IEEE Cluster Conference 2022 in Heidelberg

Die internationale High-Performance- und Cluster-Computing-Community kommt aktuell in Heidelberg zur IEEE Cluster Conference zusammen. Vom 6. bis 9. September bietet das etablierte Konferenzformat ein umfangreiches Programm mit Impulsvorträgen, Workshops, Tutorials und Diskussionen rund um Hardware, Middleware, Software und wissenschaftliche Cluster-Computing-Anwendungen. Das Forum bringt Forschende, Entwickler:innen und Anwender:innen aus Wissenschaft, Industrie, Labs sowie Wirtschaft zusammen, um sich über aktuelle Trends und technologische Entwicklungen im Cluster-Computing und dessen Nutzung für wissenschaftliche und kommerzielle Anwendungen auszutauschen. Die thematischen Schwerpunkte sind „Anwendungen, Algorithmen und Bibliotheken“, „Architektur, Netzwerke/Kommunikation und Management“, „Programmierung und Systemsoftware“ sowie „Daten, Speicherung und Visualisierung“. Ein wichtiges Element der Konferenz ist ein Mentoring-Programm zur Förderung des wissenschaftlichen Nachwuchses, das Einstiegsmöglichkeiten in das wissenschaftliche und berufliche Netzwerk bietet. Die diesjährige Edition wird federführend von Prof. Dr. Abhinav Bhatele (University of Maryland, College Park) und Prof. Dr. Felix Wolf (TU Darmstadt) organisiert; der lokale Gastgeber ist Prof. Dr. Holger Fröning (Universität Heidelberg) und Dr. Felix Zahn (CERN). Verschiedene GA-Mitglieder beteiligen sich wie folgt am Konferenzprogramm: ● Das JSC beteiligt sich mit einem Keynote-Vortrag über die Integration von Quantencomputern in HPC-Infrastrukturen und ist Mitorganisator des Workshops „HPC for International Collaboration between Europe and Latin America“, in dem Arbeiten aus dem EU-Projekt RISC2 vorgestellt werden. ● Das HLRS ist auf dem Workshop „Monitoring and Analysis for HPC Systems Plus Applications“ mit einem Vortrag zum Thema “Towards Real-Time Classification of HPC Workloads via Out-of-Band Telemetry” vertreten. ● Die Aktivitäten des IT Centers der RWTH Aachen University umfassen das Tutorial-Angebot „Introduction to Research Data Management (RDM) with Hands-On for HPC Use Cases“, eine Poster-Präsentation zum Thema „H2M: Towards Heuristics for Heterogeneous Memory“ sowie Co-Chair-Funktionen im Tutorial- sowie Workshop-Programm. Weitere Informationen: clustercomp.org/2022/ (Kontakt: Felix Wolf, TU Darmstadt)

Verstetigung für die KI-Forschung in Deutschland

Die deutsche KI-Forschung erhält seit dem 1. Juli einen kräftigen Schub: Fünf der insgesamt sechs deutschen KI-Kompetenzzentren werden mit einer institutionellen Förderung durch das BMBF und die beteiligten Bundesländer als dauerhafte Einrichtungen verstetigt. Bund und Sitzländer stellen insgesamt bis zu 100 Mio. Euro pro Jahr bereit. Ziel ist es, europäisch und international wettbewerbsfähige Forschung zu etablieren, um Deutschland als führenden Standort für Forschung, Lehre und Technologietransfer in der KI nachhaltig zu stärken. Die fünf Zentren wurden 2021 durch ein internationales Expertengremium evaluiert und erhalten mit der Verstetigung ein neues Fundament. Zusammen mit dem Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) bilden sie den Nukleus der deutschen KI-Forschung: das Berlin Institute for the Foundations of Learning and Data (BIFOLD), das Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig (ScaDS.AI Dresden/Leipzig), das Munich Center for Machine Learning (MCML), das Tübingen AI Center (TUE.AI) und das Kompetenzzentrum Maschinelles Lernen Rhein-Ruhr (ML2R), das zum „Lamarr Institute for Machine Learning and Artificial Intelligence“ wird. Weitere Informationen: bmbf.de/bmbf/shareddocs/pressemitteilungen/de/2022/07/0107-KI-PM.html.

ScaDS.AI Dresden/Leipzig – feierliches Kickoff

Zum Start der institutionellen Förderung des KI-Kompetenzzentrums „ScaDS.AI Dresden/Leipzig – Center for Scalable Data Analytics and Artificial Intelligence“ durch das BMBF und den Freistaat Sachsen fand am 5. September eine feierliche Kickoff-Veranstaltung an der TU Dresden statt. Neben dem sächsischen Staatsminister für Wissenschaft, Sebastian Gemkow, sowie den Rektorinnen der TU Dresden und der Universität Leipzig kamen zu diesem Anlass rund 200 geladene Vertreter:innen der deutschen KI-Kompetenzzentren und der Sächsischen Wirtschaft in Dresden zusammen. Auf dem Programm standen neben Vorträgen zu zentralen Themen des Kompetenzzentrums auch die Living-Labs-Einweihung des ScaDS.AI an der TU Dresden sowie eine Poster-Session zu aktuellen Forschungsaktivitäten. Weitere Informationen: scads.ai (Kontakt: Gina Valentin, ScaDS.AI Dresden/Leipzig)

Weltweit erstes NVIDIA ARM Cluster am KIT

Der Großteil der HPC-Systeme weltweit setzte im letzten Jahrzehnt auf nur zwei unterschiedliche Hardware-Architekturen: x86-Prozessoren von Intel/AMD und NVIDIA-GPUs. Dies erleichtert den Nutzenden den Wechsel zwischen Systemen, allerdings bleibt das Po-

tenzial alternativer Architekturen ungenutzt. Daher hat das NHR@KIT als zentralen Bestandteil die sogenannte „Future Technologies Partition“ (FTP) eingerichtet – ein Hard- und Software-Testbett für neue, disruptive Technologien, die in den großen Systemen noch nicht verfügbar sind. NVIDIA plant ab 2023 die Herstellung eigener Hochleistungsprozessoren auf Basis der Arm-Architektur. Diese sollen zusammen mit der nächsten GPU-Generation mit dem Codenamen Hopper in zukünftigen Supercomputern zum Einsatz kommen. Um bereits heute die Portierung von Anwendungen zu ermöglichen, bietet NVIDIA seinen Partnern Entwicklungssysteme an. Diese sind mit je einem Ampere Altra Arm-Prozessor, zwei NVIDIA A100 GPUs und einer NVIDIA BlueField-2 Data Processing Unit ausgestattet. Da die meisten Berechnungen auf den großen Systemen wie z. B. HoreKa am KIT auf mehr als nur einem Serversystem gleichzeitig laufen, hat NHR@KIT als einziger Standort weltweit ein Cluster aus diesen Systemen beschafft und in die FTP integriert. Weitere Informationen: scc.kit.edu/15764.php (Kontakt: Simon Raffeiner, KIT)

Schneller als das biologische Vorbild

Die Simulation biologisch plausibler, neuronaler Netzwerke in den Neurowissenschaften stellt Anforderungen an Latenzzeiten für Kommunikations- und Speicherzugriffe wie wenige andere Anwendungen. Selbst einfache Netzwerke mit etwa 80.000 Neuronen und 300 Mio. Synapsen (dieses „Potjans-Diesmann-Microcircuit“ genannte Modell entspricht etwa ein Millionstel der Hirnrinde) konnten bisher auf HPC-Systemen − aber auch dedizierten Neuromorphic-Computing-Systemen − bestenfalls in annähernd biologischer Realzeit simuliert werden. Für die zukünftige Erforschung von Lernvorgängen in deutlich komplexeren Netzwerken sind signifikant schnellere Simulationen jedoch unabdingbar. In einem von der Helmholtz-Gemeinschaft geförderten Kooperationsprojekt ist es am Forschungszentrum Jülich (FZJ) erstmals gelungen, das Microcircuit auf einem IBM-Neural-Supercomputer viermal schneller als in Realzeit zu simulieren. Obwohl diese FPGA-basierte Maschine und ihr Kommunikationssystem für völlig anders geartete, künstliche neuronale Netze optimiert ist, konnten neuartige Ansätze validiert werden, die zur Entwicklung besser angepasster Hardwaresysteme für die eigentliche Erforschung von Lernvorgängen dienen können. Neben dem Einsatz in der neurowissenschaftlichen Forschung könnten solche neuromorphen Maschinen zukünftig auch als zusätzliche Module in modulare Supercomputer integriert werden. Die Originalveröffentlichung zu diesen Arbeiten ist in „Frontiers in Neuroscience“ erschienen. Weitere Informationen: go.fzj.de/nc-interview (Kontakt: Arne Heittmann, PGI-10@FZJ)

Neue Kernkomponeten für die Open Science Cloud

Die Entwicklung einer offenen Arbeitsumgebung für öffentlich finanzierte, vertrauenswürdige Forschungsdaten und damit verbundener Dienste bis 2027 ist Ziel der European Open Science Cloud (EOSC). Im Rahmen der Umsetzungsphase werden dafür in dem aktuellen dreijährigen Projekt FAIRCORE4EOSC seit 1. Juni neun neue Kernkomponenten entwickelt. Neben einer EOSC-Forschungssoftware-Infrastruktur und persistenten Identifikatoren gehören dazu auch Komponenten für die gemeinsame Nutzung bzw. Verknüpfung verschiedener Metadatenschemata sowie forschungsorientierte Suchdienste für alle EOSC-Repositorien. Koordiniert vom finnischen IT Center for Science arbeiten in dem Projekt 22 Partner von Dienstleistern, Forschungseinrichtungen und globalen Expertennetzwerken zusammen, darunter die beiden Gauß-Allianz-Mitglieder GWDG und DKRZ. Die GWDG ist federführend für die Entwicklung von Diensten im Kontext von persistenten Identifikatoren verantwortlich. Das DKRZ bringt die Anforderungen der Klimamodellierung insbesondere für die Datensuche und für persistente Identifikatoren ein. Weitere Informationen: dkrz.de/de/faircore4eosc. (Kontakt: Hannes Thiemann, DKRZ)

NHR-Stipendienprogramm des NHR-Vereins

Die NHR Graduate School ist die gemeinsame Graduiertenschule der neun NHR-Zentren. Das Stipendienprogramm richtet sich an Wissenschaftler:innen, die im Bereich des Hochleistungsrechnens promovieren möchten. Die in das Programm aufgenommenen Promovierenden werden ab 1. April 2023 für einen Zeitraum von bis zu drei Jahren gefördert. Die Betreuung erfolgt an einem NHR-Zentrum, das den jeweiligen Forschungsschwerpunkt der wissenschaftlichen Arbeit anbietet. Ein sechsmonatiger Aufenthalt an einem der anderen Zentren vermittelt Einblicke in die divergente Arbeitsweise im NHR-Verbund. Zur jährlich stattfindenden Summer School treffen sich alle NHR-Stipendiat:innen zum Austausch und gemeinsamen Lernen. So werden die Promovierenden Teil einer bundesweiten, hervorragend vernetzten Wissenschaftsgemeinschaft. Nachdem der erste Jahrgang im Frühjahr erfolgreich gestartet ist, läuft bis zum 15. September die zweite Ausschreibung. Weitere Informationen: nhr-verein.de/en/application (Kontakt: scholarship@nhr-verein.de)

Veranstaltungen

- 26.-27.09.2022: Machine Learning with AMD Instinct™ GPUs and ROCm™ Software, HLRS@GCS, online
- 29.-30.09.2022: AMD Instinct™ GPU Training, HLRS@GCS, online
- 17.-21.10.2022: CFD with OpenFOAM®, HLRS@GCS, Stuttgart

HPC-Kalender der Gauß-Allianz:
hpc-calendar.gauss-allianz.de