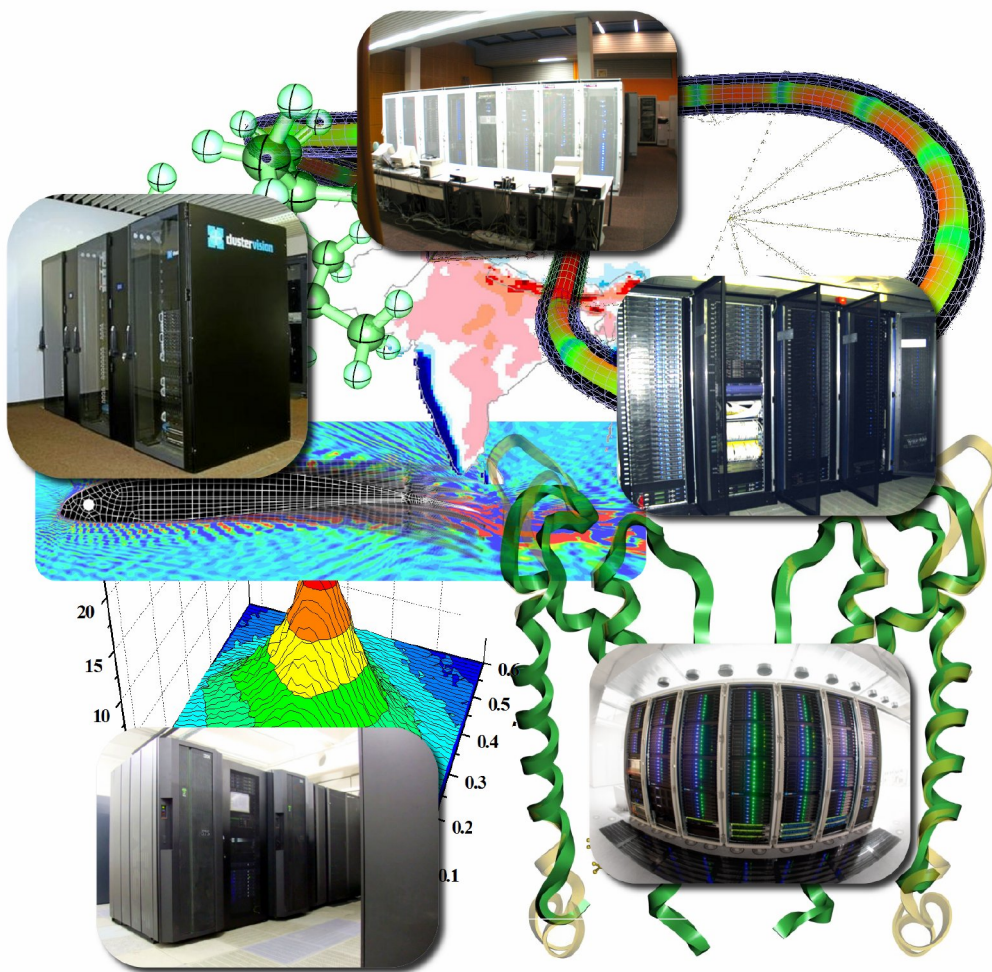


Auszug
aus dem

Research Report

High-Performance Computing in Hessen 2009



29. März 2010

Inhaltsverzeichnis

1	Vorwort	5
2	Einleitung	6
3	Technik und Organisation	7
3.1	Der HHLR am Forschungszentrum Computational Engineering (CE) in Darmstadt	7
3.1.1	Systemaufbau	7
3.1.2	Auslastung des HHLR	7
3.1.3	Parallelität der Jobs	8
3.1.4	HPC-Ausbildung an der TU Darmstadt	8
3.1.5	Verweise	9
3.2	Linux Cluster am Center for Scientific Computing (CSC) der Uni Frankfurt	10
3.3	HPC-Cluster der Justus-Liebig-Universität Gießen	12
3.3.1	Kennzahlen	12
3.3.2	Hardware	12
3.3.3	Software	13
3.4	Linux Cluster im IT-Servicezentrum der Universität Kassel	14
3.4.1	Auslastung	14
3.4.2	SuGI Grid-Cluster	15
3.5	Der Linux-Cluster MaRC an der Philipps-Universität Marburg	16
3.5.1	Systembeschreibung	16
3.5.2	Beschaffung in zwei Ausbaustufen	16
3.5.3	Hardware	16
3.5.4	Nutzung	17
3.5.5	Verweise	17
4	Projektberichte	18
4.1	Chemie	18
4.1.1	Chemoselective Boron–Carbon Bond Cleavage by Hydroboration of Borirenes	18
4.1.2	Quantum Chemical Studies on Transition Metal Compounds $[M(ZnR)_n]$ ($M=Mo$, $n=12$; $M=Ru$, $n=10$; $M=Rh$, $n=9$; $M=Ni, Pd, Pt$, $n=8$) [1]	20
4.1.3	The Dewar-Chat-Duncanson Model Reversed – Bonding Analysis of Group-10 Complexes	22
4.1.4	N-Heterocyclic Carbene Stabilized Digermanium(0)	24
4.1.5	Theoretical Study of Divalent Si(0) Compounds	26
4.1.6	Difference in solid-liquid interfacial tension between hydrophobic and hydrophilic solid surfaces by Molecular Dynamics simulations	28
4.1.7	Band Gap Tuning in Nanodiamonds: First Principle Computational Studies	30
4.1.8	Computational characterization of novel phenylhydroxycarbenes	32
4.1.9	Modellierung von Atom-Molekülreaktionen	34
4.1.10	Structural Analyses of N-Acetyl-DMAP salts	36
4.1.11	Phosphine und Phosphinoxide	37
4.1.12	Investigation of the Addition of C and Si to CH_4 and SiH_4 The Role of Quantum Mechanical Tunneling in the Reaction of $C + CH_4$ and $C + H_2O$	38

4.1.13	Photochemische Umwandlungen von Molekülen der Formel CH ₂ OS	39
4.1.14	Computational Studies on Asymmetric Hydrogenation of Ketimines with Trichlorosilane	41
4.1.15	Molecular Dynamics Simulation of the Kcv Potassium Channel	43
4.1.16	Control of the Transmembrane Potential in K ⁺ Channel Simulations	44
4.1.17	Quantum Chemistry in Solution by Combining 3D Integral Equation Theory with a Cluster Embedding Approach	46
4.1.18	Neuartige Hydroxycarbene und deren fundamentale Reaktionen	47
4.2	Physik	48
4.2.1	Phonon Sidebands in the Luminescence Spectrum	48
4.2.2	Modelling of the optical properties of semiconductor gain materials for laser applications	49
4.2.3	Photon-correlation spectroscopy in strong-coupling semiconductor quantum-dot systems	51
4.2.4	Towards understanding of C ₆₀ -based nanowire growth with anomalous anisotropy	52
4.2.5	Stability of a nanofractal on surface	54
4.2.6	Phase transitions in fullerenes	56
4.2.7	Phase transitions in proteins	58
4.2.8	Impurity effect on the melting of nickel clusters as seen via molecular dynamics simulations	60
4.2.9	Defects in epitaxial graphene layers: Density functional theory studies	62
4.2.10	<i>Ab initio</i> calculations of ZnMgO lattice parameters	63
4.2.11	Metallization of the $\sqrt{3}\times\sqrt{3}$ 4-mercaptopyridine SAM on Au(111)	64
4.2.12	Numerical simulations of the magnetic properties of nanoscale alloys	67
4.2.13	Dileptons and charm as probes of the strongly interacting quark-gluon plasma within parton-hadron-string-dynamics (PHSD) transport.	69
4.2.14	Exotische hadronische Materie und Atomkerne	70
4.2.15	Strongly interacting Fermi systems in 1/N expansion: From cold atoms to color superconductivity	72
4.2.16	Design Studies on a High Current Storage Ring	73
4.2.17	Ratio Fluctuations in Nucleus-Nucleus Collisions from Microscopic Transport Approach	74
4.2.18	Microscopic description of the phase transition from hadronic to partonic matter and dynamics of hadronization	75
4.2.19	Orbital-dependent exchange-correlation energy functionals	77
4.2.20	Ultrakalte Atome in optischen Gittern & Nanostrukturen	79
4.2.21	Pion induced coherent strangeness production	83
4.2.22	Parton-Hadron-String Dynamics	86
4.2.23	Effects from transverse momentum and mass distributions of partons in the Drell-Yan process at PANDA energies	88
4.2.24	Simulationsrechnungen zur Photoproduktion von ω -Mesonen	89
4.2.25	Superscaling in lepton-nucleus scattering	91
4.2.26	Korrelationen in exotischer Kernmaterie	93
4.2.27	Neutrino-nucleus interactions in a coupled-channel hadronic transport model	94
4.2.28	Thermodynamics of the quark-gluon plasma	96
4.2.29	In-Medium Properties of Vector Mesons	98
4.2.30	R-Matrix-Rechnungen zur Photoionisation atomarer Ionen	100
4.2.31	Simulation of the FRANZ bunchcompressor	101
4.2.32	About the origin of the mass of the nucleon in a linear sigma model	102
4.3	Ingenieurwissenschaften	104
4.3.1	Geometrische Struktur kleinskaliger Turbulenz	104

4.3.2	DFG-Graduiertenkolleg 1344: „Instationäre Systemmodellierung von Flugtriebwerken“ A4: „Vermischungsvorgänge in der Turbine“	106
4.3.3	Wall Normal Rotating Channel Flow: Direct Numerical Simulation, Modeling and Lie Group Analysis	108
4.3.4	FSI-basierte Optimierung von Profilstrukturen	110
4.3.5	Parallel Non-Linear Finite Elements for Micropolar Continua	111
4.3.6	Grobstruktursimulation der Strömung über periodische Hügel bei hohen Reynolds-Zahlen	112
4.3.7	Parallel Computation of Fluid Structure Interaction in Labyrinth Seals of Jet Engines	113
4.3.8	Grobstruktursimulation turbulenter Strömungen	115
4.3.9	Aeroacoustics of turbulent flows at low Mach number	117
4.3.10	Flow optimization using parallel solved sensitivity equations	119
4.3.11	Multigrid Methods applied to Fluid-Structure Interaction	121
4.3.12	Numerical flow control and optimization	123
4.3.13	Simulation turbulenter Fluid-Struktur-Interaktion	125
4.3.14	Evolutionary Optimization Methods for flow Shape Optimization	126
4.3.15	Modelling of solidification of binary fluids with non-linear viscosity models	128
4.3.16	Simulation and Control of Drop Size Distributions in Stirred Liquid/Liquid Systems	130
4.3.17	Elastohydrodynamische Mehrkörpersimulationstechnik	131
4.4	Mathematik & Informatik	133
4.4.1	Adaptive numerical wavelet frame methods for elliptic and parabolic operator equations and inverse problems	133
4.4.2	Parallel peer methods for initial value problems	136
4.4.3	Grid Workflow Modeling and Execution	137
4.4.4	Multimedia Computing	140
4.4.5	Turbulence Simulations on Clusters and Desktop Computers	144
4.4.6	Virtualized Cluster Computing	147
4.4.7	PASTHA – Parallelizing Stencils in Haskell	149
4.4.8	Erfassung großer Datenmengen auf Hochleistungsrechnern	151
4.4.9	On Whitehead’s asphericity conjecture	153
4.5	Biologie & Medizin	154
4.5.1	Entwicklung von Methoden und Algorithmen für genomweite Assoziationsstudien	154
4.5.2	Weiterentwicklung und Implementation von Methoden der Kopplungsanalyse für die genetische Kartierung komplexer Krankheiten	155
4.5.3	Simulationen mit SAS	156
4.5.4	Biophysical Annotation of Molecular Evolution	157
4.5.5	Molecular Structure of Gas Vesicle Forming Proteins – work in progress	159
4.5.6	Surface reconstruction using Delaunay triangulations for application in life sciences	160
4.6	Sonstige	162
4.6.1	Calculating NMR Order Parameters from MD Simulations	162
4.6.2	Ursachen der unerwarteten Stabilität von RNA-Fluorbasen	163
4.6.3	Global water resources in a changing world	165
4.6.4	WaterGAP: <i>Water</i> - Global Assessment and <i>Prognosis</i>	167
4.6.5	Regional climate simulations over South Asia	169

1 Vorwort

Die erfolgreiche Durchführung innovativer Forschungsprojekte zur Sicherung des Technologiestandortes Deutschland erfordert in nahezu allen Bereich und speziell bei biologischen, chemischen, mathematischen, physikalischen und ingenieurwissenschaftlichen Fragestellungen neben der experimentellen Untersuchung einen weiterhin steigenden Einsatz computerunterstützter Simulationen. Mittels numerischer Algorithmen können dabei einerseits tiefere wissenschaftliche Erkenntnisse in experimentell schwer zugänglichen Bereichen gewonnen und andererseits Analysen zur gezielten Ausrichtung und folglich zur Reduzierung kostenintensiver Experimente durchgeführt werden. Die dabei stets steigenden Anforderungen führen zu einer wachsenden Komplexität bei den zugrundeliegenden mathematischen Modellen, die nur durch eine Synergie von Soft- und Hardwareentwicklungen bewältigt werden können. Der vorliegende Forschungsbericht liefert in diesem Kontext einen eindrucksvollen Nachweis der existenziellen Relevanz hessischer Hochleistungsrechner bei der wissenschaftlichen Forschungs- und Entwicklungsarbeit an allen beteiligten hessischen Standorten. Erst der kontinuierliche Ausbau der zur Verfügung stehenden Ressourcen ermöglicht dabei eine anhaltende Wettbewerbsfähigkeit der hessischen Hochschulen im nationalen wie auch internationalen Vergleich.

Die präsentierten Forschungsprojekte zeigen an dieser Stelle die ausgezeichnete Qualität der hierzulande realisierten computerunterstützten Forschung. Ihre internationale Anerkennung wird hierbei sowohl durch die zahlreichen Publikationen in renommierten Zeitschriften als auch durch Beiträge im Rahmen angesehener Fachtagungen belegt. Dabei zeigt sich in beeindruckender Weise die Nutzung der hessischen Hochleistungsrechner bei der notwendigen Grundlagenforschung einerseits und der Umsetzung der hierbei gewonnenen Erkenntnisse innerhalb industrierelevanter Anwendungsszenarien andererseits.

Prof. Dr. Andreas Meister
(Vorsitzender des Hessischen Beirats für Hochleistungsrechnen)

2 Einleitung

Die Bedeutung des Wissenschaftlichen Rechnens als eine Schlüsseltechnologie für nahezu alle Bereiche von Wissenschaft und Technik ist heutzutage unumstritten. Trotz der enormen Fortschritte im Bereich der Rechnertechnologie ist hierbei in vielen Anwendungsfällen der Einsatz von Hochleistungsrechnern unerlässlich, nicht zuletzt auch aufgrund der stetig wachsenden Komplexität der Problemstellungen. Das Hochleistungsrechnen trägt entscheidend zur Reduktion von Entwicklungszeiten neuer Methoden und Technologien bei. So können Probleme angegangen werden, die wegen der hohen Anforderungen an die Rechenleistung anderweitig einer numerischen Simulation nicht zugänglich sind.

Der Verfügbarkeit einer adäquaten Hochleistungsrechnerkapazität kommt damit für den Wissenschaftsstandort Hessen eine große Bedeutung zu, da diese die Grundlage für eine national und international konkurrenzfähige Forschung im Bereich des Wissenschaftlichen Rechnens darstellt. Der Erfolg von Sonderforschungsbereichen, Forschergruppen, Graduiertenkollegs und einer Vielzahl unterschiedlicher Einzelprojekte hängt entscheidend davon ab.

Um dem Rechnung zu tragen, wurden nicht nur die beiden großen Hessischen Rechner in Darmstadt und Frankfurt weiter ausgebaut, sondern auch weitere meist schwächer vernetzte Cluster an den Universitäten Gießen, Kassel und Marburg auf- oder ausgebaut.

Durch die Diversifizierung der Leistung soll den stark unterschiedlichen Anforderungen der Nutzer besser Rechnung getragen werden. Problemstellungen mit moderatem Ressourcenbedarf können auf den kleineren System kostengünstig und unbürokratisch abgearbeitet werden, während die Ressourcen der stärker vernetzten großen System für Anwendungen mit entsprechenden Anforderungen geschont werden.

Dem Darmstädter Forschungszentrum CE und dem Frankfurter CSC, die beide assoziierte Mitglieder der Gauss-Allianz sind, kommt die Funktion von Kompetenzzentren im Bereich des Hochleistungsrechners zu. Die notwendige fachübergreifende Kompetenz ist durch die interdisziplinäre Zusammensetzung der beiden Zentren, sowie durch entsprechende Aktivitäten der Mitglieder in Forschung und Lehre gewährleistet. Damit leisten beide Zentren einen Beitrag zu den folgenden Aufgaben:

- Entwicklung bzw. Weiterentwicklung von Anwendungssoftware für Hochleistungsrechner in den verschiedenen Anwendungsbereichen
- Unterstützung von Nutzern bei der Portierung von Anwendungssoftware
- Ausbildung von wissenschaftlichem Nachwuchs im Bereich des Wissenschaftlichen Rechnens
- Technologietransfer in die Industrie im Rahmen von Kooperationsprojekten
- Organisation von regelmäßigen Benutzer-Kolloquien, die den Erfahrungsaustausch zwischen allen Nutzern der jeweiligen Rechner ermöglichen
- Kontaktpflege und Zusammenarbeit mit anderen im Bereich des Hochleistungsrechnens tätigen Arbeitsgruppen im In- und Ausland (z.B. Workshops, Forschungsprojekte)

Die Rechner sind für Nutzer aller hessischen Universitäten zugänglich. Die einzelnen Rechenzeitkontingente richten sich nach den finanziellen Beteiligungen der Hochschulen, Fachbereiche bzw. Fachgebiete.

3 Technik und Organisation

3.1 Der HHLR am Forschungszentrum Computational Engineering (CE) in Darmstadt

3.1.1 Systemaufbau

Seit November 2008 besteht der HHLR aus einem Cluster von 14 Rechenknoten mit jeweils 32 Power6-Prozessoren. Die Prozessoren haben eine Taktfrequenz 4,7 GHz und können vier Floatingpoint Operationen (FLOP) pro Takt verarbeiten. Das gesamte System kam somit auf eine Peakperformance von 8,5 TFLOP/s. Die Rechenknoten sind mit einem Arbeitsspeicher von 128 GB (einer sogar mit 256 GB) ausgestattet. Auch wenn bei weitem nicht alle Problemstellungen soviel Speicherplatz benötigen, ist es wichtig, dass die zentralen Rechencluster diese Ressourcen vorhalten, da Problemstellungen,



gen, die diese Anforderungen haben, auf konventioneller Hardware nicht lauffähig sind. Durch die 32 in einem System integrierten Prozessoren können auch sehr kommunikationslastige Programme bis zu 32 CPUs nutzen. Alle Knoten des Clusters sind latenzarm mit acht DDR-Infiniband-Lanes vernetzt. Aufgabenstellungen mit moderateren Anforderungen an die Kommunikationsleistung können so auch auf mehrere Knoten verteilt werden.

Das neue System ist von der Systemarchitektur her identisch mit dem Entwicklungscluster (JUMP) des Bundeshöchstleistungsrechners in Jülich. Auch am Höchstleistungsrechner in Garching (RZG MPI/IPP) und am Deutschen Klimarechenzentrum (DKRZ) kommt die gleiche Technik zum Einsatz. Der HHLR ist damit bestens geeignet, als Brücke zwischen den PC-Clustern der Institute und den deutschen Höchstleistungsrechnern zu fungieren.

Im Dezember 2009 wurde das System noch einmal um vier Knoten auf jetzt 18 Maschinen erweitert. Damit besteht der HHLR heute aus 576 eng vernetzten Power6-CPU's und erreicht eine Peakperformance von fast 11 TFLOP/s.

3.1.2 Auslastung des HHLR

Beim Umbau des HHLR Ende 2008 wurden die bisherigen Power5-Rechner durch deutlich leistungsfähigere Power6-Systeme ersetzt. Obwohl sich dabei die Anzahl der CPU's (und damit die verfügbaren CPU-Stunden) etwas verringert haben, steht wegen der gestiegen Leistungsfähigkeit insgesamt mehr Rechenkapazität zur Verfügung. Der in Abbildung 1.2.1 erkennbare Abfall der abgegebenen Rechenzeit zum Jahreswechsel 2008/2009 ist darauf zurückzuführen.

Das Verhältnis von der abgegebenen Rechenzeit zur maximal verfügbaren Rechenzeit des HHLR lag im Jahresmittel bei 90%. Das System ist damit vollständig ausgelastet.

Die Zuweisung der Rechenzeit an die Wissenschaftler der einzelnen Standorte richtet sich nach der Beteiligung der jeweiligen Hochschule an den Investitionskosten. Die TU Darmstadt hat ein Kontingent von 67,5%, der Uni Gießen und Marburg stehen 11,3% bzw. 10,8% zu und die Universitäten in Frankfurt und Kassel haben einen Anspruch auf je 5,2% der Rechenzeit. Wenn ein Standort sein Kontingent nicht ausschöpft, können andere Standorte diese Rechenzeit anteilig nutzen.

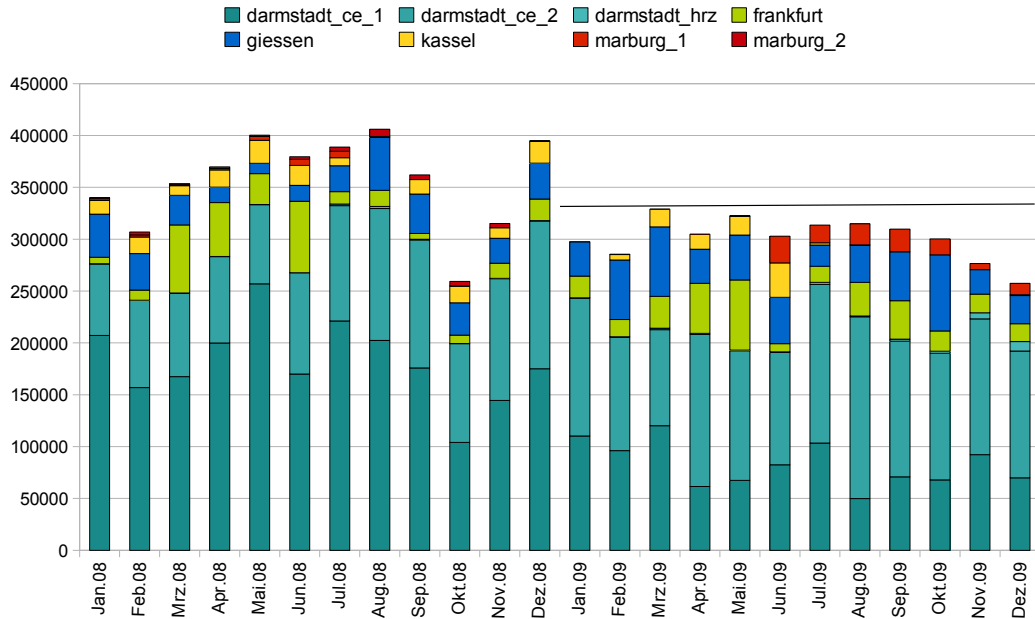


Abbildung 1.2.1: Verteilung der monatlichen Rechenzeit auf die einzelnen Universitäten

Die Nutzung des Rechners durch die Universität Marburg ist 2009 wieder gestiegen, sie liegt aber nach wie vor merklich unter ihrem Kontingent. Im Jahresdurchschnitt haben marburger Wissenschaftler knapp 3,3% der insgesamt abgegebenen Rechenzeit genutzt. Auch die Uni Kassel hat mit 2,5% nur die Hälfte ihres Kontingents abgerufen.

Von dieser Zurückhaltung profitieren die anderen Standorte. Die TU Darmstadt nutzte 71% des Rechners, die Uni Frankfurt 9,2% und die Uni Giessen annähern 14%.

3.1.3 Parallelität der Jobs

Betrachtet man die Entwicklung der Jobgrößen (Parallelität) (s. Abbildung 1.3.1) so fällt auf, dass sich im Laufe des Jahres 2009 das Maximum der abgegebenen Rechenzeit von 8-fach parallelen zu 32-fach parallelen Jobs verschoben hat. Zu erklären ist diese Verschiebung durch den Umbau des Rechners Ende 2008. Bei diesem Umbau wurde die Knotengröße von in der Regel 8 CPUs auf 32 CPUs erhöht.

Trotz einer vollständigen Infiniband-Vernetzung werden nach wie vor die meisten Jobs innerhalb eines Knotens gerechnet. Das legt die Vermutung nahe, dass zukünftig noch größere SMP-Systeme oder sehr viel schnellere interne Vernetzungen gebraucht werden, um diese Anwendungen effektiv nutzen zu können.

3.1.4 HPC-Ausbildung an der TU Darmstadt

Mit dem Bachelor- und Master-Studiengang Computational Engineering sowie der Graduiertenschule Computational Engineering wurde an der TU Darmstadt ein durchgängiges konsekutives Konzept zur interdisziplinären Ausbildung im Bereich des High Performance Computing (HPC) etabliert. Der Bachelor-/Master-Studiengang wird vom fachbereichsübergreifenden Studienbereich Computational Engineering organisiert und betreut, der von den Fach- und Studienbereichen Mathematik, Mechanik, Bauingenieurwesen und Geodäsie, Maschinenbau, Elektrotechnik und Informationstechnik, Informatik und dem Forschungszentrum Computational Engineering getragen wird. Den Studierenden werden mathematische, informationswissenschaftliche und ingenieurwissenschaftliche Kenntnisse, mit denen sie physikalische und technische Systeme modellieren und simulieren

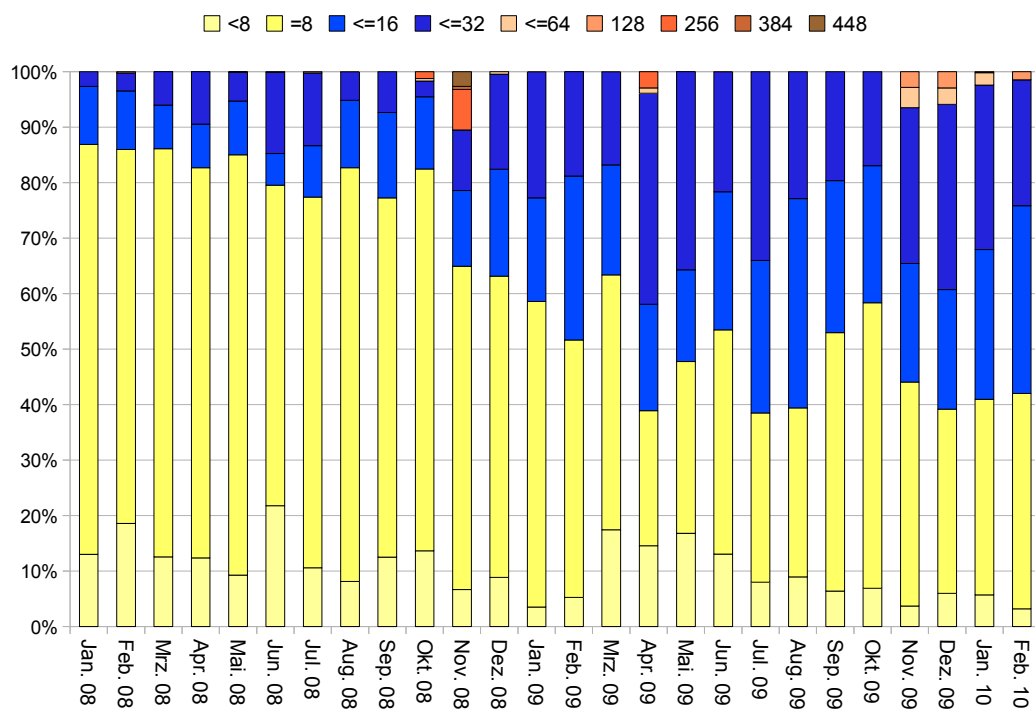


Abbildung 1.3.1: Verteilung der monatlich abgegebenen Rechenzeit auf die Parallelität der Jobs. Mit dem Umbau des HHLRs auf 32-fach SMPs Ende 2008 hat sich das Maximum von 8-fach parallelen Jos zu 32-fach parallelen Jobs verschoben.

können, vermittelt. Methoden des HPC sind hierbei ein wesentlicher Bestandteil. Das Ausbildungskonzept wird durch die Graduiertenschule Computational Engineering komplettiert, in welche die Studierenden bereits nach dem ersten Masterjahr eintreten können und diese nach weiteren vier Jahren mit einer Promotion abschließen können. HPC ist ein zentrales Querschnittsthema der Forschungsarbeiten an der Graduiertenschule.

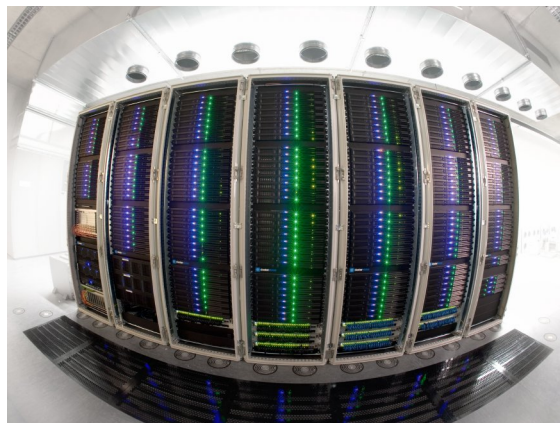
3.1.5 Verweise

1. Weitere Informationen zum HHLR:
<http://www.hhlr.tu-darmstadt.de>
2. Weitere Informationen zum Forschungszentrum CE:
<http://www.rc.ce.tu-darmstadt.de/>

3.2 Linux Cluster am Center for Scientific Computing (CSC) der Uni Frankfurt

Das Center for Scientific Computing (CSC) der Goethe-Universität betreibt einen der beiden hessischen Hochleistungsrechner. Das CSC wurde im Jahr 2002 als gemeinsame Initiative von 15 Forschungsgruppen der naturwissenschaftlichen Fachbereiche der Universität Frankfurt ins Leben gerufen. Ausgangspunkt für seine Gründung war die Erkenntnis, dass in den modernen Naturwissenschaften die Bewältigung komplexer numerischer Aufgaben eine zentrale Rolle einnimmt.

Das CSC bietet daher neben der Bereitstellung einer umfangreichen Rechenkapazität für die Projekte der beteiligten Forschergruppen ein fachübergreifendes Forum auf dem Gebiet Computational Science in Form von interdisziplinären Vorträgen. Darüber hinaus bietet das CSC einen zweijährigen englischsprachigen Masterstudiengang in Computational Science an. Der Studiengang ist in modularer Form konzipiert, so dass alle Studienleistungen in Form von international anerkannten Credit Points gemäß dem European Credit Transfer System bewertet werden. Die durchweg in englischer Sprache gehaltenen Kurse eröffnen auch ausländischen Studierenden eine interdisziplinäre Ausbildung in allen Bereichen der Computer Simulation.



Die Forschungsaktivitäten der am CSC arbeitenden Wissenschaftlern umfassen eine große Zahl von Projekten mit hohem numerischen Aufwand. Die Projekte, die auf dem CSC-Cluster bearbeitet werden, umfassen eine weite Spanne aktueller Themenbereiche. Sie erstrecken sich über so unterschiedliche Gebiete wie die Untersuchung der Strukturen von Proteinen in der Biochemie und die Eigenschaften von Kristallen unter höchstem Druck in der Geophysik. Die Dynamik der Grundbausteine der Materie, der Quarks, und Gluonen, im Urknall und bei Hochenergieexperimenten werden in der theoretischen Physik auf dem CSC simuliert. Wissenschaftler des Frankfurt Institute for Advanced Studies modellieren komplexe atomare Strukturen und untersuchen die Möglichkeiten der Krebstherapie mit Schwerionenstrahlen.

Der Hochleistungsrechner des CSC besteht derzeit aus 2 Linux MPP-Clustern (CSCIII und Erweiterung).

Das Clustersystem CSCIII wurde in 2006 zur Benutzung freigegeben und Anfang 2007 ausgebaut. Es besteht aus 251 Knoten, die jeweils mit 2 DualCore 2GHz Opteron CPUs ausgestattet sind. Jeder Knoten verfügt über 8 GByte Hauptspeicher, zusätzlich sind 44 Knoten auf 16 GByte aufgerüstet worden. 64 der Knoten (und damit 256 Prozessorkerne) sind über Myrinet schnell vernetzt. Ein Raid-System stellt 25 TByte Speicherplatz zur Verfügung.

Der Ende 2009 in Betrieb genommene Ausbau des Clusters umfasst 194 Knoten mit jeweils 2 HexaCore 2,2 Ghz Opteron-Prozessoren, die mit DDR-ConnectX-Infiniband vernetzt sind. Jeder Knoten ist mit 32 GByte Hauptspeicher ausgestattet, 40 Knoten zusätzlich mit 64 GByte. Ein paralleles Dateisystem mit einer Lese- und Schreibrate von mehr als 4 GByte/s steht hier den Benutzern als schnelles globales Scratch-Dateisystem zur Verfügung. Es besteht aus 10 Servern und hat eine Speicherkapazität von 56 TByte. Untergebracht sind die Rechenknoten und Server dieses Clusters in 3 wassergekühlten und 2 luftgekühlten Racks. Die Peak-Rechenleistung des Ausbaus beträgt 20 TFlop/s.

Die Rechenknoten und Server des CSCIII-Systems werden in Kürze in die Erweiterung integriert, so dass aus Benutzersicht die Cluster als einheitliche Rechenumgebung bestehen.

Durch den stetigen Ausbau verfügt das CSC damit über ein MPP-System aus insgesamt 3332 Prozessorkernen mit einer theoretischen Rechenleistung von 24 TFlop/s.

In Kürze steht schon der nächste Ausbau des Clusters an: Zusätzlich werden noch weitere 104 Rechenknoten mit 2 HexaCore 2,2 Ghz Opteron-Prozessoren mit 32 GByte Speicherausbau installiert, ebenso 8 Knoten mit 4 HexaCore 2,2 Ghz Opteron-Prozessoren und 128 GByte Hauptspeicher. Durch weitere 10 Storage-Server wird die Kapazität des parallelen Dateisystems auf 136 Tbyte vergrößert. Die Rechenleistung des Systems wird damit auf 37 TFlop/s erhöht.

Seit 2009 steht allen Mitgliedern des CSC im Rahmen ihrer jeweiligen Rechenkontingente ein kleiner GPGPU-Cluster (Scout) als experimentelles System zur Verfügung. Das System besteht aus 9 Recheneinheiten, wobei jede Einheit aus zwei CPU- und drei GPU- Knoten besteht. Die CPU-Systeme bestehen aus zwei QuadCore Xeon CPUs mit 16 Gbyte Hauptspeicher. Die GPU-Knoten sind Tesla S1070 Systeme von Nvidia. Jeder GPU-Knoten leistet 4 Tflop/s single precision (sp), bzw 345GFlop/s double precision (dp), so dass das Gesamtsystem eine Spitzenleistung von 108 TFlop/s sp, bzw 9.3 Tflop/s dp erreicht. Die Kommunikation innerhalb einer Recheneinheit erfolgt über PCI Express, die Kommunikation zwischen den Recheneinheiten wird ab 2010 durch ein Infiniband-Netzwerk erweitert. Ziel des Scout ist es, allen interessierten Arbeitsgruppen Erfahrungen im Einsatz von GPGPU-Systemen zu ermöglichen.

Den Nutzern wird eine Vielzahl von Softwarepaketen und Compilern zur Verfügung gestellt. Die Datensicherung aller Systeme erfolgt über das Backup-System des Hochschulrechen-zentrums der Universität. Mehr als 500 Wissenschaftler aus über 90 Arbeitsgruppen aus dem Bereich der Naturwissenschaften, Mathematik und Informatik sorgen für eine vollständige Auslastung des Rechners über das ganze Jahr. Etwa ein Viertel der verfügbaren Rechenzeit wird von Gruppen aus Darmstadt, Marburg, Gießen und Kassel verbraucht.

Die Finanzierung der Anlage erfolgte überwiegend durch Beiträge Frankfurter Hochschullehrer und -Institutionen, sowie Zuschüsse des Landes Hessen und des Bundes (im Rahmen des HBFG und Forschungsgroßgeräte-Programms). Über den Verbund der Hessischen Hochleistungsrechner (HHLR) haben Wissenschaftler aller hessischen Hochschulen Zugang zu den Computersystemen des CSC.

3.3 HPC-Cluster der Justus-Liebig-Universität Gießen

Nachdem es an der Universität Gießen mehrere Jahre lang keine eigenen Möglichkeiten für Wissenschaftliches Rechnen gab, hat sich die Situation in 2009 grundlegend verändert. Nach der Bereitstellung von Geldmitteln und einer europaweiten Ausschreibung sowie Zuschlagserteilung in 2008 wurde 2009 ein neues HPC-Cluster installiert und am 25.05.2009 in Betrieb genommen. Die Lieferung der Hardware sowie die Installation und Basiskonfiguration erfolgten durch die Firma Clustervision.

3.3.1 Kennzahlen

- 3 Serracks
- 2 Kühleinheiten
- 40 Rechenknoten mit je zwei Quad-Core CPUs
- 2 GB Hauptspeicher/Core
- 320 Cores
- 2,94 TFLOPs
- 8 TB Plattenplatz für Home-Verzeichnisse
- 18 TB temporärer Lustre-Speicherplatz
- Gigabit-Ethernet und Infiniband (4x DDR)
- Scientific Linux, C und FORTRAN Compiler
- Queuing-System SUN Grid Engine



3.3.2 Hardware

Die einzelnen Komponenten des Clusters sind in 3 Serracks im Rechnerraum des HRZ untergebracht. Zwischen diesen Racks sind zur Kühlung des Clusters Luft-/Wasser-Wärmetauschereinheiten (Fabrikat Rittal, LCP plus) angeordnet. Diese werden über einen ebenfalls neu installierten Wasser-/Wasser-Wärmetauscher im Keller des HRZ-Gebäudes versorgt, der wiederum an die Kaltwasserversorgung der Stadtwerke angeschlossen ist. Zur Stromversorgung wurde die bestehende USV-Anlage durch eine zusätzliche zentrale USV-Anlage ergänzt.

Das Cluster besteht aus den Funktionseinheiten: Frontend-Rechner, Rechenknoten, Kommunikationssystem und einem Lustre-Dateisystem.

Die Frontend-Rechner stellen den Login- und Filetransfer-Service sowie das Queuing-System zur Verfügung. Sie sind als HA-Verbund (aktiv/passiv) ausgelegt. Die Hardware besteht aus zwei Supermicro 2HE-Chassis mit H8DMU+-B Motherboard und redundantem Netzteil. In jedem Frontend-Rechner kommen zwei AMD Opteron 2376 Quad-Core-Prozessoren mit 2,3 GHz zum Einsatz. Die sonstigen Daten der Frontend-Rechner sind: 16 GB Hauptspeicher und zwei interne 250 GB Platten (gespiegelt) für das Betriebssystem. Zu der Funktionseinheit Frontend-Rechner gehört weiterhin ein externes RAID-System (DELL Power Vault MD3000 mit 15 1TB 7,2k SATA-Platten). Es wird mit RAID6 plus Hotspare betrieben und stellt ca 12 TB netto als Plattenspeicherplatz zur Verfügung. Dieser wird für die Home-Verzeichnisse der Benutzer (ca. 8 TB) und einen zentralen Ort zur temporären Zwischenspeicherung von Daten (ca. 4 TB) genutzt.

Die Rechenknoten sind 1HE-Twin-Chassis von Supermicro mit je zwei H8DMT-INF+ Motherboards. Je Rechenknoten sind zwei AMD Opteron 2376 Quad-Core-Prozessoren mit einer Taktrate

von 2,3 GHz, 16 GB Hauptspeicher (2 GB pro Core) und zwei 500 GB Platten (gespiegelt) installiert. Die internen Platten dienen als lokaler Scratch-Speicher insbesondere für parallele Prozesse, die lokal auf einem Rechenknoten ablaufen können. Bei den Rechenknoten wurde auf ein redundantes Netzteil verzichtet. Insgesamt sind 20 Twin-Chassis entsprechend 40 Rechenknoten mit je 2 Quad-Core-Prozessoren installiert. Das ergibt zusammen 320 Cores mit einer theoretischen Rechenleistung von 2,94 TFLOPs.

Das Kommunikationssystem besteht aus einem Management-, einem Ethernet- und einem Infiniband-Datennetz. Die Verwaltung des Clusters erfolgt über den in jeder Rechnerkomponente installierten IPMI Port. Diese sind über ein 100 MBit-Ethernet Netz miteinander verbunden. Für die Kommunikation der Knoten untereinander und mit den Frontend-Rechnern wird ein 1 GBit-Ethernet Netz bestehend aus drei in einem Stack betriebenen DELL Power Connect 6248 Switches eingesetzt. Ein DDR-Infiniband Netzwerk mit einem bis auf 144 Ports ausbaubaren Flextronics (Mellanox) Switch dient zur schnellen Interprozesskommunikation und für den schnellen Datentransfer mit dem Lustre-Dateisystem.

Das Lustre-Dateisystem besteht aus zwei in Failover-Konfiguration betriebenen Lustre MDS Servern und zwei OSS Fileservern. Die Hardware der MDS Server entspricht dem der Frontend-Rechner, der Hauptspeicher ist allerdings wegen der Funktion als Verzeichnissever mit 32 GB doppelt so groß ausgelegt. Die Verzeichnisdaten liegen auf einem externen RAID-System (DELL Power Vault MD3000 mit 15 146 GB 15k SAS Platten), das mit RAID10 plus Hotspare betrieben wird. Für den Lustre-Speicherplatz sorgen zwei OSS Server jeweils bestehend aus einem 4 HE Supermicro Chassis mit H8DME-2 Motherboard mit je zwei AMD Operon 2376 Quad-Core-Prozessoren mit einer Taktrate von 2,3 GHz, 8 GB Hauptspeicher und zwei gespiegelten internen 250 GB Platten. Der eigentliche Speicher ist pro OSS Server aus 12 internen 1 TB 7,2k SATA-Platten realisiert, die in RAID6 plus Hotspare Konfiguration betrieben werden. So stehen insgesamt ca. 18 TB netto an Lustre-Speicherplatz zur Verfügung.

In 2009 wurde der weitere Ausbau des Clusters in Auftrag gegeben. Dieser beinhaltet 40 weitere Rechenknoten in 20 Supermicro 1HE-Twin-Chassis mit je zwei D8DMT-IBXF Motherboards. Die Rechenknoten entsprechen den schon installierten Rechenknoten, sind allerdings mit je zwei AMD Opteron 2431 Hexa-Core-Prozessoren mit einer Taktrate von 2,4 GHz und 32 GB Hauptspeicher (2,66 GB pro Core) bestückt. Die Kommunikationsnetze und insbesondere der Infiniband Switch werden entsprechend ausgebaut werden. Zur Unterbringung der neuen Hardware werden ein weiteres Serrerrack und zwei weitere Kühleinheiten installiert werden. Durch die neuen Rechenknoten erhöht sich die Anzahl der Cores auf 800 und die theoretische Rechenleistung auf 7,5 TFLOPs.

3.3.3 Software

Das Cluster wird mit der von Clustervision entwickelten Verwaltungssoftware CusterVisionOS betrieben. Auf den einzelnen Knoten läuft Scientific Linux als Betriebssystem. Zur Programmentwicklung können die Programmiersprachen C und FORTRAN eingesetzt werden. Dazu stehen die Compiler von GNU, Intel und PGI zur Verfügung. Für die Unterstützung zur Optimierung paralleler Prozesse wurde das Tool Alinea Opt installiert. Als Queuing-System wird SUN Grid Engine (SGE) eingesetzt.

3.4 Linux Cluster im IT-Servicezentrum der Universität Kassel

Das Linux Cluster besteht zur Zeit aus 150 Maschinen mit insgesamt 1080 Prozessoren. Jede Maschine hat mindestens 2 AMD-Opteron Prozessoren und 8 GB Hauptspeicher. 62 dieser Systeme sind mit einer Infiniband-Vernetzung ausgestattet. Das Cluster läuft unter dem Betriebssystem CENTOS 5.3 mit PBS-Torque Resource Manager und Maui Cluster Scheduler. 10 Maschinen stehen für interaktive Nutzung zur Verfügung. Der Zugriff auf Plattenspeicher erfolgt über GPFS-Dateisysteme. Von den 1080 Prozessoren wurden 336 Prozessoren aus Mitteln eines Fachgebietes im Fachbereich Naturwissenschaften beschafft und stehen deshalb nur den Mitarbeitern dieses Fachgebietes zur Verfügung. 40 Doppelprozessorsysteme mit je zwei AMD-Sechskernprozessoren und 64 bzw. 128 GB Hauptspeicher wurden gerade erst in Betrieb genommen. Die Geräte wurden nach einen gemeinsamen Großgeräteantrag gemäß Artikel 91b GG von sechs Forschungsgruppen an der Universität beschafft.



Anwendungssoftware: Abaqus-6.5-6 (x86-64), ACML 4.1.0, jrMan, Gaussian 03, GROMACS 3.3.3-1, Matlab R2008b, Mathematica 7.0, MD Nastran 2.1 (x86-32), MD Patran 2.1 (x86-32), Meep-0.20.3, MPI: mpich-1.2.7, MPI: LAM/MPI version 7.1.2, MPI: mpich2-1.0.5p4 , MPI: mvapich2-1.0.1, mpb-1.4.2, NWChem 5.1, OpenFOAM-1.5, Pixie-2.2.4, R-2.8.1 (Rmpi mit lme4)

3.4.1 Auslastung

Im zweiten Halbjahr 2009 gab es ca. 60 aktive Nutzer. In diesem Zeitraum wurden 1742961 CPU-Stunden zur Verfügung gestellt, von denen 821245 Stunden abgerufen wurden. Die größten Nutzergruppen kommen aus dem Fachbereichen Physik, Maschinenbau und Elektrotechnik.

3.4.2 SuGI Grid-Cluster

Im Rahmen eines Aufstockungsantrages für das SuGI-Projekt wurde im Dezember 2007 ein zweites Linux Cluster beschafft, das primär für Grid-Computing Anwendungen reserviert ist. Bei diesem Cluster handelt es sich um 21 Doppelprozessorsysteme mit Intel Quad-Core CPU's, 32 GB Hauptspeicher und 10 TB Plattenspeicher. Das Cluster ist seit Januar 2008 im Nutzerbetrieb und soll insbesondere für Grid-Anwendungen in der Chemie genutzt werden. Dazu wurde es als Ressource in mehreren Virtuellen Organisationen des D-Grid angemeldet. Das Cluster läuft unter dem Betriebssystem CENTOS 5 und ist damit kompatibel zum Linuxcluster des HRZ. Für Grid-Anwendungen ist die Middleware Gobus Toolkit auf beiden Linux Clustern installiert. Zudem gibt es für das Grid-Cluster ein UNICORE 5 und ein UNICORE 6 Frontend. Das Cluster ist mit hochschulinternen Nicht-Grid-Anwendungen ebenfalls gut ausgelastet.

3.5 Der Linux-Cluster MaRC an der Philipps-Universität Marburg

3.5.1 Systembeschreibung

Der Linux-Cluster MaRC (**Marburger RechenCluster**) besteht aus insgesamt 144 Knoten mit jeweils zwei DualCore AMD Opteron-Prozessoren. Arbeitsgruppen aus verschiedenen Fachbereichen waren maßgeblich an der Finanzierung beteiligt und können MaRC entsprechend nutzen, ebenso Anwender aus anderen hessischen Hochschulen. MaRC wird vom Hochschulrechenzentrum (HRZ) der Philipps-Universität Marburg betrieben.



Abbildung 5.1.1: MaRC, Ansicht von vorne

3.5.2 Beschaffung in zwei Ausbaustufen

Die erste Ausbaustufe von MaRC wurde nach längeren Planungen noch im Rahmen eines HBF-G-Antrags im Dezember 2005 beschafft und Anfang 2006 in Betrieb genommen.

Eine zweite Ausbaustufe konnte im März 2007 über eine Mischfinanzierung realisiert werden. Neben Restmitteln aus dem HBF-G-Antrag wurden dabei weitere Bundesmittel eingesetzt; ein Teil (19) der Knoten sind über das D-Grid-Integrations-Projekt finanziert und daher organisatorisch in das D-Grid integriert.

3.5.3 Hardware

Die insgesamt 144 Knoten von MaRC verteilen sich auf fünf Racks. Sie sind mit 2.0 GHz (1. Ausbaustufe) bzw. 2.4 GHz (2. Ausbaustufe) getaktet. Die 57 neueren Knoten verfügen jeweils über 16 GByte DDR2 Hauptspeicher, die älteren Knoten sind noch mit DDR Hauptspeicher (8 GByte und zum Teil 16 GByte) bestückt. Die innere Vernetzung ist doppelt via GigaBit Ethernet ausgelegt, 24 der neueren Knoten sind zusätzlich über ein InfiniBand-Netzwerk verbunden. Ein Knoten dient als (NFS-)Fileserver und stellt über ein RAID-Subsystem 4 TByte Plattenplatz (brutto) bereit. Seit Mitte 2009 wird zusätzlich über die InfiniBand-Knoten ein verteiltes (GlusterFS-)Filesystem realisiert. Alle Compute-Knoten haben jeweils eine interne 250 GByte SATA-Festplatte.

3.5.4 Nutzung

Das Nutzungsprofil variiert sehr stark in Abhängigkeit von der Benutzergruppe. Während ein Teil der Benutzer ausschließlich fachspezifische Standardprogramme einsetzt, entwickeln und nutzen andere eigene Software; zum Teil ist auch die Organisation paralleler und verteilter Prozesse selbst Gegenstand der Forschung. Durch die Kooperation aller Beteiligten konnten mehrfach Lösungen für besondere Anforderungen gefunden werden, ohne den Produktionsbetrieb signifikant zu beeinträchtigen.

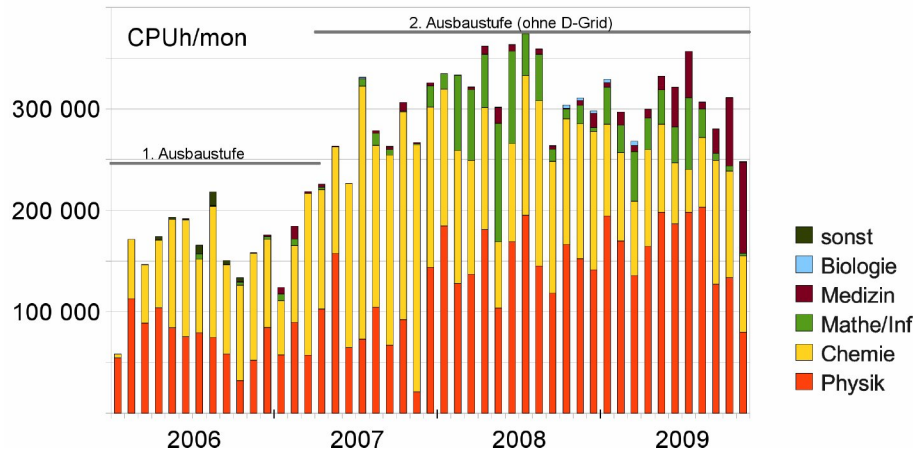


Abbildung 5.4.1: Nutzung von MaRC durch verschiedene Fachbereiche. Knoten der 2. Ausbaustufe wurden ihrer höheren Taktfrequenz entsprechend mit dem Faktor 1.2 gewichtet.

Die Nutzung von MaRC durch serielle und parallele Programme hält sich in etwa die Waage. Bei der parallelen Nutzung dominieren Jobs mit 2-8 Prozessen; nur ganz selten werden Jobs mit 64 und mehr Prozessen gestartet. Insbesondere viele serielle Jobs haben einen hohen Hauptspeicherbedarf, so dass manchmal nicht die CPU-Leistung, sondern die Verfügbarkeit von Hauptspeicher die Nutzung des Clusters limitiert; dies ist der Hauptgrund für die niedrigeren Balken in der CPU-Nutzungsstatistik, die in Abb. 5.4.1 dargestellt ist. Die Nutzung des Clusters durch SMP-Jobs ist durch die Architektur der Knoten auf maximal vier Cores bzw. parallele Threads limitiert. Für höhere Anforderungen eignet sich z.B. der HLR in Darmstadt. Durch den Betrieb von MaRC reduzierte sich dort andererseits der Marburger Anteil der HLR-Nutzung, so dass wertvolle Ressourcen gezielter eingesetzt werden konnten.

Die User-Dokumentation von MaRC wird im Sinne eines Gemeinschaftsprojekts in der Form eines Wikis gepflegt.

3.5.5 Verweise

1. MaRC Homepage: <http://www.uni-marburg.de/hrz/infrastruktur/zserv/cluster>
2. User-Dokumentation zu MaRC im Wiki: <http://clust-doc.hrz.uni-marburg.de>