

NHR - Graduate School: Research Fields 2025

List of Contents

NHR4CES: RWTH Aachen University and Darmstadt University	3
1#NHR4CES: Data Management and Data Science and Machine Learning.....	3
2#NHR4CES: Parallelism and Performance.....	4
3#NHR4CES: Visualization	4
4#NHR4CES: Energy Conversion	4
5#NHR4CES: Fluids	5
6#NHR4CES: Materials Design	5
7#NHR4CES: Digital Patient	5
NHR@FAU –Friedrich-Alexander-Universität Erlangen-Nürnberg	7
1#FAU: Improving Scalable Solvers by AI	7
2#FAU: Performance Engineering from Single Core to Highly Parallel Architectures	7
3#FAU: Molecular Simulations of Bio-Molecular Complexes	7
5#FAU: High-throughput Simulations for a Mechanistic Understanding of Crystallization Processes.....	8
GWDG – NHR@GÖTTINGEN –Göttingen.....	9
1#Göttingen: HPC-Cloud-Convergence	9
2#Göttingen: Large Language Models in HPC.....	9
3#Göttingen: Data-Center Twin	9
4#Göttingen: Scalable Big Data Analytics.....	10
NHR@KIT – Karlsruhe Institute of Technology	11
1#KIT: High-performance computations for Multiphysics models.....	11
2#KIT: Combined numerical and parallel methods.....	11
3#KIT: Biologically inspired artificial intelligence.....	12
4#KIT: Machine learning in Earth system models	12
NHR@PC2 – Paderborn Center for Parallel Computing – University of Paderborn	13
1#PC2: Sparse Linear Algebra with FPGAs	13
3#PC2: Methods for Atomistic Simulations on Heterogeneous Computer Architectures	14
4#PC2: Methods for compressing smooth volumetric data on grids	14
ZIH – NHR@TUD – Dresden University.....	16
1#TUD: Software Tools for Parallel Performance Analysis and Optimization	16
2#TUD: HPC Workflows for Big Life Science Data.....	16
3#TUD: High-Performance Storage Systems	17
4#TUD: Machine Learning and Training Data Preparation	17
5#TUD: Energy Efficient Computing	17

6#TUD: Machine Learning for Research Data Management.....	18
NHR@SW – Goethe University Frankfurt, Johannes Gutenberg University Mainz, University of Kaiserslautern-Landau, Saarland University	
1#SW: Algorithmic Differentiation in HPC Environments	19
2#SW: Compilation Techniques and Domain-Specific Languages for Heterogeneous Systems	19
3#SW: Expressive and Efficient Compilation of Flexible Compute Pipelines and Graphs for GPUs	20
4#SW: GPU-Bio: Accelerated algorithms and machine learning for modern bioinformatics on GPU-accelerated HPC systems.....	20
5#SW: Modular and Automated Workload Analysis for HPC-Systems	21
NHR@ZIB – Berlin	
1#ZIB: Hybrid Quantum-Classical Algorithms for Combinatorial Optimization.....	22
2#ZIB: Large-scale Explainable AI	22
3#ZIB: AI-Accelerated Molecular Dynamics.....	22
4#ZIB: Energy-Efficient Computing on Data-Parallel Platforms	23
5#ZIB: Semantic I/O with Object Stores in HPC	23
6#ZIB: Massive Simulation of Agent-Based Models.....	24

NHR4CES: RWTH Aachen University and Darmstadt University

In NHR4CES, the Rheinisch-Westfälische Technische Hochschule Aachen (RWTH Aachen University) and Darmstadt University (TuDa) join forces to combine their strengths in HPC applications, algorithms, and methods, and the efficient use of HPC hardware. Our goal is to create an ecosystem combining best practices of HPC and research data management to address questions that are of central importance for technical developments in the economy and society. We support the development of scalable algorithms and software for the investigation, development, design, construction, evaluation, and production of engineering applications with a particular focus on engineering and materials science as well as engineering-oriented physics, chemistry, and medicine.

1#NHR4CES: Data Management and Data Science and Machine Learning

Data-intensive applications are getting more and more important in the context of HPC infrastructure. First, classical HPC users who in the past used simulation-based and thus computation-intensive methods rely more and more on data-driven methods such as ML and AI. Second, as another trend, we see that also the user base of HPC infrastructure is growing since many disciplines that did not use HPC in the past, plan to use the HPC infrastructure for tackling data-intensive problems using scalable resources. As such, the goal of the Cross-Sectional Groups (CSGs) *Data Management* and *Data Science and Machine Learning* is to develop novel methods to enable classical HPC users but also new user groups to implement data-intensive applications on top of the scalable HPC infrastructure. For this, the Cross-Sectional Groups aims to develop a range of different methods that can broadly be classified into two main directions:

- **Cross-Sectional Group Data Management:** The existing data services such as long-term data storage are today completely decoupled from the HPC infrastructure. As such the first goal is to better integrate the existing data services into the available HPC infrastructure. Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

- Prof. Dr Matthias Müller, RWTH Aachen University, Chair of Computer Science (High-Performance Computing)
- Prof. Dr Carsten Binnig, Darmstadt University, Professor of Computer Science (Data and AI Systems)
- Prof. Dr. Christian Bischof, Darmstadt University, Professor of Computer Science (Scientific Computing).
- **Cross-Sectional Group Data Science and Machine Learning:** we aim to develop novel methods to enable users without data science skills to build data-intensive applications by automating data engineering tasks such as data transformation, logging, and cleaning as well as automating the ML model building e.g. via neural architecture search and making ML and process mining approaches differentiable. By automating data engineering and ML model building, we enable classical HPC users as well as new user groups who are no experts in data science and ML to build data-driven applications much faster and with much lower (manual) overhead than today. Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

- Prof. Bastian Leibe, RWTH Aachen University, Chair of Computer Vision
- Prof. Wil van der Aalst, RWTH Aachen University, Chair of Process and Data Science

- Prof. Kristian Kersting, Darmstadt University, Computer Science Department, Artificial Intelligence and Machine Learning (AIML).

2#NHR4CES: Parallelism and Performance

Performance and Parallelism-related research is being executed at RWTH and TUDa in several cases in joint projects. The Cross Section Group "Parallelism and Performance" invents methods and algorithms to assess and improve the correctness, performance, and parallelism of HPC applications and develop tools (*MUST*, *Archer*, *Extra-P*, *Disco-POP*) to put theory into practice. Based on work with real-world applications, they propose new approaches to program modern HPC architectures, bring these into standardization bodies such as *MPI* and *OpenMP*, and extend the corresponding runtime systems. In doing so, they regularly explore and characterize new system and microprocessor architectures. Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

Prof. Felix Wolf, Darmstadt University, Department of Computer Science, Department Chair (2021-2023)

Dr. Christian Terboven, RWTH Aachen University, IT Center

Dr. Sandra Wienke, RWTH Aachen University, IT Center.

3#NHR4CES: Visualization

Visualization allows for efficient data exploration and analysis in virtually all domains. The rapid development of hardware and simulation software, however, introduces a variety of challenges for scientific visualization. This includes finding new approaches for the visualization of complex and uncertain data types, dealing with large data, or analyzing running simulations with in-situ methods. The goal is to provide researchers with efficient tools to gain better insights into their data and techniques may range from providing immersive environments to defining new data features. Further information is provided on our [website](#).

Supervisor and location (depending on the topic of the doctoral thesis):

Prof. Torsten Kuhlen, RWTH Aachen University, IT Center.

4#NHR4CES: Energy Conversion

Within the Simulation and Data Lab "Energy Conversion", we develop HPC-ready reactive CFD (rCFD) software and methods through a co-design process. Such rCFD software is key for exploring physics and efficient virtual designs of energy conversion devices as part of the energy transition, a global challenge with major economic and social impact. Highly optimized numerical approaches are developed, tested, validated, and packaged in different forms for dissemination. Resulting in specialized HPC modules that are highly parallelized and optimized for the targeted Tier II architectures, also providing efficient usage on Tier I machines. Software quality is ensured together with the CSGs by Continuous Integration (CI).

Two simulation frameworks will be used, the DNS code Ciao developed at RWTH Aachen University and the open-source code OpenFOAM, customized for LES and 3D URANS applications at Darmstadt University (TUDa). Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

Prof. Heinz Pitsch, RWTH Aachen University, Institute for Technical Combustion

Prof. Christian Hasse, Darmstadt University, Institute of Simulation of Reactive Thermo-Fluid Systems, Head of the Institute

Dr. Holger Marschall, Darmstadt University, Head of Research Group Computational Multiphase Flow.

5#NHR4CES: Fluids

The detailed analysis of flow problems is essential in energy conversion, environmental sciences, climate research, etc. Besides experiments, numerical investigations are to be performed to solve the scientific challenges that are related to multi-scale multi-physics problems such as particle-laden flows, gas-liquid flows, flows of complex and particulate matter, aerodynamics, aeroacoustics, and combustion. To capture the details of such problems high-performance computing is a must. Furthermore, model reduction for multi-query applications, and data- and physics-based model integration are emerging topics. An entire spectrum of scalable discretization methods – finite volumes, finite elements, and Discontinuous Galerkin / high-order Schemes – is considered within the Simulation and Data Lab “Fluids”. Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

Prof. Wolfgang Schröder, RWTH Aachen University, Chair of Fluid Mechanics and Institute of Aerodynamics

Prof. Jeanette Hussong, Darmstadt University, Head of Institute for Fluid Mechanics, and Aerodynamics

Appl. Prof. Dr. Suad Jakirlic, Darmstadt University, Fluid Mechanics, and Aerodynamics

Prof. Marek Behr, RWTH Aachen University, Chair for Computational Analysis of Technical Systems

Prof. Martin Oberlack, Darmstadt University, Institute of Fluid Dynamics.

6#NHR4CES: Materials Design

Molecular dynamics simulations, using high-performance computing architectures, allow the explore materials properties and processes with high spatial and temporal resolution.

While running multi-million particle simulations nowadays can be done on a routine basis, the analysis of all trajectories is still a challenging matter. With the advent of novel data analysis methods based on machine learning algorithms, new and improved analysis methods are on the horizon. In our working group Material Design, which is part of the Simulation and Data Labs at NHR4CES, we address the development and implementation of analysis methods as well as the integration into community-based software tools. Possible projects include the simulation and structure recognition of grain boundary phases, glasses, and complex microstructures. Further information is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

Prof. Karsten Albe, Darmstadt University, Materials Modelling

Prof. Korte-Kerzel, RWTH Aachen University, Physical Metallurgy and Metal Physics Faculty of Georesources and Material Engineering

Prof. Jochen Schneider, RWTH Aachen University, Chair for Materials Chemistry

Prof. Bai-Xiang Xu, Darmstadt University, Division Mechanics of Functional Materials.

7#NHR4CES: Digital Patient

Efficient computational methods for the development of digital patients representing disease progression, therapeutic options, and efficiency of personalized therapies are essential for future digitalization in medicine, both in research and clinical health care. Digital patients require an optimal adaption of model structures to the medical needs and availability of mechanistic knowledge and data. Moreover, digital patients exhibit multi-scale structures ranging from the molecular up to the multi-organ level. They can be realized by hybrid integration of physics-based models, machine learning, and molecular simulations. Molecular simulations as well as identification of digital patients from data require efficient adaption of HPC methods, e.g., Markov-Chain Monte Carlo simulations, to the specific requirements of

hybrid ML-mechanistic models. These methods need to be adapted to the specific heterogeneous structures of digital patients and data.

Further information on the working group "Digital Patient" of the Simulation and Data Lab at NHR4CES is provided on our [website](#).

Supervisors and locations (depending on the topic of the doctoral thesis):

[Prof. Giulia Rosetti](#), Forschungszentrum Jülich, Computational Biomedicine

[Prof. Volkmar Schulz](#), University Hospital RWTH Aachen, Chair of Experimental Physics III B

[Prof. Andreas Schuppert](#), University Hospital RWTH Aachen, Chair of Computational Biomedicine.

NHR@FAU –Friedrich-Alexander-Universität Erlangen-Nürnberg

The Erlangen National High-Performance Computing Centre ([NHR@FAU](#)) is part of the [Friedrich-Alexander-Universität Erlangen-Nürnberg](#) with a main focus on: performance engineering, scalable solvers and atomistic structure simulations in chemistry, life science, materials science, and physics.

1#FAU: Improving Scalable Solvers by AI

In many HPC applications, solving sparse linear systems is the fundamental performance bottleneck. Only a few algorithmic principles, such as, e.g., multigrid, are available to construct scalable solvers. Further acceleration is possible with techniques that avoid storing the system matrix. Such matrix-free techniques are one focus of research at [NHR@FAU](#). These methods require that the solver is co-designed with the discretization method. However, since the optimal choice of components and parameters for discretization and solvers are highly problem-dependent, it is difficult to find them manually for more complex matrices. Therefore, surrogate models e.g. based on physically inspired neural nets (PINNs) or genetic programming are used to improve the numerical methods. However, these data-driven methods can have high computational costs in the training phase.

To address this issue, another central component in building highly efficient and easily usable implementations in this co-design process is modern code-generation (meta-programming) techniques which form another research focus of [NHR@FAU](#).

Supervisor & Location:

[Prof. Harald Köstler](#), FAU University, Head of Research Erlangen National High-Performance Computing Center & Chair of Computer Science 10 (System Simulation).

2#FAU: Performance Engineering from Single Core to Highly Parallel Architectures

Performance Engineering (PE) is a structured, performance-model-based process for the structured optimization and parallelization of basic operations, algorithms, and application codes for modern computing architectures. Performance models and performance tools are the core components of the PE process. The group uses and further develops the execution cache memory (ECM) model, which is a generalization of the well-known Roofline model and allows predictions of single-core performance and scaling within a multi-core chip. The interaction of node-level bottlenecks with highly parallel code execution is another recent area of research. Furthermore, simple-to-use performance analysis and modeling tools such as LIKWID or OSACA are developed at [NHR@FAU](#). Hardware-efficient building blocks for sparse linear algebra and stencil solvers are strong topical foci of the PE activities of the center.

Supervisor & Location:

[Prof. Gerhard Wellein](#), FAU University, Department of Computer Science, Professorship for High-Performance Computing & Head of [NHR@FAU](#).

3#FAU: Molecular Simulations of Bio-Molecular Complexes

Bio-molecular complexes such as protein-protein, protein-DNA, or (other) protein-ligand complexes are flexible entities whose conformational dynamics are important for their function. By molecular dynamics simulations, using high-performance computing architectures, these dynamics can be explored in atomic detail. However, depending on the process to be investigated, advanced methods such as enhanced sampling techniques or approaches with hybrid quantum mechanical and classical (QM/MM) potentials are needed, often also in combination.

We are using a catalog of molecular simulation techniques and data analysis methods to study bio-molecular complexes involved in important processes such as DNA repair, enzymatic reaction pathways, ligand-induced signaling, or immune processes at the plasma membrane.

Supervisors and locations (depending on the topic of the doctoral thesis):

Prof. Petra Imhof, FAU University, Computer Chemistry Center, Technical Director

Prof. Heinrich Sticht, FAU University, Professorship of Bioinformatic

Prof. Rainer Böckmann, FAU University, Professorship for Computational Biology.

4#FAU: Development of new electronic structure methods combining high accuracy with computational efficiency

The workhorse for electronic structure calculations in chemistry, physics, and materials science are methods based on density-functional theory (DFT). Despite their enormous success in practical application, presently available DFT methods suffer from several shortcomings. Their accuracy does not reliably reach what is called chemical accuracy, i.e., an accuracy of 1 kcal/mol in reaction energies, and non-covalent interaction cannot be treated. A promising new type of DFT method relies on the adiabatic-connection fluctuation-dissipation (ACFD) theorem. ACFD methods exhibit unprecedented accuracy and applicability. These methods need to be implemented in highly efficient codes making full use of the opportunities high-performance computing offers.

Supervisor & Location:

Prof. Andreas Görling, FAU University, Department of Chemistry and Pharmacy, Chair of Theoretical Chemistry.

5#FAU: High-throughput Simulations for a Mechanistic Understanding of Crystallization Processes

The formation of material structure can be rationalized, to a first approximation, from geometric considerations. But while crystal structure is generally well understood, dynamic aspects of crystallization are much less accessible. The prevailing strategy in computational materials science is to simulate with ever higher accuracy. Here, we propose to investigate crystallization processes systematically across parameter spaces and aim at a mechanistic understanding that correlates the geometry of matter (structure) and interaction (chemistry) to details of ordering pathways at full atomic resolution. This goal is accomplished by establishing a crystallographic structure analysis and simulation infrastructure that rival's experimental synthesis and structure solutions from scattering data. Analysis of trajectory data addresses long-standing open questions in the context of multi-step nucleation and directional growth. Applications are phase diagrams of intermetallic, industrial crystal growth, and fractional crystallization in geology.

Supervisor & Location:

Prof. Michael Engel, FAU University, Department of Chemical and Biological Engineering, Professorship for Modeling of Self-Organization Processes.

The Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen (GWDG) is a joint computing center and research facility of the Max Planck Society and the Göttingen University. It provides infrastructures for data management and analysis, collaboration, and computing for various national projects. This national high-performance data center with the Emmy supercomputer is a tier 2 center and offers advice and support to a diverse portfolio of users and is itself involved in research in application-related projects. The NHR centre NHR@Göttingen is part of the National High-Performance Computing (NHR) and a member of the NHR network.

1#Göttingen: HPC-Cloud-Convergence

In recent years, classic HPC users have seen an ever-increasing interest in the public cloud that is used as part of traditional HPC workflows. There are many reasons for this, e.g. the usage of special hardware components in the cloud, or that data products created in a data centre are shared with a scientific community - that continues to process it in the cloud using High-Performance Data Analytics. At the same time, a technology initially developed for the cloud is increasingly used in HPC centres and vice versa. Hence, cloud and HPC systems are converging. As part of this topic, synergistic effects in compute, storage, and management are researched.

Supervisor & Location:

Prof. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

2#Göttingen: Large Language Models in HPC

In recent years, LLMs such as ChatGPT-4 have revolutionized the field of natural language processing (NLP) by demonstrating exceptional performance in various tasks, including text generation, sentiment analysis, and machine translation. They bear chances to be utilized for various tasks in HPC as well, for instance, for logfile analysis, for analysing system faults, debugging and performance analysis of parallel applications, and for providing a rich variety of user support. Recognizing the untapped potential of LLMs in High-Performance Computing (HPC) environments, this PhD project aims to explore their application in enhancing HPC operations and user experience through innovative problem-solving techniques. Additionally, platforms to deploy inference models efficiently in an HPC centre are explored.

Supervisor & Location:

Prof. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

3#Göttingen: Data-Center Twin

Data centres contain various infrastructures to, e.g., steer airflow and cooling of housed systems. Understanding and optimizing the infrastructure but also the whole life cycle of IT components is important to improve efficiency and make computing greener. The modeling and simulation of the technical processes and relevant characteristics allow investigation of the impact of various tuning options and enable what-if analysis to optimize planning in each data centre but also in future data centres.

Supervisor & Location:

Prof. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

4#Göttingen: Scalable Big Data Analytics

Machine learning methods are omnipresent in our daily lives and are also increasingly used in scientific domains such as life science and the digital humanities. However, in the data centre, HPC workflows could still benefit more from opportunities for machine learning and deep learning methods. As part of this topic, we research:

- How scalable data analytics workflows can be established for relevant scientific domains.
- How existing workflows can be augmented using deep learning, machine learning, and High-Performance Data Analytics, in general.

Supervisor & Location:

Prof. Julian Kunkel, Göttingen University, Professor in High-Performance Computing & Deputy Head High-Performance Computing at GWDG.

Karlsruhe Institute of Technology (KIT) is a university of the state of Baden-Württemberg as well as a research center in the Helmholtz Association and creates synergies from serving both purposes. The NHR centre NHR@KIT addresses four user communities (1) earth system science, (2), materials science, (3) engineering for energy and mobility research, and (4) particle physics. Core HPC method competencies are data-intensive computing, numerical algorithms, and sustainable software development. In HPC operations, NHR@KIT focuses on secure authentication and authorization.

1#KIT: High-performance computations for Multiphysics models

The numerical solution of Multiphysics models enables predictions of coupled processes in diverse fields of engineering applications such as material and geophysical science. For many applications, high-performance computing (HPC) is crucial to allow large-scale simulations in representative volume elements, which are required to predict physical mechanisms and processing routes. Such numerical simulations are based on the coupled solution of complex partial differential equations comprising e.g. mass and heat transport, flow, or thermo-chemo-mechanical interactions. The consideration of physically meaningful domain sizes is limited by the computational cost of the simulation. Therefore, large-scale simulations require HPC employing parallelized software that uses e.g. domain decomposition techniques based on MPI or OpenMP. Thereby, the computational efficiency of the numerical algorithms and the Multiphysics coupling of these parallel solvers is of key importance. This can be achieved by using vectorization as well as methods of load-balancing between different parallel processes. The work aims to improve the computational efficiency of a Multiphysics framework in the HPC context using such techniques, and thus, enhance the applicability of simulation methods for large and complex problems.

Supervisors & Location (depending on the topic of the doctoral thesis):

Prof. Britta Nestler, Karlsruhe Institute of Technology, Professor for microstructure simulations
Dr. Martin Reder, Karlsruhe Institute of Technology, Institute for Digital Materials Research.

2#KIT: Combined numerical and parallel methods

The efficient use of high-performance computers is a central challenge when investigating coupled evolution processes on large computational domains such as those arising in multiscale problems featuring different length scales that must be resolved with different accuracy. The solution of the resulting large systems of equations requires a combination of efficient numerical solvers with parallel computation methods. Most solution approaches based purely on local information show a market deterioration of their performance with the size of the computational domains, making methods such as adaptive and multigrid methods based upon the introduction of coarser auxiliary problems highly attractive. While numerically highly efficient, this can raise additional difficulties to the efficiency of the parallelization. The research work aims to implement a suitable parallelization of these numerical methods via MPI and OpenMP and to create a systematic hierarchical approach with a uniform simulation framework for multiscale processes that includes programming interfaces for data processing and numerical approaches in combination with a consistently parallel solution with high scaling properties.

Supervisors & Location (depending on the topic of the doctoral thesis):

Prof. Britta Nestler, Karlsruhe Institute of Technology, Professor for Microstructure Simulations
Dr Martin Reder, Karlsruhe Institute of Technology, Institute for Digital Materials Research.

3#KIT: Biologically inspired artificial intelligence

Biologically inspired computing aims at solving complex computational problems by mimicking biological processes or structures. Especially in complex optimization problems found in many artificial intelligence and machine learning approaches to applications from numerous scientific and engineering domains, bio-inspired algorithms are recognized as state-of-the-art approaches. Examples of these algorithms include genetic and particle swarm optimization, spiking neural networks, or feedback-free learning. NHR@KIT is looking for projects targeting, but not limited to, biologically inspired neural architecture search, hyperparameter optimization, learning algorithms for (deep) neural networks, and neuromorphic computing. The project may target either methodological advances from the computer sciences and/or novel and innovative solutions to complex computational problems from the natural or engineering sciences. A strong emphasis is put on the scalability of the researched biologically inspired algorithms to large-scale computational applications, exposing inherent parallelism.

Supervisors & Location (depending on the topic of the doctoral thesis):

Dr. Arnd Koeppel, Karlsruhe Institute of Technology, Microstructure Modelling and Simulation, Group Leader.

Dr. Michael Selzer, Karlsruhe Institute of Technology, Microstructure Modelling and Simulation.

4#KIT: Machine learning in Earth system models

Classical modelling has long been the standard method for long- and short-term predictions in climate science, and models such as the KIT-developed ICON-ART rank among the power users of large HPC systems.

With the rapid developments in hardware and the rise of the Exascale era, scientists can process ever-growing amounts of data and refine their simulations to improve results. The downsides of this evolution are long computing times and significant power consumption. However, in recent years scientists pursued a new strategy by adapting large neural networks and using the improvements in machine learning and artificial intelligence. This ansatz yielded encouraging results while reducing both compute time and power consumption. The goal of this research work is to combine both, classical modelling, and cutting-edge AI techniques.

Supervisors & Location (depending on the topic of the doctoral thesis):

Prof. Martin Frank, Karlsruhe Institute of Technology, Director Scientific Computing Center (SCC) & Professor for Computational Science and Mathematical Methods, Department of Mathematics

Dr. Ole Kirner, Karlsruhe Institute of Technology, Department of Scientific Computing & Mathematics, Scientist in the Simulation and Data Lab Earth System Science

Dr. Jasmin Hörter, Karlsruhe Institute of Technology, Head of the Department of Scientific Computing & Mathematics

Dr. Jörg Meyer, Karlsruhe Institute of Technology, Head of the Department of Data Analytics, Access, and Applications.

The Paderborn Center for Parallel Computing (PC2) is a central research institute and HPC center at Paderborn University with the mission to conduct interdisciplinary research at the intersection of computational science, computer science, and innovative computer architecture. Within the NHR Alliance, PC2 focuses on applications in computational physics and chemistry; numerical methods and libraries for heterogeneous computing systems; and energy-efficient computer architectures such as FPGAs and GPUs. The supervision of the proposed doctoral topics with a strong connection to natural sciences is carried out jointly with corresponding chairs from the Faculty of Natural Sciences.

1#PC2: Sparse Linear Algebra with FPGAs

For addressing large-scale computational science and engineering problems, the use of sparse matrix/vector data representations is essential for keeping the memory and computational demand of the simulations tractable. The efficient execution of linear algebra operations on sparse data structures is however challenging because the predominant current computer architectures are optimized for dense rather than sparse linear algebra and suffer from a widening gap between computational and memory throughput. This is problematic for memory-bound sparse operations. Therefore, CPUs or GPUs typically achieve only a small fraction of their peak performance for sparse linear algebra.

FPGAs can close this gap by customizing the internal memory architecture, tailoring the sparse matrix storage format, and chaining algebraic operations to avoid repeated transfers from/to external DRAM memory. First studies have demonstrated the potential of FPGAs for sparse operations for selected operations but so far there is no cohesive design method or library of composable matrix operations for FPGAs. Examples of Ph.D. projects in this area are:

- the development of composable building blocks for sparse linear algebra
- the development of novel techniques for integrating the building blocks into more complex algorithms (e.g. solvers)
- the study of arithmetic customization (low/mixed-precision computing)
- integration into widely- used open-source application codes.

Supervisor & Location:

Prof. Christian Plessl, Paderborn University, Institute of Computer Science.

2#PC2: Scalable and Performance-Portable Scientific FPGA Applications

The use of FPGAs as accelerators has so far mostly focused on strong scaling applications for single nodes. The next step for the adoption of FPGAs in HPC is to build scalable, parallel, multi-node applications for scientifically relevant application domains. For highly scalable applications, suitable network architectures and communication libraries are required that are usable directly from FPGA applications without routing all communication through the host. Also, runtime systems and middleware for parallel FPGA systems are still a wide-open research field. Examples of Ph.D. projects in this area are:

- the development of communication infrastructures and libraries that are suitable for direct use from FPGA applications.
- the performance portability of FPGA applications among different FPGA types or to other architectures such as CPUs and GPUs.

While OpenCL and SYCL-based tools have brought great progress in terms of functional portability, performance optimization techniques still differ widely. In the context of scalable applications with FPGAs directly communicating via common standards, performance portability

can also be important in heterogeneous execution modes where different application kernels are flexibly mapped to the most suitable device. Further examples of PhD projects in this area are:

- the improvement of performance portability through code generation or domain-specific languages
- exploration of configurable and composable architecture-optimized libraries or template-based architecture adaptation
- techniques for heterogeneous execution modes combining streaming and dataflow execution modes with block-based or bulk synchronous execution modes.

Supervisor & Location:

Prof. Christian Plessl, Paderborn University, Institute of Computer Science.

3#PC2: Methods for Atomistic Simulations on Heterogeneous Computer Architectures

The simulation of atomic interactions at different levels of physical accuracy and numerical demand ranging from force-field molecular dynamics over static and dynamic density functional theory up to quantum chemistry is a cornerstone for computational science in physics, chemistry, material sciences, and life sciences. Because of the extremely widespread use and computational demand of these methods, improvements in efficiency and performance in the underlying methods, algorithms, and implementations have a large practical impact and can reduce the carbon footprint of the computation. We are particularly interested in the development of improved simulation methods for electronic structure computation that target heterogeneous computer architectures comprising GPUs and FPGAs.

By exploiting massive parallelism, mixed and low-precision computing, and hardware customization and operating each architecture at its sweet spot, substantial gains in efficiency and performance can be achieved. Recent examples of this co-design approach are our highly scalable and efficient submatrix method and our FPGA-accelerated computation of electronic repulsion integrals for quantum chemistry, which have been integrated into CP2K. Examples of Ph.D. projects in this area are the development of novel, highly scalable, and efficient methods for atomistic simulations with DFT-variants on modern hardware accelerators, with a combination of GPUs and FPGAs; investigation of approximate computing methods that combine low and high-precision arithmetic to reduce time/energy to solution while limiting the effect on the quality of results.

Supervisors & Location (depending on the topic of the doctoral thesis):

Prof. Christian Plessl, Paderborn University, Institute of Computer Science

Prof. Martin Brehm, Paderborn University, Institute of Chemistry;

(dependent on the specific focus of the Ph.D. thesis).

4#PC2: Methods for compressing smooth volumetric data on grids

For a few decades, modern experimental and computational science has been more and more confronted with the "big data" issue", i.e. the necessity to handle - and possibly store - huge amounts of intermediate raw data to obtain the desired results. One strategy to mitigate such issues is to compress the raw data suitably to reduce bandwidth and storage requirements.

The efficiency of a compression algorithm depends a lot on the precise nature of the data. If the general structure of the input data is well known, tailor-suited compression algorithms can be designed which exploit that structure. We are developing such tailor-suited approaches to compress time series of volumetric data on grids that are smooth both in space and time. For the application of storing total electron density data from ab initio calculations, a lossless

compression ratio of 40:1 can be reached: (<https://pubs.acs.org/doi/10.1021/acs.jcim.8b00501>). However, for other categories of volumetric data such as wave functions or molecular orbitals, slightly different fine-tuning will be required to achieve a similar compression efficiency. Examples of Ph.D. projects in this area:

- Designing compression algorithms for other types of volumetric data time series
- Optimizing the compression and decompression code for more efficiency
- Parallelization of the compression and decompression operations.

Supervisor & Location:

Prof. Martin Brehm, Paderborn University, Institute of Chemistry.

The Dresden University is one of the largest technical universities in Germany. It has a long history of running HPC installations and has been specialized for data intensive HPC for a long time. Its NHR centre's focus topics are:

- Methods for Big Data and data analysis as well as management
- Machine Learning
- (Tiered storage architectures and I/O optimization
- Performance and energy efficiency analysis and optimization.

The application area focus topics are:

- life sciences and
- Earth system science.

The NHR@TUD centre at TU Dresden cooperates closely with the Center for Scalable Data Analytics and Artificial Intelligence - ScaDS.AI Dresden/Leipzig. For further questions on the research fields offered by NHR@TUD, please contact Dr. Matthias Lieber via e-mail, mentioning the research field title.

1#TUD: Software Tools for Parallel Performance Analysis and Optimization

Parallel performance, efficiency, and scalability are primary concerns for the HPC community. Dedicated software tools for performance analysis and optimization are indispensable to cope with several challenges for HPC programmers and cluster operators. Several performance aspects are waiting for new software tool approaches for measuring, analyzing, and steering parallel performance, among others: new heterogeneous hardware architectures; new and evolving parallel programming models; more and more complex application codes; Big Data and Machine learning frameworks in HPC environments; AI analysis of monitoring data; as well as high scalability and low overhead of the tools themselves. This topic offers the opportunity to join an established performance tools research group working on well-known software tools renowned in the HPC community.

Supervisor & Location:

Prof. Wolfgang E. Nagel, Dresden University, Chair of Computer Architecture at the Institute of Computer Engineering & Director of the Center for Information Services and High-Performance Computing (ZIH).

2#TUD: HPC Workflows for Big Life Science Data

Single-cell omics and live imaging of developing embryos and organoid cultures generate big data in the life sciences. We need to analyze these big life science data to crack the code of biological self-organization and morphogenesis, as started by Alan Turing with his work on "The Chemical Basis of Morphogenesis". This approach integrates HPC workflows, mathematical modeling, and multi-scale simulations with interdisciplinary thinking to develop and apply novel algorithms. For example, the hybrid simulation of a developing organism combines a mechanistic model of a subsystem of interest, like cell-to-cell signaling and gene regulatory networks, with a processed stream of measurement data for the other subsystems, like the segmentation, tracking, and interpolation of 3D time-lapse microscopy data of the cell positions, shapes, and contacts in a growing and deforming embryo. The results of such hybrid simulations can then be compared to measurement data for the entire system to perform robust parameter estimation and model selection for the subsystem of interest. Such novel HPC workflows will be widely applicable across the life sciences and beyond.

Supervisor & Location:

Prof. Andreas Deutsch, Dresden University, Head of Center for Information Services and High-Performance-Computing.

3#TUD: High-Performance Storage Systems

Work on new concepts and implementations for highly efficient data workflows and storage systems that connect HPC and AI. This will combine methods from the different scopes which offer mutual advantages. For example, combine HPC technologies like Infiniband, RDMA, NVMeoF, GPU direct, and others with different storage options. It involves managing workflow approaches compatible with the respective HPC and AI environment. It may also focus on data and metadata management integration in data-driven workflows. Finally, it may concern performance and scalability improvements for the most demanding processing stages. Prototyping high-end storage systems for data processing and realizing demanding workflows might be part of the assignment.

Supervisor & Location:

Prof. Wolfgang E. Nagel, Dresden University, Chair of Computer Architecture at the Institute of Computer Engineering & Director of the Center for Information Services and High-Performance Computing (ZIH).

4#TUD: Machine Learning and Training Data Preparation

In many use cases, rapidly increasing data generation rates require constant updates of the trained machine and deep learning models at a faster rate, making it vital to pre-process the incoming data and update the model in close time. This requires adapting systems and models to handle the data in the form of data streams using online/active learning. Often the data is generated in real-time at external experiments, requiring immediate analysis, especially for mission-critical applications, e.g., in medical scenarios or safety domain. Modern HPC infrastructures can play a vital role in achieving this requirement. Here, the techniques for data reduction, data transfer, and data pre-processing with efficient use of resources are an ongoing area of research. Furthermore, current research involves investigation into the scaling of these methods using distributed computing frameworks w.r.t. different loads, storage systems, computing systems, IO, energy consumption, etc., while aiming for low latency and high throughput computation. In addition to exploring different data quality and data integration techniques, the current work could also involve exploring secure and privacy-aware processing techniques for highly sensitive data. Within the national AI competence center ScaDS.AI Dresden/Leipzig, ZIH and its partners are involved in research on efficient distributed learning approaches and performance optimization of machine learning applications on scalable HPC infrastructures and alternative computing architectures for the future.

Supervisor & Location:

Prof. Wolfgang E. Nagel, Dresden University, Chair of Computer Architecture at the Institute of Computer Engineering & Director of the Center for Information Services and High-Performance Computing (ZIH).

5#TUD: Energy Efficient Computing

Limited power and energy budgets have become major challenges in many areas of (scientific) computing. ZIH has developed a strong expertise in power and energy measurements, addressing important aspects such as accuracy, scalability as well as temporal and spatial granularity. This work is complemented by the development of methods such as modeling and optimizing the energy efficiency of parallel applications. Benefiting from several decades of work in performance benchmarking processor and memory architectures, we are developing a growing body of Open-Source software tools to address optimization challenges, e.g., BenchIT

for low-level benchmarking, lo2s for extremely lightweight node-level energy and performance monitoring, and MetricQ for highly scalable, distributed metric data processing. We also consider system cooling to be an integral part of energy efficiency optimization, and waste heat reuse to mitigate unavoidable detrimental effects on the environment. As such, our approach to energy efficiency research is truly holistic, spanning from the single processor core to the data center.

Supervisor & Location:

Prof. Wolfgang E. Nagel, Dresden University, Chair of Computer Architecture at the Institute of Computer Engineering & Director of the Center for Information Services and High-Performance Computing (ZIH).

6#TUD: Machine Learning for Research Data Management

Data management (DM) is often perceived as an additional effort detaining researchers from their desired job, doing research. Nevertheless, good scientific practice and traceability of research outcomes demand DM throughout the research process. Especially working with large amounts of data requires enhanced techniques to keep DM tasks in check and streamline to the research process. Automatic detection of discrete datasets within a data collection and its significant metadata may reduce individual DM efforts significantly. However, depending on the precise research question, similar data could be handled and analyzed very differently. Machine learning algorithms could help the researcher to structure large amounts of data and perform important DM tasks in time, after a learning phase, e.g. on discrete dataset definition, important metadata, and project schedule. Ideally, researchers can hand over large parts of time-consuming DM tasks after a learning phase and become DM supervisors rather than performing DM tasks themselves.

Supervisor & Location:

Prof. Wolfgang E. Nagel, Dresden University, Chair of Computer Architecture at the Institute of Computer Engineering & Director of the Center for Information Services and High-Performance Computing (ZIH).

NHR@SW – Goethe University Frankfurt, Johannes Gutenberg University Mainz, University of Kaiserslautern-Landau, Saarland University

The NHR South-West Centre (NHR@SW) is a collaboration of [Goethe University Frankfurt](#), [Johannes Gutenberg University Mainz \(JGU Mainz\)](#), [University of Kaiserslautern-Landau \(RPTU KL\)](#) and [Saarland University](#). Research activities at NHR@SW focus on performance engineering, artificial intelligence, life sciences, and physics.

1#SW: Algorithmic Differentiation in HPC Environments

The landscape of HPC architectures, special accelerators, and communications models has evolved over the last decades and will continue to do so in the future. This implies that new libraries and programming models are used to fully facilitate the peak performance of the new systems. For all these architectures and libraries, it is an ongoing challenge to provide automatically created derivative code via the theory of Algorithmic Differentiation (AD). Examples are the PGAS programming model, CUDA, ROC, FPGAs, TPUs, or GPUs. Each library needs to be analysed for its specific challenges and extended such that a differentiated version can be used in the context of AD.

Another aspect is the formulation of problems in domain-specific languages (DSLs) which makes the formulation more concise. Furthermore, the generated code can be optimized for different hardware architectures and run-time environments. Formulating problems also requires the use of derivatives or the sensitivities for the full problem formulation are required in a design optimization context. Algorithmic Differentiation (AD) techniques can be used to enhance DSL frameworks with the capability to additionally generate the derivative code of the problem formulation. The generated derivative code is then again optimized for different hardware architectures and run-time environments.

Supervisors & Locations (depending on the topic of the doctoral thesis):

[Prof. Nicolas Gauger](#), University of Kaiserslautern-Landau, Chairholder for Scientific Computing & Director of Computing Center (RHRZ).

[Prof. Sebastian Hack](#), University of Saarland, Professor of Computer Science at Saarland Informatics Campus.

2#SW: Compilation Techniques and Domain-Specific Languages for Heterogeneous Systems

Future HPC systems will become more and more heterogeneous including accelerators of various kinds. GPUs are already widely adopted but, in the future, more specialized accelerators such as TPUs or FPGAs could make their way into HPC systems because they promise better energy efficiency. However, hardware heterogeneity poses a significant challenge to programming these systems because it deviates from standard programming models. Our group does basic research on domain-specific languages (DSLs) and compiler techniques to enable the programming of future heterogeneous systems. We work on supporting next-generation vector instruction sets within the LLVM compiler toolchain and on the DSL framework AnyDSL that allows for simple and effective DSL embedding using partial evaluation and compilation to a wide range of accelerators including NVIDIA and AMD GPUs, FPGAs, CPUs with various vector extensions.

Supervisor & Location:

[Prof. Sebastian Hack](#), University of Saarland, Professor of Computer Science at Saarland Informatics Campus.

3#SW: Expressive and Efficient Compilation of Flexible Compute Pipelines and Graphs for GPUs

GPUs have become a key element of almost all high-performance computing systems. However, the programming models for GPU are still very limited, and support for efficiently mapping and optimal scheduling a pipeline or graph of different compute tasks onto more and more complex GPU architectures and their increasing set of special functional units is essentially non-existent.

Current GPUs support two separate operation modes with two different and quite limited operating environments: Firstly, in Graphics Mode they support the traditional (graphics) pipeline with shaders for some predefined programmable stages as well as so-called compute shaders. This supports optimal mapping and scheduling but is limited to the traditional and linear graphics pipeline. Secondly, in Compute Mode they offer a more flexible programming model but access to the special function units is very limited (if available at all), and automatic mapping and scheduling is not supported and must be done by hand. Most GPU programming environments target the latter mode, but even that still greatly limits the expressivity because widely used programming features are not supported (recursion, indirect function calls, stacks, exceptions, and many more).

With *Shady* we have developed the prototype of a compiler that eliminates all these limitations. It adds a novel model for control flow that clearly defines the divergence/convergence behaviour of parallel HW threads. This allows for completely avoiding the many bugs in current GPU compilers concerning convergence behaviour and can be implemented even within current GPU programming environments. We have started to work with the GPU vendors through Khronos to discuss how extensions could make this approach even more efficient with upcoming extensions. Significant research is still needed on the one side to map this model to the different GPU architectures while optimally exposing their specific functional unit, and on the other hand to best expose this to higher-level programming environments.

Based on the above approach, we will continue developing a high-level and flexible way to define compute pipelines and graphs as well as ways to allow for automatically mapping them as optimal as possible to different GPU architectures while taking their different functional units into account. This approach is based on clearly defined communication between the tasks even then being mapped to the available communication mechanisms on GPUs (registers, shared memory, global memory, CPU memory, and eventually even network connections). We are already working with GPU vendors to best support this approach, plan to provide our tool openly to NHR and the HPC community at large, and publish the results in suitable academic journals and conferences.

Supervisors & Locations (depending on the topic of the doctoral thesis):

Prof. Philipp Slusallek, Saarland University, Professor for Computer Graphics & Scientific Director at the German Research Center for Artificial Intelligence (DFKI)

Prof. Sebastian Hack, Saarland University, Professor of Computer Science at Saarland Informatics Campus.

4#SW: GPU-Bio: Accelerated algorithms and machine learning for modern bioinformatics on GPU-accelerated HPC systems

Recent years have seen a tremendous increase in the volume of data generated in the life sciences, The abundance of generated data allows us to address a plethora of questions that could not be answered previously with numerous applications in precision and personalized medicine as well as drug discovery. However, the analysis of large-scale data sets poses difficult computational challenges and is thus a highly active field of research. For example, as protein

sequence database growth outpaces Moore's law, and increased use of advanced machine learning (ML) applications such as AlphaFold2 require increasingly more multiple sequence alignments (MSAs) to be generated. Thus, corresponding MSA-generating tools need to be innovative quickly enough to cope with increased complexity. The explosive growth of life science datasets and the rapid evolution of sequencing technologies thus motivates the development of algorithms and tools that can scale towards large-scale input datasets. This in turn requires the efficient exploitation of modern high-performance computing (HPC) systems that are moving towards increased levels of parallelism using many-core architectures such as GPUs with thousands of cores per chip. The goal of this project is the design and implementation of highly efficient massively parallel algorithms optimized for CUDA-enabled GPUs with associated software tools for relevant analysis tasks in the life sciences. This project will comprehensively tackle efficiency, sensitivity, and precision bottlenecks in a variety of applications such as protein structure prediction and metagenomics. Our approach will be based on developing libraries of key algorithms in a massively parallel way. They will then be used in several applications to gain significant speedups compared to current approaches while scaling to large data sets. Thus, this project will boost the downstream analysis of sequencing and structural datasets for many important tasks in bioinformatics.

Supervisor & Location:

Prof. Bertil Schmidt, Johannes Gutenberg University Mainz, Professor for High-Performance-Computing at the Institute of Computer Science.

5#SW: Modular and Automated Workload Analysis for HPC-Systems

Given the complexity of modern HPC systems, achieving theoretical peak performance depends on a myriad of parameters and system configurations. To optimize the system performance and efficiently use the underlying resources, various methods can be applied, including simulation, benchmarking, and monitoring. However, these methods and the tools are oftentimes not compatible with each other and only consider a selection of performance factors such as network, I/O, resource allocation, or parallel execution. Yet, each of these approaches generates knowledge that can be applied to similar problems or system configurations. To avoid such knowledge being collected only for one-time purposes, and to also support other users, this knowledge must be easily accessible and available to the community. The MAWA-HPC (Modular and Automated Workload Analysis for HPC Systems) approach aims to develop a universal workflow and tool infrastructure (starting with the execution environment up to the evaluation dashboard) that can be applied to different use cases and workloads from various science domains. Through its modular architecture, the workflow will enable the support of different community tools, which will increase the compatibility of the individual tools and cover new use cases. By utilizing the support of established monitoring and profiling tools, node-level performance engineering tools, network benchmarks, and microbenchmarks for different parallel programming models, we want to integrate different performance models into MAWA-HPC such as a multi-dimensional Roofline model and a holistic performance scoring system. The holistic performance score will consider multiple performance dimensions, including network, compute, parallel I/O, and time. Our previous work has introduced the I/O Roofline model, which is tool and platform-agnostic. For example, including time as an additional dimension, the Roofline model can provide insights into an application's performance over time, enabling the identification and understanding of performance anomalies. Possible PhD projects include the generation of user hints, bottleneck detection, pattern analysis, performance prediction, and performance optimization, possibly using ML and AI mechanisms.

Supervisor:

Prof. Sarah Neuwirth, Johannes Gutenberg University Mainz, Professor for Computer Science.

At Zuse Institute Berlin (ZIB), model-based simulation and optimization, data-driven problem solving, and advanced computing meet with high-performance research services and transfer to diverse application fields and industries. The NHR center at ZIB is part of a research institute with a focus on Scientific Computing and a broad spectrum of interdisciplinary research projects and about 300 scientists and research infrastructure specialists. The NHR center at ZIB will take the leadership in selected application domains, e.g., life sciences with a focus on model-driven simulations, advanced integration of artificial intelligence/machine learning and simulation, or heterogeneous HPC architectures jointly with other NHR centers.

1#ZIB: Hybrid Quantum-Classical Algorithms for Combinatorial Optimization

Quantum computing has the potential to revolutionize HPC in the long term. At ZIB, we aim to develop a workflow from the design of hybrid quantum-classical algorithms up to their simulation utilizing quantum simulators on HPC hardware. We consider this specifically in the context of combinatorial optimization problems, because of the industrial relevance as well as the first promising results from quantum computing for this problem class. The problem comprises three sub-aspects:

- Identification of combinatorial optimization problems that can be solved using NISQ and QAOA in combination with traditional approaches.
- development and design of hybrid quantum-classical approaches for selected examples of the problems identified in (1)
- simulation and verification of quantum-classical hybrid algorithms. Since we do not expect that general-purpose quantum computing will be available soon and with sufficient scaling, a preliminary verification by quantum computing simulators running on classical HPC hardware is essential.

Supervisor & Location:

Dr. Patrick Gelß, Zuse Institute Berlin, Department for AI in Society, Science, and Technology & Head of the IOL.Quant group.

2#ZIB: Large-scale Explainable AI

AI has the potential to revolutionize how we do science, and it will be at the forefront of the next wave of innovation. Methods capable of learning from large-scale combined with methods from explainable AI (XAI) will play an important role in future applications. This requires new methodological approaches, such as e.g., new optimization methods and more generally new learning algorithms. At the same time, it requires new approaches to handling and processing large-scale data. At ZIB we develop both the underlying methodologies for large-scale learning and XAI as well as the necessary HPC approaches to effectively execute these approaches. This includes the full design stack from high-level end-to-end learning methodologies to lowest level efficient data handling procedures.

Supervisor & Location:

Contact person: Dr. Thomas Steinke, Zuse Institute Berlin, Head of Department.

3#ZIB: AI-Accelerated Molecular Dynamics

The convergence and combination of classic HPC methods with approaches from artificial intelligence (AI) is an ongoing trend in science. Molecular dynamics (MD) benefits from AI as well: High-dimensional neural networks potentials enable the accurate modeling of many-body

interaction in materials research. They allow for predictions of undiscovered chemical species, describing the chemical space of a given structure, or precisely predict the pathways between folded and unfolded conformations of a protein. To effectively support molecular dynamics research with these promising AI-based models, they need to be combined with existing HPC methods and codes. Also, accelerator hardware is required to fulfill the computational demands of such research. ZIB is working closely with researchers and software developers from the MD domain to enable such combinations of AI and HPC. In addition, a focus of ZIB's research in the supercomputing department is on the cross-platform support of AI methods on accelerators by using modern programming languages and frameworks.

Supervisors & Location (depending on the topic of the doctoral thesis):

Dr. Marcus Weber, Zuse Institute Berlin, Research Group Leader Computational Molecular Design

Dr. Nicholas Charron, Zuse Institute Berlin, Department of Parallel and Distributed Computing.

4#ZIB: Energy-Efficient Computing on Data-Parallel Platforms

Data-parallel (GPU) and data-flow (FPGA, Next Silicon) accelerators play an important role in energy-efficient computing. The more domain-specific algorithms can be effectively mapped on these accelerators, the more energy-efficient data processing can be realized in the future. Whereas for dense and partially sparse linear algebra advantages of mixed precision and approximate computing strategies have been demonstrated on these devices, there are many questions in other areas of numerical mathematics like PDE solvers unanswered. Further, from the technical point of view, some challenges need to be tackled. The current "art of coding" is under question as with the diversity of accelerators one must select the associated programming models. At ZIB, we want to continue our interdisciplinary work in areas of mathematical numeric's and modern software engineering by exploring the potential of mixed-precision algorithms, e.g., on unstructured grids on current and next-generation data-parallel / data-flow accelerators. For that, our emphasis is on standard programming models (OpenMP, SYCL including oneAPI) being portable across vendors and processor architectures.

Supervisor & Location:

Dr. Steffen Christgau, Zuse Institute Berlin, Department of Parallel and Distributed Computing.

5#ZIB: Semantic I/O with Object Stores in HPC

Object stores like S3 are a common storage infrastructure in cloud environments. Nowadays, object stores become increasingly popular in HPC as an alternative concept to store data as they overcome scalability issues of POSIX-complaint file systems used traditionally in HPC configurations. Examples of HPC-relevant objects store are based on distributed key/value stores. This offers the possibility to integrate application-specific metadata directly into the managed objects at the low level. Moreover, with the flexibility of a key/value store the efficiency of I/O operations can be further increased as the logical structure of the data to be stored can be mapped to an optimized schema of the key/value backend. As such, these data containers are natively supported by the storage system with best performance for storing and retrieving the data. At ZIB, we perform research with one of the current implementations of a distributed object store, the Distributed Asynchronous Object Store (DAOS). Using its infrastructure, the potential benefit of taking the semantic knowledge of data into account will be investigated.

Supervisor:

Dr. Steffen Christgau, Zuse Institute Berlin, Department of Parallel and Distributed Computing.

6#ZIB: Massive Simulation of Agent-Based Models

Agent-based models (ABMs) have become very popular for describing the interaction of many similar entities/agents (human beings, households, institutions, or the like) by a set of rules describing the changing location of agents and their interaction with each other. Interacting agents can alter their own or the other agent's properties, for example, by changing an opinion or transmitting an infection. Information about agents' mobility and status dynamics is typically deduced from available data sets and expert knowledge. Due to their versatility, ABMs are very attractive for modeling spatiotemporal population dynamics in various disciplines, ranging from social process models via traffic or mobility transition models to opinion spreading and infection dynamics. ABMs with large numbers of agents (millions or larger) require high-performance computing infrastructure for the required massive simulations, especially if scenario studies with changing parameters or different external controls are investigated. AT ZIB, we are developing tools for such massive ABM simulations on HPC systems.

Supervisor & Location:

Contact person: Dr. Thomas Steinke, Zuse Institute Berlin, Head of Department.