

## Termine

- **Forschung Live! Tag der offenen Tür in Garching:** Am 25.10. findet der Tag der offenen Tür der TUM in Garching statt. Prof. Bode hält im Rahmen der Veranstaltung einen Vortrag über Supercomputer. Weiterhin wird das Infiband-Cluster seines Lehrstuhls präsentiert. KONWIHR wird zusammen mit der **abayfor** auf der Ausstellung vertreten sein.
- Am 25. Oktober findet im Großraum Erlangen-Nürnberg die "Lange Nacht der Wissenschaften" statt. (Vgl. Beitrag weiter vorne im *Quartl*.)
- "Objekt-orientierte Programmierung in C++" lautet der Titel eines Kurses, der vom 19.-22. Januar am LRZ stattfindet.
- Anfang Februar werden dann am LRZ in drei Kurse *HPC-Grundlagen* sowie Details zu *Totalview* und *VAMPIR* vorgestellt werden.

## Nichts los im Land?

Erlangen, im Januar 2004

Wir trauern um das  
**KONWIHR-Quartl**  
(ehem. **FORTWIHR-Quartl**)

Mangels Beiträgen aus Netzwerk-Kreisen konnte das *Quartl* nach 36 erfolgreichen Ausgaben nicht mehr weiter künstlich am Leben gehalten werden.

In stiller Trauer  
*Quartl-tz*

Wollen Sie diese Todesanzeige auf dem nächsten *Quartl* sehen?

Nein — dann helfen Sie mit. Greifen Sie zur Feder oder besser zum Computer und fangen Sie an zu schreiben ...

**PS:** Das rechtzeitige Abschicken an [konwihr@lstm.uni-erlangen.de](mailto:konwihr@lstm.uni-erlangen.de) bitte nicht vergessen!

## Bitte notieren:

Die letzte planmäßige KONWIHR-Antragrunde wurde im September eingeläutet. Kümmern Sie sich daher ggf. um geeignete Anschlußfinanzierungen außerhalb von KONWIHR!

Für den 10. Dezember ist die diesjährige **KONWIHR-Mitgliederversammlung** geplant. Die Einladungen werden zusammen mit der geplanten Tagesordnung rechtzeitig durch die Geschäftsstellen an die Projektleiter verschickt werden.

### Impressum

KONWIHR *Quartl*\*

– das offizielle Mitteilungsblatt des Kompetenznetzwerks für technisch-wissenschaftliches Hoch- und Höchstleistungsrechnen (KONWIHR) – erscheint jeweils zum Quartalende.

#### Herausgeber:

Prof. Dr. A. Bode, Sprecher des KONWIHR,  
Prof. Dr. Dr. h.c. F. Durst, stellv. Sprecher

#### Redaktion:

Dipl.-Ing. Th. Zeiser  
Lehrstuhl für Strömungsmechanik  
Cauerstraße 4, D-91058 Erlangen  
Tel./Fax: ++49-9131-85 28280 / 29503  
**e-mail:** [konwihr@lstm.uni-erlangen.de](mailto:konwihr@lstm.uni-erlangen.de)

Dipl.-Inf. A. Schmidt  
Institut für Informatik, TU München  
D-85748 Garching bei München  
Tel./Fax: ++49-89-289 17680 / 17662  
**e-mail:** [konwihr@in.tum.de](mailto:konwihr@in.tum.de)

#### WWW:

<http://www.konwihr.uni-erlangen.de/>  
<http://konwihr.in.tum.de/>

#### Redaktionsschluss

für die nächste Ausgabe: **14.12.2003**  
**Nächste Ausgabe** des *Quartls*:  
**Januar 2004**

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

# KONWIHR *Quartl*

(36. Ausgabe) 3/2003

Quartl-online: <http://www.konwihr.uni-erlangen.de/quartl/>

## Editorial

Nach den sprachwissenschaftlichen Ausführungen in der letzten Ausgabe mit der ausführlichen Diskussion von allerlei Kreativitäten rund um den geliebten „Brückentag“ sehen wir uns schon wieder mit einer Wortneuschöpfung konfrontiert – dem reflexiven Verb „sich einquarteln“. Kurze Denkpause gefällig? Die Auflösung ist simpel: Jemand – in zunehmendem Maße offenbar vorzugsweise akademische Würdenträger – „quartelt sich ein“, wenn er sich gewissen Editorial-verdächtigen Schreiberlingen gegenüber gezielt mit einer handfesten Sottise in Szene zu setzen und somit seine Erwähnung im *Quartl* gewissermaßen gewaltsam herbeizureden sucht. Da gab es zum Beispiel die Münchener Spektabilität, die in vorösterlicher Tollkühnheit auf ihren (in der Tat beinahe zitierwürdigen) verbalen Erguss gleich noch ein „Ich muss Ihnen doch Nachschub für Ihre Kolumne liefern“ draufsetzte und ihre Absichten gar nicht erst zu kaschieren versuchte. Oder dann in der Hitze dieses Sommers jener TU-Vizepräsident (Semi-Magnifizenz?), der in sein seelenruhig bis salbungsvoll vortragenes „Ja, ja – die Schwaben haben eine Neigung zu abartigen Entscheidungen“ ein solch gerüttelt Maß an Lebenserfahrung hineinlegte, dass diese Bemerkung in früheren Zeiten die sofortige Mobilmachung westlich von Ulm zur Folge gehabt hätte. Dass aber nicht einmal die Sekundanten sich bemühen mussten, hat zwei Gründe. Erstens fühlt man sich als in Baden Geborener, in Württemberg Aufgewachsener, in Oberbayern Ausgebildeter und dann in Bayrisch- sowie schließlich Württembergisch-Schwaben (dem einzig echten Schwaben übrigens) Berufstätiger nicht ganz so heftig auf den Schlipps getreten wie vielleicht ein richtiger Herr Häberle (also ein Vertreter jener Landsmannschaft, die man im Allgäu schon mal liebevoll „Neckar-Preußen“ nennt). Zweitens – und dies ist vielleicht sogar der wichtigere Grund – hat der gute Mann unter Beachtung des konkreten Kontexts durchaus Recht.

Denn worüber war gesprochen worden? Über die *Qualitätsoffensive für die Universitäten des Landes Baden-Württemberg* – zunächst also ganz und gar nichts Abartiges, sondern ein (zumindest in der Theorie) segensreiches Werkzeug. Wir machen uns schlau und lesen: „*Im nationalen und internationalen Wettbewerb um herausragende (Nachwuchs-) WissenschaftlerInnen will das Ministerium für Wissenschaft, Forschung und Kunst durch die Bereitstellung zusätzlicher Ausstattungsmittel Spitzenberufungen fördern. Es sollen für besonders herausragende WissenschaftlerInnen attraktive Arbeitsbedingungen geschaffen werden, die im Rahmen der mit den eigenen Mitteln des Landes sonst üblichen Konditionen nicht zu erreichen sind. Hierdurch werden Impulse für innovative neue Arbeitsgebiete an den Universitäten des Landes, insbesondere im Bereich der zentralen Schlüsseltechnologien und auf besonders zukunftssträchtigen, risikoreichen Forschungsgebieten mit der notwendigen kritischen Masse an Ressourcen erwartet.*“

Im Klartext: Endlich eine Chance, der hierzulande oftmals beklagten Ohnmacht gegenüber den Berufungsverhandlungspraktiken vieler US-amerikanischer sowie anderer Eliteunis nach der Art „Ich ging ’rein, der Rektor drückte mir einen dicken Scheck in die Hand und gab mir 30 Sekunden Bedenkzeit, und ich musste einfach zusagen“ zu entkommen – ohne das ansonsten übliche Gezerre zwischen Fakultät und Hochschule, wer denn nun das fehlende Neuntel der Sekretariatsstelle etc. zu finanzieren habe. Das Ganze ist insbesondere auch dafür gedacht, Deutsche Wissenschaftler im Ausland zur Rückkehr zu bewegen (natürlich nur die Vertreter der leuchtenden Spitzen-Exzellenz!). Jetzt unternehmen wir an

unserer Fakultät derzeit einen solchen Versuch einer „Heimholung“ – das Match lautet gewissermaßen Universität Stuttgart gegen MIT. Und so freuten wir uns natürlich über diesen von der Obrigkeit in weiser Verwendung von Privatisierungserlösen angelegten Topf, sammelten eifrig Argumente, schrieben Stellungnahmen noch und nöcher – um dann in den ersten Julitagen von der Hochschulverwaltung zu vernehmen, dass die Deadline der Antragseinreichung beim Ministerium der 30. Juni gewesen sei. Der Antrag werde dann zum 30. 9. (2003, immerhin) bearbeitet und weitergeleitet. Also nichts mit dem Tempovorteil, nichts mit der Überholspur – zurück ins Glied, Sommerloch abwarten, und dann wird die Begutachtung angestoßen. Na ja, so fließen wenigstens ein paar Zinsgroschen in den Qualitäts-Sparstrumpf.

Nun gut – der eine schüttelt den Kopf über leichtfertig vertane Chancen, der andere wundert sich eben über abartige schwäbische Praktiken. Wie’s ausgeht, wird man sehen. Wie immer an dieser Stelle natürlich der Link für Interessierte:

[http://mwk.baden-wuerttemberg.de/Aktuelles/Ausschreibungen/qualitaetsoff/Ausschreibung\\_kurz.html](http://mwk.baden-wuerttemberg.de/Aktuelles/Ausschreibungen/qualitaetsoff/Ausschreibung_kurz.html)

Lieber Herr Vizepräsident, vielen Dank für Ihre schwer zu toppende Steilvorlage – Sie setzen Maßstäbe! Wundern werden Sie sich über das alles freilich kaum – schließlich sind Ihnen die Neigungen der Schwaben offensichtlich ja bestens bekannt. Doch nun viel Spaß bei der Lektüre der neusten Ausgabe des *Quartls!*

Hans-Joachim Bungartz

(1,3 GHz), 112 Gbyte gemeinsamem Hauptspeicher und fast 3 TByte Plattenplatz.



Abb. 1: SGI-Altix System

Aufgrund der bisherigen guten Erfahrungen mit der Itanium2 CPU ist zu erwarten, dass die Anlage auch anspruchsvollsten numerischen Aufgaben gewachsen sein wird: Benchmarkläufe auf einem baugleichen 32-CPU System in Manchester (allerdings mit den langsameren 900 MHz-Prozessoren) haben für Lattice-Boltzmann-Codes eine Performance ergeben, die einen NEC SX-6 Vektorprozessor um einen Faktor 3 und einen IBM p690 Knoten mit ebenfalls 32 CPUs um einen Faktor 2 hinter sich lässt.

Dazu kommt die besondere Eigenschaft des Shared Memory, die das System auch für Anwendungen interessant macht, die mit den Eigenheiten eines „echten“ Clusters nicht kompatibel sind. Z. B. können OpenMP-Codes natürlich vom großen gemeinsamen Hauptspeicher profitieren, aber auch die MPI-Performance kann sich sehen lassen: Der bekannte Pallas-MPI-Benchmark liefert eine Ping-Pong Bandbreite von über 1,1 GByte/s und eine Latenz von ca. 1,7  $\mu$ s. Das sind Werte, die übliche Standard-

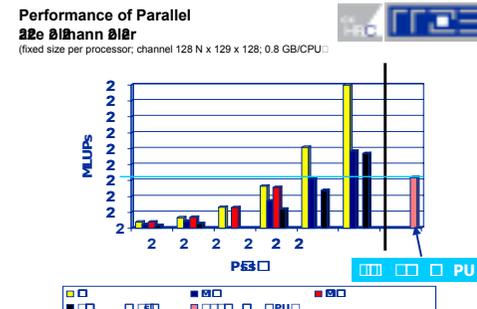


Abb. 2: Lattice Boltzmann Benchmarkergebnisse auf verschiedenen Systemen

Hochgeschwindigkeits-Interconnects wie z. B. Myrinet (ca. 250 MByte/s Bandbreite bzw. 10–15  $\mu$ s Latenz) bei weitem nicht erreichen. Mit dem Redhat Linux Betriebssystem werden sich viele Anwender schnell zurecht finden, zudem hat SGI die Softwareausstattung mit Performancetools und optimierten numerischen Bibliotheken komplettiert. Das neue System wird voraussichtlich in der zweiten Oktoberhälfte installiert.

Zusätzlich wurde eine Ergänzung des IA32-Clusters um weitere 13 Rechenknoten veranlasst. Damit stehen den Kunden auf dieser Maschine in Kürze insgesamt 172 Xeon-CPU's zur Verfügung. Im Lichte der Auslastungszahlen während der ersten 5 Betriebsmonate wird diese Maßnahme von vielen Anwendern begrüßt. Das RRZE verfügt nun über einen Park von Hochleistungsrechnern, der die verschiedensten Bedürfnisse abdeckt und sich ideal zwischen Institutsressourcen und Höchstleistungsrechenzentren platziert [3].

[georg.hager@rrze.uni-erlangen.de](mailto:georg.hager@rrze.uni-erlangen.de)

- [1] <http://www.rrze.uni-erlangen.de/dienste/hpc/cluster32/>
- [2] <http://www.sgi.com/servers/altix/>
- [3] <http://www.rrze.uni-erlangen.de/dienste/hpc/Rechner/>

## Fazit und Ausblick

Nach dem zunächst enttäuschenden Start der Itanium Prozessor-Familie kann man jetzt endlich sagen, dass diese Architektur im Bereich des Hochleistungsrechnens einsatzfähig ist. Hierzu haben auch Fortschritte bei der Compiler-Entwicklung erheblich beigetragen. Dennoch verbleibt bei der Software sicher noch Raum für eine bessere Ausschöpfung der vollen Leistungsfähigkeit dieses Prozessors. Für den Einsatz am ganz oberen Ende – also für wissenschaftliche Problemstellungen, die traditionell auf Vektor-Architekturen implementiert werden mussten – wird man über die Standard-Chipsätze hinaus deutlich verbesserte Speicher-Interfaces verwenden müssen. Der NUMA-Ansatz, wie er von der Firma SGI in ihren Itanium-basierten Altix-Systemen verwendet wird ist sicher eine vielversprechende Entwicklung.

Aus Sicht des LRZ wird sich der Wechsel auf die 64-Bit Architektur mit Sicherheit bezahlt machen: Auf den jetzt verfügbaren Systemen kann im Vergleich zu IA32-kompatiblen Systemen mit deutlich höherer Applikationsleistung gerechnet werden; der Wegfall der Ein-

schränkungen bei der Speicheradressierung gestattet es, völlig neue Problemklassen anzugehen. Mit dem Itanium-Cluster steht nunmehr am LRZ ein weiterer Parallelrechner zur Verfügung, der bezüglich Programmierung und Skalierbarkeit den Weg zu den Systemen der nächsten Generation ermöglicht; so wird dieses Rechensystem ein unverzichtbares Hilfsmittel für die Konzeption der Benchmark-Suite zur Beschaffung des nächsten Bundeshöchstleistungsrechners sein.

Wir sind schon gespannt, was die Quantenchemiker bei ihrem nächsten Besuch am LRZ sagen werden . . .

## Danksagungen

Mein Dank geht an Herrn Dr. Ludger Palm (LRZ) für die Gaussian 98 Testrechnungen und Leistungsdaten, sowie an Herr Marek Behr, Ph.D. (LNL der TU München) für die Resultate und Abbildungen seiner Rechnungen mit XNS. Herrn Dr. Gerhard Wellein (Regionales Rechenzentrum Erlangen) danke ich für seine Hinweise zur Optimierung der Triadenleistung auf der Itanium-Architektur.

reinhold.bader@lrz-muenchen.de

## Erlanger Prozessorgeflüster

### Beschaffung eines Altix-Systems von SGI am RRZE und Ausbau des IA32-Clusters

von Georg Hager (HPC-Gruppe am RRZE)  
Nachdem die erste Stufe der Ausschreibung für ein Compute-Cluster am RRZE in der Beschaffung eines 32-Bit-Systems der Firma Transtec [1] gipfelte, fiel nun Anfang September auch die Entscheidung für die zweite Stu-

fe, die mit 64-Bit-CPU's ausgestattet werden sollte. Den Zuschlag erhielt aufgrund der Beurteilung mittels Benchmark-Tests das System mit dem besten Preis-/Leistungsverhältnis: ein Altix 3700 Rechner der Firma SGI.

Obwohl die Ausschreibung ein Cluster mit Hochgeschwindigkeitsnetzwerk spezifizierte, wurde bewusst auch der Weg für reine Shared-Memory Lösungen offen gehalten. Die angebotene Altix 3700 [2] ist ein ccNUMA-System mit 28 Itanium2-Madison Prozessoren

Erlangen / Nürnberg, 25. Okt. 2003

## Lange Nacht der Wissenschaften

Am 25. Oktober wird es soweit sein. Im Großraum Nürnberg–Fürth–Erlangen wird die erste *Lange Nacht der Wissenschaften* stattfinden. Über 300 Bildungs- und Informationsangebote werden von rund 150 Forschungseinrichtungen und Unternehmen präsentiert. Sicherlich etwas außergewöhnliches!

“*Supercomputing in Wissenschaft und Technik*” — unter diesem Motto präsentieren sich KONWIHR und das Regionale Rechenzentrum Erlangen (RRZE) mit einem Gemeinschaftsstand im Hörsaal-Mensa-Gebäude am Roten Platz der Technischen Fakultät in Erlangen. Das Motto der TechFak “Wir gestalten Zukunft”, trifft sicherlich auch auf KONWIHR zu.

## Von Dienstleistungen und Pizzaschachteln

Wenn dem PC die Luft ausgeht, kommen Wissenschaftler der nordbayerischen Universitäten meist zur HPC-Gruppe am RRZE, wo sie in der HPC-Gruppe kompetente Ansprechpartner, Beratung und Hilfe, aber auch modernste Rechner finden, deren Leistungsfähigkeit typischerweise um einen Faktor 100–200 größer ist als die eines modernen PCs. Was Technik in der Form einer größeren “Pizzaschachtel” zu leisten vermag oder wie das Innenleben der neuesten 64-Bit-Rechner aussieht, zeigen Dr. Wellein und seine Mitarbeiter vom KONWIHR-Projekt cxHPC.

**Numerische Strömungsmechanik — ohne Supercomputer geht nichts**  
Supercomputer leisten das 1000- bis 10 000-Fache eines modernen PCs! Solche Superrechner liefern die technische Voraussetzung für komplizierte Simulationsrechnungen wie die Vorhersage

von aero- und hydrodynamischen Eigenschaften. Mitarbeiter aus dem Lehrstuhl für Strömungsmechanik von Prof. Durst demonstriert verschiedene Einsatzmöglichkeiten von Höchstleistungsrechnern in der numerischen Strömungsmechanik, u.a. aus den KONWIHR-Projekten BESTWIHR, Flusib und DiSiVGT. Anhand eines eindrucksvollen Experiments wird der LSTM-Erlangen auch die experimentelle Sichtbarmachung von Strömungsphänomenen zeigen.

## Sensoren und anderes — Elektrodynamische Autolautsprecher aus dem Computer

Schnellschaltende Magnetventile für die Automobilindustrie, Leistungsultraschall für die Medizintechnik oder auch elektrodynamische Autolautsprecher sind Beispiele für hochkomplexe Sensor-Aktor-Systeme. Mitarbeiter vom Lehrstuhl für Sensorik (Prof. Lerch) zeigen interaktiv, wie sie vor dem Bau eines Prototyps Design, Leistungsfähigkeit und Produktionsprozess am Computer optimieren. Strömungsinduzierter Lärm, wie er im KONWIHR-Projekt FlowNoise untersucht wird, muss dabei natürlich auch mit berücksichtigt werden.

## Metallschäume sind Metallträume — Werkstoffkunde & Informatik

Metallschäume sind neue Werkstoffe, von denen Ingenieure im Automobil- und Flugzeugbau träumen: leicht und stabil! Das KONWIHR-Projekt FreeWiHR (Dr. Körner vom Lehrstuhl für Werkstoffkunde und Technologie der Metalle, sowie der Lehrstuhl für Systemsimulation von Prof. Rüde) beschäftigt sich mit der Simulation von derartigen Metallschäumen. Da der Produktionsprozess der Schäume noch nicht vollständig verstanden ist, sollen mit Hilfe numerischer Simulationen neue Erkenntnisse ge-

wonnen und die Produktionsparameter wie Druck, Gaskonzentration und die Beschaffenheit des Rohmaterials optimiert werden. Ein Teilaspekt dieses Projekts ist die Erzeugung aussagekräftiger Visualisierungen, um die simulierten Vorgänge intuitiv erfassen zu können.

### Theoretische Physik — Vielteilchensysteme, chaotische Dynamik

Auch im Institut für Theoretische Physik II wird viel Rechenleistung benötigt, wenn es um die Simulation der Dynamik von Vielteilchensystemen im Detail geht. Die Anwendungsbeispiele liegen in verschiedenen Gebieten: Plasmaphysik, große Moleküle, Nanoteilchen oder Atomkerne. Die Methoden sind entsprechend vielfältig und umfassen sowohl klassische Molekulardynamik als auch aufwändige quantenmechanische Simulationen. In allen Fällen tritt als grundlegender Aspekt chaotische Dynamik auf, typisch für Systeme mit vielen Freiheitsgraden und nichtlinearer Kopplung. Anhand einfacher Modellsysteme werden die überraschenden Eigenschaften chaotischer Dynamik demonstriert.

Am Gemeinschaftsstand von KONWIHR und RRZE gibt es also viel zu sehen. Neben Postern und Animationen wird es kompetente Gesprächspartner und interaktive Computersimulationen aber auch Handfestes (aktuelle Rechner und "Pizaboxen" zum Bestaunen, ein Experiment zur Strömungssichtbarmachung, sowie die eine oder andere Überraschung) geben.

**Wo:** Uni-Erlangen, TechFak, Mensa-Hörsaal-Gebäude, Erwin-Rommel-Straße  
**Wann:** Sa 25.10.2003, 19:00–01:00 Uhr

Aber auch an anderen Stellen wird Ihnen KONWIHR und das Hochleistungsrechnen begegnen ...

### Computer & Chemie

Computer spielen unter anderem auch in der modernen Chemie eine wichtige Rolle. Das Computer-Chemie-Centrum in Erlangen ist die einzige Einrichtung seiner Art in Europa. Dabei geht es u.a. um die Entwicklung von Methoden zur Vorhersage von chemischen Eigenschaften, die dann benutzen werden, um Arzneistoffe und Werkstoffe zu entwickeln. Auch die Wirkungsweise von Enzymen kann nur mit Hilfe leistungsfähiger Rechner bestimmt werden, wie dies beispielsweise im KONWIHR-Projekt Enzymech gemacht wurde.

**Wo:** Uni-Erlangen, Computer-Chemie-Centrum, Nägelsbachstraße 25

**Wann:** Sa 25.10.2003, 19:00–01:00 Uhr

Und auch an der FH-Nürnberg treffen Sie wieder auf KONWIHR ...

### Ein Chip entsteht

so lautet das Motto des Lehrstuhls für Elektronische Systeme (ELSYS) bei der Langen Nacht der Wissenschaften. U.a. wird dabei die *Intelligente Kamera* von Prof. Urbanek aus dem KONWIHR-Projekt FPGA zeigen, wie leicht oder wie schwierig die Erkennung von Verkehrszeichen ist.

**Wo:** FH-Nürnberg, Maschinenhalle

**Wann:** Sa 25.10.2003, stündlich ab 19:00

Machen Sie sich auf den Weg!

[www.nacht-der-wissenschaften.de](http://www.nacht-der-wissenschaften.de)

**PS:** Die Eintrittskarten gelten im gesamten VAG/VGN-Bereich von Samstag 12:00 bis Sonntag 8:00 Uhr auch als Fahrausweise in den öffentlichen Nahverkehrsmitteln.

Durch die Umstellung auf die Winterzeit bleibt Ihnen nach der Langen Nacht sogar noch eine Stunde mehr zum Ausschlafen.

Quartl-tz

Auf der SR8000 gibt es zwei Kommunikationsparadigmen: Mit "MPP" ist die Verteilung von MPI-Prozessen auf einzelne Prozessoren gemeint, während im "COMPAS"-Modus ein MPI-Prozess einen 8-Prozessor-Knoten umfasst. Die MPI-Implementierung von Hitachi ist offenbar auf letzteren Modus hin optimiert (insbesondere kann nur in diesem Modus MPI das Betriebssystem umgehen); dessenungeachtet ist der Myrinet-Interconnect bezüglich der Latenz sogar im Vergleich mit dem COMPAS-Modus deutlich besser. Im Gegensatz zum COMPAS-Modus kann man bei Nutzung von MPP auf der Hitachi die eigentlich verfügbare Hardware-Bandbreite nur zu einem Bruchteil nutzen.

### XNS: Ein Navier-Stokes Löser

Das Programmpaket XNS, entwickelt an der Rice-Universität sowie an der TU München, simuliert u. a. anderem dreidimensionale inkompressible Strömungen; hierbei werden die Lösungen der Navier-Stokes-Gleichungen auf einem unregelmäßigen, an die Geometrie der umströmten Gegenstände angepassten Gitter numerisch approximiert. Das Programmpaket ist mit MPI parallelisiert worden und läuft schon seit längerem auf klassischen Massiv-Parallelrechnern wie etwa der auf dem Alpha Prozessor basierenden Cray T3E. Eines der am LRZ-Cluster gerechneten Beispiele beschreibt die Umströmung eines U-Bootes der Los-Angeles Klasse; das Gitter wurde mittels eines Octree-Gittergenerators erzeugt und enthält 153000 Tetraeder-Elemente. Der obere Teil von Abbildung 4 zeigt, wie sich das Gitter rund um das U-Boot anschmiegt: Dort sind die größten Variationen in der Strömung zu erwarten. Der untere Teil zeigt, welche Rechenleistung (in Zeitschritten pro Stunde) das LRZ-Cluster im Vergleich zu

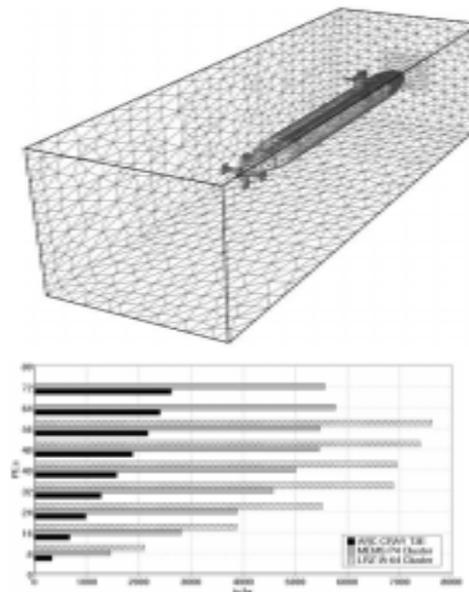


Abbildung 4: Simulation der Umströmung eines U-Bootes

anderen Parallelrechnern erzielt.

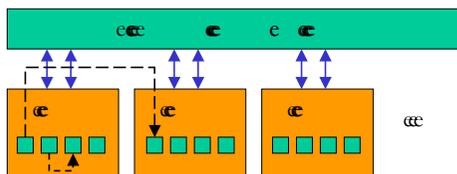
Mit "MEMS P4-Cluster" ist hierbei ein mit 1.7 GHz getaktetes Pentium-4 Cluster bezeichnet, das als Interconnect ebenfalls Myrinet 2000 verwendet, und "ARC CRAY T3E" gibt die Leistung für die T3E-1200 an. Letzterer Rechner zeigt aufgrund der inzwischen veralteten Prozessoren eine geringe Grundrechenleistung, aber der auch nach heutigem Stand der Technik sehr leistungsfähige Interconnect (die Latenzen liegen unterhalb von 3 Mikrosekunden!) gewährleistet eine sehr gute Skalierung des Codes auf hohe Prozessorzahlen. Im Gegensatz hierzu stößt man bei den Myrinet-basierten Systemen in der Gegend von 32 CPUs auf eine Skalierbarkeitsgrenze. Trotz der höheren Taktung der Pentium-4 Systeme liefern die Itanium-CPU's eine um etwa ein Drittel höhere Applikationsleistung. Für größere Probleme erhält man jedoch deutlich verbessertes Skalierungsverhalten.

hier betrachteten Systemen abheben. Beschränkt man sich andererseits auf rein cache-orientierte Probleme, so legen die Leistungszahlen einen Vergleich zwischen der konventionellen RISC-Bauweise der SR8000 und der neuen EPIC-Bauweise des Itanium-2 mit zusätzlicher Parallelität in der Instruktionsverarbeitung nahe: Wäre nämlich der SR8000-Knoten vergleichbar schnell getaktet wie eine Itanium-CPU, so entsprächen 8 Hitachi CPUs in etwa 4 Itanium-CPU's. Für den hier betrachteten Anwendungsbe- reich bringt also EPIC einen Leistungs- vorteil um einen Faktor 2.

Durch geeignete Angaben bei der Job- Konfiguration ist es im Übrigen möglich, die vier CPUs eines 4-Wege Systems als einen Rechenknoten zu betreiben. Ob- wohl Gaussian 98 nicht perfekt paralle- lisiert ist, so erhält man bei den hier ge- rechneten Tests Effizienzen von 60–75%, mit vier Prozessoren erhält man also im- merhin noch bis zur dreifachen Rechen- leistung der Einzel-CPU.

### Der Myrinet-2000 Interconnect

Die Itanium-Systeme sind untereinan- der (und mit dem Storage-System) über einen Myrinet-Interconnect ver- bunden. Das führt für parallele An- wendungen im Vergleich zu Standard- Netzwerktechnologien zu wesentlich ver- besserter Skalierbarkeit und Performan- ce. Um einen im Verhältnis zur hohen Knotenleistung gut balancierten Paral- lelrechner zu erhalten, ist jeder 4-Wege Knoten über zwei Myrinet-Links an den Interconnect angeschlossen.



Die Abbildung illustriert, wie parallele Programme nun entweder innerhalb eines Knotens (im shared Memory) oder über Knotengrenzen hinweg (über den Myrinet Switch) miteinander kommunizieren (gestrichelte Pfeile). Als Software-Schnittstelle für die Kommunikation wird das Message Passing Interface (MPI) verwendet; es kommt eine von der Firma Myricom für die Myrinet-Hardware optimierte (quelloffene) MPICH-Variante zum Einsatz. Parallele Programme werden ohne weitere Vorkehrungen zunächst vier MPI Prozesse pro Knoten starten; da jeder Myrinet-Link mehrere Kommunikationskanäle zulässt, funktioniert das auch. Allerdings könnte es sich je nach Anwendung herausstellen, dass es effizienter ist, mehrere CPU's zu einem Rechenknoten zusammenzufassen. Innerhalb eines Rechenknotens wird dann nicht mit MPI kommuniziert, sondern der Compiler sorgt für automatisch oder durch Unterstützung mit OpenMP-Direktiven parallelisierten Code. Inwie- weit dieser Einsatz hybrid paralleler Pro- gramme sinnvoll ist, muss in jedem Ein- zelfall untersucht werden; dies ist ggf. mit nicht unbeträchtlichem Arbeitsaufwand verbunden.

Die Tabelle gibt einen Überblick über die Eckdaten für die parallele Kommu- nikation und einen Vergleich zur Hitachi SR8000; sie wurden aus dem sogenannten Pallas MPI Benchmark gewonnen, wo- bei der MPI-Aufruf "Sendrecv" gemessen wurde, bei dem bidirektionaler Datenaus- tausch stattfindet.

Konfiguration	Latenz ( $\mu$ s)	Bandbreite (MBytes/s)
<b>Myrinet</b>		
IA64 Node-Intern	5,3	~1000
IA64 Switch	11,9	422
<b>SR8000</b>		
MPP Node-Intern	25	210-230
MPP Crossbar	35	220
COMPAS Crossbar	18-21	~1900

## Strömung in porösen Medien Vergleich LB-Simulation – NMR-Messung

Moderne experimentelle Techniken so- wie die stetig steigende Rechenleistung in Verbindung mit neuen numerischen Verfahren erlauben immer öfter die *lo- kale* Betrachtung von komplexen geo- metrischen Strukturen (wie z.B. Zu- fallsschüttungen) und der darin ablaufen- den Transportvorgänge. Hierdurch wird es erst möglich, komplexe Phänomene richtig zu verstehen und ihre Wirkung in der technischen Anwendung zu berück- sichtigen bzw. sogar gezielt auszunutzen. Experimentelle Verfahren und numeri- sche Simulationsmethoden können sich bei dieser Erforschung gut ergänzen. Al- lerdings werden diese Synergieeffekte bis- her oftmals noch nicht ausgenutzt.

Genau die Nutzung derartiger Synergie- effekte ist wesentlicher Bestandteil eines kürzlich begonnenen DFG-Projekts, in dem der Lehrstuhl für Mechanische Ver- fahrenstechnik und Mechanik von Prof. Buggisch aus Karlsruhe und der LSTM- Erlangen sowie die TC1-Erlangen zusam- menarbeiten.

### Motivation

Durchströmte Packungen, Festbettreak- toren, gepackte Kolonnen und andere poröse Strukturen sind in der chemischen Verfahrenstechnik in vielfältiger Weise im Einsatz, um beispielsweise Reaktio- nen durchzuführen, Stoffe zu mischen, zu trennen, abzuscheiden oder zu filtern und Wärme zu- oder abzuführen.

Das globale Verhalten kann hierbei durch Messung der Zu- und Abflüsse relativ leicht an vorhandenen Anlagen bestimmt werden. Rückschlüsse auf die im In- neren ablaufenden Vorgänge sind hier- mit jedoch nur sehr begrenzt möglich.

Auf Grund der fehlenden Informati- on über die lokal ablaufenden Prozesse sind Modellgleichungen, deren Modellpa- rameter an globale Messdaten angepasst wurden, nur für den vermessenen Be- reich der Prozessbedingungen (Parame- terbereich, Geometrie) anwendbar. Ins- besondere sind somit keine zuverlässigen prädiktiven Simulationen (z.B. für ein scale-up oder bei Geometrieänderungen) möglich.

Werden die lokal ablaufenden Pro- zesse richtig verstanden und in ge- eigneten (mehrdimensionalen) Gleichun- gen berücksichtigt, so ist zu erwar- ten, dass eine deutlich bessere Model- lierung und Vorhersage des Betriebsver- haltens möglich wird. Um die lokalen Vorgänge zu untersuchen, sind nicht- invasive Messmethoden, aber auch di- rekte numerische Simulationen prinzipiell gut geeignet. Bisher scheitert dies jedoch noch häufig, weil die entsprechenden Me- thoden noch nicht existieren oder aber zu aufwändig sind.

Optische Methoden, wie z.B. die Laser- Doppler-Anemometrie (LDA), die bereits vereinzelt zur Geschwindigkeitsbestim- mung innerhalb von Kugelschüttungen angewandt wurde (z.B. [3, 5]), erfordern eine aufwändige Anpassung der optischen Eigenschaften von fester und fluider Phase. Die Notwendigkeit optischer Transpa- renz und übereinstimmender Brechungs- indices schränkt die möglichen Materialien und Fluideigenschaften stark ein.

### Ortsaufgelöste NMR

*Ortsaufgelöste NMR* (magnetic reso- nance imaging) wurde seit ihrer Ent- wicklung 1972 mit großem Erfolg in der Medizin eingesetzt und ist in den letzten Jahren für Ingenieure eine zu- nehmend wichtige Messmethode gewor- den. Eine weitere eher nichtklassische

Anwendung der NMR ist die *Beobachtung von Transportprozessen*, nämlich die Messung von Selbstdiffusionskoeffizienten oder Strömungsgeschwindigkeiten [2]. Die Kombination von beiden Techniken, d.h. die *orts aufgelöste* Messung von *Transportprozessen* mit NMR, stellt für Untersuchungen in komplexen Strukturen eine ideale Messmethode dar. Ihr Vorteil liegt in der Möglichkeit, nicht-invasive Messungen an optisch nicht-transparenten Systemen vorzunehmen.

Bisher können dreidimensionale Spindichte-Bilder zur Erfassung *geometrischer Strukturen* verwendet werden sowie *ein- und zweidimensionale stationäre Geschwindigkeitsprofile* gemessen werden. Die höchste räumliche Auflösung der dreidimensionalen Spindichte-Bilder beträgt  $(40 \mu\text{m})^3$  und erfordert eine Messzeit von mehreren Stunden. Die Aufnahme eines eindimensionalen stationären Geschwindigkeitsprofils einer Geschwindigkeitskomponente mit einer räumlichen Auflösung von etwa  $200 \mu\text{m}$  benötigt etwa eine Stunde, die Messung zweidimensionaler Geschwindigkeitsprofile mehrere Stunden bis Tage. Um den Messaufwand in Grenzen zu halten erscheint es sinnvoll, begleitende Simulationen durchzuführen.

### Transportvorgänge & Längenskalen

In porösen Strukturen und Packungen treten stets unterschiedliche Längenskalen auf, die auf der Nano-Ebene (z.B. Hohlräume von Zeolithen) beginnen können und im Falle der Geologie bei großräumigen Reservoirs enden [4]. Im Folgenden sollen nur noch die Ebenen betrachtet werden, auf denen in der verfahrenstechnischen Anwendung hydrodynamische Prozesse ablaufen. Dies ist zum einen die *mikroskopische Ebene*, die durch die Freiräume zwischen den Füllkörpern gebildet wird und

als charakteristische Länge die Abmessung eines Füllkörpers besitzt. Die hier ablaufenden Prozesse lassen sich durch die *elementaren Transportgesetze* unter Verwendung von *physikalischen Stoffdaten* beschreiben. Prinzipiell wäre diese Ebene für alle Betrachtungen wünschenswert, was jedoch in der Regel am enormen Simulationsaufwand und der Unkenntnis der lokalen Strukturen scheitert. Daher werden insbesondere in der praktischen Anwendung *makroskopische Ansätze* verwendet, die eine Ebene höher ansetzen und auf der Annahme lokaler Homogenitäten beruhen. Die Transportvorgänge werden hierbei durch *effektive* oder *gemittelte Parameter* beschrieben, welche nicht nur von den Stoffeigenschaften sondern vor allem von der zugrundeliegenden — jetzt aber homogenisiert betrachteten — Struktur abhängen. A priori lassen sich diese Parameter daher nur selten vorhersagen.

### 3D numerische Simulation mit Lattice Boltzmann Verfahren

Detaillierte dreidimensionale CFD-Simulationen können — im Gegensatz zu den klassischen experimentellen Methoden — direkt Informationen über die Zustände in den Porenräumen liefern.

Bei den Lattice Boltzmann Verfahren [1] handelt es sich um einen einfachen gas-kinetischen Ansatz. Die Modellierung der Fluidbewegung setzt hier auf der mikroskopischen bzw. mesoskopischen Ebene an, wobei jedoch durch geschickte Vereinfachungen (im wesentlichen durch eine physikalische Diskretisierung des Problems) die Zahl der Freiheitsgrade gegenüber klassischen gas-kinetischen Verfahren deutlich reduziert wurde, wodurch eine äußerst effiziente Simulation möglich wird. Nichts destotrotz werden aber Hochleistungsrechner für die Simulation benötigt.

welches Leistungsfenster typische HPC-Applikationen hineinfallen. Es wird damit auch klar, dass das System für den am LRZ geplanten Zweck sinnvoll einsetzbar ist: Das in der Quantenchemie sehr häufig eingesetzte Programmpaket Gaussian 98 ist auf optimale Ausnutzung der Caches hin programmiert worden.

Untypisch im Vergleich zu herkömmlichen Cache-basierten Systemen ist jedoch, dass sich die Triadenleistung durch händisches Eingreifen am Code deutlich verbessern lässt: Die mit “Unroll4” bezeichnete Kurve in Abbildung 2 ergibt sich durch Angabe einer Compiler-Direktive vor der Triadenschleife, und die Kurve “Unroll4\_Memlayout” erhält man darüber hinaus erst durch eine nicht triviale Veränderung des Codes, durch den die Vektoren in ganz bestimmter Weise im Speicher angeordnet werden. Hier muss konstatiert werden, dass der Compiler noch nicht von sich aus optimal laufenden Maschinencode erzeugt; der Leistungsunterschied beträgt im L2-Cache fast ein Drittel! Für Anwendungen muss dieser Effekt natürlich nicht unbedingt in hier vorliegender Weise auftreten, es ist jedoch damit zu rechnen, dass optimale Rechenleistung auch hier erst nach sorgfältigem Tuning des Codes erzielt werden kann.

### Quantenchemie mit Gaussian 98

Wenden wir uns nun dem Bereich zu, für den die Cluster-Erweiterung hauptsächlich – wenn auch nicht ausschließlich – gedacht ist: Mit dem schon erwähnten Quantenchemie-Paket Gaussian 98 – das neueste Release Gaussian 03 ist im Übrigen am LRZ-Cluster ebenfalls schon verfügbar – wurden die im folgenden beschriebenen Messungen an verschiedenen Architekturen ausgeführt. Und zwar wurden sogenannte “single-point” Molekülrechnungen mit kleinen

Basissätzen an unsymmetrischen halogenierten Kohlenwasserstoffen  $C_nH_{2n}ClF$  mit variabler, durch den Parameter  $n$  gekennzeichnete Molekülgröße durchgeführt. Abb. 3 zeigt die Gesamtrechnenzeiten für verschiedene Architekturen.

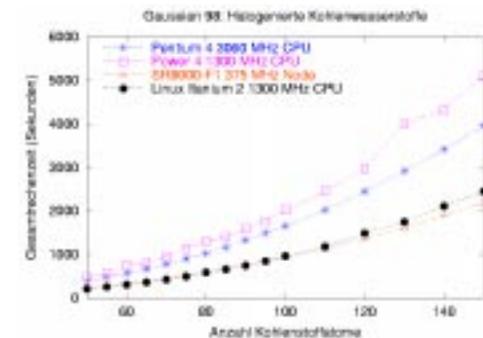


Abbildung 3: Rechenzeit für Gaussian 98 Tests auf verschiedenen Architekturen

Es ergeben sich folgende Aussagen:

1. Der Power-4 Prozessor ist auf Grund seines nur mit ein Drittel der Systemfrequenz getakteten L3-Caches in diesem Anwendungsbereich dem Itanium-2 System deutlich unterlegen.

2. Trotz der deutlich geringeren Taktung gegenüber dem Pentium-4 liefert das Itanium-2 System etwa ein Drittel mehr Rechenleistung bei gleicher Problemgröße. Die Entscheidung seitens des LRZ, im SMP-Bereich auf 64-Bit Architekturen umzusteigen ist damit auch von der Applikationsleistung her gerechtfertigt.

3. Der Vergleich mit einem 8-CPU Knoten des Höchstleistungsrechners Hitachi SR8000 zeigt, dass dieser vom technischen Standpunkt durchaus nicht ungeeignet für quantenchemische Aufgabenstellungen ist. Auf Grund des Speicher-Interfaces, das mit 32 GByte/s soviel leistet wie der Cache eines Itanium-Systems dürfte sich der Hitachi-Knoten sogar bei sehr großen Problemstellungen sehr vorteilhaft von allen anderen

kurrent zu etwa den Produkten der Firma NAG geworden. Selbstverständlich können auf Anforderung weitere Pakete nachinstalliert werden; eine Anfrage an [lxadmin@lists.lrz-muenchen.de](mailto:lxadmin@lists.lrz-muenchen.de) genügt um dies zu veranlassen.

### Leistungsmessungen

Eine erste Bewertung der Rechenleistung der neu beschafften Systeme basiert zum Einen auf den abnahme-relevanten Benchmark-Läufen, zum Anderen auf den ersten Testläufen der Benutzerprogramme.

### Die Itanium Cache-Hierarchie

Wenn ein Mikroprozessor Rechnungen ausführt, wird er nur auf solchen Daten unmittelbar operieren können, die in den Prozessor-Registern liegen. Davon gibt es nur wenige, sodass Zugriffe auf größere Datenmengen Zugriffe in das ggf. außerhalb der eigentlichen CPU liegende Speichersystem erfordern. Beim Itanium-Prozessor besteht dieses aus einem 3-stufigen, mit voller Systemfrequenz getakteten Cache, und dem Hauptspeicher, der deutlich langsamer getaktet ist und damit nur eine Bandbreite von 6.4 GByte/s im Gegensatz zu den 32 GByte/s etwa des L3 Caches aufweist. Zugriffe auf den Hauptspeicher bremsen also die CPU gründlich aus, egal wie schnell sie auch getaktet sein mag. Zur Illustration dieses Sachverhalts betrachten wir die Rechenleistung der Vektor-Triade, die durch das Code-Fragment

```
DO I=1, N
  A(I)=B(I)*C(I)+D(I)
END DO
```

definiert sei. Die Rechenleistung der Triade für verschiedene Vektorlängen N wird durch die mit "Compiler" bezeichnete Kurve in Abbildung 2 illustriert; der Grund für diese Bezeichnung ist, dass der Compiler den Code ohne weitere Eingriffe durch den Benutzer optimiert.

Man sieht, wie bei einer Vektorlänge von etwa 10000 – hier wird die Grenze des L2-Caches erreicht – die Rechenleistung deutlich einbricht; von den 32 GByte/s nomineller Bandbreite werden im L3-Cache knapp 20 GByte/s wirklich erreicht. Bei einer Vektorlänge von etwa 106 beginnt der Prozessor schließlich mit direkten Zugriffen auf den Hauptspeicher; dort werden von 6.4 GByte/s nominaler Bandbreite 4.0 GByte/s tatsächlich erreicht. Von der Spitzenleistung des Prozessors, nämlich 5.2 GFlop/s – das ist wegen der zweifach vorhandenen FMA-Einheiten das Vierfache der Taktfrequenz – werden dann noch 250 MFlop/s, also 4.8% erzielt.

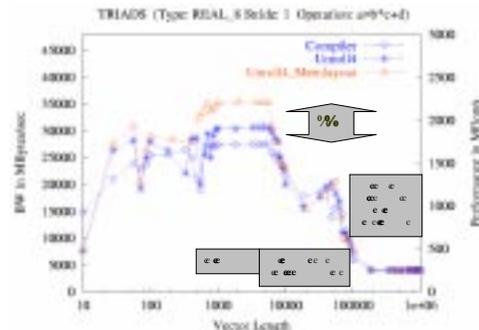


Abbildung 2: Vektor-Triade auf dem mit 1.3 GHz getakteten Itanium-2 Prozessor

Das hier geschilderte Verhalten ist durchaus typisch für cache-basierte Rechensysteme; Variationen beschränken sich in der Regel etwa darauf, einen größeren aber dafür langsamer getakteten L3-Cache zu verwenden; auch die Zahl der Prozessor-Register wird variiert. Nur wenige Hersteller stellen ein schnelleres oder parallel zugreifbares Speicher-Interface zur Verfügung, da dies den Preis des Gesamtsystems stark verteuert. Die tatsächlich zu erzielende Rechenleistung ist damit stark von der jeweiligen Applikation abhängig: In Abbildung 2 ist durch die Text-Kästchen angedeutet, in

### Ergebnisse und Ziele

Bereits ein erster quantitativer Vergleich zwischen den Simulationsergebnissen und NMR-Geschwindigkeitsmessungen in verschiedenen Ebenen einer Kugelschüttung (Abb. 1) zeigt eine erstaunlich gute Übereinstimmung (Abb. 2).

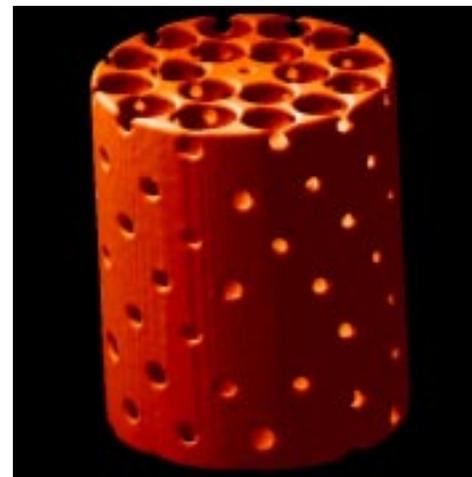


Abb. 1: Beispiel einer untersuchten porösen Struktur: Kugelschüttung aus 4mm Kugeln in einem Rohr mit 20mm Durchmesser.

### NMR-Geschwindigkeitsmessung

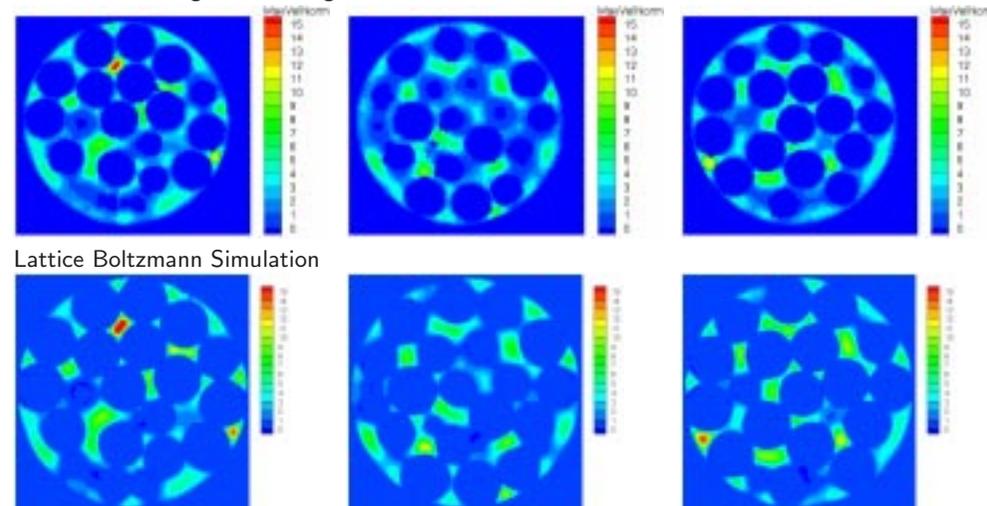


Abbildung 2: Geschwindigkeitsverteilung in verschiedenen Ebenen der Kugelschüttung. (Vgl: Farbabbildung in der Online-Ausgabe!)

In der Simulation können wesentlich leichter als bei den NMR-Messungen Reynoldszahlen variiert und lokale Informationen im gesamten dreidimensionalen Gebiet geliefert werden. Daher soll nur durch Simulationen die Reynoldszahlabhängigkeit der lokalen Strömungsfelder untersucht werden.

Bei der Auswahl von Ebenen für die Messungen soll ebenfalls auf die Simulationsergebnisse zurückgegriffen werden, um gezielt solche Stellen zu identifizieren, die charakteristische Besonderheiten erwarten lassen.

Sowohl experimentell als auch numerisch wird also eine breite Datenbasis für die weitere Analyse der Strömungsfelder und daraus bestimmbarer Größen (z.B. radiale und axiale Profile, Häufigkeitsverteilungen, innerer Längenskalen) bereitgestellt.

Häufigkeitsverteilungen der Geschwindigkeit eines Newtonfluids zeigen für Lochplattenanordnungen (Abb. 3) in den Lücken (d.h. Ebenen zwischen den Loch-

blechen) für positive Geschwindigkeiten einen exponentiell abfallenden Verlauf, ähnlich wie dies in Kugelschüttungen zu beobachten ist. Ein kleiner Anteil des Gesamtvolumenstroms strömt entgegen der Hauptströmungsrichtung (negativer Geschwindigkeitsanteil).

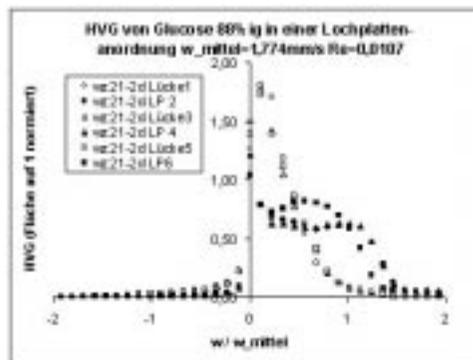


Abbildung 3: Experimentell ermittelte Geschwindigkeitshäufigkeitsverteilung in einer Lochplattenanordnung.

Um aus einer Analyse von Strömungsfeldern innere Längenskalen zu ermitteln, soll systematisch untersucht werden, welche Information in der Messvolumen-größenabhängigkeit der Geschwindigkeitshäufigkeitsverteilung steckt: Wenn das Messvolumen klein ist im Vergleich zu den Hohlraumabmessungen der Schüttung, sollte sich eine Dirac-Funktion ergeben. Bei sehr großen Volumina erhält man nach Literaturangaben (und nach Messungen der Arbeitsgruppe von Prof. Buggisch) eine abklingende Exponentialfunktion. In der Messvolumen-größenabhängigkeit der Geschwindigkeitshäufigkeitsverteilung müsste also Information über innere Längenskalen der Packung stecken. Die spannende Frage, welcher Art diese Information ist und wie man sie praktisch nutzen kann, soll durch Simulation und Experiment untersucht werden.

Außer zur Messzeitersparnis eignen sich die komplementäre numerische Untersuchungen aber insbesondere auch, um ideales Verhalten und reale Artefakte gegenüberzustellen. So ist beispielsweise langfristig zu erwarten, dass durch den Vergleich von idealisiertem Verhalten in der Simulation und realem Verhalten im Experiment die Bedeutung der Berücksichtigung nicht-newtonschen Verhaltens oder der Einfluss von Gedächtnis-Effekten quantifiziert werden kann. Bereits jetzt werden in der NMR-Anlage in Karlsruhe nicht-newtonsche Fluide erfolgreich vermessen. Um nicht-newtonsches Verhalten in der Simulation berücksichtigen zu können, sind noch umfangreiche Erweiterungen des implementierten Lattice Boltzmann Modells nötig. Dies könnte auch Gegenstand einer längerfristigen Weiterführung des Projekts sein, wobei jedoch aus heutiger Sicht *nicht* zu erwarten ist, dass komplexes nicht-newtonsches Verhalten, insbesondere Gedächtnis-Effekte leicht in die Simulation zu integrieren sind. An dieser Stelle sind die experimentellen Untersuchungen den Simulationen sicherlich ein gutes Stück voraus.

thomas.zeiser@lstm.uni-erlangen.de

- [1] G. Brenner, T. Zeiser & F. Durst. Simulation komplexer fluider Transportvorgänge in porösen Medien. *Chem.-Ing.-Tech.*, **74** (2002) 1533–1542.
- [2] P. Callaghan. *Principles of NMR Microscopy*. Oxford Univ. Press, 1993.
- [3] K. Rottschäfer. *Geschwindigkeitsverteilungen in durchströmten Füllkörperschüttungen*. Doktorarbeit, TUM, 1996.
- [4] M. Sahimi. *Flow and Transport in Porous Media and Fractured Rock*. VCH Verlagsgesellschaft, Weinheim, 1995.
- [5] M. Steven. Detaillierte Simulation und Analyse der Struktur von Katalysatorschüttungen und lokaler Transportprozesse. Dipl.Abeit, Uni-Erlangen, 2001.

Für die Pentium- und Itanium-Cluster wurden im Übrigen getrennte LINPACK-Leistungsmessungen vorgenommen; mit 350 GFlop/s konnte sich das IA32-Segment sogar noch auf Platz 353 der aktuellen TOP500-Liste platzieren. 64 CPUs des Itanium-Clusters erreichten eine LINPACK-Leistung von 260 GFlop/s. Die Spitzenrechenleistung des Gesamtclusters übertrifft jetzt 1 TeraFlop/s.

### Betriebsmodell und Systemsoftware am LRZ-Cluster

Alle Pentium-4 Knoten werden zumindest in der Anfangsphase als Pool für rein serielle Durchsatz-Jobs betrieben. Hierbei dient die bisher schon bewährte Sun Grid-Engine (SGE) als Batch-Queueing System. Zehn der Knoten stehen hierbei für bis zu zwei Wochen lang laufende Jobs zur Verfügung; solche Jobs haben dann den Knoten für sich allein. Alle Übrigen werden mit je zwei Jobs aufgefüllt, die bis zu 6 Tagen dort verbleiben können (das entspricht etwas mehr als 3 CPU-Tagen).

Ein Rechenknoten des Itanium-Clusters steht als Interaktivmaschine – insbesondere also zum Compilieren und Testen von Programmen – zur Verfügung; alle längeren Rechnungen und auch parallele Programmläufe sollten jedoch grundsätzlich ebenfalls als Batch Jobs über die SGE abgesetzt werden. Zwei der sechzehn Batch-Knoten werden hierbei grundsätzlich nur für serielle Jobs eingesetzt; auf den übrigen können parallele Programme auf bis zu 56 Prozessoren laufen. Für Programme, die den Myrinet-Interconnect verwenden, ist die Rechenzeit auf 48 Stunden begrenzt, alle anderen dürfen bis zu 6 Tagen im Cluster verbleiben.

Aus Sicherheitsgründen ist der interaktive Zugang zum Cluster nur über verschlüsselte Verbindungen (“Se-

cure Shell”) möglich; insbesondere ist der Zugriff vom Cluster ins Internet (“http-Protokoll”) unterbunden.

Das gesamte Cluster wird unter dem freien Betriebssystem Linux betrieben, wobei im IA32-Segment die von der Nürnberger Firma SuSE hergestellte Distribution zum Einsatz kommt, auf dem Itanium-Cluster die von Red Hat (USA). Auf Grund der Besonderheiten der LRZ-Umgebung – zum Beispiel der Einsatz des verteilten Dateisystems AFS – und aus Gründen der Betriebsstabilität des Gesamtclusters werden jedoch LRZ-intern gebaute Linux-Kernel eingesetzt. Aus Benutzersicht von größerem Interesse sind die Compiler und Tools, welche für die Programm-Entwicklung zur Verfügung stehen. An erster Stelle steht hier die Compiler-Suite von Intel, die insbesondere im Bereich der Fortran-Programmierung inzwischen das Produkt der Wahl ist; darüber hinaus ist auch das neueste Release 3.3.1 der GNU Compiler Collection (gcc) installiert worden. Noch nicht verfügbar sind modernste Tools zur Optimierung, Profilierung und automatischen Parallelisierung von Programmen; insbesondere auf der Itanium-Architektur dürfte es noch mindestens ein halbes Jahr dauern bis hier die ersten fertigen Produkte installiert sind. Schon jetzt verfügbar sind die sogenannten “Math Kernel Libraries” der Firma Intel, die für Grundoperationen aus der linearen Algebra, der Fourier-Transformation und Signal-Verarbeitung speziell auf die jeweiligen Architekturen optimierte Bibliotheken bereitstellt. Darüber hinaus sind vom LRZ eine ganze Reihe von frei verfügbaren Bibliotheken auf dem Cluster installiert worden; für den Bereich des wissenschaftlichen Rechnens ist insbesondere die GNU Scientific Library (GSL) inzwischen ein ernsthafter Kon-

Systemen gab es technisch bedingte Verzögerungen, dafür lieferte MEGWare im August 2003 Systeme mit der neuesten Itanium 2-Generation (mit 1.3 GHz getaktete "Madison"-CPUs) aus.

#### Leistungsdaten der neuen Systeme

Im Unterschied zum bisherigen Cluster sind die neuen Systeme in einem Rack montiert; so zeigt Abbildung 1 die drei Schränke mit jeweils 30 Pentium-4 Knoten, jeweils zwei Knoten sind nebeneinander in einem eine Höheneinheit messenden Einschub untergebracht.



Abbildung 1: Drei MEGWare Racks mit Pentium-4 Knoten

Unten im Rack befinden sich neben einem Gigabit-Ethernet Switch, der für die Netzwerk-Verbindung der Knoten untereinander und mit der Außenwelt sorgt, auch der sogenannte ClustSafe-Switch, mit dem sich jeder Knoten einzeln elektrisch fernsteuern (also etwa herunter- oder hochfahren) lässt. Die Pentium-4 Systeme sind mit 3.06 GHz getaktet und erreichen aus dem Hauptspeicher Bandbreiten von bis zu 3.0 GByte/s (formale Spitzenleistung ist 4.2 GByte/s). Die 17 Itanium-2 Systeme der neue-

sten Generation sind in zwei Racks untergebracht; jeder 4-Wege Multiprozessor benötigt vier Höheneinheiten, sodass maximal 8 bis 9 Boxen in ein Rack passen, wenn auch noch die verbleibende Infrastruktur untergebracht werden soll. Über den Myrinet-Interconnect wird später noch zu reden sein; für die Verbindung zur Außenwelt sind auch die Itanium-Rechner mit Gigabit Ethernet ausgestattet. Alle neuen Systeme enthalten pro CPU 2 Gigabyte Hauptspeicher; insbesondere kann man auf den Itanium-Systemen jetzt für ein einzelnes Programm ggf. deutlich mehr als 2 GByte Hauptspeicher anfordern.

Nicht vergessen werden sollte auch der Hintergrund-Speicher: Sechs leistungsfähige IA32-basierte Fileserver in Verbindung mit der parallelen Dateisystem-Software "PVFS" (parallel virtual file system) sorgen dafür, dass sowohl serielle als auch parallele Programme auf dem Linux-Cluster ihre Ausgabedaten mit hoher Bandbreite sowohl herauschreiben als auch wieder einlesen können und alle Daten global auf den jeweiligen Clustern zugänglich sind. Insbesondere auf den Itanium-Knoten erhält man wegen der Myrinet-Anbindung I/O Bandbreiten von bis zu mehr als 500 MByte/s. Und auch für seriellen Zugriff erhält man eine I/O-Leistung, die sehr deutlich oberhalb des für eine lokale Platte erreichbaren Wertes liegt. Insgesamt stehen – nach Abzug der für einen sicheren Betrieb notwendigen Reserven – im Moment etwa 1.5 TByte an Netto-Plattenplatz für das gesamte Linux-Cluster zur Verfügung; jedoch ist die Lebensdauer der Daten dort durch eine Hochwasserlöschung begrenzt. Diesem Mechanismus kann der Benutzer jedoch durch gezieltes Archivieren längerfristiger benötigter Daten entgegenwirken.

## Marie-Curie Research-Training Network Computational Seismology

Die Gruppe im Bereich computational seismology des Departments für Geo- und Umweltwissenschaften an der Ludwig-Maximilians-Universität München wurde von der EU in Brüssel zu Vertragsverhandlungen bezüglich eines Marie-Curie Research Training Networks im Rahmen des FP6-Programms eingeladen. Wenn alles gut geht, wird das Netzwerk mit dem Namen SPICE (Seismic wave Propagation and Imaging in Complex media: a European network) im Januar 2004 gestartet. Am Netzwerk mit einer Laufzeit von 4 Jahren sind 14 Institute (u.a. in Zürich, Paris, Rom, Oxford, Trieste, Prag, Dublin, Oslo) mit Schwerpunkten in geologischen, physikalischen, mathematischen und rechen-technischen Aspekten der numerischen Wellenausbreitung beteiligt, wobei die Münchner Gruppe die Koordination übernimmt.

Ziel dieses Netzwerks ist zum einen, die 3D Methodik zur Wellenausbreitung in allen Skalenbereichen (Laborskala, Sedimentbecken zur Simulation von

Erdbebenszenarien, Vulkane, Planeten etc.) weiterzuentwickeln, zum anderen diese Methodik in das Inversionsproblem mit einzubinden, um zum Beispiel die räumliche Auflösung tomographischer Images des Erdinnern zu verbessern. Ein zentrales Element soll die Entwicklung gemeinsamer Standards zur Datenvisualisierung, Verifikation, mass-storage, online Datenaccess usw. für die Seismologie sein.

Die Forschung, die im Rahmen des KONWIHR Projekts NBW (Numerische Berechnung des seismischen Wellenfeldes) gemacht wurde, zusammen mit den exzellenten Rahmenbedingungen, die München für die Koordination eines solchen Netzes bietet, waren mit Sicherheit für die positive Begutachtung des Antrags von Bedeutung. Der Koordinator möchte sich in diesem Sinne für die Unterstützung durch das KONWIHR Programm in den letzten Jahren besonders bedanken!

Prof. Dr. Heiner Igel  
Sektion Geophysik, Department für Geo- und Umweltwissenschaften, LMU München

---

### Jubiläumskurs zum Vormerken ...

#### Kurzlehrgang NUMET 2004

Nach dem großen Erfolg in den vergangenen Jahren wird der Kurzlehrgang NUMET (Numerische Methoden zur Berechnung von Strömungs- und Wärmeübertragungsproblemen) auch 2004 wieder in Erlangen stattfinden. Der Lehrstuhl für Strömungsmechanik an der Universität Erlangen-Nürnberg, führt diesen Kurs unter Mitwirkung von KONWIHR-Mitarbeitern nun bereits zum zehnten Mal (*Jubiläum !!!*) durch. Auch für NUMET 2004 konnten inter-

national anerkannte Experten auf dem Gebiet der Numerischen Strömungsmechanik (u.a. Prof. Bungartz, Prof. Durst, Prof. Lerch, Prof. Perić, Dr. Menter, Prof. Schäfer) für die Vorlesungen gewonnen werden. Die Vortragenden besitzen Erfahrung sowohl in der Grundlagenforschung zur Entwicklung von Berechnungsverfahren als auch in der Erstellung und Anwendung von Software für die industrielle Praxis. Zwei der Vortragenden sind Leiter der Entwicklungsabteilungen führender CFD-Software-Anbieter.

Der Kurzlehrgang gibt eine Einführung in die numerische Lösung von

Strömungs- und Wärmeübertragungsproblemen. In 21 einstündigen Einzelvorlesungen und zusätzlichen Computerdemonstrationen werden die Grundlagen der Finite-Volumen-, der Finite-Differenzen-, der Finite-Elemente- und der Lattice-Boltzmann-Berechnungsverfahren vermittelt und verschiedene Anwendungsmöglichkeiten vorgestellt. In den Vorträgen werden alle Schritte von der Diskretisierung über den Einbau der Randbedingungen bis zur gekoppelten Lösung der Erhaltungsgleichungen behandelt. Neben den zugrundeliegenden Ideen werden die gebräuchlichen Approximationsmethoden und Diskretisierungsschemata im Detail beschrieben. Ergänzend zur Erläuterung der Grundgleichungen der Strömungsmechanik wird die Modellierung und numerische Implementierung typischer Phänomene wie Turbulenz oder Wärme- und Stoffübertragung dargestellt. Ferner werden zwei Vorlesungen über gekoppelte Probleme (Fluid-Struktur-Wechselwirkung und Aeroakustik) angeboten.

Um aufzuzeigen, wie die Effizienz der zum Teil sehr aufwendigen Strömungsberechnungen erhöht werden kann, wird die Mehrgittertechnik vorgestellt und eine Einführung in die Diskretisierung mittels dünner Gitter gegeben. Eine besondere Bedeutung kommt in diesem Zusammenhang der Portierung von Strömungsprogrammen auf Hoch- und Höchstleistungsrechner zu, die aus diesem Grund in einer eigenen Vorlesung behandelt wird. Das technisch-wissenschaftliche Hochleistungsrechnen wird vielfach als Voraussetzung dafür angegeben, komplexe dreidimensionale Strömungssimulationen in Rechenzeiten durchzuführen, die für die industrielle Praxis akzeptabel sind. Die Weiterentwicklung und Etablie-

rung dieser Disziplin wird an der Universität Erlangen-Nürnberg von Wissenschaftlern verschiedener Fachrichtungen verfolgt, um Anwendern aus Forschung und Industrie geeignete Hilfsmittel zur Verfügung stellen zu können. Die aus diesen Arbeiten resultierenden Ergebnisse und der aktuelle Wissensstand über die Eigenschaften und die Einsetzbarkeit verschiedener Rechnerarchitekturen werden u.a. Thema der im Rahmen des Lehrgangs stattfindenden Computerdemonstrationen sein.

Der Kurzlehrgang wendet sich an all diejenigen, die selbst Rechenprogramme für Strömungssimulationen entwickeln oder aber kommerzielle Softwarepakete anwenden. Programmentwickler finden durch die detaillierte Behandlung aller Einzelschritte, die zur Programmerstellung notwendig sind, wertvolle Hilfestellung. Anwender haben die Gelegenheit, die wesentlichen Bestandteile eines kompletten Berechnungsverfahrens und dessen Funktionsweise kennenzulernen; dies ist sowohl für die Anwendung als auch für die kritische Interpretation der Ergebnisse von Softwarepaketen unverzichtbar. Besonders nützlich und reizvoll für alle Teilnehmer sollte nicht zuletzt die Möglichkeit sein, durch Diskussionen mit den Vortragenden Informationen über die neuesten Entwicklungen auf dem Gebiet der *Numerischen Strömungsmechanik* aus erster Hand zu bekommen. Weitere Infos und Anmeldeformulare finden Sie auf der unten angegebenen Web-Page.

Termin 8.–11. März 2004  
 Ort LSTM Erlangen  
 Anmeldung ab sofort  
 (bis 16.1.2004 Preisnachlaß)  
 Leitung Priv.-Doz. Dr. M. Breuer  
[numet2004@lstm.uni-erlangen.de](mailto:numet2004@lstm.uni-erlangen.de)  
[www.lstm.uni-erlangen.de/numet2004](http://www.lstm.uni-erlangen.de/numet2004)

## Münchener Prozessorgeflüster

### Erste Erfahrungen mit dem Itanium-2 Cluster & den neuen Pentium-4 Systemen am LRZ

von Dr. Reinhold Bader  
 (Hochleistungssysteme am LRZ)

#### Motivation & Beschaffungsverfahren

Etwa ein Jahr nach Inbetriebnahme des Höchstleistungsrechners in Bayern, einer Hitachi SR8000 mit einer Spitzenrechenleistung von mehr als 1 TFlop/s, bekamen wir Besuch von einigen Quantenchemikern. Sie waren sehr unzufrieden.

– *Dieser neue Rechner, den ihr da habt ... wir verstehen nicht, was ihr da gekauft habt. Für uns ist er unbrauchbar ...*

– *Stimmt. Wir haben ihn ja auch nicht für euch gekauft!*

– *Das finden wir aber nicht gut! Wir haben nämlich jede Menge Bedarf an Rechenzeit. Jetzt solltet ihr auch mal für uns etwas tun, sonst ...*

– *Na gut. Wir sehen zu, dass wir für euch etwas Geeignetes herbekommen.*

Der nachfolgende Bericht möchte schildern, was das LRZ in Folge der obigen Unterhaltung unternommen hat, um die artikulierten Bedürfnisse abzudecken. Mitte 2001, nach Abschaltung des IBM SP2 Parallelrechners (dessen Ressourcen insbesondere an Hauptspeicher schon längst nicht mehr den Anforderungen der Nutzer genügten) war auch aus LRZ-Sicht klar, dass im Bereich des sogenannten *Durchsatz-Rechnens*, bei dem zahlreiche zumeist serielle Rechnungen etwa im Rahmen von Parameter-Studien durchzuführen sind, eine Versorgungslücke entstanden war. Diese Lücke konnte durch das damals schon verfügbare Linux-Cluster nur unzureichend abgedeckt werden. Im Rahmen eines HBFVG Beschaffungsverfahrens konnten daraufhin mehr als 1 Million Euro an Investitionsmit-

teln erfolgreich beantragt werden; auf Grund der Anforderungs-Profile aus den Münchner Forschungsinstituten ermittelte das LRZ folgende Zielvorstellung für die Erweiterung des bestehenden Linux-Clusters:

- 90 Einzel-CPU Systeme mit möglichst schnell getakteten IA32-kompatiblen Prozessoren (“Los 1”)
- 16 4-Wege Shared-Memory Systeme mit schnellem Myrinet-Interconnect für mäßig parallele Programme (“Los 2”)
- Etwa 2 TByte hochwertiger (d. h. breitbandig anbindbarer) Hintergrund-Plattenspeicher für die intermediäre Datenthaltung (“Los 3”)

Während im seriellen Bereich schon bald klar war, dass nur die Pentium-4 Architektur von Intel in der Lage war, die hohen Anforderungen des LRZ an Spitzenleistung und Speicherbandbreite zu erfüllen, sorgte die Preisentwicklung bei den modernen 64-Bit SMP-Systemen im Laufe des Sommers 2002 für Überraschungen: Das LRZ sah sich veranlasst, kurzfristig seine Ausschreibung von Xeon-basierten IA32 4-Wege Prozessoren auf die neuen Architekturen umzustellen; damit kamen auch mehr Kandidaten als bislang erwartet zur Konkurrenz: In erster Linie die Itanium-Architektur von Intel, daneben aber auch die neue Power4-Architektur von IBM sowie potentiell auch der x86-64 Prozessor (“Hammer”) der Firma AMD.

Die nach Leistungs/Preisverhältnis bewertete Ausschreibung wurde dann für die Lose 1 und 2 von der Firma MEG-Ware Computers (die im 64-bit Bereich ein Itanium-Angebot machte) gewonnen; im Bereich der Storage-Lösung trug die Firma Advanced UniByte den Sieg davon. Noch im Frühjahr 2003 konnten dann die IA32-Systeme in Betrieb genommen werden; bei den Itanium-