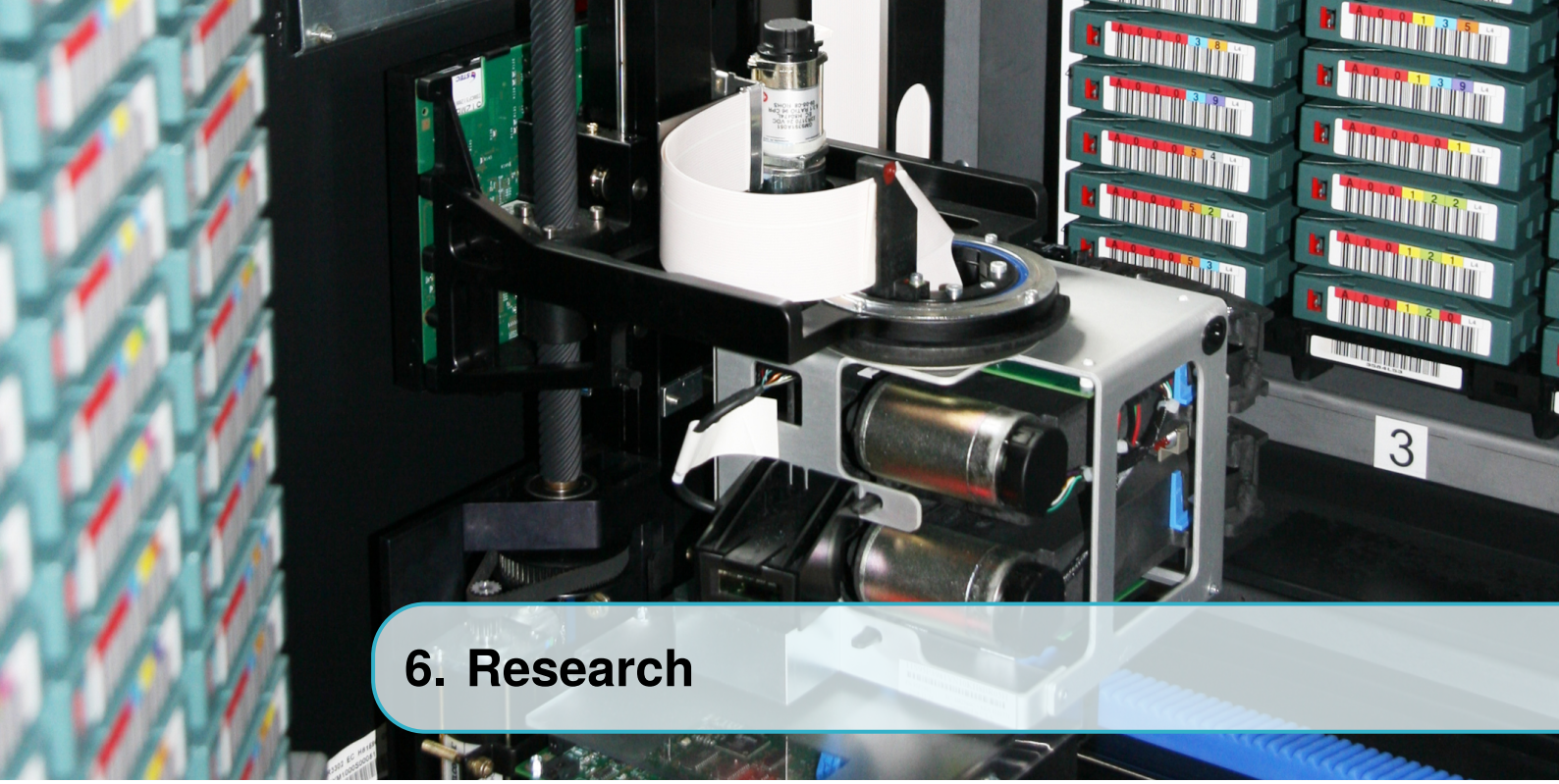
The *Software & Tools* division also hosts and maintains additional services for usage within NHR@FAU:

- A LimeSurvey instance for online surveys (https://survey.nhr.fau.de)
- A Discourse forum software instance for documentation and knowledge transfer (https://community.nhr.fau.de)

All services are running in the same Docker setup on a RRZE VM using a reverse proxy as frontend. They are integrated into the RRZE service infrastructure and use the RRZE database and backup offerings.

5.3 Talks

- J. Eitzinger, T. Wilde: BMBF Green HPC project Quelloffene Lösungsansätze für Monitoring und Systemeinstellungen für energieoptimierte Rechenzentren (EE-HPC). Online talk at Powerstack Seminar (2022), November 10–11, 2022, Texas, USA.
- T. Gruber: Using the *LIKWID* tool suite. Webinar at the RSE Seminars. University of Cambridge, UK, November 3, 2022.



6. Research

Division head: *Prof. Dr. Harald Köstler*

Prof. Dr. Harald Köstler

The overarching goal of the research activities at NHR@FAU is to introduce a systematic and model-guided performance engineering (PE) process into all performance analysis and optimization activities within the field of computing. This process helps developers and performance analysts on all levels of expertise to understand observed performance and find optimization opportunities. Our research is focused on developing and applying analytic performance models, tools, and libraries that support this mission. We participate in projects where we can leverage our expertise to help our partners develop deeper insights into performance issues and how to mitigate them. Additionally, we combine our HPC expertise with software engineering and code generation technology to provide software frameworks for a variety of applications in the field of computational science and engineering.

6.1 Performance modeling and performance engineering

People

Dominik Ernst, Ayesha Afzal

GPU performance modeling

The paper *Analytical Performance Estimation during Code Generation on Modern GPUs* was accepted for publication in a special issue of the "Journal of Parallel and Distributed Computing." It models the complex memory hierarchy of NVIDIA A100 GPUs and presents a method for accurately estimating the data transfer volumes in the different levels of the cache hierarchy based solely on high level code features without executing and not even compiling the code. These data are fed into an extended performance model that enables predictions and insight into performance behavior of algorithms.

In the context of the European EoCoE-II project (see Sect. 6.5.4), we identified and implemented further improvements in a Navier-Stokes matrix assembly routine of the finite element code *Alya*. We found out that the key changes to the GPU-specialized variant using OpenACC are applicable to the CPU variant as well, which makes a (re-)unified code base possible. A detailed analysis and measurements of the different stages of optimizations for both CPUs and GPUs have been detailed in a paper that is currently in the final steps of preparation before submission.

Performance of Memory-bound Applications: The Role of Idle Waves

In prior years, using a variety of HPC platforms and diverse benchmarking scenarios, we published our work on developing and validating the analytic model for the propagation speed of idle waves[1] and incorporating noise effects into analytical performance models that describe how these idle waves interact nonlinearly within a parallel code on a cluster and how they decay due to system noise, system topology, and application load imbalance. Further, we analyzed how memory bandwidth bottlenecks affected the propagation of idle waves, which suggested that the subsequent desynchronization might speed up code execution via overlapping the productive work with the communication overhead.

In 2022, four papers co-authored by Afzal et al. were published, and a fifth paper (which is an extension of one of these four papers) was submitted for publication in 2023. The details of these five papers are listed below.

Analytic performance model for parallel overlapping memory-bound kernels
: In January 2022, our article was published in the "Computation and Concurrency: Practice and Experience (CCPE)" journal and released with open access [1]. These findings expand on earlier research, presenting an analytical performance model that explains how various compute kernels share memory bandwidth in a ccNUMA contention domain.

The Role of Idle Waves, Desynchronization, and Bottleneck Evasion in the Performance of Parallel Programs
: This paper was accepted for publication in the "IEEE Transactions on Parallel and Distributed Systems (TPDS)" Journal in October 2022 [3]. It investigates how our previous findings, which were primarily based on microbenchmarks and toy codes, could be applied to study the performance implications of desynchronization and automatic asynchronous communication for a variety of application properties and parameters, including saturation point, matrix structures, domain decomposition, and communication concurrency.

Addressing White-Box Modeling and Simulation Challenges in Parallel Computing
: In June 2022, this paper was published at the "ACM SIGSIM Conference on Principles of Advanced Discrete Simulation (PADS'22)," which was held in New York, NY, USA [6]. These results were presented virtually and guide the selection of parameter changes for optimizing program performance by locating performance-relevant bottlenecks and the dynamics of massively parallel programs in a white-box (first-principles) setting.

[1] Idle waves emerge from delays in computation or communication on specific processes and propagate throughout the parallel program like a train delay that makes other trains wait and thus "ripples" through the schedule.

Exploring Techniques for the Analysis of Spontaneous Asynchronicity in MPI-Parallel Applications
In September 2022, this paper was published at the International Conference on Parallel Processing and Applied Mathematics (PPAM'22)," in Gdansk, Poland [5]. This work was initiated in April 2022 during the visit of associate professor Stefano Markidis from KTH Royal Institute of Technology in Stockholm to NHR@FAU. His visit was supported by a grant from the HPC-Europa3 Transnational Access program, which enables European researchers to collaborate on HPC projects and access the top-level HPC systems. This study explores the utility of using principal component analysis, clustering techniques, correlation functions, and a novel phase space plot for identifying, classifying, and characterizing the dynamics of large-scale MPI-parallel programs.

Making Applications Faster by Asynchronous Execution: Slowing Down Processes or Relaxing MPI Collectives
In October 2022, the PPAM'22 conference invited selected papers to submit expanded versions in a special issue of the "Future Generation Computer Systems (FGCS)" Journal. This manuscript was submitted to the FGCS Journal as an extended version of the aforementioned paper [5] and is anticipated to be published in 2023 [0]. This study extends the utility of phase-space graphs as a new tool to visualize parallel program dynamics and sheds light on the fact that noise, independent of its source, is not always detrimental and how it can be leveraged for performance improvements if certain conditions are met.

Asynchronicity and Load Balancing Techniques: A Trade-Off

To improve resource utilization, we began looking quantitatively into the interaction of advanced load balancing methods and desynchronization to analyze whether load balancing might be ineffective when modest imbalances allow MPI processes to desynchronize. The research is currently ongoing in collaboration with the High Performance Computing research group at the University of Basel's Department of Mathematics and Computer Science, and it is anticipated that the findings will be published in 2023.

Posters and prizes

In 2022, the following two posters were presented:

White-box Modelling of Parallel Computing Dynamics
This poster was presented virtually in January 2022 at the 5th international conference on "High Performance Computing in Asia-Pacific Region (HPC Asia)" [4].

DisCostiC: A DSL-based Parallel Simulation Framework using First-Principles Analytic Performance Models
This poster was presented in June 2022 at the "Platform for Advanced Scientific Computing Conference" (PASC) in Basel, Switzerland [2].

NHR@FAU also contributed to the "International HPC Summer School" (IHPCSS) that has been held in Athens, Greece in 2022 with a tutorial by Dominik Ernst about "GPU Performance Analysis" (for details on this tutorial, please see Section 4.1.1). The IHPCSS is jointly organized by HPC institutions from the USA, Canada, Europe, and Japan.

Together with the Chair for Hardware Architecture at the Department of Computer Science at FAU, NHR@FAU co-hosted a student team that participated in the "Student

Cluster Competition" at the "Supercomputing Conference 2022." NHR@FAU supported the team with access to their complete hardware arsenal, system administration and application expertise, vendor contacts, and most importantly, with their team advisor Dominik Ernst. The team placed fifth overall, and won the "Highest HPCG Score" Award.

6.2 Performance tools

Our developments and research in the field of performance related tools is bundled in the *Software & Tools* division (for details see Section 5.1), which collaborates closely with the *Research* division.

6.3 Building blocks for sparse linear algebra and stencil solvers

People

Christie Alappat, Jan Laukemann

Lightweight MPI+X-parallel sparse matrix-vector multiplication library

We have initiated the development of a lightweight and user-friendly library specifically designed to optimize the performance of the sparse matrix-vector multiplication (SpMV) operation. This operation is widely used in numerous applications and often becomes the primary performance bottleneck. The implementation is based on a hardware-efficient data storage format called SELL-C-σ, which is proven to deliver excellent performance on CPUs, GPUs, and wide-SIMD manycore accelerators. The library will employ hybrid MPI+X parallelization, enabling efficient distribution of computational tasks across multiple nodes or processors. Furthermore, it will provide advanced features such as communication hiding and domain partitioning, further optimizing the overall performance and usability of the SpMV operation.

Performance optimization of matrix power kernels

In 2021, we initiated a proof-of-concept and commenced our research on leveraging the analytical graph concepts embedded within our RACE framework to employ cache blocking for sparse matrix-power-vector multiplication (MPK) kernels. Building upon this initial groundwork, throughout 2022, we conducted extensive testing, performance analysis, and optimization to refine and finalize a fully functional cache-blocked MPK implementation.

The results of our research demonstrated that cache blocking techniques on modern multicore processors can achieve a remarkable performance improvement of up to 5× in MPK computations. To disseminate our discoveries, we documented our findings in a research paper titled "Level-Based Blocking for Sparse Matrices: Sparse Matrix-Power-Vector Multiplication," which has been accepted for publication by the "IEEE Transactions on Parallel and Distributed Systems" journal [7].

Speeding up iterative solvers using RACE

The success of our cache-blocked MPK led us to start research in the direction of integrating the optimized MPK into applications. We find that many iterative solvers and their components use MPK; using the cache-blocked variant of RACE can yield a significant performance boost for these applications. Some initial feasibility tests show promising results. Our future plan is to integrate our optimized MPK into the Trilinos framework, which will allow us to experiment with various solvers.

Distributed-memory parallel matrix power kernels

Another line of research which we have just kicked off is to enable distributed, MPI-level parallelism for cache-blocked MPK. This will allow the use of our cache blocking optimization on large-scale systems. An initial implementation and investigation in this regard has already been started in the form of a master thesis by Dane Lacey.

Building blocks for sparse linear algebra and stencil solvers: Adaptive linearized storage of sparse tensors

In collaboration with Intel and the University of Oregon, Jan Laukemann continued his work on the mode-agnostic Adaptive Linearized Tensor Order (ALTO) format for sparse tensor decomposition, which is an important method for extracting latent information from high-dimensional (multi-modal) sparse data. They introduced an opportunistic conflict resolution algorithm, in which threads collaborate instead of contending on memory access to discover and resolve their conflicting updates on-the-fly. As a result, the framework delivers superior in-memory performance compared to prior state-of-the-art, and is the only framework capable of processing out-of-memory tensors, achieving up to 33.35x speedup over the state-of-the-art mixed-mode compressed sparse fiber (MM-CSF) on a range of real-world sparse tensors [15].

6.4 Software engineering for HPC and data analytics

The research presented in this area is a joint work of NHR@FAU and the "Lehrstuhl für Systemsimulation" (LSS), where mainly the group of Prof. Köstler is involved.

People

Rafael Ravedutti, Sebastian Kuckuk, Sam Jacob, Sara Faghih-Naini, Kajol Kulkarni

Generation of Higher-Order Discretizations Deployed as ExaSlang Specifications (GHODDESS)

In 2022, the paper about our novel parameter-free adaptivity indicator was published [11]. Furthermore, we completed and submitted our paper about the masked block-structured approach to Advances in Water Resources [12]. In addition, we finished the implementation of the distributed p-adaptive algorithm and intensively analyzed the performance. Several optimizations were implemented into *GHODDESS* and *ExaStencils* in this regard. A paper about the results was submitted to ACM "Transactions on Mathematical Software" [10]. In the *KONWIHR* project *Continuous Benchmarking for the GHODDESS framework*, we established a CI/CB pipeline to track the performance over time, e.g., to achieve direct feedback after code changes.

Lastly, we combined *GHODDESS* with pystencils and StencilStream to generate discontinuous Galerkin shallow water simulations for FPGAS. We compared performance results to an earlier FPGA implementation of our unstructured code UTBEST. The results are published in [8].

Performance engineering and code generation for molecular dynamics and particle simulations

Within the NHR Alliance central project *Optimierung von Bibliotheken für datenparallele Prozessorarchitekturen*, Rafael Ravedutti Lucio Machado provided support in analyzing *libxc*, a library for the computation of exchange-correlation and kinetic energy functionals

for density-functional theory. The analysis focused on the overall workflow, the compute kernel composition, the Maple-to-C scripts, and potential for optimization. Based on this, he provided suggestions and a proof-of-concept implementation by extracting the parts of interest from *libxc* and providing an easy way to replace and evaluate the correctness of these kernels with the new implementation using Python and SymPy. A part of the results was presented in a project poster accepted to ISC'22.

Moreover, we worked on the development of *MD-Bench* (https://github.com/RRZE-HPC/MD-Bench), a performance-focused prototyping harness to evaluate the performance of molecular dynamics (MD) kernels with several strategies and configurations on different machines. Recently, the implementation of the GROMACS M×N pair list scheme was introduced into *MD-Bench*, making it the first proxy app with such an optimization. This is particularly important for studying the performance improvements obtained with it and evaluate whether this optimization can be leveraged in other particle simulation scenarios such as Discrete Element Methods (DEM), Smoothed-Particle Hydrodynamics (SPH), and other MD simulations [14] beyond GROMACS (such as the material modeling simulation cases that currently can only be carried out in LAMMPS, for instance). *MD-Bench* was also used for teaching purposes during the MuCoSim lecture, where students evaluated the parallel performance from its CPU kernels on the *Fritz* cluster, and also implemented the CUDA variants for its kernels and evaluated their performance in the GPU nodes from the *Alex* cluster. *MD-Bench*'s first paper entitled "MD-Bench: A generic proxy-app toolbox for state-of-the-art molecular dynamics algorithms" was published in the "14th International Conference on Parallel Processing and Applied Mathematics", in Gdansk, Poland, and received the Best Paper Award (https://hpc.fau.de/2022/09/14/md-bench-best-paper-award/), an extended version of the paper should be published in a "Special Issue of the Future Generation Computer Systems" (FGCS) journal. The work presents the features from *MD-Bench*, as well as different use-cases and interesting performance insights from MD short-range force calculation kernels.

Finally, Rafael also develops *P4IRS* (https://github.com/rafaelravedutti/pairs), a framework for the code generation of particle simulation kernels in C++, targeting multi-CPU and multi-GPU clusters using OpenMP, MPI, and CUDA. The framework uses a simple symbolic description in Python to setup the simulation and to describe the potentials and/or force fields, and then generates optimal code for the chosen hardware using domain knowledge from such simulations achieved in *MD-Bench* experimentation and performance studies.

Surrogate models for computational fluid dynamics

The analysis of the flow field using the Navier-Stokes equation is a computationally expensive and complex task. Currently, there is a dearth of a generalized 3D dataset that may be utilized as a common reference. This study focuses on generating and validating a generalized fluid dynamic dataset for 3D artificial single geometries and performing supervised training using a CNN-based surrogate model to infer average velocity components. The generated dataset can be used as a benchmark to serve as a basis of comparison for evaluating the performance and comparing the results of different algorithms or models and the trained weights can also be used for pre-training.

The *WaLBerla Framework* is used to produce a total of 6,260 random single geometry simulations up to Reynold's number of 10,000, with objects positioned at various locations in the domain using a Python based automation script. Several tests are carried out in order to validate the obtained dataset. The work goes on to predict average fluid

velocity components by training the deep-learning model with the generated dataset. The study also focuses on experimenting with the pre-processing of data, investigating the influence of data size on model training, the impact of model components on model performance, and hyperparameter tuning to achieve better accuracy. The best model produced a 0.0073 mean absolute error and 0.0081 root mean squared error in the test dataset.

6.5 Projects

The Research division of NHR@FAU led and participated in several research projects in 2022. Funding sources were the NHR Alliance, the DFG, and the EU *Horizon 2020* program.

6.5.1 NHR Alliance central projects with NHR@FAU partners

Cx as a service for sustainable HPC research software engineering (PI: Harald Köstler)
A large number of active developers contributing to the same code base increases the probability of new bugs being introduced because not all possible configurations are being tested. This either decreases the quality of the code or requires significant amounts of human resources to be spent on integration tests before a new release is made. Continuous Integration (CI) can reduce both the amount of time necessary for these tests, as well as provide near instant feedback after bugs that cause build failures have been introduced, by automatically building all possible configurations at given points in time, e.g., after every source code commit. In addition, Continuous Benchmarking (CB) frames not only unified testing but also a continuous benchmark environment. Finally, Continuous Deployment (CD) enables scientists to provide efficient, reliable, and sustainable HPC software.

Optimization of Software Libraries for Data-Parallel Processor Architectures (PI: Bernd Meyer)
The community of developers of atomistic simulations code created over the last decade versatile usable libraries for the computation of electronic structures of molecules and solid-state materials based on density functional theory (DFT). This project aims to migrate and optimize functionality of the libraries libxc, libint, libcint, and FFTXlib to support data-parallel processors (manycore CPUs, GPUs, FPGAs). Both, a manual, proof-of-concept optimization approach as well as automated code-generation techniques are explored.

PerfLab: NHR Performance Lab (PI: Gerhard Wellein)
Modern HPC architectures with specialized hardware and software are forcing code developers/users to acquire a deeper understanding of the hardware-software relationship, together with knowledge modeling methods and tools, to be able to apply manual or (semi-)automated code optimizations. The NHR centers at ZIB, FAU, UPB, RWTH and GWDG are combining their activities in the "NHR Performance Lab," with a long-term focus on helping the NHR community to broadly exploit the potential of modern computing architectures.

Standards and interfaces for a system-wide job-specific performance monitoring (PI: Jan Eitzinger)
There already exist a number of solutions for the system-wide, continuous performance monitoring at the NHR centers. These solutions have been developed

partly in cooperation and partly as isolated efforts. This project aims to create interoperability and modularity among these solutions. The goal is, on the one hand, a common solution in which some components remain adaptable or interchangeable and, on the other hand, common standards on which data is collected in which form and granularity. From July 18–19, 2022, NHR@FAU organized an in-person monitoring workshop in Erlangen[2]. 2022 was the last year of funding but for 2023, the follow-up NHR pathfinder project PathoJobs (coordinated by TU Darmstadt) was accepted.

Quantum Computing and Integer Programming (PI: Alexander Martin)
The aim of the project is to investigate the potential for simulation of QC systems in the context of integer programming. In particular, we focus on the QUBO problems as they are natural benchmark problems for QC and well-studied in integer programming. To this end, we will also develop a GPU-based reference algorithm for QUBOs to be able to evaluate the approaches of QCs to performance, both real and simulated.

6.5.2 DFG project

Dynamic HPC software packages: Seamless integration of existing software packages and code generation techniques
Complex phenomena in the natural and the engineering sciences are increasingly being studied with the help of simulation techniques. This is facilitated by a dramatic increase of the available computational power, and Computational Science and Engineering (CSE) is emerging as a third fundamental pillar of science. CSE aims at designing, analyzing, and implementing new simulation methods on high-performance computing systems such that they can be employed in a robust, user-friendly, and reliable fashion to a wide variety of scientific and engineering problems. Considering the high cost of supercomputing, reaching the best possible computational efficiency becomes a primary criterion and is central to the research agenda of CSE. Furthermore, HPC software must support a range of increasingly complex applications on modern heterogeneous and volatile hardware platforms, where often many different algorithms are combined to model interacting physical processes. For that, HPC software often has to be modified extensively to make full use of the additional performance of newly released architectures. Our main goal is to provide a new class of dynamic software frameworks to the HPC users that combine existing and established HPC frameworks with current code generation technology to increase the productivity when introducing new applications or porting to new platforms. We will show the benefits of this approach for three real-world multi-physics applications: (1) optimization of wind turbines and wind farms, (2) formation and dynamics of dunes that occur in many environmental systems such as riverbeds, and (3) simulation of charged particles in microfluidic flows.

6.5.3 BMBF projects

DAREXA-F: Data reduction for exascale applications in fusion research
For the efficient usage of HPC applications in the exascale era, we need to improve the scalability on large and heterogenous systems. This requires a variety of components, such as efficient processing, data storage, software, and algorithms.

[2] https://hpc.fau.de/2022/07/07/nhr-monitoring-workshop-2022/

The goal of this BMBF-funded project is to develop new methods for reducing data traffic between compute nodes with distributed memory and storage in file systems on supercomputers. For this purpose, a co-design approach will be used to develop solutions for variable-precision computation, data compression, and novel data formats. These solutions will be used to improve GENE, a program used worldwide for the simulation of plasma turbulence, and will be validated using GENE (http://genecode.org/). This would be a breakthrough in plasma physics with global visibility and many novel application possibilities, resulting in an acceleration of fusion research. Furthermore, new insight and methods of the *DAREXA-F* project can be transferred to other research areas and be made available for teaching and research of a broad audience. The project is a collaboration between Max Planck Computing and Data Facility (MPCDF), Garching, Max-Planck-Institut für Plasmaphysik (IPP), Garching, Technische Universität München, Garching, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, and ParTec AG, München.

EE-HPC: Energy efficient HPC

The energy consumption of HPC data centers is a decisive factor in the procurement and operation of the systems. EE-HPC (https://eehpc.clustercockpit.org/) achieves a more efficient energy use of HPC systems by targeted job-specific control and optimization of the hardware configuration as well as of settings of the runtime environments.

The aim of the project is the automated optimization of the energy efficiency of HPC systems. An innovative monitoring system is to contribute to reducing energy consumption while simultaneously increasing computing performance. This goal is to be achieved by new software-based control mechanisms of system parameters. The adjustment of system parameters, such as the utilization of computing nodes, is to take place automatically. A monitoring software coupled with a novel user interface shall provide the user with a transparent platform to also decide on the energy efficiency part of the computing load. This holistic approach ensures flexible and broad use for a wide range of applications.

StrömungsRaum - Novel Exascale-Architectures with Heterogeneous Hardware Components for Computational Fluid Dynamics Simulations

For applications to efficiently exploit the power of exascale systems, scalability must be improved on very large and heterogeneous systems. A variety of components are required for modern high-performance computing: from processors to data storage and file systems to software and algorithms. All of these components also require new technologies and adaptations to specific applications and interfaces. The goal of the project StrömungsRaum (https://gauss-allianz.de/de/project/title/StroemungsRaum) is to improve the scalability of the open source software FEATFLOW from the field of computational fluid dynamics (CFD) for application on exascale architectures with heterogeneous hardware components. This should enable finer-resolution and more complex computations and improve energy efficiency by reducing computation time. The core of the work consists of the development of novel numerical solution methods, such as so-called multigrid solvers and highly scalable domain decomposition methods, which will be tested and validated within the project. The innovative core is a novel, scalable solution for flow simulations with subsequent implementation in modern heterogeneous (exascale) architectures. Due to the improved scalability with simultaneous

increase in efficiency, high-resolution simulations can be generated for industrial use. One example is flow simulations in chemical reactors to achieve the most complete reaction and high yield. The open-source approach additionally provides a high degree of broad-scale effectiveness.

6.5.4 EU project EoCoE-II

Many pressing questions about the future global energy supply lead to highly complex scientific problems that are increasingly being researched with the help of simulations on supercomputers. The scientific topics range from photovoltaics to the use of geothermal energy or the design of wind farms to plasma physics for the possible future use of fusion energy. Simulations replace and complement expensive and lengthy experiments. Together with several partners from different European countries (France, Germany, Italy, Belgium, Great Britain, Spain, Poland), a continuation application to the expiring EU project "EoCoE" (Energy oriented Centre of Excellence) was submitted in 2018, funded by the *Horizon 2020* project (https://www.eocoe.eu) and ended in 2021. An extension as EoCoE-III is requested currently and would start 2024.

At FAU, in addition to the NHR@FAU, the "Chair of System Simulation" (Prof. Rüde) is also involved. The HPC group supports the application developers from the other project parts in the area of performance engineering. This includes, among others, the organization of courses and tutorials but also "hackathons", where project collaborators can apply advanced performance analysis techniques using their own simulation codes.

These codes include *Alya* for simulating multi-physics problems, *GYSELA*, a code for simulating plasma turbulence in fusion reactors, and *EURAD-IM*, a program for predicting and analyzing air quality. In the case of *Alya*, GPU optimization strategies developed using a simple code version ("proxy app") could be applied to many parts of the production code, resulting in a massive speedup. *EURAD-IM* and *GYSELA* benefited from detailed performance analysis and code quality improvement, especially with respect to vectorization.

6.6 PhD defenses

- Julian Hammer, *Design and Implementation of an Automated Performance Modeling Toolkit for Regular Loop Kernels*[3]
- Faisal Shahzad, *Efficient Application-level Fault Tolerance Methods for Large Scale HPC Applications*[4]

6.7 Publications

[1] A. Afzal, G. Hager, and G. Wellein. "Analytic performance model for parallel overlapping memory-bound kernels." *Concurrency and Computation: Practice and Experience* 34.10 (2022), e6816.

[3] https://hpc.fau.de/2022/12/18/we-congratulate-dr-julian-hammer
[4] https://hpc.fau.de/2022/06/15/we-congratulate-dr-faisal-shahzad-on-his-doctorate

[2] A. Afzal, G. Hager, and G. Wellein. *DisCostiC: A DSL-based Parallel Simulation Framework using First-Principles Analytic Performance Models*. Poster at the Platform for Advanced Scientific Computing (PASC'22) Conference, June 27–29, 2022, Basel, Switzerland. June 27–29, 2022.

[3] A. Afzal, G. Hager, and G. Wellein. "The Role of Idle Waves, Desynchronization, and Bottleneck Evasion in the Performance of Parallel Programs." *IEEE Transactions on Parallel and Distributed Systems, TPDS* (2022).

[4] A. Afzal, G. Hager, and G. Wellein. *White-box Modelling of Parallel Computing Dynamics*. Poster at the 5th International Conference on High Performance Computing in Asia-Pacific Region (HPC Asia'22), Jan 12-14, 2022, Virtual, Online. Jan. 12–14, 2022.

[5] A. Afzal, G. Hager, G. Wellein, and S. Markidis. "Exploring Techniques for the Analysis of Spontaneous Asynchronicity in MPI-Parallel Applications." *Proceedings of the International Conference on Parallel Processing and Applied Mathematics* (Gdansk, Poland). PPAM'22. Sept. 11–14, 2022.

[6] A. Afzal, G. Wellein, and G. Hager. "Addressing White-Box Modeling and Simulation Challenges in Parallel Computing." *Proceedings of the 2022 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation*. SIGSIM-PADS '22. Atlanta, GA, USA: Association for Computing Machinery, 2022, pages 25–26.

[7] C. Alappat, G. Hager, O. Schenk, and G. Wellein. "Level-Based Blocking for Sparse Matrices: Sparse Matrix-Power-Vector Multiplication." *IEEE Transactions on Parallel and Distributed Systems* 34.2 (2023), pages 581–597.

[8] C. Alt, T. Kenter, S. Faghih-Naini, J. Faj, J.-O. Opdenhövel, C. Plessl, V. Aizinger, J. Hönig, and H. Köstler. "Shallow Water DG Simulations on FPGAs: Design and Comparison of a Novel Code Generation Pipeline." *High Performance Computing*. Edited by A. Bhatele, J. Hammond, M. Baboulin, and C. Kruse. Cham: Springer Nature Switzerland, 2023, pages 86–105.

[9] S. Ejima, F. Lange, and H. Fehske. "Photoinduced metallization of excitonic insulators." *Phys. Rev. B* 105 (24 June 2022), page 245126.

[10] S. Faghih-Naini, V. Aizinger, R. Angersbach, S. Kuckuk, and H. Köstler. "p-adaptive discontinuous Galerkin method for the shallow water equations on heterogeneous computing architectures." *submitted to TOMS* (2023).

[11] S. Faghih-Naini and V. Aizinger. "p-adaptive discontinuous Galerkin method for the shallow water equations with a parameter-free error indicator." *International Journal on Geomathematics* 13.18 (Oct. 2022).

[12] S. Faghih-Naini, S. Kuckuk, D. Zint, S. Kemmler, H. Köstler, and V. Aizinger. "Discontinuous Galerkin method for the shallow water equations on complex domains using masked block-structured grids." *submitted to Advances in Water Resources* (2022).

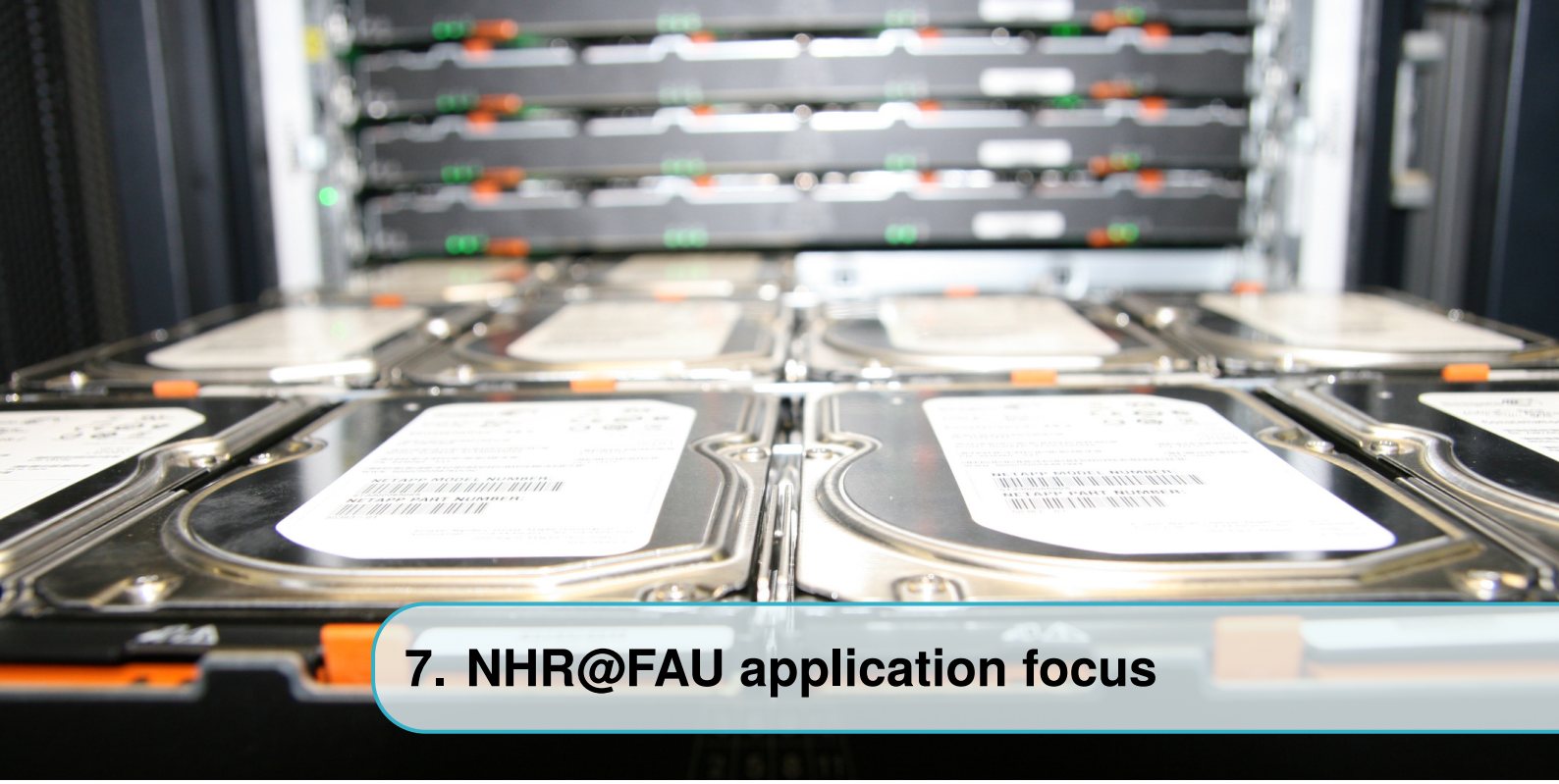
[13] A. Filusch and H. Fehske. "Tunable valley filtering in dynamically strained $\alpha - \mathcal{T}_3$ lattices." *Phys. Rev. B* 106 (24 Dec. 2022), page 245106.

[14] R. R. L. Machado, J. Schmitt, S. Eibl, J. Eitzinger, R. Leißa, S. Hack, A. Pérard-Gayot, R. Membarth, and H. Köstler. "tinyMD: Mapping molecular dynamics simulations to heterogeneous hardware using partial evaluation." *Journal of Computational Science* 54 (2021), page 101425.

[15] A. Nguyen, A. E. Helal, F. Checconi, J. Laukemann, J. J. Tithi, Y. Soh, T. Ranadive, F. Petrini, and J. W. Choi. "Efficient, out-of-memory sparse MTTKRP on massively parallel architectures." *Proceedings of the 36th ACM International Conference on Supercomputing*. 2022, pages 1–13.

6.8 Talks

- C. L. Alappat. "Performance Engineering for Sparse Matrix-Vector Multiplication with the Recursive Algebraic Coloring Engine." Online talk at the *NHR PerfLab Seminar*, February 1, 2022.
- A. Afzal. "The Role of Idle Waves in Modeling and Optimization of Parallel Programs." Online talk at the *NHR PerfLab Seminar*, April 26, 2022.
- T. Gruber. "Improving the hardware-software-mapping on EoCoE-II applications." EoCoE-II, final meeting, June 21, 2022.
- C. L. Alappat. "RACE: Speeding Up Sparse Iterative Solvers Using Level-Based Blocking Technique." Talk at MS41, ECCOMAS 2022, Oslo, Norway, June 8, 2022.
- C. L. Alappat. "RACE: Speeding Up Sparse Iterative Solvers Using Level-Based Blocking Technique." Talk at Sparse Days 2022, Saint-Girons, France, June 20, 2022.
- C. L. Alappat. "RACE: Speeding up Iterative Solvers and Spectral Clustering using Level-Based Blocking Techniques." Talk at MS5D, PASC 2022, Basel, Switzerland, June 29, 2022.
- A. Afzal. "Exploring Techniques for the Analysis of Spontaneous Asynchronicity in MPI-Parallel Applications." Paper presentation at PPAM 2022, the 14th International Conference on Parallel Processing and Applied Mathematics, Gdansk, Poland, September 11–14, 2022.
- G. Hager. "Spontaneous Asynchronicity—Parallel Programs out of Lockstep." Keynote at PPAM 2022, the 14th International Conference on Parallel Processing and Applied Mathematics, Gdansk, Poland, September 11–14, 2022.
- R. Ravedutti Lucio Machado, J. Eitzinger, H. Köstler, and G. Wellein. "MD-Bench: A generic proxy-app toolbox for state-of-the-art molecular dynamics algorithms." PPAM 2022, the 14th International Conference on Parallel Processing and Applied Mathematics, Gdansk, Poland, September 11–14, 2022
- F. Lange. "Infinite MPS simulations for 1D systems in and out of equilibrium." Talk at the group seminar of the Chair for Quantum Theory, Department pf Physics, FAU Erlangen-Nürnberg, October 26, 2022.
- G. Wellein. "Level-based Blocking for Sparse Matrix-Power-Vector Multiplication." CS Seminar, Berkeley National Laboratory, USA, August 19, 2022.
- G. Wellein. "Power, Energy, and HPC." Invited talk at the Workshop on Sustainability and Computational Science 2022, Lund University, Lund, Sweden, November 24, 2022.



7. NHR@FAU application focus

NHR@FAU offers the whole spectrum of both atomistic simulation methods and their fields of application: Scientists at FAU use methods from classical MD to quantum chemistry, and everything in between. Atomistic simulation methods are applied in chemistry, biology, physics, medicine, materials science, and engineering.

We have established a Germany-wide unique interdisciplinary competence center that helps users to select and use atomistic simulation methods in an HPC environment and actively accompanies and coordinates the development of high-performance simulation codes. An interdisciplinary approach promises not only synergy effects through the exchange and joint development of simulation and evaluation tools but a cross-fertilization of materials and life sciences, which often use the same or very similar simulation techniques.

A further core project is the education and lifelong training of scientists and engineers. The close cooperation between theory, simulation, and experiment ensures that the training is not aimed specifically at modelers but that it is made available to experimental colleagues as well. NHR@FAU makes an important contribution to the key technologies of scientific computing and software development through sustained concentration of methodological competence both in application and development of computer codes and their hardware-related optimization.

7.1 Atomistic Simulation Center (ASC)

The NHR centers in Paderborn, Berlin, and Erlangen have a strong focus on atomistic simulations. Researchers develop methods for and perform interdisciplinary simulations with various applications in the materials and the life sciences, spanning the areas of physics, biology, and chemistry. In order to join forces, share expertise, and bundle competence, the centers have formed the *Atomistic Simulation Center*. This umbrella organization serves as a platform providing guidance in tackling research questions by simulations.

Liaison scientists at the three centers, who are local experts in software and application domains, assist in providing individual support through advice and training on methods, software, tricks and pitfalls, best-suited architectures, and best practices.

The Atomistic Simulation Center held its second symposium on November 28–29, 2022. In 15 talks, PIs from the three centers and invited speakers presented their research in the different areas of atomistic simulations: Machine learning (ML) and neural networks, Quantum methods for many body systems, Many Body and Complex systems, Slow Dynamics in Liquid Matter, and ML for Biomolecular Simulations.

7.2 NHR@FAU Project Support by liason scientists

The NHR@FAU project ''b127dc - AOTTP-DFG18-1 - Characterization of molecular diffusion in electrolyte systems'' was supported by Marius Trollmann and Anna Kahler concerning simulation parameters and performance optimization.

From the project application

The main aim of the intended research project is to gain a fundamental understanding of the diffusive mass transport in electrolyte systems. For this, DLS experiments are combined with Equilibrium Molecular Dynamics (EMD) simulations to characterize the diffusive mass transport by the determination of molecular diffusion coefficients of systematically selected electrolyte systems. Here, diffusion coefficients for mixtures containing ions and molecules with varying size, structure, and charge distribution are intended to be studied over a broad temperature and composition range at macroscopic thermodynamic equilibrium. For electrolytes containing ions with a permanent electrostatic charge, special focus lies on the evaluation of the influence of intermolecular electrostatic interactions on the diffusive process. While EMD simulations are able to quantify this contribution, they strongly rely on the underlying force fields (FFs). In the further development and testing of existing FFs in this project, a special focus lies on the combination of compatible FFs for the different species of interest allowing a sound description of the interactions relevant in electrolyte mixtures. To access accurate diffusion coefficients by DLS, which also serve for the validation and development of the simulations, the technique should be further developed for the reliable measurement of diffusivities in electrolyte systems.

HPC necessity as given in the application

To obtain a statistically sound enough data point, comparatively large simulation times about 200–300 nano seconds must be performed. For this task, the GROMACS software package, which has shown to be one of best performing MD simulations software package for HPC systems, is used. Since GROMACS employs enhanced parallelization algorithms, MD simulations can efficiently be implemented in HPC clusters. In addition, GROMACS developers have been able to offload time-consuming calculation like non-bonded interactions to GPUs, making it more attractive to be used in HPC clusters like *Alex*. However, to improve the performance of our simulations in ns/day, benchmarking must be performed. Thus, to improve the performances and/or scalability of our simulations, we are highly interested in seeking the support of NHR@FAU. Based on benchmark simulations, HPC clusters *Alex* or *Fritz* will be selected for production simulations. The thermophysical properties of interest for the current research project, which are, namely, density, mutual diffusivities, thermodynamic factor and viscosities, are obtained in a post- processing manner from trajectories. For this, post-processing scripts, which take up to about 45 minutes, were developed and parallelized on multiple threads/nodes for an efficient use of compute time. These postprocessing scripts are planned to be run on the RRZEs woody cluster.

Problem definition

The usage of molecular dynamics (MD) simulations in the proposed project appears to be reasonable; however, several aspects should be taken into careful consideration beforehand. Firstly, the addition of new molecules to already existing force fields (FFs) or the combination of parameters from different FFs requires a careful evaluation and testing of the new parameter set. The parametrization strategy as well as the compatibility of the individual FFs must be considered in all cases. Secondly, the simulation of charged species requires reasonable settings for the simulation parameters to handle the electrostatic and the van der Waals interactions, as well as neighbor list searches (e.g., cutoffs, frequency). Arbitrary choices can lead to numerically unstable simulations and unphysical results. Thirdly, GROMACS has a plethora of implemented methods for MD simulations ranging from different timestep integration algorithms over several barostats and thermostats to controlling constraints placed on bonds. In addition, only some of these MD methods are implemented to run on GPU; thus, temperature coupling methods or constraint settings could be adjusted to gain performance without influencing the dynamics of the system. And finally, GROMACS runs out of the box on a large variety of hardware with relatively good performance but there are certain runtime parameters that can speed up the simulation significantly if chosen correctly. Furthermore, depending on post-production analyses, output control parameters have to be adjusted to avoid re-running of simulations in case certain data are needed.

Methods

To support the group on the aforementioned points as best as possible, simulation and post-processing protocols, previous publications, and related literature were requested. The setup of the production simulations was straightforward with 20,000 interaction sites, an integration time step of 2 fs, temperature and pressure setting in line with experimental setups, and a simulation length of about 300 ns. OPLS-AA was chosen as force field and bond-constraints were set on all-bonds using the LINCS algorithm. The equilibration step, performed with a 1 fs time step before the production, was split into two phases: 0.5 ns in the canonical ensemble with different initial velocities followed by 3.0 ns with isothermal-isobaric ensemble settings to attain the right pressure.

Main properties of interest, calculated in post-processing, were: Maxwell-Stefan diffusivities (requires center of mass (COM) positions of the molecules), thermodynamic factor (requires radial distribution functions related to COM of molecules), and viscosity (requires output from energies files and position of atoms).

The group sent four scientific papers detailing simulation parameters and calculation approaches:

- Klein et.al., published by the group, describes the determination of mutual diffusivity directly from MD simulations. This paper gives all related fundamentals as well as the simulation protocols together with involved parameters. Klein, et al., *J. Phys. Chem. B* 125, 5100–5113 (2021).
- García-Melgarejo et.al. describes a systematic way to develop OPLS non-bonded FF parameters of solvents used in electrolytes, enabling some level of reliable determination of transport properties. García-Melgarejo, et al., *J. Phys. Chem. B* 124, 4741–4750 (2020).
- Kirshnamoorthy et.al. is relevant for studying structural as well as some dynamic properties of electrolyte solution. Kirshnamoorthy, et al., *Phys. Chem. Chem. Phys.* 20, 25701–25715 (2018).

Table 7.1: Performance data obtained for the benchmark set from the NHR@FAU project ''b127dc - AOTTP-DFG18-1 - Characterization of molecular diffusion in electrolyte systems".

	allbonds_ nosehoover	hbonds_ nosehoover	allbonds_ vrescale	hbonds_ vrescale
A40	480 ns/day	586 ns/day	759 ns/day	775 ns/day
1 Fritz node	382 ns/day	417 ns/day	369 ns/day	418 ns/day
2 Fritz nodes	581 ns/day	598 ns/day	562 ns/day	608 ns/day
3 Fritz nodes	645 ns/day	724 ns/day	645 ns/day	720 ns/day

- Liu et.al. discusses the Maxwell-Stefan diffusivities of electrolytes systems consisting of ionic liquids. Liu, X., et al., *J. Phys. Chem. B* 115, 8506–8517 (2011).

Since the literature provided convincing evidence that the protocols and force field choice lead to reliable results for this type of simulation, it was decided to focus on finding the optimal thermostat and best constraint settings to yield maximum performance.

The group agreed to send simulation files that allowed us to setup different types of simulation benchmarks: (1) all-bond constraints with Nosé-Hoover thermostat, (2) all-bond constraints with v-rescale thermostat, (3) h-bond constraints with Nosé-Hoover thermostat, and (4) h-bond constraints with v-rescale thermostat. The reason for running four benchmarks is that not all available algorithms in an MD step are ported to GPUs, i.e., data has to be exchanged between GPU and CPU more often. For common simulation settings, calculations concerning the short-range non-bonded interactions, long-range non-bonded interactions (Particle Mesh Ewald), bonded interactions, and execution of update and constraints can be off-loaded to a GPU.

In this case, simulations conducted with the Nosé-Hoover thermostat require the calculation of update and constraints on the CPU, which can lead to overall performance decrease due to a larger communication overhead. Furthermore, the constraints setting in the GROMACS molecular simulations parameter file controls which bonds in the topology will be converted to rigid holonomic constraints; most often, constraints are only put on bonds containing hydrogen atoms because these bonds vibrate with a very short oscillation period of less than 10 fs. Converting all bonds between atoms to constraints might ensure atomic positions in relation to each other but leads to longer computation times and can, depending on the force field used, yield unreliable results. Covalent bonds are usually parameterized so that atoms have an optimal distance depending on the chemical environment.

Support & Performance analysis

Four benchmarks on one A40 GPU of *Alex* and on up to three *Fritz* nodes with GROMACS 2021.5 were performed. According to our blogpost (https://hpc.fau.de/2022/02/10/gromacs-performance-on-different-gpu-types/), systems with a relatively small number of atoms (up to 20,000 or 30,000 atoms) do not benefit from the professional A100 GPUs, thus we did not run benchmarks on this type of hardware. The data from the benchmark runs is listed in Table 7.1.

Notably, a slight decrease in performance is expected due to the overhead of writing to output files, as the mdp files of the production runs require ''nstenergy=10" (energies are written every 10 MD steps). The scaling on *Fritz* appears reasonable for all benchmarks; however, the performance of the v-rescale benchmarks is higher on the A40 compared to a total of three *Fritz* nodes. The performance of the hbonds benchmarks is higher

compared to the allbonds benchmarks and that is consistent over all choices of hardware and thermostats. Both equilibration steps are lower in performance on CPU and GPU (data not shown); allbonds_nosehoover shows the lowest performance values on GPU, followed by the hbonds_nosehoover. On *Fritz*, all benchmark versions reach a similar performance for the equilibration steps. For running on the A40 GPUs, we recommended setting the following environment variables to increase performance:

```
export GMX_GPU_PME_DECOMPOSITION=1
export GMX_USE_GPU_BUFFER_OPS=1
export GMX_DISABLE_GPU_TIMING=1
export GMX_ENABLE_DIRECT_GPU_COMM=1
export GMX_GPU_PME_PP_COMMS=true
export GMX_GPU_DD_COMMS=true
```

For the Nosé-Hoover thermostat, the update step has to be put back on the CPU but for v-rescale, the update step can be offloaded to the GPU; the benchmarks on GPU ran for 200,000 steps. The command used is: `gmx mdrun -s ⟨benchmark⟩ -nb gpu -pme gpu -bonded gpu -update ⟨gpu or cpu⟩ -ntomp 16 -ntmpi 1 -pin on -pinstride 1`

For running on *Fritz* nodes, the optimal runtime parameters have to be checked for each system and these are different for the four benchmark systems and the number of used nodes, so it would be necessary finding these parameters after deciding which contraints and thermostat will be used in this project. The command used on CPU is: `gmx_mpi mdrun -npme ⟨npme⟩ -resethway -nsteps 30000 -noconfout -maxh 0.1 -ntomp ⟨ntomp⟩ -dlb yes`

Moreover, links to an example GROMACS script on *Alex* (https://hpc.fau.de/systems-services/documentation-instructions/special-applications-and-tips-tricks/gromacs/#collapse_3), to the general Slurm documentation (https://hpc.fau.de/systems-services/documentation-instructions/batch-processing/), and to the batch processing on *Alex* (https://hpc.fau.de/systems-services/documentation-instructions/clusters/alex-cluster/#batch) were provided. Since the group has previous experience with running MD simulation on CPU on the *Emmy* cluster, they were informed about changes of available environment variables between Torque and Slurm. A comprehensive overview of how the commands have changed from Torque to Slurm can be found on the Woody transition announcement page (https://hpc.fau.de/2022/07/17/transition-from-woody-with-ubuntu-18-04-and-torque-to-woody-ng-with-almalinux8-and-slurm/).

As a final measure, the group was informed that the latest version of the GROMACS 2021 series was available on *Alex* (2021.6, gromacs/2021.6-gcc11.2.0-mkl-cuda), that it was released mid 2022 and fixed a (very rare) bug in the pressure coupling of simulations if ''update gpu" was enabled (https://manual.gromacs.org/2021-current/release-notes/2021/2021.6.html. The results of the benchmark tests should be not affected by switching to the 2021.6 release.

7.3 Activities of liason scientists

7.3.1 Dr. Frank Beierlein

Computer-Chemie-Centrum (CCC) and NHR@FAU

NHR activities, projects, and support

Frank Beierlein was part of the organizing team of the NHR Atomistic Simulation Center (ASC) Symposium 2022, which took place online from Nov 28–29, 2022. Here, he helped Petra Imhof with the preparations and was in close contact with the other ASC centers

(Paderborn and Berlin). Frank Beierlein also prepared a poster for the "HPC Village" at the Lange Nacht der Wissenschaften 2022. The regular NHR *liaison scientist* seminars and the AG Imhof group seminar serve as further opportunities to present and discuss current and planned activities. Close contact with the NHR@FAU HPC group is further ensured by regularly attending the NHR@FAU HPC Café.

As one of the two *liaison scientists* (LS) responsible for Amber, Frank Beierlein provides support for Amber customers of NHR@FAU, together with Anselm Horn. Frank Beierlein trained scientists for performing standard MD and thermodynamic integration free energy simulations with Amber on the NHR@FAU GPU cluster *Alex*. Jorge Amador Balderas from the Imhof group uses Frank Beierlein's scripts and protocol to investigate DNA repair mechanisms, while Eduard Neu from the Gmeiner group investigates opioid receptor ligands, and compares the Amber results to GROMACS. The Gmeiner group (NHR project <*Medchem-Dynamics*>, Susanne Gleich, Nico Staffen, Eduard Neu) is also supported in their effort to perform high-throughput docking of a large ligand database into the NK1 receptor, looking for agonists. Watching the Amber mailing list and the Amber literature ensures being up-to-date on what happens in the Amber community. Together with the NHR@FAU HPC group (Thomas Zeiser), Frank Beierlein optimized the setup of Amber 20/22 free energy thermodynamic integration simulations for efficient use of the new GPU cluster *Alex*.

Frank Beierlein is an expert in parameterizing nonstandard residues for Amber. Together with Anselm Horn (Sticht group), Frank Beierlein compares different approaches for the parameterization of non-standard amino acids.

In his current research on DNA repair mechanisms, Frank Beierlein carries on performing conventional, unbiased MD as well as alchemistic free energy methods, like thermodynamic integration, a technique that is increasingly used in pharmaceutical research. A HPC user report describing this work was published on the NHR@FAU HPC web pages. Frank Beierlein's NHR proposal (b106dc/DNARepairTDG) for computing time and storage on *Alex* was successful, which enables him and Petra Imhof's group to continue working on DNA repair by Thymine DNA Glycosylase. Additionally, Frank Beierlein supports a new DFG project in DNA repair of Petra Imhof (IM141/1-3), which was granted by DFG in 2022.

Petra Imhof and Frank Beierlein also closely collaborate with a large consortium around Andriy Mokhir (Organic Chemistry, FAU) in the preparation of a new DFG-funded Research Training Group (prospective new RTG 2885 "Organic Chemistry in Living Cells" – LiveOrgChem). In the past years, Frank Beierlein has published several papers with Andriy Mokhir which can be considered as preparative work and Frank Beierlein, Petra Imhof, and Andriy Mokhir currently write two publications on cancer-specific prodrug activation.

In their teaching (see below) and research activities, Frank Beierlein and Petra Imhof take care that documentation, scripts, and inputs are as user-friendly as possible to ensure that even unexperienced users or potential NHR customers can use advanced simulation techniques on HPC systems.

Teaching and training

Frank Beierlein is closely involved in teaching computational chemistry courses/hands on user trainings for the FAU students at the Computer-Chemistry-Center (CCC). These were in the winter semester (WS 21/22): *Moderne Softwareapplikationen* (2 courses), *Molecular Modeling* and in the summer semester (SS 22): *Computational Chemistry* (2 courses), *Molecular Modeling*, Bio-Organic & Bio-Inorganic LAB.

Publications

F. Beierlein, S. Volkenandt, P. Imhof, *Oxidation Enhances Binding of Extrahelical 5-Methyl-Cytosines by Thymine DNA Glycosylase*, *J. Phys. Chem. B* (2022), 126, 1188–1201.

Posters

NHR@FAU Computer-Chemie-Centrum/AK Imhof, *DNA-Reparatur im Supercomputer*, Lange Nacht der Wissenschaften, Erlangen, May 21, 2022.

Outreach

The NHR Atomistic Simulation Center (ASC) Inauguration Symposium 2022, held online on November 28–29, 2022, was not only a highly interesting and very successful scientific conference but it also offered the opportunity to inform a broader audience about NHR. The Molecular Modeling Workshop, which is annually organized in Erlangen by the Molecular Graphics and Modeling Society—Deutschsprachige Sektion e.V., was planned to take place September 13–15, 2022 but due to Corona, the meeting had to be shifted to spring 2023. Frank Beierlein is an active member in this society and supports the managing board as helping hand and cash auditor („Kassenprüfer“).

He also contributed scientifically by giving a talk and presenting a poster, using this opportunity to emphasize the excellent computing facilities and support of NHR@FAU, in addition to the scientific results. Together with Petra Imhof and the HPC staff, Frank Beierlein prepared a poster for the Lange Nacht der Wissenschaften in May 2022.

Collaboration with other liaison scientists and HPC staff

Together with Anselm Horn, Frank Beierlein is the responsible *liaison scientist* for Amber. Frank Beierlein also closely collaborates with Anselm Horn on the parameterization of non-standard amino acids and nucleic acids. Together with Thomas Zeiser, the Amber performance was optimized on the NHR@FAU HPC systems. Together with Anna Kahler, Harald Lanig and Thomas Zeiser, the Gmeiner group is supported. Together with Petra Imhof and the HPC staff, Frank Beierlein prepared a poster for the Lange Nacht der Wissenschaften 2022. Together with Petra Imhof, the other LS and Harald Lanig, Frank Beierlein was part of the organizing team of the NHR Atomistic Simulation Center (ASC) Symposium 2022.

7.3.2 PD Dr. Anselm Horn

Department of Medicine (Professorship for Bioinformatics) and NHR@FAU

NHR activities, projects, and support

Anselm Horn was involved in a cooperative project with the workgroup of Prof. Dr. Dr. Meinel (Institut für Pharmazie und Lebensmittelchemie, Julius-Maximilians-Universität Würzburg). His expertise on Amber force field parameterization allowed him to develop ff14SB-compatible parameters for a PEG3 and a PEG6 spacer unit. The project about the dynamics of a PEG-containing peptide linker covalently coupled to insulin-like growth factor was finalized and results were published in a joint research paper. [1] These simulations also served as preparatory work for a DFG proposal, which plans to use NHR@FAU for its computational aspects.

In addition, results of a cooperative project with Dr. Link-Paulus from the workgroup of Prof. Dr. Boßerhoff (Institut für Biochemie, Friedrich-Alexander-Universität Erlangen-Nürnberg) including MD simulations of a new splice variant of the regulatory protein Argonaute 2, which plays a role in melanoma, were published in a joint research paper. [2] Moreover, the project results were presented on a conference poster. [3]

Since Anselm Horn had been granted an early-user-access for the new clusters *Alex* and *Fritz* in order to especially test the large number of A40 cards in simulations, he designed the following project in 2022: the project aims at investigating the feasibility of modeling the binding of a ligand to a single chain antibody fragment (scFv) starting from many different initial ligand orientations to ensure truly independent simulation runs. Each of the 300 MD runs for the two protein systems covered a simulation time of 1 µs and also served as an AMBER burn-in test for *Alex*. The experience in job farming and data organization may be most valuable for other cluster users and NHR@FAU projects.

Anselm Horn participated in promoting NHR@FAU resources. These efforts allowed for example to attract the workgroup of Prof. Dr. Hildebrand (University Leipzig) as new user of NHR@FAU (Project short name: GPCRSCOMPEVO).

Performance Tests

Some initial performance tests for different Amber CPU-modules were performed on *Fritz*. It was found that the module impi-intel yielded the best performance for a medium sized system (175K atoms) for 1 to 4 nodes in MPI parallel execution.

A first performance test for the new Amber22 version compared to the older Amber20 version on *Alex* showed almost no performance gain in the newer version. Interestingly, for different systems from the standard Amber benchmark set the performance between the A40 and the A100 GPU cards was quite different: While in small systems (like JAC-NPT, 4 fs, ca. 24K atoms, opt) the speedup was less than 10%, large systems (like STMV-NPT, 4 fs, ca. 1,076K atoms, opt) experienced a speedup by a factor of ca. 1.5.

Teaching and training

In order to provide a first help for new users, Anselm Horn continued to work on an extension of the *Tips&Tricks* for Amber with a practical focus. His experiences, especially from the early-access-project, have been shared especially within other groups in Erlangen.

Outreach

Anselm Horn was part of the core organizational committee realizing the NHR Atomistic Simulation Center (ASC) Symposium 2022, which took place as online conference at 28–29 November, 2022. [4] He continued support on the Amber mailing list, helping Amber users all over the world to overcome problems, especially in the area of parameter generation. Thus, he helped increase the publicity of the abbreviation NHR@FAU as a kind of brand and the NHR initiative in general. [5]

Anselm Horn attended the German Conference of Cheminformatics (GCC) 2022 in Garmisch [6] together with Harald Lanig and presented both a scientific as well as an organizational poster describing the NHR@FAU capabilities within the ASC. The advertisement included the opportunity for potential users to obtain easy access to the NHR@FAU cluster machines and a voucher for free computational resources on *Alex* or *Fritz*.

For the public event Lange Nacht der Wissenschaften May 25, 2022, Anselm Horn prepared two posters. On those, science on NHR@FAU machines was explained to the interested public. Due to Corona issues, however, he could not join the event in person but a master student from the Bioinformatics group volunteered to do the actual on-site demonstration. [7]

Together with Harald Lanig, the head of the Molecular Graphics and Modeling Society—Deutschsprachige Sektion e.V., Anselm Horn performed initial planning of a Molecular Modeling Workshop (MMWS). Due to the pandemic, however, this conference was canceled by a decision of the board of directors. [8]

Exchange and Collaboration with other liaison scientists

Anselm Horn participated in an on-site exchange meeting with scientists from the Leibniz Rechenzentrum (LRZ) in Garching (July 6, 2022). He presented the idea of the ASC, in which the three NHR locations Erlangen, Paderborn, and Berlin join their expertise with respect to atomistic simulations.

He attended the first version of the online C++-CUDA-programming course held by Sebastian Kuckuk. After finishing the course and obtaining the official certificate, he provided detailed feedback in order to further improve the course.

Anselm Horn provided some support to Frank Beierlein with regards to his NHR proposal. Furthermore, he engaged in a collaboration project with Frank Beierlein with the aim of a joint publication of the two NHR@FAU PI research groups Imhof and Sticht. This MD simulation project also benefit from his expertise in AMBER parameter generation.

Administrative tasks

Anselm Horn acted as technical reviewer for NHR@FAU proposals.

References

[1] M. Beudert, L. Hahn, A. H. C. Horn, N. Hauptstein, H. Sticht, L. Meinel, R. Luxenhofer, M. Gutmann, T. Lühmann. "Merging bioresponsive release of insulin-like growth factor I with 3D printable thermogelling hydrogels," *Journal of Controlled Release* (2022), 347, 115–126.

[2] L. Linck-Paulus, A. H. C. Horn, A. Matthies, S. Fischer, G. Meister, H. Sticht, M. Kappelmann-Fenzl, A. Bosserhoff. "A previously unknown Argonaute 2 variant positively modulates the viability of melanoma cells," *Cellular and Molecular Life Sciences* (2022), 79, 475.

[3] L. Linck-Paulus, A. H. C. Horn, A. Matthies, S. Fischer, G. Meister, H. Sticht, M. Kappelmann-Fenzl, A. Bosserhoff. "A new Argonaute 2 splice variant influences microRNA function and melanoma cell viability" to be presented at the 73rd Mosbacher Kolloquium "The World of RNAs—Principles & Applications," Mosbach/Baden, Germany; March 31 to April 2, 2022.

[4] https://www.atomistic-simlab.hpc.fau.de/asc-symposium-2022/

[5] http://archive.ambermd.org/

[6] https://veranstaltungen.gdch.de/tms/frontend/index.cfm?l=10916&mod us=

[7] https://www.fau.de/2022/05/news/lange-nacht-der-wissenschaften-2022-highlights-an-der-fau/

[8] https://mmws2022.mgms-ds.de/

7.3.3 Dr. Sebastian Kuckuk

Chair of Computer Science 10 (System Simulation) and NHR@FAU

NHR activities, projects, and support

Sebastian Kuckuk has been heavily involved in providing support in efficiently utilizing the NHR@FAU GPGPU cluster *Alex* where his efforts span from micro-benchmarks to whole program performance evaluation. His benchmarks include data transfers between CPU and GPU, between different GPUs within a node and between compute nodes. Parts of his findings were used to optimize the system settings of *Alex*. After

observing performance issues of particular codes with the common pattern of reductions, Sebastian Kuckuk analyzed and compared different algorithmic approaches on GPUs. He presented his findings to colleagues and could provide recommendations on optimizing performance. Lastly, he conducted a performance evaluation of the widely used lattice QCD code *grid* after a user from the University of Regensburg observed unsatisfactory performance on *Alex*. Different configurations were measured and a simplified performance model was set up, which is in part based on earlier results of conducted micro-benchmarks. After improving performance, mostly by tuning job scripts, a good agreement of observed timings and the theoretical limit based on performance model estimations was found.

Furthermore, Sebastian Kuckuk continued to provide consultation for the group of Vadym Aizinger (University Bayreuth) on numerical solvers, computations on block-structured grids, and the efficient usage of GPUs. This included various virtual and in-person meetings as well as contributions to journal papers currently under review and a poster presentation at the *Platform for Advanced Scientific Computing (PASC)* conference 2022.

Teaching and training

Sebastian Kuckuk is a certified *NVIDIA Deep Learning Institute* (DLI) university ambassador. He further extended the range of courses he is certified to teach beyond *Fundamentals of Accelerated Computing with CUDA C/C++* to now also include *Fundamentals of Accelerated Computing with CUDA Python*, *Accelerating CUDA C++ Applications with Multiple GPUs*, and *Scaling CUDA C++ Applications to Multiple Nodes*. This allowed him to additionally be granted a specialization title in *Accelerated Computing* that, as of 2022, only a small number of people world-wide have.

In 2022, he organized and delivered a total of five courses at the NHR@FAU (Fundamentals with C++ in March, April and December, and Fundamentals with Python in August and September). An additional joint course with colleagues from the LRZ was conducted in November. All courses were delivered online and had a total of 175 registrations (with 36 no-shows) from 20 different academic institutions. In feedback forms, participants recommend the respective NHR@FAU course to colleagues with an average score of 4.8 out of 5 (55 total votes).

He also continued to further his knowledge about GPU programming and associated topics by attending the *GPU Technology Conference (GTC)* as well as courses and bootcamps on specific programming techniques including OpenACC, OpenMP and SYCL, and on broader fields such as deep learning and scientific machine learning.

His proficiency in GPU programming allowed him to contribute CUDA hands-on exercises to the popular tutorial on *Node-Level Performance Engineering* (NLPE). Moreover, work on a new course teaching and comparing different approaches to accelerator programming has begun and a first delivery is planned for 2023 alongside with deliveries of multi-GPU and multi-node DLI courses.

Meetings, travel, and community outreach

- Joint discussion on topics in training and support with colleagues from NHR@FAU and LRZ (Garching, July 6, 2022)
- Attendance in the workshop *Deep Learning and GPU programming using OpenACC* at HLRS conducted by colleagues from HLRS and LRZ; discussion about joint GPU programming courses (Stuttgart, July 12-14, 2022)

- LSS retreat meeting, participation and organization (Waischenfeld, July 18–20, 2022)
- Presentation of the results of the *Performance Optimization and Productivity (POP)* Assessment of GHODDESS (May 9, 2022, online)
- Presentation of *Fast GPU Reductions on Alex* (July 11, 2022, online)

Collaboration with other liaison scientists

- Sebastian Kuckuk continued to work with Rafael Ravedutti to provide consultation on code generation and performance optimization for the *libxc* framework in the scope of the project *Optimierung von Bibliotheken für datenparallele Prozessorarchitekturen*.
- Rafael Ravedutti supported Sebastian Kuckuk in the delivery of GPU courses as a teaching assistant providing (technical) support to attendees and answering participant questions.
- Sebastian Kuckuk supported Anna Kahler (NHR@FAU) in conducting an in-depth profiling of the widely used molecular dynamics software GROMACS on GPUs.
- Together with Dominik Ernst (NHR@FAU), Sebastian Kuckuk provided support for Vadym Aizingers group (University Bayreuth) in understanding and eliminating performance anomalies stemming from different types of memory allocation in generated ocean flow simulation codes.

Publications

- D. Zint, R. Grosso, V. Aizinger, S. Faghih-Naini, S. Kuckuk, H. Köstler. "Automatic Generation of Load-Balancing-Aware Block-Structured Grids for Complex Ocean Domains." Proceedings of the 2022 SIAM International Meshing Roundtable (SIAM IMR22).
- R. Angersbach, H. Köstler, S. Kuckuk. "Fusion of Massively-Parallel Simulation Frameworks and Code Generation Methodologies for Lattice Boltzmann and Multigrid Applications." Poster presentation at Platform for Advanced Scientific Computing (PASC) Conference 2022.
- S. Faghih-Naini, D. Zint, S. Kuckuk, S. Kemmler, V. Aizinger, R. Grosso, H. Köstler. "Shallow Water Simulations on Complex Ocean Domains using Block-Structured Grids." Poster presentation at Platform for Advanced Scientific Computing (PASC) Conference 2022.
- R. Angersbach, S. Kuckuk, H. Köstler. "Generating Coupling Interfaces for Multi-physics Simulations with ExaStencils and waLBerla." Accepted for publication in the proceedings of the 24th IEEE International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC 2023).

7.3.4 Dr. Samaneh Nasiri

Institute for Materials Simulation (WW8) and NHR@FAU

NHR activities, projects, and support

In 2022, Samaneh Nasiri was a Liaison scientist at NHR@FAU from January 1st to September 30th, 2022.

In March 2022, Samaneh Nasiri and Michael Zaiser submitted a DFG-funded proposal entitled "Molecular simulation of deformation processes in composites of light

metals and covalent nanoparticles." This project aims to find out how embedded nanoparticles change the deformation and fracture properties of the resulting nanocomposites and identify optimal parameters for mechanical property enhancement. This study involves large-scale atomistic simulations, which are planned to be carried out using the computation resources at NHR@FAU. To use the NHR@FAU resources, Samaneh Nasiri and Michael Zaiser prepared and submitted a proposal for Tier-2 HPC Access to NHR@FAU resources entitled "Probing the interplay between different deformation and failure mechanisms in lightweight metals reinforced by covalently bonded nanoparticles with molecular dynamics simulation."

The research will be conducted in collaboration with colleagues in the UK and China who investigate the same metal-carbon systems experimentally and with other simulation methods such as density functional theory. Samaneh Nasiri has performed preliminary investigations and proof-of-concept simulations required for the DFG-project proposal and the NHR@FAUproposal. This Project will be performed using *LAMMPS*, a classical molecular dynamics code with a focus on materials modeling. The current version of LAMMPS is installed and tested on the *Emmy* and *Fritz* cluster.

Samaneh Nasiri performed benchmark simulations to study the scalability of LAMMPS code on both *Emmy* and *Fritz* parallel cluster provided by NHR@FAU. A short summary of this study in given in the following part. To benchmark the scalability of LAMMPS code on the *Fritz* parallel cluster provided by NHR@FAU, we performed molecular dynamics simulation of tensile tests of Al samples with sizes between 10^6 and 10^9 atoms on a varying number of compute nodes (one to 64) of *Fritz* parallel cluster. The EAM potential of Ercolessi and Adams is employed to describe the interatomic interactions between Al atoms. We changed the number of atoms and number of nodes (N) on the *Fritz* cluster and calculated the corresponding scale factor.

The scaling factor (SF), a measure of scaling behavior, can be defined as $SF = \frac{N}{N_0}\left(\frac{P_0}{P_N}\right)$, where P_0 represents the performance of the simulation on N_0 number of nodes (reference value), while P_N represents the performance of the same simulation on N nodes. An example of "perfect" scaling would be when SF is close to 1. The lower SF value is a result of a weak scaling.

For systems with 10^6, 10^7, and 10^8 atoms, $N_0 = 1$, while for larger simulation samples with 10^9 atoms, $N_0 = 16$. Figure 7.1 shows the scaling factor versus the number of nodes. For the smallest system size, i.e., the system with 10^6 atoms, simulation efficiency decreases as the number of nodes increases from one to 16. The maximum number of four nodes should suffice for a one million atoms simulation, while for a simulation of ten and hundred million atoms, 16 and 32 nodes are still an efficient choice. For the most extensive simulation with one billion atoms (10^9 atoms), we used 16, 32, and 64 nodes on *Fritz*. The red data point in Figure 7.1 shows that the scaling factor for simulations of one billion atoms is very close to 1. Therefore, using 64 nodes on *Fritz* for our molecular simulation with one billion atoms has a 100% efficiency of HPC usage.

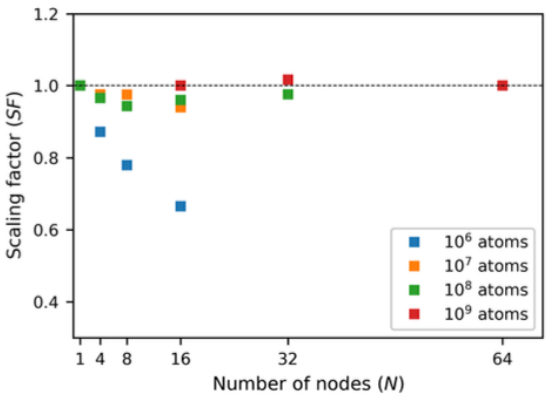


Figure 7.1: Scalability of our atomistic simulations using LAMMPS on *Fritz* cluster.

For this project, data visualization and post-processing will be performed using *OVITO* (https://www.ovito.org). OVITO is a scientific data visualization and analysis software for molecular and other particle-based simulation models. Samaneh Nasiri wrote a python script for data visualization and analysis which uses the Dislocation Extraction Algorithm (DXA) in OVITO to recognize partial dislocations and certain secondary grain boundary dislocations. The python script is tested and working on *Fritz* cluster. One of the main challenges in this project is to create the initial configuration of the composite systems with high density of defects. Samaneh Nasiri wrote a script which uses *Atomsk* (https://atomsk.univ-lille.fr) to create initial atomic configurations. This script works perfectly on *Fritz*.

Samaneh Nasiri supervised a Master thesis entitled "Investigation of the interface energy between Mg and SiC using DFT calculations." As part of the supervision, she trained the master student to get the DFT codes running using Quantum Espresso package on *Emmy* and *Meggie* clusters at NHR@FAU. Samaneh Nasiri revised and published one pre-reviewed journal paper on the atomistic simulations of metal-carbon nano-particles systems.

Meetings, travel, and community outreach

Samaneh Nasiri organized recurring monthly online meetings with the collaborators in the UK and China to discuss the preliminary investigation and progress of the above-mentioned project. She also contributed to the NHR@FAU newsletter for the April 2022 edition.

7.3.5 Rafael Ravedutti

Chair of Computer Science 10 (System Simulation) and NHR@FAU

NHR activities, projects, and support

Rafael works on the development of *MD-Bench* (https://github.com/RRZE-HPC/MD-Bench), a performance-focused prototyping-harness for short-range molecular dynamics kernels that comprises several optimization algorithms in the field. Recently, the implementation of the Cluster Pair algorithm from GROMACS was introduced into MD-Bench, making it the first application with a clean and simpler implementation of such kernels. This is particularly important for performance engineering efforts, which is the main focus of the project. A paper describing the tool with a few performance engineering show cases was published in the 14th International Conference on Parallel Processing and Applied Mathematics (PPAM 2022) conference and received the Best Paper Award for the Main Track. An invitation to publish an extended version of the paper was received and it is expected that this version will be published in the Future Generation Computer Systems (FGCS) journal in 2023. Despite this, many efforts on implementing such algorithms in the GPU were made along the year of 2022, which also led to teaching and advising activities for projects in the MuCoSim lecture and master thesis.

Rafael also develops *P4IRS* (https://github.com/rafaelravedutti/pairs), a framework for the code generation of particle simulation kernels targeting multiple hardware such as CPUs and GPUs. The framework uses a simple symbolic description in Python to setup the simulation and to describe the potentials and/or force fields, and then generates optimal code for the chosen hardware using domain-knowledge from such simulations achieved in MD-Bench experimentation and performance studies.

Teaching and training

In 2022, Rafael was responsible for the *Exercises for High-end Simulation and Practice (HESP)* lecture from the Chair of Computer Science 10 (System Simulation). The exercises consist of implementing particle simulations (MD, DEM and SPH) to execute efficiently on GPUs with domain-specific optimization algorithms such as the Linked Cells and Verlet Lists. He also supervises the final projects for the lecture which focus on extending the MD implementations from the exercises to run DEM and SPH simulations.

Rafael was also a tutor and responsible for the exercises in the *Practical Parallel Algorithms with MPI (PAMPI)* lecture in 2022. During the lecture, he provided support for students regarding MPI programming, performance evaluation, and optimization techniques for distributed-memory parallel applications.

Rafael also supervises the master project from Arsenii Andropov, which is still ongoing and focuses on extending the Cluster Pair algorithm in MD-Bench with the super-clustering strategy, leading to a more efficient use of the algorithm in a GPU by better exploiting the spatial locality of the atom clusters. The thesis should be finished by June 2023.

Meetings, travel, and community outreach

- Weekly meetings of NHR Research Day
- Talk *Introducing GROMACS optimizations into MD-Bench* (28.03.2022, virtual)

Collaboration with other liaison scientists

- Joint work with Sebastian Kuckuk on *Optimierung von Bibliotheken für datenparallele Prozessorarchitekturen* to provide consultation on code generation and performance optimization for the libxc framework.
- Rafael Ravedutti supported Sebastian Kuckuk in the delivery of GPU courses as a teaching assistant providing (technical) support to attendees and answering participant questions.

Publications

R. Ravedutti Lucio Machado, J. Eitzinger, H. Köstler, and G. Wellein. "MD-Bench: A generic proxy-app toolbox for state-of-the-art molecular dynamics algorithms." Accepted for PPAM (https://ppam.edu.pl/) 2022, the 14th International Conference on Parallel Processing and Applied Mathematics, Gdansk, Poland, September 11–14, 2022. PPAM 2022 Best Paper Award. Preprint: https://doi.org/10.48550/arXiv.2207.13094

7.3.6 Marius Trollmann

Department of Biology (Professorship for Computational Biology) and NHR@FAU

NHR activities, projects, and support

Marius Trollmann is a *liaison scientist* in the field of molecular dynamics (MD) simulations with a special focus on the software package *GROMACS*. His expertise lies in the simulation of bio-molecular systems employing GPUs or CPUs. He has assisted NHR funded projects to maximize the utility of the available resources. This includes, for example, the consultation of projects in the choice of suitable and efficient parameters for their simulations (e.g. thermostats) (Project ID: b127dc). He has also been involved in preparing run scripts to load multiple simulations on a GPU with the NVIDIA MPS server (Project ID: b118bb) and in the planning and execution of benchmarks for large-scale computing time projects (Project ID: b134dc, b174dc). Moreover, he is involved in the resolving of minor issues with the available GROMACS installations.

Scientific results

As a PhD-student, he applies and deepens his knowledge about the available resources in his own research projects. In a recent study, the lipid nanoparticles (LNPs) employed by BioNTech/Pfizer to transfer bioactive mRNA in human cells were characterized using atomistic MD simulations. Resources of NHR@FAU were used to setup the first atomistic simulation of a whole lipid nanoparticle (≈ 7.2 million atoms). Results of the work were presented at several conferences (poster and contributed talks, see list below) and were published in the Biophysical Journal (see Figure 7.2). A current project of him deals with the investigation and in-silico design of a novel antimicrobial peptide (manuscript in preparation).

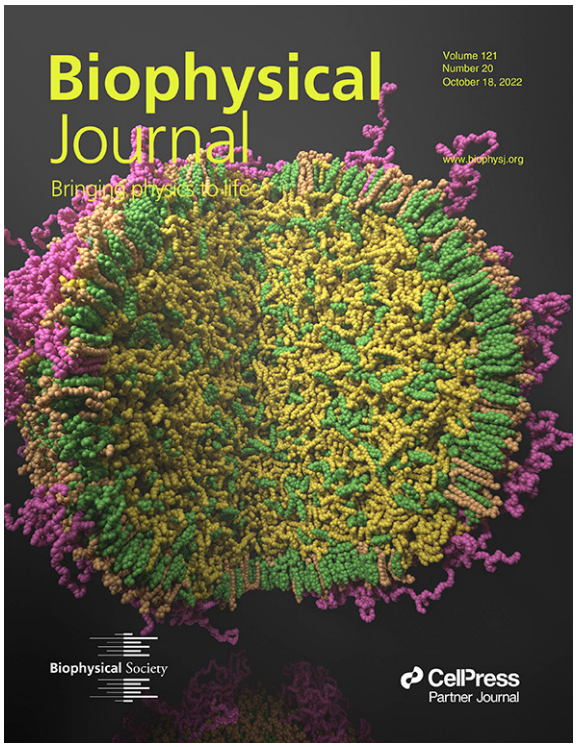


Figure 7.2: Atomistic structure of a whole lipid nanoparticle at the cover of the Biophysical Journal. The coloring of the molecules corresponds to the different types of lipids employed in the composition. Trollmann and Böckmann. *Biophys. J.* 121, 3927–3929 (2022).

Outreach

He contributed to the organization and execution of the second NHR Atomistic Simulation Center (ASC) Symposium by maintaining the virtual meeting during the symposium. A list of his further outreach activities can be found below.

Teaching and training

Marius Trollmann is involved in the preparation and execution of several modules focusing on programming or MD simulations. He contributed to the following modules:

- Orientierungsmodul Strukturbiologie I, Master Biology/Integrated Life Sciences
- Orientierungsmodul Strukturbiologie II, Master Biology/Integrated Life Sciences
- Python for Bioinformatics and Data Analysis, Master Biology/Integrated Life Sciences/Integrated Immunology
- Fachmodul Strukturbiologie, Bachelor Biology/Integrated Life Sciences

In the "Orientierungsmodul Strukturbiologie I" and "Orientierungsmodul Strukturbiologie II" courses, students are given the opportunity to prepare and conduct their own molecular dynamics (MD) simulations to work on a realistic research hypothesis. The course "Python for Bioinformatics and Data Analysis" offers an introduction to the popular programming language Python, highlighting its applications in natural sciences. In the "Fachmodul Strukturbiologie" course, students receive a one-week introduction to GROMACS, will cover the fundamentals of atomistic simulations, and are given examples of simple applications. Participants learn how to prepare and run simulations of bio-molecular systems, including membranes and proteins, at both atomistic and

coarse-grained levels of resolution. Additionally, common tools and workflows for processing and analyzing simulation trajectories are covered in the tutorial. For the first time, participants from NHR were enabled to take part in this course.

In response to the growing demand for knowledge on running simulations at high-performance clusters, a virtual three-day course tailored specifically to NHR participants is planned for October 2023. This course will provide in-depth instruction on utilizing GROMACS to run simulations at high performance clusters, ensuring that participants are equipped with the skills and knowledge necessary to excel in their research.

Meetings, travel, and community outreach

- Weekly virtual NHR@FAU liaison scientist meeting
- Thematic meeting of the Biophysical society—"Biophysics at the Dawn of Exscale Computers" (May 16–20, 2022, Hamburg, poster presentation)
- Lange Nacht der Wissenschaften (May 21, 2022, Erlangen, public event)
- German Biophysical Society Meeting 2022, (September 25–28, 2022, Konstanz, contributed talk)
- NHR Atomistic Simulation Center Symposium, (November 28–29, 2022, online)

7.3.7 Dr. Egor Trushin

Chair of Theoretical Chemistry and NHR@FAU

NHR Activities, projects, and support

Egor Trushin has been involved in the development and implementation of quantum chemistry and density functional methods in a number of computational quantum chemistry packages. The main contribution has been made to the Molpro Quantum Chemistry Software [1]. Egor Trushin contributed computational methods to Molpro, in the development of which he has been involved for several years. These programs will be available to users in the next Molpro release. σ-functional [2,3] calculations are also available in the latest version of the ADF [6] package, to which Egor Trushin contributed part of the code.

- Random Phase Approximation and σ-functional program [2,3], which allows accurate calculation of various chemical properties such as, e.g., reaction energies and barrier heights of chemical reactions or determination of geometries and vibrational properties of molecules.
- Numerically stable optimized effective potential (OEP) exact-exchange (EXX) program [4], which makes numerically stable OEP-EXX calculations feasible even with basis sets from standard basis set libraries, which was mostly impossible before.
- Kohn-Sham inversion program [5], which solves the fundamental problem of Kohn-Sham density functional theory to construct Kohn-Sham exchange-correlation potentials corresponding to given electron densities.

In 2022, Egor Trushin was involved in the publication of two articles

- J. Erhard, S. Fauser, E. Trushin, A. Görling. "Scaled σ-functionals for the Kohn-Sham correlation energy with scaling functions from the homogeneous electron gas." *J. Chem. Phys.* 157 (2022) 114105.

- J. Erhard, E. Trushin, A. Görling. "Numerically stable inversion approach to construct Kohn-Sham potentials for given electron densities within a Gaussian basis set framework." *J. Chem. Phys.* 156 (2022) 204124.

References

[1] Molpro Quantum Chemistry Software https://www.molpro.net/)

[2] E. Trushin, A. Thierbach, A. Görling. "Towards chemical accuracy at low computational cost: Densityfunctional theory with σ-functionals for the correlation energy." *J. Chem. Phys.* 154 (2021) 014104.

[3] S. Fauser, E. Trushin, C. Neiss, A. Görling. "Chemical accuracy with σ-functionals for the Kohn-Sham correlation energy optimized for different input orbitals and eigenvalues." *J. Chem. Phys.* 155 (2021) 134111.

[4] E. Trushin, A. Görling. "Numerically stable optimized effective potential method with standard Gaussian basis sets." *J. Chem. Phys.* 155 (2021) 054109.

[5] J. Erhard, E. Trushin, A. Görling. "Numerically stable inversion approach to construct Kohn-Sham potentials for given electron densities within a Gaussian basis set framework." *J. Chem. Phys.* 156 (2022) 204124.

[6] ADF (Amsterdam Density Functional) https://www.scm.com/doc/ADF/

7.3.8 Dr. Joachim Paier

Department of chemistry and pharmacy and NHR@FAU

NHR project

Joachim Paier is a *liaison scientist* in the field of *ab initio* methods for electronic structure computations. He was involved in the project *Implementation of σ-functionals into VASP*.

Proposed work

Orbital-dependent correlation energy functionals (*aka* implicit density functionals) are likely candidate functionals achieving chemical accuracy, i.e., 1 kcal/mol or 4.18 kJ/mol errors in actual calculations. Practical examples are the description of binding energies or energy barriers for gas-phase reactions on one hand or reactions in surfaces on the other hand. Specifically, chemical accuracy implies error bars within 1 kcal/mol when theory is compared with observed results. Recently, the group around Prof. Görling developed functionals based on the so-called *random phase approximation* (RPA) to improve its methodological accuracy at unaltered computational cost. This is remarkable because going beyond RPA based on quantum-mechanical reasoning involves substantial additional computational workloads (e.g., [1–2]). These functionals have been introduced into the literature as σ-functionals [1]. Assessments have successfully shown that molecular atomization energies and energy barriers (as well as many other properties with emphasis on main-group chemistry) significantly improve. Thus, σ-functionals de facto achieve chemical accuracy [1, 3–4].

With the aim in mind to extend their application towards the solid state, σ-functionals have been implemented into the *Vienna ab initio simulation package*, VASP [5], and they have been successfully tested (*vide infra*). The proposer (JP) has been financially supported by the NHR *liaison scientist* funds. VASP is a plane-wave-based commercial *ab initio* molecular dynamics code having a variety of many-body theories (e.g., RPA) implemented to solve electronic structure problems subject to periodic boundary conditions.

Table 7.2: Atomization energies of CH_2 und H_2O (kcal/mol) compared with experiment.

	Expt.	**RPA**		**σ-functional (PBE_W1)**	
		GTO [1, 4]	PW [7]	GTO [1, 4]	PW [7]
CH_2 ($^3B_1$)	189	179 (-10)	180 (-9)	187 (-2)	184 (-5)
H_2O	233	222 (-11)	222 (-11)	234 (+1)	230 (-3)

VASP is going to be used by many users of the Regionales Rechenzentrum Erlangen (RRZE, within the NHR association) and therefore herein discussed implementations may be beneficial to these users. Moreover, VASP is a worldwide employed code and—due to technical accuracy—considered as standard in solving the electronic structure problem underlying density functional theory (DFT) and beyond. A report of the present work accomplished within this *liaison-scientist* funded project is described and summarized in the following sections.

Implementation and results

1. Atomization energies of small molecules: Table 7.2 shows atomization energies in kcal/mol for CH_2 ($^3B_1$) and H_2O molecules obtained using Gauss-type orbitals (GTO) and plane waves (PW) as basis sets. Errors or deviations with respect to experiment (corrected for zero-point energy effects) are given in parentheses.

GTO results have been recently published by Görling and coworkers [1, 4]. Geometric structures of these molecules have been taken from the literature and were used in PW-based calculations. Note that (conventional) RPA has been employed, i.e., no-selfconsistency in the orbitals and orbital energies, which were obtained deploying the Perdew, Burke, Ernzerhof (PBE [6]) exchange-correlation functional. Regarding molecules, the herein tested σ-functional uses the parametrization W1 [1].

Atomization energies obtained using PW were calculated in a cell with dimensions of $8 \times 9 \times 9.5$ Å^3 and a PW cut-off to expand the orbitals of 1000 eV. Additional tests showed that extrapolation of the cell volume towards infinity based on a $1/V^2$ decay minorly contributes within 1 kcal/mol.

2. Energy versus volume curves: A commonly applied test case for a readily computable crystalline semiconductor is the dependence of energy (per cell) versus volume of silicon (Si), for instance in diamond structure. These are single point calculations using RPA, and σ-functionals based on W1, S1, and S2 parametrization [1, 4] to the PBE functional. These energies were fitted to Murnaghan's equation of state [8] and summarized in Table 7.3. Corresponding graphs to the numerical data are shown in Figure 7.3.

Total energies in above-mentioned calculations were converged using a $8 \times 8 \times 8$ k-point mesh and a 800 eV plane-wave cut-off. Note that the standard cut-off for the Si pseudopotential is 254 eV.

Discussion

Computed atomization energies of the open-shell CH_2 radical (high-spin triplet state) and the close-shell water molecule can be discussed in terms of three aspects. First, from a methodological point of view the RPA-calculations show that the effect of the basis set used is negligibly small. Discrepancies between GTO reference values and PW-based

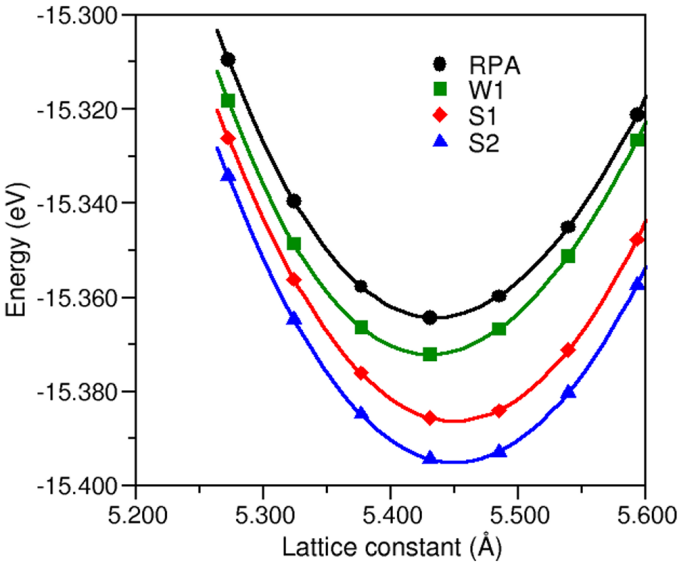


Figure 7.3: Total energy (eV) per Si- atom in the unit cell computed for several lattice constants centered around the equilibrium. These points were fitted according to Murnaghan's equation of state. (lines). Legends for the W1, S1, and S2 parametrization [1, 4] based on the PBE functional are given as inset.

results are within 1 kcal/mol. Results obtained using the σ-functional (for instance) parametrized in PBE_W1 show slightly larger deviations in terms of magnitude ranging between 3 to 4 kcal/mol. Current experience shows that finer frequency meshes most likely rectify this error. Nonetheless, these discrepancies are significantly smaller than deviations between RPA and observed results being ca. 10 kcal/mol. Furthermore, one can conclude that the σ-functional using PBE_W1 parametrization alleviates the pronounced underestimation of molecular atomization energies obtained using the RPA. It thus allows for an (on average) more accurate description of atomization energies. Using a converged GTO-basis set leads to a mean absolute error (MAE) of 1.5 kcal/mol, using basis-set-superposition-error-free plane-waves and pseudopotentials yields an MAE of 4 kcal/mol. These results clearly improve upon RPA, which yields an MAE of 10.5 kcal/mol (Table 7.2).

First applications of σ-functionals to solids, for instance crystalline silicon (Table 7.3; Figure 7.3) show that W1, S1, as well as S2 parametrizations to the PBE functional improve upon RPA lattice constants. These results are compared to observed results corrected for finite temperature anharmonic expansion effects [9]. The overestimation of 3% predicted by RPA is substantially reduced to within a range between 0.2 and 0.5%. Roughly speaking, bulk moduli are related to the curvature of the equation of state at equilibrium and is therefore a measure for the quality of total energies. The σ-functional

Table 7.3: Equilibrium lattice constants (Å) and bulk moduli (GPa) for Si in cubic diamond structure obtained using RPA and σ-functionals in W1, S1, and S2 parametrization to the PBE functional. Relative errors (%) with respect to experiment are given in parentheses.

	Expt.	**RPA**	**σ-functional**		
			W1	S1	S2
a_0 (Å)	5.421[a]	5.435 (+3.3)	5.430 (+0.17)	5.449 (+0.54)	5.449 (+0.54)
B_0 (GPa)	99.2	98.3 (-0.91)	100.5 (+1.3)	98.8 (-0.40)	97.8 (-1.4)

[a] Corrected for anharmonic effects [9].

based on the S1 parametrization (Table 7.3) improves upon RPA. Importantly, the error in the lattice constant, as well as in the bulk modulus decreases.

Summary and conclusions

Quantitative agreement between computed atomization energies for two example molecules and results published in the literature shows that σ-functionals were successfully implemented into the VASP code. Moreover, these calculation results show that σ-functionals within a specific parametrization describe molecular atomization energies better than the underlying RPA correlation energy. Therefore, σ-functionals have become available to widespread applications by VASP users. This project was supported by the NHR *liason-scientist* funding.

References

[1] E. Trushin; A. Thierbach; A. Görling, *J. Chem. Phys.* 154 (2021) 014104.

[2] A. Görling, *Phys. Rev. B* 99 (2019) 235120.

[3] J. Erhard; S. Fauser; E. Trushin; A. Görling, *J. Chem. Phys.* 157 (2022) 114105.

[4] S. Fauser; E. Trushin; C. Neiss; A. Görling, *J. Chem. Phys.* 155 (2021) 134111.

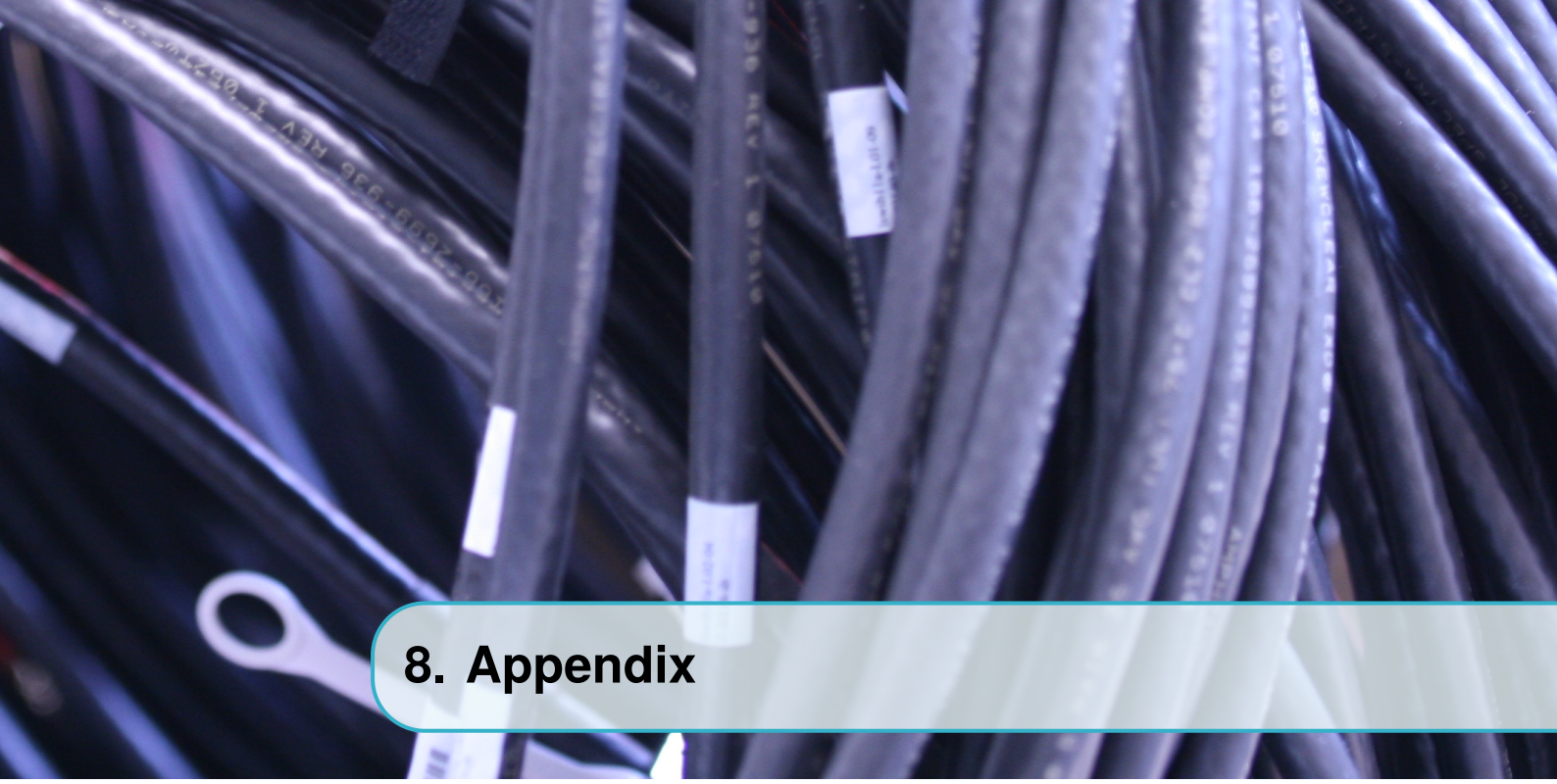
[5] VASP, The VASP manual. www.vasp.at (accessed 06/06/2020).

[6] J. P. Perdew; K. Burke; M. Ernzerhof, *Phys. Rev. Lett.*, 77 (1996) 3865–3868.

[7] Implemented and computed within NHR@FAU, 2022.

[8] F. D. Murnaghan, *Proc. Natl. Acad. Sci. USA* 30 (1944) 244–247.

[9] J. Harl; L. Schimka; G. Kresse, *Phys. Rev. B* 81 (2010) 115126.



8. Appendix

8.1 Publications which used NHR@FAU Resources

The following list contains publications from outside FAU that used NHR@FAU resources.

[1] K. Breitwieser, R. Dorta, D. Munz. "On the Magic Iodide- and Acid Effects in the Metolachlor Process." *Organometallics* (2022), 41, 3801.

[2] S. Frieß, A. Benyak, A. Herrera, A. M. Escalona, F. W. Heinemann, J. Langer, D. Fehn, D. Pividori, D. Munz, K. Meyer, R. Dorta. "Sulfoxide Pincer Supported Iridium Complexes: Facile Three-Electron Oxidative Addition of Iodine to Ir(I) and Its Reductive Elimination from Ir(IV)." *Inorg. Chem.* (2022), 61, 1236.

[3] A. Cadranel, L. Gravogl, D. Munz, K. Meyer. "Intense Photoinduced Intervalence Charge Transfer in High Valent NHC Iron Complexes." *Chem. Eur. J.* (2022), e2022002.

[4] F. Neururer, S. Liu, D. Leitner, M. Baltrun, K. Fisher, H. Kopacka, K. Wurst, L. Daumann, D. Munz, S. Hohloch. "A Mesoionic Carbene Complex of Manganese in Five Oxidation States." *Chem. Commun.* (2022), 58, 6096.

[5] S. Reith, S. Demeshko, B. Battistella, A. Reckziegel, C. Schneider, F. Meyer, D. Munz, C. G. Werncke. "Between Nitrene and Imidyl – A Trigonal High-Spin Iron Imide in Two Oxidation States." *Chem. Sci.* (2022), 13, 7907.

[6] W. Mao, D. Fehn, F. W. Heinemann, A. Scheurer, D. Munz, K. Meyer. "Umpolung in a Pair of Co-balt(III) Terminal Imido/Imidyl Complexes." *Angew. Chem. Int. Ed.* (2022), e202206848.

[7] A. Grünwald, B. Goswami, D. Momper, M. Gimferrer, B. Morgenstern, F. W. Heinemann, C. Kay, D. Munz. "Palladium Terminal Imido Complexes with Singlet Nitrene Character." *J. Am. Chem. Soc.* (2022), 144, 8897.

[8] K. Breitwieser, H. Bahmann, R. Weiss, D. Munz. "Gauging Radical Stabilization with Carbenes." *Angew. Chem. Int. Ed.* (2022), e202206390.

[9] D. Sarkar, L. Groll, F. Hanusch, D. Munz, S. Inoue. "Ligand Assisted CO2 Sequestration and Catalytic Valorization by an NHI-Stabilized Stannylene." *ChemCatChem* (2022), e202201048.

The following list contains publications from inside FAU that used NHR@FAU resources.

[1] I. Aillaud, S. Kaniyappan, R. R. Chandupatla, L. M. Ramirez, S. Alkhashrom, J. Eichler, A. Horn, M. Zweckstetter, E. Mandelkow, H. Sticht, and S. A. Funke. "A novel D-amino acid peptide with therapeutic potential (ISAD1) inhibits aggregation of neurotoxic disease-relevant mutant Tau and prevents Tau toxicity in vitro." *Alzheimer's Research and Therapy* 14 (2022).

[2] A. S. Ambekar, C. Schwarzmeier, U. Rüde, and V. Buwa. "Particle-resolved turbulent flow in a packed bed: RANS, LES, and DNS simulations." *AIChE Journal* (2022).

[3] M. Ammon, A. Raabgrund, and M. A. Schneider. "Adsorption, self-assembly and self-metalation of tetra-cyanophenyl porphyrins on semiconducting CoO(100) films." *Surface Science* 720 (2022).

[4] M. Andernach, J. Turton, and T. Mölg. "Modeling cloud properties over the 79 N Glacier (Nioghalvfjerdsfjorden, NE Greenland) for an intense summer melt period in 2019." *Quarterly Journal of the Royal Meteorological Society* (2022).

[5] S. Bachhuber, D. Weber, I. Weygers, and T. Seel. "RNN-based Observability Analysis for Magnetometer-Free Sparse Inertial Motion Tracking." *2022 25th International Conference on Information Fusion, FUSION 2022* (Linkoping). Institute of Electrical and Electronics Engineers Inc., 2022.

[6] M. Beudert, L. Hahn, A. Horn, N. Hauptstein, H. Sticht, L. Meinel, R. Luxenhofer, M. Gutmann, and T. Luehmann. "Merging bioresponsive release of insulin-like growth factor I with 3D printable thermogelling hydrogels." *Journal of Controlled Release* 347 (2022), pages 115–126.

[7] P. Boonsawat, A. H. C. Horn, K. Steindl, A. Baumer, P. Joset, D. Kraemer, A. Bahr, I. Ivanovski, E. M. Cabello, M. Papik, M. Zweier, B. Oneda, P. Sirleto, T. Burkhardt, H. Sticht, and A. Rauch. "Assessing clinical utility of preconception expanded carrier screening regarding residual risk for neurodevelopmental disorders." *npj Genomic Medicine* 7 (2022).

[8] M. Conrad, C. Söldner, and H. Sticht. "Effect of Ions and Sequence Variants on the Antagonist Binding Properties of the Histamine H1 Receptor." *International Journal of Molecular Sciences* 23 (2022).

[9] P. M. Dinh, M. Vincendon, F. Coppens, E. Suraud, and P.-G. Reinhard. "Quantum Dissipative Dynamics (QDD): A real-time real-space approach to far-off-equilibrium dynamics in finite electron systems." *Computer Physics Communications* 270 (Jan. 2022).

[10] R. Eschenbacher, T. Xu, E. Franz, R. Löw, T. Moje, L. Fromm, A. Görling, O. Brummel, R. Herges, and J. Libuda. "Triggering the energy release in molecular solar thermal systems: Norbornadiene-functionalized trioxatriangulen on Au(111)." *Nano Energy* 95 (2022).

[11] F. Fan, B. W. Kreher, H. Keil, A. Maier, and Y. Huang. "Fiducial marker recovery and detection from severely truncated data in navigation assisted spine surgery." *Medical Physics* (2022).

[12] G. Fickenscher, L. Fromm, A. Görling, and J. Libuda. "Atomic Layer Deposition of HfS2 on Oxide Interfaces: A Model Study on the Initial Nucleation Processes." *Journal of Physical Chemistry C* 126 (2022), pages 21596–21605.

[13] L. Folle, K. Tkotz, F. Gadjimuradov, L. Kapsner, M. Fabian, S. Bickelhaupt, D. Simon, A. Kleyer, G. Krönke, M. Zaiß, A. M. Nagel, and A. Maier. "Towards Super-resolution CEST MRI for Visualization of Small Structures." *Bildverarbeitung für die Medizin 2022*. Edited by K. Maier-Hein, T. M. Deserno, H. Handels, A. Maier, C. Palm, and ThomasTolxdorff. Proceedings, German Workshop on Medical Image Computing, Heidelberg, June 26–28, 2022. Wiesbaden: Springer Vieweg, 2022, pages 210–215.

[14] E. Franz, D. Krappmann, L. Fromm, T. Luchs, A. Görling, A. Hirsch, O. Brummel, and J. Libuda. "Electrocatalytic Energy Release of Norbornadiene-Based Molecular Solar Thermal Systems: Tuning the Electrochemical Stability by Molecular Design." *Chemsuschem* (2022).

[15] M. Freiberger, M. Minameyer, I. Solymosi, S. Frühwald, M. Krug, Y. Xu, A. Hirsch, T. Clark, D. M. Guldi, M. von Delius, K. Amsharov, A. Görling, M. E. Perez-Ojeda Rodriguez, and T. Drewello. "Two Rings Around One Ball: Stability and Charge Localization of [1 : 1] and [2 : 1] Complex Ions of [10]CPP and C60/70[*]." *Chemistry - A European Journal* 29 (2022).

[16] L. Fusek, M. Kastenmeier, E. Franz, L. Fromm, A. Görling, O. Brummel, and J. Libuda. "Anchoring of porphyrins on atomically defined cobalt oxide: In-situ infrared spectroscopy at the electrified solid/liquid interface." *Surface Science* 718 (2022).

[17] M. Gambietz, M. Nitschke, J. Miehling, and A. Koelewijn. "What should a metabolic energy model look like? Sensitivity of metabolic energy model parameters during gait." *9th World Congress of Biomechanics 2022 Taipei* (Taipei). 2022.

[18] C. Ge, S. Weisse, B. Xu, D. Dobritzsch, J. Viljanen, J. Kihlberg, N.-N. Do, N. Schneider, H. Lanig, R. Holmdahl, and H. Burkhardt. "Key interactions in the trimolecular complex consisting of the rheumatoid arthritis-associated DRB1*04:01 molecule, the major glycosylated collagen II peptide and the T-cell receptor." *Annals of the Rheumatic Diseases* (2022).

[19] C. B. Gerber, A. Fliedner, O. Bartsch, S. Berland, M. Dewenter, M. Haug, I. Hayes, P. Marin-Reina, P. R. Mark, F. Martinez-Castellano, I. Maystadt, D. Karadurmus, K. Steindl, A. Wiesener, M. Zweier, H. Sticht, and C. Zweier. "Further characterization of Borjeson-Forssman-Lehmann syndrome in females due to de novo variants in PHF6." *Clinical Genetics* (2022).

[20] F. Hemauer, U. Bauer, L. Fromm, C. Weiß, A. Leng, P. Bachmann, F. Düll, J. Steinhauer, V. Schwaab, R. Grzonka, A. Hirsch, A. Görling, H.-P. Steinrück, and C. Papp. "Surface Chemistry of the Molecular Solar Thermal Energy Storage System 2,3-Dicyano-Norbornadiene/Quadricyclane on Ni(111)." *ChemPhysChem* (2022).

[21] C. J. K. Hewa Kankanamge, F. Lenahan, T. Klein, and A. P. Fröba. "Viscosity and Interfacial Tension of Binary Mixtures Consisting of an n-Alkane, Branched Alkane, Primary Alcohol, or Branched Alcohol and a Dissolved Gas Using Equilibrium Molecular Dynamics Simulations." *International Journal of Thermophysics* 43 (2022).

[22] L. Hümbs, A. Martin, and L. Schewe. "Exploiting complete linear descriptions for decentralized power market problems with integralities." *Mathematical Methods of Operations Research* (2022).

[23] E. Kataev, L. Fromm, Q. Tariq, D. Wechsler, F. J. Williams, N. Tsud, S. Franchi, H.-P. Steinrück, A. Görling, and O. Lytken. "Anchoring of phthalic acid on MgO(100)." *Surface Science* 720 (2022).

[24] M. Kerscher, T. Klein, P. Preuster, P. Wasserscheid, T. M. Koller, M. H. Rausch, and A. P. Fröba. "Influence of dissolved hydrogen on the viscosity and interfacial tension of the liquid organic hydrogen carrier system based on diphenylmethane by surface light scattering and molecular dynamics simulations." *International Journal of Hydrogen Energy* 47 (2022), pages 39163–39178.

[25] T. Klein, F. Lenahan, Z. Zhai, M. Kerscher, J. H. Jander, T. M. Koller, M. H. Rausch, and A. P. Fröba. "Viscosity and Interfacial Tension of Binary Mixtures Consisting of Linear, Branched, Cyclic, or Oxygenated Hydrocarbons with Dissolved Gases Using Surface Light Scattering and Equilibrium Molecular Dynamics Simulations." *International Journal of Thermophysics* 43 (2022).

[26] E. Kropac, T. Mölg, and S. Teichert. "The potential of coralline algae as SST proxy and for climate model evaluation: A New Zealand case study." *EGU General Assembly 2022* (Wien). 2022.

[27] F. Lenahan, M. Piszko, T. Klein, and A. P. Fröba. "Diffusivities in Binary Mixtures of n-Decane, n-Hexadecane, n-Octacosane, 2-Methylpentane, 2,2-Dimethylbutane, Cyclohexane, Benzene, Ethanol, 1-Decanol, Ethyl Butanoate, or n-Hexanoic Acid with Dissolved He or Kr Close to Infinite Dilution." *Journal of Chemical and Engineering Data* 67 (2022), pages 622–635.

[28] F. Lenahan, Z. Zhai, C. J. K. Hewa Kankanamge, T. Klein, and A. P. Fröba. "Viscosity and Interfacial Tension of Ternary Mixtures Consisting of Linear Alkanes, Alcohols, and/or Dissolved Gases Using Surface Light Scattering and Equilibrium Molecular Dynamics Simulations." *International Journal of Thermophysics* 43 (2022).

[29] L. Linck-Paulus, A. Horn, A. O. Matthies, S. Fischer, G. Meister, H. Sticht, M. Kappelmann-Fenzl, and A. K. Boßerhoff. "A novel splice variant of Argonaut 2 affects microRNA target genes and cell viability of melanoma cells." *JOURNAL DER DEUTSCHEN DERMATOLOGISCHEN GESELLSCHAFT*. HOBOKEN: WILEY, 2022, pages 36–36.

[30] L. Linck-Paulus, T. Meißgeier, K. Pieger, A. Horn, A. O. Matthies, S. Fischer, G. Meister, H. Sticht, M. Kappelmann-Fenzl, and A. K. Boßerhoff. "A previously unknown Argonaute 2 variant positively modulates the viability of melanoma cells." *Cellular and Molecular Life Sciences* 79 (2022).

[31] J. Link, T. Perst, M. Stöve, and B. Eskofier. "Wearable Sensors for Activity Recognition in Ultimate Frisbee Using Convolutional Neural Networks and Transfer Learning." *Sensors* 22 (2022), page 2560.

[32] J. Link, L. Schwinn, F. Pulsmeyer, T. Kautz, and B. Eskofier. "xLength: Predicting Expected Ski Jump Length Shortly after Take-Off Using Deep Learning." *Sensors* 22 (2022).

[33] A. Meier, E. Bänsch, and F. Frank. "Schur preconditioning of the Stokes equations in channel-dominated domains." *Computer Methods in Applied Mechanics and Engineering* 398 (2022).

[34] T. Mölg and C. Pickler. "A mid-troposphere perspective on the East African climate paradox." *Environmental Research Letters* 17 (2022).

[35] M. Nitschke, M. Mayer, E. Dorschky, and A. Koelewijn. "How many sensors are enough? Trajectory optimization using sparse inertial sensor sets." *9th World Congress of Biomechanics 2022 Taipei* (Taipei). 2022.

[36] M. Piszko, F. Lenahan, L. Friedl, T. Klein, and A. P. Fröba. "Mutual and Thermal Diffusivities in Mixtures of Cyclohexane, n-Hexadecane, n-Octacosane, or n-Hexanoic Acid with Carbon Dioxide Obtained by Dynamic Light Scattering and Molecular Dynamics Simulations." *Journal of Chemical and Engineering Data* 67 (2022), pages 3059–3076.

[37] K. Platzer, H. Sticht, C. Bupp, M. Ganapathi, E. M. Pereira, G. Le Guyader, F. Bilan, L. B. Henderson, J. R. Lemke, H. Taschenberger, N. Brose, R. Abou Jamra, and S. M. Wojcik. "De Novo Missense Variants in SLC32A1 Cause a Developmental and Epileptic Encephalopathy Due to Impaired GABAergic Neurotransmission." *Annals of Neurology* (2022).

[38] M. J. Poller, S. Bönisch, B. Bertleff, J. C. Raabe, A. Görling, and J. Albert. "Elucidating activating and deactivating effects of carboxylic acids on polyoxometalate-catalysed three-phase liquid–liquid-gas reactions." *Chemical Engineering Science* 264 (2022).

[39] J. Reichstein, S. Schötz, M. Macht, S. Maisel, N. Stockinger, C. Cuadrado Collados, K. Schubert, D. Blaumeiser, S. Wintzheimer, A. Görling, M. Thommes, D. Zahn, J. Libuda, T. Bauer, and K. Mandel. "Supraparticles for Bare-Eye H2 Indication and Monitoring: Design, Working Principle, and Molecular Mobility." *Advanced Functional Materials* 32 (2022).

[40] P.-G. Reinhard, P. M. Dinh, D. Dundas, E. Suraud, and M. Vincendon. "On the stability of hole states in molecules and clusters." *European Physical Journal - Special Topics* (2022).

[41] P.-G. Reinhard and W. Nazarewicz. "Information content of the differences in the charge radii of mirror nuclei." *Physical Review C - Nuclear Physics* 105 (2022).

[42] P.-G. Reinhard, X. Roca-Maza, and W. Nazarewicz. "Combined Theoretical Analysis of the Parity-Violating Asymmetry for Ca 48 and Pb 208." *Physical Review Letters* 129 (2022).

[43] C. Rettinger, S. Eibl, U. Rüde, and B. Vowinckel. "Rheology of mobile sediment beds in laminar shear flow: effects of creep and polydispersity." *Journal of Fluid Mechanics* 932 (2022), A1.

[44] C. Rettinger and U. Rüde. "An efficient four-way coupled lattice Boltzmann – discrete element method for fully resolved simulations of particle-laden flows." *Journal of Computational Physics* 453 (2022).

[45] P. Schmidt, M. Kerscher, T. Klein, J. H. Jander, F. Bioucas, T. Rüde, S. Li, M. Stadelmaier, S. Hanyon, R. Fathalla, A. Bösmann, P. Preuster, P. Wasserscheid, T. M. Koller, M. H. Rausch, and A. P. Fröba. "Effect of the degree of hydrogenation on the viscosity, surface tension, and density of the liquid organic hydrogen carrier system based on diphenylmethane." *International Journal of Hydrogen Energy* 47 (2022), pages 6111–6130.

[46] R. Simpetru, M. Oßwald, D. Braun, D. Souza de Oliveira, A. Cakici, and A. Del Vecchio. "Accurate Continuous Prediction of 14 Degrees of Freedom of the Hand from Myoelectrical Signals through Convolutive Deep Learning." *44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,*

EMBC 2022 (Scottish Event Campus, Glasgow, UK). Institute of Electrical and Electronics Engineers Inc., 2022, pages 702–706.

[47] E. Socher, L. Heger, F. Paulsen, F. Zunke, and P. Arnold. "Molecular dynamics simulations of the delta and omicron SARS-CoV-2 spike – ACE2 complexes reveal distinct changes between both variants." *Computational and Structural Biotechnology Journal* 20 (2022), pages 1168–1176.

[48] D. Thönnes and U. Rüde. "Performance Modeling of Generated Stencil Kernels within the HyTeG Framework." *Platform for Advanced Scientific Computing (PASC) Conference* (Basel, Switzerland). 2022.

[49] M. Trollmann and R. Böckmann. "mRNA lipid nanoparticle phase transition." *Biophysical Journal* (2022).

[50] E. Vintschger, D. Kraemer, P. Joset, A. Horn, A. Rauch, H. Sticht, and R. Bachmann-Gagescu. "Challenges for the implementation of next generation sequencing-based expanded carrier screening: Lessons learned from the ciliopathies." *European Journal of Human Genetics* (2022).

[51] S. Wagner, R. Sheikhi, F. Kayatz, M. Münsch, M. Hauptmann, and A. Delgado. "Fluid-structure-interaction simulations of forming-air impact thermoforming." *Polymer Engineering and Science* (2022).

[52] H. Wittkämper, S. Maisel, M. Moritz, M. Grabau, A. Görling, H.-P. Steinrück, and C. Papp. "Temperature-dependent XPS studies on Ga-In alloys through the melting-point." *Surface Science* 717 (2022).

[53] F. Xiang, S. Maisel, S. Beniwal, V. Akhmetov, C. Ruppenstein, M. Devarajulu, A. Dörr, O. Papaianina, A. Görling, K. Amsharov, and S. Maier. "Planar π-extended cycloparaphenylenes featuring an all-armchair edge topology." *Nature Chemistry* 14 (2022), pages 871–876.

[54] T. Yang, J. Yang, X. Deng, E. Franz, L. Fromm, N. Taccardi, Z. Liu, A. Görling, P. Wasserscheid, O. Brummel, and J. Libuda. "Modifying the Electrocatalytic Selectivity of Oxidation Reactions with Ionic Liquids." *Angewandte Chemie International Edition* (2022).

[55] Z. Yang, M. Markl, and C. Körner. "Predictive simulation of bulk metallic glass crystallization during laser powder bed fusion." *Additive Manufacturing* 59 (2022).

[56] M. Zürl, P. Stoll, I. Brehm, R. Raab, D. Zanca, S. Kabri, J. Happold, H. Nille, K. Prechtel, S. Wünsch, M. Krause, S. Seegerer, L. von Fersen, and B. Eskofier. "Automated Video-Based Analysis Framework for Behavior Monitoring of Individual Animals in Zoos Using Deep Learning—A Study on Polar Bears." *Animals* 12 (2022), page 692.

8.2 Compute Projects

In this section we list all *compute projects* that used NHR@FAU resources in 2022, sorted by fields of science. To give a short overview of the diverse scientific disciplines, the projects are listed according to DFG subject areas—information that is added to the application by the PI. Abstracts are truncated as indicated by square brackets; full-length abstracts can be found at https://hpc.fau.de/about-us/nhr-compute-time-projects/.

8.2.1 Humanities and Social Sciences

Linguistics

Pose22: Pose Estimation on Russian International News Media (Large scale)

[. . .] studying the mechanisms used for disinformation, in particular viewpoint manipulation, by Russian state-sponsored media.[. . .] create a dataset fully annotated with pose information of the English-language programs found on the now-banned YouTube presence of RT (Russia Today) [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

ProtCTRL: A conditional transformer for à la carte protein sequence generation

[. . .] protein design process relied on searching the global minima of multidimensional energy functions, [. . .] required significant computational times for each run. Recent advances in NLP have produced protein language models capable of generating fit protein sequences within seconds.[. . .]

University: University of Bayreuth
Target system: GPGPU cluster Alex

8.2.2 Life Sciences

Basic Research in Biology and Medicine

Dynasome3: Exploring Protein Dynamics Space to Improve Protein Function Prediction

The function of proteins is determined by their amino acid sequence and tertiary structure, but nevertheless the particular function of most proteins is unknown.[. . .] we explore to what extent protein function can be predicted by protein dynamics, and explore the space of protein dynamics in general.[. . .]

University: Georg-August-Universität Göttingen
Target system: parallel computer Fritz

ImmunoDomains: Interplay of immune receptors and lipid environment in signaling (Large scale)

[. . .] immune receptors expressed on the surface of a variety of cells is typically characterized by the sensing of an external signal, followed by signal modulation and transmission into the cell.[. . .] affected by the composition, structure, and characteristics of the plasma membrane [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

GPCRSIM: Metadynamics simulations of ligand binding/unbinding and receptor activation/deactivation for G-protein coupled receptors (Large scale)

GPCRSIM uses classical (force-field) molecular dynamics simulations to determine binding sites, binding free energies and activation/deactivation free-energy profiles for predominantly class A G protein coupled receptors.[. . .] observe rare events such as binding or activation [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

CoupledFoldBind: Conformational presentation switching processes studied by Molecular Simulations

[. . .] follow conformational changes in proteins at atomic resolution and at high time resolution.[. . .] Especially in case of conformational switching processes such as binding induced folding an understanding of the process requires the analysis of intermediate states and driving forces for conformational changes.[. . .]

University: Technical University of Munich
Target system: GPGPU cluster Alex

SimMediSoft: Biomolecular simulations for the efficient design of lipid nanoparticles

[. . .] tool to deliver RNA to target cells thereby providing promising perspectives to combat life-threatening diseases such as Amyloidosis or COVID-19.[. . .] to resolve the structure of these clinically relevant particles and [provide] molecular insights how the RNA cargo is distributed inside the [lipid nanoparticles] LNPs.[. . .]

University: University of Augsburg
Target systems: parallel computer Fritz & GPGPU cluster Alex

GPCRSCOMPEVO: Computational models of structure, dynamics and evolution of GPCRs (Large scale)

[. . .] R*-Gs/i/o arrestin complexes resolved so far do not provide a clear explanation for G protein coupling specificity.[. . .] existence of transient complexes between the R* and GTP-bound G protein [. . .] several novel intermediates on the way to the formation of GαsGTP and may contribute to coupling specificity.

University: Leipzig University
Target systems: parallel computer Fritz & GPGPU cluster Alex

Antivirals: Structure-based design and optimization of ligands for novel antiviral strategies

[. . .] complexes between antibodies and the viral fusion proteins from HIV-1 and CoV-2 are analyzed to identify energetic hot-spots of the interaction.[. . .] design of antibody-derived peptides that bind to viral fusion proteins thereby blocking viral infection.[. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

DNARepairTDG - DNA Repair by Thymine DNA Glycosylase

Thymine DNA glycosylase [. . .] involved in DNA repair [. . .] removes mispaired or modified DNA bases [. . .] ensures genetic integrity.[. . .] investigate the possible role of imino-tautomeric forms [. . .] [and the] effect of different protonation states [. . .] [and the effect of] an important histidine residue in the binding pocket.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

Resolving the Structure of mRNA-Vaccine Lipid Nanoparticles (early access)

Lipid nanoparticles (LNPs) are very successfully employed as novel transport vehicles for mRNA vaccines.[. . .] lack of a molecular picture and molecular insight into LNPs.[. . .] aim to provide unique insight at the atomistic scale into the structure and mechanisms of these carriers.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

A deep unsupervised Model for Protein Design (early access)

[. . .] we have pre-trained a generative language model on the entire protein sequence space. We find that our language model, ProtGPT2, effectively speaks the protein language and can generate de-novo sequences with natural properties in a matter of seconds.

University: University of Bayreuth
Target system: GPGPU cluster Alex

Dynamics of B2AR-Gs(GTP) (early access)

University: Leipzig University
Target system: parallel computer Fritz

Plant Sciences

CEC: Convergent evolution of carnivorous plants

[. . .] By comparison of the genetic and transcriptomic landscape between carnivorous and non-carnivorous plants, we may be able to identify the common genetic elements required for plant carnivory. This will not only grant us insights into plant carnivory, but into plant evolution as a whole.

University: University of Würzburg
Target system: parallel computer Fritz

Medicine

MASCARA: Molecular Assessment of Signatures ChAracterizing the Remission of Arthritis

[. . .] we will investigate the generation of synthetic chest X-rays using latent diffusion models. This way, we will be able to generate large amount of data necessary to train chest X-ray classification networks while preventing leakage of patient identity.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: parallel computer Fritz

PatRo-MRI-2: Pathology-robust image reconstruction in Magnetic Resonance Imaging

[. . .] beneficial for speedy imaging protocols that prioritize patient comfort. However, this reduction in data acquisition could cause generic image reconstruction techniques to obscure disease markers, replacing pathological features with typical healthy image features derived from the training data.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

GastroDigitalShirt: Development and test of deep neural network models for the automatic detection of body sounds to monitor digestion in control group and patients with intestinal disorders

[. . .] develop an unobtrusive wearable technology for long-term digestion monitoring, termed GastroDigitalShirt. We investigate low-amplitude bowel sounds (BS) as indicators of digestive disorders, including chronic inflammatory bowel diseases (IBD).[. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

FPRMetaD: Investigating binding pathways for a diverse set of ligands with biased and unbiased simulation of the Formyl Peptide Receptor
Our project includes unbiased and biased molecular dynamics simulations of the FPR receptor class and its vast array of ligands that include modified peptides as well as non-modified peptides and small molecules like Lipoxin A4 or the circular peptide Ciclosporin A.[. . .]

University: Westfälische Wilhelms Universität Münster
Target systems: parallel computer Fritz & GPGPU cluster Alex

InTimeVRSimulPatMod: In-time Virtual Reality Simulation Patient Models: Machine Learning and immersive-interactive Modeling of Virtual Patient Bodies
[. . .] provide high-quality body models given medical imaging data by the segmentation of relevant structures.[. . .] current machine learning methods and atlas-based methods are to be compared for their segmentation proposals [. . .]

University: Aalen UAS
Target systems: parallel computer Fritz & GPGPU cluster Alex

Neurosciences

HPC-MarkovModelling: Single-channel Markov modelling of voltage-gated ion channels with simulations and implementation of the 2D-Fit algorithm on High Performance Computing Cluster
[. . .] modelling single-channel patch-clamp data with Markov models. The 2D-Dwell-Time fit with simulations of time series captures gating kinetics with a high background of noise and can extract rate constants beyond the recording bandwidth.[. . .] exceptionally valuable for relating ion-channel kinetics [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

8.2.3 Natural Sciences

Molecular Chemistry

SpectroscopicProperties: Spectroscopic properties of molecules with unusual electronic structures
[. . .] predicted how to harness the peculiar properties of carbene decorated diradicals in solar cells and demonstrated their use as singlet fission molecules. Thus, our calculations helped to discover a new class of molecules of use for solar energy conversion, quantum computing, or organic light emitting diodes (OLEDs).

University: Saarland University
Target system: parallel computer Fritz

MoTrNanoMat: Molecular transport in nanoporous materials
[. . .] interconnected channels can be used for "flow-through" applications such as purification of drinking water or nanoseparation of proteins or organic solvents. Here, we will evaluate the impact of (i) nanomaterial kind, (ii) pore size, (iii) pore shape, and (iv) solvent polarity on the material's permeability [. . .]

University: University of Stuttgart
Target systems: parallel computer Fritz & GPGPU cluster Alex

Chemical Solid State and Surface Research

SurfCatal_AIMD_MLFF: Computational modeling of new surface catalysis systems by means of ab initio methods as well as novel machine-learning force-field approaches
[...] Periodic DFT simulations are able to shed a light on the exact processes taking place at the catalyst. Recently, a new machine-learning force-field (ML-FF) was developed which is able to efficiently learn on the fly from DFT data, leading to a high-level FF for metal surfaces in contact with other phases [...]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: parallel computer Fritz

Physical and Theoretical Chemistry

AKES: Chemical Modelling of Processes in Pharmaceutical Chemistry
Covalent inhibitors currently experience a renaissance in medicinal chemistry due to their various advantages, including prolonged residence times, lower sensitivity against pharmacokinetic aspects, and high efficacy. Our work addresses the reaction mechanisms of cysteine protease rhodesain with covalent inhibitors.[...]

University: Julius-Maximilians-Universität Würzburg
Target system: GPGPU cluster Alex

Ion-catch: Molecular Modelling based design of ligand shells to functionalize magnetic nanoparticles for the removal of heavy metal pollutants from water
[...] designing tailor-made functionalization of magnetite nanoparticles to bind heavy metal ions and related organometallic compounds by means of molecular modelling and simulation.[...] developing a model-based search strategy for identifying suitable constituents and structures as guides to syntheses.[...]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

CatalAcetylen: Acetylene selective hydrogenation to ethylene catalyzed by bi- and trimetallic alloys: AI search for new catalysts
Ethylene (C_2H_4) is a building block for the production of polymers.[...] C_2H_4 is usually produced by cracking light alkanes and contains acetylene (C_2H_2) which deactivates the catalyst. Therefore, it is important to reduce the amount of it before starting the polymerization reaction.[...]

University: Freie Universität Berlin, Technische Universität Berlin
Target system: parallel computer Fritz

Condensed Matter Physics

NQSdynamics: Neural quantum states for dynamical processes in quantum matter
The efficient numerical simulation of nonequilibrium real-time evolution in isolated quantum matter constitutes a key challenge for current computational methods.[...] we employ the neural quantum state, which [...] has shown great potential in studying various quantum phenomenon.[...]

University: University of Augsburg
Target system: parallel computer Fritz

DMFT2TBLG: DMFT study of a heavy-fermion model for twisted bilayer graphene
Twisted bilayer graphene (TBLG) has recently captivated the interest of the condensed matter community, for its capability of hosting a wide variety of peculiar phenomena such as superconductive and correlated insulating phases, as well as topological features.[...]

University: Universität Würzburg
Target system: parallel computer Fritz

FRG: Functional Renormalization Group calculations for material analysis
Current ab-initio theory for solid state materials excels in the prediction of electronic band structures. The secondary part of any full description—the interaction between electrons—is beyond the scope of those methods.[...] (FRG) has the expressed goal of deriving effective low-energy interaction models.[...]

University: RWTH Aachen
Target system: parallel computer Fritz

ALFQMCsim: Emergent and critical phenomena in correlated electron systems: Quantum Monte Carlo simulations (Large scale)
[...] a general implementation of the so called auxiliary field quantum Monte Carlo algorithm.[...] triggered at solving systems of correlated electrons that couple to bosonic modes such as lattice vibrations.[...] allows us to compute properties of systems in thermodynamic equilibrium at polynomial cost.[...]

University: Julius-Maximilians-Universität Würzburg
Target system: parallel computer Fritz

Functional renormalization group calculations (early access)
University: RWTH Aachen
Target system: parallel computer Fritz

Massively parallel simulation via Markov Chain Monte Carlo techniques (early access)
University: Julius-Maximilians-Universität Würzburg
Target system: parallel computer Fritz

Particles, Nuclei, and Fields

CLSfiniteV: Finite volume study of 2 + 1f QCD from lattice simulations
High precision calculations from lattice QCD are needed nowadays [...] investigate the systematic effects introduced by the finite volume of the system, which is important in particular for quantities like low energy constants, pseudoscalar masses and decay constants, or the axial charge of the nucleon.

University: Universität Regensburg
Target system: parallel computer Fritz

Statistical Physics, Soft Matter, Biological Physics, Nonlinear Dynamics

EnSimTurb: Towards ensemble simulations of fully developed turbulence
[...] developing new theoretical and computational approaches to better understand and model fully developed turbulence.[...] capture the large-scale dynamics of turbulent flows. Both projects use our code TurTLE, a pseudo-spectral solver of the Navier-Stokes equations which also features particle tracking capabilities.

University: University of Bayreuth
Target systems: parallel computer Fritz & GPGPU cluster Alex

Astrophysics and Astronomy

HESSML: Advanced Machine Learning Analysis for the H.E.S.S. Telescopes
[. . .] With this project, we aim to develop a variety of new analysis methods for H.E.S.S. data, including improved identification of muonic events, exploration of geometric deep learning techniques, fast simulation approaches and new data augmentation strategies.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

Atmospheric Science, Oceanography and Climate Research

ATMOS: Numerical atmospheric modeling for the attribution of climate change and for model improvement
[. . .] aims to explore a novel climatic indicator, namely crustose coralline algae that grow in shallow ocean waters, for the purpose of improved global climate model evaluation.[. . .] improvement potential with regard to sea surface temperatures in the Southern Ocean and the effect of these ocean conditions [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: parallel computer Fritz

8.2.4 Engineering Sciences

Mechanics and Constructive Mechanical Engineering

FRASCAL-FE: Computational continuum mechanics simulations at LTM for FRASCAL
[. . .] Specifically, project P8 (“Fracture in Polymer Composites: Meso to Macro”) aims to study the influence of different mesoscopic parameters, including microstructure morphology, on the macroscopic fracture properties of nano-particle reinforced polymers.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: parallel computer Fritz

FRASCAL-MD: Particle-based computing at LTM for FRASCAL
[. . .] In particular, sub-project P6 (“Fracture in Thermoplastic Polymers: Discrete-to-Continuum Coupling”) provides a link between the level of atoms and the continuum with specific interest in the multiscale modelling and simulation of polymer fracture.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

Fluid Mechanics, Technical Thermodynamics and Thermal Energy Engineering

AOTTP-DFG18-1: Characterization of molecular diffusion in electrolyte systems
Electrolytes conduct electric current by the movement of ions while blocking the free movement of electrons. For applications of electrolytes as working fluids, the transport of ions is important and can be subdivided into diffusion, convection, and, in the presence of an electric field, migration.[. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target systems: parallel computer Fritz & GPGPU cluster Alex

Flow around a Wind Turbine Blade at Reynolds Number 1 Million (early access)
[...] carry out high-fidelity numerical simulations of the flow around a wind turbine blade at a realistic Reynolds number to get a deeper insight into this phenomenon and especially the transition process under different levels of the turbulence intensities of the approaching flow.

University: Helmut Schmidt University Hamburg
Target system: parallel computer Fritz

Evolution of drops in homogeneous isotropic turbulence (early access)
University: University of Bremen
Target system: GPGPU cluster Alex

Materials Science

Materials4.0 - AITDB: Ab initio thermodynamic database development
[...] focus will be on phase stabilities of various phases, including dynamically unstable ones. Having acquired such a database, the phase stabilities can be put into practice by re-parametrizing binary phase diagrams and studying the implications on multicomponent phase diagrams.

University: Universität Stuttgart
Target system: parallel computer Fritz

Computer Science

digiOnko: digiOnko – Mit digitaler Medizin gegen Brustkrebs, AP5 - Histopathologie
[...] improving the screening, early detection, diagnosis, treatment and aftercare for women with breast cancer.[...] Deep learning methods will be used to analyze digital Whole Slide Images (WSI), in particular to automatically calculate scores and improve treatment decisions.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

DeepPano: Erzeugung von Panoramabildern aus 3D-Laser-Punktewolken und Kamerabildern
[...] data obtained by the 3D-indoor-scanning systems [...] who can deliver high-quality 3D point clouds and registered input images.[...] focus on novel neural rendering techniques [...], on the generation of panoramic images, and [...] to achieve free-viewpoint video, based on this data, in real-time.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

HEISSRISSE: Massively Parallel Simulation of the Melt Pool Area during Laser Beam Welding using the Lattice Boltzmann Method
Using abstraction layers and code generation concepts, software developed with waLBerla is sustainable [...] One of the significant aspects [...] lies in the validation of newly implemented models and algorithms, and in the interoperability with models from the partners in the research unit.[...]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

IRRW: Scaling Inverse Rendering to the Real World

How do we best represent objects and their variations for inverse rendering? Can a combination of classical and novel techniques increase photorealism whilst retaining a low dimensional and interpretable representation? And given such object models: How do we efficiently infer the scene graph [. . .]

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

ICETHICKNESS: Machine learning-based retrieval of ice thickness/internal structures from radargrams

[. . .] we aim at using and modifying machine learning techniques from medical imaging as well as natural language processing and apply those to glaciological radargrams to extract information on ice thickness and internal structures of ice bodies.

University: Friedrich-Alexander-Universität Erlangen-Nürnberg
Target system: GPGPU cluster Alex

Editor:

Prof. Dr. G. Wellein

Editorial staff:

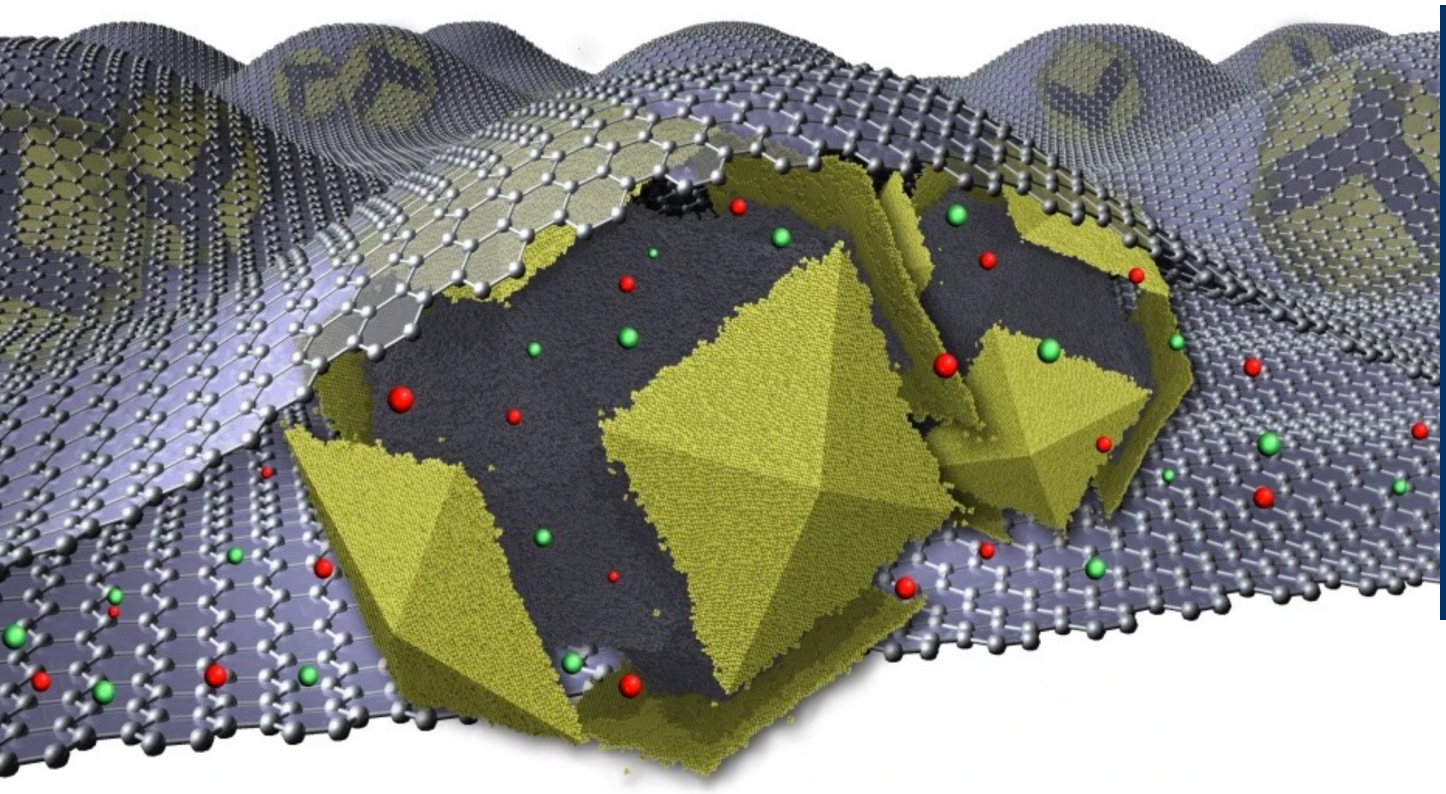
Katja Augustin
Dr. Georg Hager
Dr. Anna Kahler
Dr. Harald Lanig

Image credits:

If not listed here, the images are the property of NHR@FAU/RRZE.
Alberto Leonardi, FAU/Nature Communications volume 11, Article number: 3041 (2020) (front and back cover)
MEGWARE Computer Vertrieb und Service GmbH (p. 5, 26)
Friedrich-Alexander-Universität Erlangen-Nürnberg, University of Greifswald, Technical University of Munich (p. 13 f.l.t.r.)
Rainer A. Böckmann, Marius F.W. Trollmann, Biophysical Journal Volume 121, Issue 20, 18 October 2022 (p. 73)

ISSN 2751-8647

Erlangen, August 2023



nhr.fau.de