Book of Abstracts

SPPEXA

SYMPOSIUM

25.-27.01.2016

Technische Universität München
Leibniz Rechenzentrum

Contents

Invited Talks

1.1 Exascale N-body Algorithms for Data Analysis and Simulation

George Biros
ICES, University Texas at Austin
201 E 24th St
Austin, TX 78712
E-Mail: biros@ices.utexas.edu

N-body algorithms are ubiquitous in science and engineering and form the core methods for many production codes in industrial, academic, and government labs. The find application in both computational physics and machine learning. Tree-based methods typically require irregular memory access patterns that result in reduced off- and on-node performance. Although significant progress has been made in improving off-node performance, on-node performance remains an open problem. This is especially true for production tree-based codes that have multi-stage computations involving data reshuffles and multiple computational kernels. This on-node utilization wall–a chronic problem since the early nineties–not only remains unresolved but has become much more acute with the emergence of deeper memory hierarchies and manycore and heterogeneous architectures. In this talk, I will outline the computational kernels used in N-body methods and I will describe the challenges in scaling them efficiently.

1.2 Beyond Peta-scale on Stencil and Particle-based GPU Applications

Takayuki Aoki
Global Scientific Information and Computing Center (GSIC)
2-12-1 i7-3 O-okayama, Meguro-ku
Tokyo 152-8550
E-Mail: taoki@gsic.titech.ac.jp

GPU (Graphics Processing Unit) has been widely used in high-performance computing as an accelerator with high computational performance and wide memory bandwidth. We have succeeded in several large-scale stencil and particle applications on a GPU supercomputer TSUBAME: a weather prediction model covering the whole Japan area with 500-m horizontal resolution, a turbulent airflow simulation with LES and Lattice Boltzmann Method for a central part of metropolitan Tokyo for 10 km x 10 km area with 1-m resolution and a phase-field simulation for the dendritic solidification of a binary alloy with 0.3 trillion cells. We also demonstrate granular DEM and fluid SPH simulations using billion particles which are quite different from N-body problems. Stencil framework approaches for practical applications have been also developed for high productivities and an extended roofline model using node performance and interconnection bandwidth instead of processor performance and bandwidth is used to discuss post peta-scale applications.

1.3 Exascale - On What Dimension?

Craig Stewart
Indiana University Pervasive Technology Institute
2709 E 10th St
Bloomington, IN 47405
E-Mail: stewart@iu.edu

US President Obama recently signed an executive order creating the *National Strategic Computing Initiative*. This executive order sets out a goal of creating an exaFLOPS computational system for the US. One of the announcements about this initiative quoted the President's Council of Advisors on Science and Technology which has previously stated that high-performance computing "must now assume a broader meaning, encompassing not only flops, but also the ability, for example, to efficiently manipulate vast and rapidly increasing quantities of both numerical and non-numerical data." This talk will discuss US efforts toward exascale computation, particularly considering the synergies between developments needed to enable exascale tightly coupled parallel workloads and exascale big data computations.

EXA-DUNE - Flexible PDE Solvers, Numerical Methods, and Applications

2.1 Advances Concerning Multiscale Methods and Uncertainty Quantification in Exa-Dune

Peter Bastian, Steffen Müthing
Interdisciplinary Center for Scientific Computing, Heidelberg University, Im Neuenheimer Feld 368, 69120 Heidelberg / Germany
E-Mail: {peter.bastian,steffen.muething}@iwr.uni-heidelberg.de

Jan Mohring, Oleg Iliev
Fraunhofer Institute for Industrial Mathematics ITWM, Fraunhofer-Platz 1, 67663 Kaiserslautern / Germany
E-Mail: {jan.mohring,oleg.iliev}@itwm.fraunhofer.de

Christian Engwer, René Milk, Mario Ohlberger, Jorrit Fahlke
Institute for Computational and Applied Mathematics, University of Münster, Orleans-Ring 10, 48149 Mnster / Germany
E-Mail: {christian.engwer,rene.milk,mario.ohlberger}@wwu.de, jorrit@jorrit.de

Dominik Göddeke
Institut für Angewandte Analysis und Numerische Simulation, University of Stuttgart, Allmandring 5b, 70569 Stuttgart / Germany
E-Mail: dominik.goeddeke@mathematik.uni-stuttgart.de

Olaf Ippisch
Institut für Mathematik, TU Clausthal-Zellerfeld, Erzstr. 1, 38678 Clausthal-Zellerfeld / Germany
E-Mail: olaf.ippisch@tu-clausthal.de

Markus Geveler, Dirk Ribbrock, Stefan Turek
Inst. for Applied Mathematics, TU Dortmund, Vogelpothsweg 87, 44227 Dortmund / Germany
E-Mail: {markus.geveler,dirk.ribbrock,stefan.turek}@math.tu-dortmund.de

In this contribution we present advances concerning efficient parallel multiscale methods and uncertainty quantification (UQ) that have been obtained in the frame of the DFG priority program 1648 *Software for Exascale Computing* (SPPEXA) within the funded project Exa-Dune.
Within Exa-Dune one of our interests is the effective numerical treatment of partial differential equations that arise in complex models with physical properties that vary rapidly on very small scales, for example pollutant transport in groundwater or single phase flow in porous media. These kinds of models regularly lead to very large scale, and numerically demanding problems to solve, with UQ, or a more general multi-query context, naturally requiring many of these solves. Faced with this potentially huge computational challenge it is paramount

that we employ both effective numerical schemes and efficient, highly scalable codes that can use multiple levels of parallelism in modern peta-scale, and even more so future exa-scale, compute cluster hardware. To that end we introduce a mathematical abstraction for a wide variety of multiscale methods which are uniquely suited for the posed models. Founded on this abstraction layer we demonstrate our hybrid shared-memory/MPI implementation of the Multiscale Finite Element Method (MsFEM) based in the well known Distributed and Unified Numerics Environment (DUNE) and using the DUNE Generic Discretization Toolbox, which yields promising strong scaling results on SuperMUC and CHEOPS. We will also show how the Multi-Level Monte-Carlo Method (MLMC) can be used to reduce the computational complexity plaguing classical UQ schemes and how we were able to easily use our interface based MsFEM solver with our very general MLMC code. We will close the presentation with scalability results for a model problem of UQ for single phase flow with random permeability fields.

2.2 Hardware-based Efficiency Advances in the EXA-DUNE Project

Steffen Müthing, Peter Bastian
Interdisciplinary Center for Scientific Computing, Heidelberg University, Im Neuenheimer Feld 368, 69120 Heidelberg / Germany
E-Mail: {steffen.muething,peter.bastian}@iwr.uni-heidelberg.de

Christian Engwer, Jorrit Fahlke, René Milk, Mario Ohlberger
Institute for Computational and Applied Mathematics, University of Münster, Orleans-Ring 10, 48149 Münster / Germany
E-Mail: {christian.engwer,jorrit.fahlke,rene.milk.mario.ohlberger}@uni-muenster.de

Markus Geveler, Dirk Ribbrock, Stefan Turek
Institute for Applied Mathematics, TU Dortmund, Vogelpothsweg 87, 44227 Dortmund / Germany
E-Mail: {markus.geveler,dirk.ribbrock,stefan.turek}@math.tu-dortmund.de

Dominik Göddeke
Institute for Applied Analysis and Numerical Simulation, University of Stuttgart, Allmandring 5b, 70569 Stuttgart / Germany
E-Mail: dominik.goeddeke@mathematik.uni-stuttgart.de

Oleg Iliev, Jan Mohring
Fraunhofer Institute for Industrial Mathematics ITWM, Fraunhofer-Platz 1, 67663 Kaiserslautern / Germany
E-Mail: {oleg.iliev,jan.mohring}@itwm.fraunhofer.de

Olaf Ippisch
Institut für Mathematik, TU Clausthal-Zellerfeld, Erzstr. 1, 38678 Clausthal-Zellerfeld / Germany
E-Mail: olaf.ippisch@tu-clausthal.de

We present advances concerning efficient finite element assembly and linear solvers on current and upcoming HPC architectures obtained in the frame of the EXA-DUNE project, part of the DFG priority program *SPPEXA*. In this presentation, we focus on node-level performance and accelerator integration, which complement the proven MPI-level scalability of the *DUNE* framework. In particular, we show different vectorisation approaches for low order Lagrange and high order sum factorised DG discretisations of convection-diffusion problems as well as GPU-accelerated preconditioners.

2.3 Firedrake and PyOP2, a Hierarchy of Composable Abstractions for Finite Element Simulation

David A. Ham
Department of Mathematics, Imperial College London / United Kingdom
E-Mail: david.ham@imperial.ac.uk

Lawrence Mitchell
Department of Computing and Department of Mathematics, Imperial College London / United Kingdom
E-Mail:

Gheorghe-Teodor Bercea, Miklós Homolya, Fabio Luporini, Paul H.J. Kelly
Department of Computing, Imperial College London / United Kingdom
E-Mail:

Florian Rathgeber
European Centre for Medium-range Weather Forecasts / United Kingdom
E-Mail:

Andrew T.T. McRae
Department of Mathematical Sciences, University of Bath / United Kingdom
E-Mail:

The complexity explosion caused by the application of advanced numerics on sophisticated parallel hardware to changing and coupled physical systems presents a critical barrier to simulation development. The key to overcoming this is to create simulation software systems which embody the abstraction and composability of the underlying mathematics. By doing this, a system is created in which mathematicians, computer scientists, and application specialists can each deploy their own expertise while benefiting from the expertise of the others. Critically, this approach minimises the extent to which individual experts need to acquire the expertise of the other contributors in order to benefit from their efforts.
In this talk I will present Firedrake and PyOP2, a composition of new and existing abstractions which creates a particularly complete separation of concerns. This enables the creation of high performance, sophisticated finite element models from a very high level mathematical specification and has enabled advances in computer science and numerics, at the same time as facilitating the creation of simulation systems for a variety of applications.

2.4 Space-time Domain Decomposition Method for a Two-phase Flow Model in Porous Media

Elyes Ahmed
Inria, Paris / France
E-Mail: Elyes.Ahmed@inria.fr

Caroline Japhet
LAGA, Université Paris Nord / France and Inria, Paris / France
E-Mail: Caroline.Japhet@inria.fr

Michel Kern
Inria, Paris / France
E-Mail: Michel.Kern@inria.fr

We consider a simplified model for two-phase flow through a heterogeneous porous medium. The porous medium is made up of different subdomains, each of them being a different rock type. Each rock is characterized by its relative permeability and capillary pressure curves that are functions of the phase saturations. Domain decomposition provide a natural approach for dealing with the different physics occurring in the different rock types, which induce a discontinuity of the saturation. We propose a global in time domain decomposition, which gives us the flexibility to use different time steps. More precisely, focusing on the capillary forces motion, we propose a new approach based on Optimized Schwarz Waveform Relaxation (OSWR), with nonlinear Robin conditions, in the context of nonlinear degenerate parabolic problem. We propose and analyze a hybridized finite volume scheme for the approximation of the multidomain solution that also leads to an existence result for the nonlinear problem.
To accelerate the convergence of the OSWR method, we reformulate the multidomain problem as a nonlinear interface problem whose unknowns are Robin data on each side of the interface, and for which more general non-linear solvers can be used (Newton-like solvers). The method is naturally parallelizable across the subdomains. The interface problem is global in time, thus different time grids are employed to adapt to different time scales in the different rock types. An additional advantage of a global in time method is that information is exchanged only at the end of each iteration, so that communication between subdomains is reduced.
Numerical results for three-dimensional problems for domains with different capillarity functions are presented to illustrate the performance of the method. More realistic prototypes for simulations for the underground storage of nuclear waste are shown. The choice of the Robin parameters, and its influence on the convergence of the domain decomposition method will be studied.

ExaStencils - Advanced Stencil-code Engineering

3.1 Performance Prediction of Multigrid-Solver Configurations

Alexander Grebhahn, Norbert Siegmund, Sven Apel
University of Passau, Innstraße 33, 94032 Passau / Germany
E-Mail: {alexander.grebhahn,norbert.siegmund,apel}@uni-passau.de

Harald Köstler
Friedrich-Alexander University Erlangen-Nürnberg,
Cauerstraße 11, 91058 Erlangen / Germany
E-Mail: harald.koestler@fau.de

Multigrid solvers are among the most effective means to solve partial differential equations. To achieve the optimal performance for a multigrid system, its components and parameters (e.g., the smoother or the number of pre-smoothing and post-smoothing steps) have to be chosen carefully. Additionally, it is necessary to consider hardware-specific properties. To this end, it is often necessary to reimplement parts of a given system to achieve the optimal performance, which produces large maintenance costs and is error-prone. To overcome the reimplementation process, configurable multigrid systems can be used. Here, only the optimal variant has to be identified, which is also a non-trivial task, but requires less maintenance.
To identify the optimal variant of a configurable multigrid system, we use a *performance-influence model*, which is learned by a unique combination of forward feature selection and linear regression. A performance-influence model presents the influences of configuration options and their interactions on performance in a comprehensible way as a linear function. Prior to learning such a model, we select a specific set of configurations out of the configuration space, using heuristics and experimental designs. We can integrate domain knowledge into our sampling and learning process, to improve accuracy of the learned model and reduce effort for measurement and learning.
In this talk, we present different kinds of domain knowledge and how they can be used to reduce the number of measurements and to improve prediction accuracy and reduce measurement and learning effort. We present experimental results for two systems. One is the code generator developed in the ExaStencils project and the other is a hierarchical hybrid-multigrid system, implemented to run on large-scale systems, such as the JuQueen. For the ExaStencils code generator, we present the applicability of the approach on a complex system without using domain knowledge. In the second experiment, we demonstrate how to integrate domain knowledge in the learning approach and how our approach can be used to validate the knowledge.

3.2 Systems of Partial Differential Equations in ExaSlang

Christian Schmitt, Sebastian Kuckuk, Frank Hannig, Jürgen Teich, Harald Köstler, Ulrich Rüde
FAU Erlangen-Nürnberg / Germany
E-Mail: {christian.j.schmitt,sebastian.kuckuk}@fau.de
hannig@cs.fau.de
teich@informatik.uni-erlangen.de
{harald.koestler,ulrich.ruede}@fau.de

Christian Lengauer
University of Passau / Germany
E-Mail: lengauer@fim.uni-passau.de

As high-performance computing systems are becoming increasingly complex, writing software that attains maximum performance and scalability and is, at the same time, easily composable by application scientists without a deep computer science background is getting more and more challenging. Often, engineers and natural scientists do not have profound enough knowledge of programming and need experts to implement their codes who, on the other hand, are likely to have trouble understanding the physical problem to be solved. One possible remedy is to exploit the potential given by technologies, such as domain-specific languages (DSLs), that provide appropriate abstractions, automatic code generation support, and auto-tuning. In the domain of geometric multigrid solvers, project ExaStencils follows this road. It aims at providing highly optimized and scalable numerical solvers, specifically tuned towards the given problem and target hardware at hand. Its programming language is the hierarchical, four-layer DSL ExaSlang. In recent work, the most concrete layer, ExaSlang 4, has been introduced. Here, we introduce data types for local vectors to ExaSlang 4 as the next step towards ExaSlang 3, in order to support computations that use point-local vectors and matrices. These data types allow an intuitive modeling of many physical problems represented by systems of partial differential equations, e.g., the simulation of flows that include vector-valued velocities.

3.3 Preparing hypre for Emerging Architectures

Ulrike Meier Yang
Lawrence Livermore National Laboratory,
Box 808, L-561, Livermore, CA, 94551 / USA
E-Mail: umyang@llnl.gov

The hypre software library provides high performance preconditioners and solvers for the solution of large sparse linear systems on massively parallel computers. One of its attractive features is the provision of conceptual interfaces, which include a structured, a semi-structured, and a traditional linear-algebra based interface. The (semi-)structured interfaces are an alternative to the standard matrix-based interface that describes rows, columns, and coefficients of a matrix. Here, instead, matrices are described primarily in terms of stencils and logically structured grids. These interfaces give application users a more natural means for describing their linear systems, and provide access to methods such as structured multigrid solvers, which can take advantage of the additional information beyond just the matrix. Since current architecture trends are favoring regular compute patterns to achieve high performance, the ability to express structure has become much more important.
In this presentation, we describe our efforts to prepare hypre for emerging architectures, which include the creation of a new structured-grid matrix class that supports rectangular matrices and constant coefficients and a semi-structured-grid matrix class that builds on the new structured-grid matrix as well as some preliminary results for structured AMG on GPUs.

3.4 Efficient Solvers for Density Functional Theory

Rochus Schmid
Universität Bochum, Universitätsstr. 150, 44801 Bochum / Germany
E-Mail: rochus.schmid@rub.de

In electronic structure codes for chemical systems the quantum mechanic problem for the electrons needs to be solved for the given nuclei position. This is usually done via the Kohn-Sham Density Functional Theory (DFT) approximation, which involves solving a non-linear second order Eigenvalue problem. In the majority of electronic structure codes, the electronic wavefunctions and densities are expanded in an analytic basis (e.g. plane waves or a linear combination of Gaussian type atomic orbitals). In contrast to that our Real Space DFT code exploits a numerical discretization and stencil based operations for e.g. the Laplacian in the operator for the electronic kinetic energy, cf. Schmid et al. [Phys. Stat. Sol. (B) 243(5): 1001–1015, 2006]. For molecular dynamics simulations we use the Car-Parrinello approach, where both wavefunctions and nuclei positions are propagated in a fictitious joint dynamic. Our code is written in a combination of C-routines (using BLAS library calls) and an object oriented framework written in Python. The code is MPI-parallel. Currently, we apply the code for simulating charged solid/liquid interfaces, which occur in electrochemical systems. Here we exploit the freedom in the choice of boundary conditions, available in the Real Space approach and couple to a (modified) Poisson-Boltzmann model to represent the electrolyte in a continuum fashion. In the presentation, an overview over the code and a critical discussion of the specific aspects of stencil operations in the context of KS-DFT will be given. In particular, the current Multigrid Poisson solver needed to recompute the Coulomb potential in each timestep will be discussed.

EXASTEEL - Bridging Scales for Multiphase Steels

4.1 High Performance Computing and Strategies for the Multiscale Modeling and Simulation of Advanced High Strength Steels

Jörg Schröder, Ashutosh Gandhi
University of Duisburg-Essen / Germany
E-Mail: {j.schroeder,ashutosh.gandhi}@uni-due.de

Daniel Balzani
TU Dresden / Germany
E-Mail: daniel.balzani@tu-dresden.de

Axel Klawonn, Martin Lanser
University of Cologne / Germany
E-Mail: {klawonn,mlanser}@math.uni-koeln.de

Oliver Rheinbach
TU Bergakademie Freiberg / Germany
E-Mail: oliver.rheinbach@math.tu-freiberg.de

In many fields of steel applications, e.g. automotive engineering or manufacturing, the optimization of the material properties is one of the main challenges. Consequently, reliable computational modeling of sheet metal forming processes is required to give access to numerical approaches. Therein, the materials microstructure needs to be included as it has a dominant influence on the mechanical behavior. We focus on a direct multiscale approach often referred to as FE2, where at each macroscopic Gauss point of a finite element calculation additional microscopic boundary problems are solved in a scale-coupling scheme. We present a method which incorporates graded properties in the individual microscopic phases enabling an accurate representation of the material behavior at the macroscale. In FE2TI, our parallel MPI based implementation of the fully coupled FE2 method, these large and complex microscopic problems are solved in parallel using iterative FETI-DP solvers, which have proved to be highly scalable and robust for structural mechanics problems. This adds a second level of parallelism to the FE2 method. We will present scalability results of the FE2TI software up to the complete Mira BlueGene/Q (ANL) and JUQUEEN BlueGene/Q (JSC Jülich). We will also also show results obtained from a production run using the complete JUQUEEN with up to 458 752 cores, simulating the deformation of a large perforated plate using realistic microstructures.

4.2 Scalable Nonlinear Domain Decomposition Methods

Axel Klawonn, Martin Lanser
Mathematisches Institut, Universität zu Köln,
Weyertal 86-90, 50931 Köln / Germany
E-Mail: {axel.klawonn,martin.lanser}@uni-koeln.de

Oliver Rheinbach
Institut für Numerische Mathematik und Optimierung,
Fakultät für Mathematik und Informatik,
Technische Universität Bergakademie Freiberg,
Akademiestr. 6, 09596 Freiberg / Germany
E-Mail: oliver.rheinbach@math.tu-freiberg.de

The simulation of deformation processes of micro-heterogeneous materials, e.g., dual-phase steels, on modern supercomputers requires efficient and highly scalable implicit solvers for nonlinear problems in structural mechanics.
In this talk, the nonlinear and nonoverlapping domain decomposition methods "Inexact Reduced Nonlinear FETI-DP" (Dual-Primal Finite Element Tearing and Interconnecting) and "Inexact Nonlinear FETI-DP" will be presented. Both methods are capable of overcoming scalability limits of classical FETI-DP domain decomposition methods, which is achieved by moving the FETI-DP coarse problem to the preconditioner. This allows the inexact solution of the coarse problem using, e.g., some cycles of an algebraic multigrid method, without losing any accuracy of the solution. Additionally, the nonlinear domain decomposition helps to resolve local nonlinearities by solving local nonlinear problems on each FETI-DP subdomain. This localizes work, significantly reduces communication, and leads to improved performance compared to classical Newton-Krylov-FETI-DP approaches.
In this talk, numerical results for heterogeneous nonlinear elasticity problems are presented. Parallel weak and strong scalability up to half a million BlueGene/Q cores will be demonstrated using Vulcan (LLNL, USA), JUQUEEN (JSC, Jülich, Germany), and Mira (ANL, USA).

4.3 Software Organization for Nonlinear Solver Composition

Matthew Knepley
Computational and Applied Mathematics, Rice University / USA
E-Mail: knepley@gmail.com

t.b.a.

4.4 Increasing Scalability in Algebraic Multigrid for Elasticity

A. H. Baker
National Center for Atmospheric Research,
1850 Table Mesa Dr, Boulder, CO, 80305 / USA
E-Mail: abaker@ucar.edu

A. Klawonn, M. Lanser
Mathematisches Institut, Universität zu Köln,
Weyertal 86-90, 50931 Köln / Germany
E-Mail: {axel.klawonn, martin.lanser}@uni-koeln.de

T. Kolev
Lawrence Livermore National Laboratory,
Box 808, L-561, Livermore, CA, 94551, USA
E-Mail: tzanio@llnl.gov

O. Rheinbach
Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg, Akademiestr. 6, 09599 Freiberg / Germany
E-Mail: oliver.rheinbach@math.tu-freiberg.de

Ulrike Meier Yang
Lawrence Livermore National Laboratory,
Box 808, L-561, Livermore, CA, 94551 / USA
E-Mail: umyang@llnl.gov

Classical algebraic multigrid (AMG) methods are very effective for systems derived from elliptic scalar partial differential equations (PDEs) since the prolongation operators accurately interpolate the nullspace of the operator, which is constant. However when dealing with systems of PDEs, the nullspace contains several vectors, e.g. in systems derived from linear elasticity, the nullspace consists of the rigid body modes, which comprise both translations and rotations. Since classical AMG, including standard approaches modified for systems of PDEs, only interpolates the translations, the result is a loss of optimality and scalability.
In this presentation, we examine several approaches that incorporate the rotations into the interpolation, the global matrix approach as well as the local neighborhood approach, and compare their performance on high performance computers to standard classical AMG approaches for systems of PDEs, such as the unknown and a hybrid approach. Our experiments demonstrate a significant increase in scalability using up to half a million MPI processes when incorporating the rotations into the prolongation operators.

EXAHD - An Exa-scalable Two-level Sparse Grid Approach for Higher-dimensional Problems in Plasma Physics and Beyond

5.1 Computational Plasma Physics at the Exascale: Goals, Obstacles, and New Ideas

Frank Jenko
UCLA Physics and Astronomy / USA
E-Mail: jenko@physics.ucla.edu

Plasma physics has been an important driver for high performance computing for decades, given its key role in fusion research (and countless other applications) as well as in space and astrophysics. The fundamental equations of plasma physics are nonlinear, and many fundamental plasma processes involve a wide range of spatio-temporal scales. Therefore, simulations employing state-of-the-art algorithms and state-of-the-art supercomputers are vital to realistic descriptions of actual physical systems. In order to guide and interpret measurements in large-scale (and expensive) future fusion experiments like ITER and space missions like THOR, exascale simulations will be required, ushering in an era of increasingly theory-driven plasma research. Extended partnerships with applied mathematics and computer science are crucial for success. In my talk, I will discuss goals, obstacles, and new ideas in this context.

5.2 Recent Developments in the Theory and Application of the Sparse Grid Combination Technique

Markus Hegland, Brendan Harding
Australian National University, Canberra / Australia
E-Mail: {markus.hegland,brendan.harding}@anu.edu.au

Christoph Kowitz
Technical University of Munich, Munich / Germany
E-Mail: kowitz@in.tum.de

Dirk Pflueger
University Stuttgart, Stuttgart / Germany
E-Mail: dirk.pflueger@ipvs.uni-stuttgart.de

Peter Strazdins
Australian National University, Canberra / Australia
E-Mail: peter.strazdins@anu.edu.au

The sparse grid combination technique is used to solve complex systems of partial differential equations. Basically, it determines a linear combination

$$u^C = \sum_k c_k u_k \tag{5.1}$$

of component solutions u_k for different grid resolutions. The main example considered here are the gyrokinetic equations of plasma physics which are derived from the Boltzmann equations for ion and electron densities and are combined with the Maxwell equations for the electromagnetic fields. The component solutions are determined with the strongly scalable code GENE.
The sparse grid combination technique provides an additional level of parallelism which is well suited for exascale computing. A new version of the combination technique where both the underlying component grids and the combination coefficients are adapted is shown to lead to scalable algorithms which moreover are algorithmically fault-tolerant and thus provide an alternative to check-point restart. New methods based on the combination technique are derived for the solution of initial value problems and eigenvalue problems.

5.3 Scalable Algorithms for the Solution of Higher-dimensional PDEs

Mario Heene, Dirk Pflüger
IPVS, Universität Stuttgart, Universitätsstr. 38, 70569 Stuttgart / Germany
E-Mail: {mario.heene,dirk.pflueger}@ipvs.uni-stuttgart.de

The solution of higher-dimensional problems, such as the simulation of plasma turbulence in a fusion device as described by the five-dimensional gyrokinetic equations, is a grand challenge for current and future high-performance computing. The sparse grid combination technique is a promising approach to the solution of these problems on large-scale distributed memory systems. The combination technique numerically decomposes a single large problem into multiple moderately-sized partial problems that can be computed in parallel, independently and asynchronously of each other. The ability to efficiently combine the individual component grids to a common sparse grid is a key to the overall performance of such large-scale computations. In this work, we present new algorithms for the recombination of distributed component grids and demonstrate their scalability to at least 90,112 cores of a Cray XC40 HPC system.

5.4 SDC-Resilient Algorithms Using the Sparse Grid Combination Technique

Alfredo Parra Hinojosa
Fakultät für Informatik, TUM, Boltzmannstr. 3, 85748 Garching / Germany
E-Mail: hinojosa@in.tum.de

Brendan Harding, Markus Hegland
Mathematical Sciences Institute, ANU, John Dedman Building 27, 2601 Canberra / Australia
E-Mail: {brendan.harding,markus.hegland}@anu.edu.au

Hans-Joachim Bungartz
Fakultät für Informatik, TUM, Boltzmannstr. 3, 85748 Garching / Germany
E-Mail: bungartz@in.tum.de

In this talk we discuss the impact of Silent Data Corruption (SDC) on our parallel implementation of the Sparse Grid Combination Technique (SGCT) algorithm. SDC is a type of *silent*, or undetected, error in the floating point data that can lead to wrong simulation results. Bit flips are a typical cause of SDC, but it is difficult to tell with confidence how SDC will originate in future exascale systems. Although SDC is not expected to occur as often as processor failure in future supercomputers, a single occurrence can ruin the whole simulation, and evidence suggests that SDC will occur frequently enough to be taken seriously.
Once we understand what SDC is, we can discuss ways to simulate it in existing machines and talk about general strategies to develop robust algorithms in its presence. Instead of implementing checksums, replicating processes, or storing checkpoints (as it is most commonly done), we use the numerical properties of the SGCT to identify, filter, and recover from SDC. The fact that the SGCT has some inherent data redundancy can be exploited to determine if, and where, SDC has occurred. We propose two ways of doing this, which we illustrate in a simple 2D scenario. Our algorithms rely on calculating norms or using voting mechanisms, and we argue that they are easy to implement and come at a negligible computational cost. At the end of the talk we discuss possible parallelization strategies, which can be easily coupled with our existing code.

Terra-Neo - Integrated Co-design of an Exa-scale Earth Mantle Modeling Framework

6.1 Resilience for Multigrid Software at the Extreme Scale

Markus Huber
Institute for Numerical Mathematics,
Technische Universität München / Germany
E-Mail: huber@ma.tum.de

Björn Gmeiner, Ulrich Rüde
Institute of System Simulation, University Erlangen-Nuremberg / Germany
E-Mail: {gmeiner,ulrich.ruede}@fau.de

Barbara Wohlmuth
Institute for Numerical Mathematics,
Technische Universität München / Germany
E-Mail: wohlmuth@ma.tum.de

Future high performance systems will be characterized by millions of compute nodes that are executing up to a billion parallel threads. This compute power will be extremely expensive not only with respect to acquisition costs but also due to the operational costs, whereby the energy consumption is becoming a major concern. The increasing system size results in a higher probability of failure of the components of the HPC-system, and thus fail-safe performance is one of the new challenges in extreme scale computing.
Here, we concentrate on fault tolerant algorithms for the numerical approximation of elliptic partial differential equations. The design of exa-scale enabled iterative solvers for these equations plays an important role. We address this challenge by combining domain partitioning with highly scalable geometric multigrid schemes to obtain fast and fault-robust solvers in three dimensions. The recovery strategy is based on a hierarchical hybrid concept where the values on lower dimensional primitives such as faces are stored redundantly and thus can be recovered easily in case of a failure. The lost volume unknowns in the faulty region are re-computed approximately with multigrid cycles by solving a local Dirichlet problem on the faulty subdomain. Different strategies are compared and evaluated with respect to performance, computational cost, and speed up. Especially effective are strategies in which the local recovery in the faulty region is executed in parallel with global solves and when the local recovery is additionally accelerated. This results in an asynchronous multigrid iteration that can fully compensate faults. Excellent parallel performance on a current peta-scale system is demonstrated.

6.2 Mantle Convection on Modern Supercomputers

Jens Weismüller
LMU München, Department of Earth and Environmental Sciences, Theresienstr. 41, 80333 Munich / Germany
E-Mail: jens.weismueller@geophysik.uni-muenchen.de

Siavash Ghelichkhan, Hans-Peter Bunge
LMU München, Department of Earth and Environmental Sciences, Theresienstr. 41, 80333 Munich / Germany
E-Mail: {siavash.ghelichkhan,hans-peter.bunge}@geophysik.uni-muenchen.de

Björn Gmeiner, Ulrich Rüde
FAU Erlangen-Nürnberg, Cauerstr. 11, 91058 Erlangen / Germany
E-Mail: {bjoern.gmeiner,ulrich.ruede}@fau.de

Markus Huber, Lorenz John, Barbara Wohlmuth
TU München, Boltzmannstr. 3, 85748 Garching / Germany
E-Mail: {huber,barbara.wohlmuth,john}@ma.tum.de

Mantle convection is the cause for plate tectonics, the formation of mountains and oceans, and the main driving mechanism behind earthquakes. The convection process is modeled by a system of partial differential equations describing the conservation of mass, momentum and energy. Characteristic to mantle flow is the vast disparity of length scales from global to microscopic, turning mantle convection simulations into a challenging application for high-performance computing.
As an application, we consider flow in the uppermost mantle layer. A variety of geologic observations point to fast flow in this region that may exceed plate tectonic velocities by an order of magnitude. At the same time there is mounting evidence from seismology for flow-like structures in the upper 100-200 km of the mantle. Using high resolution simulations, we present a set of geodynamic simulations to link these observations. We include low viscosity channels of varying thickness, with an extreme case of 100 km and a significant viscosity contrast of up to 4 orders of magnitude relative to the deeper mantle. We obtain an increase in velocity by a factor of 10 between a 1000 km thick and the very thin channel, translating into velocities of $\sim$ 20 cm/a within the narrow channel.

6.3 Krylov Subspace Methods and Exascale Computations - A Good Match or Lost Case?

Zdeněk Strakoš
Charles University in Prague, Faculty of Mathematics and Physics, Sokolovska 83, 186 75 Prague 8 / Czech Republic
E-Mail: strakos@karlin.mff.cuni.cz

This contribution will attempt to point out possible strengths, weaknesses and open questions related to using Krylov subspace methods in massively parallel computational environments.
Krylov subspace methods can be seen as highly nonlinear model reduction that can be very efficient in some cases and too costly in other. Convergence behaviour is well understood for Hermitian (normal) matrices, where we can conveniently rely on the spectral decomposition. That does not have a parallel in non-normal cases. Construction and analysis of efficient preconditioners is therefore complicated and it is often based on a simplified view to Krylov subspace methods as linear contractions. Computational efficiency requires tight, reliable and cheap a-posteriori error estimators that would allow to construct stopping criteria lowering the cost of the whole computational process. Understanding numerical stability issues is crucial and this becomes even more urgent with increased parallelism where the cost of communications becomes a prohibitive factor. Finally, we emphasize importance of links between the infinite dimensional problems, their discretization, construction of preconditioners, and algebraic Krylov subspace methods computations.
The contribution will present results from the monographs coauthored with Josef Malek [Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs, SIAM Spotlights, SIAM, Philadelphia, 2015] and Jörg Liesen [Krylov Subspace Methods, Principles and Analysis, Oxford University Press, Oxford, 2013].

6.4 Numerical Methods in the Finite Element Mantle Convection Code ASPECT

Wolfgang Bangerth
Texas A&M University
E-Mail: bangerth@math.tamu.edu

Timo Heister
Clemson University
E-Mail: heister@clemson.edu

This talk focusses on the numerical challenges involved with the simulation of large scale mantle convection problems based on the experience developing the open source, community code ASPECT (Advanced Solver for Problems in Earth ConvecTion, `https://aspect.dealii.org/`) based on the finite element library deal.II (`https://dealii.org`).
We will discuss the usage of accurate finite element discretization and stabilization schemes for the underlying nonlinear PDEs. Large demands on the accuracy of the solution require the combination of adaptive mesh refinement techniques and large scale, parallel computations. We developed robust, efficient, and scalable linear solvers based on block preconditioners and algebraic multigrid for the resulting linear systems. The main focus of this talk is on the successful integration of all these different numerical challenges.
Finally, we will cover some advanced applications and highlight strategies in software design we employ to develop a flexible, reusable code with an active community.

ExaFSA - Exascale Simulation of Fluid-structure-acoustics Interactions

7.1 Partitioned Multi-Physics on Distributed Data via preCICE

Benjamin Uekermann, Hans-Joachim Bungartz, Alexander Shukaev
Scientific Computing in Computer Science, Technical University of Munich, Boltzmannstraße 3, 85748 Garching / Germany
E-Mail: {uekerman,bungartz}@in.tum.de

Florian Lindner, Miriam Mehl, Klaudius Scheufele
Institute for Parallel and Distributed Systems, University of Stuttgart, Universitätsstraße 38, 70569 Stuttgart / Germany
E-Mail: {florian.lindner,miriam.mehl,klaudius.scheufele}@ipvs.uni-stuttgart.de

One of the great prospects of exa-scale computing is to simulate challenging highly complex multi-physics scenarios with different length and time scales. A modular approach re-using existing software for the single-physics model parts has great advantages regarding flexibility and software development costs. At the same time, it poses challenges in terms of numerical stability and parallel scalability. The coupling library preCICE provides communication, data mapping, and coupling numerics for surface-coupled multi-physics applications in a highly modular way. In this talk, we give a brief introduction to preCICE and focus mainly on the library's realization and performance on distributed data.

7.2 Efficient Coupling of Acoustic and Flow on Massively Parallel Systems

Verena Krupp, Harald Klimach, Sabine Roller
Simulation Techniques & Scientific Computing, Hölderlinstr. 3, 57076 Siegen, Germany
E-Mail: {verena.krupp, harald.klimach,sabine.roller}@uni-siegen.de

Fluid-acoustic interaction is an important milestone towards full FSA, and yields a multi-scale problem including different lengths and timescales, which is numerically challenging as computation on the finest resolution for all parts of the domain is not feasible.
Our approach is a partitioned coupling, where we split the multi-scale problem into the single physical phenomena and realize their interaction via a coupling library. Though, each phenomenon can be simulated with an optimized solver using its specific equations, resolution and numerical method. Each component of such a coupled approach, needs to be highly scalable to not pose a bottleneck. Therefore, we are using the high order solver Ateles, which is proven to be highly scalable on distributed parallel systems. The opensource coupling library preCICE[1] has recently been parallelized with a peer-to-peer approach, and thus, is perfectly suited for the coupling on massively parallel machines.
With both, a scalable solver and a scalable coupling implementation, we present the scaling behavior for a two-field simulation on up to four islands with 32,768 MPI processes on the IBM DataPlex system SuperMUC at the Leibniz Computing Centre (LRZ) in Munich.
For an ideal load balancing, especially in such a multi-component approach, no component should wait for another and hence produce idle nodes. Using a parallel explicit (non-staggered) coupling method and on heuristic based static load balancing enable a large scale simulation of fluid-acoustic interaction, where both partitions, flow and acoustic, only spend 5% and 7% respectively in the coupling tool (including idling times due to imbalances).
The presented coupling approach allows for solving large multi-scale problems on massively parallel machines where the numerical method as well as the computational resources can be perfectly tailored to the physics and thereby archiving suitable load balancing.

[1] `http://www.precice.org`

7.3 Nonlinear Preconditioning and Multiphysics

Lulu Liu, David Keyes
Extreme Computing Research Center
King Abdullah University of Science and Technology / Saudi Arabia
E-Mail: {lulu.liu,david.keyes}@kaust.edu.sa

The multiplicative Schwarz preconditioned inexact Newton (MSPIN, 2014) algorithm, as a complement to additive Schwarz preconditioned inexact Newton (ASPIN, 2001), provides a Gauss-Seidel-like way to improve the global convergence of systems with unbalanced nonlinearities, preserving the root and transforming the convergence landscape while reusing the user-supplied functions and Jacobians of the original problem. We review the algorithmic and theoretical background of MSPIN, illustrate with various model problems and speculate on fruitful extensions to multiphysics problems, such as fluid-structure interaction.

7.4 A Code Transformation Approach to Achieving High Performance Portability

Hiroyuki Takizawa
Tohoku University, Japan
E-Mail: takizawa@cc.tohoku.ac.jp

Daisuke Takahashi
University of Tsukuba, Japan
E-Mail: daisuke@cs.tsukuba.ac.jp

Reiji Suda
The University of Tokyo, Japan
E-Mail: reiji@is.s.u-tokyo.ac.jp

Ryusuke Egawa
Tohoku University, Japan
E-Mail: egawa@cc.tohoku.ac.jp

Even today, system-aware performance optimizations are imperative to exploit the potential of a modern HPC system. As HPC system architectures are getting more complicated and diversified, the importance of system-aware performance optimizations will further increase in an upcoming extreme-scale computing era. One severe problem in system-aware performance optimizations is that, as the name implies, an HPC application code is optimized only for a particular system architecture and thus the performance may not be portable to another system. HPC application development and maintenance will become to require much more time and efforts due to the complexity and diversity of future HPC system architectures. To tackle this challenging problem, our research project is developing a code transformation framework, *Xevolver*, which allows users to define their own code transformation rules and also to customize typical rules for individual applications and systems. By expressing system-awareness as code transformations, users do not need to directly modify the application code for system-aware performance optimizations. This prevents the application code from being specialized for a particular HPC system, and thereby the application can achieve a high performance portability across diverse HPC systems. This talk introduces our research activities and some case studies to discuss the strengths and limitations of our approach to separation of system-awareness from application codes. Using only standard technologies and tools, Xevolver can achieve important system-aware performance optimizations, such as loop optimizations and data layout optimizations, without complicating the original application code.

ESSEX - Equipping Sparse Solvers for Exascale

8.1 Performance Engineering for Algorithmic Building Blocks in the GHOST Library

Georg Hager, Moritz Kreutzer
Erlangen Regional Computing Center (RRZE),
Martensstr. 1, 91058 Erlangen / Germany
E-Mail: {georg.hager,moritz.kreutzer}@fau.de

Gerhard Wellein
Department of Computer Science, University of Erlangen-Nuremberg,
Martensstr. 3, 91058 Erlangen / Germany
E-Mail: gerhard.wellein@fau.de

The GHOST (General, Hybrid, Optimized Sparse Toolkit) library is a collection of building blocks for algorithms dealing with large sparse matrices. It supports heterogeneous parallelism across architectural boundaries via the MPI+X paradigm and provides facilities for affinity-aware programming and fine-grained resource arbitration. GHOST was developed to support the algorithms and applications layers in the ESSEX (Equipping Sparse Solvers for Exascale) project, but its components are widely applicable in other fields as well. The library follows the concept of "holistic performance engineering," which means that the performance properties of its kernels are well understood in terms of their relevant bottlenecks on all architectures. This talk gives an overview of GHOST and the performance engineering process, presents instructive examples, and points to future work within the ESSEX project.

8.2 Fine-Grained Parallel Algorithms for Incomplete Factorization Preconditioning

Edmond Chow
School of Computational Science and Engineering,
Georgia Institute of Technology, Atlanta, Georgia / USA
E-Mail: echow@cc.gatech.edu

An incomplete LU (ILU) factorization is generally computed by a Gaussian elimination process where certain nonzeros are neglected in some way. Here, we present a completely different method for computing ILU factorizations. The new method is iterative in nature, is highly parallel, and benefits from asynchronous computation. The fact that the method is iterative means that the *exact* ILU factorization does not need to be computed, and we show experimentally that roughly converged factorizations can be very effective preconditioners. Furthermore, the number of iterations required is typically very small (less than 5). It is also easy to show theoretically that these iterations converge asymptotically.

Parallelizing the sparse triangular solves with the ILU factors is another challenge. Here, we advocate using Jacobi iterations to solve with the triangular factors, which is related to using Neumann series approximations which have also been advocated in the past, but have not been found to be robust. Such an approach also seems disadvantageous because the matrices are already triangular. However, just like in the iterative computation of the ILU factors, the motivation to use an iterative approach here is again parallelism and the lack of a need to solve these equations very accurately. Because the matrices are triangular, the Jacobi method is guaranteed to converge asymptotically. To improve robustness, a block Jacobi splitting may be used to improve the normality of the iteration matrix. Here, it is useful to find reorderings that produce diagonal blocks with large norms, but at the same time do not adversely affect ILU preconditioning.

8.3 Recent Advancements and Future Plans for Next-Generation Sparse Solvers in Trilinos

Heidi Thornquist
Electrical Modeling and Simulation, Sandia National Laboratories / USA
E-Mail: hkthorn@sandia.gov

With the availability and diversity of powerful computational resources, including multi-core CPU and GPU technology, there is significant interest in numerical software libraries that allow a developer to optimize the trade-off between effort and impact. In this talk we will discuss the current and ongoing efforts by which Trilinos is providing enabling technologies for the development of academic and industrial software targeted at next-generation architectures. The focus of this talk will be on the sparse linear solvers (Belos) and eigensolvers (Anasazi) packages in Trilinos. Since the scalable performance of these solver packages fundamentally hinges upon efficient linear algebra and preconditioners, the discussion will also include linear algebra libraries (Tpetra/Kokkos) and node-level solvers (ShyLU/Basker).

8.4 Feeding of the Thousands – Leveraging the GPU's Computing Power for Sparse Linear Algebra

Hartwig Anzt
Innovative Computing Laboratory, University of Tennessee,
1122 Volunteer Blvd, Knoxville, TN 37996 / USA
E-Mail: hanzt@icl.utk.edu

Graphics processing units provide very good floating point performance for compute-intense kernels that provide fine-grained parallelism. Significant acceleration potential is available for dense linear algebra, where compute-intense kernels may achieve flop rates close to the theoretical peak. The situation is different for sparse linear algebra, as for those the memory bandwidth usually becomes the performance-limiting factor. Composing algorithms out of the standard BLAS operations often results in unnecessary data movement, and performance capabilities remain unused. In many cases, exploiting data locality by fusing multiple arithmetic operations into one kernel allows for much better GPU utilization, but toolkits for solving sparse linear algebra problems have to balance between performance optimization and flexibility. In this talk, we introduce the concept of kernel fusion, evaluate the benefits for different sparse linear algebra solvers, and quantify the solver's efficiency with respect to theoretical bounds. We also investigate the potential of non-traditional approaches to sparse linear algebra problems, that aim at compensating numerical drawbacks with enhanced parallelism.

8.5 Can Block Iterative Eigensolvers Fulfill their Performance Promise?

Jonas Thies
DLR, Simulation and Software Technology,
Linder Höhe, 51147 Cologne/Germany
E-Mail: jonas.thies@dlr.de

Melven Röhrig-Zöllner
DLR, Simulation and Software Technology,
Linder Höhe, 51147 Cologne/Germany
E-Mail: melven.roehrig-zoellner@dlr.de

Iterative algorithms for sparse linear systems and eigenvalue problems typically face three bottlenecks on a supercomputer: the bandwidth of the main memory and network, and the network latency for global reduction operations. It is well known that block variants of eigensolvers can overcome these bottlenecks to some extent, by performing sparse matrix-vector multiplication (spMVM) on several vectors at once and collecting several scalar products into a single reduction. We give an overview of useful implementation details to make such methods work efficiently and achieve the expected block speed-up—from basic BLAS-like operations and using higher precision arithmetic in memory bounded situations to choices in algorithms and data structures to exploit these highly optimized kernels.

DASH - Hierarchical Arrays for Efficient and Productive Data-intensive Exascale Computing

9.1 A Tiling Based Programming Model and Its Supportive Tools

Didem Unat
Koc University, Sariyer, Istanbul / Turkey
E-Mail: dunat@ku.edu.tr

With the current trends in system design, the next generation of high-performance computing systems will have node architectures based on many-core processors and non-uniform memory access (NUMA) designs. In response to these architectural challenges, we are adopting a tiling-based programming model to expose additional parallelism and enhance data locality. In this talk, I will describe the programming model approach and present performance results on AMR frameworks. I will also discuss our future plans for supportive compiler and asynchronous execution of tiles and further data locality optimization on heterogeneous systems.

9.2 HPX - A Task Based Programming Model in a Global Address Space

Thomas Heller
Friedrich-Alexander-Universität Erlangen-Nürnberg / Germany
E-Mail: thomas.heller@cs.fau.de

t.b.a.

9.3 Expressing and Exploiting Multidimensional Locality in DASH

Tobias Fuchs, Karl Fürlinger
Ludwig-Maximilians-Universität (LMU) Munich, MNM Team, Oettingenstr. 67, 80538 Munich/Germany
E-Mail: {tobias.fuchs,karl.fuerlinger}@nm.ifi.lmu.de

DASH is a realization of the PGAS (partitioned global address space) programming model in the form of a C++ template library. It provides a multidimensional array abstraction which is typically used as an underlying container for stencil- and dense matrix operations. Efficiency of operations on a distributed multidimensional array highly depends on the distribution of its elements to processes and the communication strategy used to propagate values between them. Locality can only be improved by employing an optimal distribution that is specific to the algorithm's data access pattern, run-time parameters such as node topology, and numerous additional aspects. Application developers have no way of knowing these implications which also might change in future releases of DASH. In this talk, we present fundamental properties of distribution patterns that identified in related work and existing HPC applications. We describe a classification scheme of multi-dimensional distributions based on these properties and demonstrate how optimal distribution patterns can be determined automatically and, to a great extent, at compile time.

9.4 Tool Support for Developing DASH Applications

Denis Hünich, Sebastian Oeste, Andreas Knüpfer
TU Dresden, Zellescher Weg 12, 01069 Dresden / Germany
E-Mail: {denis.huenich,sebastian.oeste,andreas.knuepfer}@tu-dresden.de

Karl Fürlinger, Tobias Fuchs
LMU München, Oettingenstr. 67, 80538 Munich / Germany
E-Mail: {karl.fuerlinger,tobias.fuchs}@nm.ifi.lmu.de

DASH is a new parallel programming model for HPC which is implemented as a C++ template library plus a runtime library which can be implemented on top of various PGAS (Partitioned Global Address Space) substrates. DASH's goal is to be an easy and efficient approach to parallel programming with C++. Part of the initial plans for DASH were supporting software tools. This includes debugging as well as performance monitoring. First, debugging is particularly necessary while adopting a new parallelization model. Second, performance assessment is naturally important in High Performance Computing. While support tools are generally an important part of a programming ecosystem, we are convinced that providing tools early brings multiple advantages. This is true for application developers using DASH but also for the DASH development itself. This talk presents the specific debugging and performance monitoring extensions for DASH as well as exemplarily use cases and an early assessment of the results.

EXAMAG - Exascale Simulations of the Evolution of the Universe Including Magnetic Fields

10.1 Cosmology on a Moving Mesh

Volker Springel
Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg / Germany
Zentrum für Astronomie, Universität Heidelberg, Astronomisches Recheninstitut, Mönchhofstr. 12-14, 69120 Heidelberg / Germany
E-Mail: volker.springel@h-its.org

Simulations of cosmic structure formation address multi-scale, multi-physics problems of vast proportions. These calculations are presently at the forefront of today's use of supercomputers, but continued future success in this field requires the development of new numerical methods that excel in accuracy, robustness, parallel scalability, and physical fidelity to the processes relevant in galaxy and star formation. In the EXAMAG project we aim to substantially improve the astrophysical moving-mesh code AREPO and extend its range of applicability, with the goal of producing a competitive application code for the upcoming large computing platforms. Here I report on some of our results for new, powerful high-order discontinuous Galerkin schemes, on anisotropic transport of heat and relativistic particles, and on improvements of the accuracy of the treatment of ideal magnetohydrodynamics. I will also highlight some recent scientfic results obtained with our new methods, including the hydrodynamical cosmological "Illustris" simulation of galaxy formation, and simulations of Milky Way-like disk galaxies that for the first time are able to make realistic predictions for the built-up of the Galaxy's magnetic field throughout cosmic history.

10.2 Progress on Higher Order Schemes for Cosmological Applications

Christian Klingenberg
Dept. of Mathematics, Würzburg University / Germany
E-Mail: klingen@mathematik.uni-wuerzburg.de

Praveen Chandrashekar
Tata Institute of Fundamental Research, Bangalore / India
E-Mail: praveen@tifrbng.res.in

Kevin Schaal, Volker Springel
Heidelberg Institute for Theoretical Studies, Heidelberg / Germany
E-Mail: Kevin.Schaal@h-its.org
volker.springel@h-its.org

Gero Schnücke
Dept. of Mathematics, Würzburg University / Germany
E-Mail: gero.schnuecke@mathematik.uni-wuerzburg.de

We will talk about our contribution to a project with the goal of a self-consistent numerical simulation of the evolution of the universe. It requires the development of new numerical methods that excel in accuracy and parallel scalability. These numerical methods are based on mathematical theory in order to guarantee the above mentioned requirements. We shall give an overview of progress we have made to that end in the EXAMAG project:

- higher order discontinuous Galerkin methods on a Cartesian grid with automatic mesh refinement for the Euler equations
- a moving mesh arbitary Lagrangian-Eulerian method for discontinuous Galerkin method for the Euler equations
- an entropy stable discontinuous Galerkin method for the Euler and MHD equations

10.3 Task-based Parallelization of a Transport Discontinuous Galerkin Solver. Applications to Plasma Models.

Philippe Helluy
IRMA, University of Strasbourg, 7 rue Descartes, Strasbourg / France
E-Mail: helluy@unistra.fr

We present an implicit Discontinuous Galerkin solver for the transport equation. Because of the upwind nature of the numerical flux the linear system in the implicit step is block triangular. The scheme is thus well adapted to a task-based implementation. We present such an implementation using the StarPU library and applications to plasma physics.

10.4 Multi-component Cosmological Simulations: Collisional and Collisionless Fluids

Naoki Yoshida
Department of Physics, University of Tokyo,
7-3-1 Hongo, Bunkyo, 113-0033 Tokyo / Japan
E-Mail: naoki.yoshida@ipmu.jp

We present the results from state-of-the-art radiation-hydrodynamics simulations of structure formation in the early universe. Our simulations follow the evolution of a cosmic primordial gas from a diffuse state to a compact star; the resulting density evolution ranges over twenty orders of magnitude. We employ a novel adaptive mass refinement technique to achieve the extreme dynamic range. Non-equilibrium chemistry and radiation transfer are directly coupled to hydrodynamics in a self-consistent manner, and in a fully cosmological context. We show how the density ripples left over from the Big Bang drive the formation of the first star in the universe. By combining a stellar evolution model, we calculate the growth of a primordial protostar through to the early stages as a star with thermonuclear burning. We utilize a number of realizations to predict the mass distribution of the first stars, which can be compared with the inferred masses from observations of very old stars in the Galaxy. Finally, we discuss the future prospects for simulating the formation of the first galaxies. As a second topic, we introduce a 'classic but novel' method for following the gravitational dynamics of a collisionless fluid such as cosmic dark matter. Unlike the conventional N-body techniques, we integrate directly the collisionless Boltzmann equation in the full six-dimensional phase space. We test our code by solving basic problems such Landau damping and the two-stream instability. We compare the test results by using a third-order positive flux conserved scheme and a fifth-order MP5 scheme. We also develop a hybrid N-body + Boltzmann solver that enables us to follow structure formation with a cold dark matter component and a hot neutrino component. While we identify significant velocity noise with the conventional particle method, we reproduce smooth but fine structures in momentum space with our Boltzmann solver using 128^6 grids. We discuss the prospects for larger-scale simulations in the exa-scale era.

FFMK - A Fast and Fault Tolerant Microkernel-based System for Exascale Computing

11.1 FFMK: Building an Exascale Operating System

Hermann Härtig
TU Dresden, 01062 Dresden/Germany
E-Mail: hermann.haertig@tu-dresden.de

Carsten Weinhold
TU Dresden, 01062 Dresden/Germany
E-Mail: carsten.weinhold@tu-dresden.de

The FFMK project designs, builds, and evaluates an operating system platform to cope with the challenges expected for Exascale supercomputing. In particular, these include performance losses caused by the much larger impact of run-time variability within applications, hardware, and the operating system. Furthermore, the vastly increased amount of components of future Exascale systems will make applications running on top more vulnerable to failures.
We believe that a systems-software design for Exascale machines that addresses these challenges must be based on a coordinated approach across all layers, including applications. We envision the platform as a whole will be managed by distributed load-balancing and fault-tolerance services, which are capable of supporting dynamically changing partitions. Monitoring and decision making is done at three levels: (1) on each multi-core node, (2) per application among nodes, and (3) based on a global view of master management nodes.
The architecture builds upon a node-local, small light-weight kernel supporting a combination of specialized runtimes and a full-blown general-purpose operating system. For FFMK, we have instantiated the architecture components with the L4 microkernel, virtualized Linux, and MPI as an application runtime for bulk-synchronous applications. We carefully split system and runtime components such that the performance of applications remains undisturbed by management-related background activities. The performance isolation and local load balancing are enforced by schedulers; the basic building blocks of global platform management are randomized Gossip algorithms and decentralized decision making. XtreemFS serves as a fault-tolerant checkpoint store that uses local memory of all nodes to achieve scalability.
In the talk, we will give an overview of these building blocks and how they are combined into a coherent architecture in the FFMK OS. We will present lessons learned and discuss results from the first phase of the project, before giving an outlook on the research challenges we plan to tackle in next funding period.

11.2 Fast In-Memory Checkpointing with POSIX API for Legacy Exascale-Applications

J. Fajerski, M. Noack, A. Reinefeld, F. Schintke, T. Schütt
Zuse Institute Berlin, Takustr. 7, 14195 Berlin / Germany
E-Mail: {fajerski,noack,reinefeld,schintke,schuett}@zib.de

Fault-tolerance is already an important issue in current supercomputers. But for the success of exascale computers it will become a decisive factor, because the high number of components and their complexity make these systems much more vulnerable. On exascale computers, system failures will become the norm rather than an exception in long-running applications.

Application-level checkpointing is a popular means to improve fault tolerance. When a crash occurs, a checkpoint is retrieved and the program execution resumes from that point on. Unfortunately, checkpoint/restart will come to the limits as the gap between the applications' memory footprint and the limited parallel I/O bandwidth becomes more pronounced. Writing a checkpoint to disk will cause the system processors to idle for a growing proportion of time and, even worse, the mean time between failure may be shorter than the time needed to save a checkpoint to disk.

We present a scheme that writes erasure encoded checkpoints into the main memory of other nodes in the HPC system. The rationale for this is twofold: first, writing to memory over the interconnect is several orders of magnitude faster than to disk and second, erasure encoded data is able to survive component failures.

We use a distributed file system with a tmpfs backend and a POSIX compliant FUSE client. Due to the unavailability of FUSE on supercomputers, we developed a client that intercepts and substitutes application calls to the file system with the LD_PRELOAD mechanism. This POSIX-compliant API allows legacy applications, which are prepared for application-level checkpoint/restart, to quickly materialize their checkpoint via the supercomputer's interconnect without the need to change the code.

Experimental results on a Cray XC40 show that our LD_PRELOAD client yields 70% better sequential bandwidth (with striping) than FUSE while still being transparent to the application. With erasure encoding the performance compared to striping is 17% to 49% worse because of the encoding effort and the additional data handling required for redundancy. These results are preliminary and may improve in the future. Even so, this already indicates that erasure-encoded memory checkpoint/restart is an effective means to improve resilience for exascale computing.

11.3 HPC Operating System Design

Rolf Riesen
Hillsboro, Oregon / USA
E-Mail: rolf.riesen@intel.com

Future exa-scale and larger systems will have many more CPU cores than today's systems, different memory types, a more complex memory hierarchy, large on-chip interconnects, and Non-Uniform Memory Access (NUMA) properties. These resources need to be carefully and efficiently managed in order to let applications scale and perform at the desired levels for which these systems are being designed.

While the hardware is getting more complex, innovative runtime systems, programming models, and the applications themselves are evolving as well. The operating system for such machines has to be able to adapt quickly to new paradigms and requirements. It needs to be easy to experiment with to try out innovative ideas. Of course, the operating system for these extreme-scale machines needs to be high performance and very scalable. And, it needs to be Linux compatible.

Our group at Intel is researching ways to combine these conflicting and difficult requirements into a new operating system that embeds a lightweight kernel into Linux. The compute cores of a node run the lightweight kernel for performance and scalability. The other cores run the Linux kernel to deliver traditional and non-time-critical services.

11.4 The Power of Performant Code Abstraction in Real World HPC Applications

Michael Bussmann, Carlchristian Eckert, Axel Huebl, Fabian Jung, René Widera, Malte Zacharias, Erik Zenker
Helmholtz-Zentrum Dresden - Rossendorf, Bautzner Landstrasse 400, 01328 Dresden / Germany
E-Mail: {m.bussmann,c.eckert,a.huebl,f.jung}@hzdr.de
{r.widera,m.zacharias,e.zenker}@hzdr.de

Guido Juckeland, Andreas Knüpfer, Wolfgang Nagel, Benjamin Worpitz
Technische Universität Dresden, Zentrum für Informationsdienste und Hochleistungsrechnen (ZIH), Helmholtzstr. 10, 01069 Dresden / Germany
E-Mail: {guido.juckeland,andreas.knuepfer,wolfgang.nagel}@tu-dresden.de

Many-core architectures are going to stay for a while. Thus, scaling real world applications on HPC systems with thousands of nodes and hybrid architectures will be important to application developers. We present two real world applications, plasma simulations and high repetition rate data analysis, that strongly benefit from speed ups seen with accelerator hardware. From our viewpoint as users with real-world problems we discuss our path towards scalable, hardware-independent, performant application codes. Our focus is not to reinvent the wheel, but to make full use of the power of existing C++ codes by fully abstracting all parts of a code that require optimization, separating concepts used in the code from their optimized implementation. We do this by providing light-weight, single-purpose, zero-overhead interfaces to central code components. In the talk I will discuss our experiences with this approach, the reusable interfaces that have come out of this and the vision behind putting everything together.

CATWALK - A Quick Development Path for Performance Models

12.1 Automatic Performance Modeling of HPC Applications

Felix Wolf
TU Darmstadt, Mornewegstr. 30, 64293 Darmstadt / Germany
E-Mail: wolf@cs.tu-darmstadt.de

Christian Bischof, Alexandru Calotoiu, Christian Iwainsky, Sergei Shudler
TU Darmstadt, Mornewegstr. 30, 64293 Darmstadt / Germany
E-Mail: {christian.bischof,christian.iwainsky}@sc.tu-darmstadt.de,
{calotoiu, shudler}@cs.tu-darmstadt.de

Torsten Hoefler, Grzegorz Kwasniewski
ETH Zürich, Universitaetsstrasse 6, 8092 Zürich / Switzerland
E-Mail: {htor,gkwasnie}@inf.ethz.ch

Bernd Mohr, Alexandre Strube
Jülich Supercomputing Centre, 52425 Jülich / Germany
E-Mail: {b.mohr,a.strube}@fz-juelich.de

Andreas Vogel, Gabriel Wittum
Goethe Universität Frankfurt, Kettenhofweg 139,
60325 Frankfurt am Main / Germany
E-Mail: {andreas.vogel,wittum}@gcsc.uni-frankfurt.de

Many existing applications suffer from inherent scalability limitations that will prevent them from running at exascale. Current tuning practices, which rely on diagnostic experiments, have drawbacks because (i) they detect scalability problems relatively late in the development process when major effort has already been invested into an inadequate solution and (ii) they incur the extra cost of potentially numerous full-scale experiments. Analytical performance models, in contrast, allow application developers to address performance issues already during the design or prototyping phase. Unfortunately, the difficulties of creating such models combined with the lack of appropriate tool support still render performance modeling an esoteric discipline mastered only by a relatively small community of experts. This article summarizes the results of the Catwalk project, which aimed to create tools that automate key activities of the performance modeling process, making this powerful methodology accessible to a wider audience of HPC application developers.

12.2 Automated Performance Modeling of the UG4 Simulation Framework

Sebastian Reiter, Arne Nägel, Andreas Vogel, Gabriel Wittum
Goethe Universität Frankfurt,
Kettenhofweg 139, 60325 Frankfurt am Main / Germany
E-Mail: sebastian.reiter@gcsc.uni-frankfurt.de
{arne.naegel,andreas.vogel,wittum}@gcsc.uni-frankfurt.de

Alexandru Calotoiu, Felix Wolf
Technische Universität Darmstadt,
Mornewegstraße 30, 64293 Darmstadt / Germany
E-Mail: {calotoiu,wolf}@cs.tu-darmstadt.de

Alexandre Strube
Forschungszentrum Jülich, Wilhelm-Johnen-Straße, 52425 Jülich / Germany
E-Mail: a.strube@fz-juelich.de

Many scientific research questions such as the drug diffusion through the human stratum corneum are formulated in terms of partial differential equations and their solution is numerically addressed using grid based finite element methods. For detailed and more realistic physical models those computational tasks become challenging and thus plenty of computational resources and complex numerical codes with good scaling properties up to largest core counts are required. The UG4 framework is employed for such simulations and prior empirical tests have shown good scaling properties. However, the sheer size of the code prevents an analytical and fine grained performance modeling. We therefore applied automated performance modeling to entire UG4 simulations in order to analyze and validate the scalability of the code. Focussing on the human skin permeation problem we present our results and show how performance bottlenecks in the numerical software are detected, predicted, removed, or ruled out using the performance modeling approach. Due to the analysis we were able to address potential future scalability issues and predict the behavior on upcoming largest machines.

12.3 Exploring Emerging Technologies in the Extreme Scale HPC Co-Design Space with Aspen

Jeffrey S. Vetter
Oak Ridge National Laboratory and Georgia Institute of Technology / USA
E-Mail: vetter@ornl.gov

Concerns about energy-efficiency and reliability have forced our community to reexamine the full spectrum of architectures, software, and algorithms that constitute our ecosystem. While architectures have remained relatively stable for almost two decades, new architectural features, such as heterogeneous processing, nonvolatile memory, and optical interconnection networks, will demand that software systems and applications be redesigned to exploit these new capabilities. A key capability of this activity is accurate modeling of performance, power, and resiliency. We have developed the Aspen performance modeling language that allows fast exploration of the design space. Aspen is a domain specific language for structured analytical modeling of applications and architectures. Aspen is designed to enable rapid exploration of new algorithm and architectures. Because of the succinctness, expressiveness, and composability of Aspen, it can be used to model many properties of a system including performance, power, and resiliency. Aspen has been used to model traditional HPC applications, and recently extended to model scientific workflows for HPC systems and scientific instruments, like ORNL's Spallation Neutron Source. Models can be written manually or automatically generated from other structured representations, such as application source code or execution DAGs. These Aspen models can then be used for a variety of purposes including predicting performance of future applications, evaluating system architectures, informing runtime scheduling decisions, and identifying system anomalies. Aspen is joint work with Jeremy Meredith (ORNL).

12.4 Algorithmic Time, Energy, and Power Trade-offs in Graph Computations

Richard Vuduc
Georgia Institute of Technology / USA
E-Mail: richie@cc.gatech.edu

t.b.a.

GROMEX - Unified Long-range Electrostatics and Dynamic Protonation for Realistic Biomolecular Simulations on the Exascale

13.1 How to do Nothing in Less Time

Ivo Kabadshow, David Haensel, Andreas Beckmann
Jülich Supercomputing Centre, Research Centre Jülich / Germany
E-Mail: {i.kabadshow,d.haensel,a.beckmann}@fz-juelich.de

In today's MD simulations the scaling bottleneck is shifted more and more from computation towards communication. Especially the calculation of long-range interactions in O(N) amplifies this imbalance even more. Within the GROMEX project we use the Fast Multipole Method to calculate Coulomb interactions in linear complexity. But how can we reduce the total runtime further in a latency-dominated environment? We have to change the communication pattern to reduce the latency and bandwidth costs along the critical communication path. However, this will lead to additional replication of computation which eventually can be overlapped with communication.
To reduce the visibility of the introduced communication layer inside the algorithm we employ several abstraction layers to hide the actual implementation. This also allows an easy exchange of the underlying communication techniques (blocking/non-blocking, one-sided/two-sided) without changing code at the algorithmic level. Additionally, this enables us to implement different parallelization schemes on top of this infrastructure.
In this talk we show the advantages of the implemented communication concept as well as the straightforward application thereof. Finally we present benchmark results of such a latency-avoiding communication scheme.

13.2 FMM Goes GPU: Smooth Trip or Bumpy Ride?

Bartosz Kohnke, Carsten Kutzner, Thomas Ullmann, Helmut Grubmüller
Max Planck Institute for Biophysical Chemistry,
Department of Theoretical and Computational Biophysics,
Am Fassberg 11, 37077 Göttingen / Germany
E-Mail: {bartosz.kohnke,ckutzne,thomas.ullmann}@mpibpc.mpg.de,
hgrubmu@gwdg.de

Andreas Beckmann, Ivo Kabadshow
Institute for Advanced Simulation (IAS), Jülich Supercomputing Centre (JSC),
Wilhelm-Johnen-Straße, 52425 Jülich / Germany
E-Mail: {a.beckmann,i.kabadshow}@fz-juelich.de

The N-body problem provides a very simple, yet scientific algorithm to utilize modern GPUs. However, the computational complexity is $O(N^2)$. An algorithm reducing runtime and complexity to optimal $O(N)$ for any required precision is the Fast Multipole Method (FMM). The FMM allows the efficient computation of the long-range interactions between particles by means of expansions of the potentials into multipole/Taylor moments. The algorithm requires several operators, partly depending up on each other, performing computations in a octree data structure. The operators shift (M2M) multipole moments from the leaves up to the root node of the tree, translate (M2L) multipole moments to Taylor moments and shift (L2L) Taylor moments down to the leaves. The accuracy of the approximate solution depends on the degree of the expansion p. The cost of applying the operators to the multipole/Taylor moments is $O(p^4)$. In this talk, we present our CUDA-enabled, templated C++ implementation of the FMM, describe the obstacles and resulting algorithmic design decisions towards our presented solution. We especially focus on employment of dynamic parallelism to reduce the massive index operation overhead of the operators that is being done with naive approach to the parallelization. We also allude to the exploitation of unified memory that minimizes the porting efforts and to utilization of hardware supported Atomics to enhance reduction steps. We will present timings/scalings for several FMM operators and will discuss different parallelization strategies of the workflow within the tree and within the operators.

13.3 High-throughput molecular dynamics simulations and Markov modeling

Frank Noé
FU Berlin, Arnimallee 6, 14195 Berlin / Germany
E-Mail: frank.noe@fu-berlin.de

By combining massively distributed or weakly coupled molecular dynamics (MD) simulations with Markov state models (MSMs) and related approaches we can now effectively sample equilibrium kinetics on millisecond timescales for moderately-sized atomistic protein models. This allows us to explore protein-ligand association, binding affinities and association/dissociation rates, and even push to protein-protein assocation. Extremely rare event sampling, such as the dissociation of a stable protein-inhibitor complex, is still out of reach for unbiased MD. We therefore show that by combining massive unbiased MD simulations by relatively little enhanced sampling data results in accurate sampling of both thermodynamics and kinetics of stable protein-inhibitor complexes [Wu et al., J. Chem. Phys. 141, 214106 (2014)].

13.4 Should I Stay or Should I Go: Proton Transfer Revisited

Paul Czodrowski
Merck KGaA, Frankfurter Strasse 250, 64293 Darmstadt / Germany
E-Mail: paul.czodrowski@merckgroup.com

Protonation changes are omnipresent events in the protein-ligand process. Experimental validation of these events is often hindered by the high demand of protein supply. pK_a calculations are therefore a resource-saving alternative and enable a large-scale study on many different systems. Besides the bound state of well studied protein-ligand complexes, results for kinases amenable for covalent attack will be presented. Understanding the protonation change is the last missing piece in a complete view of the molecular recognition process.

ExaSolvers - Extreme Scale Solvers for Coupled Problems

14.1 Hierarchical Low Rank Approximation for Extreme Scale Uncertainty Quantification

Lars Grasedyck
RWTH Aachen / Germany
E-Mail: lgr@igpm.rwth-aachen.de

We consider the problem of uncertainty quantification for extreme scale parameter dependent problems where an underlying low rank property of the parameter dependency is assumed. For this type of dependency the hierarchical Tucker format offers a suitable framework to approximate a given output function of the solutions of the parameter dependent problem from a number of samples that is linear in the number of parameters. In particular we can a posteriori compute the mean, variance or other statistical quantities of interest. In the extreme scale setting it is already assumed that the underlying fixed-parameter problem is distributed and solved for in parallel. We provide in addition a parallel evaluation scheme for the sampling phase that allows us on the one hand to combine several solves and on the other hand parallelise the sampling.

14.2 Scalable Shape Optimization Methods for Structured Inverse Modeling in 3D Diffusive Processes

Arne Nägel
Goethe Universität, Kettenhofweg 139, 60325 Frankfurt a.M./Germany
E-Mail: naegel@gcsc.uni-frankfurt.de

Volker Schulz, Martin Siebenborn
Trier University, FB IV, Department of Mathmatics, 54286 Trier / Germany
E-Mail: {volker.schulz,siebenborn}@uni-trier.de

Gabriel Wittum
Goethe Universität, Kettenhofweg 139, 60325 Frankfurt a.M. / Germany
E-Mail: wittum@gcsc.uni-frankfurt.de

We consider inverse modeling of the shape of cells in the outermost layer of the human skin. We propose a novel algorithm that combines mathematical shape optimization with high-performance computing. Our aim is to fit a parabolic model for drug diffusion through the skin to data measurements. The degree of freedom is not the permeability itself, but the shape that distinguishes regions of high and low diffusivity. These are the cells and the space in between. The key part of the method is the computation of shape gradients, which are then applied as deformations to the finite element mesh, in order to minimize a tracking type objective function. Fine structures in the skin require a very high resolution in the computational model. We therefore investigate the scalability of our algorithm up to millions of discretization elements.

14.3 Parallel Solvers in Space and Time

Martin Neumüller
Institute of Computational Mathematics, Johannes Kepler University, Altenberger Str. 69, 4040 Linz / Austria
E-Mail: neumueller@numa.uni-linz.ac.at

We present and analyze a new space-time parallel multigrid method for parabolic equations. The method is based on arbitrarily high order discontinuous Galerkin discretizations in time, and a finite element discretization in space. The key ingredient of the new algorithm is an inexact block Jacobi smoother. By using local Fourier mode analysis we determine asymptotically optimal smoothing parameters, a precise criterion for semi-coarsening in time or full coarsening, and give an asymptotic two grid contraction factor estimate. We then explain how to implement the new multigrid algorithm in parallel, and show with numerical experiments its excellent strong and weak scalability properties. This space-time multigrid method is also used to solve an optimal control problem and the Stokes equations.

14.4 Solving Big Problems in Neuroscience

Markus Breit
G-CSC, University of Frankfurt, 60325 Frankfurt a.M. / Germany
E-Mail: markus.breit@gcsc.uni-frankfurt.de

Gillian Queisser
Temple University, 1805 N. Broad Street, Philadelphia, PA 19122-6094 / USA
E-Mail: gillian.queisser@temple.edu

Martin Stepniewski
G-CSC, University of Frankfurt, 60325 Frankfurt a.M. / Germany
E-Mail: martin.stepniewski@gcsc.uni-frankfurt.de

Computation-based neuroscience is gradually moving from analogy-driven modeling to detailed modeling and simulation that is based on physical first principles. Novel three-dimensional models of biochemical and electrical signals, that incorporate detailed three-dimensional and ultra-structural geometry information in the computational domain, are pushing computational problem sized towards the exa-scale.
This talk will introduce big problems in neuroscience and will highlight how progress in scalable numerical methods is – and in the future will – significantly change the interdisciplinary field of computational neuroscience. In part of the talk will discuss discretization methods for handling complex three-dimensional morphologies and for coupling hybrid-dimensional models and will apply these to electrodiffusion models defined by the Poisson-Nernst-Planck equations and large-scale networks simulations.

New Projects

15.1 Advanced Data Placement via Ad-hoc File Systems at Extreme Scales (ADA-FS)

Andreas Knüpfer, Wolfgang E. Nagel
TU Dresden / Germany
E-Mail: {andreas.knuepfer,wolfgang.nagel}@tu-dresden.de

André Brinkmann
Johannes Gutenberg University Mainz / Germany
E-Mail: brinkman@uni-mainz.de

Achim Streit
Karlsruhe Institute of Technology / Germany
E-Mail: achim.streit@kit.edu

The ADA-FS project aims at improving the effective I/O performance of highest-scale parallel application programs on the next generation of Leadership-class supercomputers.
The envisioned architectures of future exascale supercomputers contain very "fat" compute nodes with much more processing elements resp. processing power per node than today's machines. So do actual announcements for upcoming 100 Petaflop/s machines. This will further reduce the ratio of node bandwidth to/from the outside vs. the number of processing elements which will heavily affect the performance of parallel file systems. The architectural solution to this will be node-local storage that works as burst buffers for the file system.
This leads to a new challenge on the software side. All parallel applications need to employ the local storage to be able to reach decent I/O performance. Yet, the additional complexity should not be addressed in an application's source code. Especially, since the local burst buffers will be machine specific with different technologies, speeds, and sizes.
The ADA-FS project plans to solve this software challenge with a newly designed overlay file system. This transparent overlay file system will be present on all compute nodes of a parallel job from before the job starts until after it finishes. The overlay file system can intercept normal I/O operations from the parallel application and modify them. This includes buffering input or output in node-local storage areas and thus transparently managing burst buffers for the parallel application.
On top of the central overly file system the ADA-FS project plans further research tasks, including the deployment of the overlay file system in a production HPC environment and the integration in job scheduling systems. Also, it will examine how a reduced POSIX I/O semantics may look like, which is general enough for most application uses cases but ignores some critical and expensive

POSIX guarantees in parallel scenarios. Furthermore, it will explore machine learning strategies to predict upcoming I/O behavior. Following that, it will investigate central planning to optimize the I/O behavior of concurrent and consecutive parallel jobs on the same supercomputer resp. on the same parallel file system.
The presentation will give an overview about the ADA-FS project which is a new project in the second phase of the SPPEXA program. The project will start in February 2016.

15.2 Advanced Computation and I/O Methods for Earth-System Simulations (AIMES)

Julian Kunkel
Deutsches Klimarechenzentrum, Bundesstr. 45a, 20146 Hamburg / Germany
E-Mail: kunkel@dkrz.de

With the Advanced Computation and I/O Methods for Earth-System Simulations (AIMES) project we will address the key issues of programmability, computational efficiency and I/O limitations that are common in next-generation icosahedral earth-system models. To enhance programmability and reach a higher-level of code design, we apply and advance concepts and tools for domain-specific languages on earth-system models. We will analyze suboptimal I/O performance delivered by the scientific file formats and push lossy data compression to reduce vast amount of data that is created with large scale simulations. Ultimately, with the project we intend to foster development of best-practices and useful norms by cooperating on shared ideas and components. During the project, we will ensure that the developed concepts and tools are not only applicable for earth-science but for other scientific domains as well.

15.3 ExaDG: A Matrix-free Multilevel Infrastructure for High Performance Discontinuous Galerkin Methods

Martin Kronbichler
Technische Universität München, Lehrstuhl für Numerische Mechanik, Boltzmannstr. 15, 85748 Garching / Germany
E-Mail: kronbichler@lnm.mw.tum.de

State-of-the-art finite-element software libraries focus on using sparse-matrix structures for their linear algebra because of their ease of use and the accessibility of highly efficient algebraic multigrid solvers. Nevertheless, the low computational intensity of less than one floating point operation per byte of memory loaded of the sparse matrix vector product and its consequences on modern computer architectures have not been addressed. In prototypical applications, tensor product based matrix-free codes with much higher computational intensity have proven superior on current architectures with high potential for the future. Therefore, we discuss the implementation of a completely matrix-free infrastructure for wide ranges of applications and efficient geometric multigrid solvers in the deal.II software library. This work encompasses efficient multithreaded implementation of local operations and embedding into the existing message passing environment.

15.4 MYX - MUST Correctness Checking of YML and XMP Programs

Matthias S. Müller
RWTH Aachen / Germany
E-Mail: mueller@itc.rwth-aachen.de

Serge Petiton, Nahid Emad
Maison de la Simulation / France
E-Mail: petiton@lifl.fr, emad@prism.uvsq.fr

Taisuke Boku
University of Tsukuba / Japan
E-Mail: taisuke@cs.tsukuba.ac.jp

Hitoshi Murai
RIKEN / Japan
E-Mail: h-murai@riken.jp

Exascale systems challenge the programmer to write multi-level parallel programs, which means employing multiple different paradigms to address each individual level of parallelism in the system. The long-term challenge is to evolve existing and develop new programming models to better support the application development on exascale machines. In the multi-level programming paradigm FP3C, users are able to express high-level parallelism in the YvetteML workflow language (YML) and employ parallel components written in the XcalableMP (XMP) paradigm. By developing correctness checking techniques for both paradigms, and by investigating the fundamental requirements to first design for and then verify the correctness of parallelization paradigms, MYX aims to combine the know-how and lessons learned of different areas to derive the input necessary to guide the development of future programming models and software engineering methods.
XMP is a PGAS language specified by Japans PC Cluster Consortium for high-level programming and the main research vehicle for Japans post-petascale programming model research targeting exascale. YML is used to describe the parallelism of an application at a very high level, in particular to couple complex applications YML provides a compiler to translate the YvetteML notation into XMP-parallel programs, and a just-in-time scheduler to manage the execution of parallel programs. The MUST correctness checker can detect a wide range of issues in MPI, OpenMP and hybrid MPI+OpenMP programs by collecting program information and aggregating this in a tree-based overlay network capable of running different types of analysis. Due to the use of the P^nMPI profiling interface, MUST can in principle trace and analyze any MPI communication either directly from the application code or any middleware library, such as the XMP runtime.

In MYX we will investigate the application of scalable correctness checking methods to YML, XMP and selected features of MPI. This will result in a clear guideline how to limit the risk to introduce errors and how to best express the parallelism to catch errors that for principle reasons can only be detected at runtime, as well as extended and scalable correctness checking methods.