

PHILIPPS - UNIVERSITÄT MARBURG
HOCHSCHULRECHENZENTRUM (HRZ)

Dr. J. Radloff, Geschäftsführender Direktor



Hochschulrechenzentrum der Philipps-Universität Marburg
Hans-Meerwein-Straße, 35032 Marburg

An die
Deutsche Forschungsgemeinschaft
Herrn Dr. J
Kennedyallee 40

53170 Bonn

Marburg, den 25.08.95

Paketpost: 35043 Marburg
Hans-Meerwein-Straße
Telefon: (06421) 28-3516/3521
Telefax: (06421) 28-6994
Email: box@hrz.uni-marburg.de
HRZ/Ra/Nau 0092

Betr.: HBFG-Antrag: **Parallelrechner (MPP-System)** vom 30.08.94

DFG Kz III P 2 - 3772-118-275

hier: **Endgültige Gerätewahl: IBM SP2**

Bezug: Ihr Schreiben vom 03.01.95, mein Bericht vom 07.02.95

DFG-Empfehlung vom 27.04.95

Mein Zwischenbericht vom 31.07.95, unser Telefonat vom 02.08.95

Anlg.: IBM SP2 Konfigurationsskizzen

Angebots-Übersicht

Überlegungen zur Gerätewahl

File-, Backup- und Archive-Service

Angebote

Sehr geehrter Herr Dr.

die Maßnahme ist mittlerweile vom Wissenschaftsrat empfohlen (Drs 2083/95 vom 21.06.95) und vom Hessischen Ministerium für Wissenschaft und Kunst bewilligt worden (Erlaß W II 5.2-986/1223-4- vom 13.07.95). Danach soll die endgültige Gerätewahl der DFG vorgelegt und erläutert werden; diese ist nach weiteren Präsentationen auf der Basis neuer Angebote von der eingesetzten Arbeitsgruppe (bestehend aus Wissenschaftlern verschiedener Fachbereiche und Mitarbeitern des HRZ) am 21.08.95 entschieden worden.

Gemäß meinem Bericht vom 07.02.95 waren nur noch die beiden Systeme CONVEX SPP/1000 und IBM SP2 verblieben; dabei hatten wir für das SPP/1000 System votiert, und zwar im wesentlichen wegen des konzeptionellen System-Ansatzes und der Unterstützung beider Programmier-Modelle (Message Passing und Shared Memory). Wir sind nach wie vor der Ansicht, daß dem Shared Memory Programmier-Modell die Zukunft gehört; dies hat die jüngste Supercomputer-Tagung in Mannheim auch bestätigt. Dennoch müssen wir unser Votum aufgrund der geplanten Veränderungen in der Hessischen Rechnerlandschaft und des momentanen SPP/1000 Erscheinungsbildes revidieren (vgl. *Überlegungen zur Gerätewahl* in der Anlage); die **Gerätewahl** ist

zugunsten der IBM SP2 entschieden worden, die von Anfang an als Auffangvariante galt. Die SP2 unterstützt z.Zt. zwar nur das Message Passing Programmier-Modell, IBM strebt aber die Shared Memory Funktionalität an (wenn auch vielleicht nur realisiert mit Hilfe von Message Passing).

Nach unserem Votum vom 07.02.95 waren die Firmen **Cray Research** und **SNI/SGI** noch einmal aufgewacht und hatten neue Systemvorschläge unterbreitet. Entsprechende Präsentationen für die Anwender fanden im Juli bzw. August statt; schließlich haben auch diese Firmen konkrete Angebote vorgelegt, die dann allerdings doch nicht die Auffangvariante IBM SP2 ausstechen konnten. An dem Systemvorschlag J90 --> T3E von Cray Research zeigten die Anwender geringes Interesse; zum einen besteht kein zusätzlicher Bedarf an Vektorrechner-Kapazität, zum anderen hat es die Firma im Laufe der gesamten Rechnerauswahl versäumt, genügend Informationen über ihre Parallelrechner-Linie zu vermitteln. Dagegen zeigten die Anwender für den Systemvorschlag PowerChallenge --> DSM-System von SNI/SGI großes Interesse; hier war aber die Kernfrage, ob es den zukünftigen Parallelrechner (mit Global Shared Memory) je geben wird bzw. wann, denn bis jetzt liegen erst wenig Informationen darüber vor; die leidvollen Erfahrungen bei der Rechnerauswahl (mit TMC, KSR, Cray) sowie die Rücknahme der Entscheidung für die CONVEX SPP/1000 haben die Risikobereitschaft sowohl bei den Anwendern als auch beim HRZ gedämpft. Die Angebote sind mit jeweils mehreren Varianten recht umfangreich ausgefallen, deswegen liegt eine *Angebots-Übersicht* als Anlage bei; darüber hinaus sind die *Überlegungen zur Geräteauswahl* in einer Anlage zusammengefaßt.

Bei der Entscheidung zugunsten der IBM SP2 ging es nicht nur darum, einen jetzt verfügbaren Parallelrechner zu beschaffen; wichtig war darüber hinaus, auch noch zukünftige Technologie erwerben zu können:

- Zielsystem ist eine IBM SP2 mit 3 vollausgebauten Frames, inkl. dem für 1996 angekündigten neuen High Performance Switch sowie Knoten neuer Technologie ab voraussichtlich 1997 in einem der Frames, vgl. *IBM SP2 Konfigurationsskizzen* in der Anlage.
- Im Rahmen der Bewilligung (von 3.23 Mio. DM) könnte jetzt eine SP2 mit insgesamt 46 Thin Nodes 2 beschafft werden (Variante 0).
- Der neue Switch (TBS anstelle HPS+) soll im ersten Halbjahr 1996 verfügbar sein; er soll zwischen je zwei Knoten die Latency von 39 μ s auf 35 μ s vermindern sowie die maximal erreichbare Point-to-Point Bandwidth von 36 MB/s auf 96 MB/s erhöhen (sustained). Um den Übergang zu dieser Technologie zu ermöglichen, soll zunächst nur Frame 1 mit einem HPS+ Switch ausgestattet werden, Frames 2 und 3 mit Ethernet; sobald der neue TBS Switch verfügbar ist, soll er in alle 3 Frames eingebaut werden. Dieser Migrationsschritt reduziert die Anzahl der Knoten auf 42 (Variante 1).
- Entgegen früheren Aussagen soll die POWER2-Architektur nun doch fortgesetzt werden; Prognosen besagen, daß die Frequenz der POWER2-Prozessoren von 66.7 MHz auf ca. 135 MHz angehoben werden soll, was insbesondere eine Verdopplung der Gleitpunkt-Leistung bedeuten würde; entsprechende Knoten dürften aber erst 1997 verfügbar werden. Um den Übergang zu dieser Technologie zu ermöglichen, soll jetzt nur ein Teil des Finanzvolumens (ca. 2.63 Mio. DM) für die Beschaffung des Parallelrechners und der Rest (ca. 0.6 Mio. DM) für die Beschaffung eines File-, Backup- und Archive-Service genutzt werden. Die Beschaffung eines derartigen Service war ursprünglich erst für 1996 geplant; wenn sie jetzt vorgezogen würde, könnte der Ausbau des Parallelrechners für Anfang 1997 eingeplant werden. Dieser Migrationsschritt bedeutet, daß zunächst nur 34 + 1 Knoten beschafft werden (Variante 2).

- Die aus den Migrationsschritten resultierenden Nachteile werden von den Anwendern bewußt in Kauf genommen. Der neue Switch ist besonders wichtig, andererseits muß aber ab Installation wenigstens für einen Teil der Knoten der jetzige Switch verfügbar sein. Bzgl. der Knoten ist Vollausbau von Anfang an nicht erforderlich, hier kann der Ausbau dem Bedarf angepaßt werden; bei Beschaffung neuer Knoten bietet IBM den Austausch der 4 Knoten in Frame 3 an, so daß dieser dann nur mit Knoten neuer Technologie bestückt wäre (Nachtrag).
- Im Rahmen des laufenden Antrags soll deshalb eine Beschaffung gemäß IBM-Angebot vom 31.07.95 (d.h. Variante 2) realisiert werden; aus dem Nachtrag vom 21.08.95 soll Alternative 2 (Austausch 4 Knoten) genutzt werden; die Feinabstimmung der Konfiguration ist in Arbeit. Für den Ausbau des Parallelrechners um max. 12 Knoten in 1997 muß ein neuer Antrag gestellt werden; einschließlich Austausch der Knoten kann damit die Gleitpunkt-Leistung des Parallelrechners um voraussichtlich 80 % gesteigert werden.

Einen zentral für die gesamte Universität bereitgestellten **File-, Backup- und Archive-Service** gibt es noch nicht. Die zentralen Server IBM RS/6000 (2 * Mod. 580 + 2 * Mod. 590) des HRZ stellen eine gewisse Plattenkapazität (insgesamt ca. 50 GB) zur Verfügung; davon werden turnusmäßig Benutzer-Files (ca. 20 GB) vom Operator auf einzelne Bänder gesichert. Das ist aber auch alles; es gibt kein automatisches Backup, keine File Migration und keinen Band-Roboter; insbesondere können keine Files der dezentralen Server (z.Zt. ca. 20 UNIX Server und 25 NOVELL NetWare Server) gesichert und archiviert werden. In diesen zentralen Dienst muß nun natürlich auch der Parallelrechner integriert werden; dies insbesondere, da auch die anderen hessischen Hochschulen diesen Rechner nutzen können. Auswahlüberlegungen sind in der Anlage *File-, Backup- und Archive-Service* zusammengefaßt.

Ich bitte um weitere zügige Bearbeitung, da die **Beschaffung noch im laufenden Haushaltsjahr** abgewickelt werden muß.

Mit freundlichen Grüßen

