

# Abschlußbericht der Projekte Multiskalen Moleküldynamik und Krustenfluide

Prof. Dr. Michael Griebel, Prof. Dr. Horst Neugebauer  
Dr. Gerhard Zumbusch

25. August 2000

## 1 Beteiligte Einrichtungen

- Universität Bonn, Institut für Angewandte Mathematik, Abt. Wissenschaftliches Rechnen (Prof. Dr. Michael Griebel)
- Universität Bonn, Institut für Geodynamik (Prof. Dr. H. J. Neugebauer)
- GMD, St. Augustin
- Forschungszentrum Jülich (Dr. Thomas Eickermann)

## 2 Zusammenfassung

Im Rahmen des Gigabit-Testbed-West wurden vom 1. Dezember 1998 bis zum 30. Juni 2000 seitens der Universität Bonn in Zusammenarbeit mit der GMD und dem Forschungszentrum Jülich die beiden Anwendungsprojekte “Multiskalen Moleküldynamik” und “Krustenfluide” realisiert. Beiden Teilprojekten unterliegt die gemeinsame Idee relevante Fragestellungen der einzelnen Forschungsgebiete mittels komplexer Computersimulationen zu modellieren. Hierbei wurden speziell Probleme betrachtet, die auf herkömmlichen Rechnerplattformen nicht oder nur mit Einschränkungen lösbar sind und somit ein verteiltes Rechnen auf einer Metacomputing-Plattform erfordern.

### **Multiskalen Moleküldynamik**

Im Projekt Multiskalen Moleküldynamik wurde die Dynamik von Makromolekülen untersucht. Als zugrundeliegendes Modell wurde die klassische Moleküldynamik verwendet. Hierbei wurden jedoch insbesondere bei der Betrachtung neuer Materialien, wie z.B. den Kohlenstoff-Nanoröhren, moderne bond-order Potentiale implementiert, die in der Lage sind, über die bisherigen Eigenschaften

klassischer Moleküldynamik hinaus auch Bindungsbrüche und Quanteneffekte zu beschreiben. Jedoch weisen diese Potentiale eine sehr viel komplexere Struktur auf als klassische Potentiale. Dies führt zu einem erhöhten Berechnungsaufwand, zumal diese Potentiale noch gekoppelt werden müssen mit den herkömmlichen van der Waals Potentialen für die intermolekularen Wechselwirkungen. Hinzu kommt, daß die betrachteten Stoffe materialimmanente Wellenlängen besitzen, die dazu führen, daß realistische Simulationen nur mit sehr großen Systemen im Bereich von mehreren Millionen Atomen durchgeführt werden können. Ähnliches gilt im Bereich der Biomoleküle. Die hier interessanten Fragestellungen, wie z.B. die Proteinfaltung oder das Schlüssel-Schloß Prinzip im Drug-Design, können nur durch Kopplung unterschiedlichster Potentiale, die teilweise auf unterschiedlichen Zeitskalen interagieren, modelliert werden. Ziel des Projektes war es diese unterschiedlichen Wechselwirkungen zu entkoppeln und auf verteilten Rechnern isoliert voneinander zu berechnen, um dann in regelmäßigen Kommunikationsschritten die Daten untereinander auszutauschen und somit das Gesamtergebnis der Simulation zu berechnen. Die Implementierungen konnten im Rahmen des Projektes erfolgreich umgesetzt werden. Mit Hilfe der entwickelten Programme konnten dann im Bereich der Kohlenstoff-Nanoröhren, der Biomoleküle und der Rißausbreitung in Metallen erfolgreich Simulationen durchgeführt werden, die experimentelle Systemgrößen im Bereich von mehreren hundert Millionen Atomen erreichten.

## **Krustenfluide**

Ziel in diesem Teilprojekt war die Erstellung eines verteilten Programmpaketes, das sich aus einzelnen Modulen zur Berechnung der Elastizitäts-, Strömungs-, Temperatur- und Stoffkonzentrationsgleichungen zusammensetzt. Die Gesamtlösung ergibt sich dann durch die geeignete (schnelle) Kopplung der einzelnen Module. Die unterschiedlichen Anforderungen an die Einzellöser machen es notwendig, hierzu jeweils optimale Rechnerarchitekturen einzusetzen. Im Rahmen der numerischen Modellexperimente sollten grundsätzliche Aussagen zum Systemverhalten der hochgradig gekoppelten Komponenten ermittelt werden. Dies betrifft die Berechnung von Spannungsfeldern in der Stoffmatrix, die durch Deformationen, thermische Einflüsse und chemische Wechselwirkungen modifiziert werden. In Bezug auf das Fluid interessieren Strömungsfelder, Temperatur- und Konzentrationsverteilungen. Aus den bisherigen Erfahrungen ist das Auftreten von Bifurkationen sowie Phasen der Oszillation in der Lösung von großem wissenschaftlichen Interesse. Hierüber werden Zugänge geschaffen zum Verständnis thermaler und hydrothormaler Systeme, chemischer Fluid-Gesteinswechselwirkungen sowie den Prozessen der langzeitigen Kompaktion und episodischen Erdbeben, sogenannten Schwarmbeben oder auch induzierten Beben. Analogieschlüsse und Korrelationen erfordern eine entsprechende Auswertung der Modellexperimente. Während das System-Zeitverhalten durch dimensionslose Kennzahlen etwa der

Rayleigh- oder Nusseltzahl erschließbar ist, erfordern die errechneten Felder der Matrixspannung, Temperatur- und Stoffkonzentration und der Strömung entsprechende dreidimensionale Repräsentation. Die erforderlichen Bildsequenzen richten sich nach der erforderlichen Auflösung der Prozeßbeschreibung, damit geht die Auswertung der Experimente in Richtung Animation. Strömungsvorgänge in der Erdkruste sind per se als ein Prozeßgefüge anzusehen, d.h. die physikalischen und chemischen Variablen des Systems sind stark untereinander gekoppelt. Das natürliche System ist offen. Eine Quantifizierung des Systemverhaltens ist jedoch notwendige Grundlage für jede zeitliche und räumliche Bewertung anthropogener Initiativen und Eingriffe. Dies betrifft u.a. die Dynamik von Reservoiren des Wassers und der Kohlenwasserstoffe wie Öl und Gas, die Anlage von unterirdischen Speichern und Kavernen und letztendlich die Verpressung von Schadwässern als Entsorgungsmaßnahme mit ungewissem Verbleib. Das Phänomen der Absenkung der Landoberfläche über einem ausgebeuteten Gasreservoir stellt beispielsweise in einem vom Einfluß des Meerwassers bedrohten Gebiet eine brisante Problematik dar.

Nach neuestem Forschungsstand ist die Zone der Wechselwirkungen auf die oberen 15 km der Erdkruste zu beziehen. In diesen Bereich dringen Oberflächenwasser ein, von unten steigen Gase und thermale Lösungen auf. Das gesamte System ist Deformationen unterworfen, die von der Langzeittektonik - mit einer Rate von Millimetern pro Jahr - bis hin zu kurzperiodischen Kräften der Erdzeiten reichen. Die Fülle der natürlichen Phänomene seitens der Kräfte der Strömungs- und Gesteinsmechanik und des Stoff- und Energiehaushalts definieren ein Grundkonzept, das neben den strömenden Fluiden die umgebende Stoffmatrix sowie den Energie- und Stoffaustausch auf Flüssigkeit und Matrix bezogen beinhaltet.

Von Seiten der numerischen Mathematik konnte das gekoppelte System Matrix-Fluid-Temperatur-Stoffkonzentration erfolgreich auf drei Dimensionen erweitert und auf dem verteilten System implementiert werden. Um die hierbei auftretenden Probleme durch die sehr großen Anforderungen an Speicherplatz und Rechenzeit zu bewältigen wurden Mehrgitterverfahren eingesetzt. Hierbei wurde insbesondere die Verwendung von robusten Mehrgittermethoden, wie Mehrgitterverfahren mit matrixabhängigen Prolongationen sowie algebraischen Mehrgitterverfahren untersucht. Eine weitere für den 3D Fall notwendige Effizienzsteigerung ist durch die Parallelisierung der Verfahren erreichbar. Hierbei wurde insbesondere das Zusammenspiel der Kopplungen der einzelnen Systemkomponenten genauer betrachtet. Wichtige Anwendungen im Bereich der Wechselwirkungen Boden-Pflanze sowie Hangstabilitätsmodelle konnten umgesetzt werden. Diese werden im Abschnitt 5 detailliert beschrieben.

### 3 Anschluß an das Testbed-West

An der Universität Bonn wurden nach Inbetriebnahme der ATM Verbindung zur GMD (St. Augustin), deren Verlängerung zum Forschungszentrum Jülich bereits länger als OC-48 Verbindung besteht, Cisco ATM Switches an der Zwischenstation des Instituts für Informatik und am Endpunkt des Institutes für Angewandte Mathematik (IAM) installiert und in Betrieb genommen. Die Anbindung des Parnass2 PC-Clusters am IAM kann grundsätzlich über zwei Wege erfolgen: Jeder Knoten wird über Ethernetverbindungen einzeln mit dem Switch verbunden, was einer aggregierten Bandbreite von 7.2 GB/s entspricht, wo dann die Umsetzung auf LAN-Emulation mit 1 GB/s und die Weiterleitung mit OC-48 erfolgt. Alternativ kann ein Gateway-Rechner des IAM genutzt werden, der mit OC-12 an den IAM ATM Switch angeschlossen ist und sowohl LAN-Emulation als auch Classical IP beherrscht. In Jülich sind die entsprechenden Cray Rechner jeweils über HIPPI/ATM Gateway-Rechner mit OC-12 Anschlüssen versehen, womit sich für alle Wege eine theoretische nutzbare Gesamtbandbreite von OC-12 mit etwa 400 MB/s ergibt. Eine Übersicht über die Anbindung findet man in den Abb. 1 und 2.

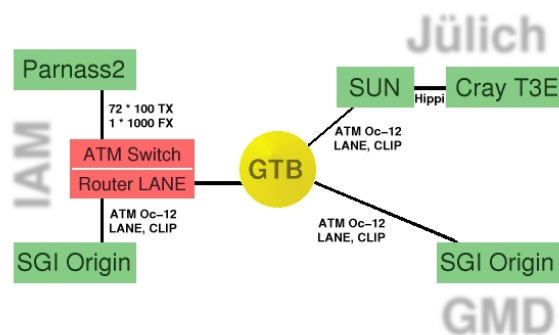


Abbildung 1: Darstellung der Anbindung der Uni-Bonn an das Gigabit Testbed

### 4 Testbed und Metacomputing

Aus einem Projekt des Argonne Nat. Lab. wurde das Metacomputing Toolkit Globus für sichere Authentifizierung, Scheduling und Kommunikation verwendet und die darauf aufbauende heterogene MPI Implementierung Mpich-G. Damit ergibt sich eine für die Anwenderprogramme konsistente Sicht des gekoppelten Meta-Computers. Nach Anfangsschwierigkeiten mit der Installation auf den Crays konnten dann verschiedene Messreihen sowohl Rohdaten als auch mit Anwen-

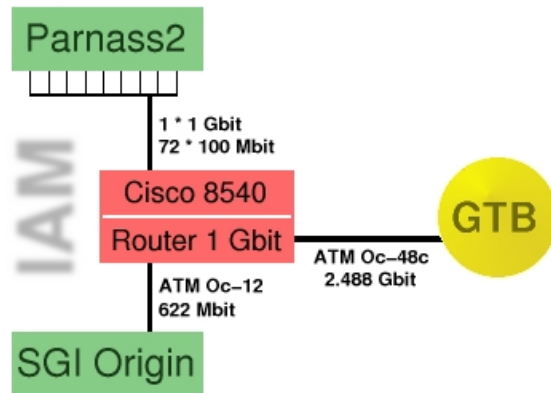


Abbildung 2: Darstellung der Anbindung der Uni-Bonn an das Gigabit Testbed

dungsprogrammen durchgeführt werden. Während sich die ATM Verbindungen zwischen IAM Rechnern und GMD sowie Gateway Rechnern des FZ Jülich als stabil und recht performant erwiesen, ergaben die letzten Meter zur Cray auffällig geringere Werte. Dabei zeigte sich für Rohdaten zwischen LAN-Emulation und Classical IP deutlich geringere Differenzen (290 Mb/s CLIP zu 300 MB/s LAN-E) als zwischen den verschiedenen ATM-IP Umsetzern (Switch 340 Mb/s zu Gateway-Rechner 140 Mb/s). Bei MPI Anwendungen fiel der erwartete Overhead relativ hoch aus, sowohl zwischen Parnass2 und anderen IAM Computervern (290 Mb/s zu 56Mb/s) als auch zur Cray (bis zu 1Mb/s). Leider konnten die auffälligen Schwankungen der Messungen nach Jülich nicht mehr untersucht werden, da die dortigen Projektmittel bereits verausgabt waren und die Leitung GMD-Jülich später ganz außer Betrieb ging. Damit wurden eine Reihe weiterer gekoppelter Simulationen über ATM Verbindungen innerhalb des IAM realisiert. Zunächst aufgetretene Probleme mit der Stabilität des Globus Metacomputing-Directory-Service sind in der aktuellen Globus Version 1.3 durch eine Dezentralisierung eliminiert worden, wobei diese Version aus erwähnten Gründen nicht mehr zum Einsatz kam.

Da das Globus System auf der CRAY T3E des Forschungszentrums Jülich und der IBM SP2 der GMD zu Beginn des Projektes noch nicht verfügbar war, mußten sich unsere ersten Tests auf die Systeme im Institut für Angewandte Mathematik beschränken. Hierbei handelt es sich um verschiedene Origin 2000 Systeme von SGI und den Parallelrechner Parnass2 der aktuell auf Platz 362 der TOP 500 Liste platziert ist. Nähere Informationen zu Parnass2 sind in Abschnitt 4.3 zu finden.

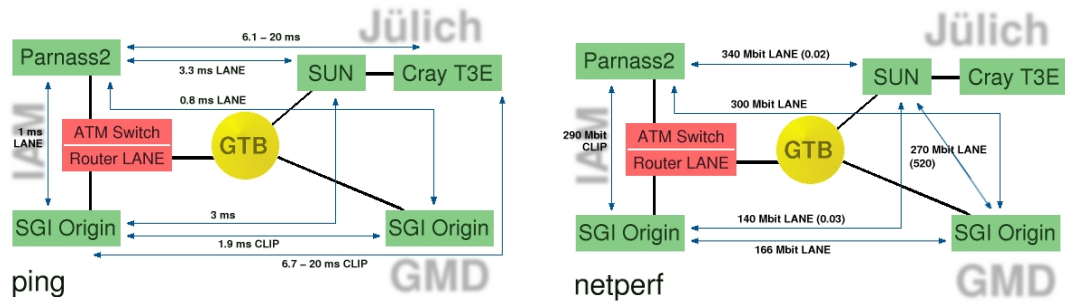


Abbildung 3: Latenzen (links mit ping) und Bandbreiten (rechts mit Netperf) im GTB.

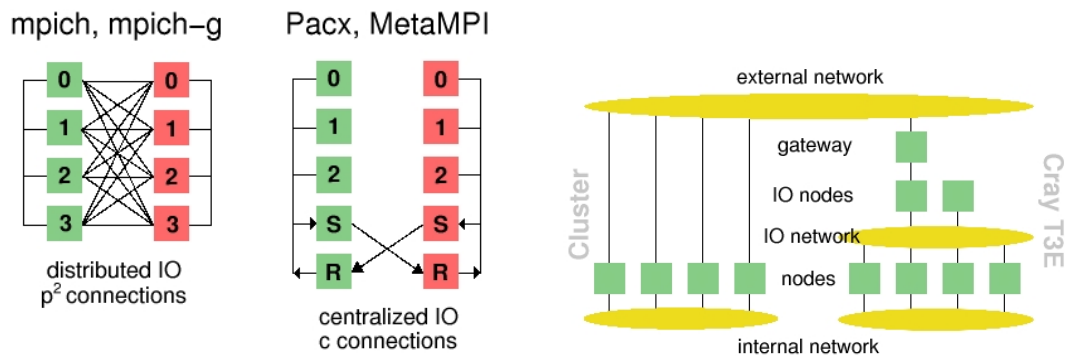


Abbildung 4: Strukturen konkurrierender globaler MPI Implementierungen

## 4.1 Netz-Performance

Durch den verspäteten Anschluß an das Gigabit stand für erste Messungen lediglich der Zeitraum 1. Jahreshälfte 2000 zur Verfügung, dabei ergaben sich folgende Werte:

Zunächst haben wir die Leistung von TCP/IP zwischen den beteiligten Rechnern gemessen, um die Netzleistung für das spätere Metacomputing abschätzen zu können.

Wir haben einerseits Latenzzeiten mit ping zwischen Rechnern und andererseits Bandbreiten mit Netperf gemessen (siehe Abb. 3). Die Latenzen waren im wesentlichen beeinflusst von der Geschwindigkeit der Router/Switche und Umsetzer sowie natürlich der physikalischen Signallaufzeiten. Dabei zeigte sich, daß lokal die Umsetzung von Ethernet auf LANE etwa 1 ms dauert, ähnliche Zeiten ergaben sich auch zur GMD, wobei deutlich wurde, daß die SGI Origin meßbar langsamer

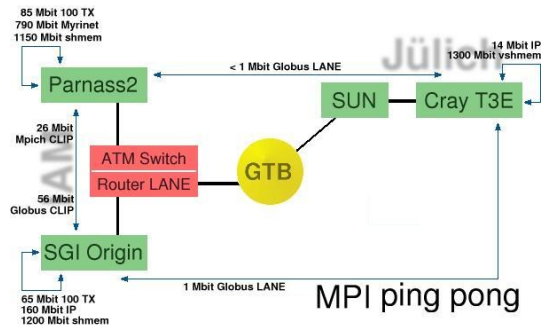


Abbildung 5: Bandbreiten für globales MPI im GTB

war als unser Router. Die Zeiten nach Jülich entsprachen der physischen Entfernung bei mindestens 3 ms zum Gateway-Rechner. Wirklich bemerkenswert waren allerdings die Messungen mit der CRAY T3E, bei der starke Schwankungen zwischen 6 und 20 ms beobachtet wurden, was auf eine starke Belastung des CRAY I/O-Knoten schließen läßt. Das bedeutet, daß der größte Anteil an der Latenz durch die CRAY verursacht wird und ein kleinerer Teil durch die Glasfaserstrecke und die verwendeten ATM-Switche.

Die Bandbreitenmessungen haben wir mit Netperf durchgeführt, um eine obere Abschätzung für die in Anwendungen maximal mögliche Bandbreite des Netzes zu erhalten. Die durchgängige OC-12 Leitung mit einer Nominalbandbreite von 622 MBit/sec. liefert unter TCP/IP protokollabhängig bis zu 400 MBit/sec. Wir waren allerdings nicht in der Lage selbst zwischen den Rechnern der GMD und Jülich diese Werte zu reproduzieren. Die besten Datenübertragungsraten hatten wir zwischen dem Parnass2 System und dem CRAY Gateway-Rechner mit bis zu 340 MBit/sec. gemessen, die allerdings lastabhängig starken Schwankungen unterworfen waren. Interessanterweise hatten wir dazu LANE und nicht CLIP verwendet. Was man auch noch deutlich erkennen kann ist, daß unsere SGI-Origin etwas geringere Datenübertragungsraten aufweist, obwohl wir für die langen Distanzen die TCP Puffergrößen optimiert hatten.

## 4.2 Globus

In Tabelle 1 findet sich eine kurze Übersicht über die aktuell verfügbaren MPI Implementierungen. Hier ist Globus als MPICH-G aufgeführt. Im wesentlichen ist Globus eine Erweiterung des MPICH Systems von Argonne National Labs um das Globus Device. MPI ist das de facto Standard-Interface zur Entwicklung von message-passing basierten parallelen Programmen. MPICH-G ermöglicht einem MPI Programm ohne jegliche Änderung am Code in einer Gitterumgebung (verteiltes heterogenes Netzwerk von Supercomputern) zu laufen.

MPI-Vers.	Plattform	Hersteller	Eigenschaften
MPI	CRAY T3E	CRAY/SGI	
MPI	SGI Origin	SGI	ARRAY
MPICH-PM	Linux	RWCP (Japan)	Anpassung an Myrinet/Shared.-Mem.
MPICH	SGI/Cray/Linux	Argonne	CH-P4 Ethernet/Fast Ethern. (IP) (rsh/rhost)
MPICH-G (Globus)	SGI/Cray/Linux	Argonne	schneller als P4 NEXUS-Device für globus (IP) (gsh), 1024 Bit encrypt.

Tabelle 1: Übersicht über die unterschiedlichen MPI-Versionen auf den einzelnen Rechnerplattformen

Das Globus Metacomputing Toolkit ist weltweit bereits über 70 mal erfolgreich installiert worden und hat zahlreiche Preise gewonnen - zuletzt auf der Supercomputing '98 in Orlando (Florida). Es besteht aus mehreren Teilen:

- Benutzerverwaltung/Authentifizierung der Rechner und der User
- 1024 Bit Verschlüsselung + Kompression
- uniformer Zugriff auf verschiedene Rechnerressourcen (gsh, globus-rcp).
- MDS - Metacomputing Directory Service ist Teil der Informations-Infrastruktur des Globus Metacomputing Toolkit. Es bietet einen Überblick über vorhandene Ressourcen (Rechner, Software, Netze, gatekeeper, Dienste etc.)
- Gatekeeper steuern die Jobs auf unterschiedlichen Queueing-Systemen auf den einzelnen Plattformen

Die Vorteile des Systems liegen neben der einfachen Portierung bereits bestehender MPI-Programme vor allem in der Optimierung für hohe Latenzen, die aus der one-sided communication Strategie resultiert. Das MDS speichert statische und dynamische Informationen über den Status einer Metacomputing Umgebung. Hierzu wird ein LDAPv3 server benutzt. Die Sicherheit über Authentifizierung und Benutzerverwaltung sowie Transparenz des Systems sind besonders hervorzuheben. Der Gatekeeper übernimmt die Kopplung der unterschiedlichen Queueing Systeme, sodaß der Benutzer unabhängig hiervon arbeiten kann.

Unsere ersten Tests mit dem Globus System bestätigen diese Vorteile. Erste Bandbreiten Vergleiche sind in den Abbildungen 6 und 7 abgebildet. Erste Ergebnisse mit unseren Anwendungen der Projekte *Krustenfluide* und *Multiskalen Moleküldynamik* sind in den entsprechenden Abschnitten aufgeführt.

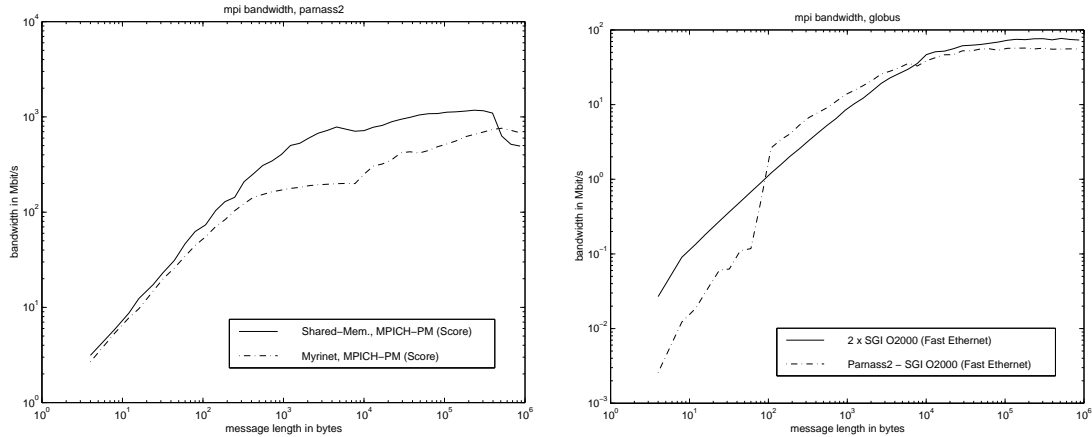


Abbildung 6: Links: Bandbreite zwischen zwei Parnass2 Prozessoren über Myrinet sowie zwischen zwei Parnass2 Prozessoren über shared memory unter MPICH-PM (SCORE). Rechts: Bandbreite zwischen Parnass2 und einer SGI O2000 im Vergleich zur Bandbreite zwischen zwei SGI O2000, jeweils über Fast Ethernet 100MBit/sec. unter Globus.

In Abbildung 6 (rechts) ist die Bandbreite unter der Verwendung des Globus Systems abgebildet. Sie ist deutlich glatter als die vergleichbaren Kurven der Standard MPI-Systeme (siehe Abb. 7 links). Auch wird die volle Fast Ethernet Bandbreite erreicht. Dies liegt mitunter daran, daß unter Standard-MPI drei verschiedene Modi Verwendung finden. Abhängig von der Größe der verschickten Pakete, kann es protokoll-bedingt zu mehreren Latenzen kommen (short, eager, rendezvous). Hier ist bei Globus die one-sided-communication deutlich weniger latenzanfällig. Die Unregelmäßigkeiten in der Globus Kurve zwischen den unterschiedlichen Systemen SGI Origin 2000 und Parnass2 beruhen höchstwahrscheinlich auf Byte swapping Effekten (Datenkonvertierung).

### 4.3 Rechnerplattform: Parnass2

Parnass2 ist ein am Institut für Angewandte Mathematik entwickelter und gebauter Parallelrechner mit 128 Intel Pentium II 400 MHz CPU's. Der Rechner ist aktuell (Juni 1999) auf Platz 362 der TOP500 Liste der schnellsten Computer der Welt.

Parnass2 ist ein PC-Cluster mit 64 Doppel-Prozessor Rechenknoten, welche jeweils mit 512 MB RAM bestückt sind. Zur Kommunikation zwischen diesen Rechenknoten wird ein Hochgeschwindigkeitsnetz der Firma Myricom verwendet. Die Myrinet-Netzwerkkarte ist eine 32-Bit PCI Karte, welche eine theoretische Übertragungsgeschwindigkeit von 1280 MBit/s voll duplex erreicht. Die Topologie

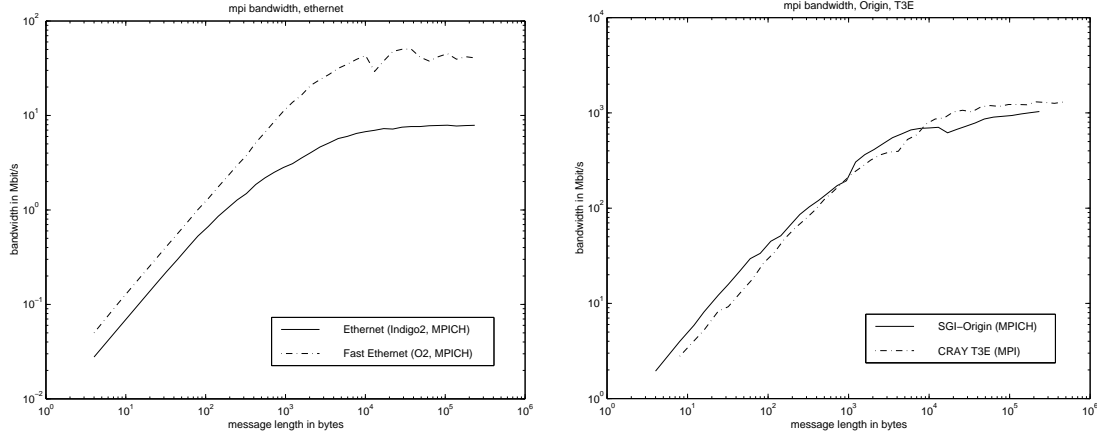


Abbildung 7: Links: Bandbreite zwischen zwei SGI Indigo2 unter MPICH über Ethernet 10Mbit/s. im Vergleich zur Bandbreite zwischen zwei SGI O2 unter MPICH über Fast-Ethernet 100Mbit/s. Rechts: Bandbreite der CRAY-T3E unter MPI im Vergleich zur Bandbreite innerhalb einer SGI Origin 2000 unter MPICH.

des verwendeten Netzes ist ein dreistufiger Fat-Tree mit voller Bisektionsbandbreite von 82 GBit/s.

Als Betriebssystem von Parnass2 wird Linux 2.2.7 eingesetzt, welches um den Hardwaretreiber für die Myrinet-Karte erweitert wurde. Dieser Treiber sowie das im Cluster verwendete message-passing System MPI-PM (SCORE-Projekt) wurde vom RWCP (Japan) entwickelt. Parnass2 erreicht eine Peak-Performance von 51.2 Gflop/s. Im LINPACK Benchmark erreicht Parnass2 eine Leistung von 29.6 Gflop/s.

## 5 Projekt *Krustenfluide*

### 5.1 Meilenstein 1:

Das dem Projekt zugrundeliegende Differentialgleichungssystem zur Beschreibung komplexer geophysikalischer Prozesse wurde auf den gesättigt-ungesättigten Bereich erweitert. Die Modellierung der Bodenzone auch oberhalb des Grundwasserleiters stellt somit kein Problem dar. Die für eine Lösung der Differentialgleichung notwendigen Rand- und Anfangsbedingungen sind bekannt. Für die Permeabilität  $K(P)$  wurde eine Approximation nach Van Genuchten angesetzt. Die Sättigungsfunktion  $S(P)$ , ebenfalls eine stark nichtlineare Funktion des Drucks, wurde hinsichtlich gegebener Konvergenzkriterien des für die Linearisierung verwendeten Newton-Verfahrens leicht durch Splines modifiziert. Numerische Experimente haben gezeigt, daß diese Näherung in Bezug auf unsere Anwendungen

gute Ergebnisse liefert.

Eine Anwendung des Projektes ist die Modellierung von transpirationsbedingten Bodendeformationen. Eine Beziehung von Transpiration von Bäumen und Bodenbewegungen ist experimentell gesichert. Laufende Messungen der Arbeitsgruppe Prof. Kümpel (Geophysik Uni Bonn) stellen die für einen Vergleich mit unserem Modell notwendigen Parameter zur Verfügung.

Die Theorie der Wasseraufnahme von Pflanzenwurzeln ist aktuelles Forschungsgebiet. In einem erstellten zweidimensionalen Modell, basierend auf der Darcy-Gleichung, wurden verschiedene in der Literatur angegebene Extraktionsfunktionen (Quellterme der Darcy-Gleichung) getestet. Die Unterschiede im Hinblick auf das Systemverhalten haben sich als signifikant herausgestellt; eine Mittelung aus der Mikroskala (Wurzel) auf eine Makroskala (Auflösung des Berechnungsgitters) unter Einbeziehung biologischer Erkenntnisse ist deshalb vermutlich nötig und soll zu einer präziseren Naturbeschreibung führen.

Ein erstes Anwendungsbeispiel ist in Abb. 8 zu sehen. Es handelt sich hierbei um Strömungs-, Transport, und Deformationsprozesse in der umgebenden Erdkruste eines Baumes.

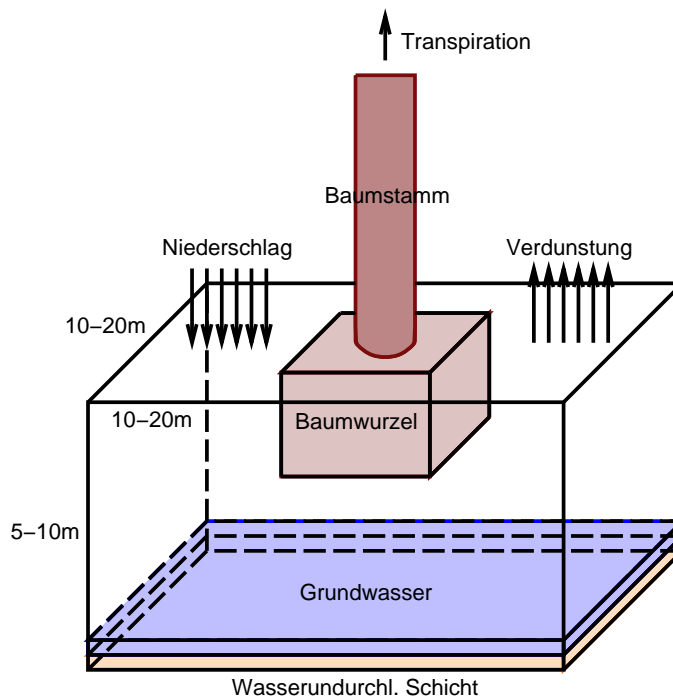


Abbildung 8: Anwendung: Erdkruste eines Baumes

Der Strömungslöser (Darcygleichung) wurde erfolgreich auf die CRAY T3E por-

tiert (vgl. Abb. 9). Jedoch ergaben sich hierbei deutliche Effizienzverluste gegenüber den Erfahrungen mit dem Code auf anderen MIMD-Systemen.

Prozessoren	Problem I [s] (3.5e+6 cells)	Problem II [s] (4.5e+5 cells)
4	80	18.5
8	42	10.4
16	21	5.3
32	11	2.3
64	5.5	1.5

Abbildung 9: Laufzeiten für zwei Testprobleme von Darcy/GP

Erste Analysen ergaben, daß hier offensichtlich das unterschiedliche I/O Handling eine Rolle spielt, sodaß nunmehr die Kommunikationsstruktur, die zunächst auf einem Master/Slave Konzept beruhte, umgestaltet werden muß.

Dies wurde mit Hilfe der guten Erfahrungen aus dem Parallelprojekt *Multiskalen Moleküldynamik* umgesetzt (s.u.).

## 5.2 Meilenstein 2:

Der Elastizitätslöser (Biot-Gleichung) und der Temperaturlöser wurden auf Parnass2 und der CRAY T3E getestet. Als Elastizitätslöser wird ein Standard-Tool eingesetzt. Als Temperatur- und Stoffkonzentrationslöser wurde am Institut für Angewandte Mathematik ein adaptiver, paralleler Code mit dynamischer Lastbalancierung mit einer auf dem Mehrgitterverfahren basierenden Diskretisierungstechnik entwickelt. Die Kopplung zwischen der O2000 und Parnass2 des Instituts für Angewandte Mathematik konnte erfolgreich umgesetzt werden. Da die Netzanbindung nach Jülich zu Beginn des Projektes noch nicht verfügbar war, wurden zunächst hausinterne Kopplungsversuche durchgeführt. Die Ergebnisse der hausinternen Kopplungsversuche unter Globus sind in Tabelle 2 aufgeführt. Vergleichszahlen unter herkömmlichem MPICH sind in Tabelle 3 zu finden. Die Zahlen spiegeln das Ergebnis der Bandbreitenmessungen aus Abschnitt 4.2 wieder. Die Zeitmessungen ergeben unter Globus einen deutlichen Effizienzgewinn, der über weite Distanzen noch größer sein dürfte.

Schließlich ist für eine schnelle Kopplung über das Gigabit-Netz eine effiziente Echtzeit-Kompression notwendig. Hierzu wurde ein am Institut für Angewandte Mathematik entwickeltes, adaptiv hierarchisches Kompressionsverfahren zur Übertragung von Video-Daten auf die spezielle Anwendungssituation im Projekt Krustenfluide - Austausch von Daten bei der Kopplung von Systemen partieller Differentialgleichungen - portiert.

Rechner	Prozessoren	Zeit[s]
Parnass2	1	60
Osiris (O2000)	1	48
Ovid (O2000)	1	54
Parnass2	2	37
Osiris	2	27
Parnass2/Osiris	1/1	32
Osiris /Ovid	1/1	34
Parnass2/Osiris	2/2	19
Osiris /Ovid	2/2	19

Tabelle 2: Zeitmessungen verschiedener Kopplungsversuche zwischen Origin O2000 Systemen und Parnass2 unter Globus. Osiris u. Ovid sind SGI O2000 Rechner.

Rechner	Verteilung	Zeit [s]
Osiris	1	60
Osiris/Ovid	1/1	35
Ovid	1/1	39
Ovid (shared mem)	1/1	35

Tabelle 3: Zeitmessungen verschiedener Kopplungsversuche zwischen Origin O2000 Systemen und Parnass2 unter Standard MPICH.

Das Verfahren wurde erfolgreich eingesetzt und erste Messungen ergaben bereits Kompressionsraten von 1 : 5 bis 1 : 10.

## 5.3 Meilenstein 3:

### 5.3.1 Anwendungsbeispiel: *Hydraulic Lift*

Wir modellieren Wasseraufnahme, Wassertransport durch das Wurzelsystem und transpirationsbedingte Bodendeformation als einen gekoppelten Prozeß. Die Wasseraufnahme von Pflanzenwurzeln ist ein mehrskaliges Phänomen. Auf einer Makroskala betrachten wir den Fluß durch die Bodenmatrix innerhalb und außerhalb des Wurzelbereiches sowie den Fluß durch das Wurzelsystem selbst. Wasseraufnahme und Transport wird als ein passiver Prozeß betrachtet, osmotische Effekte werden vernachlässigt. Auf der Mesoskala wird die Druckverteilung um die einzelne Wurzel, d.h. den Fluss in die Wurzel hinein, berechnet. Durch diese lokale Betrachtung wird berücksichtigt, daß die Wurzeln nicht einen gemittelten Druck

innerhalb einer makroskopischen Zelle 'spüren', sondern den lokalen Druck an der Wurzeloberfläche. Auf der Makroskala besteht der Modellansatz aus einer zweidimensionalen Richards-Gleichung (Boden) bzw. einer modifizierten Darcy-Gleichung (Wurzel), auf der Mikroskala wird die Druckverteilung in eindimensionaler Näherung berechnet. Das gekoppelte nichtlineare Gleichungssystem wird numerisch gelöst. Das Modell zeigt gute Übereinstimmungen mit Meßdaten zu hydraulic lift. Bei hydraulic lift handelt es sich um ein seit etwa 15 Jahren bekanntes Phänomen, daß viele Bäume und Sträucher Wasser nicht nur aufnehmen, sondern in zum Teil erheblichem Maße auch in den Boden abgeben können.

### Gleichungen:

#### Boden (Globales Gitter [i,j])

$$\phi \frac{\partial S(P)}{\partial t} - \nabla [K_f(P) \{ \nabla P + 10^4 e_z \}] = -2\pi (P_{rhi} - P_x) a_r K_r S_{rhi}$$

#### Wurzel (Globales Gitter [i,j])

$$\phi_x \beta_P \frac{\partial P_x}{\partial t} - \nabla [a_r K_a \{ \nabla P_x + 10^4 e_z \}] = 2\pi (P_{rhi} - P_x) a_r K_r S_{rhi}$$

#### Boden Mikroskala:

$$\phi \frac{\partial S_{loc}(P_{loc})}{\partial t} - \nabla [K_f(S_{loc}) \{ \nabla P_{loc} \}] = \nabla [K_f(P) \{ \nabla P + 10^4 e_z \}]$$

S(i, j) : Sättigung(loc: Mikroskala; rhi: Rhizosphäre)

P(i, j) : Boden Wasserpotential

$a_r$  [m<sup>2</sup>/m<sup>3</sup>] : Wurzeldichte

$K_r$  [m/Pa s] : radiale Leitfähigkeit ( $\approx 10^{-12}$ )

$K_a = \frac{\pi r_x^4}{8\eta}$  [m<sup>4</sup>/Pa s] (axiale Leitfähigkeit (Hagen-Poiseuille))

$\phi$  : Porosität ( $\approx 0.5$ )

$\phi_x$  : Xylemvolumen innerhalb Zelle/Zellvolumen ( $\approx 10^{-4}$ )

$\beta_P = 5 * 10^{-10}$  [Pa<sup>-1</sup>] : Kompressibilität von Wasser

#### Fluß durch Bodenmatrix, Makroskala:

Wir gehen von einem rein diffusivem Fluß aus, ein Makroporenfluß (Wurmlöcher, Klüfte...) wird nicht berücksichtigt. Als Extraktionsfunktion (siehe letzter Bericht) wird ein modifizierter Ansatz von Herkelrath (1977) verwendet, mit zwei

wichtigen Modifikationen: Erstens ist der Wurzelndruck (Xylemdruck)  $P_x$  nicht vorgegeben, sondern selber Lösung einer partiellen Differentialgleichung (Gleichung 2). Diese Erweiterung ist sinnvoll, da der Xylemdruck nicht bzw. nur sehr ungenau meßtechnisch erfaßbar ist. Zweitens ist  $P_{Rhi}$  auf der rechten Seite nicht identisch mit dem makroskopischen Druck (auf der linken Seite), sondern ist selber Lösung einer weiteren, lokalen Diffusionsgleichung. Dieser Ansatz ist insofern biologisch motiviert, als der Druckunterschied Wurzel - Wurzeloberfläche für die Größe der Wasseraufnahme entscheidend ist. Die radiale Leitfähigkeit (radial conductivity) ist ein Maß für die Durchlässigkeit der Wurzelmembran. Die axiale Leitfähigkeit kann nach dem Gesetz von Hagen-Poiseuille approximiert werden. Für beide Parameter stehen Meßdaten zur Verfügung. Die Gleichung wird mit finiten Differenzen diskretisiert. Als Gleichungslöser wird der Mehrgitterlöser verwendet. Die Linearisierung erfolgt mit einem Fixpunktverfahren.

### **Fluß durch das Wurzelsystem:**

Den Fluß durch das gesamte Wurzelsystem einer Pflanze mit allen Wurzelsegmenten mehrerer Ordnungen zu modellieren, ist aus Rechenzeit- und Speicherplatzgründen nicht möglich, wir hatten gehofft eine schnellere Verbindung zur CRAY nach Jülich herstellen zu können, um dann in Vollast beide Maschinen CRAY und Parnass2 für diese Aufgabenstellung nutzen zu können, jedoch konnte dies bis zum Abschalten der Leitung im Frühling 2000 nicht umgesetzt werden.

Deshalb kann innerhalb einer Modellierung bislang nur ein stark reduziertes, artifizielles Wurzelsystem betrachtet werden. In unserem Modell wird das Problem dadurch gelöst, daß die einzelnen Wurzeln durch eine Wurzelichte [in  $m/m^3$ ] ersetzt werden und der Xylemdruck Lösung einer gesättigten Darcy-Gleichung ist. Die Verwendung der Darcy-Gleichung zur Berechnung des Fluidtransport durch die Wurzel setzt voraus, daß das betrachtete Wurzelsystem einigermaßen 'gleichförmig' angeordnet ist, denn für das Wurzelsystem besteht i.a. kein direkter Druckausgleich verschiedener Wurzelsegmente, die sich innerhalb einer Zelle des Berechnungsgitters befinden (im Gegensatz zu einem porösen Medium, für welches üblicherweise die Darcy-Gleichung verwendet wird).

### **Fluß durch Bodenmatrix, Mikroskala:**

Die Berechnung des Drucks und der Sättigung unmittelbar an der Wurzeloberfläche erfolgt in eindimensionaler Näherung. Hierzu wird die Wurzelverteilung innerhalb einer Zelle des globalen Gitters als homogen angenommen. Der Radius des zu berechnenden Zylinders um eine einzelne Wurzel (das Wurzeinzugsgebiet) ist eine Funktion der Wurzelichte. Erste Untersuchungen zeigen, daß es nicht grundsätzlich nötig ist, die lokale Druckverteilung zu berechnen. Bei hoher Sättigung z.B. ist der lokale Druckgradient um die Wurzel herum zu vernachlässigen. Als Gleichungslöser für diese eindimensionale Gleichung wird eine Gauß

Elimination verwendet. Die Berechnung der Druckverteilung in der Mikroskala ist rechenzeit- und speicherplatzintensiv, da für jede Zelle des globalen Gitters innerhalb des Wurzelbereichs eine Differentialgleichung gelöst werden muß mit einer möglichst feinen Diskretisierung. Beispiele für die Wahl der Anfangs- und Randbedingung für die Gleichungen 1 und 2 sowie für die Wurzelichte sind in Abbildung 11 dargestellt.

### Hydraulic lift:

Abbildung 10 zeigt, wie die radiale Konduktivität  $K_r$  als Funktion der Richtung des Druckgradienten um die Wurzel angegeben werden kann. Bei hydraulic lift (Graph 1) ist  $K_r$  unabhängig von der Richtung des Druckgradienten. Ohne hydraulic lift wird  $K_r$  0 im Falle  $P_s < P_x$ . Der lineare Teil von Graph 2 mit  $0 < P_s - P_x < P_0$  ist nötig, um numerische Oszillationen zu reduzieren.

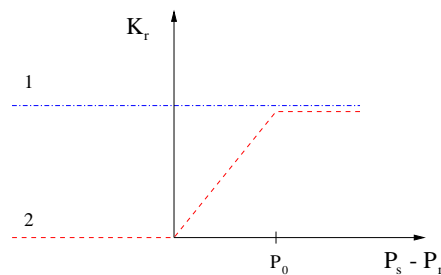


Abbildung 10: Die Abhängigkeit der radialen Leitfähigkeit von der Richtung des Wasserpotentialgradienten um die Wurzel. Graph 1: Hydraulic lift. Graph 2: Ohne hydraulic lift.

### Ausblick:

Die in Abbildung 11 gezeigte Approximation des Wurzelsystems durch Wurzelichten ist eine nicht wünschenswerte Vereinfachung der Problemstellung. Insbesondere für die Modellierung von Wurzelwachstumsprozessen wäre eine Auflösung der einzelnen Wurzel notwendig. Bei nichtadaptiver Auflösung, einem makroskopischem Gebiet von nur einem Meter und einem Wurzeldurchmesser von einem Millimeter wäre eine Diskretisierung von  $1000^3$  Punkten notwendig. Hinzu käme die Berechnung des Fluidflusses durch das Wurzelsystems durch ein gekoppeltes eindim. Differentialgleichungssystem (gesättigte Darcygleichung, evtl. nichtlinear). Dies ist in der gegebenen Situation der Übertragungsgeschwindigkeiten unseres Testbeds noch nicht sinnvoll umzusetzen.

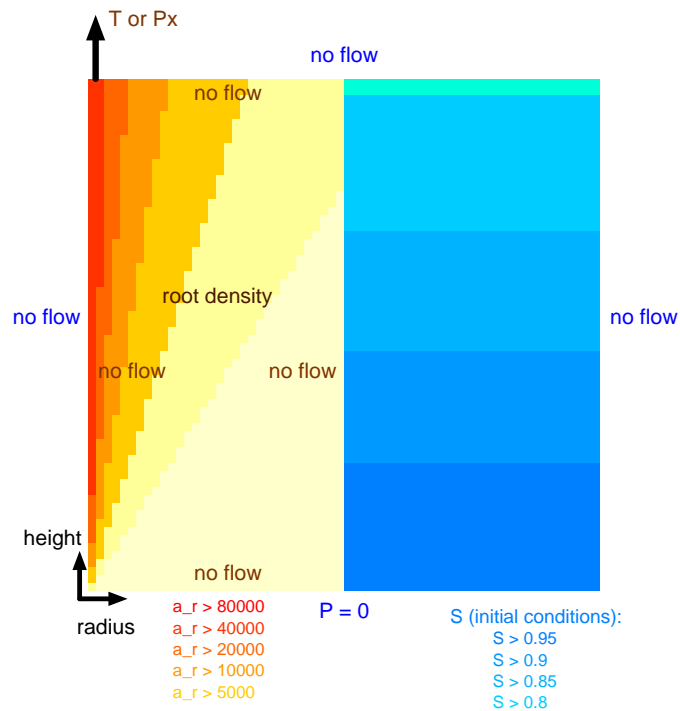


Abbildung 11: Wurzeldichte, Rand- und Anfangsbedingungen für die Sättigung (Beispiel) - Gleichungen 1 und 2. Höhe und Radius des Gebiets sind 5 m.

### 5.3.2 Anwendung: Bodendeformation als Folge der Transpiration

Die Bodendeformation als Folge der Transpiration wurde für den elastischen Fall berechnet. Die Rückkopplung der Bodendeformation auf die Fluidbewegung wurde als gering abgeschätzt und vernachlässigt. Als Randbedingungen wurde ein freier Rand an der Erdoberfläche sowie in vertikale Richtung gleitende Flächen im Zentrum und am Rand des Gebietes angesetzt. Der untere Rand wurde festgehalten. Die elliptische Differentialgleichung wurde mit finiten Differenzen diskretisiert. Die berechneten Verschiebungen liegen bei realistischer Wahl der Lamekoefizienten in einer Größenordnung, die von einem Tildmetermeßgerät aufgelöst werden können.

Die Animation in Abb. 12 zeigt die räumliche und zeitliche Verteilung der Wasseraufnahme bzw. Abgabe durch ein Wurzelsystem eines transpirierenden Baumes, der hydraulic lift betreibt. Durch die Pfeile ist die horizontale Bodenverschiebung gekennzeichnet, die mit der Elastizitätsgleichung berechnet wurde. Als Randbe-

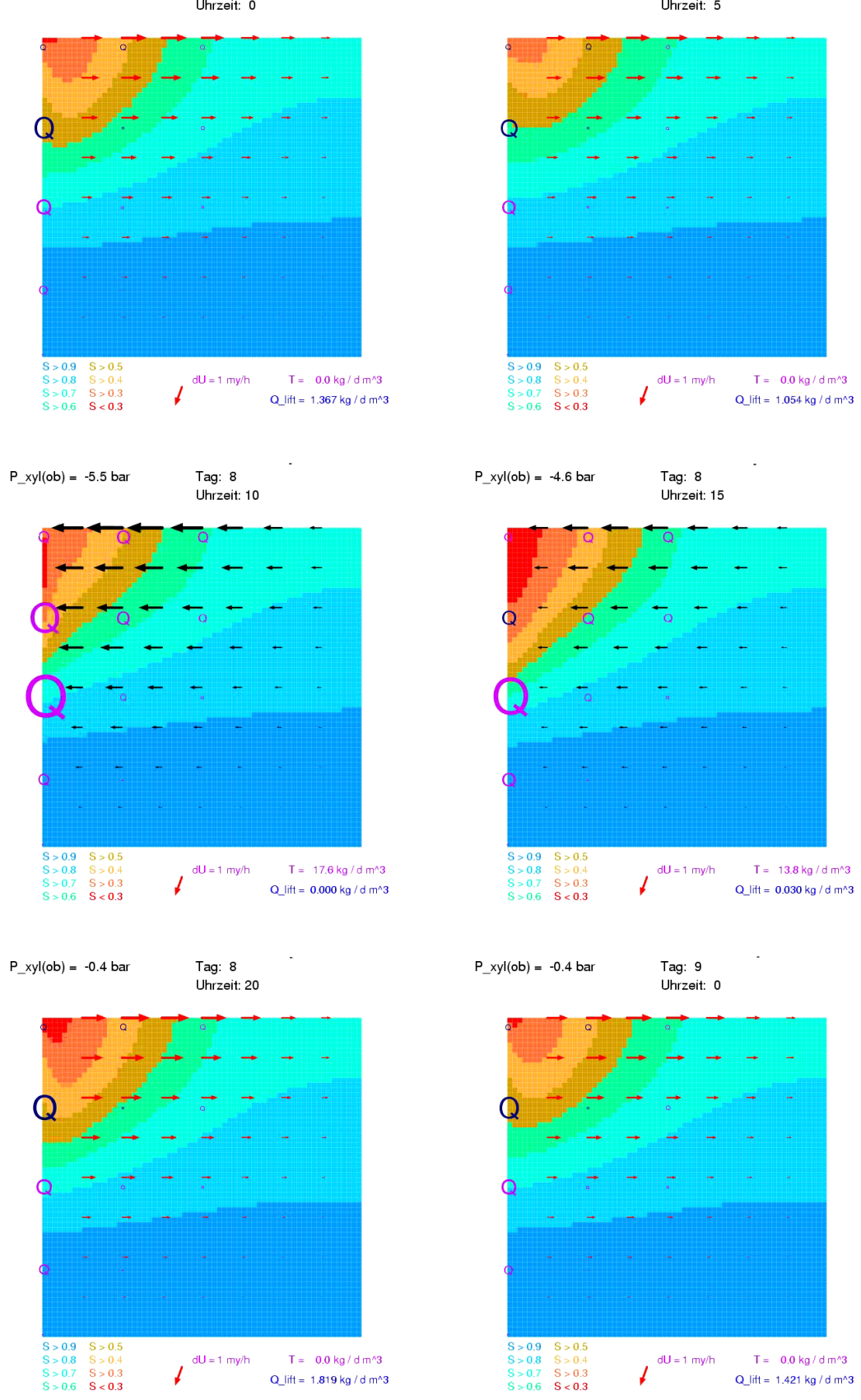


Abbildung 12: Animation der Bodendeformation als Folge der Transpiration

dingungen für die Gleichungen 1 und 2 wurden die in Abb. 11 angegebenen verwendet. Die Q's repräsentieren die Menge des an den Wurzeln aufgenommenen Wassers. Lila bedeutet Wasseraufnahme, blau Wasserabgabe an den Boden. Wir sehen, daß die Wasseraufnahme im Tagesverlauf räumlich variiert. Der Baum zieht gegen morgen zunächst in der oberflächennahen Schicht. Die Wasseraufnahme verlagert sich gegen abend zunehmend in tiefere, noch feuchtere Zonen. Die Bodendeformation ist ebenfalls eine Funktion der Tiefe; eine Messung mit einem Tildmeter ist deswegen möglich. Das System in Abbildung 12 befindet sich in der Nähe eines stationären Zustandes.

## Ergebnisse:

Für den Fall eines Sommertages wurde für einen tiefwurzelnden Baum hydraulic lift simuliert. Je nach Wahl der Parameter für den Boden und für die Pflanze ergeben sich während der Nacht umgeschichtete Wassermengen, die zwischen 10% und 50% der Transpirationsrate liegen. Dies liegt in der gleichen Größenordnung von Messungen von Emerman and Dawson 1996 an einem Zuckerahornbaum ( $100 \pm 50$  l/Tag). Das unterstützt die Vermutung, daß hydraulic lift ein rein passiver, physikalischer Prozess ist, also keine osmotische Komponente benötigt. In Abbildung 13 ist die Schwankung des Bodenwasserpotentials (in bar) als Folge der Transpiration dargestellt. Im Falle von hydraulic lift ist die tägliche Amplitude um etwa eine Zehnerpotenz größer als ohne hydraulic lift; dies ermöglicht uns das Aufstellen eines expliziten Meßverfahrens für diesen Effekt; die Transpirationsrate wurde im vorliegenden Fall sinusförmig angesetzt mit einer Gesamttranspiration von 100 l/Tag. Mit hydraulic lift kann die Pflanze offensichtlich mit geringeren Saugspannungen die gleiche Transpirationsrate erbringen. Dies stellt für die Pflanze insofern eine Optimierung dar, als geringere Saugspannungen mit geringerer Emboliegefahr verbunden sind. Unter Pflanzenphysiologen ist bislang die Bedeutung von hydraulic lift umstritten.

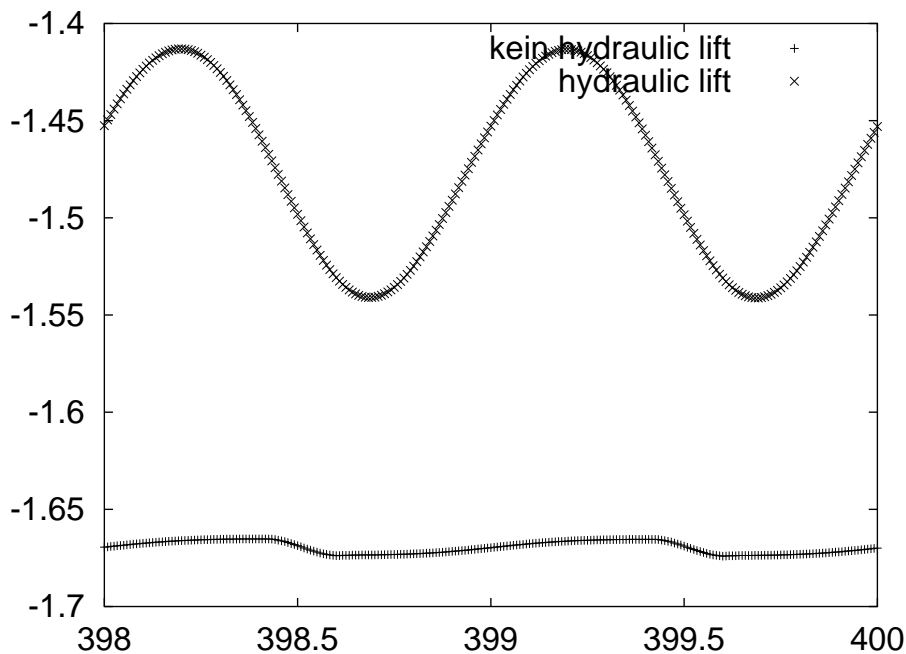


Abbildung 13: Schwankung des Bodenwasserpotentials (in bar) als Folge der Transpiration (Radius:1m, Tiefe:0.8m).

### 5.3.3 Numerische Umsetzung

### 5.3.4 Effizienzanalyse der Kopplung

Wir haben den Krustenfluid-Code in einer Gebietszerlegungs-Parallelisierung auf das gekoppelte System von zwei Parallel-Rechnern mit einer globalen MPI-Implementierung portiert. Dabei verwenden wir eine dreidimensionale Partitionierung des Gebiets, wobei unterschiedlich große Teile auf Parnass2 und die SGI abgebildet werden. Die MPI Implementierung MPICH liefert dazu ein transparentes Bild eines einzigen globalen Meta-Computers, inklusive der notwendigen Datenkonvertierung und Kommunikation zwischen den Rechnern. In diesem Fall wird im wesentlichen lokale Kommunikation verwendet, bei der jeder Prozessor mit seinen Nachbarn in allen drei Raumdimensionen Daten austauscht. Die Referenzmessung fand mit *einem* Parnass Prozessor und *einem* SGI Prozessor statt. Nachfolgend haben wir dazu die Rechenzeiten mit entsprechend mehr Prozessoren auf den einzelnen Maschinen in Relation gesetzt. Gemessen wurde die gemittelte Rechenzeit für einen einzelnen Zeitschritt, wobei natürlich für eine vollständige Simulation weit mehr als 10.000 Zeitschritte notwendig wären.

Wie zu erwarten war, skaliert die Rechenzeit bei einer Erhöhung der Prozessoren, was wir im Falle der SGI bis zu drei Prozessoren und im Falle von Parnass bis zu 16 Prozessoren sehr eindrucksvoll sehen (siehe Abb. 14 links). Allerdings führt eine Erhöhung der Zahl der Prozessoren über die angegebenen Werte hinaus zu keiner weiteren Verbesserung, im Falle der SGI sogar zu einer deutlichen Verschlechterung der Rechenzeit. Das ist zum einen auf die limitierte Bandbreite der Verbindung zwischen den beiden Rechnern zurückzuführen, zum anderen wurde im Falle der SGI ein gekoppeltes System verwendet, bei dem wir bei vier und mehr Prozessoren, mehrere SGI Origins verwendet haben. Damit kommt nicht mehr die Effizienz der shared-memory Kommunikation, sondern TCP/IP zum Einsatz, was sowohl zu einem Leistungseinbruch als auch zu Schwankungen in den Meßergebnissen führt. Es bleibt anzumerken, daß die Simulation größerer Beispiele, die damit den Speicher des Referenzsystems sprengen würden, effizientere Skalierungen für große Prozessorzahlen liefern würden.

### 5.3.5 Datenkompression

Zu Beginn des Projektes wurde ein am Institut für Angewandte Mathematik entwickeltes Echtzeit-Kompressionsverfahren eingesetzt (s.o.). Hierbei ist zwischen verlustfreier und verlustbehafteter Kompression zu unterscheiden. Daher wurde nun zusätzlich noch die verlustbehaftete Kompression mit biorthogonalen Wavelets implementiert. Diese brachte eine zusätzliche Effizienz von 10-20%. Ferner wurde der trade-off zwischen verlustfreier und damit normaler Konvergenz des Verfahrens sowie verlustbehafteter und damit langsamerer Konvergenz des Verfahrens untersucht. Es hat sich herausgestellt, daß verlustfreie Kompression in

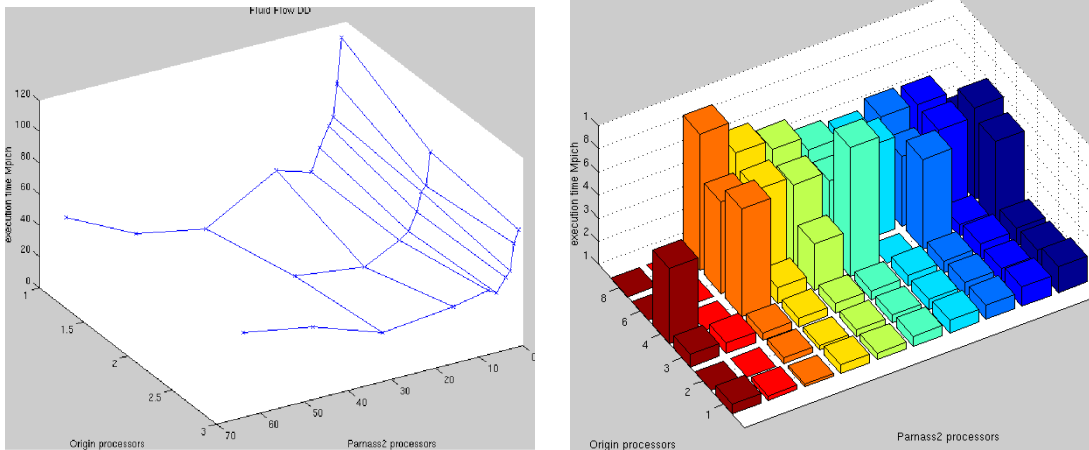


Abbildung 14: Rechenzeit eines Zeitschrittes einer gekoppelten Strömungssimulation zwischen Parnass2 und einer SGI Origin. Abszissen stellen die Zahl der Prozessoren dar, 1-64 auf Parnass2 sowie 1-8 auf der Origin. Die Ordinate bildet die Ausführungszeiten in Sek. ab.

unseren Anwendungen vorzuziehen sind, da sich die Genauigkeit und Konvergenz sonst dramatisch verschlechtern.

Daraufhin haben wir symmetrische Kompressionsverfahren untersucht unter anderem Entropieverfahren, Differenzenkodierung sowie Huffmankodierung. Dabei ergab sich, daß die einfachsten Verfahren die Besten waren, da die CPU auf Sender- und Empfängerseite nicht unnötig belastet wurde.

#### 5.4 Meilenstein 4:

Das gekoppelte System wurde nunmehr auf den oberflächennahen Bereich erweitert. Besondere Schwierigkeiten ergeben sich dabei durch die Berücksichtigung der ungesättigten Zone und des Bodenreliefs als neue Grenzfläche. Das resultierende Modell erlaubt schließlich die Beschreibung von Bodendeformationen unter dem Einfluß der Fluidbewegung im ungesättigten und gesättigten Bereich. Durch die Hinzunahme eines Ansatzes zur Rissausbildung eignet sich somit das gekoppelte System zur physikalisch-basierten Analyse von Hangstabilitäten. Von Seiten der numerischen Mathematik wurde das gekoppelte System Matrix-Fluid-Temperatur-Stoffkonzentration auf sehr große Anforderungen an Speicherplatz und Rechenzeit erweitert. Hierbei wurden nunmehr robuste Mehrgittermethoden zur Lösung der entstehenden Gleichungssysteme eingesetzt. Hierbei zeigte sich, daß die algebraischen Mehrgittermethoden Vorteile haben gegenüber den geometrischen Mehrgittermethoden mit matrixabhängigen Prolongationen.

Die Numerischen Ergebnisse finden sich in Abbildung 15. In der Simulation

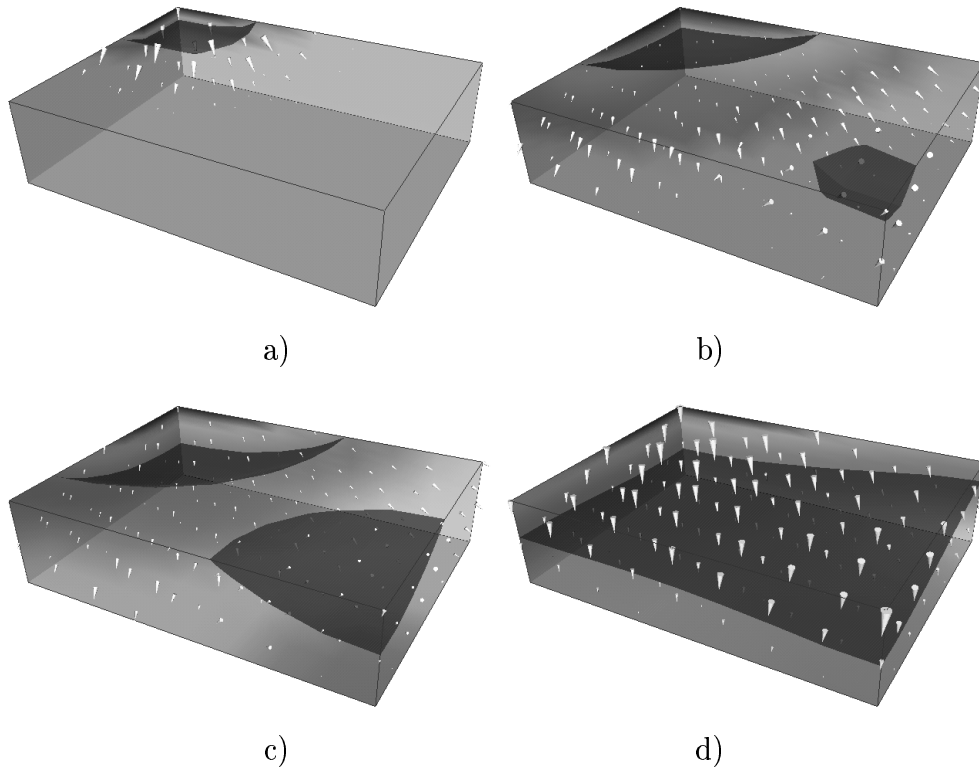


Abbildung 15: Verschiedene Zeitschritte einer Grundwasser-Simulation in der gesättigt/ungesättigten Zone. Dargestellt sind Isoflächen des Drucks sowie die Richtung des Verschiebungsvektorfeldes der Bodendeformation (stark überhöht).

ändern sich die Randbedingungen des Drucks an der Oberfläche (Regen/Verdunstung) lokal und temporal. Beginnend mit einem hydrostatischen Ausgangszustand setzt zunächst Niederschlag in einem kleinen Teilgebiet links hinten ein (a). Der Niederschlag wandert dann nach rechts vorne (b) während links hinten zunächst leichte Verdunstung eintritt (c). Schließlich beginnt auch dort wieder Niederschlag welcher bis zum Ende der Simulation (d) anhält. Man sieht die Isoflächen des Drucks an der Grenze zwischen gesättigter und ungesättigter Zone sowie Pfeile welche die Richtung des Verschiebungsvektorfeld der Bodendeformation anzeigen.

## 6 Projekt *Multiskalen Moleküldynamik*

### 6.1 SuParCup 99

Das Programm TreeMol war Teil eines Beitrages zum SuParCup 1999. Dieser Beitrag wurde mit einem 2. Preis ausgezeichnet. Den ersten Preis hat eine Forschergruppe der Sandia National Labs gewonnen.

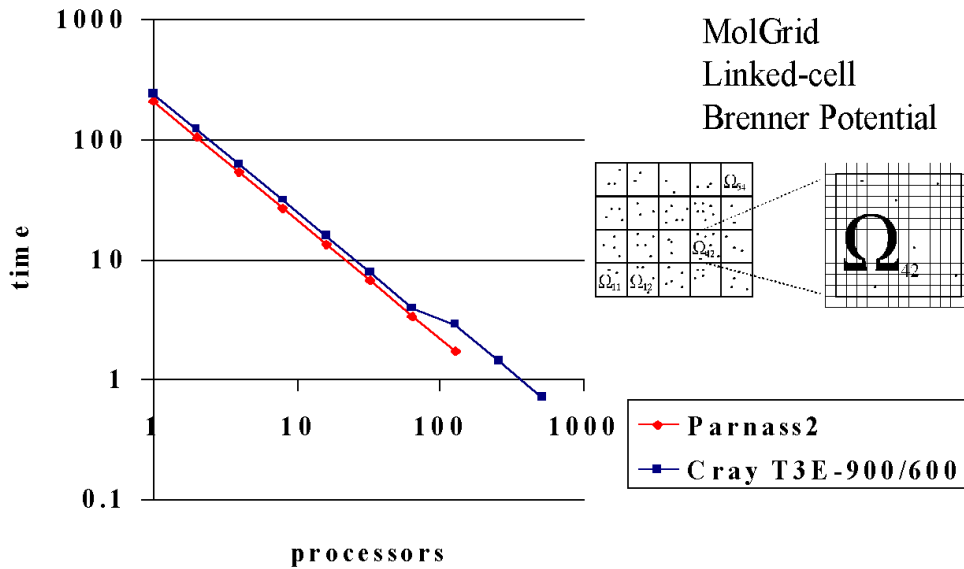


Abbildung 16: Skalierung des gekoppelten *MolScale* auf Parnass2 im Vergleich zur CRAY T3E

## 6.2 Meilenstein 1:

Die Portierung und Anpassung einer heterogenen MPI-Version in Form des Globus-Metacomputing-Systems wurde bereits in den vorangegangenen Abschnitten detailliert ausgeführt, sodaß hier nunmehr auf die Entwicklung des Softwarepaketes *MolScale* (vgl. 6.2.1) sowie die Entwicklung der Kommunikationsmodelle (vgl. 6.2.2) eingegangen wird.

### 6.2.1 MolScale

Die beiden Programmodule MolGrid und TreeMol wurden im Moleküldynamikpaket *MolScale* vereint. Hierzu wurden die einzelnen Module hinsichtlich des Einsatzes auf Metacomputern optimiert. Dies wurde unter dem Aspekt sehr komplexer Anwendungen - die ja im Metacomputing Projekt im Vordergrund stehen - erfolgreich umgesetzt. Das Programmpaket *MolScale* ist nunmehr in der Lage biomolekulare Anwendungen mit bis zu 170 Millionen Atomen - gegenüber 16 Millionen bei Antragstellung - zu simulieren.

Folgende Anwendungen konnten erfolgreich auf der CRAY T3E und Parnass2 umgesetzt werden:

- Biochemie: Biomolekulare Anwendungen mit 64 Mio. Atomen (vgl. Abb. 20, 23 (rechts)), darunter auch BPTI.

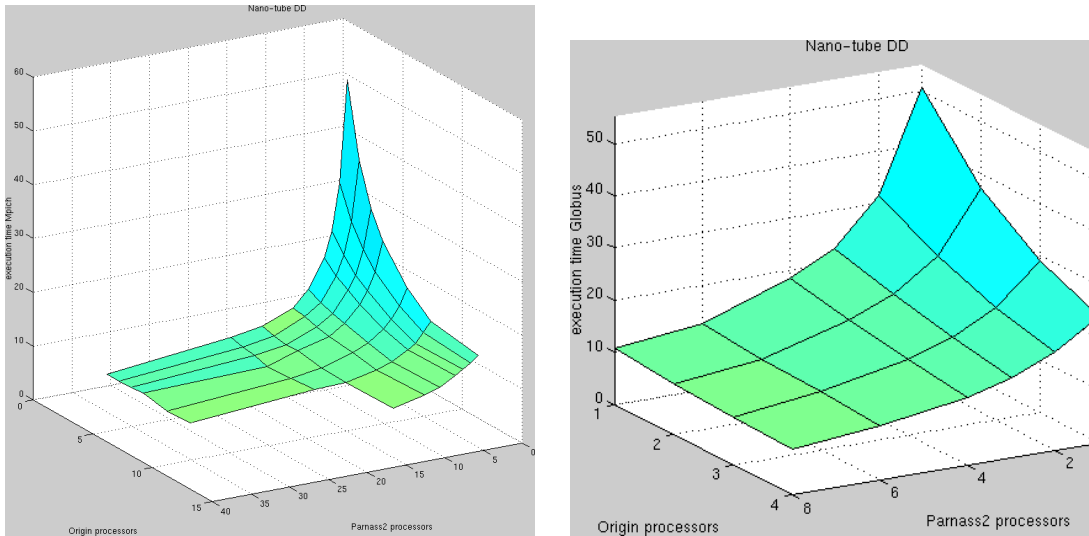


Abbildung 17: Rechenzeit einer gekoppelten Simulation auf Parnass2 und einer Origin mit MPICH (links) bzw. MPICH-G (GLOBUS) (rechts)

- Condensed Matter: Argon Simulation mit 170 Mio. Atomen
- Festkörper: Mikrorisse in Metallen mit bis zu 1 Million Atomen (vgl. Abb. 25)
- Nanotechnik: Nanotubes mit bis zu 30 Millionen Atomen können nun mit reaktiven Potentialen, die im naiven Ansatz eine Komplexität von  $O(N^4)$  haben, simuliert werden. (vgl. Abb. 21,22,23 (links))
- Fluide: Die Rayleigh-Taylor Instabilität konnte mit 1 Mio. Atomen gerechnet werden. (vgl. Abb. 24)

### 6.2.2 Kommunikationsmodelle

Für die Realisierung dieser Anwendungen mußten die Kommunikationsmodelle effizienter gestaltet werden. Hierbei wurden unterschiedliche Kommunikationsmodelle für das TreeMol-Modul auf der Cray T3E implementiert und getestet. So wurde das bisher verwendete Master/Slave Konzept aufgegeben - dabei wurde die Gebietszerlegung über raumfüllende Kurven in jedem Zeitschritt über den Masterprozess gesteuert. Dies erwies sich für komplexe Anwendungen als *Flaschenhals*, so daß nunmehr die Gebietszerlegung vollständig parallel läuft.

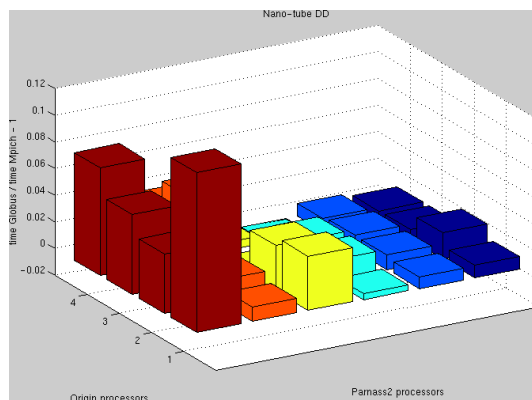


Abbildung 18: Vergleich der relativen Rechenzeiten aus der gekoppelten Simulation

Außerdem konnte eine Reduktion der gespeicherten Kopien von Partikeln anderer Prozessoren dadurch erreicht werden, daß nunmehr nur noch Pseudopartikel als Kopien zulässig sind. Dennoch ist das Ergebnis der effizienten Speicherausnutzung noch nicht zufriedenstellend, so daß im nächsten Berichtszeitraum näher analysiert werden muß, an welcher Stelle hier weiter optimiert werden kann, um höhere Partikelzahlen zu erzielen.

Ferner stellt sich bei derart komplexen Systemen die Notwendigkeit einer effizienten Partikelinitialisierung. Die Partikel wurden bisher immer von einem Prozess aus initialisiert, was bei einer Datenmenge von mehreren Millionen Partikeln unmöglich ist. Daher wurde nun eine parallele Dateninitialisierung implementiert. Die Speedups der weiterentwickelten Softwarepakete werden im folgenden Abschnitt besprochen.

### 6.3 Meilenstein 2:

Ein wesentlicher Aspekt der ersten Projektphase war laut Planung die Optimierung und Abstimmung der einzelnen Module TreeMol und MolGrid. Hierzu mussten die Module angepaßt werden, um:

1. hohe Skalierbarkeit auf der heterogenen Plattform
2. Realisierbarkeit komplexer Anwendungen

zu erzielen. Hierzu wurde in MolGrid innerhalb dieser Projektphase ein vollständig neues Anwendungsmodul für die Simulation von Kohlenstoff Nanotubes implementiert. Hierbei stellen Kohlenstoff Nanotubes eine der herausforderndsten Problemstellungen unserer Zeit dar. Diese 1991 von dem Japaner Iijima entdeckten

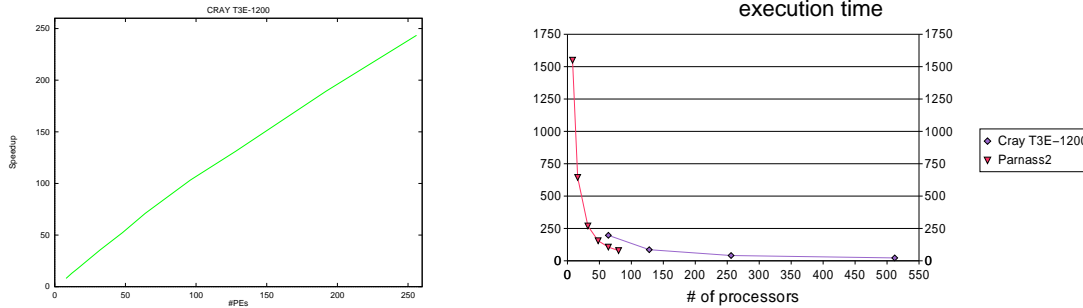


Abbildung 19: Links: Speedup von TreeMol auf der CRAY T3E-1200 bis 256 Prozessoren. Rechts: Laufzeitverhalten in Abhängigkeit von Prozessorzahl bis zu 512 Prozessoren auf der CRAY T3E im Vergleich zu Parnass2

Gebilde versprechen in naher Zukunft die Nanotechnologie zu revolutionieren. Erste Prototypen, wie z.B. Feld-Emissions Displays oder Epoxid-Harz Legierungen mit Kohlenstoff-Nanotubes existieren bereits.

Die Materialeigenschaften dieser Nanotubes sind unübertroffen. Einerseits sind sie extrem elastisch andererseits haben sie Materialeigenschaften, die sie mit Diamant vergleichbar machen. Die Computer Simulation von Kohlenstoff-Nanotubes stellt höchste Anforderungen an die Rechenleistung dar, da die zugrundeliegenden Algorithmen wesentlich komplexer sind als bei herkömmlichen Biopolymer Anwendungen. Hier ist nun in dieser Projektphase eine Simulation mit 100 Millionen Atomen gelungen, dies ist die weltweit größte Simulation mit Nanotubes. Tabelle 4 und 5 geben die Speedups dieser Simulationen wieder. Jeweils auf den beiden Plattformen Cray und Parnass2. Wie aus der Tabelle ersichtlich skaliert der Code hervorragend bis zu 512 Prozessoren und erreicht eine parallel Effizienz von bis zu 98%. Die Speedups wurden nach der üblichen Formel  $S(p) = T(1)/T(p)$  erstellt. Die Scaleups geben die Laufzeiten bei gleicher Last pro Prozess wieder und sind in den Tabellen durch Fettdruck hervorgehoben.

Auch das Modul TreeMol weist sehr gute Skalierungseigenschaften auf. Ein Speedupvergleich für die unterschiedlichen Plattformen CRAY und Parnass2 findet sich in Abbildung 19. Da dieses Modul die langreichweitigen elektrostatischen Wechselwirkungen berechnet, ist es vor allem für Biopolymer Anwendungen geeignet. Hier konnte durch verschiedene Optimierungen, wie sie bereits im Abschnitt 6.2.2 erläutert wurden, nunmehr wesentlich komplexere Anwendungen durchgeführt werden. Ein Beispiel ist in Abbildung 23. Hier handelt es sich um eine Dimethylcarbonat Simulation mit bis zu 64 Millionen Atomen. Das Laufzeitverhalten hierzu ist in Tabelle 6 abgebildet. Die bereits verbesserten Ergebnisse auf den lokalen Plattformen ermöglichen jedoch noch nicht die für das Projekt angestrebten komplexen biomolekularen Anwendungen. Diese benötigen die volle

time particles	processors								processors			
	CRAY T3E-900								CRAY T3E-600			
	1	2	4	8	16	32	64	65	128	256	512	
10024	15.3	7.63	3.89	1.95	0.99	0.52						
20020	30.1	15.2	7.81	3.89	1.95	0.99	0.52	0.75				
40012	<b>60.2</b>	30.4	15.6	7.75	3.89	1.95	1.00	1.35	0.73			
80024	120.9	<b>60.8</b>	31.1	15.6	7.75	3.90	1.97	2.74	1.45	0.73	0.38	
160020	240.6	122.4	<b>62.0</b>	31.0	15.5	7.77	3.89	5.57	2.86	1.44	0.73	
320012				<b>62.0</b>	31.0	15.5	7.77	11.1	5.66	2.86	1.44	
640024					<b>62.1</b>	31.0	15.5	22.4	11.3	5.66	2.86	
1320984						<b>62.9</b>	31.5	45.3	22.9	11.5	5.74	
2641968							<b>62.9</b>	90.6	46.1	23.0	11.4	
5283936								<b>182.2</b>	91.4	45.7	22.9	
10567872								363	<b>185.3</b>	91.4	45.6	
21135744									364.4	<b>183.7</b>	91.4	
29061648										251.0	<b>125.3</b>	
59994984											258	
99991640											430	

	processors							
	1	2	4	8	16	32	64	
Speed up	1.00	1.97	3.88	7.76	15.52	30.97	61.85	
Efficiency	1.00	0.99	0.97	0.97	0.97	0.97	0.97	

Tabelle 4: Oben: Ausführungszeiten des Moduls MolGrid für Nanotubes mit dem Brenner Potential und einem Beeman-Verlet Integrationsschema 5. Ordnung auf der CRAY T3E. Unten: Speed-Up und parallele Effizienz für Nanotubes mit 160020 Kohlenstoff-Atomen

Verfügbarkeit des Metacomputers und sind erst nach Kopplung der beiden Module TreeMol und MolGrid auf der CRAY und Parnass2 ausführbar. Die Testläufe auf der CRAY mußten zwischenzeitlich unterbrochen werden, da ein Betriebssystemupdate zu Problemen führte, die laut Aussage des ZAM auf fehlerhaftes I/O Handling zurückzuführen waren. Dies hat u.a. zu Systemabstürzen geführt, welche die gesamte Maschine blockiert haben. Das Problem wurde seitens des ZAM durch Bereitstellung eines weiteren OS-PE's für I/O Aufgaben gelöst.

### 6.3.1 Globus Messungen

Die Tabellen 7 und 8 geben die Ergebnisse des Vergleichs zwischen Globus und MPICH wieder. Die ersten Testläufe wurden zwischen verschiedenen Plattformen zunächst unter getrennter Verwendung der Module MolGrid und TreeMol durchgeführt. Das Modul MolGrid konnte problemlos unter Globus kompiliert

time particles	Parnass2 processors								
	1	2	4	8	16	32	64	96	126
10024	<b>13.16</b>	6.56	3.34	1.67	0.85	0.44	0.22	0.17	0.17
20020	26.3	<b>13.16</b>	6.72	3.34	1.68	0.85	0.44	0.33	0.28
40012	52.84	26.27	<b>13.4</b>	6.67	3.34	1.69	0.84	0.59	0.48
80024	<b>105.3</b>	52.6	26.8	<b>13.4</b>	6.69	3.34	1.68	1.15	0.90
160020	210.6	<b>105.2</b>	53.6	26.8	<b>13.4</b>	6.70	3.34	2.23	1.72
320012	423.0	210.4	<b>107.3</b>	53.6	26.8	<b>13.4</b>	6.71	4.48	3.45
640024	846.7	421.3	214.7	<b>107.3</b>	53.7	26.8	<b>13.4</b>	8.90	6.83
1320984				217.1	<b>108.8</b>	54.5	27.2	18.2	<b>13.9</b>
2641968					218.3	<b>109.1</b>	54.5	36.4	27.6
5283936						218.2	<b>108.9</b>	72.7	55.2
10567872							218.3	144.8	<b>110.3</b>

	processors								
	1	2	4	8	16	32	64	96	126
Speed up	1.00	2.00	3.93	7.86	15.72	31.43	63.05	94.44	122.44
Efficiency	1.00	1.00	0.98	0.98	0.98	0.98	0.99	0.98	0.97

Tabelle 5: Oben: Ausführungszeiten des Moduls MolGrid für Nanotubes mit dem Brenner Potential und einem Beeman-Verlet Integrationsschema 5. Ordnung auf Parnass2. Unten: Speed-Up und parallele Effizienz für Nanotubes mit 160020 Kohlenstoff-Atomen

time particles $N$	CRAY T3E-1200 processors $P$		
	128	256	512
1 mio.	16.12	8.18	5.30
2 mio.	35.40	16.94	
4 mio.	<b>86.12</b>	40.95	22.40
8 mio.	207.80	<b>91.56</b>	48.07
16 mio.	<b>511.42</b>	214.04	<b>99.65</b>
32 mio.	1763.53	<b>644.61</b>	
64 mio.			<b>746.88</b>

Tabelle 6: Ausführungszeit des Moduls TreeMol auf der CRAY T3E-1200

und eingesetzt werden. Das TreeMol Modul verwendet *long-long* integer zur Indizierung von Hashknoten, die momentan von Globus nicht unterstützt werden. Hier mußten einige Anpassungsarbeiten vorgenommen werden. Wie schon in den vorangegangenen Abschnitten für die Krustenfluide (vgl. 6.2.2) und die Bandbreitenmessungen (vgl. 4.2) unter Globus berichtet, ist auch hier die Effizienz des Globus Systems gegenüber herkömmlichem MPICH deutlich erkennbar.

Prozessoren auf		Laufzeit (sec)	
Ovid (O2000)	Plinius (O2000)	MPICH	GLOBUS
1	0	4.823179	3.958883
2	0	2.810561	2.410061
3	0	2.362030	1.881435
4	0	1.955958	1.706795
1	1	2.875025	2.429457
2	2	1.934255	1.594707
3	3	1.721953	1.397702
4	4	1.348921	1.323624

Tabelle 7: Testmessungen des Moduls TreeMol unter Globus

Rechner	Proz.	MPICH	GLOBUS
Ovid	1	25.0	22.2
	2	12.8	11.1
	4	6.8	5.65
Plinius	1	24.8	22.1
	2	12.46	11.1
	4	6.4	5.6
Ovid+Plinius	8	3.3	3.10
Ovid(3)+Plinius(3)+Parnass2(2)	8		3.1
Parnass2+Ovid	2		11.0
	4		5.67
Parnass2	1	15.4	16.2
	2	7.69	9.2

Tabelle 8: Testmessungen des Moduls MolGrid unter Globus. Ovid, Plinius sind SGI Origin 2000 Systeme

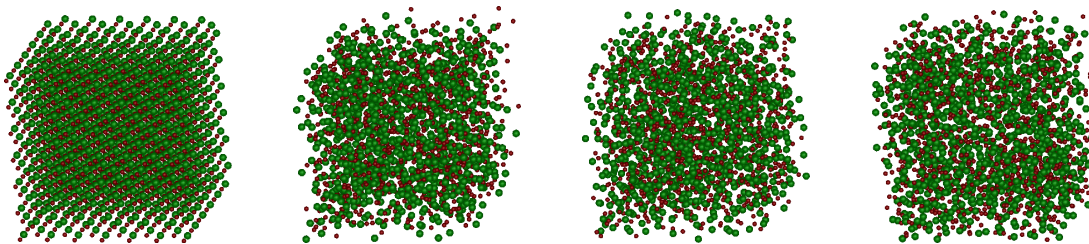


Abbildung 20: NaCl-Schmelzen: Startkonfig.: uniformes Gitter, 1014 Na, 1014 Cl Ionen, Maxwell-Boltzmann Geschwindigkeitsverteilung bei 1800 K, Partikelkonfiguration für NaCl zu den Zeiten 0.00, 0.47, 0.94 and 1.41 (Zeiteinheit:  $1.03e-15$  sec)

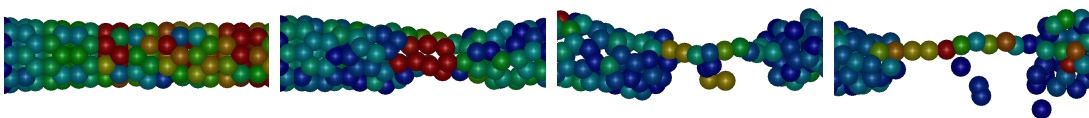


Abbildung 21: Zugexperimente mit Kohlenstoff-Nanotubes

### 6.3.2 Integrationsverfahren und Kopplung

Schließlich wurden im Hinblick auf die Kopplung der beiden Programmmodule TreeMol und MolGrid verschiedene Integrationsverfahren implementiert und getestet.

Insbesondere bei den Nanotubes zeigte sich, daß die Genauigkeit der herkömmlichen Integrationsverfahren nicht ausreicht. Nach verschiedenen Tests erwies sich das Beeman-Verlet Verfahren 5. Ordnung als zufriedenstellend. Anhand der ersten Erfahrungen mit der Kopplung und Synchronisation der Kraftberechnung müssen die Integrationsverfahren entsprechend angepaßt werden. Dies wurde zunächst mit kleinen Testanwendungen durchgeführt.

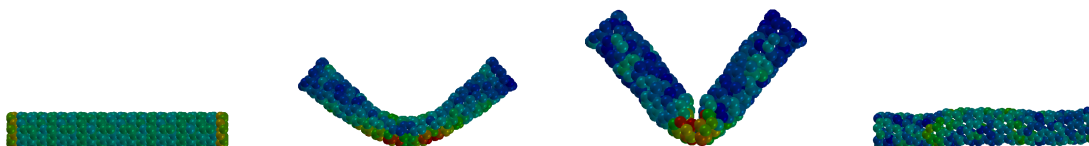


Abbildung 22: Biegen von Nanotubes

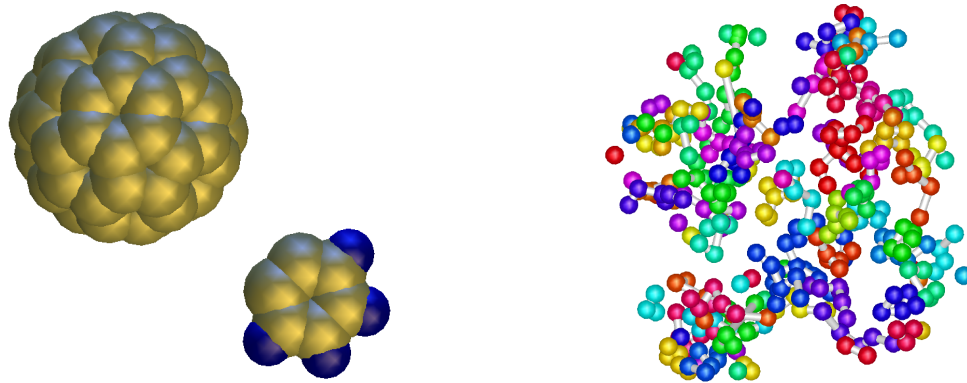


Abbildung 23: Kollision eines  $C_{60}$  Buckyballs mit Dehydrobenzol (links), Moleküldynamiksimulation von Dimethylcarbonat (rechts)

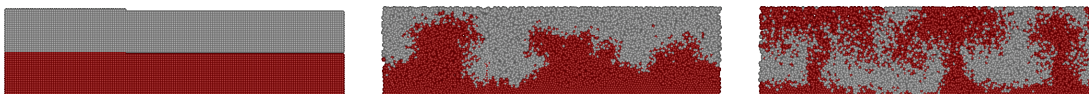


Abbildung 24: Rayleigh-Taylor Instabilität

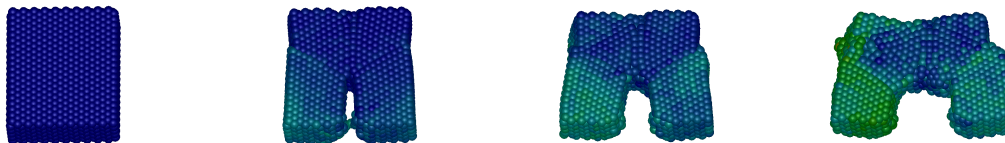


Abbildung 25: Mikrorisse in harten Materialien (3D)

## 6.4 Meilenstein 3:

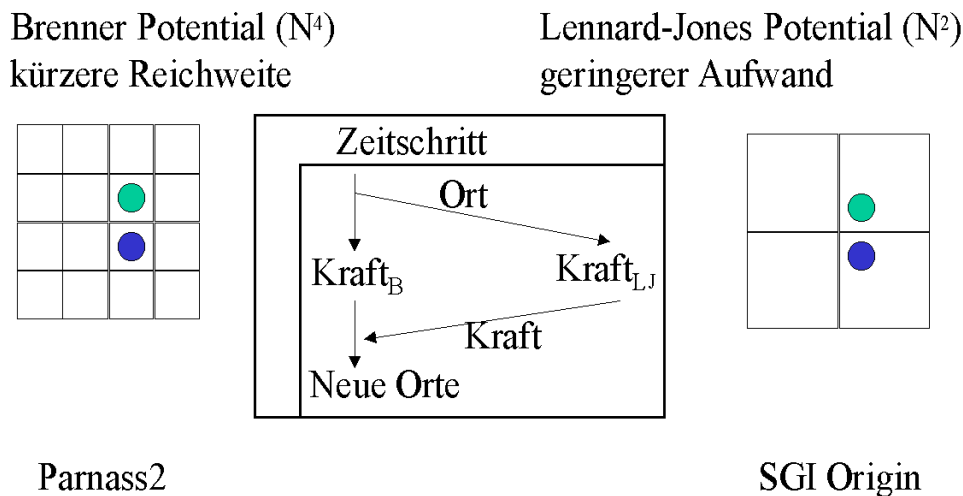


Abbildung 26: Neue Kopplungstechnik: Operatorsplitting

Das bisher durch Kopplung der beiden Module *TreeMol* und *MolGrid* entstandene Softwarepaket *MolScale*, konnte in der letzten Projektphase wesentlich verbessert und weiterentwickelt werden.

Hierbei wurden die Erfahrungen aus der ersten Projektphase genutzt, um durch Umstrukturierung und Verwendung verbesserter Datenstrukturen die Effizienz deutlich zu steigern.

### 6.4.1 Geometric Decomposition

Abbildung 16 zeigt die Speedups und die Parallelisierungsstrategie. Wir erkennen, die exzellente Skalierungseigenschaft des *MolScale* Codes auf beiden eingesetzten Plattformen, nämlich Parnass2 und CRAY, weiterhin sehen wir, daß die Geschwindigkeit pro Prozessor vergleichbar ist, daß sogar Parnass2 etwas schneller ist, was im Vergleich zu den in der CRAY eingesetzten COMPAQ Alpha Prozessoren bemerkenswert ist. Angegeben ist auch hier jeweils die Rechenzeit für einen einzelnen Zeitschritt und es liegt nahe, daß sowohl die Zahl der Partikel als auch die Zahl der Zeitschritte durch die vorhandenen Rechenressourcen limitiert sind. Die Kopplung der beiden Maschinen ist daher für größere Simulationen erforderlich.

In den Abbildungen 17, 18 sind die Messergebnisse für den ersten Kopplungstest abgebildet. Hier wird mit einer Gebietszerlegungsstrategie parallelisiertes *MolScale* auf dem virtuellen, globalen Parallelrechner, der aus den lokalen Maschinen

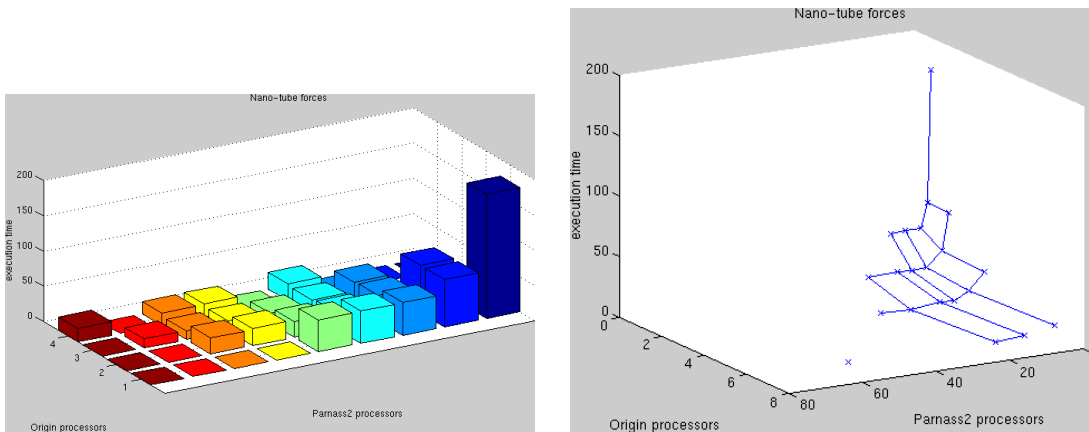


Abbildung 27: Rechenzeiten für neue Kopplungstechnik (Operatorsplitting) auf Parnass2 und SGI Origin

Parnass2 und einer SGI Origin besteht eingesetzt. Abgebildet sind wieder Rechenzeiten für einen Zeitschritt. Die Rechner sind mit ATM über LANE gekoppelt. Zum Einsatz kommen die MPI Implementierungen MPICH und MPICH-G, wobei wir in diesem Fall beide Implementierungen vergleichen wollen. Als Referenzrechnung dient uns in diesem Fall eine Simulation mit *einem* Parnass2 und *einem* SGI Origin Prozessor. Die gemessenen Rechenzeiten zeigen eine sehr gute Skalierbarkeit sowohl in der Zahl der bis zu 36 Parnass2 Prozessoren als auch in der Zahl der bis zu 12 Origin Prozessoren. Wir beobachten dieses Verhalten für beide MPI Implementierungen und wir sehen eine durchgängige Verringerung der Rechenzeiten bei größeren Prozessorzahlen im Gegensatz zu den vorangegangenen Ergebnissen mit dem Krustenfluid-Code. Dies kann man unter anderem darauf zurückführen, daß die *MolScale* Simulation ein günstigeres Verhältnis von Rechenzeit zu Kommunikation aufweist.

Die Abbildung 18 zeigt noch einmal den detaillierten Rechenzeitvergleich beider MPI Implementierungen, wobei wir hier deutlich sehen, daß MPICH-G bis zu 8% Prozent langsamer ist als MPICH, dies läßt sich insbesondere für sehr kleine Zahlen von Parnass2 Prozessoren nachweisen. Dies steht im Gegensatz zu den reinen Bandbreiten Messungen, bei denen MPICH-G um einen Faktor 2 besser abgeschnitten hatte als MPICH und läßt sich auf die Implementierungen der Datenkonvertierung zurückführen.

#### 6.4.2 Operator Decomposition

Ferner wurde das in der ersten Projektphase fertiggestellte Modul für die Large-Scale Simulation von Kohlenstoff-Nanotubes nun um ein Modul für langreichweitige Wechselwirkungen erweitert.

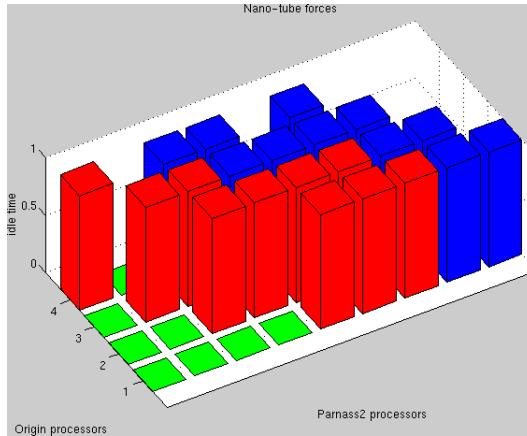


Abbildung 28: Idle Zeiten während der gekoppelten Simulation

Damit können nun erstmals gekoppelt auf verschiedenen Rechnerplattformen das komplexe Brenner-Potential und langreichweitige Potentiale **parallel** berechnet werden.

Kopplungsversuche mit diesem neuen Tool konnten erfolgreich umgesetzt werden. Hierbei wurde auch schon die Zerlegung der Potentiale durch Operatorsplitting erfolgreich umgesetzt. Siehe Abb. 26

Die Effizienz wird in Abbildung 27 ersichtlich. Dort ist ein Kopplungsexperiment mit 44800 Atomen zwischen Parnass2 und einer SGI Origin dargestellt. Hierbei wird das langreichweitige Potential auf Parnass2 berechnet und das kurzreichweitige Brenner Potential auf einer Origin. Verwendet wird die MPI Implementierung in MPICH. Wir vergleichen die Rechenzeit für einen einzigen Zeitschritt einer Konfiguration mit einem Prozessor Parnass2 sowie einem Origin Prozessor mit größeren Zahlen von Prozessoren auf den jeweiligen Maschinen. In diesem Fall ist zu beachten, daß es nicht sinnvoll ist, die Zahl der Prozessoren auf den einzelnen Rechnern unabhängig voneinander zu variieren, da die Last, die durch das Operatorsplitting gegeben ist, auf beide Rechner in einem festen Verhältnis abgebildet wird. Dadurch ist es notwendig auch die Zahl der Prozessoren in eben diesem Verhältnis zu skalieren. Dabei wurden die Parameter so gewählt, daß deutlich mehr Parnass2 Prozessoren als Origin Prozessoren zum Einsatz kommen. Wenn wir das geeignete Verhältnis fixieren, erhalten wir entlang der Diagonalen der Abbildungen 27 eine sehr gute parallele Skalierbarkeit.

Es ist klar, daß der Einsatz von mehr Prozessoren auf einem der beiden Rechner gegenüber dieser Ideallinie keine weitere Verbesserung an Rechenzeit bringt und den betreffenden Rechner entsprechend lange leerlaufen läßt, wie deutlich in Abbildung 28 zu sehen ist. Hier sieht man auch noch mal deutlich das geeignete experimentell zu ermittelnde Verhältnis der Prozessoren.

Die Anwendungsbreite des verbesserten Softwarepaketes *MolScale* erstreckt sich nunmehr auch auf mehrwandige Nanotubes. Hierbei werden die Wechselwirkungen zwischen den Tubes (langreichweitiges Potential) auf einer Rechnerplattform behandelt, während die Wechselwirkungen innerhalb der Nanotubes (Brenner Potential) auf der anderen Rechnerplattform parallel hierzu erfolgen. Eine Animation eines solchen Simulationslaufes ist in Abb. 29 zu sehen. Dort wird ein vierwandiger Nanotube jeweils an den äußersten Ringen longitudinal gestreckt. Hierbei kann man einen interessanten Effekt beobachten: die inneren Röhren erfahren bereits eine plastische Deformation, während die äußeren Tubes weiterhin elastisch bleiben. Dies bestätigt die Vermutung, daß das Elastizitätsverhalten der Nanotubes unmittelbar vom Durchmesser der jeweiligen Röhren abhängt.

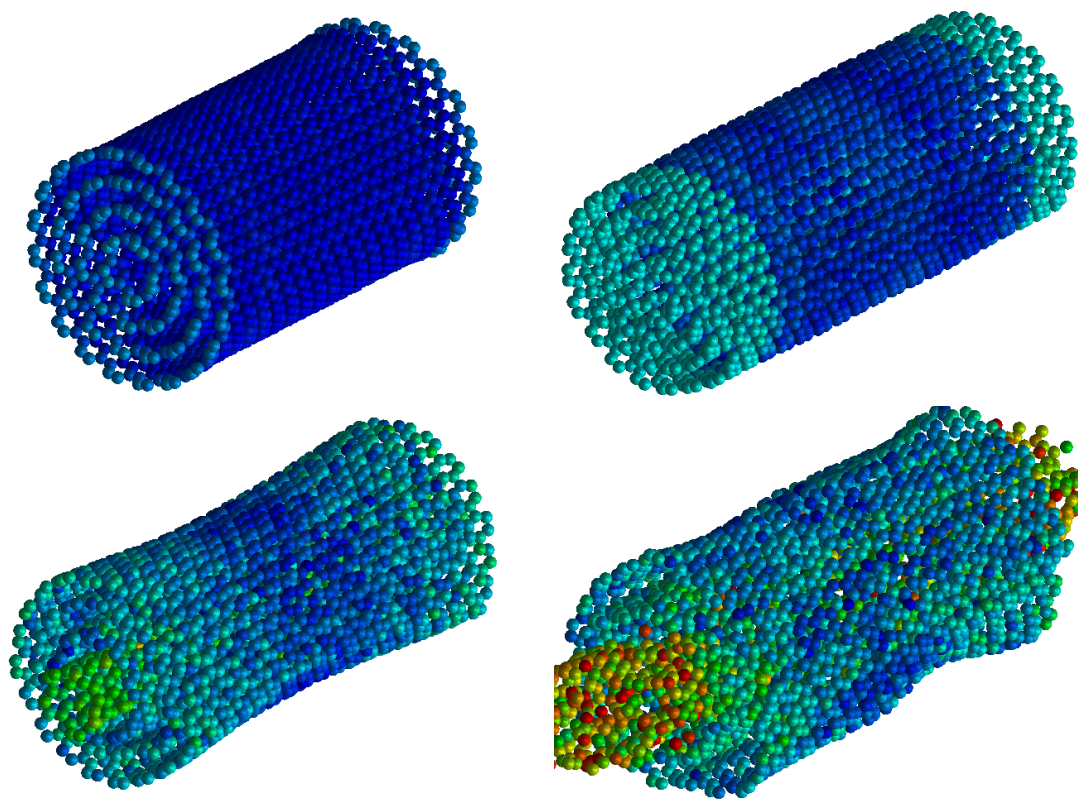


Abbildung 29: Eine Animation mit mehrwandigen Kohlenstoff-Tubes

Ferner konnten unsere Simulationen mit Proteinen verbessert werden. Eine erste Animation ist in Abbildung 30 zu sehen. Wir haben konnten eine Schnittstelle an die Brookhaven Protein Data Bank implementieren, die es uns ermöglicht nahezu sämtliche bekannte Proteine zu simulieren.

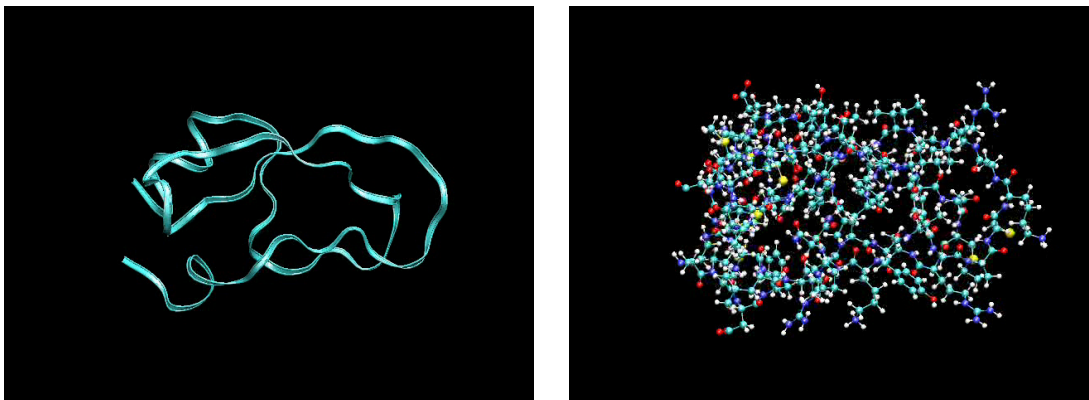


Abbildung 30: Eine Animation mit Proteinen

## 6.5 Meilenstein 4:

Die Simulationen konnten abschließend selbst für komplexe Systeme mit komplizierten Potentialen auf bis zu 170 Millionen Atome erweitert werden. Dies wurde durch eine erneute Verbesserung der Speicherausnutzung erreicht. Anwendungsseitig konnten zusätzlich zu den ursprünglich geplanten Simulationen weltweit erstmalig ein Komposit aus Nanotubes und Polyäthylen mit mehr als einer Million Atome im Rechner simuliert werden. Dabei gelang es erstmalig einen Lasttransfer von der Polyäthylen Matrix auf den Nanotube bei einem Poisson Verhältnis von 0.3 nachzuweisen. Abbildung 31 zeigt oben den Startzustand eines Kompositmaterials bestehend aus einer Polyäthylen Matrix und einem (10,10) Nanotube der an den Enden geschlossene Kappen besitzt. Die Abbildung unten zeigt einen späteren Zeitschritt unter tensilem Stress, bei dem bereits deutliche Deformation des Nanotubes zu erkennen ist, somit ist der Lasttransfer von der Matrix auf den Tube gelungen. Kompositmaterialien mit Nanotubes sind aktueller Forschungsgegenstand der NASA (Houston). Die Ergebnisse aus diesem Berichtszeitraum werden bei der NanoTech 2000 in Houston vorgestellt werden. Darüberhinaus wurden im Bereich der Biomoleküle neue Simulationen erfolgreich umgesetzt. Mit der nunmehr erfolgreich abgeschlossenen Anbindung an die Brookhaven Protein-Datenbank, konnten mit Hilfe dieser Anbindung erstmals Simulationen mit HIV-Protease durchgeführt werden.

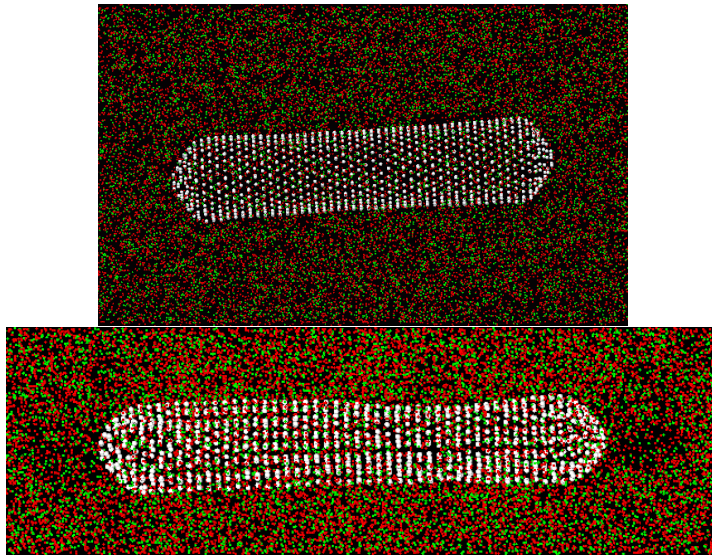


Abbildung 31: Eine Animation eines Kompositmaterials aus Nanotubes und Polyäthylen unter tensilem Stress, oben: Startzustand, unten: Ein Zeitschritt, bei dem bereits eine Deformation des Tubes erfolgt ist; somit ist Last von der Matrix auf den Tube übertragen worden.