

Erläuterungen zum HBF - Antrag  
vom 30.08.94

## Parallelrechner (MPP-System)

Der beantragte Parallelrechner soll den Einstieg in die Parallelverarbeitung ermöglichen, von der Anwendungs-Entwicklung bis zum Einsatz fertiger Anwendungen; er soll nicht nur für Nutzer aus der Philipps-Universität Marburg bereitgestellt, sondern in den Verbund hessischer Hoch/Höchstleistungsrechner eingebracht werden und damit von allen hessischen Hochschulen genutzt werden können. Der Parallelrechner soll im Hochschulrechenzentrum der Philipps-Universität Marburg installiert werden und dort 2 ältere Rechner ablösen (CONVEX C230 und DEC VAX 6000-420).

### Inhaltsverzeichnis

1.	Übersicht.....	3
1.1	Zentrale Einrichtung .....	3
1.2	Verantwortlicher Ansprechpartner .....	3
1.3	Beantragte Arbeitsplatzrechner .....	3
1.4	Beantragter Parallelrechner.....	3
1.5	Gesamtbetrag der Anmeldung (inkl. MwSt) .....	3
1.6	Nutzer .....	3
1.7	Konzept.....	4
2.	Komponenten des beantragten Parallelrechners.....	6
3.	Integrationskizze.....	7
4.	Einsatz.....	8
4.1	Projekte des FB Mathematik .....	8
4.2	Projekte des FB Physik.....	10
4.3	Projekte des FB Physikalische Chemie .....	14
4.4	Projekte des FB Chemie.....	15

4.5	Projekte des Wissenschaftlichen Zentrums für Materialwissenschaften.....	16
4.6	Projekte der Universität Gießen.....	16
4.7	Begründung der Notwendigkeit der Beschaffung.....	19
5.	Vorhandene Ausstattung und Nutzung.....	20
5.1	Zentrale Server und Großrechner .....	20
5.2	Kommunikationsnetz .....	20
5.3	Dezentrale Rechner .....	24
5.4	Nutzung der Großrechner.....	24
5.5	Nutzung der hessischen Hoch/Höchstleistungsrechner.....	25
6.	Geprüfte Alternativen und Begründung der Auswahl.....	32
6.1	EG-Ausschreibung und Angebots-Übersicht.....	32
6.2	Katalog klassifizierender Charakteristika .....	34
6.3	Benchmarks .....	38
6.4	Bewertung der Angebote.....	39
7.	Literatur .....	40
8.	Tabellen, ergänzende Unterlagen.....	41
8.1	Klassifizierende Charakteristika.....	41
8.2	NAS Parallel Benchmark.....	53
8.3	Leistungsbeschreibung für die Angebotsabgabe.....	55
8.4	Teilnahmewettbewerb der Firmen.....	63
8.5	Bekanntmachung der Ausschreibung.....	64

## 1. Übersicht<sup>1</sup>

### 1.1 Zentrale Einrichtung

Hochschulrechenzentrum (HRZ) der Philipps-Universität Marburg in Hessen  
Stellen: 31, davon 15 für wiss. Mitarbeiter und 2 für Programmierer.

### 1.2 Verantwortlicher Ansprechpartner

Dr. Jürgen Radloff, Geschäftsführender Direktor  
des Hochschulrechenzentrums der Philipps-Universität Marburg  
Hans-Meerwein-Straße  
35032 Marburg

Telefon: (06421) 28-3521  
Telefax: (06421) 28-6994

### 1.3 Beantragte Arbeitsplatzrechner

Keine; Arbeitsplatzrechner sind vorhanden bzw. werden im Rahmen von CIP und WAP beschafft (vgl. 5.3).

### 1.4 Beantragter Parallelrechner

CONVEX SPP1000/XA mit  
4 Hypernodes à 8 CPUs (100 MHz), 2 FDDI Interfaces  
insgesamt 6.4 GFLOP/s Peak Performance  
2 GB Arbeitsspeicher, 20 GB Plattenspeicher, DAT Laufwerk  
SPP-UNIX, NQS, Sub-Complex Manager  
FORTRAN inkl. Application Compiler und C  
Parallelized Scientific Subroutine Library  
Parallelisierungs-Werkzeuge (Debugger, Analyzer, Trace Tool)

### 1.5 Gesamtbetrag der Anmeldung (inkl. MwSt)

3.229.124 DM

### 1.6 Nutzer

Der Parallelrechner soll von allen Professoren (ca. 410), wiss. Mitarbeitern (ca. 1 400) und Studenten (ca. 18 700) der Philipps-Universität Marburg genutzt werden können; dies werden in erster Linie Angehörige der folgenden Fachbereiche sein (Stand: 31.12.93):

	Professoren	wiss.Mitarbeiter	Studenten
Mathematik (inkl. Fachgebiet Informatik)	20	20	670
Physik	30	71	392
Physikalische Chemie	10	42	zus.mit Chemie
Chemie	20	96	761

Darüber hinaus soll der Parallelrechner über das WIN von den Angehörigen aller anderen hessischen Hochschulen genutzt werden können; über die Zugangsmodalitäten entscheidet der wiss. Beirat Hessischer Höchstleistungsrechner (vgl. 1.7).

Der geplante Einsatz des Parallelrechners wird in Abschnitt 4 anhand ausgewählter Projekte aus der Philipps-Universität Marburg sowie der Justus-Liebig-Universität Gießen erläutert.

<sup>1</sup> Gliederung analog den Leitfäden zu CIP- und WAP-Antrag

## 1.7 Konzept

Die beiden Großrechner CONVEX C230 und DEC VAX 6000-420 werden z.Zt. noch ausgiebig für numerisch-intensive Anwendungen genutzt (vgl. 5.4). Ihr Betrieb ist (bei ca. 500.000 DM Folgekosten/Jahr) mittlerweile jedoch völlig unwirtschaftlich geworden; sie sollen deshalb durch einen leistungsfähigen **Compute Server** abgelöst werden:

- Der Mini-Supercomputer CONVEX C230 hat Ende 1989 für die Nutzer der Universität Marburg den Einstieg in Vektorarithmetik, UNIX und Internet-Services gebracht (Parallelprogrammierung dieses SMP-Systems blieb ungenutzt). Der Rechner wurde schwerpunktmäßig für Anwendungen im Fachbereich Chemie (Leibnizpreisträger) beschafft und von Anfang an voll genutzt.
- Rechner auf der Basis von VAX/VMS hat es im HRZ der Universität Marburg seit 1984 gegeben; die DEC VAX 6000-420 wurde Anfang 1990 als zweiter Zentraler Rechner beschafft (neben einer IBM 4381-R23, die inzwischen stillgelegt ist); seit 1991 ist klar, daß mit diesem Rechner die Unterstützung von VAX/VMS und DECnet durch das HRZ ausläuft. Die Nutzung des Rechners ließ sich langsam an; insbesondere bereitete der Vektorprozessor Probleme.

Informations-, Kommunikations- und Software-Server sind vorhanden (vgl. 5.1); Fileserver-Dienste werden von diesen Servern (soweit möglich) mit übernommen. Gesamtinstallationswert und Folgekosten reduzieren sich durch die Ablösung der beiden Großrechner auf etwa die Hälfte; gleichzeitig steigen die entsprechenden Werte beim Kommunikationsnetz (vgl. 5.2).

Mit der Ablösung der alten Rechner soll zugleich der Einstieg in eine neue Technologie - die Parallelverarbeitung - erfolgen. **Vektorrechnerkapazität** gibt es nämlich in Hessen gegenwärtig in ausreichendem Umfang; 1991/92 wurden 4 Vektorrechner der gleichen Systemfamilie von SNI/Fujitsu beschafft, die unter UNIX betrieben werden und von allen hessischen Hochschulen genutzt werden können:

- Hessischer Höchstleistungsrechner (HHLR) an der TH Darmstadt; S400/40 mit 4 skalaren und 2 vektoriellen Einheiten, 2 GB Arbeitsspeicher und insgesamt 5 GFLOP/s Peak Performance.
- Hochleistungsrechner (HLR) an den Universitäten Frankfurt, Gießen und Kassel; S200/10 bzw. S100/10 mit je 1 skalaren und 1 vektoriellen Einheit, 256 MB Arbeitsspeicher und 1 bzw. 0.5 GFLOP/s Peak Performance.

Marburger Nutzer haben von Anfang an auf allen 4 Vektorrechnern gearbeitet und dabei bis Juni 1994 insgesamt 30,2 % der Rechenzeit genutzt (vgl. 5.5).

Gerade in Betrieb gegangen ist (im August 1994) an der TH Darmstadt auch noch ein Vektor-Parallelrechner VPP500/4+1 von SNI/Fujitsu, als Abrundung der Maßnahme von 1991/92; der Rechner enthält 4 Prozessorelemente mit je 1.6 GFLOP/s Peak Performance und 256 MB Arbeitsspeicher, die über einen Crossbar gekoppelt sind; er benötigt die S400/40 als Frontend-System und kann nur im Batch-Betrieb (via NQS) genutzt werden. Für 1996 ist der Austausch durch einen Nachfolgerechner VPPC/4 geplant, der stand-alone betrieben und auch interaktiv genutzt werden kann.

Betrieb und Benutzung der hessischen Hoch/Höchstleistungsrechner sind in einer Vereinbarung geregelt [HMWK 91]; die Koordination liegt bei einem **Wissenschaftlichen Beirat**, dem aus jeder Universität 2 Wissenschaftler und der HRZ-Leiter angehören sowie je 1 Vertreter der Fach- und Kunsthochschulen, der GSI in Darmstadt und des Ministeriums für Wissenschaft und Kunst. Diesem Beirat obliegen Aufgaben wie z.B. die Fortschreibung des Landeskonzepts bzgl. Höchstleistungsrechnerkapazität, die Abgabe von Empfehlungen zur Hardware- und Software-Ausstattung sowie die Kontingentierung von Rechenzeit. Der Beirat hat die Beschaffung eines Parallelrechners an der Universität Marburg in mehreren Sitzungen behandelt und diese Initiative begrüßt; der Parallelrechner soll in den Verbund der hessischen Hoch/Höchstleistungsrechner eingebracht werden.

Es geht also nicht darum, Vektorrechnerkapazität durch Parallelrechnerkapazität zu ersetzen; es geht vielmehr darum, neben Vektorrechnerkapazität (z.Zt. insgesamt ca. 13.4 GFLOP/s für ca. 27.2 Mio. DM in 1991/92) auch **Parallelrechnerkapazität** (insgesamt 6.4 GFLOP/s für 3.2 Mio. DM in 1994/95) zur Verfügung zu stellen. Für Wissenschaftler in Hessen sollen Voraussetzungen geschaffen werden, parallele Anwendungen entwickeln zu können, und dies nach Möglichkeit auf der Basis unterschiedlicher Programmiermodelle und Programmiersprachen. Schließlich sollen

fertige, parallelisierte Anwendungen, insbesondere aus dem Bereich Chemie (Hauptnutzer der Universität Marburg) eingesetzt werden können; echte Produktion heißt das Ziel.

Innerhalb der Universität ist die Entscheidung zur Beschaffung eines Parallelrechners vom Ständigen Ausschuß für Datenverarbeitung am 01.07.93 gefällt worden. Die **Anforderungen an den Parallelrechner** resultieren aus allgemeinen Trends (vgl. 7.) sowie spezifischem Bedarf und lauten kurz gefaßt:

- MIMD-Architektur
- möglichst ohne Frontend
- interaktiver Zugang, verteilte Verarbeitung und Batch-Nutzung
- Shared Memory und Message Passing Programmiermodell
- hohe Rechenleistung (> 3 GFLOP/s) und großer Arbeitsspeicher (> 1GB)
- möglichst hohe Knotenleistung (> 100 MFLOP/s)
- leistungsfähiges Verbindungsnetzwerk
- ein Betriebssystem für den gesamten Rechner, Benutzersicht UNIX und ausgereifte Systemmanagement-Werkzeuge
- FORTRAN und C inkl. Parallelisierungs-Hilfen, Laufzeitbibliotheken und ausgereifte Programmentwicklungs-Werkzeuge
- PVM, PARMACS, MPI
- Verfügbarkeit von NAG und IMSL
- Verfügbarkeit von GAUSSIAN, TURBOMOLE, ACES II, GAMESS und MOLPRO

M wie massiv ist mit dem Begriff MPP verbunden geblieben, trotz Schwerpunktverlagerung von der SIMD-Architektur mit einfachen Knoten zur MIMD-Architektur mit "fetten" Knoten. M wie moderat würde die Anforderungen besser kennzeichnen; wichtig sind jedoch vor allem Parallelität und Skalierbarkeit (deshalb ist der Begriff MPP-System eingeklammert).

Die Anforderungen sind in eine Leistungsbeschreibung (vgl. 8.3) eingeflossen; in der Zeit vom 21.02. bis 25.04.94 wurde eine EG-weite Ausschreibung durchgeführt (vgl. 6.1; dort ist auch erläutert warum), aus der insgesamt **7 Angebote** hervorgegangen sind:

- |            |                   |   |
|------------|-------------------|---|
| - CONVEX:  | SPP 1000/XA       | * |
| - Cray:    | Y-MP + T3D        |   |
| - IBM:     | SP2               | * |
| - Intel:   | Paragon XP/S - 5+ |   |
| - nCUBE:   | nCUBE 3           |   |
| - SNI/KSR: | KSR2              | * |
| - TMC:     | CM-5E             |   |

Eine Übersicht über alle Angebote ist in einer Tabelle zusammengefaßt (vgl. 6.1). Mit der Leistungsbeschreibung wurden die Firmen gleichzeitig aufgefordert, Charakteristika ihrer Parallelrechner anzugeben (vgl. 6.2); dabei wurde auf einen Katalog aus Köln/Jülich zurückgegriffen [Bön]; die Angaben sind (soweit verwertbar) ebenfalls in Tabellen zusammengefaßt (vgl. 8.1). Darüber hinaus wurde den Firmen in der Leistungsbeschreibung angekündigt, daß Benchmarks aus dem Bereich Chemie vorgesehen seien.

Übrig geblieben sind schließlich nur **3 Angebote (\*)**: Das Angebot von Cray war nicht ganz ernst gemeint, Intel hat nach Übergabe der Chemie-Benchmarks verzichtet und TMC ist dem Shakeout erlegen; den Rechner von nCUBE gibt es noch nicht (vgl. 6.1). Damit hat sich eine Reduzierung auf Rechner in engerer Wahl erübrigt.

Bei der **Bewertung der angebotenen Parallelrechner** (vgl. 6.4) darf das Preis/Leistungsverhältnis nicht im Vordergrund stehen; vorrangig aus Marburger Sicht<sup>2</sup> sind gegenwärtige Stellung und Zukunftschancen von Hersteller und System. Der Shakeout am MPP-Markt ist in vollem Gange [FLASHs, PAMG 94]; bei den Ereignissen um TMC wurden auch andere Firmen genannt. CONVEX hat in den 80ern als einziger Hersteller den Kampf am Mini-Supercomputer-Markt überlebt<sup>3</sup>; die

<sup>2</sup> mit 20jähriger TR4- bzw. TR440-Erfahrung

<sup>3</sup> Bei der Marburger Rechnerauswahl in 1989 war Multiflow noch mit im Rennen; im Frühjahr 1990 mußte die Firma dann aufgeben.

nächsten Monate werden zeigen, wie erfolgreich CONVEX am MPP-Markt sein wird. Die Finanzprobleme von KSR Ende 1993 sowie die Ausrichtung auf den kommerziellen Markt Mitte 1994 sind hinlänglich bekannt. Daß IBM sich aus dem techn./wiss. Markt noch einmal zurückziehen wird, darf bezweifelt werden; beachtenswert ist, daß die SP-Reihe nicht IBM's einzige MPP-Initiative ist.

Hinsichtlich Architektur, Betriebssystem und Programmiermodellen entsprechen am ehesten die beiden Systeme von CONVEX und SNI/KSR obigen Anforderungen an den Parallelrechner; beide unterstützen Shared Memory und Message Passing Programmiermodell. Beim Shared Memory Modell erleichtert der globale Adreßraum Fehlersuche und Programmanalyse; die Parallelisierung kann schrittweise erfolgen, da die Daten nicht von Anfang an verteilt werden müssen. Das System von IBM unterstützt nur das Message Passing Modell, basiert aber auf besonders leistungsfähigen Prozessoren; dies zeigen sowohl die NAS Parallel als auch die Chemie-Benchmarks. Kein System ist in jeder Hinsicht überlegen; durch die Bewertung der Stärken und Schwächen (vgl. 6.4) lassen sich jedoch Prioritäten setzen.

**Folgerung:** Zum Zeitpunkt der Antragstellung wird dem System SPP 1000/XA der Firma CONVEX die höchste Priorität eingeräumt; dementsprechend wird dieses System beantragt (vgl. 2.). Damit weitere Entwicklungen berücksichtigt werden können - z.B. neuere Ankündigungen, geänderte Preise, Entwicklungen am Markt und Erfahrungen aus anderen Rechenzentren -, soll die Auswahlentscheidung zum Zeitpunkt der Beschaffung überprüft und ggf. korrigiert werden.

## 2. Komponenten des beantragten Parallelrechners

Beantragt wird ein SPP 1000/XA System von CONVEX. Diesem System wird zum Zeitpunkt der Antragstellung die höchste Priorität eingeräumt; die Auswahlentscheidung wird zum Zeitpunkt der Beschaffung überprüft und ggf. korrigiert (vgl. 1.7).

		Listenpreis/DM	
1	SPP1000 XA-32	SPP 1000/XA-System with 32 CPUs (100 MHz) 1.024 MB Memory, I/O Prozessor, I/O chassis 2 Hypernode Towers, Tower and Test Station 4 Hypernode chassis, Power Supplies	3.577.650,--
4	MPU1000-QH	Hypernode Memory Upgrade, 256 to 512 MB	327.680,--
1	DAR-1000	Base Disk and DAT subsystem includes: SCSI-2 Fast/Wide SBUS I/O Controller 2 x 2 GB Disk Drive 1 x DDS-2 DAT Tape drive	44.340,--
2	DAR-1100	High Capacity Disk add on includes: SCSI-2 Fast/Wide SBUS I/O Controller 4 x 2 GB Disk Drive	137.140,--
2	FDC 1000	FDDI Fiber Optic Controller	32.060,--
1	SPP UX-16	SPP-UNIX Operating System for up to 16 concurrent users	16.420,--
1	NQS-64 CPU	CONVEX NQS+ Batch Processing Manager-up to 64 CPUs	52.300,--
1	APC-1	Application Compiler for 1 user	10.750,--
1	FC-5	CONVEX Parallelizing Fortran for up to 5 users	46.240,--
1	CC-5	CONVEX Parallelizing C-Compiler for up to 5 users	46.240,--
1	CXDB-5	Parallel optimized code Debugger for up to 5 users	8.400,--

1	CXPA-5	Parallel Performance Analyzer for up to 5 users	8.400,--
1	CX Tools-5	Parallel Application Development Environment up to 5 users	34.050,--
1	MLIB-8CPU	CONVEX Parallelized Scientific Subroutine Library up to 8 CPUs	29.030,--
1	CXTRACE-5	CONVEX Trace Tool + PVM Library for up to 5 users	21.020,--
1	SCM-64CPU	Sub-Complex Manager for up to 64 CPUs	<u>42.230,--</u>
Summe Listenpreise			4.433.950,--
abzgl. Forschungsrabatt für die Universität Marburg			<u>1.626.016,--</u>
Endpreis			2.807.934,--
<b>Endpreis inkl. 15% MwSt</b>			<b>3.229.124,--</b>

### 3. Integrationskizze

Die Skizze zeigt die Integration des Parallelrechners in die Server-Ausstattung des HRZ; Skizzen zu den Parallelrechnern befinden sich bei den Angeboten. Der FDDI-Ring für NIS, NFS, etc. ist vorhanden; der Benutzer-Zugang auf die Server erfolgt zur Zeit noch über Ethernet, das FDDI-Backbone befindet sich im Aufbau (vgl. 5.2). Der Router mit dem 2 MBit/s Win-Anschluß ist vorhanden.

