

3.5 Parallelrechner in Marburg

In Marburg ist ab 1993 die Beschaffung eines Parallelrechners geplant worden, als Nachfolge des "Mini-Supercomputers" CONVEX C230 (der von Dez. 1989 bis Aug. 1995 betrieben wurde). Ausschreibung, Auswahl und HBFG-Antrag erfolgten im Frühjahr/Sommer 1994; Begutachtung durch die DFG, Bewilligung und endgültige Gerätewahl dauerten bis Sommer 1995. Die Universität hatte lange auf ein Virtual Shared Memory System gesetzt (von CONVEX), sich dann aber doch für ein Distributed Memory System IBM RS/6000 SP entschieden; bevorzugtes Programmiermodell dieses Systems ist Message Passing, auf der Basis von Bibliotheken wie MPI oder PVM.

Die IBM RS/6000 SP ist Anfang Oktober 1995 geliefert und anschließend im Hochschulrechenzentrum installiert worden; die offizielle Freigabe für Benutzer erfolgte am 15.12.95. Im Berichtsjahr 1996 ist der Parallelrechner mit einem neuen Switch ausgestattet worden; deshalb war auf die Beschaffung des alten HPS+ Switches für das gesamte System in 1995 verzichtet worden (vgl. vorangegangenen Jahresbericht). Dokumentationen zur Hardware- und Softwareausstattung zum Benutzer-Zugang und zum Betrieb, etc. wurden von Anfang an im WWW bereitgestellt (vgl. <http://www.uni-marburg.de/hrz/sp/welcome.html>). Darüber hinaus werden laufend Workshops für Benutzer angeboten.

Der Parallelrechner kann von allen hessischen Hochschulen genutzt werden; aufgrund der Finanzierung kann die Universität Marburg 60 % der Rechenzeit für ihre Nutzer beanspruchen. Es können sowohl fertige parallele Anwendungen eingesetzt als auch neue entwickelt werden; schließlich können einzelne Knoten des Parallelrechners auch seriell genutzt werden.

Das Accounting ist Anfang März 1996 angelaufen, so daß erstmalig mit diesem Jahresbericht auch Betriebs- und Nutzerstatistiken vorliegen.

3.5.1 Rechnerausstattung

Die IBM RS/6000 SP (wie Scalable POWERparallel System) besteht aus Knoten, die über einen Switch verbunden sind. Die Knoten entsprechen RS/6000 Workstations mit POWER2 Prozessoren; der Switch verbindet jeden Knoten mit jedem anderen Knoten.

IBM RS/6000 SP

Anzahl Frames	3	
Anzahl Knoten/Frame	max. 16	
Anzahl Knoten insgesamt	35	
Arbeitsspeicher insgesamt	8.2	GByte
Plattenspeicher intern insgesamt	152.2	GByte
Plattenspeicher extern insgesamt	90.0	GByte
Peak Floating Point Performance insgesamt	9.3	GFLOP/s

Ausstattung der Knoten

- 16 Thin Nodes 2 à 128 MB Arbeitsspeicher und 2.2 GB Plattenspeicher
- 14 Thin Nodes 2 à 256 MB Arbeitsspeicher und 4.5 GB Plattenspeicher, davon 10 mit 2 MB Level 2 Cache
- 4 Thin Nodes 2 à 512 MB Arbeitsspeicher, 2*4.5 GB Plattenspeicher und 2 MB Level 2 Cache
- 1 Wide Node mit 512 MB Arbeitsspeicher und 4*4.5 GB Plattenspeicher

Spezifikation der Knoten

Wide Nodes und Thin Nodes 2 gehören beide zur superskalaren POWER2-Architektur mit 1 Instruction, 2 Floating Point und 2 Integer Units, bis zu 6 Instructions/Cycle, 66.7 MHz und 266.7 MFLOP/s Peak Performance. Unterschiede sind:

	Wide Node	Thin Node 2
Daten Cache	256 KB	128 KB
Level 2 Cache	-	opt. 2 MB
Prozessor-Cache Bus	256 Bit	256 Bit
Arbeitsspeicher	64 MB - 2 GB	64 - 512 MB
Cache-Arbeitsspeicher Bus	256 Bit	128 Bit

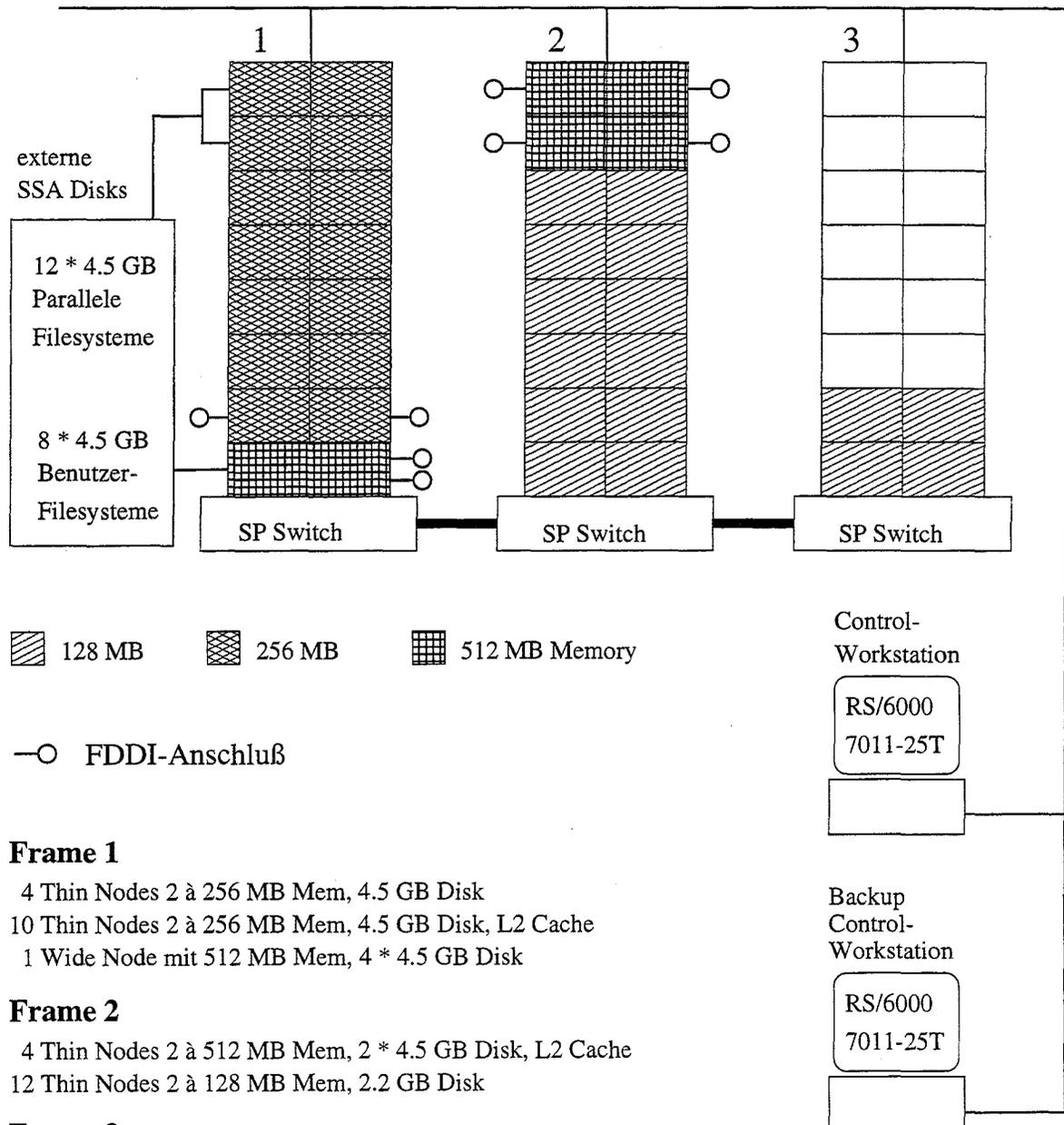
Switch

Any-to-any Multi-Stage-Switch: 8 Bit parallel, bidirektional, max. 96 MB/s je Richtung. Der Übergang vom Vorgängermodell HPS+ in Frame 1 zum SP Switch in allen drei Frames erfolgte Anfang Juli 1996.

HRZ Uni Marburg

Parallelrechner IBM SP

Ethernet mit allen Knoten



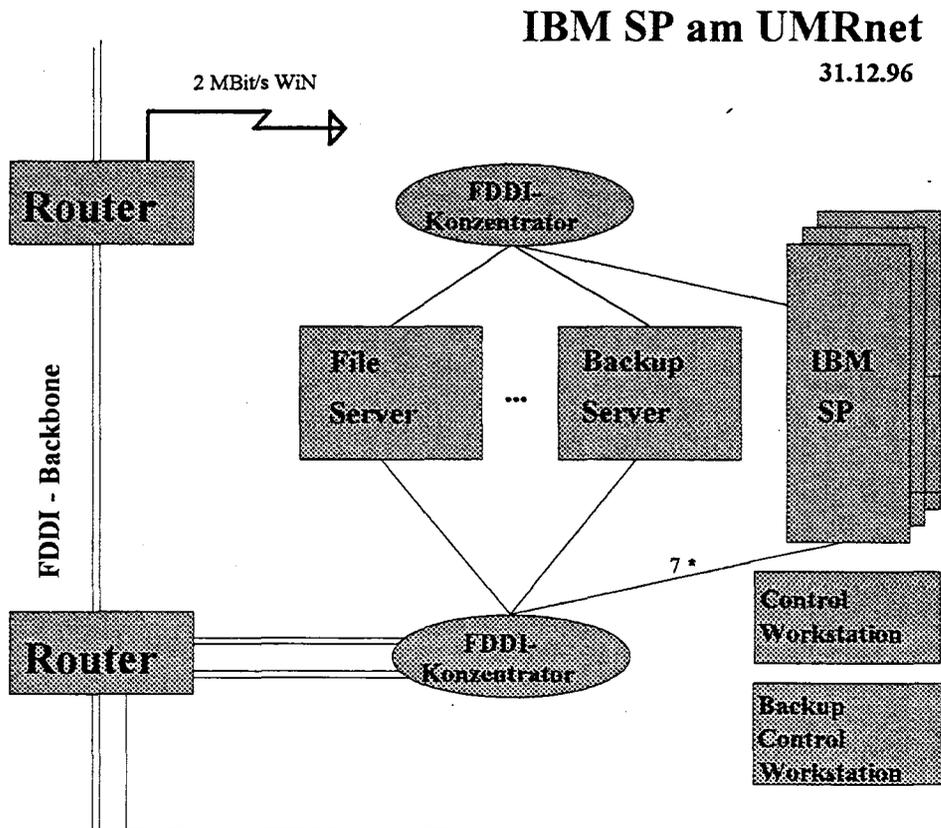
Stand: 15.07.96

3.5.2 Netzanbindung

Die Netzanbindung der IBM RS/6000 SP erfolgt durch den Netzanschluß einzelner Knoten:

- Der Wide Node und 6 Thin Nodes 2 sind in ein FDDI-LAN integriert, welches an das FDDI-Backbone des Hochschulnetzes UMRnet angeschlossen ist.
- Der Wide Node ist darüber hinaus in ein eigenständiges FDDI-LAN mit File-, Backup- und Archive-Server integriert.
- Alle Knoten sind mit den Control Workstations zu einem Ethernet-LAN verbunden (inkl. Anschluß an das FDDI-Backbone).

Das UMRnet war in 1996 mit 2 MBit/s an das WiN angeschlossen und damit in das Internet integriert. Der Zugang zum Parallelrechner erfolgte aus der Universität Marburg so wie aus allen anderen hessischen Hochschulen mit Hilfe der Internet-Dienste.



3.5.3 Softwareausstattung

Die Software für die IBM RS/6000 SP wird auf allen Knoten bereitgestellt (meistens lokal, zum Teil via NFS vom Wide Node).

Systemsoftware

AIX	4.1.4	AIX Version 4 SPO
PSSP	V2.1	Parallel System Support Programs System Administration, Monitoring und Data Repository Resource Manager und Switch Support
PE	V2.1	Parallel Environment inkl. MPL und MPI (s.u.) Parallel Operating Environment (POE) Parallel Debugger (PDBX) Parallel Profiling (prof, gprof) Visualization and Performance Monitoring Tool (VT)
RVSD	V1.1	Recoverable Virtual Shared Disk
LL	V1.2.1	LoadLeveler (inkl. NQS Interface)
PIOFS	V1.1.	Parallel I/O File System
MPE	1.0.11	Multiprocessing Environment Library

Message Passing Bibliotheken

PVMe	V2.1	Parallel Virtual Machine
MPI 1.1		Message Passing Interface

Sprachprozessoren

FORTRAN	V3.2.3	XL FORTRAN for AIX V3 & V4
HPF	V1.1.0	XL High Performance Fortran
C++	V3.1.3	C++ for AIX, inkl. C for AIX
PASCAL	V1.2.4	XL Pascal

Mathematische Bibliotheken

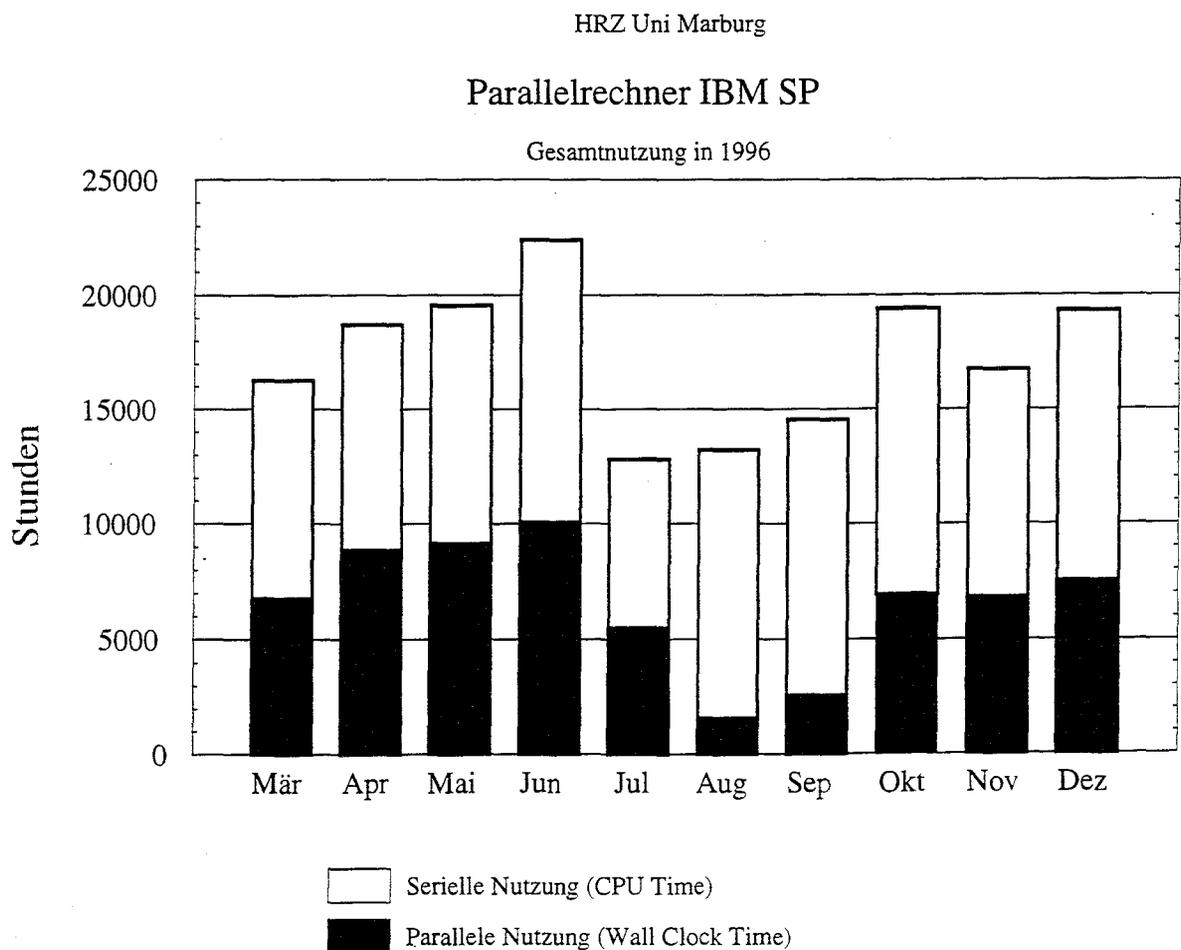
OSLp	V1.1.1	Parallel OSL Optimization Subroutine Library
PESSL	V1.1	Parallel ESSL Engineering and Scientific Subroutine Library
NAG	Mark 17	NAG Library (optimiert für POWER2-Architektur)

In den Monaten Juli bis September konnten aufgrund von diversen Softwarefehlern während insgesamt 28 400 Knotenstunden keine parallelen Jobs bearbeitet werden.

3.5.5 Nutzerstatistik

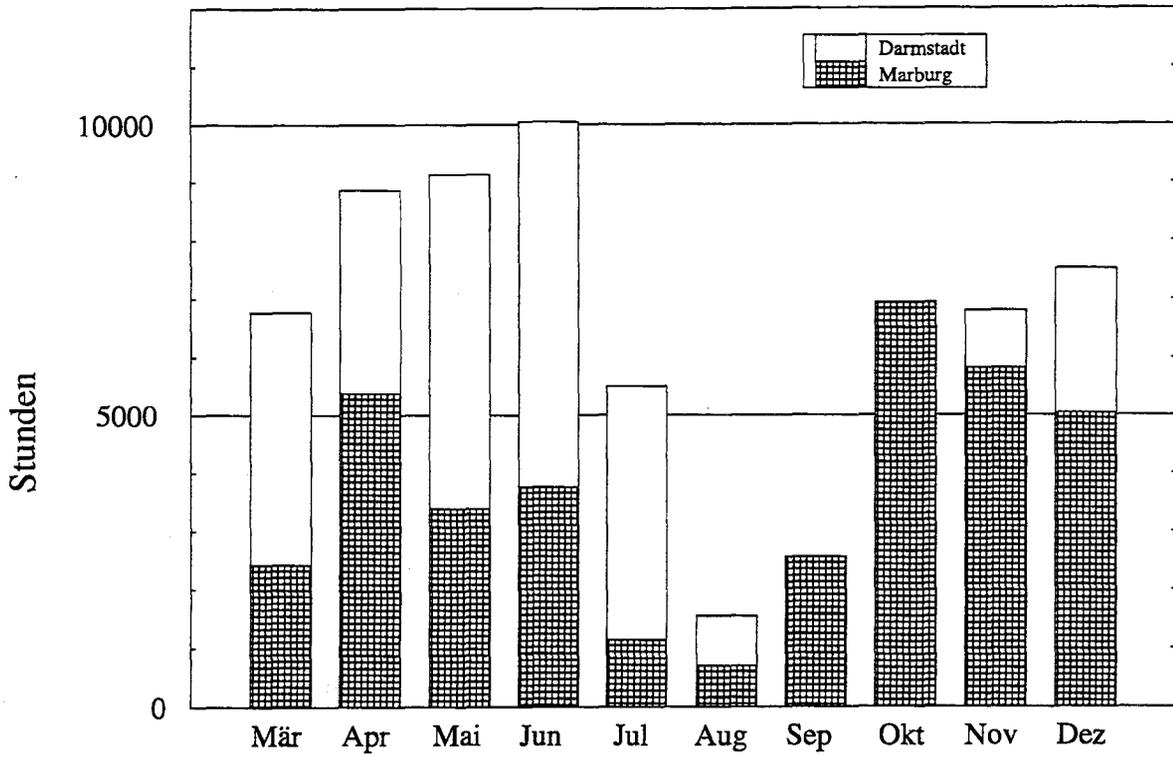
Beim parallelen Batchbetrieb werden die Knoten dediziert an die Jobs vergeben; die Nutzung wird deshalb in Knoten-Stunden gemessen (Wall Clock Time * Anzahl Knoten). Beim seriellen Batchbetrieb laufen auf den Knoten maximal 2 Jobs (Time Sharing), so daß die Nutzung in CPU-Stunden gemessen wird.

Das nachfolgende Diagramm zeigt die Gesamtnutzung der IBM RS/6000 SP, aufgeschlüsselt nach paralleler und serieller Nutzung. Die weiteren Diagramme zeigen die Nutzung durch die einzelnen Hochschul-Standorte bzw. durch die Arbeitsgruppen mit dem höchsten Bedarf an Rechenzeit, verteilt über alle Standorte.



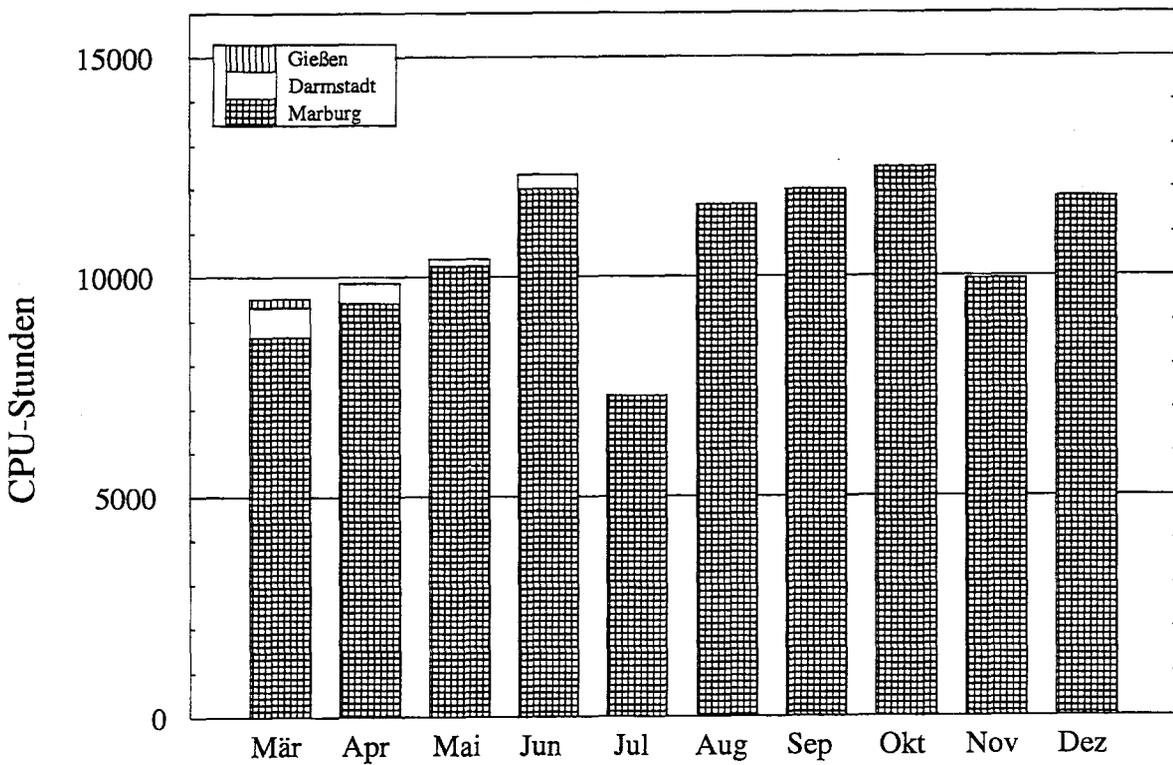
Parallelrechner IBM SP

Parallele Nutzung durch Standorte in 1996



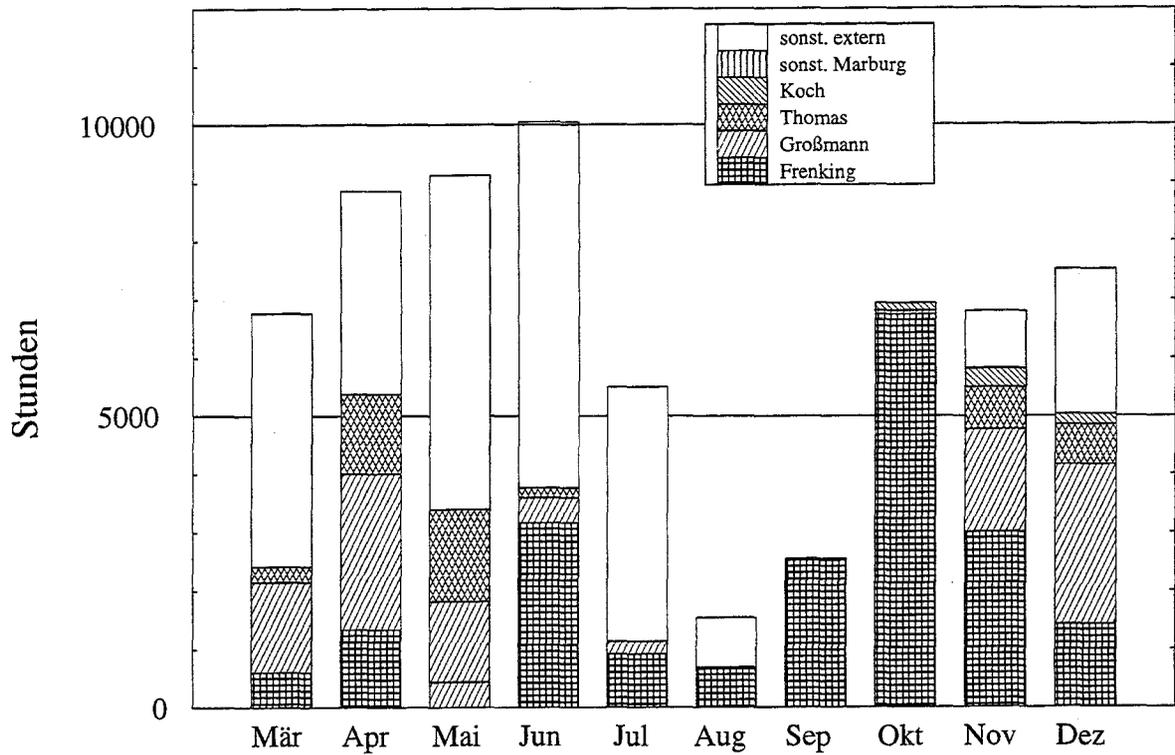
Parallelrechner IBM SP

Serielle Nutzung durch Standorte in 1996



Parallelrechner IBM SP

Parallele Nutzung durch Arbeitsgruppen in 1996



Parallelrechner IBM SP

Serielle Nutzung durch Arbeitsgruppen in 1996

