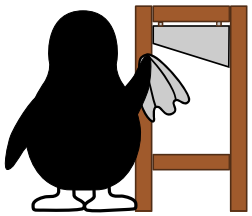
Quartl* - Impressum
Herausgeber:
Prof. Dr. A. Bode, Prof. Dr. H.-J. Bungartz, Prof. Dr. U. Rüde
Redaktion:
S. Herrmann, S. Reiz, Dr. S. Zimmer
Technische Universität München
School of Computation, Information and Technology
Boltzmannstr. 3, 85748 Garching b. München
Tel./Fax: ++49-89-289 18611 / 18607
e-mail: herrmasa@in.tum.de,
https://www.cs.cit.tum.de/sccs/startseite/
Redaktionsschluss für die nächste Ausgabe: **01.09.2024**

* Quartel: früheres bayerisches Flüssigkeitsmaß,
→ das Quart: 1/4 Kanne = 0.27 l
(Brockhaus Enzyklopädie 1972)

Quartl

2/2024
110. Ausgabe

Inhalt



Das Quartl erhalten Sie online unter https://www.cs.cit.tum.de/sccs/weiterfuehrende-informationen/quartl/

KONWIHR

Elitenetzwerk Bayern

Das Quartl ist das offizielle Mitteilungsblatt des *Kompetenznetzwerks für Technisch-Wissenschaftliches Hoch- und Höchstleistungsrechnen in Bayern* (KONWIHR) und der *Bavarian Graduate School of Computational Engineering* (BGCE)

Editorial

Viel ist in den vergangenen Wochen über die diversen im Netz kursierenden Videos von der Insel Sylt berichtet und diskutiert worden. Am ersten Freitag danach ereilte mich ein Anruf unseres Pressesprechers: auf einem der Videos sei wohl einer unserer Studierenden identifiziert worden, der Name kursiere bereits in den Social Media, und es mehrten sich Kommentare zwischen „Was gedenkt die TUM zu tun?“ und „Hängt ihn höher!“. Schnell war allen Verantwortlichen klar, dass man zunächst einmal die polizeilichen Ermittlungen abwarten müsse und werde, bevor irgendeine Maßnahme anstehe. Und so gab es eine unmissverständliche Ansage der TUM, wie die Universität mit ihrem Deutschland-weit einmaligen Reichtum an internationaler Vielfalt zu verabscheuungswürdigem Bockmist à la „Deutschland den Deutschen, Ausländer raus!“ stehe, ich informierte unsere School-Leitung sowie den Betreuer des betreffenden Studierenden, aber das war’s erst mal.

Bereits kurz darauf war in diversen Medien zu lesen, eine Studierende einer Hamburger Hochschule habe umgehend Hausverbot erhalten, und ein paar andere „Darsteller“ in den Videos seien sehr schnell gefeuert worden. Wow – das nenne ich schnelles und konsequentes Handeln. Bei uns gingen dagegen weiter Anfragen ein „Und was macht die TUM?“

Sorry, aber wo sind wir eigentlich? Gilt nicht in Deutschland nach wie vor die Unschuldsvermutung? Ist es nicht in unserem Rechtssystem eine bewährte Praxis, erst mal die zuständigen Behörden in der Angelegenheit ermitteln zu lassen und insbesondere alle irgendwie Beschuldigten zu befragen und Stellung nehmen zu lassen? Seit wann ersetzt eine Welle der Empörung bei Instagram & Co. eine polizeiliche Ermittlung oder gar ein gerichtliches Verfahren? Und warum ist jedes beliebige Taka-Tuka-Video plötzlich über jeden Zweifel erhaben? Man fühlt sich unweigerlich an so manchen B-Western erinnert, wo ein wütender Lynch-Mob einen vermeintlichen Pferdedieb als prophylaktische Maßnahme mal aufhängt. Was sich genau mit dem Gaul zugetragen hat, kann man dann ja noch später in Ruhe klären. Na bravo.

.

.

Ich kann ja die individuelle Empörung nachvollziehen – zugegebenermaßen schoss auch mir spontan ein „so jemand hat bei uns nichts verloren“ durch den Kopf. Ich kann auch noch nachvollziehen, dass in einer Zeit, wo viele offensichtlich jede kleinste Gefühlsregung öffentlich zur Schau stellen, man es hier nicht mit einem „dislike“ bewenden lassen will, sondern auch noch einen geharnischten (und wie immer wohl durchdachten ...) Kommentar dazu absetzt. Dass aber davon ausgegangen wird, dass man sich als irgendwie betroffene Institution solchen Blödsinn zu eigen macht; bzw. dass einzelne Institutionen dem Druck des digitalen Lynch-Mobs auch spontan nachgeben, geht für mich eindeutig zu weit.

Ein paar Tage später erreichte dann eine Nachricht des Betreffenden die TUM. Darin distanziert er sich in aller Deutlichkeit von besagtem Gegröle und liefert eine schon ziemlich andere Darstellung der Geschehnisse. Es könnte also komplizierter sein, als der erste Eindruck nahelegt. Auch das zeigt uns, dass es aus gutem Grund Ermittlungsbehörden und nicht nur Taka-Tuka-Blockwarte gibt!
Insofern werden wir jetzt das Ergebnis der Ermittlungen abwarten und dann entscheiden, ob und ggf. welche Maßnahmen als angemessen erscheinen. Denn eins ist sicher: Solchen dummen, diskriminierenden und einen guten Teil unserer Gesellschaft unmittelbar bedrohenden Parolen darf man keinen Raum gewähren, nicht den kleinsten!

In Sachen Lynch-Mob erfuhr ich dann ein paar Tage später bei einem Treffen der Graduate Deans der EuroTech-Universitäten eine weitere verstörende Nachricht. Mein Kollege vom Technion in Haifa berichtete, dass sich nach dem Hamas-Überfall am siebten Oktober am Technion spontan „studentische Gerichte“ gebildet hätten, vor denen Verfahren gegen arabische Mitstudierende durchgeführt wurden. Dazu muss man wissen, dass das Technion rund 30% arabische Studierende hat. Eine dreistellige Zahl von „Anklageschriften“ wurde dann der Hochschulleitung und der Staatsanwaltschaft übergeben, bei ersterer mit der Erwartung der sofortigen Exmatrikulation. Natürlich ist man diesem dreisten Ansinnen nicht nachgekommen. Bei den anschließenden Ermittlungen der Staatsanwaltschaft ergab sich übrigens in lediglich zwei

(!) Fällen ein erhärteter Anfangsverdacht. Alle anderen Verfahren wurden zurückgewiesen bzw. umgehend eingestellt.
Auch hier sind Überreaktionen, Rachegelüste etc. individuell verständlich, nach einem derart brutalen Überfall und in einer solch akuten Bedrohungslage. Aber ein kollektives Vorgehen mit standrechtlichen Schauprozessen geht gar nicht. Das erinnert an finsterste Exzesse aufgebrachter Sansculotten in der Hochphase der französischen Revolution. Damals wurde auch zuerst die Guillotine angeworfen und anschließend (vielleicht) nachgedacht oder ermittelt. Kleine Anmerkung am Rande: Die Guillotine ist übrigens nach einem Arzt benannt ...

Doch nun wünscht Ihnen die gesamte Quartl-Redaktion einen unbeschwerten und hoffentlich nicht zu nassen, nicht zu trockenen, nicht zu kühlen, aber auch nicht zu heißen Sommer, und zunächst natürlich viel Spaß mit der neuesten Ausgabe Ihres Quartls!

Hans-Joachim Bungartz.

* Notiz * Notiz * Notiz *

Termine 2024

- **Upcoming SIAM Conferences & Deadlines**
  `https://www.siam.org/conferences/calendar`
- **Supercomputing 2024:**
  The International Conference for High Performance Computing, Networking, Storage and Analysis (SC23) – SC 24 in Atlanta, GA, USA: 17.11.-22.11.2023 `https://sc24.supercomputing.org/`
- **KONWIHR News** `https://www.konwihr.de/`

Figure 3: Juniors discussing their deliverables and expectations ; © Michael Wiedenmann

As an improvement from last year, we were pleasantly surprised by the large number of excellent candidates and we fully enjoyed the interaction of the Opening Weekend with the new intake. We are very much looking forward to working with before-last batch of BGCE students (BGCE will formally stop in March 2027 due to the regulations of the ENB) for another eventful BGCE year.

Qunsheng “Keefe” Huang, Tobias Neckel

Iterationsschleife

N=50
06.06.2024

In der Nähe der chinesischen Stadt Xian steht – vergraben in der Erde – eine tönerne Armee. Ihre Aufgabe ist es, den toten Kaiser Qin Shi Huang Di[a] zu beschützen. Qin Shi Huang Di gilt als der erste historisch belegte chinesische Kaiser und Begründer der Qin Dynastie. Sein Grabmal befindet sich westlich hinter der Terra-Cotta-Armee. Die Armee liegt östlich des Grabmals, sodass sie den Kaiser vor seinen aus dem Osten angreifenden Feinden beschützen kann.

Mit der Erstellung der Terra Cotta Armee begann der Kaiser bereits bei seiner Machtergreifung als König von Qin im Jahr 246 v. Chr. mit 13 Jahren seinem Vater auf den Thron des Reiches Qin nachfolgte. Die Arbeiten dauerten 38 Jahre und wurden etwa 2 Jahre nach dem Tod des Herrschers beendet.
Jeder einzelne Angehörige der Armee wurde nach einem realen Vorbild modeliert. Jedes Gesicht ist individuell, jeder Körperbau ist individuell, jede Kleidung ist individuell und weist darauf hin, aus welchem Teil des Reiches der abgebildete Soldat kam. An den Haarknoten der Soldaten erkennt man, ob sie Linkshänder oder Rechtshänder waren. Dieser Kaiser wollte also im Tod nicht von irgendwelchen Figuren, sondern von den Abbildern seiner realen Armee geschützt werden.
Eine Kommandozentrale für die Armee wurde ebenfalls gefunden und ausgegraben. Sie enthält alle notwendigen Einrichtungen für eine Kommandozentrale sowie eine Reihe von Offizieren und Soldaten. Ein Oberkommandierender hingegen fehlt völlig. Die Erklärung der Historiker für das Fehlen des Oberkommandanten ist klar: der Kaiser wollte keinen ständigen Oberbefehlshaber haben und hat daher für jeden Feldzug einen anderen General – und nur für diesen Feldzug – zum Oberkommandierenden ernannt. So wollte er es wohl auch über den Tod hinaus handhaben. Schon 210 v. Chr. wollte ein alter Herrscher über den Tod hinaus, alle Zügel in der Hand behalten. Eine Tradition die wir heute noch immer fortsetzen.

M. Resch

[a] Si Huang Di wurde wohl 259 v. Chr. geboren und starb am 10. September 210 v. Chr. Er gilt als erster Reichseiniger der historisch belegt ist.

Numerische Modellierung der kardialen Elektrophysiologie auf zellulärer Ebene

Die Fußball-EM zeigt es wieder einmal deutlich: Das menschliche Herz schlägt in der Regel ständig, mitunter besonders heftig. Statistisch signifikant wird das Herz offenbar bei solchen Ereignissen wie der EM auch aus dem Rhythmus gebracht. Dabei ist der dauerhafte regelmäßige Herzschlag schon nicht leicht zu verstehen, noch weniger, welche Faktoren den Herzschlag aus dem Rhythmus bringen können oder sogar im schlimmsten Fall zum Stillstand führen.

Genau wie technische Geräte nutzen biologische Zellen Elektrizität für eine schnelle Kommunikation. Herzmuskelzellen erzeugen elektrische Ströme, um die Aktion intern zu koordinieren und mit ihren Nachbarzellen zu kommunizieren, damit der gesamte Herzmuskel als Einheit arbeiten kann. Dieses Koordinierungssystem beruht auf leitenden Verbindungen zwischen den Zellen, die durch strukturelle Muskelerkrankungen beeinträchtigt werden können, häufig mit Herzrhythmusstörungen als erstem Symptom. Herzrhythmusstörungen sind für etwa 15% aller Todesfälle verantwortlich, ein großer Teil der Ursachen davon sind ungeklärt. Die Verhinderung dieser Todesfälle ist eine große gesellschaftliche Herausforderung. Um sie zu verstehen, benötigen wir einen detaillierten Einblick in die Elektrophysiologie des Herzens. Wir wollen insbesondere erkennen, welchen Einfluss einzelne, durch Alterung oder Krankheit beschädigte Zellen auf den Herzschlag haben.

Heutige Computermodelle betrachten das Herz als homogenisierten Raum, in dem an jedem Punkt, sowohl Zellen als auch umgebender extrazellulärer Raum, die entsprechenden elektromagnetischen Größen existieren. So kann das gesamte Herz mit einigen Millionen Bildpunkten, welche jeweils hunderte Zellen zusammenfassen, dargestellt und der Herzschlag simuliert werden. Solche Modelle bilden die Funktionalität des Herzen auf makroskopischer Ebene akkurat ab und sind ein wichtiges Forschungswerkzeug. Zum Verständnis struktureller Herzerkrankungen könnten jedoch geschädigten

After a informal introduction to the roots of the BGCE program by our very own Tobias Neckel, the new BGCE batch attended the Soft Skill Seminar “When teamwork works”, lead by the two coordinators Jan Seydel and Michael Wiedenmann.
Later that evening, we enjoyed an in-depth exploration of the future prospects of algorithmic improvement of Machine Learning algorithms in a talk by Prof. Hans-Joachim Bungartz in the traditional “Kaminabend”.

Figure 2: Seniors prepare to “Step-Out” ; © Jan Seydel

On Saturday, the BGCE students took part in the Soft Skills courses “When Teamwork Works” and “Step Out”, designed to foster some Esprit de corps and teach teamwork skills. While the juniors focused on fundamental teamwork, the seniors were tasked with execution and were able to produce completely original music video.
The final day of the Opening Weekend took place on Sunday the 13th of April. There, the juniors and seniors worked through the “Group Challenge” activity together, where they were presented with a complicated task to complete together in a minimal stint of time.
This finally culminated in the final “Consulting Circle’, where the seniors could share their gathered wisdom to the juniors in a guided, informal session.

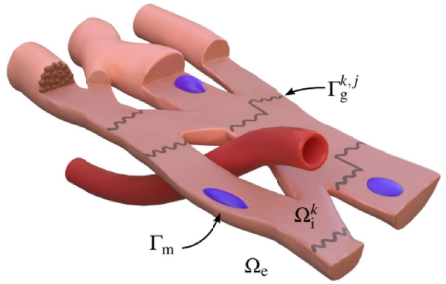
BGCE: Opening Weekend

As a tradition, the BGCE family welcomes the newest cohort of the Bavarian Graduate School of Computational Engineering (BGCE) program in its yearly spring gathering. This year, we see a large intake of 19 new juniors that started in the Winter Semester 2023/24!

Figure 1: Participants of the Bavarian Graduate School of Computational Engineering pose triumphantly after a successful team-building activity; © Michael Wiedenmann

Verbindungen zwischen einzelnen Zellen ausschlaggebend sein - eine vieldiskutierte Ursache tödlicher Herzrhythmusstörungen. Deshalb benötigen wir den Schritt von homogenisierten Computermodelle auf makroskopischer Ebene, die räumliche Skalen weit jenseits der biologischen Zellgröße betrachten, zu einem Modell auf mikroskopischer Ebene, das einzelne Zellen detailliert in Betracht ziehen kann.

Ein solches Modell ist das Extracellular, Membrane and Intracellular (EMI) Modell, das statt mit hunderten Zellen pro Bildpunkt mit tausenden Bildpunkte pro Zelle arbeitet. In jedem Bildpunkt existiert hier entweder intrazellulärer Raum, extrazellulärer Raum oder Zellmembran. Simulationen mit diesem Modell sind ungefähr um den Faktor 10,000 feingranularer als mit einem homogenisierten Modell und benötigen dementsprechend große Datenmengen und einen enormen Rechenaufwand. Da wir nun nicht mehr homogenisierten Raum betrachten, stellen uns außerdem die detaillierte Darstellung von intrazellulären und extrazellulären Potentialdifferenzen über die Zellmembran (transmembrane Voltages), zwischen den Zellen (Gap Junctions) und in den Ionenkanälen durch die Zellmembran vor numerische Herausforderungen (Abb. 1). Um diese Herausforderungen zu meistern, benötigen wir numerische Verfahren und ein skalierbares Software-Ökosystem, um hochparallele Computer effizient zu nutzen.



$$
\begin{cases}
\nabla \cdot (G_i^k \nabla \phi_i^k) = 0, & \text{on } \Omega_i^k, \\
\nabla \cdot (G_e \nabla \phi_e) = 0, & \text{on } \Omega_e, \\
V_m^k = \phi_i^k - \phi_e, & \text{on } \Gamma_m^k, \quad \text{(Transmem. voltages)} \\
-G_i^k \nabla \phi_i^k \cdot \mathbf{n} = G_e \nabla \phi_e \cdot \mathbf{n} = C_m \partial_t V_m^k + I_{ion}(V_m^k, \mathbf{y}), & \text{on } \Gamma_m^k, \quad \text{(Transmem. currents)} \\
-G_i^k \nabla \phi_i^k \cdot \mathbf{n} = G_i^j \nabla \phi_i^j \cdot \mathbf{n} = \kappa(\phi_i^k - \phi_i^j), & \text{on } \Gamma_g^{k,j}, \quad \text{(Gap junctions)} \\
\partial_t \mathbf{y} = F(V_m, \mathbf{y}), & \text{on } \Gamma_m^k
\end{cases}
$$

Abbildung 1: Schematische Darstellung einiger Zellen im EMI-Modell mit allen in Betracht gezogenen Größen.

MICROCARD ist ein europäisches Forschungsprojekt, in dem ein Team aus Biomedizintechnik-Ingenieuren, Mathematikern und Informatikern eine ausgereifte Plattform zur Simulation der Elektrophysiologie des menschlichen Herzens im Mikrometermaßstab entwickelt. Software-seitig werden die Simulationen im openCARP Software-Stack realisiert, der vom Karlsruher Institut für Technologie und der Universität Freiburg entwickelt wird. Um Simulationen mit bis zu 100,000,000 Herzzellen zu ermöglichen, wird die Ginkgo Open-Source Bibliothek verwendet, die von Prof. Hartwig Anzt und seinem Team entwickelt wird. Ginkgo verfügt über ein großes Spektrum numerischer Verfahren, die sowohl auf Multicore-CPUs als auch GPU-Architekturen von AMD, Intel und NVIDIA ausgeführt werden können. In der Vergangenheit hat die Verwendung von Ginkgo im Rahmen des US Exascale Computing

- *GPU Performance and Feature Enhancement of the Earthquake Cycle Simulation Software Tandem*
  by Prof. Dr. habil. Alice-Agnes Gabriel
  Department of Earth and Environmental Sciences, Institute of Geophysic, Group of Earthquake Physics, LMU

- *Compute Cloud access solutions for FAIR re-use of HPC research data*
  by Prof. Christian Stemmer
  Chair of Aerodynamics and Fluid Mechanics, TUM

You can find more details about these projects at

`https://www.konwihr.de/konwihr-projects/`

KONWIHR: New projects from round 2024-1

The competence network for scientific high-performance computing in Bavaria welcomes the new projects that succeeded in the application round 2024-1. In every round, we accept proposals for “normal” (up to 12 months) and “small” (up to 3 months) projects, as well as “basis” projects to establish contact points. In this round, the following projects were funded:

- *Optimization of Material Evaluation in a Radiance Transport Simulation*
  by Prof. Dr.-Ing. Kai Selgrad,
  Computer Science and Mathematics / Computer Graphics, Ostbayerische Technische Hochschule (OTH) Regensburg

- *Optimization of LuKARS*
  by Prof. Dr. habil. Gabriele Chiogna,
  Applied Geology and Modeling of Environmental Systems, FAU

- *LexicoLLM – Leveraging Large Language Models for Lexicography*
  by Prof. Dr. Peter Uhrig,
  Digital Linguistics with a focus on Big Data, FAU

- *Dynamics of Complex Fluids*
  by Prof. Dr. Jens Harting,
  Department of Chemical and Biological Engineering and Department of Physics, FAU
  and Helmholtz Institute Erlangen-Nürnberg for Renewable Energy

Projektes bereits die Skalierbarkeit auf bis zu 16,000 GPUs demonstriert, jetzt werden numerische Verfahren speziell auf die Verwendung im MICROCARD Projekt zugeschnitten.

Die einzelnen Zellen bieten im EMI-Modell gewissermaßen natürliche Teilgebiete oder Subdomains, die sich mit Hilfe einer Domain-Decomposition-Methode separat voneinander betrachtet werden können. Im MICROCARD Projekt haben wir uns für die Verwendung von Balanced Domain Decomposition by Constraints (BDDC) als Vorkonditionierer entschieden [1,2]. Bei diesem Verfahren werden Subdomains unabhängig voneinander gelöst und die lokalen Lösungen über sogenannte Constraints mit einem kleinen, globalen linearen System aneinander gekoppelt (Abb. 2). Constraints können die Gleichheit der lokalen Lösungen in einem Bildpunkt erzwingen (Stetigkeits-Constraints) oder aber schwächere Anforderungen stellen, wie zum Beispiel, dass der Durchschnitt über alle Bildpunkte auf der Grenze zwischen zwei Subdomains gleich ist. Würden wir überall Gleichheit fordern, wäre BDDC ein exakter Löser und sehr teuer aufzustellen, mit Durchschnitts-Constraints können wir die Lösung annähern und die Kosten zum Aufstellen des Vorkonditionierers niedriger halten.

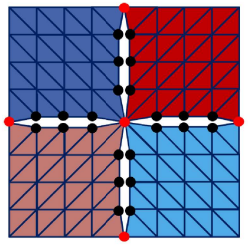


Abbildung 2: In BDDC betrachten wir Subdomains unabhängig voneinander und koppeln die lokalen Lösungen über sogenannte Constraints (rote und schwarze Punkte) miteinander. Hier stellen rote Punkte Stetigkeits-Constraints dar und schwarze Punkte schwächere Durschnitts-Constraints.

In Wirklichkeit sehen Zellen natürlich nicht so schön gleich und quadratisch aus wie in Abb. 2. Eine realistischere Darstellung gibt Abb. 3. Hier sind 10 Zellen und ihre umgebenden extrazellulären Subdomains dargestellt. Die Größe dieser Subdomains variiert enorm: die größten extrazellulären Subdomains sind bis zu 10-mal größer als die kleinste Zelle, und auch zwischen den Zellen variiert die Größe bis zu einem Faktor von 10. Das führt bei parallelen Berechnungen zu Load-Balancing Problemen, d.h. während ein Prozessor an einer großen Subdomain arbeitet, sind Prozessoren, die für kleinere verantwortlich waren, schnell fertig und warten für eine lange Zeit ohne zu arbeiten. Um diese Ineffizienz zu vermeiden, verwenden wir eine Implementierung, die auf Tasking basiert. Hier werden Arbeitspakete (Tasks) in einen Taskpool gegeben, von dem sich Prozessoren, die gerade nichts zu tun haben, Arbeit holen können. Für das Beispiel mit 20 Subdomains aus Abb. 3 kann so die gleiche Arbeit in lediglich 9% längerer Zeitdauer mit 4 statt 20 Prozessoren erledigt werden – eine enorme Ressourcenersparnis.

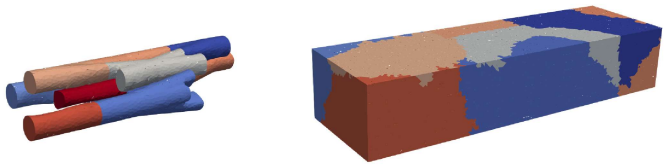


Abbildung 3: 10 Zellen und ihre umgebenden extrazellulären Subdomains.

KONWIHR: Apply now!

Do you want to optimize your code, but you need some help (and funding)? The competence network for scientific high-performance computing in Bavaria accepts applications for scientific HPC projects (up to 50 000 €), as well as for "basis" projects to establish contact points. **Tip:** Increase the success of your proposal and your project by contacting your local computing center before applying, to discover optimization potential in your code.

The main objective of KONWIHR is to provide technical support for the use of high-performance computers and to expand their deployment potential through research and development projects. Close cooperation between disciplines, users, and participating computer centers as well as efficient transfer and fast application of the results are important.
Find already funded projects at

`https://www.konwihr.de/konwihr-projects/`

Read more on konwihr.de and join the konwihr-announcements mailing list.

Contact KONWIHR

For any KONWIHR inquiries, you only need one address:

info@konwihr.de

Your email will be read carefully and answered by your contact people in the Bavarian North and South. Together with Prof. Gerhard Wellein and Prof. Hans-Joachim Bungartz (who you can also reach using the same address), we collect and process your proposals two times per year (1st of March and 1st of September). Learn more about how you can apply for funding at:

`https://www.konwihr.de/how-to-apply/`

LRZ-Insight: Dr. Margarita Egelhofer

Der Traum von einem Beruf wird oft schon in jüngsten Jahren gelegt – durch Geschichten und durch Spiele: Margarita Egelhofers Vater nahm gerne als Alien am Telefon Kontakt mit ihr auf und weckte so in seiner Tochter die Leidenschaft für unendliche Weiten, grundsätzliche Fragen und zur Astrophysik. Zwar rückte die Wissenschaftlerin von ihrem Wunsch ab, Astronautin zu werden, aber sie erforschte intensiv das All und seine Universen am Max-Planck-Institut in Garching und an der Universität Bologna. Inzwischen unterstützt sie Forschende am Leibniz-Rechenzentrum (LRZ) und hilft ihnen, Codes zu optimieren und auf Supercomputer zu implementieren. So will die Astrophysikerin Wissenschaft und Erkenntnis beschleunigen: „Als Betreuerin von Forschenden“, sagt sie, „schätze ich es sehr, mit spannenden Menschen zusammenarbeiten zu dürfen, so bekomme ich auch noch die neuesten Erkenntnisse aus unterschiedlichsten Forschungsbereichen mit.“ Mehr über Dr. Margarita Egelhofer und ihre Arbeit im Computational X Support (CXS) lesen Sie hier `https://www.lrz.de/presse/ereignisse/2024-04-28-DrMEgelhofer-dt/`.

Susanne Vieser

Im September 2024 wird das MICROCARD Projekt enden, die Forschung aber im Rahmen eines neu gegründeten „EuroHPC Center of Excellence“ weitergeführt. Das Herz von EuroHPC wird dann auch auf dem Standort TUM Campus Heilbronn schlagen – hoffentlich ohne Rhytmusstörungen.

Hartwig Anzt

Literatur

[1] `https://epubs.siam.org/doi/10.1137/S1064827502412887`

[2] `https://arxiv.org/pdf/2212.12295`

Durham HPC/AI Days 2024

The week before ISC, i.e. from 7 May 2024 to 10 May 2024, we hosted the Durham HPC/AI Days 2024. The location of choice might not come as a surprise: Durham. A similar event was on the year before, and it was successful and well-attended. Nevertheless, the interest this time blew us away: More than 170 participants joined us for the four days, contributing towards two parallel sessions. On Tuesday, we kicked off with two tutorials—one led by NVIDIA showcasing their new Grace Hopper nodes that we recently installed in Durham, and one led by colleagues from Cerebras who made their AI-focused chip available to the participants—before the afternoon provided a platform to the ExCALIBUR projects to showcase their insights. ExCALIBUR is the UK's exascale software program and a particular focus of the Tuesday session was HPC in combination with AI. Parallel to the ExCALIBUR talks, the UK's Lustre User Group did meet. The evening closed with beer (not stale and not warm) and fish 'n' chips. Local cuisine.

On Wednesday, DiRAC, the theory computing community of the UK, met, while the main paper track covered a wide variety of topics ranging from workflows, benchmarking to the retaining and identification of talent. We also saw a two-hour session on SYCL. A highlight of the day—besides proper pizza from a proper wood-fired oven, wine and posters in the evening—certainly was the two keynotes: Hatem Lteief discussed energy-aware mixed precision computing (and gave a preview of their Gordon Bell submission this year), while Luigi del Debbio provided an overview over EuroHPC JU, the access workflows and the available resources; a timely talk given that the UK re-joined EuroHPC a week later.

Eine optimale Umgebung für Container finden

Hochleistungsrechnen ist ohne Container nicht mehr denkbar. Diese digitalen Boxen enthalten alle Werkzeuge, um einen wissenschaftlichen Code auszuführen, sie bringen aber nach Erkenntnissen am Leibniz-Rechenzentrum (LRZ) nicht auf allen Computersystemen die gleiche Leistung. LRZ-Forscher Max Höb geht für seine Doktorarbeit den Gründen nach: „Ich arbeite an einem Modell und einer Methode, dieses zu evaluieren. Konkret heißt das, ich bestimme und vergleiche unterschiedlichste Leistungsmerkmale von Containern", erklärt Höb. „Die Anzahl an heterogenen Computerarchitekturen wächst und stellt uns vor die Frage, auf welchem System eine Applikation am besten ausgeführt werden sollte, ob die zur Verfügung gestellten Ressourcen tatsächlich benötigt werden und ob die Nutzung optimiert werden kann. Daher können auch Container nicht auf jedem Hochleistungsrechner die gleichen Leistungen erbringen. Mit seiner Arbeit will Höb Kennzahlen und Argumente zusammentragen, damit Rechenzentren leichter entscheiden können, welche Containertechnologie am besten zu ihren Ressourcen passt `https://www.lrz.de/presse/ereignisse/2024-03-25-Container-im-HPC/`.

Susanne Vieser

Probezeit eines Supercomputers

Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften

Höchstleistungsrechner werden eigens für spezielle Anforderungen aufgebaut und sind daher Unikate. Vor dem eigentlichen Betriebsstart stehen daher umfangreiche Experimente und Funktionstests, vor allem aber Teamarbeit: Vor dem Betriebsstart von SuperMUC-NG Phase 2 – dem Ergänzungssystem für den LRZ-Supercomputer, in dem neben CPU erstmals Graphics Processing Units (GPU) installiert wurden – arbeiteten Forschende und LRZ-Spezialistinnen Hand in Hand mit Technologie-Unternehmen zusammen. Sie testeten nicht nur Funktionsweisen, sondern passten wissenschaftliche Codes an das beschleunigte System an und implementierten außerdem erste Modelle für statistische Anwendungen der Künstlichen Intelligenz (KI). Einen Bericht über die Pilotphase sowie die technischen Details von SuperMUC-NG Phase 2 finden Sie hier `https://www.lrz.de/presse/ereignisse/2024-05-06-SNG-2-Pilotphase/`.

Susanne Vieser

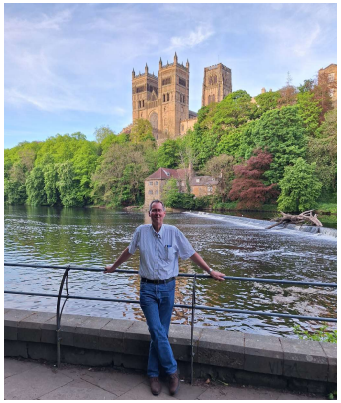


Figure 1: Hans poses in front of Durham Cathedral down at the River Wear. That's by the way also Durham's graduation hall.

The Friday kicked off with a talk by Alison Kennedy who chairs the UK exascale procurement panel and closed with a keynote by Johannes Doerfert. Alison highlighted how ethically problematic it has become to drive supercomputers through solar panels if these are predominantly manufactured through slave labour, while Johannes took the audience on a tour-de-force through the latest LLVM developments. In-between, we had several talks orbiting around aspects of green HPC in data centres. Without a doubt, the talk by Thomas Gruber introducing Cluster Cockpit was a highlight and sparked some follow-up discussions at ISC one week later.

So what happened on Thursday? While the majority of participants had been from the UK, this day was all shaped by international guests and collaborations. In the morning, Thomas Hauser gave a keynote presenting the multifaceted research done at NCAR in the US. After that, a two hour session featured various spotlight talks from colleagues who participated in Durham's professional HPC training. In the AY 23/24, we ran training on C/C++, cor-

Figure 2: Evening lecture on Thursday: Not everybody might have been happy about the analysis of the status-quo—who doesn't like their A100 under their desk—but a lot of people had a lot of fun and food for thought.

rectness and debugging, GPU programming and performance profiling. Our friends from RWTH Aachen (Joachim Jenke) and the VI-HPS/POP gang around Brian Wylie had been absolutely instrumental to make this happen. The feedback was all very positive though the spotlight talks provided food for thought how we can make these courses more accessible and even more useful to colleagues.

After an afternoon session organised by Women in HPC (WHPC), we offered participants the opportunity to visit Durham Castle. Durham Castle is usually not open to the public, as it also serves as our oldest college, i.e. provides accommodation to students. So this was a unique opportunity to see the old Norman Chapel, the extraordinary architecture of the late medieval rooms and notably the large dining hall—which, according to an Urban legend, was inspirational to the dining hall in Harry Potter but turned out to be too expensive for the studio to hire over a longer period. They therefore rebuilt it in the studio (we have no shortage of other places where Harry Potter

„weltliche“ (non-HPC) Anwendungen bis hin zu hervorragender Speicherbandbreite und Write-Allocate Evasion. Vor allem im Zusammenspiel mit GPUs aus dem eigenen Hause, welche über eine blitzschnelle Anbindung inklusiver geteiltem Addressraum verfügen, hat Intel es schwer, hier mitzuhalten oder gar innovativ voranzugehen. Nichtsdestotrotz zeigt sich, auch geschuldet dem noch in den Kinderschuhen steckenden Softwareumfeld und dem bisher überschaubarem Interesse an Optimierungsarbeit für Arm-Chips, dass der Sapphire Rapids vor allem im HPC-Bereich mit Vektorisierung und reiner single-Core-Performance punkten kann.
Wir sind gespannt, in welche Richtung sich die HPC-Community entwickeln wird und ob sich die „Alte Dame“ Intel von den „New Kids On The Block“ vertreiben lässt.

Jan Laukemann, Georg Hager

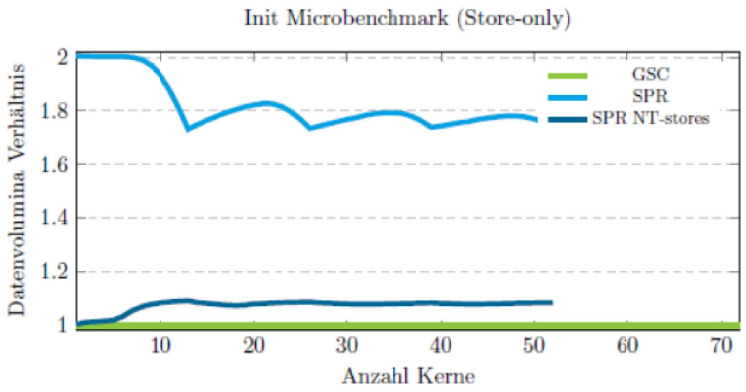


Abbildung 1: Store-only Benchmark auf je einem Sockel von GSC (72 Kerne) und SPR (52 Kerne). Die Y-Achse zeigt das Verhältnis von gemessenem Datentransfer vs. erwartetem Datentransfer (d.h. nur Speicherdatenvolumen).

zwar keinen Automatismus, dafür aber eine Instruktion, die eine Cachezeile explizit im Cache allokiert statt sie zu laden. Wer mehr über dieses faszinierende aber zugegeben nerdige Thema wissen möchte, dem sei unser kürzlich erschienenes IPDPS-Paper[1] ans Herz gelegt. Intel-Dokumentation dazu ist leider praktisch nicht vorhanden, aber ein schon etwas älterer Blog-Post[2] führt die Basics vor.

Zusammenfassend kann man sagen, dass der neue NVIDIA Grace Superchip viele Dinge richtig macht, von Energieeffizienz über skalare Performance für

[1] J. Laukemann, T. Gruber, G. Hager, D. Oryspayev, and G. Wellein: *CloverLeaf on Intel Multi-Core CPUs: A Case Study in Write-Allocate Evasion*. Accepted for publication at IPDPS 2024, the 38th IEEE International Parallel & Distributed Processing Symposium. Preprint: arXiv:2311.04797, DOI: 10.1109/IPDPS57955.2024.00038

[2] `https://blogs.fau.de/hager/archives/8997`

eventually had been shot, so not too annoying for us in Durham).
After this nice social event, we returned to the hard business. A few glasses of wine, and Hans started his evening speech explaining in great detail why every newly appointed professor nowadays insists on their own GPU cluster, who manages these clusters (John Smith, a long-term research group associate who has nothing else to do), and what this means both for the ethics of this area (I still wonder who pays the energy bill), the self-understanding of the involved colleagues, and the long-term evolution of this kind of research. It was very interesting to see a little bit of history from the early CFD and FEM era repeating.

You can get an impression of the event from https://scicomp.webspace.durham.ac.uk/events/code_performance_series/durham-hpc-days-2024/, and you can already block your calendar for 2025. We plan to repeat this event again in the week pre-ISC. I cannot disclose big news yet (wait for the next Quartl maybe), but it will be grand. And Durham is always worth a visit!

Tobias Weinzierl

Cholesterin macht Zellmembranen gleichzeitig flexibler und stabiler

Membranen, bestehend aus einer Lipid-Doppelschicht, spielen eine sehr wichtige Rolle als Barriere zwischen dem Zellinneren und der äußeren Umgebung. Sie beherbergen eine ganze Reihe von Proteinen, die einen kontrollierten Austausch von Molekülen (z.B. Salze, Nährstoffe etc.), eine Weiterleitung von Signalen und vieles mehr regulieren. Es ist schon seit längerem bekannt, dass kleine Moleküle die Eigenschaften von Membranen beeinflussen können. Cholesterin gehört ebenfalls zu dieser Gruppe von Molekülen, nimmt jedoch mit einem Anteil von bis zu 40% in der äußeren Zellmembran von Säugetieren anscheinend eine essenzielle Rolle für die Struktur und Stabilität von biologischen Membranen ein. In früheren Experimenten und Simulationen konnte gezeigt werden, dass Cholesterin Membranen dicker und damit auch undurchlässiger und stabiler macht. Wie sich dagegen die Flexibilität von Membranen durch Zugabe von Cholesterin verändert, wurde in den vergangenen 2–3 Jahren sehr kontrovers diskutiert.

Mittels Moleküldynamik-Simulationen von sehr großen Membranstücken (bis zu 10^4 nm^2) und Lipidbizellen (Frisbee-ähnliche Membransysteme, die frei fluktuieren können) sollte ein kohärentes Bild des Einflusses von Cholesterin auf die strukturellen und insbesondere die mechanischen Eigenschaften von Membranen entwickelt werden. Um die für die Berechnung der Elastizität relevanten Zeit- und Längenskalen zu erreichen, wurde ein vergröbertes bzw. coarse-grained Kraftfeld (MARTINI) verwendet, in dem Gruppen von Atomen zu sogenannten „Superbeads“zusammengefasst werden.
Alle Simulationen wurden auf den beiden Clustern des Zentrums für Nationales Hochleistungsrechnen Erlangen (NHR@FAU) mithilfe des Software-Pakets GROMACS durchgeführt. Kleinere oder mittelgroße Systeme wurden auf dem GPU-Cluster Alex gerechnet, da GROMACS bereits auf einer NVIDIA-GPU (hier A40) eine ausreichend hohe Performance erzielt. Für das größte System (ca. 1.7×10^6 Partikel) wurde auf das CPU-Cluster Fritz

beim Marktführer massiv ist, beim GSC aber nur marginal. Wem die reine Energieeffizienz am TDP-Limit wichtig ist, der kommt mit Grace (15 Gflop/W) deutlich besser weg als mit Sapphire Rapids (10 Gflop/W); beide werden da natürlich von jeder aktuellen GPU — egal ob vom Typ „Country“ oder „Western“ — in den Schatten gestellt.

Write-Allocate Evasion
Ein lange bekanntes „Problem“ bei Intel-Chips ist das nötige Write-Allocate (WA, auch Read-for-Ownership genannt) wenn Daten, die nicht schon im Cache liegen, in den Speicher geschrieben werden. Vor der Generation Ice Lake benötigten Intel-CPUs also einen Lesezugriff vor dem Schreiben, denn die Kerne können nur mit dem Cache reden, nicht mit dem Hauptspeicher direkt — auch wenn die Cachezeile später nicht wieder benutzt wird. Solange der Programmierer nicht mitdenkt und dies mit sogenannten „Non-temporal Stores“ umgeht, verliert man so bis zu 50% der Bandbreite durch unnötige Datentransfers. Seit Ice Lake geht Intel dieses Problem jedoch an und führte mit „SpecI2M“ ein hardwareseitiges Feature ein, das kontinuierliche Schreibzugriffe bei ausreichender Nutzung der Speicherbandbreite erkennt und das Write-Allocate vermeiden kann, was wir als *Write-Allocate Evasion* bezeichnen.

Das Verhältnis von tatsächlichem Datenvolumen zu erwartetem Datenvolumen bei einem „Init“-Benchmark (reines Schreiben eines großen Arrays) ist in Abb. 1 zu sehen (Mit 1 als bestmögliches Verhältnis und 2 als schlechtestes). SPR zeigt hier — zumindest für diesen spezifischen Fall — eher eine Verschlechterung zum Vorgänger und kann bei Weitem nicht mit den „traditionellen“ aber eben nicht automatischen NT-Stores mithalten. Vor allem bei nicht-vollen NUMA Domänen sinkt die Effizienz von SpecI2M. Noch krasser wird der Unterschied zum GSC, der keinerlei Write-Allocates verbucht und es somit einfach macht, das optimale (= minimale) Datenvolumen zu erreichen. Das Ganze ist übrigens gar nicht neu: Arm-Chips kennen das Feature schon lange und wer jemals mit dem Marvell ThunderX2 experimentiert hat, sollte auch im HPC-Umfeld darüber gestolpert sein. Der Fujitsu A64FX hat

sierende Codes gebaut ist, sondern eben auch für generische Workloads, für die SIMD ein Buch mit sieben Siegeln ist und wo der Compiler eben nicht alle Register ziehen kann (pun intended).

	GSC (Neoverse V2)	SPR (Golden Cove)
Anzahl Ports	17	12
Max. SIMD-Breite	16 B	64 B
Integers	6 (2 multi-cycle + 4 single-cycle)	5
FP-Vektor-Einheiten	4	3
Loads/cy	3 × 128 B	2 × 256 B or 2 × 512 B
Stores/cy	2 × 128 B	2 × 256 B

Tabelle 1: Vergleich einiger wichtiger Parameter der GSC- und SPR-Cores

	GSC	SPR (Xeon Platinum 8470)
Kerne	72	52
Frequenz (max. / base)	3.4 GHz / 3.4 GHz	3.8 GHz / 2.0 GHz
Theor. DP Peak	3.92 Tflop/s	6.32 Tflop/s
Prakt. DP Peak	3.82 Tflop/s	3.49 Tflop/s
Max. Stromaufnahme	250 W	350 W
Cachegröße (L1/L2/L3)	64 KB / 1 MB / 234 MB	48 KB / 2 MB / 105 MB
Hauptspeicher	240 GB LPDDR5X	512 GB DDR5
ccNUMA-Domänen	1	1, 2 oder 4 (SNC-mode)
Max. Speicherbandbreite (theor. / messbar)	546 GB/s / 467 GB/s	307 GB/s / 273 GB/s

Tabelle 2: Vergleich einiger wichtiger Parameter der GSC- und SPR-Chips. Bei beiden ist der L3-Cache von allen Kernen geteilt, L1 und L2 sind jedoch privat pro Core.

Tabelle 2 zieht den Vergleich Sockel für Sockel. Mehr Cores (72 statt 52), weniger Verlustleistung und die trotzdem um 30% höhere Speicherbandbreite sprechen alle für Grace, da hilft auch die etwas höhere Turbofrequenz bei Intel nichts, vor allem weil der Taktratenverlust bei Nutzung viele Kerne

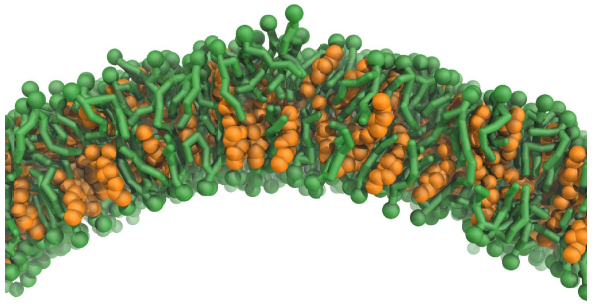


Abbildung 1: Die Abbildung zeigt ein Computermodell einer ungesättigten Membran. Diese besteht aus einer Lipid-Doppelschicht. Beim Verbiegen der Membran wandert das Cholesterin (orange) in die innere, negativ gekrümmte Lipidschicht. Diese Beweglichkeit von Cholesterin führt dazu, dass die Membran weicher wird, sich also leichter verbiegen lässt.

gewechselt, um von einer besseren Skalierung der Performance über mehrere Rechenknoten zu profitieren.
Die Ergebnisse der Simulationen deuten auf einen lipidabhängigen Effekt von Cholesterin hin: In gesättigten Lipid-Membranen konnte die erwartete Versteifung der Membran durch Cholesterin beobachtet werden. Umgekehrt zeigte sich jedoch eine sogar vergrößerte Flexibilität bei mehrfach ungesättigten Lipid-Membranen. Diese ungewöhnliche Eigenschaft von Cholesterin konnte auf dessen höhere Mobilität in ungesättigten Membranen sowie die Präferenz von Cholesterin für negative gekrümmte Membranbereiche zurückgeführt werden (sog. „diffusional softening“).
Mithilfe dieser Beobachtung können nun viele biologische Prozesse an der Zellmembran aus einer neuen Perspektive betrachtet werden: So erfordern die Aufnahme von Stoffen, die Kommunikation zwischen Zellen, und insbesondere die Membranfusion eine Umformung der Zellmembran. Ist diese deutlich weicher, geht der hohe Cholesterinanteil von Zellmembranen (ca.

40%!) mit deutlich verringerten energetischen Barrieren für diese Prozesse einher.
Das Projekt wurde teilweise von der DFG im Rahmen des SFB1027, Physical Modeling of Non-Equilibrium Processes in Biological Systems (Projekt C6), finanziell unterstützt. Die Publikation ist im OpenAccess Journal Nature Communications verfügbar: Matthias Pöhnl, Marius F. W. Trollmann, Rainer A. Böckmann. Nature Communications 2023, 14(1), 8038, DOI: 10.1038/s41467-023-43892-x.

Marius Trollman und Rainer Böckmann
Computational Biology, FAU Erlangen-Nürnberg

Gnadenschuss für Intel?

Alle Welt, nicht nur die KI-Fraktion, schreit nach GPUs, aber es gibt immer noch einen Trupp unbeugsamer ~~Gallier~~ Franken, der sich mutig gegen den Trend stellt. Glücklicherweise kommt gerade eben von NVIDIAs Gnaden der „Grace Superchip“ (GSC) daher und segnet uns mit vielen CPU-Kernen, anständiger Speicherbandbreite und tolerierbarer Leistungsaufnahme. Was ist also dran an dem Ding, und wie vergleicht es sich mit anderen populären Server-CPUs wie dem Intel Sapphire Rapids (SPR)? Hier machen wir ein kleines „Shootout“ zwischen dem Grace Superchip und einem Xeon Platinum 8470, wie er z.B. in der Large-Memory-Partition des „Fritz“-Clusters am NHR@FAU zum Einsatz kommt.
Wir haben natürlich eine Menge vortrefflicher Benchmarks vorbereitet, die das Thema in aller Breite abdecken; leider ist ein Quartl zu schmal, um sie alle zu fassen. Hier also ein schneller Überblick, der aber dennoch die wichtigsten Unterschiede zwischen GSC und SPR aufzeigen sollte.

Basics
Das SPR nutzt Intels „Golden Cove“-Architektur, GSC ist ein „Neoverse V2“-Core. In Tabelle 1 sind einige Eigenschaften der beiden Kontrahenten geenüber gestellt. Was auffällt, ist die größere Instruktionsparallelität beim GSC, was sowohl in der Zahl der Ports als auch bei den Vektoreinheiten ersichtlich ist. Auf der anderen Seite ist die SIMD-Vektorbreite beim GSC nur 128 bit; insgesamt hat ein Core deshalb nur die halbe Gleitkomma-Spitzenperformance wie ein SPR-Core, aber mehr Durchsatz bei skalarem oder 128-bittigem SIMD-Code. Insgesamt kann Intel also, wenn man die volle SIMD-Breite nutzen kann, in den meisten Disziplinen wie Load/Store-Durchsatz (incl. Gather) und arithmetischer Performance gut punkten — außer bei Gleitkomma-Divisionen, wo der Arm-Core fast dreimal schneller ist. Bei den Instruktionslatenzen hingegen hat der Arm-Core fast überall die Nase vorn. Das Ganze zeigt, dass der Neoverse V2 nicht nur für gut vektori-