



Paralleles Höchstleistungsrechnen

Einführung

Peter Bastian

Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Universität Heidelberg
Im Neuenheimer Feld 368
D-69120 Heidelberg
email: `Peter.Bastian@iwr.uni-heidelberg.de`

2. Oktober 2008



Organisatorisches

Motivation

Einige Beispielanwendungen

Superrechner

Inhalt der Vorlesung



- ▶ Dozent
 - ▶ Peter Bastian
 - ▶ zu finden in: INF 368, Raum 420
 - ▶ email: `Peter.Bastian@iwr.uni-heidelberg.de`
- ▶ Vorlesung
 - ▶ Di 9-11 in 368/532, Do 9-11 in 350/U014
 - ▶ Skript, Folien (falls verfügbar), Übungsaufgaben:
`http://hal.iwr.uni-heidelberg.de/phlr.html`



- ▶ Leiter: Christian Engwer
- ▶ findet statt im CIP-Pool des IWR, Otto-Meierhof-Zentrum
- ▶ Zeit: **Noch zu vereinbaren!**
- ▶ **Anmeldung zur Übung**
- ▶ Schein auf Übungsaufgaben
- ▶ Bachelor/Master: mündliche Prüfung (bzw. Klausur bei großem Andrang)
- ▶ Voraussetzung für die Teilnahme an der mündlichen Prüfung/Klausur: 50% der Übungsaufgaben
- ▶ C/C++ Kenntnisse erforderlich



Organisatorisches

Motivation

Einige Beispielanwendungen

Superrechner

Inhalt der Vorlesung



Warum parallel Rechnen ?

- ▶ Nebenläufigkeit
 - ▶ Abarbeitung mehrerer Prozesse auf einem Prozessor
 - ▶ Multi-Tasking Betriebssysteme seit den 60er Jahren
 - ▶ Bedienung mehrerer Geräte und Benutzer
 - ▶ Ziel: Steigerung der Auslastung
 - ▶ “Hyperthreading”: Nutze Wartezeiten des Prozessors
 - ▶ Mehrere Dinge gleichzeitig : Web-Browser, Desktop
 - ▶ Koordinationsproblematik tritt bereits hier auf
- ▶ Verteilte Anwendungen
 - ▶ Datenbasis ist inhärent verteilt: betriebswirtschaftliche Software, Warenfluß in großen Unternehmen
 - ▶ Hier wichtig: plattformübergreifende Kommunikation, Client-Server Architekturen
 - ▶ Auch wichtig: Sicherheit, VPN, etc. (behandeln wir nicht)

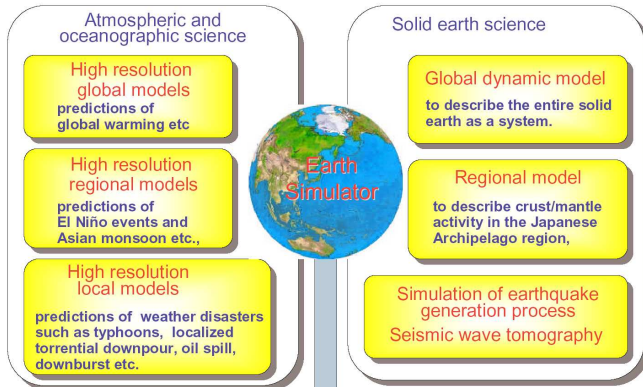


High Performance Computing

- ▶ Treibt Rechnerentwicklung seit Anfang 1940er
- ▶ Anwendungen: mathematische Modellierung und numerische Simulation
- ▶ ungestillter “Hunger” nach Rechenzeit:
 - ▶ Modellfehler → detailliertere Physik
 - ▶ Approximationsfehler → mehr Freiheitsgrade
- ▶ Grand Challenges:
 - ▶ Kosmologie, z. B. Galaxiendynamik
 - ▶ Proteinfaltung
 - ▶ Erdbebenvorhersage
 - ▶ Klimaentwicklung
- ▶ ASCI Program (Advanced Simulation and Computing)
Programm amerikanischer Parallelrechner seit 1992,
Funding 2004: 200 Mio US-\$ DoE, 87 Mio US-\$ NSF



Earth Simulator (Japan)



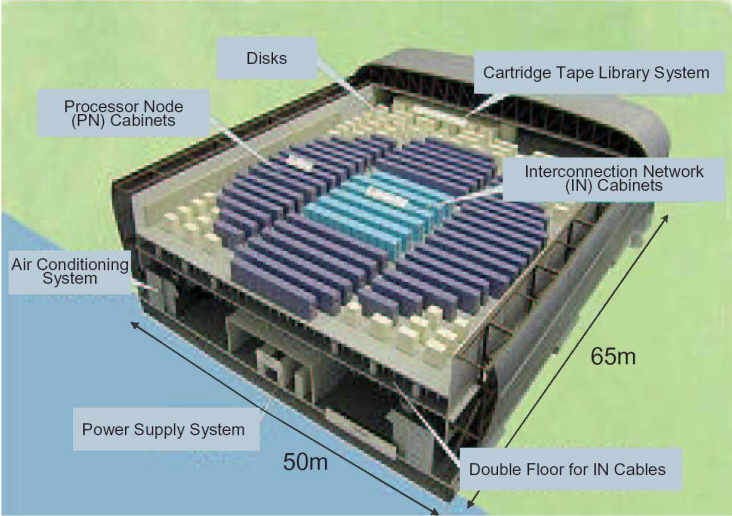
- ▶ Juni 2002 bis Juni 2004 der leistungsfähigste Computer der Welt
- ▶ Kosten: 350 bis 500 Mio US-\$
- ▶ Folie (und die folgenden) von J. Dongarra



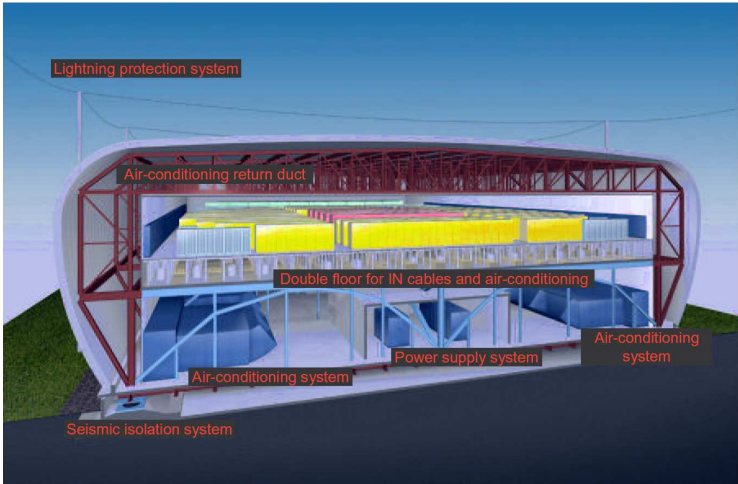
Earth Simulator

- Based on the NEC SX architecture, 640 nodes, each node with 8 vector processors (8 Gflop/s peak per processor), 2 ns cycle time, 16GB shared memory.
 - Total of 5104 total processors, 40 TFlop/s peak, and 10 TB memory.
- It has a single stage crossbar (1800 miles of cable) 83,000 copper cables, 16 GB/s cross section bandwidth.
- 700 TB disk space
- 1.6 PB mass store
- Area of computer = 4 tennis courts, 3 floors

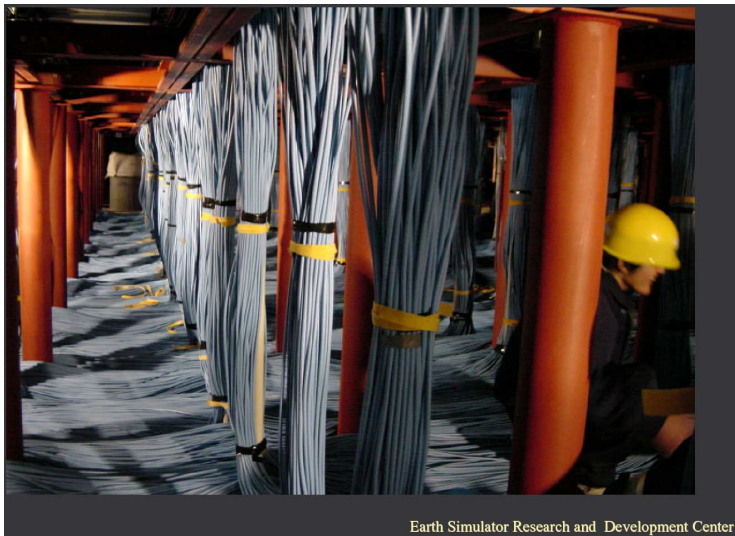
Earth Simulator Gebäude



Earth Simulator Gebäude



... die Kabel



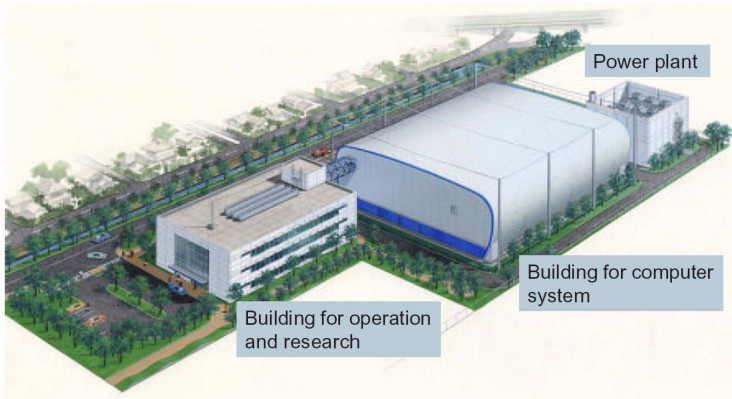
Earth Simulator Research and Development Center

... die Kabel



Earth Simulator Research and Development Center

... und ein Kraftwerk



Terascale Simulation Facility

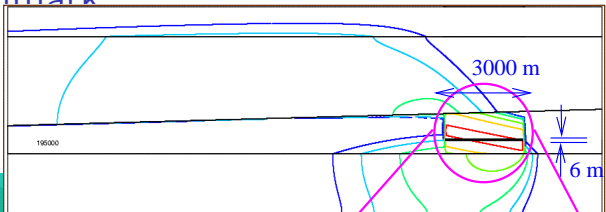


- ▶ Lawrence Livermore National Laboratory
- ▶ BlueGene/L 212992 Prozessoren, 478 TFLOPS seit 2007

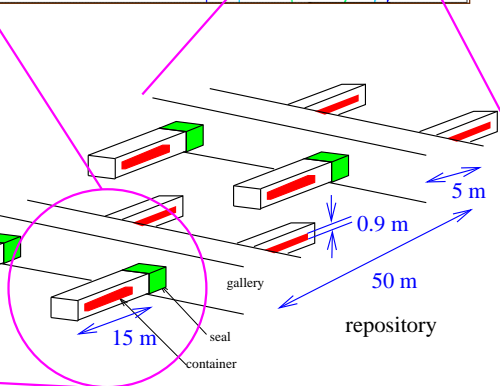
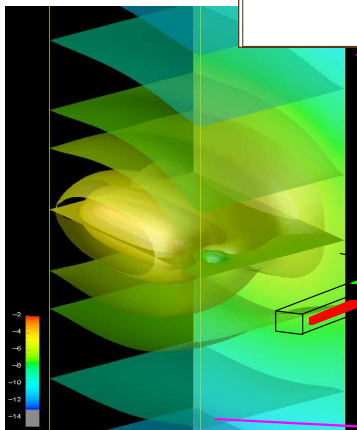
Couplex Benchmark



far field



near field



Couplex1, Couplex2



Organisatorisches

Motivation

Einige Beispielanwendungen

Superrechner

Inhalt der Vorlesung



Couplex Transport Model

Flow equation:

$$\nabla \cdot u = f \quad \text{in } \Omega, \quad u = -K\nabla H$$

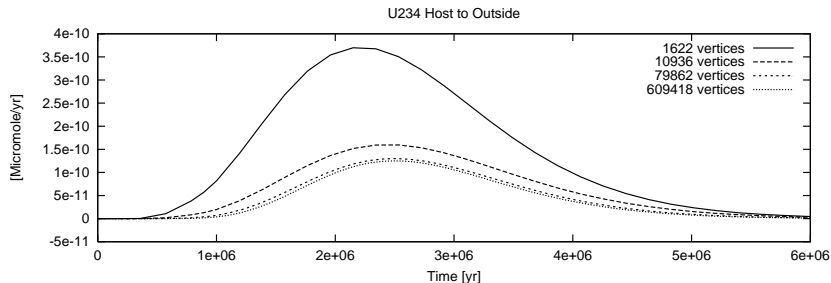
Transport equations $i = \text{silica}, {}^{135}\text{Cs}, {}^{238}\text{Pu}, {}^{242}\text{Pu}, {}^{234}\text{U}, {}^{238}\text{U}$:

$$R\omega \left(\frac{\partial C_i}{\partial t} + \lambda C_i \right) + \nabla \cdot j_i = q(C_1, \dots, C_n) \quad \text{in } \Omega,$$
$$j_i = uC_i - D(u)\nabla C_i$$

Processes:

- ▶ Convection-dispersion (Scheidegger)
- ▶ Radioactive decay ${}^{238}\text{Pu} \rightarrow {}^{234}\text{U}$, ${}^{242}\text{Pu} \rightarrow {}^{238}\text{U}$
- ▶ Precipitation: ${}^{238}\text{Pu}$, ${}^{242}\text{Pu}$, ${}^{234}\text{U}$, ${}^{238}\text{U}$ as dissolved components and (immobile) solid phase

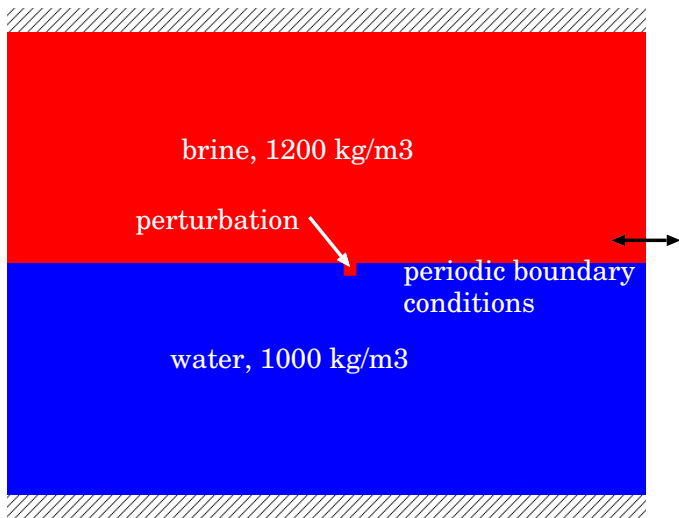
Couplex 2



P	Unbekannte	Par. Rechenzeit [h]	Seq. Rechenzeit [h]
1	120296	2.4	2.4
8	878482	6.3	21.8
64	6703598	6.4	226.3

(Gemeinsame Arbeit mit Stefan Lang)

Dichteströmung



Salzstock, Ozean, Atmosphäre, Sterninneres, Erdmantel, Kaffeetasse



Density driven flow in Porous Media

- ▶ Flow equation (Bussinesq approximation):

$$\nabla \cdot u = f, \quad u = -\frac{K}{\mu}(\nabla p - \varrho(C)g), \quad \varrho(C) = C\varrho_b + (1-C)\varrho_0 \quad \text{in } \Omega$$

- ▶ Transport equation:

$$\frac{\partial(\Phi C)}{\partial t} + \nabla \cdot j + q = 0, \quad j = Cu - D\nabla C$$

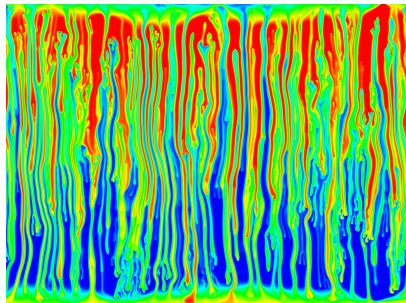
- ▶ *Elliptic* flow equation coupled to *parabolic/hyperbolic* transport equation
- ▶ Cell centered finite volume discretization, second order Godunov scheme for convective terms
- ▶ Cell centered multigrid method
- ▶ Decoupled solution scheme

Simulation Results

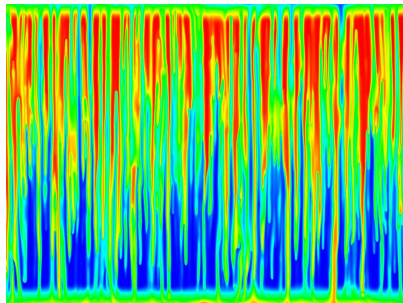


Comparison of two-dimensional and three-dimensional simulation

2d

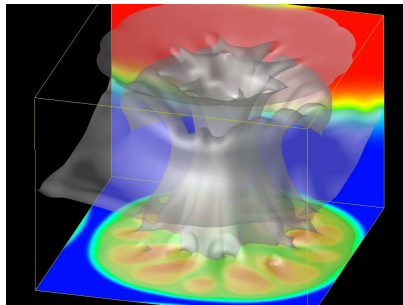
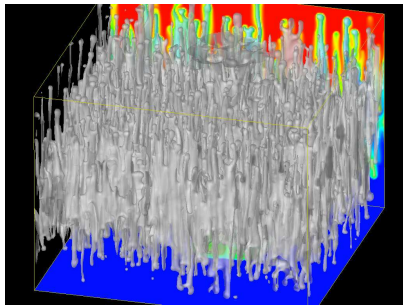


3d



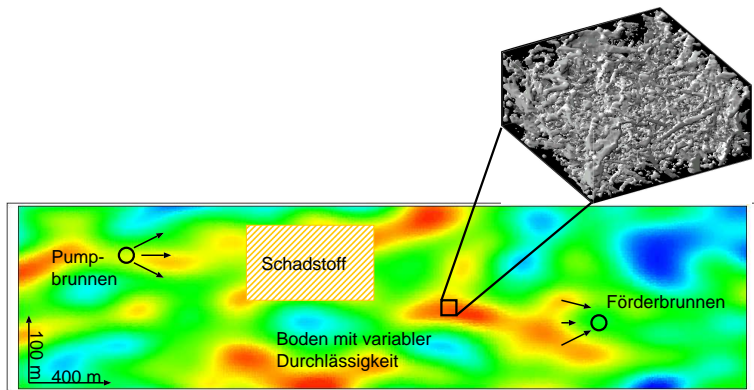


- ▶ Grid size: $1024 \times 1024 \times 768$, about 9000 time steps
- ▶ Solves linear system with $8 \cdot 10^8$ unknowns per time step (requires 30 seconds on 384 processors)
- ▶ Data produced per time step: ≈ 3.2 GByte in binary float
- ▶ Data produced by whole simulation: ≈ 29 TByte
- ▶ Visualization requires intelligent ways of data reduction and/or parallelized visualization tools
- ▶ In addition I/O capabilities of HELICS should be improved in the future

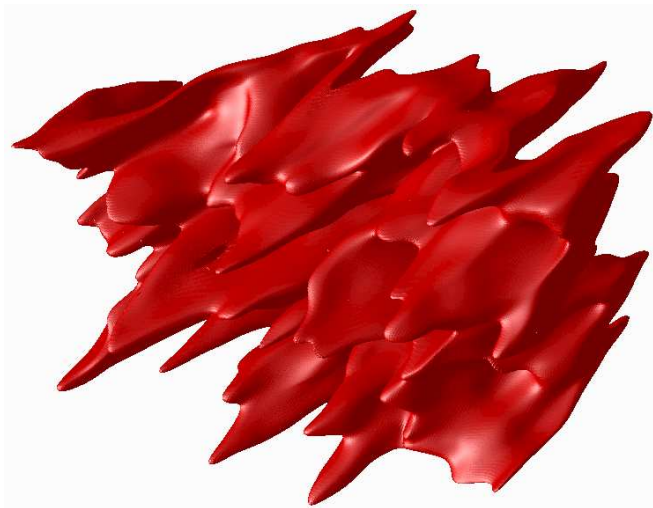


- ▶ Verändere physikalische Parameter (links wenig Diffusion, rechts viel Diffusion)
- ▶ Visualisierung großer Datenmengen ist ein Problem
- ▶ Alle Komponenten einer Simulation müssen parallel bearbeitet werden

In-Situ Reinigungsverfahren



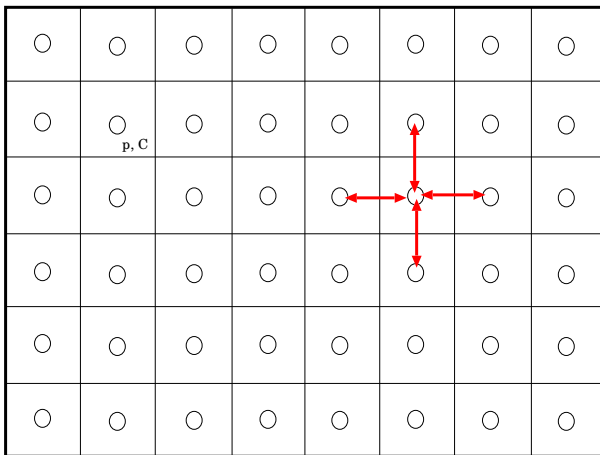
3D Transport in Heterogenem Medium



Zellenzentrierte Finite-Volumen Verfahren (CCMG, HOG)

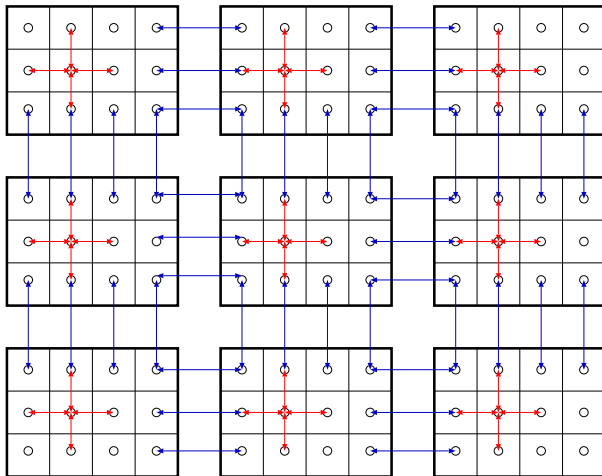
$4604 \times 384 \times 256 \approx 4.5 \cdot 10^8$ Zellen, $\approx 3 \cdot 10^4$ time steps

Numerische Strömungsberechnung



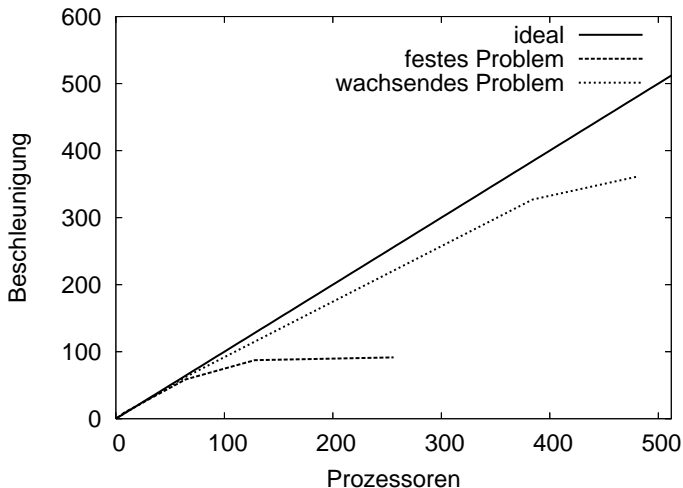
- ▶ Pro Zelle: Phys. Größen wie Druck, Temperatur, Konzentration
- ▶ Je mehr Zellen desto genauer das Ergebnis
- ▶ Lokale Datenabhängigkeiten (Differentiation!)

Datenparallelität



- ▶ Datenzerlegung/Datenaufteilung
- ▶ Datenaustausch (muss minimiert werden)

Beschleunigung



- ▶ Gegebenes Problem immer schneller rechnen
- ▶ Mit mehr Prozessoren immer größere Probleme rechnen



Organisatorisches

Motivation

Einige Beispielanwendungen

Superrechner

Inhalt der Vorlesung

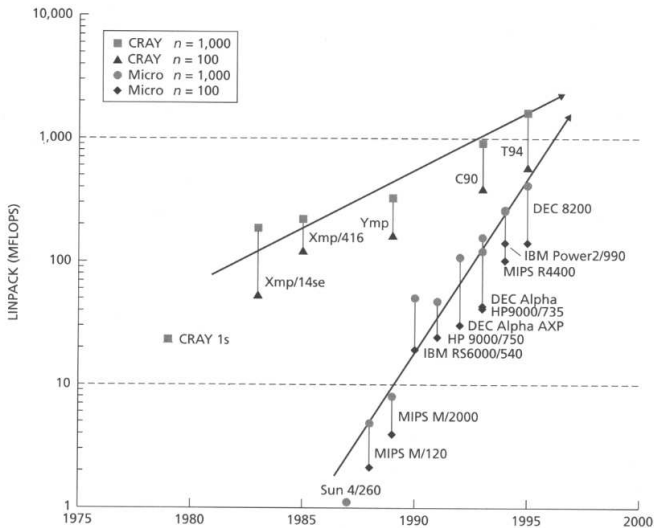


Was ist ein Superrechner ?

Gerät	Jahr	\$	Takt [MHz]	MBytes	MFLOPS
CDC 6600	1964	7M \$	10	0.5	3.3
Cray 1A	1976	8M \$	80	8	20
Cray X-MP/48	1986	15M \$	118	64	220
C90	1996		250	2048	5000
ASCI Red	1997		220	1212000	2388000
Pentium 4	2002	1500	2400	1000	1000
Core 2 XE QX6800	2007	2000	2933	>4000	25400

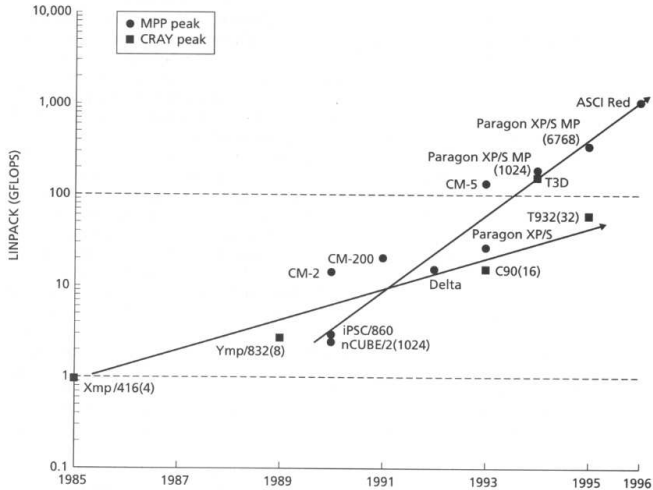
- ▶ X-MP/48: P=4, C90: P=16, ASCI Red: P=9152
- ▶ PC von heute = Superrechner von gestern

Entwicklung Einzelprozessoren in den 90ern¹



¹Culler/Singh, Parallel Computer Architecture.

Entwicklung Multiprozessoren²



²Culler/Singh, Parallel Computer Architecture.

Entwicklung seit den 90er Jahren: Cluster Computer



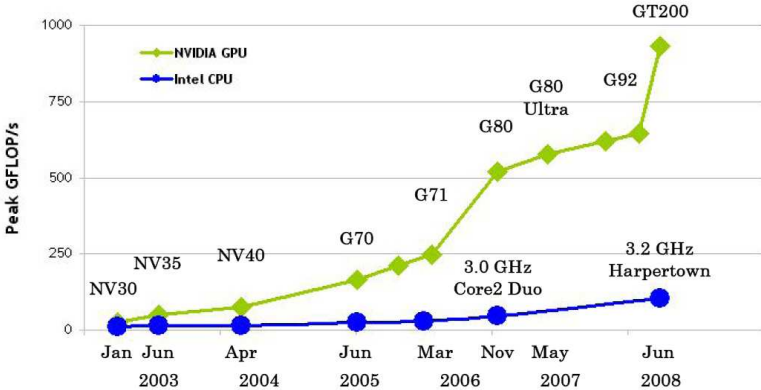
Aufbau aus Standardbauteilen



- ▶ Multicore Prozessoren
- ▶ Coprozessoren Cell, Clearspeed
- ▶ General Purpose GPU-Computing
- ▶ Petaflop Computer: 10^{18} Gleitkommaoperationen pro Sekunde erreicht im Juni 2008 der IBM Roadrunner
- ▶ 122400 Cores, Opteron mit Cell Coprozessoren



General Purpose GPU-Computing



Quelle: NVIDIA



Effiziente Algorithmen

- ▶ Schnelle Rechner sind nicht alles!
- ▶ Rechenzeiten zur (approximativen) Lösung spezieller linearer Gleichungssysteme:

N	Gauß ($\frac{2}{3}N^3$)	Mehrgitter ($100N$)
1.000	0.66 s	10^{-4} s
10.000	660 s	10^{-3} s
100.000	7.6 Tage	10^{-2} s
$1 \cdot 10^6$	21 Jahre	0.1 s
$1 \cdot 10^7$	21.000 Jahre	1 s

(für einen Rechner mit 1 GFLOP/s)

- ▶ Parallelisierung kann einem ineffizienten Verfahren nicht helfen!
- ▶ \Rightarrow Schnelle Rechner **und** schnelle Algorithmen



Organisatorisches

Motivation

Einige Beispielanwendungen

Superrechner

Inhalt der Vorlesung



Inhalt der Vorlesung

I Hardware

- ▶ Prozessorentwicklung, Pipelining, SIMD, MIMD, Caches, Cachekonsistenz, Multicore
- ▶ Verbindungsnetzwerke
- ▶ Grafikprozessoren

II Programmierung

- ▶ gemeinsamer Speicher: Locks, Semaphore
- ▶ PThreads, OpenMP
- ▶ verteilter Speicher: MPI-I, MPI-II
- ▶ RPC, Corba
- ▶ CUDA

III Algorithmen

- ▶ Bewertung von Algorithmen, prinzipielles Vorgehen, Lastverteilung
- ▶ dichtbesetzte Matrizen
- ▶ dünnbesetzte Gleichungssysteme
- ▶ Sortieren, N-Körper-Problem



- ▶ Vorlesung
<http://hal.iwr.uni-heidelberg.de/phlr.html>
- ▶ ASCI Program
<https://asc.llnl.gov/>
- ▶ Seymour Cray
<http://research.microsoft.com/users/gbell/craytalk/>
- ▶ TOP 500 Supercomputer Sites
<http://www.top500.org/>