

---

# Inhaltsverzeichnis

---

## Teil I Grundlagen

---

<b>1</b>	<b>Von Megahertz zu Gigaflops</b> .....	3
1.1	Ochsen oder Hühner? .....	3
1.2	Rechenleistung im Takt .....	5
1.3	Von-Neumann-Flaschenhals .....	7
1.4	Benchmarks .....	8
1.5	Amdahls Gesetz .....	10
1.6	Granularität .....	13
1.7	Parallele Leistungsmetriken .....	16
1.8	Notizen .....	18
<b>2</b>	<b>Parallelrechner</b> .....	21
2.1	Gemeinsamer oder verteilter Speicher .....	21
2.2	Verbindungsnetzwerke .....	24
2.3	Cluster-Computer .....	27
2.3.1	Das Beowulf-Projekt .....	27
2.3.2	Standard-Hardware .....	28
2.3.3	Linux, freie Software und offene Standards .....	30
2.3.4	Anwendungsgebiete .....	31
2.4	SETI@home und Grid-Computing .....	32
2.5	Notizen .....	34
<b>3</b>	<b>Programmieransätze</b> .....	37
3.1	Datenparallelität .....	37
3.2	Threads .....	39
3.3	Nachrichtentransfer .....	44
3.4	Notizen .....	47

---

**Teil II Technik**

---

<b>4</b>	<b>Cluster-Design</b> .....	51
4.1	Grundlegende Komponenten .....	51
4.2	Anforderungen an einen High-Performance-Cluster .....	52
4.3	Netzwerktechnik .....	53
4.3.1	Netzwerkhardware .....	54
4.3.2	Netzwerktechnologien .....	57
4.4	CPU-Architektur .....	61
4.5	Arbeitsspeicher .....	62
4.6	Massenspeicher .....	63
4.6.1	Hardware .....	63
4.6.2	Clusterdateisysteme .....	64
4.7	Diskless nodes .....	65
4.8	Hardware-Monitoring .....	66
4.9	Unterbringung, Klima und Kühlung .....	67
4.10	Cluster now! .....	69
4.11	Besonderheiten von Knoten-Hardware .....	70
4.12	Schrauben oder kaufen? .....	71
4.12.1	Do it your self .....	72
4.12.2	Schlüsselfertig .....	72
4.13	Notizen .....	73
<b>5</b>	<b>PCs vernetzen</b> .....	75
5.1	TCP/IP-Grundlagen .....	77
5.2	Calculus – Der Beispiel-Cluster .....	80
5.3	Erstinstallation .....	81
5.3.1	Hardwareaufbau .....	81
5.3.2	Linux-Installation .....	82
5.3.3	Kernel-Anpassungen und Hardware-Monitoring .....	84
5.3.4	Hardware-Monitoring per Init-Skript starten .....	89
5.4	Netzwerk-Basiskonfiguration .....	90
5.4.1	Netzwerkkartentreiber laden .....	90
5.4.2	IP-Adresse zuweisen und Gateway einrichten .....	91
5.4.3	Dauerhafte Speicherung der Netzwerkkonfiguration .....	92
5.4.4	Namensauflösung .....	94
5.4.5	Tuning .....	94
5.5	SystemImager .....	97
5.5.1	Installation .....	97
5.5.2	Vollautomatische Installation .....	98
5.5.3	Updates .....	102
5.5.4	Sicherheit .....	103
5.5.5	Einschränkungen und Alternativen .....	104
5.6	Wichtige Netzdienste .....	104

5.6.1	Network File System	105
5.6.2	Network Information Service	108
5.6.3	Berkeley-r-Kommandos und die Secure Shell	112
5.6.4	Network Time Protocol	116
5.6.5	Network Address Translation	118
5.6.6	Dynamic Host Configuration Protocol	121
5.7	Channel bonding	125
5.8	Diskless nodes	129
5.8.1	Überblick	129
5.8.2	Booten über das Netzwerk mit PXE	130
5.8.3	Kernelkonfiguration	131
5.8.4	Server-Konfiguration	131
5.8.5	Alternativen und weitere Quellen	139
5.9	Notizen	140
<b>6</b>	<b>Cluster-Dienste</b>	<b>143</b>
6.1	LAM/MPI	143
6.1.1	Installation	145
6.1.2	MPI-Programme kompilieren und starten	147
6.2	Jobverwaltung und Batch-Systeme	151
6.2.1	Funktionsweise	151
6.2.2	Scheduling-Strategien	152
6.2.3	Checkpointing	153
6.2.4	Batch-Systeme im Überblick	154
6.2.5	Anwendungsbeispiel: Sun Grid Engine	157
6.3	Cluster-Management-Systeme	160
6.3.1	OSCAR	162
6.3.2	Rocks Cluster Distribution	163
6.4	Notizen	163

---

**Teil III MPI**

---

<b>7</b>	<b>Grundlagen</b>	<b>167</b>
7.1	Das Minimalgerüst	167
7.2	Senden und Empfangen von Nachrichten	170
7.3	Kollektive Kommunikation	173
7.3.1	Daten verteilen mit MPI_Bcast	173
7.3.2	Synchronisation	175
7.3.3	Kollektive Varianten	177
7.3.4	Mehr Variabilität mit v	179
7.3.5	Daten zusammenfassen mit MPI_Reduce	182
7.4	Anatomie der Nachrichtenübertragung	189
7.4.1	Blockierender Nachrichtentransfer	190
7.4.2	Laufzeitmessungen	194

7.4.3	Deadlocks	197
7.4.4	Fairness	201
7.5	Nicht blockierender Nachrichtentransfer	205
7.5.1	Rechnen statt Warten	205
7.5.2	Zeitlicher Ablauf	207
7.5.3	Flexibles Warten	209
7.5.4	Persistente Kommunikation	210
7.6	Der MPI-Standard	213
7.6.1	MPI-1 und MPI-2	213
7.6.2	FORTRAN-Schnittstelle	215
7.6.3	C++-Schnittstelle	216
7.6.4	Profiling-Interface	217
7.6.5	Dynamische Prozesserzeugung	218
7.6.6	Einseitige Kommunikation	219
7.7	Notizen	220
<b>8</b>	<b>Fortgeschrittene Techniken</b>	<b>223</b>
8.1	Kommunikator- und Gruppenmanagement	223
8.1.1	Motivation	223
8.1.2	Undurchsichtige Objekte	223
8.1.3	Kommunikatoren kopieren und teilen	225
8.1.4	Prozessgruppen	226
8.2	Fehlerbehandlung	230
8.3	Nutzerspezifische Datentypen	234
8.3.1	Motivation	234
8.3.2	Der Aufbau von MPI-Datentypen	235
8.3.3	MPI-Funktionen für nutzerspezifische Datentypen	236
8.3.4	Senden und Empfangen von nutzerspezifischen Datentypen	238
8.3.5	Beispiele	240
8.4	Parallele Ein- und Ausgabe	244
8.4.1	Motivation	244
8.4.2	Definitionen und Konzepte	245
8.4.3	Grundfunktionen	246
8.4.4	Schreib- und Leseoperationen	249
8.4.5	Fehlerbehandlung	253
8.5	Notizen	256
<b>9</b>	<b>Parallelisierungstechniken</b>	<b>259</b>
9.1	Perfekte Parallelisierung	259
9.1.1	Peinliche und weniger peinliche Probleme	259
9.1.2	Statische Lastverteilung	260
9.1.3	Dynamische Lastverteilung, Master-Worker-Schema	265
9.1.4	Eine Master-Worker-Bibliothek	267
9.1.5	Kombinatorische Probleme	275
9.2	Geometrische Parallelisierung	289

9.2.1	Probleme auf Gittern .....	289
9.2.2	Gebietszerlegung .....	290
9.2.3	Schwingende Saite .....	291
9.2.4	Kommunikatoren und Topologien .....	298
9.2.5	Zelluläre Automaten .....	302
9.3	Notizen .....	309

---

**Teil IV Praxis**

---

<b>10</b>	<b>Debuggingmethoden und Entwicklungswerkzeuge für MPI-Programme</b> .....	313
10.1	Kontrollausgaben .....	313
10.2	Debugger .....	317
10.3	Profiler .....	322
10.4	XMPI .....	324
10.5	Namen für MPI-Objekte .....	328
10.6	Notizen .....	330
<b>11</b>	<b>Bibliotheken</b> .....	333
11.1	Überblick .....	334
11.1.1	Allgemeines .....	334
11.1.2	Lineare Algebra .....	335
11.1.3	Differentialgleichungen .....	339
11.1.4	Hydrodynamik .....	341
11.1.5	$N$ -Körperprobleme und Molekulardynamik .....	342
11.1.6	Fouriertransformation .....	343
11.1.7	Optimierung .....	344
11.1.8	Pseudozufallszahlen .....	345
11.1.9	Visualisierung .....	346
11.2	APPSPACK .....	346
11.2.1	Installation .....	346
11.2.2	Das Thomson-Problem .....	346
11.2.3	APPSPACK als Bibliothek .....	350
11.3	FFTW .....	351
11.3.1	Schnelle Fouriertransformation .....	351
11.3.2	Eindimensionale Transformationen .....	352
11.3.3	Multidimensionale Transformationen .....	356
11.3.4	Rücktransformation .....	356
11.4	Tina's Random Number Generator Library .....	356
11.4.1	Pseudozufallszahlen .....	356
11.4.2	Die TRNG-Bibliothek .....	359
11.4.3	Selbstvermeidende Zufallswege .....	361
11.5	Notizen .....	366

<b>12</b>	<b>Benchmarks</b>	367
12.1	Highly-Parallel LINPACK	367
12.1.1	Algorithmus	368
12.1.2	Installation	372
12.1.3	Konfiguration	375
12.1.4	Optimierung	379
12.1.5	Ergebnisse	380
12.2	Intel MPI Benchmarks	385
12.2.1	Installation	385
12.2.2	MPI-1-Benchmarks	387
12.3	Notizen	392
<b>13</b>	<b>Checkpoint-Restart</b>	393
13.1	Überblick	393
13.1.1	Anforderungen an Checkpoint-Restart-Systeme	393
13.1.2	Implementationen von Checkpoint-Restart-Systeme	395
13.2	Ckpt	397
13.3	Berkeley Lab Checkpoint/Restart	401
13.3.1	Installation	402
13.3.2	Anwendung	403
13.3.3	Checkpointing von Multithread- und MPI-Programmen	405
13.4	Notizen	408

---

## Teil V Anhang

---

<b>14</b>	<b>Die C-Schnittstelle des MPI-Standards</b>	411
14.1	Konstanten	411
14.2	MPI-Umgebung	413
14.3	Blockierende Punkt-zu-Punkt-Kommunikation	415
14.4	Nicht blockierende Punkt-zu-Punkt-Kommunikation	417
14.5	Persistente Kommunikation	420
14.6	Abgeleitete Datentypen	420
14.7	Kollektive Kommunikation	424
14.8	Einfache Gruppen, Kontexte und Kommunikatoren	431
14.9	Kommunikatoren mit Topologie	437
<b>15</b>	<b>Argumente aus der Kommandozeile einlesen</b>	439
	<b>Literaturverzeichnis</b>	443
	<b>Sachverzeichnis</b>	449



<http://www.springer.com/978-3-540-42299-0>

Cluster Computing

Praktische Einführung in das Hochleistungsrechnen auf  
Linux-Clustern

Bauke, H.; Mertens, S.

2006, XII, 458 S., Hardcover

ISBN: 978-3-540-42299-0