

Big Data meistern

„Die Informationsflut steigt. Unsere Forschung hilft, aus Big Data rasch und effizient nutzbare Informationen zu gewinnen.“

Die Fortschritte im Supercomputing ermöglichen immer komplexere Anwendungen. Das ist wichtig, damit wir zu neuen Erkenntnissen kommen. Doch immer realistischere Simulationen haben ihren Preis: wachsende Datenmengen, die wir bewältigen müssen, beispielsweise bei Simulationen in der Klimaforschung, der physikalischen Grundlagenforschung und in den Neurowissenschaften. Auch mit dem Human Brain Project, das auf die Simulation des menschlichen Gehirns abzielt, brechen wir in Bezug auf die zu erwartende Datenmenge in neue Dimensionen auf. Das Problem von Big Data: Das Speichern und Übertragen von großen Datenmengen dauert und bremst Anwendungen aus. Gefragt sind Ansätze, mit denen wir Daten schnell und effizient nutzen können. Das heißt, Wichtiges herausfiltern, gegebenenfalls aus unterschiedlichen Quellen zusammenführen und dann rasch auswerten. Diese Herausforderung betrifft nicht nur das Supercomputing, sondern die gesamte Informations- und Kommunikationstechnologie.

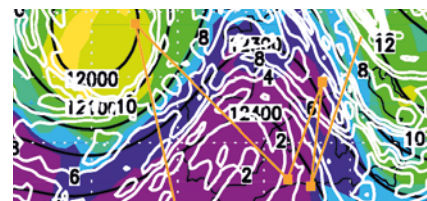
Um Lösungen zu finden, arbeiten wir eng mit Partnern aus Wissenschaft und Wirtschaft zusammen. So haben unsere Forscher im Exascale Innovation Centre gemeinsam mit Experten von IBM ein neues Speicherkonzept entwickelt: das Blue Gene Active Storage (siehe Seite 2). Es ermöglicht Supercomputern, große Datensätze zu verarbeiten und zu reduzieren, bevor diese auf Datenträger geschrieben werden. Dadurch können datenintensive Aufgaben deutlich schneller erledigt werden. Damit die Industrie direkt von unseren Forschungsergebnissen profitiert, beteiligt sich das Forschungszentrum Jülich am neuen Smart Data Innovation Lab. Das Vorhaben vereint Unternehmen und Forschungseinrichtungen, damit große Datenmengen aus Industrieprozessen künftig effizienter genutzt werden (siehe Seite 4). Beide Vorhaben sind wichtige Fortschritte, um aus Big Data nutzbare Informationen zu gewinnen.

Prof. Achim Bachem
*Vorstandsvorsitzender
des Forschungszentrums Jülich*

IN DIESER AUSGABE

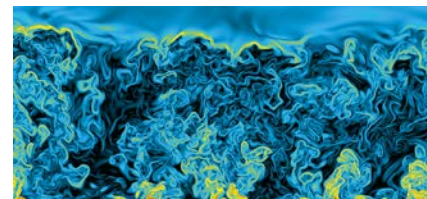
SEITE 2:

Mit virtueller Realität zu besseren Wetterprognosen



SEITE 3:

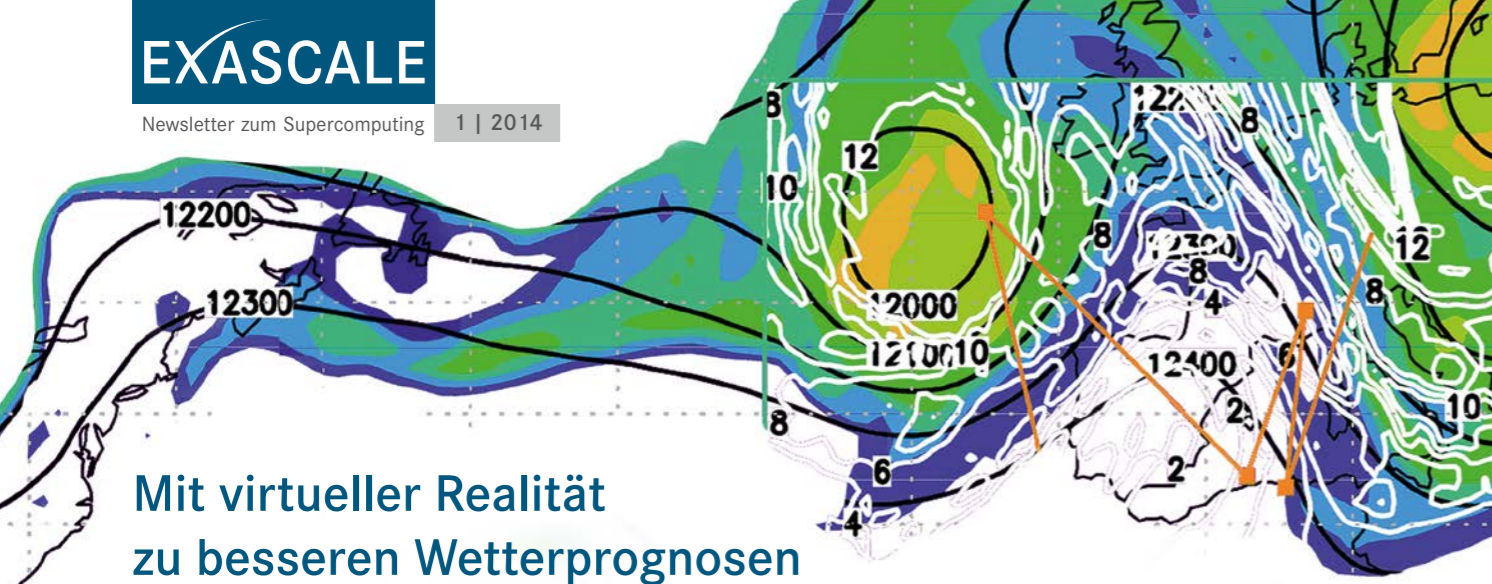
Turbulenzen: Chaos am Himmel



SEITE 4:

Kurznachrichten
Termine
Impressum

www.fz-juelich.de/ias/jsc



Mit virtueller Realität zu besseren Wetterprognosen

Auf dem Jülicher Supercomputer JUQUEEN entsteht eine virtuelle Flusslandschaft. Eine DFG-Forschergruppe wird das rund 150 Kilometer mal 200 Kilometer große Einzugsgebiet des Neckars auf dem Rechner simulieren, um Wasser- und Energieflüsse besser zu verstehen. Mit den Erkenntnissen wollen die Forscher Vorhersagen für Wetter und Hochwasser verbessern. Die Deutsche Forschungsgemeinschaft (DFG) fördert das Vorhaben „Data Assimilation for Improved Characterisation of Fluxes across Compartmental Interfaces“ in den kommenden drei Jahren mit rund 2,1 Millionen Euro.

Die Wissenschaftler haben verschiedene Modelle, die sich auf einzelne Aspekte wie Hydrologie oder Atmosphäre konzentrieren, zu einem Gesamtmodell vereinigt. „Dadurch können wir bislang vernachlässigte Wechselwirkungen einbeziehen, etwa zwischen Grundwasser und Atmosphäre. Das hilft uns, die Stoffflüsse besser zu verstehen“, erklärt Mitkoordinator Prof. Harrie-Jan Hendricks-Franssen vom Jülicher Institut für Bio- und Geowissenschaften. Grundlage für die Berechnungen sind Messdaten aus dem Neckareinzugsgebiet. Aus Modellen und Daten entsteht eine virtuelle Realität, die die Forscher

wie ein reales Untersuchungsgebiet behandeln wollen. So werden sie virtuelle Messungen etwa zu Verdunstung und Bodenfeuchte durchführen. Die Ergebnisse dieser Messungen nutzen die Wissenschaftler, um Modell und Vorhersagen immer wieder zu korrigieren. Datenassimilation nennt die Forschung dieses Vorgehen. Zugleich erhoffen sich die Projektpartner Rückschlüsse darauf, welche Messungen wichtig sind, um eine bestimmte Vorhersage zu verbessern – beispielsweise Niederschlagsmessungen für Prognosen von Grundwasserneubildungen.

Forscher simulieren ein ganzes Flusseinzugsgebiet, um Energie- sowie Wasserflüsse besser zu verstehen und Wettervorhersagen zu optimieren.

www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/DE/2013/13-12-02-dfg.html

Intelligenter Speicher-Booster für große Datenmengen

Dank einem neuen Speicherkonzept, dem Blue Gene Active Storage (BGAS), können Supercomputer datenintensive Aufgaben künftig noch schneller bewältigen. Das IT-Unternehmen IBM und das Forschungszentrum Jülich stellten die weltweit erste Installation des BGAS-Systems im November 2013 auf der größten Supercomputing-Konferenz vor, der SC13 in Denver/USA. Das System läuft am Jülich Supercomputing Centre (JSC) auf dem Superrechner JUQUEEN. Von der Installation sollen hochkomplexe Simulationen profitieren, insbesondere aus der Hirnforschung.

Solche Anwendungen erzeugen enorme Datenmengen. Die Daten auf großen Festplattensystemen außerhalb des High-Performance-Clusters abzuspeichern, hat einen schwerwiegenden Nachteil: Die Nutzer können nur eingeschränkt auf sie zugreifen. Das liegt an den langsamen Übertragungsraten solcher Systeme. „Umfangreiche Datenmengen, wie sie in der Hirnforschung erzeugt

werden, lassen sich mit diesen konventionellen Methoden nicht mehr verarbeiten. Neue Ein-/Ausgabekonzepte wie das Active Storage sind daher eine wichtige Voraussetzung, um mit Hilfe von Höchstleistungsrechnern neue Forschungsergebnisse auf diesem Gebiet zu erzielen“, erläutert JSC-Direktor Prof. Thomas Lippert.

Das Konzept des „aktiven Speichers“ haben Jülicher Wissenschaftler und Experten von IBM im gemeinsam betriebenen Exascale Innovation Centre (EIC) entwickelt. BGAS verfügt über eigene Prozessoren, was den Zugriff deutlich beschleunigt. Denn dadurch ist das System in der Lage, große Datensätze zu verarbeiten und zu reduzieren, bevor diese auf Datenträger geschrieben werden. Ergänzend kommen extrem schnelle, nichtflüchtige Flash-Speichermedien zum Einsatz.

www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/DE/2013/13-11-18bgas.html



Mit dem Blue Gene Active Storage-System kann der Jülicher Supercomputer JUQUEEN datenintensive Aufgaben noch schneller bewältigen. Davon profitiert als Erstes die Hirnforschung.

Turbulenzen: Chaos am Himmel

Fast jeder Flugzeugpassagier hat es schon einmal erlebt. Mitten im Steig- oder Sinkflug fängt die Maschine plötzlich an zu vibrieren. Schuld sind Verwirbelungen der Luft. Solche Turbulenzen sind insbesondere beim Eintritt in die erdnahe Atmosphäre aus der relativ ruhigen darüberliegenden Troposphäre anzutreffen. Dr. Juan Pedro Mellado vom Hamburger Max-Planck-Institut für Meteorologie will mit Hilfe von Simulationen auf Jülicher Höchstleistungsrechnern neue Erkenntnisse über diese Phänomene gewinnen. Das John von Neumann-Institut für Computing hat sein Vorhaben „Direkte Numerische Simulation turbulenter Mischungsprozesse in der planetaren Grenzschicht“

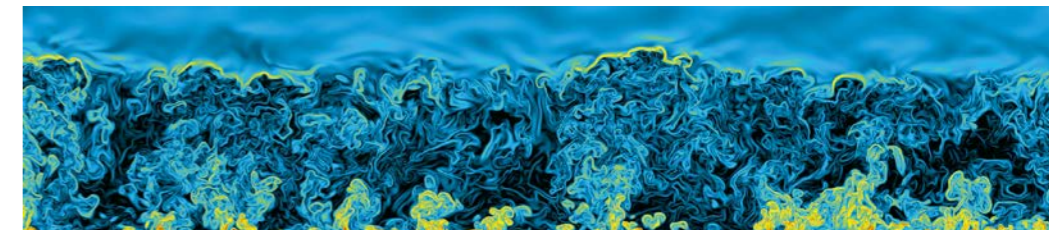
als „John von Neumann Exzellenzprojekt 2013“ ausgezeichnet.

Die planetare Grenzschicht reicht bis etwa 2.000 Meter Höhe. In diesem Übergangsbereich zwischen Erdoberfläche und Atmosphäre sind die Luftschichten ständig in Bewegung. Außerdem beeinflussen Faktoren wie Berge oder Gebäude die Luftströmungen. Wird eine Strömung zu stark gestört, vermischen sich unterschiedliche Luftmassen, es kommt zu Verwirbelungen. Das ist auch für das Klima von Bedeutung. Denn in der Grenzschicht finden wichtige Austauschprozesse statt, etwa von Wärme und Wasser. Das Problem: Turbulenzen entwickeln sich

chaotisch und verändern sich rasch. „Erst Fortschritte im Supercomputing haben es in den letzten Jahren ermöglicht, die Mischungsprozesse genauer zu studieren und dadurch einige ihrer grundlegenden Prinzipien zu verstehen“, sagt Mellado. In seinem Projekt will er unter anderem herausfinden, wie sich unterschiedliche Untergründe, etwa eisbedeckter und eisfreier Ozean, auf die Verwirbelung in den darüberliegenden Luftschichten auswirken. Außerdem untersucht er die Wechselwirkungen zwischen Wolken und Turbulenzen. Diese komplexen Zusammenhänge können in Klimamodellen noch nicht ausreichend berücksichtigt werden.

In der erdnahen Atmosphäre (untere Bildhälfte) wirbeln Luftmassen wild umher.

www.fz-juelich.de/nic/Projekte/mellado.html



Der schnellste Wasserkocher der Welt

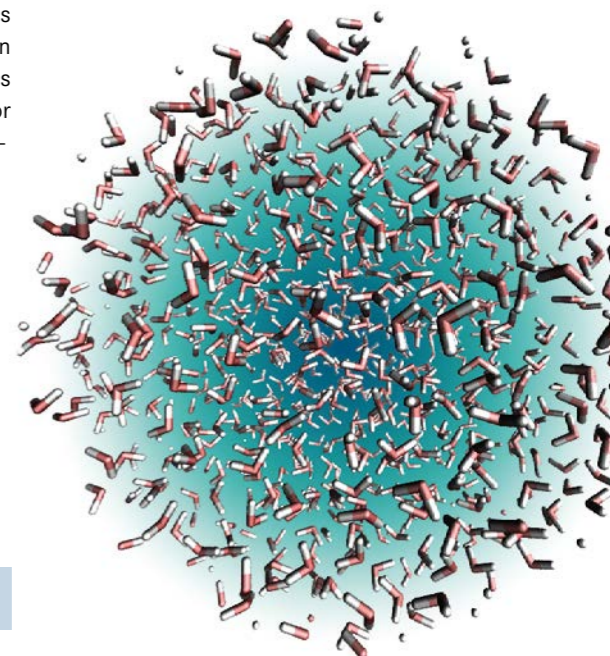
Hamburger Forscher haben eine Methode gefunden, Wasser in weniger als einer billionstel Sekunde zum Kochen zu bringen. Damit hätten sie den schnellsten Wasserkocher der Welt entwickelt. Das Verfahren wurde auf dem Jülicher Supercomputer JUROPA simuliert, aber noch nicht experimentell umgesetzt.

Bei der Simulation beschießen die Wissenschaftler vom Center for Free-Electron Laser Science (CFEL) Wasser mit einem konzentrierten Blitz aus sogenannter Terahertz-Strahlung. Die Frequenz dieser elektromagnetischen Wellen liegt zwischen Radiowellen und Infrarotstrahlung. Der Blitz bringt Wassermoleküle zum Vibrieren. Die Wasserstoffbrückenbindungen, die Wassermoleküle besonders eng zusammenhalten, lösen sich. Es entsteht eine heiße Miniwolke, die in weniger als einer tausendstel Sekunde

auseinanderfliegt. Diese Zeit reicht den Forschern, um wichtige Vorgänge in thermischen Reaktionen zu beobachten. Rund 200.000 Stunden Prozessorzeit auf JUROPA waren notwendig, um die Wechselwirkung des Terahertz-Blitzes mit dem Wasser zu berechnen. Dabei wurde ein Nanoliter – das ist ein milliardstel Liter – in nur einer halben Pikosekunde auf rund 600 Grad Celsius erwärmt. Auf einem einzelnen Prozessor hätten die Berechnungen rund 20 Jahre gedauert.

Die Arbeit eröffnet neue Experimentiermöglichkeiten mit erhitzten chemisch oder biologisch relevanten Proben. „Wasser ist das bedeutendste Medium, in dem chemische und biologische Prozesse stattfinden. Es stabilisiert bestimmte chemische Verbindungen und manche Reaktionen

werden durch Wasser überhaupt erst möglich“, erklärt Dr. Oriol Vendrell vom CFEL. Die Forscher untersuchen nun, wie der Terahertz-Blitz auf verschiedene im Wasser gelöste Molekülararten wirkt.



Heiße Miniwolke mit kurzer Lebensdauer: Forscher haben am Superrechner Wasser in weniger als einer billionstel Sekunde zum Kochen gebracht.

Angewandte Chemie – International Edition: DOI: 10.1002/anie.201305991

KURZNACHRICHTEN

Mehr Leistung mit weniger Strom

Neben DEEP Extended Reach (DEEP-ER, siehe Nr. 3/2013) hat im Oktober 2013 ein zweites neues EU-Projekt zum Exascale Computing seine Arbeit aufgenommen: Mont-Blanc 2. Das Projekt ist eine Erweiterung des bereits laufenden EU-Projekts Mont-Blanc. In dem Vorhaben, an dem auch das Jülich Supercomputing Centre (JSC) beteiligt ist, geht es um einen energie-sparenden Superrechner der Zukunft. Mit einer Trillion Rechenoperationen pro Sekunde wird die Exascale-Generation 100 bis 1.000-mal schneller sein als derzeitige Petascale-Systeme. Die Partner von Mont-Blanc entwickeln eine Rechnerarchitektur, die zugleich 15 bis 30-mal energieeffizienter ist als heutige Supercomputer. Dafür setzen sie auf stromsparende Prozessoren aus dem Mobiltelefonsektor. JSC-Forscher passen in Mont-Blanc 2 das Performance-Analyse-Softwaretool Scalasca an Hardware-Prototypen an, die mit solchen Prozessoren ausgestattet sind.

www.deep-er.eu
www.montblanc-project.eu

Industriedaten besser nutzen

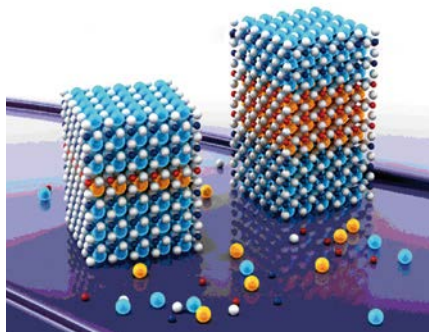
Unternehmen und Forschungseinrichtungen haben im Januar 2014 mit dem Smart Data Innovation Lab (SDIL) eine Plattform gegründet, um große Datenmengen aus Industrieprozessen künftig besser zu managen und effizienter zu nutzen. Dafür entsteht in Karlsruhe eine Höchstleistungsinfrastruktur. SDIL soll neueste Forschungsergebnisse rasch an die Industrie weitergeben und ihr so Wettbewerbsvorteile verschaffen. Das Jülich Supercomputing Centre leitet zusammen mit dem Chemiekonzern Bayer das Forschungsfeld „Personalisierte Medizin“. Dabei soll die Behandlung einer Krankheit mit Arzneimitteln auf den jeweiligen Patienten zugeschnitten werden. Weitere Forschungsfelder sind die Energiewende, Smart Cities und die sogenannte Industrie 4.0, die industrielle Fertigung und Informationstechnologie miteinander verknüpft.

www.sdil.de

SimLab für Chemie und Physik

Seit Ende 2013 gibt es ein weiteres Simulationslabor (SimLab) am Jülich Supercomputing Centre: das JARA-HPC Simulation Laboratory „Ab-Initio Methods in Chemistry and Physics“. Es unterstützt Nutzer der Jülicher Supercomputerressourcen, die sogenannte Ab-initio-Simulationen in Physik, Chemie, den Nanowissenschaften und den Materialwissenschaften durchführen. Bei diesen Simulationen fließen nur solche Parameter in Berechnungen ein, die nicht durch Experimente gewonnen wurden. Das SimLab kümmert sich insbesondere um Forschungsvorhaben, die sich mit Algorithmenentwicklung, Optimierung und Leistungssteigerung beschäftigen.

www.jara.org/index.php?id=192&S=0&L=0



Jetzt auch für Smartphone und Tablet!

Exascale Newsletter
www.exascale-news.de

„FORSCHEN in Jülich“ (als Tablet-Magazin)
www.fz-juelich.de/app



iOS (iPad)



Android

IMPRESSUM

EXASCALE NEWSLETTER des Forschungszentrums Jülich **Herausgeber:** Forschungszentrum Jülich GmbH 52425 Jülich **Konzeption und Redaktion:** Dr. Anne Rother (v.i.S.d.P.), Tobias Schlöber, Christian Hohlfeld **Grafik und Layout:** Grafische Medien, Forschungszentrum Jülich **Bildnachweis:** S.1: Forschungszentrum Jülich, ©fotolia.com/adimas/vladgrin; S. 2, 3 o.: Forschungszentrum Jülich; S. 3 u.: Oriol Vendrell/DESY; S. 4: JARA **Kontakt:** Geschäftsbereich Unternehmenskommunikation | Tel.: 02461 61-4661 Fax: 02461 61-4666 | E-Mail: info@fz-juelich.de **Druck:** Schloemer & Partner GmbH **Auflage:** 700

TERMINE

GPU Programming

7. – 9. April 2014
 am Jülich Supercomputing Centre

In dem PRACE Advanced Training Centre (PATC)-Kurs lernen die Teilnehmer CUDA, OpenCL und Multi-GPU-Programmierung kennen. Schwerpunkte sind Optimierung und Tuning von wissenschaftlichen Anwendungen.

Dozenten: Dr. Jan Meinke, Jochen Kreutz, Peter Philippen, Dr. Andrej Adinets, Anke Zitz, JSC; Jiri Kraus, NVIDIA

www.fz-juelich.de/ias/jsc/events/gpu

Advanced GPU Programming

5. – 6. Mai 2014
 am Jülich Supercomputing Centre

Der Kurs beschäftigt sich mit der Suche und Beseitigung von Engpässen, der Profilerstellung und mit fortgeschrittenen Programmier Techniken bei der GPU-Programmierung.

Dozenten: Dr. Jan Meinke, Jochen Kreutz, Peter Philippen, Willi Homberg, Dr. Andrej Adinets, JSC; Suraj Prabhakaran, GRS; Jiri Kraus, NVIDIA

www.fz-juelich.de/ias/jsc/events/advgpu

Programmierung und Nutzung der Supercomputerressourcen

19. – 20. Mai 2014
 am Jülich Supercomputing Centre

Neue Nutzer der Jülicher Supercomputerressourcen können sich in dem Kurs einen Überblick über die vorhandenen Systeme verschaffen. Unter anderem lernen sie, bewilligte Rechnerressourcen optimal zu nutzen.

Dozenten: Firmenmitarbeiter von IBM, Intel und ParTec; Mitarbeiter des JSC

www.fz-juelich.de/ias/jsc/events/sc-may

Paralleles I/O und portable Datenformate

21. – 23. Mai 2014
 am Jülich Supercomputing Centre

Performanceverluste von Parallelrechnern vermeiden, Simulationen auf unterschiedlichen Systemen durchführen: Der PATC-Kurs stellt paralleles I/O sowie portable Datenformate wie HDF5 und netCDF vor.

Dozenten: Wolfgang Frings, Dr. Michael Stephan, Dr. Florian Janetzko, JSC

www.fz-juelich.de/ias/jsc/events/parallelio

Eine Übersicht über die Veranstaltungen am Jülich Supercomputing Centre finden Sie unter:

www.fz-juelich.de/ias/jsc/events