Zugang zu deutschen Höchstleistungsrechnern

Ein Angebot der HPC Beratung am Regionalen Rechenzentrum Erlangen

Höchstleistungsrechner: Definition

- Versuch einer Definition:
  - Leistungsmerkmale => zeitabhängig
  - Architekturen => zeitabhängig
  - Verfügbarkeit / Aufwand ?

- „A supercomputer is a computer that is just one generation behind the requirements of the large scale users“ (Neil Lincoln, CDC CYBER 205)

Höchstleistungsrechner : Anforderungen an den Benutzer

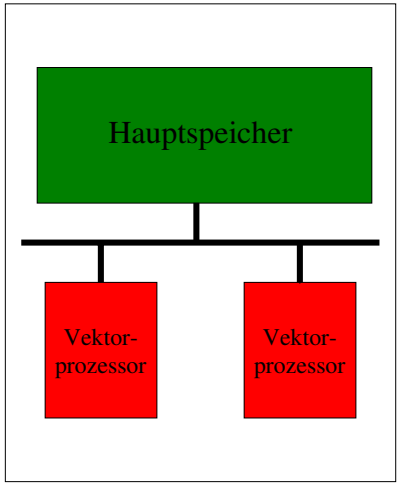
- **Vektorisierung** (Vektorlängen / Rekursionen)

  und / oder

- **Parallelisierung** (Load Balance / Kommunikation)

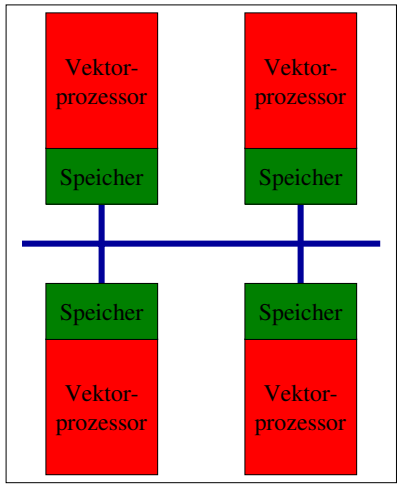
- Geduld
- Flexibilität

Höchstleistungsrechner: Vektorparallelrechner (PVP)



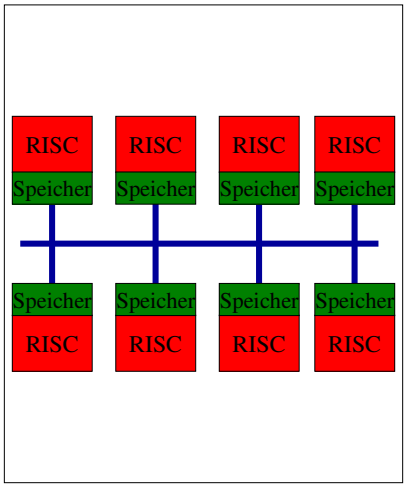
- Spezifikation:
  - Vektorprozessoren
  - Gemeinsamer Hauptspeicher
- Beispiele:
  - CRAY T90, NEC SX
- Konfigurationen:
  - 1 - 8 Prozessoren
  - 1 - 2 GB / Prozessor
  - 1 - 2 GFlops / Prozessor
- **Anforderungen:**
  - **! Vektorisierung !**
  - **(Parallelisierung)**

Höchstleistungsrechner: Parallele Vektorrechner (VPP)



- Spezifikation:
  - Vektorprozessoren
  - Verteilter Speicher
- Beispiele:
  - Fujitsu VPP300/VPP700
- Konfigurationen:
  - 1 - 52 Prozessoren
  - 2 GB / Prozessor
  - 2 GFlops / Prozessor
- **Anforderungen:**
  - **! Vektorisierung !**
  - **Parallelisierung** (MPI)

Höchstleistungsrechner: Massiv parallele Rechner (MPP)



- Spezifikation:
  - RISC-Prozessoren
  - Verteilter Speicher
- Beispiele:
  - IBM SP2, CRAY T3E
- Konfigurationen:
  - 8-512 Prozessoren
  - 128-512 MB / Prozessor
  - 200-400 MFlops / Prozessor
- **Anforderungen:**
  - **skalare Optimierung**
  - **! Parallelisierung !** (MPI)

Höchstleistungsrechner (HLR) in Deutschland

HLR für den
akademischen Sektor

HLR in Deutschland: Auszug aus TOP500 (Nov. 1998)

Rang	RZ	Typ	Prozessoren	Linpack
13	DWD	T3E1200	484	401 GFlops
15	Garching	T3E900	812	355 GFlops
16	**Stuttgart**	**T3E900**	**540**	**341 GFlops**
20	**Jülich**	**T3E600**	**540**	**235 GFlops**
30	**Jülich**	**T3E900**	**268**	**169 GFlops**
36	**Berlin**	**T3E900**	**228**	**144 GFlops**
50	**München**	**VPP**	**52**	**106 GFlops**

Leibniz RechenZentrum München

http://www.lrz.de

- IBM SP2 (MPP)
- CRAY T90 PVP
- Fujitsu VPP700 VPP

- Antragsverfahren:
  - H.W. Bohne (RRZE)
  - Dauer: 1 Tag
- Abrechnungsverfahren:
  - Quota für FAU
  - Tägl. Inkrement
- Sonstiges:
  - Umfangreiche Software
  - Informative WEB-Seiten

IBM SP2 (MPP)

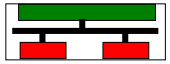
- 77 POWER2 Prozess.
- 16.7 GB Speicher (verteilt)
- 20.7 GFlops
- 334 GB Platten

- Serielle Queues:
  - 1 - 48 h
  - 120 - 1000 MB
- Parallele Queues:
  - 1 - 48 h
  - 8 - 32 Prozessoren mit 128 MB/Prozess.

Zuverlässiger „Rechenknecht“

Hoher serieller Durchsatz

CRAY T90 (PVP)

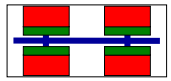


- 4 Vektorprozessoren
- 1 GB Speicher (gemeinsam)
- 7.2 GFlops
- 146 GB Platten

- High / Low Priority Queues
- Sehr guter Vektorcompiler
- Gute Performance auch bei kleinen Vektorlängen

Stabiler Produktionsrechner für rechenintensive Probleme

Fujitsu VPP700 (VPP)



- 52 Vektorprozessoren
- 104 GB Speicher (verteilt)
- 114 GFlops
- 907 GB Platten

- Serielle Queues:
  - 300 - 1800 MB
  - 1 - 24 h
- Parallele Queues:
  - 4 - 16 Prozessoren
  - 1 - 24 h
- Performance Tools

Stabiler Produktionsrechner für rechenintensive und speicherintensive Probleme

HLR in Deutschland: NIC Jülich

John von Neumann - Institut für Computing

John von Neumann Institut für Computing

http://www.kfa-juelich.de/nic

- CRAY T90 (PVP/10)

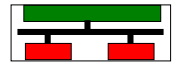
CRAY T3E (MPP):

- 512 Proz./64 GB
- 256 Proz./32 GB
- 256 Proz./128 GB

- Antragsverfahren:
  - Antrag zweimal jährlich (1.1. / 1.7.)
- Abrechnungsverfahren:
  - CPU Kontingent pro Nutzer und Periode
- Sonstiges:
  - Umfangreiche Software
  - Schulungsprogramm

HLR in Deutschland: NIC Jülich : CRAY T90

John von Neumann - Institut für Computing



CRAY T90 (PVP)

- 16 Vektorprozessoren
- 8 GB Speicher (gemeinsam)
- 30 GFlops
- (Gigaring)

- High/Low Priority Queues:
  - max. 4 GB / 10 h
- Bis vor kurzem sehr instabil (Hardwareprobleme)

⇒ (Stabiler) Produktionsrechner für rechenintensive Probleme

HLR in Deutschland: NIC Jülich : CRAY T3E

John von Neumann - Institut für Computing

CRAY T3E (MPP)

CRAY T3E600 : **512 Proz**
64 GB (verteilt) / 300 GFlops
CRAY T3E900 : **256 Proz.**
32 GB (verteilt) / 230 GFlops
CRAY T3E1200 : **256 Proz.**
128 GB (verteilt) / 300 GFlops

- Batch: 512 Proz. /4 h
- Interaktiv: 128 Proz.
- Gute Entwicklungs-/ Testumgebung.
- Globaler Adressraum („shmem“)

⇒ Stabiler (Entwicklungs- und) Produktionsrechner für gut skalierende Probleme

HLR in Deutschland: HLRS

H L R S

Höchstleistungsrechenzentrum Stuttgart

http://www.hlrs.de

- NEC SX4 (PVP/32)
- CRAY T3E (MPP/512)

- Hitachi SR2201 (RUS)
- Intel Paragon (RUS)

- Antragsverfahren:
  - Online Proposal Submission
- Abrechnungsverfahren:
  - Kontingent in DM (Euro)
- Sonstiges:
  - BundeshöchstleistungsRZ
  - Umfangreiche Software
  - Schulungsprogramm

HLR in Deutschland: HLRS : NEC SX4

NEC SX4 (PVP)

- SX4/32:
- 32 Vektorprozessoren
- 64 GFlops
- 24 GB Speicher
- 338 GB Platten

- Serielle Queues:
  - 1-8 GB / 24 - 48 h
- Parallele Queues:
  - 2-16 Prozessoren / 24 h
  - 8 GB Speicher
- Gemeinsamer Speicher / hohe Vektorleistung

Stabiler Produktionsrechner für rechenintensive und speicherintensive Probleme

HLR in Deutschland: HLRS : CRAY T3E

CRAY T3E900 (MPP)

- 512 Prozessoren
- 64 GB (128 MB / Proz.)
- 461 GFlops
- 507 GB Platten

- Queues:
  - Express / Normal
  - 32 - 512 Proz.
  - 1 - 12 h
- Gute Entwicklungs-/ Testumgebung.

Stabiler (Entwicklungs- und) Produktionsrechner für gut skalierende Probleme

HLR in Deutschland: SSC Karlsruhe

Scientific Supercomputing Center Karlsruhe

http://www.ssc.uni-karlsruhe.de

- IBM RS6000/SP : Leistungsstärkster IBM Rechner in Europa

- Antragsverfahren:
  - „Schnupperkontingent“ (1 Woche)
  - Online Proposal Submission
- Abrechnungsverfahren:
  - Kontingent in DM (Euro)
- Sonstiges:
  - Umfangreiche Software

HLR in Deutschland: SSC Karlsruhe: IBM RS6000/SP

IBM RS6000/SP (MPP)

- 256 Prozessoren
- 130 GB Speicher (verteilt)
- 107 GFlops
- 2100 GB Platten

- Produktionsqueues:
  - max. 160 Prozessoren
  - max. 12 h
- Leistungsfähige Knoten:
  - 512 MB / 2 GB Speicher
  - ~ 400 MFlops
  - lokale Platte

Stabiler Produktionsrechner für rechenintensive und speicherintensive Probleme

HPC Beratung am RRZE

Ein Angebot an Wissenschaftler mit
rechen- und/oder speicherintensiven
Problemstellungen

HPC Beratung am RRZE: Angebote und Aufgaben

- *HPC Programmentwicklung*
  - **Parallelisierung** / **Vektorisierung**
  - Fehler-/Performanceanalyse
    - *Zugang zu Höchstleistungsrechnern*
      - Portierung / Antragsverfahren
      - Rechnerspezifische Unterstützung / Optimierung
        - *Koordinierung der HPC Aktivitäten an der FA*
          - HPC Angebot des RRZE
          - Interessen der FAU gegenüber externen RZ

HPC Beratung am RRZE: Beispiel WW VI : Übersicht

- *Problemstellung*
  - CFD Code aus Werkstoffwissenschaften VI
  - Rechenzeit für min. Problemgröße: ca. 2 Tage PII
    - *Ziel*
      - Verringerung der Rechenzeit bei gleichzeitiger Erweiterung der Problemgröße
        - *Vorgehen*
          - Vektorisierung + Zugang VPP / T90
          - Parallelisierung

HPC Beratung am RRZE: WW IV : Performanceergebnisse

Vergleich der seriellen Performance (1)

Vektorisierter Code auf Fujitsu VPP700/ CRAY T90/ NEC SX4/ (IBM SP2)

Benchmark 1: Zeitabhängiges Problem / 56300 CV

Rechner	*CPU*	*Performance*
PII (400MHz)	48 h	~35 MFlops
IBM PowerII	25 h	~72 MFlops
VPP700	4.0 h	450 MFlops
T90	3.5 h	510 MFlops
SX4	3.0 h	600 MFlops

HPC Beratung am RRZE: WW IV : Performanceergebnisse

Vergleich der seriellen Performance (2)

Vektorisierter Code auf Fujitsu VPP700/ CRAY T90/ NEC SX4/ (IBM SP2)

Benchmark 2: 500 Zeitschritte / 126350 CV

Rechner	*CPU*	*Performance*
PII (400MHz)	---	---
IBM PowerII	>48 h	67 MFlops
VPP700	8.1 h	460 MFlops
T90	5.4 h	630 MFlops
SX4	4.5 h	760 MFlops

HPC Beratung am RRZE

Ansprechpartner:
Dr. Gerhard Wellein
Martensstr. 1
91058 Erlangen
Tel.: 09131 85 28737

E-Mail: Gerhard.Wellein@rrze.uni-erlangen.de
http://www.uni-erlangen.de/docs/RRZE/dienste/hpc/index.html

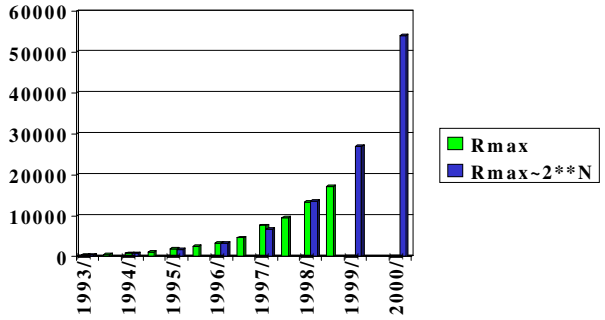
Trends im HPC

•Entwicklungen in den TOP500 und den USA

•**Bundeshöchstleistungsrechenzentrum München**

Trends im HPC: Tendenzen in den TOP500

- Aufnahmekriterium Rmax [MFlops]für die TOP500



- Projektion für 2004: Aufnahmekriterium Rmax: 1 TFlops

Trends im HPC
Die TOP10 (Stand: Nov. 1998)

Rang	Land	Herst.	Rechner	Prozess.	Typ	TFlops
1	USA	Intel	ASCI Red	9152	MPP	1.34
2	USA	SGI	T3E	1084	MPP	0.89
3	USA	SGI	T3E	1324	MPP	0.82
4	USA	SGI	ASCI Blue	6144	CSM	0.69
5	UK	SGI	T3E	876	MPP	0.55
6	USA	IBM	SP Silver	1952	MPP	0.55
7	UK	SGI	T3E	612	MPP	0.51
8	USA	IBM	ASCI Blue	1344	MPP	0.47
9	USA	SGI	T3E	700	MPP	0.45
10	USA	SGI	T3E	1084	MPP	0.45

TFlops: LINPACK (Rmax)
MPP: Massive Parallel Processing
CSM: Clustered Shared Memory (Blue Mountain: 48 x Origin2000)

Trends im HPC:
ASCI in den USA

USA / Department of Energy:
Accelerated Strategic Computing Initiative (ASCI)

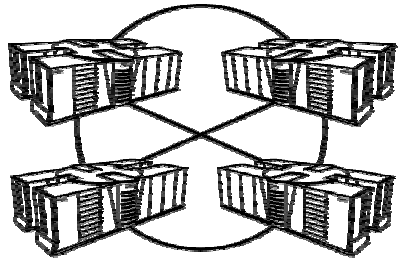
ASCI	Vendor	Peak	Inst.-Year	Inst.-Sit
Red	Intel	1.8 TFlops	1997	Sandi
Blue Mountain	SGI	3.0 TFlops	1998/99	Los Alamos
Blue Pacific	IBM	3.8 TFlops	1998/99	Livermore
White	IBM	10 TFlops	2000	???
???	???	30 TFlops	2002	???

Trends im HPC:
BHLRZ München

Bundeshöchstleistungsrechenzentrum München

- 22.1.: Positive Entscheidung des Wissenschaftsrates:
  - Einrichtung eines BHLRZ in München
  - Installation einer „moderat parallelen Vektoranlage“ innerhalb eines Jahres
  - Bearbeitung von „Grand Challenge“ Vorhaben
  - Bundesweiter Zugang
- 26.1.: Bay. Staatsministerium für Wissenschaft:
  - Anschaffung eines TFlops Rechners

Trends im HPC:
TFlops Rechner: NEC SX5



1 TFlops = 8 SX-5M Knoten
+ IXS Superswitch

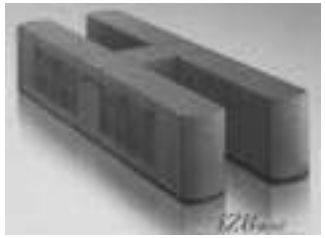
- 8 GFlops pro Vektorprozessor
  (7.5 GFlops LINPACK)
- 1 SX-5M Knoten:
  - max. 16 Vektorprozessoren
  - max. 128 GB gemeinsamer Speicher
- IXS Superswitch:
  - Crossbar: Verbindung zwischen mehreren SX-5M Knoten

Trends im HPC: TFlops Rechner: CRAY SV1

- 4 GFlops MultiStreaming Processor (MSP) ⟺ 4 x 1 GFlops
- 1 SMP Knoten:
  - max. 6 MSP + 8 x 1 GFlops Proz.
  - max. 32 GB gemeinsamer Speicher
- Superclustering der SMP Knoten:
  - max. 32 SMP Knoten

1 TFlops = 32 SMP Knoten

Trends im HPC: TFlops Rechner: Hitachi SR8000



- 8 GFlops COMPAS (CO-operative Micro-Processors in single Address Space)
- 1 Knoten:
  - 8 GFlops
  - max. 8 GB Speicher
- Max. 128 Knoten:
  - LINPACK (20 Knoten): 144 GFlops

1 TFlops = 128 Knoten

Rechenkapazitäten am RRZE

Eine Bestandsaufnahme
der verfügbaren Resourcen und der Benutzeranforderungen

Zentrale Rechner am RRZE: Bestandsaufnahme

- **CSSUN : 8 UltraSparc (250 MHz) + 8 GB gemeinsamer Speicher**
- **VPP300: 6 Vektorprozessoren (2 GFlops) + 12 GB verteilter Speicher**

- Convex: 48 Prozessoren (120 MHz) + 6 GB Speicher
- HP Cluster: 13 HP9000/7xx
- CRAY YMP/EL

➡ Ablösung der älteren Systeme ?!