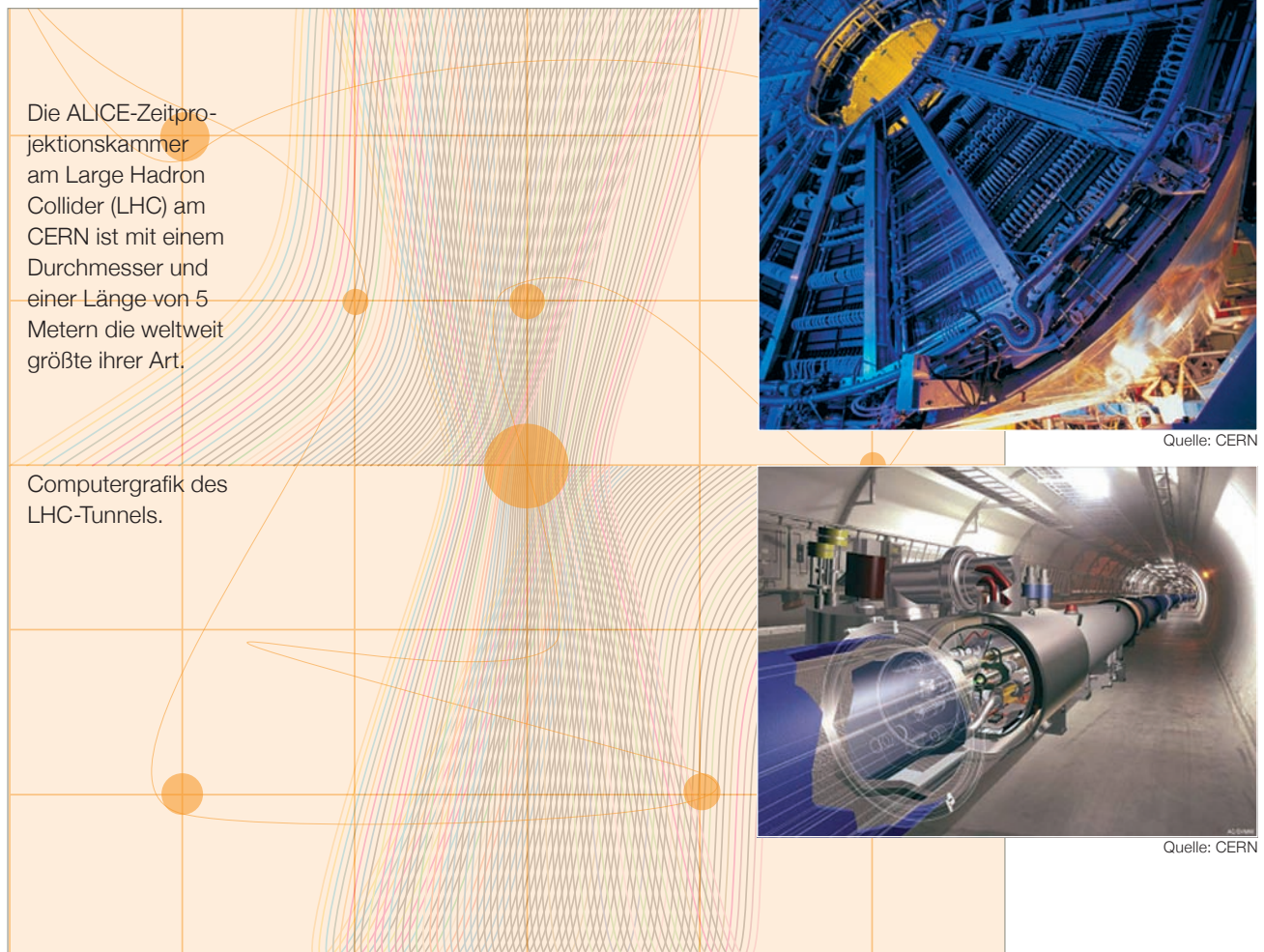


Steinbuch Centre for Computing NEWS SCC

Weltweite Testumgebung für Cloud Computing

SCC Mitbegründer der Gauss-Allianz

Datenanalyse für den leistungsstärksten Teilchenbeschleuniger der Welt



INHALT

4
SCC mit HP, Intel und Yahoo! an weltweiter Testumgebung für Cloud Computing beteiligt
SCC betreibt eines von sechs weltweiten „Center of Excellence“

6
SCC Mitbegründer der Gauss-Allianz Deutschlands namhafteste Rechenzentren bündeln ihre Ressourcen

8
SCC speichert und analysiert Daten des leistungsstärksten Teilchenbeschleunigers der Welt

12
Grid Computing Centre Karlsruhe (Gridka) Hardwareabnahme mit Mikroskopiedaten von Zebrafischembryonen
Statt 20 CPU-Jahren nur wenige Wochen Rechenzeit

14
Rekonstruktion der embryonalen Entwicklung mit DSLM

16
Grids mit EGEE III auf der Überholspur
Internationale Forschungsgemeinschaft setzt auf weltweit vernetzte Rechnerressourcen

18
Baden-Württemberg Grid - Regionale Grid-Infrastruktur für Baden-Württemberg

20
GridKa School 2008
Applikationen - Middleware - Techniken

21
Integrated Site Security for Grids
Projekt zur modellierten Zusammenführung von Grid und Site Security erfolgreich abgeschlossen

22
Das SCC stellt sich vor
In dieser Ausgabe: Die Abteilung Integration und Virtualisierung (IVI)

27
Mitarbeiter der Abteilung IVI

29
Netzwerksimulator optimiert Simulation drahtloser Netzwerke

30
SCC betreibt Intrusion Prevention-System mit 10 GBit/s Bandbreite
Neueste McAfee-Plattform erstmals im Produktionsnetz

31
SCC auf der Internationalen Supercomputing Conference „ISC 2008“ in Dresden
KIT belegt vier Plätze in der „TOP500“

32
Neues Buch zum Thema Netzwerk- und IT-Sicherheitsmanagement
Die Autoren Jochen Dinger und Hannes Hartenstein beleuchten Bezug zwischen theoretischen und betrieblichen Aspekten

34
Neuer „Masterkurs Parallele und Verteilte Systeme“

34
Lizhe Wang bester chinesischer Doktorand im deutschsprachigen Raum

35
MATLAB - Numerische Berechnungen einfach erledigen
Lizenzen für Universitätsangehörige und Studierende kostenfrei nutzbar

36
Zentrale Plattform für Team- und Projektarbeit
SCC bietet Sharepoint-Technologie innerhalb des KIT an

37
Spamfilterumstellung Campus-Nord
Neues System hat sich bewährt

37
Anregungen erwünscht
IT-Services werden weiter optimiert

38
Projekt „Gefahrstoffe an Schulen“ erfolgreich beendet

39
Einführungskurs SAS



EDITORIAL

Liebe Leserinnen und Leser,

„Cloud Computing“ heißt ein neues Forschungsfeld in der Computertechnologie. Dabei werden die kombinierten Ressourcen weltweit operierender IT- und Daten-Zentren in einer Rechner-Wolke im Internet zur Verfügung gestellt. Mit Hilfe dieser neuen Methode können nicht nur wissenschaftliche Anwendungen dynamisch betrieben werden, sondern auch Investitionen sowie Energie- und Personalkosten eingespart werden, da der Anwender selbst keine lokalen Systeme mehr betreiben muss.

Zusammen mit namhaften internationalen Partnern wie HP, Intel und Yahoo! richtet das SCC nun als eines von sechs Exzellenzzentren eine global verfügbare Testumgebung ein. Die Zusammenschaltung mit den anderen Testumgebungen wird noch in diesem Jahr erste weltweite Forschungsvorhaben zum Cloud Computing ermöglichen. Mit dieser gemeinsamen Initiative hat sich das SCC insbesondere die intensive Erforschung von Software, Hardware und Management innerhalb der „Cloud“ zum Ziel gesetzt, um so einen wesentlichen Beitrag zur Optimierung vorhandener IT-Infrastrukturen und internetbasierten Zusammenarbeit von Forschung, Wissenschaft und Industrie zu leisten.

In diesem Jahr wird auch der riesige Teilchenbeschleuniger (Large Hadron Collider, LHC) am Europäischen Forschungszentrum CERN in Genf seinen Betrieb aufnehmen. Die Rechenanforderungen stellen mit einem erwarteten Datenaufkommen von ca. 15 Petabyte pro Jahr eine noch nie dagewesene Herausforderung dar. Wir freuen uns daher umso mehr, dass das am SCC angesiedelte Grid Computing Centre Karlsruhe (GridKa) als eines von 11 weltweiten „Tier 1“-Grid-Rechenzentren maßgeblich an der Analyse und Speicherung der LHC-Daten beteiligt sein wird.

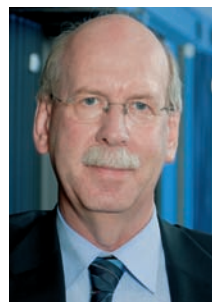
Mit Gründung der „Gauss-Allianz“ schließlich, in der sich im Juni Deutschlands führende Supercomputing-Zentren zusammengeschlossen haben, ist man der nachhaltigen Versorgung der computergestützten Wissenschaften in Deutschland mit HPC-Ressourcen der obersten Leistungsklassen ein großes Stück näher gekommen. Damit wird nicht nur das interdisziplinäre Wissenschaftsthema High Performance Computing (HPC) als eigenständige strategische Forschungsaktivität gestärkt, sondern auch die internationale Sichtbarkeit der deutschen Forschungsanstrengungen erhöht.

Viel Vergnügen bei der Lektüre wünschen Ihnen

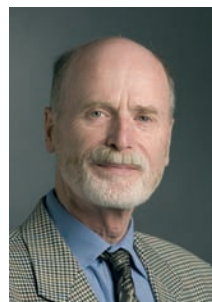
Hannes Hartenstein, Wilfried Juling und Klaus-Peter Mickel



Prof. Dr. Hannes Hartenstein
Foto: Privat



Prof. Dr. Wilfried Juling
Foto: Privat



Klaus-Peter Mickel
Foto: Privat

IMPRESSUM

September 2008

Herausgegeben im Auftrag des Direktoriums des Steinbuch Centre for Computing (SCC) von der Stabsstelle Öffentlichkeitsarbeit und Kommunikation

Anschrift: Steinbuch Centre for Computing (SCC)

Redaktion SCC-News

Zirkel 2

76128 Karlsruhe bzw.

Hermann-von-Helmholtz-Platz 1

76344 Eggenstein-Leopoldshafen

Fax: 0721/32550

<http://www.rz.uni-karlsruhe.de/publikationen/scc-news.php>

Redaktion:

Ursula Scheller (verantwortlich)

Telefon: 0721/608-4865

E-Mail: ursula.scheller@kit.edu

Layout und Bildredaktion: John Atkinson

Redaktionell bearbeitete Texte werden mit (red) gekennzeichnet. Nachdruck und elektronische Weiterverwendung von Texten und Bildern nur mit ausdrücklicher Genehmigung der Redaktion.

SCC mit HP, Intel und Yahoo! an weltweiter Testumgebung für Cloud Computing beteiligt

Steinbuch Centre for Computing (SCC) betreibt eines von sechs weltweiten „Center of Excellence“

HP, Intel und Yahoo! haben eine weltweite, offene Testumgebung für das Cloud Computing geschaffen. Innerhalb dieser Partnerschaft sollen mehrere Rechenzentren, unter anderem das SCC, das Cloud Computing in Forschung und Lehre weiter vorantreiben. Ziel der Initiative ist es, bei der Forschungsarbeit mit datenintensiven Internet-basiertem Computing finanzielle und logistische Hürden abzubauen und so die Zusammenarbeit von Industrie, Hochschulen und Regierungsbehörden zu fördern.

Mit ihrer Testumgebung stellen HP, Intel und Yahoo! ein über das Internet weltweit verfügbares Testumfeld zur Verfügung. Die Intention ist, Software, Hardware und das Management von Rechenzentren im Hinblick auf Cloud Computing so intensiv zu erforschen wie nie zuvor. Zudem wird die Initiative auch Forschungen im Bereich Anwendungen und Services für Cloud Computing unterstützen.

Partner der Initiative sind das Steinbuch Centre for Computing (SCC), die Infocomm Development Authority in Singapur (IDA) sowie die Universität von Illinois mit Sitz in Urbana-Champaign in Zusammenarbeit mit der National Science Foundation.

Zunächst wird das Testumfeld aus sechs so genannten „Center of Excellence“ bestehen. Diese sind bei der Infocomm Development Authority, der Universität von Illinois, dem SCC des KIT, den HP Labs sowie bei Intel Research und Yahoo! angesiedelt. Jeder Standort wird eine Infrastruktur für Cloud Computing bereitstellen – größtenteils basierend auf HP-Hardware und Intel-Prozessoren. Jeweils 1.000 bis 4.000 Prozessor-Kerne unterstützen die datenintensive Forschung. Alle Testzentren sollen noch in diesem Jahr voll betriebsfähig sein und dann ausgewählten Forschern in aller Welt zur Verfügung stehen.

„Die Cloud Computing-Initiative eröffnet uns und unseren Kunden völlig neue Möglichkeiten, wissenschaftliche Anwendungen dynamisch zu betreiben“, sagt Professor Wilfried Juling, geschäftsführender Direktor des Steinbuch Centre for Computing (SCC). „Die aktuelle Forschung ist ein strategisch wichtiger Schritt, Service-orientierte Architekturen auf Basis der nächsten Internet-Generation weiterzuentwickeln.“

Yahoo! möchte mit dem Testumfeld unter anderem seine technische Vorreiterrolle bei Open-Source-Projekten auf Basis von Apache Hadoop und anderen Open-Source-Plattformen wie Pig weiter ausbauen. Apache Hadoop ist ein Open-Source-Projekt für Distributed Computing der Apache Software Foundation. Bei Pig handelt es sich um eine von Yahoo! Research entwickelte Programmiersprache.

„Mit der Testumgebung von HP, Intel und Yahoo! verstärken wir unser Engagement für eine weltweite Forschungsgemeinschaft mit dem Ziel, die neuen Möglichkeiten, die das

Internet bietet, für wissenschaftliche Zwecke zu nutzen“, erklärt Prabhakar Raghavan, Leiter von Yahoo! Research. „Die Zentren ermöglichen Forschern nicht nur, Anwendungen zu testen, die über das Internet bereitgestellt werden. Zudem erhalten sie auch Zugriff auf das zugrundeliegende IT-System. So können sie besser nachvollziehen, wie Systemsoftware und -hardware in einer Cloud-Umgebung funktionieren.“

Forscher der HP Labs, HPs zentraler Forschungs- und Entwicklungseinrichtung, werden die Testumgebung dazu nutzen, um verstärkt Forschung auf den Gebieten der intelligenten Infrastruktur und dynamischer Cloud Services zu betreiben. Seit Kurzem unterstützen die HP Labs verstärkt HP und seine Kunden dabei, Cloud Computing effizient zu nutzen. Dies ist eine der treibenden Kräfte von HPs „Everything as a Service“. Hinter dieser Strategie steht das Ziel, dass IT-Komponenten und Services in Cloud-Umgebungen nahtlos zusammenarbeiten. Auf diese Weise stehen diese Services Unternehmen und Einzelpersonen abhängig von Ort, Zeit und Präferenzen und je nachdem, welcher Community sie angehören, genau so zur Verfügung, wie diese sie benötigen.

„Um das volle Potenzial von Cloud Computing auszuschöpfen, muss die Technologiebranche Cloud als eine Plattform betrachten, auf der sich neue Services und Erfahrungen schaffen lassen. Dies erfordert einen völlig neuen Ansatz im Hinblick darauf, wie wir Cloud-Infrastrukturen und -Services planen, implementieren und verwalten“, sagt Prith Banerjee, Senior Vice President of Research bei HP und Direktor der HP Labs. „Die Testumgebung für Cloud Computing von HP, Intel und Yahoo! ermöglicht uns, die fähigsten Köpfe aus Wirtschaft, Universitäten und Behörden einzubinden, um die Innovation in diesem Bereich voranzutreiben.“

Intel ist ein führender Anbieter von Plattform-Technologien. Dazu gehören Prozessoren, Chipsets, Netzwerkprodukte und Festkörperspeicher (SSD, Solid State Drives) für Cloud Computing-Rechenzentren. Aktuelle Plattform-Features wie das Data Center Management Interface (DCMI), Node Manager (NM) und Virtualisierung wurden entwickelt, um die Verwaltung und Energieeffizienz von Rechenzentren zu verbessern. Dieser offene, kollaborative Forschungsansatz erlaubt Wissenschaftlern vollen Zugang zur System-Hardware, um bestehende und zukünftige Plattform-Features weiter zu entwickeln.



„Cloud Computing“ eröffnet ganz neue Möglichkeiten, Rechner- und Daten-Ressourcen zu erschließen.

„Wir freuen uns über die Kooperation mit der akademischen Forschungs-Community – offene Zusammenarbeit mit Wissenschaftlern ist Teil der DNA von Intel Research“, bemerkt Andrew A. Chien, Vice President und Direktor von Intel Research. „Es ist wichtig, umfangreiche Testumgebungen zu entwickeln, weil sie Innovationshemmnisse verringern und die Möglichkeit bieten, im entsprechenden Maßstab zu experimentieren und zu lernen. Intels Unterstützung für Tashi, ein Open-Source Cluster Management-System für Cloud Computing und nun die Testumgebung für Cloud Computing von HP, Intel und Yahoo! – dies ist die logische Weiterentwicklung unserer fortlaufenden, für alle Seiten fruchtbaren Partnerschaften mit der Forschungs-Community, wie zum Beispiel im Fall der Universal Parallel Computing Research Centers.“

Die Infocomm Development Authority wird die Forschung in der Testumgebung fördern, indem sie den Anwendern die erforderlichen Rechenkapazitäten zur Verfügung stellt, um Software und Anwendungen für Cloud Computing zu entwickeln. Sie wird außerdem die Testumgebung und die Partnerschaften mit der Industrie nutzen, um Studenten und Fachkräfte vor Ort zu Technologien und Programmen im Zusammenhang mit Cloud Computing zu schulen.

„Auf Basis der in Singapur verfügbaren Ressourcen für die Forschung und Entwicklung rund um Cloud Computing können wir mit gleich gesinnten Partnern dieses Gebiet gemeinsam vorantreiben“, bemerkt Khoong Hock Yun, Assistant Chief Executive der Infrastructure Development Group der Infocomm Development Authority in Singapur. „Für viele ist Cloud Computing die nächste Evolutionsstufe in der IT – und sie ist möglicherweise die nächste ‚Plattform‘ für innovative Ökosysteme. Partnerschaften wie diese werden es Singapur ermöglichen, den neuen Ansatz für ein besseres wirtschaftliches und soziales Wachstum zu nutzen.“

(red)

GAUSS-ALLIANZ



Foto: Matthias Müller, TU Dresden

Kräfte gebündelt: Die Vertreter der führenden deutschen Rechenzentren und Protagonisten des High Performance Computing gründeten in Dresden die Gauss-Allianz.

SCC Mitbegründer der Gauss-Allianz

Deutschlands namhafteste Rechenzentren bündeln ihre Ressourcen

Am Rande der Internationalen Supercomputing Konferenz in Dresden wurde am 18. Juni die Vereinbarung zur Gründung der Gauss-Allianz unterzeichnet. Damit schaffen die zwölf beteiligten Supercomputerzentren in Deutschland einen weltweit einzigartigen Rechnerverbund im Bereich des High Performance Computing (HPC).

„Die damit verbundenen Möglichkeiten für Wissenschaft und Wirtschaft werden im internationalen Vergleich ein entscheidendes Wettbewerbsvorteil für Deutschland sein“, sagte Thomas Rachel, Parlamentarischer Staatssekretär im Bundesministerium für Bildung und Forschung, zur vereinbarten Zusammenarbeit.

Fragestellungen in Wissenschaft, Forschung, Technik und Gesellschaft verlangen nach immer umfassenderen und genaueren Modellbildungen für komplexe Systeme und nach immer leistungsfähigeren Methoden zur Problemlösung. Dabei ist der einfache Zugang zu Hoch- und Höchstleistungsrechnern und deren Nutzung für das wissenschaftliche Arbeiten mittlerweile für viele Forschungsbereiche unverzichtbar. Die Simulation mit Supercomputern hat darüber hinaus in Wirtschaft und Industrie eine strategische Bedeutung, denn viele Spitzenprodukte lassen sich bereits heute ohne Modellbildung und High Performance Computing nicht entwerfen.

„Ziel der Gauss-Allianz ist die nachhaltige Versorgung der computergestützten Wissenschaften in Deutschland mit HPC-Ressourcen der obersten Leistungsklassen. Dabei sollen insbesondere technisch-wissenschaftliche und wirtschaftsbezogene Anwendungsprobleme berücksichtigt werden“, erklärt Prof. Wilfried Juling, geschäftsführender Direktor des SCC.

Daneben fördert und koordiniert die Gauss-Allianz die HPC-nahen Forschungsaktivitäten in Deutschland. „Das interdisziplinäre Wissenschaftsthema ‚HPC‘ soll damit als eigenständige strategische Forschungsaktivität gestärkt und die internationale Sichtbarkeit der deutschen Forschungsanstrengungen weiter verbessert werden“, ergänzt Klaus-Peter Mickel, Technisch-Wissenschaftlicher Direktor des SCC.

Die Gauss-Allianz soll zudem Wissenschaft, Wirtschaft, Bund, Länder und Trägerorganisationen bei der Erschließung und Weiterentwicklung des Wissenschaftlichen Rechnens und des Hoch- und Höchstleistungsrechnens beraten.

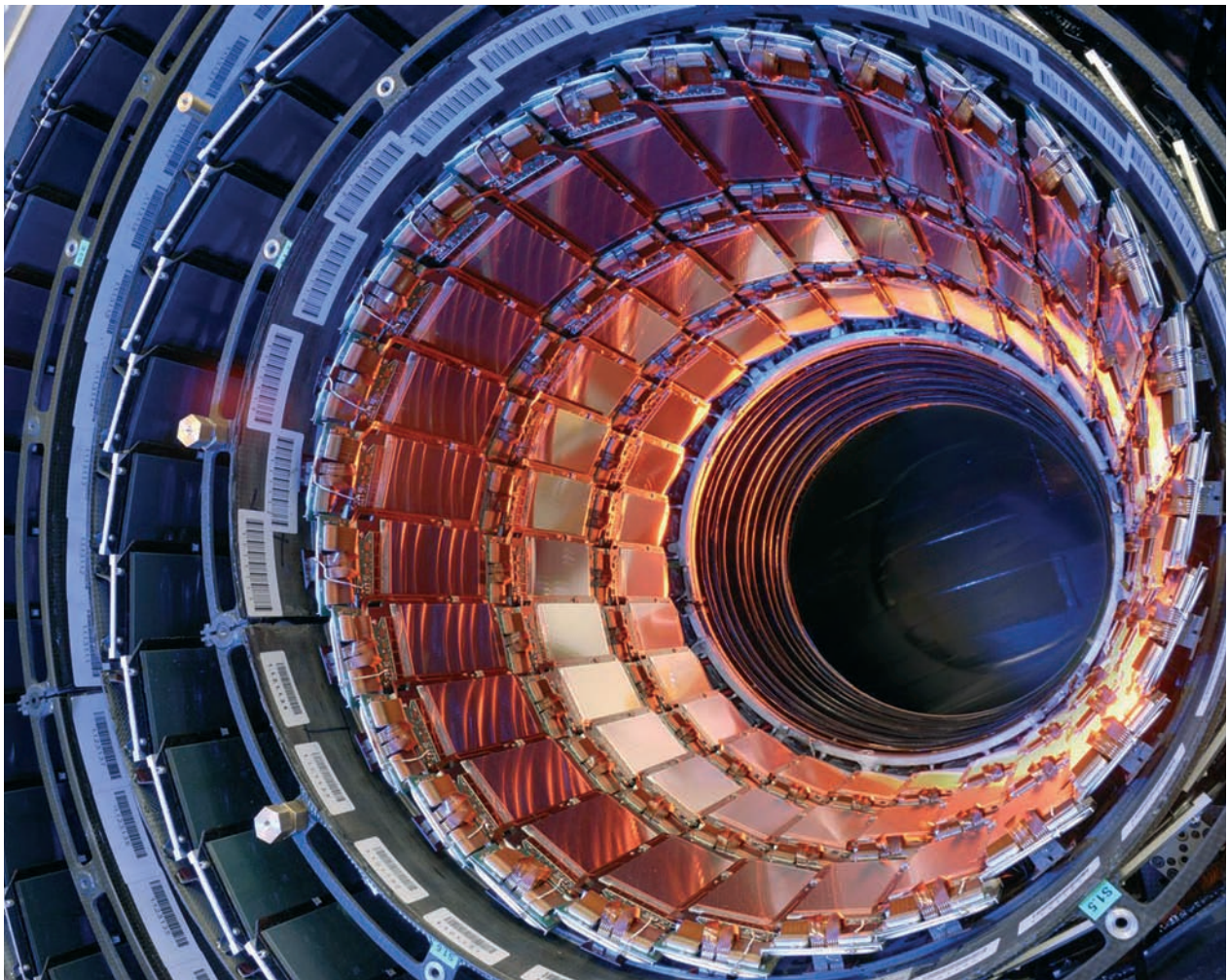
„Nach der vom BMBF in 2006 initiierten Etablierung des Gauss Centre for Supercomputing (GCS) ist dies der zweite Schritt zur Bildung eines Europäischen HPC-Ecosystems. Deutschland stärkt mit der Unterzeichnung der Gauss-Allianz nachhaltig seine führende Rolle im europäischen Supercomputing“, sagte der Vorsitzende des GCS, Prof. Achim Bachem.

Mit der Unterzeichnung setzen die Gründer die in der BMBF-Studie „High Performance Computing in Deutschland - Argumente zur Gründung einer strategischen Allianz“ vorgeschlagenen Empfehlungen um und geben ihrer langjährigen engen Zusammenarbeit unter dem Namen „Gauss-Allianz“ eine eigene organisatorische Basis.

Die 12 Unterzeichner der Vereinbarung sind das Gauss Centre for Supercomputing (GCS), das Rechen- und Kommunikationszentrum Aachen, der Norddeutsche Verbund für Hoch- und Höchstleistungsrechnen (HLRN) - vertreten durch die beiden Betreiberrechenzentren dem Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) und dem Regionalen Rechenzentrum für Niedersachsen (RRZN), das Zentrum für Informationsdienste und Hochleistungsrechnen (ZIH), das Regionale Rechenzentrum Erlangen (RRZE), das Goethe-Zentrum für Wissenschaftliches Hochleistungsrechnen (G-CSC), das Rechenzentrum Garching der Max-Planck-Gesellschaft (RZG), das Deutsche Klimarechenzentrum (DKRZ), das Steinbuch Centre for Computing (SCC) in Karlsruhe, der Deutsche Wetterdienst (DWD) und das Paderborn Center for Parallel Computing (PC²).

(red)

SCC speichert und analysiert Daten des leistungsstärksten Teilchenbeschleunigers der Welt



Mit dem CMS-Detektor am LHC können die unterschiedlichen Teilchen, die bei Zusammenstößen der Protonen entstehen, besonders gut vermessen werden. Die Abbildung zeigt die erste Hälfte der inneren Spurtrommel.

Quelle: CERN

In Kürze nimmt der Large Hadron Collider (LHC) am Europäischen Forschungszentrum CERN seinen Betrieb auf. Damit die riesigen Datenmengen dieses Teilchenbeschleunigers ausgewertet werden können, ist ein globales Netz von Computern geknüpft worden. Das am Steinbuch Centre for Computing (SCC) angesiedelte Grid Computing Centre Karlsruhe (GridKa) ist als eines von elf weltweiten „Tier 1“-Grid-Rechenzentren an der Datenanalyse des LHC beteiligt.

Pro Jahr 15 Millionen Gigabyte Datenvolumen

Der LHC ist ein kreisförmiger Teilchenbeschleuniger mit einem Umfang von 27 Kilometern. In vier riesigen unterirdischen Detektoren, den „Experimenten“ ALICE, ATLAS, CMS und LHCb, bringen die Wissenschaftler am CERN gegenläufig umlaufende Protonenpakete mit einer Schwerpunktsenergie von 14 Tera-Elektronvolt (TeV) zur Kollision, um Fragen nach dem Ursprung der Materie, der Symmetrie von Materie und Antimaterie, den Eigenschaften Dunkler Materie und Dunkler Energie sowie der Vereinigung der vier bekannten fundamentalen Kräfte der Physik zu einer Universalwechselwirkung zu untersuchen. Etwa sechs Wochen im

Jahr werden hierfür sogar Bleikerne mit einer Schwerpunktsenergie von mehr als 1.000 Tera-Elektronvolt (5,7 TeV pro Nukleon) zur Kollision gebracht. „Einschließlich der für die Auswertung notwendigen Simulationen erwartet man am CERN ein Datenaufkommen von etwa 15 Petabyte pro Jahr, das sind 15 Millionen Gigabyte“, erklärt Dr. Holger Marten, bisheriger Projektleiter des am SCC angesiedelten Tier 1-Grid Computing Centre Karlsruhe (GridKa) und Leiter der Abteilung „Verteilte Systeme und Grid“.

Weltweite Grid Computing-Infrastruktur zur Auswertung der Daten

Um diese riesigen Datenmengen zu verarbeiten, wurde 2001 das „Worldwide LHC Computing Grid Project“ (LCG) als weiteres Teilprojekt des LHC gegründet. „Ziel des LCG ist, eine weltweite Computing-Infrastruktur zu schaffen, die mehreren tausend beteiligten Wissenschaftlern aus über 50 Nationen die transparente Speicherung und Auswertung der LHC-Daten über Grid-Methoden ermöglicht“, erklärt Holger Marten weiter. Als Basis-Struktur dient hierfür ein Schichtenmodell (Multi-Tier) aus Rechenzentren mit unterschiedlichen Aufgaben. Das CERN hat als Tier-0 die Aufgabe, die Rohdaten jedes der vier Experimente zu speichern und die erste Rekonstruktion dieser Daten durchzuführen. Eine zweite Kopie der Rohdaten sowie die Ergebnisse der ersten Rekonstruktion werden an insgesamt 11 nationale oder multinationale Tier 1-Zentren verteilt, so auch ans GridKa. Dort finden im Verbund zentral gesteuerte, reguläre Datenanalysen, weitere Rekonstruktionen sowie die Speicherung der Daten aus den Tier 2-Zentren statt. Darüber hinaus bieten die Tier 1-Zentren regionale, technische Unterstützung für die Benutzer und Tier 2-Zentren. Die Hauptaufgabe der ca. 100 weltweit verteilten Tier 2-Zentren besteht wiederum darin, Monte-Carlo-Simulationen durchzuführen und Kapazitäten für die Datenanalysen der Endbenutzer bereitzustellen.

Das LCG-Projekt erstreckt sich im Wesentlichen über drei Phasen: Die Entwicklung einer prototypischen Grid-Infrastruktur bis 2005, die Hauptausbauphase 2006 bis 2007 sowie die Produktionsphase ab 2008 bis zum Ende der LHC-Datenauswertung. Neben dem Auf- und Ausbau der reinen Hardwareinfrastruktur (Rechner, Datenspeicher, Netzwerke mit adäquater Bandbreite) lag das Hauptaugenmerk des LCG-Projektes während der ersten beiden Phasen in der Entwicklung robuster und skalierbarer Massenspeichersysteme sowie einer Grid-Infrastruktur, die einen sicheren und transparenten Zugriff auf alle Ressourcen im internationalen Kontext erlaubt. „Allein die Organisation einer vertrauensvollen Zusammenarbeit von über hundert Rechenzentren unter Berücksichtigung individueller Datenschutzerfordernungen sowohl der Nutzer als auch der Zentren selbst stellte neben der bloßen technischen Aufgabe, täglich bis zu 100.000 Jobs zu verteilen, zu verfolgen und zu verarbeiten, eine besondere Herausforderung dar“, sagt Holger Marten.

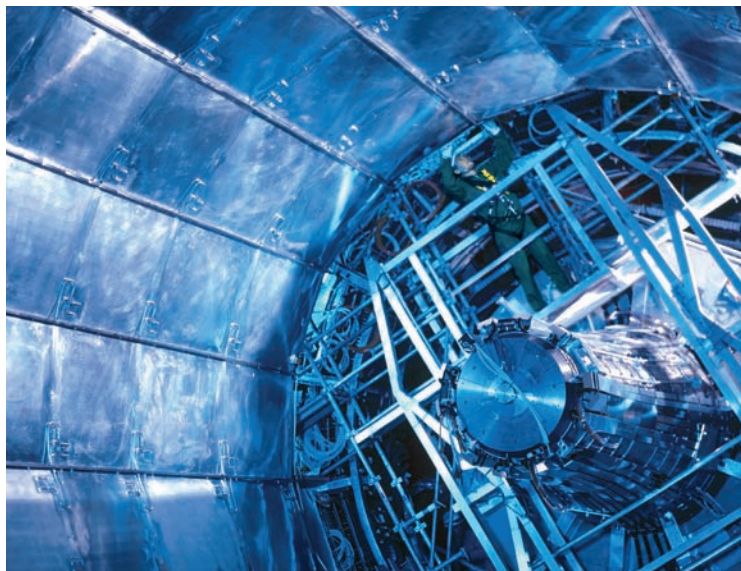


Der innere Spurdetektor des ALICE-Experiments wird integriert.

Quelle: CERN

Gesamtsystem erfolgreich getestet

Um die technischen Voraussetzungen zu schaffen, wurden hauptsächlich drei Grid-Produkte über gemeinsame Verwaltungs- und Monitoring-Strukturen miteinander verbunden beziehungsweise vom LCG in ihrer Entwicklung durch ent-



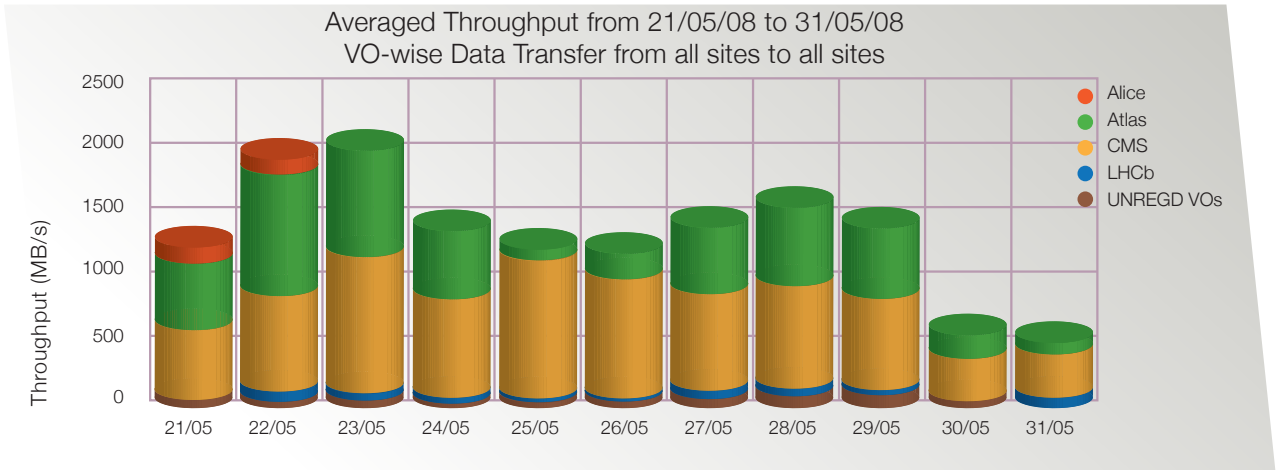
Das Experiment ALICE ist für Kollisionen von Schwerionen optimiert.

Quelle: CERN

sprechende Anforderungen maßgeblich beeinflusst: LCG/gLite der europäischen DataGrid und EGEE-Initiativen (Enabling Grids for E-Science), das amerikanische OSG (Open Science Grid) sowie die ARC Software (Advanced Resource Connector) des NorduGrid-Projektes mehrerer nordeuropäischer Länder. Parallel zum Aufbau dieser Computing-Infrastruktur haben die LHC-Experimente die Entwicklung der Software für Detektor- und Teilchensimulationen sowie für die Datenauswertung vorangetrieben. Die ersten beiden Phasen des LCG-Projektes waren daher geprägt von mehreren so genannten „Data Challenges“ wachsender Komplexität und wurden mit zwei „Combined Computing Readiness Challenges“ abgeschlossen, in denen schließlich das Gesamtsystem aus Grid-Infrastruktur und Datenprozessierung im internationalen Kontext hinsichtlich Robustheit und Skalierbarkeit mit nahezu Designkapazität getestet wurde. Bis zu 60 Zentren pro Experiment waren daran beteiligt, innerhalb weniger Wochen über mehrere Millionen Jobs hunderte von Terabytes an Daten zu simulieren, zu verteilen und auszuwerten.

Mehr als 15 Tier 2-Zentren in Europa an GridKa angeschlossen

Das Forschungszentrum Karlsruhe ist seit 2001 mit dem Aufbau und der Entwicklung des Grid Computing Centre Karlsruhe (GridKa) am LCG-Projekt beteiligt. Die Hauptziele dieses von der deutschen Kern- und Elementarteilchenphysik-Community initiierten Projektes bestanden darin, ein regionales Grid-Daten- und -Rechenzentrum der Kategorie Tier 1 für die künftige Auswertung der LHC-Daten zu entwickeln. Mehr als 15 Tier 2-Zentren in sechs europäischen Ländern sind mit GridKa als ihrem zugeordneten Tier 1-Zentrum verbunden, speichern simulierte Daten und beziehen Expe-



Datentransfer innerhalb LCG zwischen allen Zentren während der „Combined Computing Readiness Challenges“ im Mai 2008.
Quelle: <http://gridview.cern.ch>

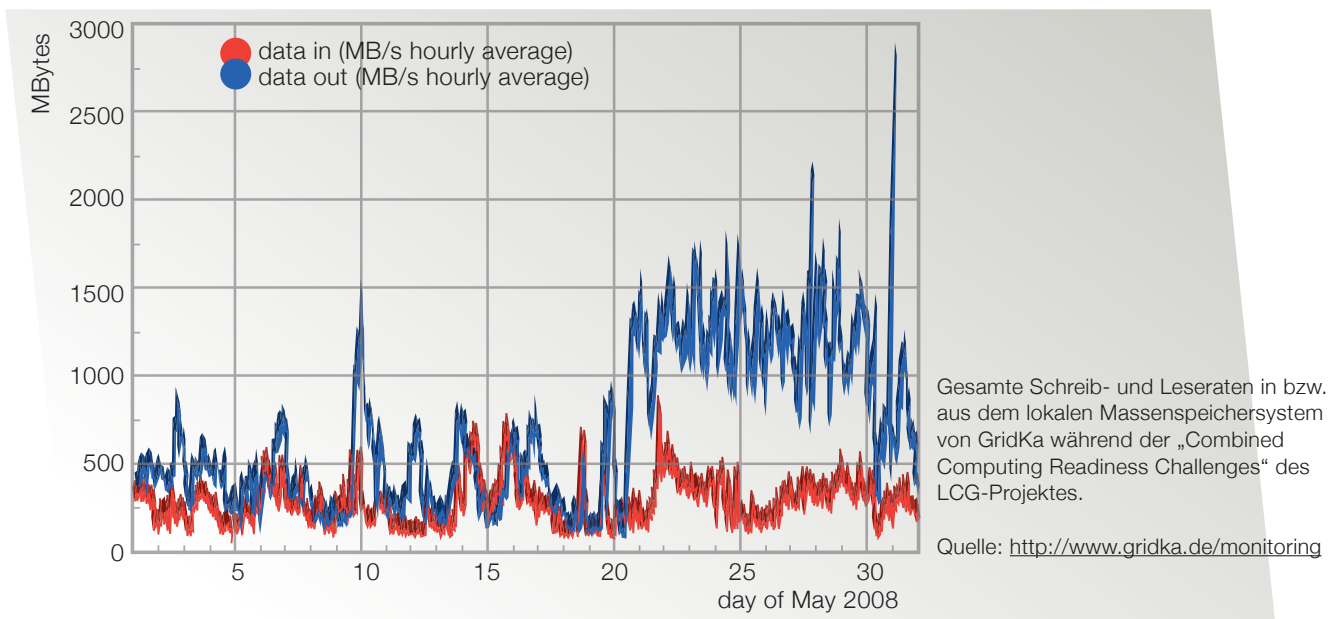
periment-Daten zur Auswertung durch die Physiker. Neben den Daten-Speichern stellt GridKa auch zentrale Grid-Dienste, wie File-Kataloge, Informationssysteme und einen File-Transfer-Service für seine Tier 2-Partner zur Verfügung. Gleichzeitig sollten mit GridKa Ressourcen für diejenigen deutschen Forschungsgruppen zur Verfügung gestellt werden, die an der Datenauswertung der bereits laufenden Beschleuniger-Experimente „BaBar“ am SLAC (Stanford Linear Accelerator Center), „CDF“ und „D0“ am FermiLab (Fermi National Accelerator Laboratory) und COMPASS am CERN beteiligt sind.

Darüber hinaus wurden im Laufe der letzten Jahre über die deutsche Grid-Initiative „D-Grid“ des Bundesministeriums für Bildung und Forschung sowie über Mittel aus dem Impuls- und Vernetzungsfonds der Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF) Ressourcen für weitere Nutzergruppen integriert. „Auf diese Weise konnte nicht nur eine optimale Synergie zwischen den Computing-Erfahrungen bereits laufender und künftiger Hochenergiephysikexperimente hergestellt, sondern die Grid-Technologie als solche - nicht zuletzt durch die Verbindung und Zusammenarbeit der beiden HGF-

Programme „Struktur der Materie“ und „Wissenschaftliches Rechnen“ - auch auf andere Wissenschaftsdisziplinen übertragen werden“, ergänzt Dr. Andreas Heiss, seit Anfang Juli 2008 neuer Projektleiter von GridKa.

Zuverlässiges Großgerät für die Wissenschaft

Das Projekt GridKa wurde seit Bestehen zweimal jährlich erweitert, die Hauptausbauphase mit dem April-Meilenstein 2008 erfolgreich abgeschlossen. Das Gesamtsystem besteht gegenwärtig aus über 1.100 Rechenknoten mit insgesamt



über 5.700 Prozessorkernen und mehr als 11 TB RAM. GridKa verfügt über ca. 11.000 Festplatten mit einer nutzbaren Gesamtkapazität von 3.600 TB und zwei Magnetbandroboter mit insgesamt 12.000 Stellplätzen und gegenwärtig 3.700 TB Speicherkapazität. Die externe Anbindung zur Datenquelle am CERN sowie zu den übrigen Tier 1- und Tier 2-Zentren wurde in enger Zusammenarbeit mit dem Verein zur Förderung eines Deutschen Forschungsnetzes (DFN-Verein) über ein redundantes Netzwerk mit einer Gesamtbandbreite von 52 Gbit/s realisiert.

„Redundant ausgelegte, zentrale Komponenten und Grid-Services für Informationssysteme, die Verteilung und Annahme von Jobs, Datentransfers, Datenspeicherungen usw. sowie eine detaillierte Überwachung von über 22.000 Systemkomponenten und Einzeldiensten sorgen dafür, dass das SCC im internationalen Grid-Verbund mit GridKa auf lange Sicht ein robustes und zuverlässiges Großgerät für die Wissenschaft anbieten kann“, so Andreas Heiss.

Die ersten LHC-Daten werden in Kürze erwartet - auf die Ergebnisse darf man gespannt sein.

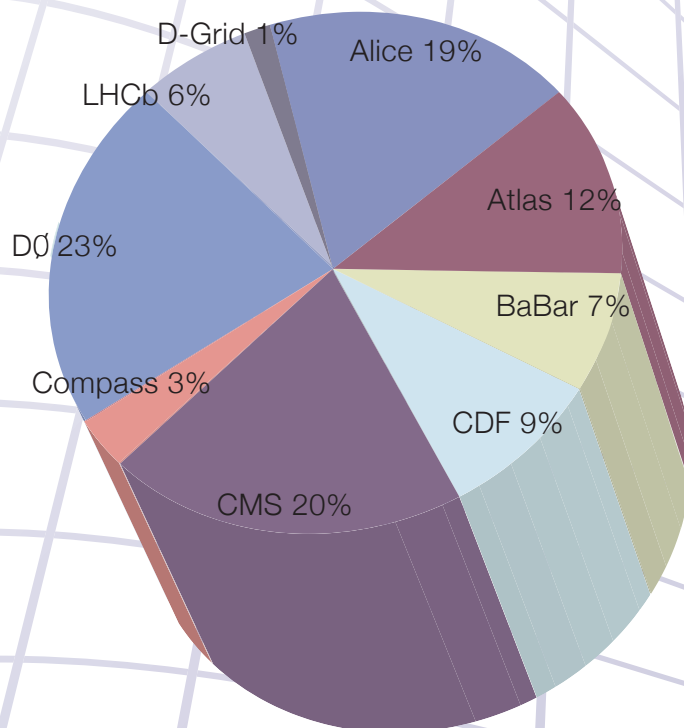
(red)

Weitere Informationen

LHC Beschleuniger
<http://lhc.web.cern.ch/lhc/>

Worldwide LHC Computing Grid
<http://lcg.web.cern.ch/LCG/>

GridKa
<http://www.gridka.de>



Verteilung der verbrauchten Rechenzeit am GridKa in 2007. Die Daten umfassen 3,54 Millionen Rechenaufträge (Jobs) mit einer Gesamtnutzungszeit von 16,4 Millionen Stunden.

Grid Computing Centre Karlsruhe (Gridka)

Hardwareabnahme mit Mikroskopiedaten von Zebrafischembryonen

Statt 20 CPU-Jahren nur wenige Wochen Rechenzeit

Durch die gleichzeitige Nutzung von bis zu 3.800 Prozessorkernen konnten die Wissenschaftler des GridKa-Projektes innerhalb weniger Wochen einen Bilddatensatz des Instituts für Toxikologie und Genetik (ITG) verarbeiten, für den ursprünglich ein Rechenzeitbedarf von 20 CPU-Jahren abgeschätzt wurde.

Seit seiner Gründung im Jahre 2001 werden Rechenleistung, Festplatten- und Magnetbandkapazitäten des GridKa im Rahmen entsprechender Projektmeilensteine zweimal jährlich erweitert. Die neue Hardware wird dann zur Qualitätssicherung vor der offiziellen Benutzungsfreigabe jeweils intensiven Abnahmetests unterzogen. Als Anfang 2008 die Clustererweiterung um ca. 480 neue Rechenknoten mit jeweils acht Prozessorkernen zum Meilenstein April 2008 installiert wurde, fragte das Institut für Toxikologie und Genetik (ITG) an, ob das SCC kurzfristig massive Rechenkapazitäten zur Verarbeitung eines Testdatensatzes bereitstellen könne. Hintergrund war, dass mit einem vom ITG und dem European Molecular Biology Laboratory (EMBL, Heidelberg) gemeinsam entwickelten, neuen Mikroskop die embryonale Entwicklung von Zebrafärblingen aufgezeichnet worden war und die gesammelten Daten (rund 1 TB) zur Auswertung anstanden.

Eine Analyse der Anfrage ergab, dass sich die auszuwertenden digitalen Bilddaten gut in kleine Portionen aufteilen ließen, die auf den lokalen Festplatten der Clusterknoten Platz hatten. Wesentlich war auch, dass diese Anwendung sehr unkritisch im Hinblick auf eventuelle Ausfälle war und deshalb die bei den brandneuen Systemen zu erwartende erhöhte Ausfallrate kein Problem darstellte. Die Art der Anwendung - kontinuierliche Zugriffe auf die Festplatten, hohe CPU-Auslastung erschien außerdem gut für Burn-In-Tests geeignet. So konnte innerhalb weniger Wochen durch gleichzeitige Nutzung von bis zu 3.800 Prozessorkernen der Testdatensatz des ITG verarbeitet werden, für den ursprünglich ein Rechenzeitbedarf von 20 CPU-Jahren abgeschätzt wurde.

„Green IT“ – Rechenleistung und Stromverbrauch

GridKa ist ein Cluster, das aus vielen, quasi handelsüblichen PCs besteht. Die jüngsten Serien haben jeweils zwei CPUs des Typs Intel Xeon E5345 (2,33 GHz) bzw. E5430 (2,66 GHz). Bei diesen CPUs handelt es sich um Quad-Core-CPU's, bei denen jeweils vier weitgehend unabhängige CPUs auf einem einzigen Chip vereinigt sind.

Jedes einzelne System ist mit 16 GB Hauptspeicher ausgestattet (2 GB pro Core) und hat zwei Festplatten von jeweils 250 GB, von denen eine für das Betriebssystem und die andere als Platten-cache für die Grid-Jobs dient. Die Rechner sind per Gigabit-Ethernet miteinander vernetzt, als Betriebssystem wird Scientific Linux eingesetzt.

Da die rasant steigenden Energiekosten auch den Rechenzentren zunehmend zu schaffen machen, hat die Industrie unter dem Label „Green IT“ inzwischen auf diesen Sachverhalt reagiert und bietet stromsparende Systeme an.

Auch für GridKa ist der Stromverbrauch und – da dieser zu fast 100% in Wärme umgesetzt wird – auch die ausreichende Kühlung der Systeme von Beginn an ein zentrales Thema. Besonders in den Jahren 2001 bis 2004 ist der Stromverbrauch vor allem der Clustersysteme drastisch angestiegen – nach detaillierten GridKa-Untersuchungen sogar wesentlich stärker, als die Rechenleistung im selben Zeitraum zugenommen hat.

Die Abbildung 1 macht diesen Zusammenhang anschaulich. Dargestellt ist die gemessene Rechenleistung pro Watt elektrischer Leistungsaufnahme für mehrere Generationen von Cluster-Systemen. Zur Messung der Rechenleistung wurde der Benchmark SPEC CPU2000 der Standard Performance Evaluation Corporation (SPEC) verwendet, ein Quasi-Industriestandard zur Performancemessung. Nachdem die Energie-Effizienz zwischen 2001 und 2004 um fast 45% zurückgegangen war, wurde seit 2005 in allen Ausschreibungen neben dem reinen Kaufpreis auch der Energieverbrauch berücksichtigt. Seitdem konnte die Energie-Effizienz bei jeder neuen Rechnergeneration gesteigert werden.

Abnahme- und Burn-In-Tests

Alle neuen GridKa-Systeme werden gründlich getestet, bevor sie zum jeweiligen Meilenstein in Produktion genommen werden. Zunächst erfolgen Abnahmetests, mit denen ermittelt wird, ob die neuen Rechner die geforderte Rechenleistung erreichen und der im Angebot genannte maximale Stromverbrauch eingehalten wird. Im Anschluss daran erfolgen so genannte Burn-In-Tests über mehrere Wochen, in denen die technische Fehlerfreiheit und Zuverlässigkeit geprüft wird, bevor die Systeme in Betrieb genommen werden.

Wie in Abbildung 2 schematisch dargestellt, sind Rechnerausfälle zwei Klassen zuzuordnen: Bei neuen Systemen ist typischerweise mit einzelnen Frühausfällen zu rechnen, die mehr und mehr nachlassen. Nach einer Betriebsdauer von erwartungsgemäß vier oder fünf Jahren treten dann vermehrt Ausfälle aufgrund alterungsbedingter Defekte auf.

Sinn der Burn-In-Tests ist es, die Frühausfälle zu erkennen und möglichst auch durch intensive Nutzung zu provozieren, bevor die Systeme in Betrieb genommen werden.

Manfred Alef, Dr. Holger Marten

Weitere Informationen

<http://www.gridka.de>

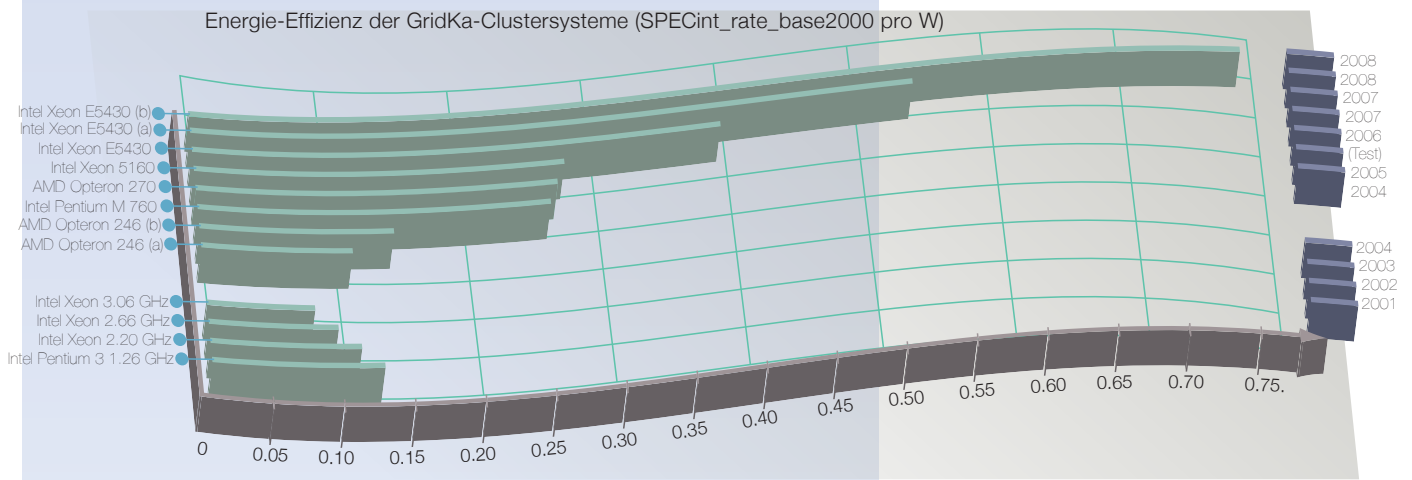


Abbildung 1:
 Energie-Effizienz (SPECint_rate_base2000 pro Watt)
 Benchmark: SPEC CPU2000 (CINT2000), 1 Kopie pro CPU/Core
 Betriebssystem: Scientific Linux 4 i386
 Compiler, Optimierungsoptionen: gcc-3.4.x -O3 -funroll-loops -march
 Messung der elektrischen Leistungsaufnahme: 1x cpuburn-1.4 pro CPU-Core

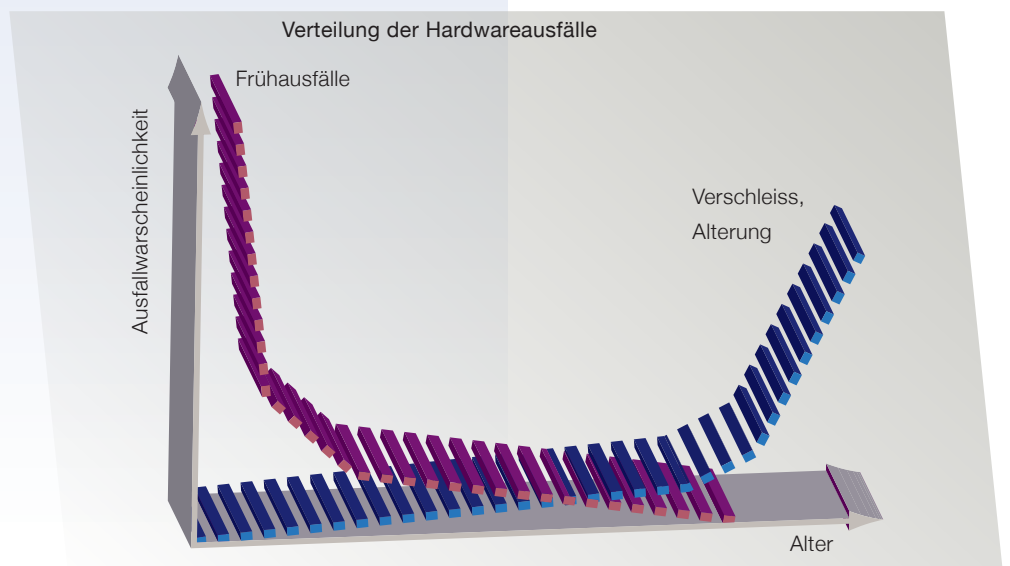


Abbildung 2:
 Zeitliche Verteilung der Hardwareausfälle

Rekonstruktion der embryonalen Entwicklung mit DSLM

Mit der am European Molecular Biology Laboratory (EMBL) Heidelberg neu entwickelten Mikroskopietechnik DSLM (Digital Scanned Laser Light Sheet Microscopy) können Embryos der zentralen biologischen Modellsysteme *Danio rerio* (Zebrafisch, auch als Zebrafisch bekannt), *Drosophila melanogaster* (Fruchtfliege) und *Mus musculus* (Maus) mit extrem hoher räumlicher und zeitlicher Auflösung dreidimensional während ihrer Entwicklung aufgenommen werden. Typischerweise startet eine solche Aufnahme zu einem Zeitpunkt, an dem der Embryo noch aus ein paar wenigen Zellen besteht (z. B. 16 oder 32), und endet mit der späten Organentwicklung. Für Zebrafisch-Embryos dauert dieses Zeitintervall 36 Stunden, in denen ca. 300.000 Bilder aufgenommen werden, was einer Datenmenge von 2 bis 3 TB entspricht. Die Bilder zeigen die dreidimensionale Helligkeitsverteilung des Embryos als Funktion der Zeit. Da solche Datenmengen schwierig zu bearbeiten sind und man eigentlich nur an den Koordinaten und der Bewegung der Zellen interessiert ist, müssen über Bildverarbeitungs-Algorithmen die wesentlichen Informationen extrahiert werden. In diesem Schritt werden die Bilddaten des Embryos in eine „digitale“ Darstellung des Embryos umgerechnet, d. h. in eine Datenbank, die präzise Informationen über die Position und Form aller Zellen im Embryo als Funktion der Zeit enthält (siehe Abbildung 1).

Die „digitalen Embryos“ erlauben es dann, einzelne Zellen als Funktion der Zeit zu verfolgen. Es ist auch möglich, die Teilung der Zellen zu erkennen, so dass das „Schicksal“ und die Herkunft aller bis zu 16.000 Zellen ziemlich genau rekonstruiert werden können. Mit Hilfe dieser Analysen lässt sich ein Abstammungsbaum aller Zellen sowie biologische Baupläne für alle wichtigen Gewebetypen und Organe bestimmen. Mit dieser Technik können nun zum Beispiel Defekte in der embryonalen Entwicklung oder in Modellsystemen für menschliche Krankheiten mit subzellulärer Auflösung in nie zuvor erreichtem Detailgrad analysiert und verstanden werden. Sowohl für medizinische Anwendungen wie auch für die biologische und biophysikalische Grundlagenforschung sind diese Datensätze von großem Wert und werden in der nahen Zukunft einen wesentlichen Beitrag zum Fortschritt in der Entwicklungsbiologie und der Erforschung und Bekämpfung schwerwiegender Krankheiten wie beispielsweise Krebs leisten.

Die Generierung der digitalen Embryos ist technologisch anspruchsvoll und auch sehr aufwendig. Die Mikroskopiedaten können aufgrund des enormen Umfangs ausschließlich über eine automatisierte Computer-Cluster-basierte Bildverarbeitungs-Pipeline rekonstruiert und in digitale Embryos konvertiert werden. Pro Embryo, d. h. pro Multi-Terabyte-Datensatz, erfordert diese Rekonstruktion im Durchschnitt eine Rechenzeit von 100.000 CPU-Stunden. Idealerweise müssen darüber hinaus Daten von vielen Embryos (ca. 10-

100 pro Spezies) erfasst werden, um die Variabilität der embryonalen Entwicklung zu berücksichtigen und statistisch signifikante Schlüsse zuzulassen. Daher ist ein leistungsfähiger Compute Cluster ein essentielles Werkzeug für die Analyse der DSLM-Daten.

In einer Zusammenarbeit zwischen dem EMBL Heidelberg und dem Steinbuch Centre for Computing (SCC) wurden bereits mehrere digitale Embryos errechnet. Dank der enormen Rechenleistung des am SCC angesiedelten GridKa-Clusters war es möglich, insgesamt mehr als fünf Tage der embryonalen Zebrafisch-Entwicklung innerhalb von 300.000 CPU-Stunden zu rekonstruieren. Da sich die Bildverarbeitung der DSLM-Daten sehr gut parallelisieren lässt, d. h. viele Computer gleichzeitig an verschiedenen Teilaspekten der Auswertung arbeiten können, war es möglich, die Rekonstruktion der über 7 TB großen Rohdaten innerhalb einer Woche auf einem 2.000-Core-Teilcluster des GridKa durchzuführen. Aufgrund dieses erfolgreichen Starts sind nun weitere DSLM-Experimente und Rekonstruktionen am SCC geplant, die zum umfassenden Verständnis der fundamentalen biologischen Prozesse in der frühen embryonalen Entwicklung deutlich beitragen werden.

Philipp J. Keller, Annette Schmidt, Jochen Wittbrodt, Ernst H. K. Stelzer / EMBL Heidelberg, Abteilungen für Zellbiologie und Entwicklungsbiologie, in Zusammenarbeit mit Manfred Alef, Steinbuch Centre for Computing (SCC) und Urban Liebel, Forschungszentrum Karlsruhe, Institut für Toxikologie und Genetik (ITG).

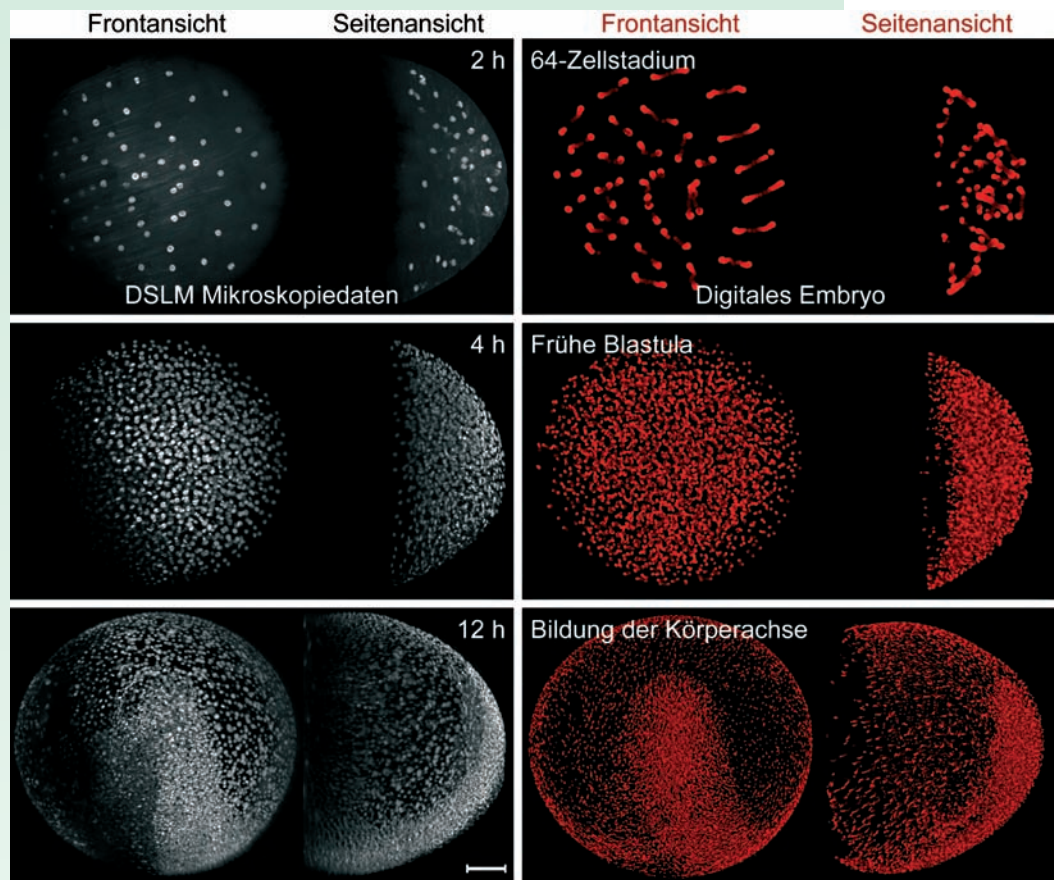


Abbildung 1: Projektionen der dreidimensionalen DSLM-Aufnahmen der Entwicklung eines Zebrafisch-Embryos (links) und Visualisierung des rekonstruierten „Digitalen Embryos“ (rechts).

Die Abbildung zeigt drei Zeitpunkte einer Aufnahme (links), die aus insgesamt 2.000 dreidimensionalen Bildstapeln besteht, die in 60-Sekunden-Intervallen aufgenommen wurden. Für die Fluoreszenzaufnahmen wurden alle Zellkerne des Fischembryos mit fluoreszent anregbaren Proteinen (GFP) markiert, so dass die Positionen aller Zellen bestimmt und in dem „digitalen Embryo“ verfolgt werden können. Die Rekonstruktionen (rechts) der Mikroskopiedaten wurden mit Hilfe eines speziell entwickelten Segmentierungs-Algorithmus errechnet. Der gesamte Datensatz besteht aus 300.000 Bildern (2,4 TB).

Grids mit EGEE III auf der Überholspur

Internationale Forschungsgemeinschaft setzt auf weltweit vernetzte Rechnerressourcen



„Weltkarte des Grid“

Anfang Mai 2008 startete das EU-Projekt „Enabling Grids for E-science (EGEE III)“. Von der Europäischen Kommission finanziert soll EGEE III Wissenschaftlern in Forschung und Industrie den Zugang zu bedeutenden Rechnerressourcen unabhängig von ihrem geografischen Standort bieten. Die Enabling Grids for E-science (EGEE)-Initiative verwaltet und betreibt die größte multidisziplinäre Grid-Infrastruktur der Welt. Das SCC ist Konsortialführer aller deutschen Projektpartner und - wie auch schon in den ersten beiden Projektphasen - im EGEE Project Management Board Sprecher für alle neun deutschen und schweizerischen Partner.

Mit EGEE III soll die bereits existierende Grid-Infrastruktur erweitert und optimiert werden. Mit einer Laufzeit von 24 Monaten, 125 Partnern aus 32 Ländern und einem Gesamtvolumen von 32 Millionen Euro ist EGEE das größte EU-Projekt im Bereich Grid Computing.

SCC „Regional Operation Centre“

Aktuelle Anwendungen stammen aus den Bereichen Hochenergie-Physik, Biowissenschaften, Geowissenschaften,

Astrophysik und Chemie. So unterstützt die Initiative Wissenschaftler am Europäischen Forschungszentrum CERN in Genf bei ihren Experimenten mit dem Teilchenbeschleuniger (Large Hadron Collider, LHC) auf der Suche nach dem Higgs-Teilchen, bietet Rechnerressourcen zur Entwicklung von lebensrettenden Verfahren in der Medizin oder zur Lösung von Energie- oder Welternährungsfragen. „Als ‚Regional Operation Centre‘ trägt das SCC maßgeblich zum Betrieb dieser Grid-Infrastruktur bei“, erklärt Klaus-

Peter Mickel, Technisch-Wissenschaftlicher Direktor des SCC und Sprecher der „German Joint Research Unit“ im EGEE Project Management Board.

250 Standorte in 48 Ländern

Die EGEE-Infrastruktur wird derzeit mehr als 150.000 mal täglich von Wissenschaftlern zur gemeinsamen Arbeit in den verschiedenen wissenschaftlichen Bereichen genutzt. Kofinanziert von der Europäischen Union arbeiten in EGEE III mehr als 120 Organisati-

onen zusammen, um der europäischen und weltweiten Forschungsgemeinschaft betriebssichere und skalierbare Ressourcen zur Verfügung zu stellen. EGEE umfasst 250 Standorte in 48 Ländern mit mehr als 60.000 CPUs und über 20 Petabyte Speicherplatz. Mehr als 8.000 Nutzer weltweit können rund um die Uhr auf diesen Ressourcenpool zugreifen.

Weltweit vereinigte Grid-Struktur angestrebt

Diese Zahlen übersteigen beträchtlich die für das Ende der ersten vier Jahre des EGEE-Programms erwarteten Ziele und zeigen die Begeisterung der Forschungsgemeinschaft für das Projekt und Grid-Lösungen. „Letztendlich strebt EGEE eine vereinigte, interoperable Grid-Infrastruktur an. Daher arbeiten die Beteiligten eng mit anderen europäischen und weltweiten Betreibern von Grid-Projekten zusammen, um bei der Definition von Standards mitzuwirken“, so Dr. Holger Marten, Leiter der Abteilung „Verteilte Systeme und Grid“ am SCC.

Einer der Gründe dafür, EGEE und das Grid ins Leben zu rufen, war die Suche nach dem Higgs-Teilchen - dem heiligen Gral der Elementarteilchenphysik. Die Rechenanforderungen des Large Hadron Collider (LHC) am CERN stellen eine noch nie dagewesene Herausforderung dar. Mehr als 15 Petabyte an Daten werden dort in Zukunft pro Jahr generiert und sollen durch tausende von Wissenschaftlern simultan analysiert werden. Das am SCC angesiedelte Grid Computing Centre Karlsruhe (GridKa) ist eines der elf weltweiten „Tier 1“-Grid-Rechenzentren, das an der Datenanalyse beteiligt ist. Mit dem für September 2008 erwarteten Anlaufen des LHC wird GridKa und die EGEE-Infrastruktur eine weitere, große Herausforderung bewältigen müssen - das Datenvolumen und damit die Menge an zu verarbeitender Information dürften sich vervielfachen.

Rechenleistung in beispiellosen Größenordnungen

Die hohen internationalen Zugriffsraten und die riesige Informationsmenge erfordern Rechenanlagen, die global auf verschiedene Orte verteilt sind. Da mit Hilfe des Grid die Arbeitslast auf alle beteiligten Rechenzentren verteilt wird, können Forscher auf der ganzen Welt gemeinsam an zentralen Fragestellungen arbeiten.

Viele verschiedene Einsatzbereiche sind denkbar. So wurde die EGEE-Infrastruktur dazu genutzt, um unter mehr als 500.000 medikamentenähnlichen Molekülen in nur wenigen Wochen diejenigen herauszusuchen, die am Besten gegen Vogelgrippe wirken könnten. „Dies bedeutet, dass auf der Grundlage des Grid potenzielle Lösungen gefunden werden können, ohne Untersuchungen im Labor durchführen zu müssen. Damit können physikalische Ressourcen und wertvolle Zeit eingespart werden. Im Hinblick auf ein mutierendes Virus kann diese Zeitersparnis Leben retten“, sagt Holger Marten.

Andere Wissenschaftler nutzen das Grid, um die Komplexität von Muskelzellen besser verstehen zu können, die Dynamik schwarzer Löcher zu berechnen, Prozesse in Zellen zu simulieren, Proteinstrukturen zu prognostizieren, die Verschmutzung der Atmosphäre zu studieren oder nach Genen in widerstandsfähigem Weizen zu suchen. EGEE eröffnet Forschern weltweit Rechenleistung in beispiellosen Größenordnungen und macht es ihnen einfach, Daten und Ergebnisse zu teilen.

Die Werkzeuge und Techniken, die in einer Disziplin genutzt werden, können sehr oft wiederverwendet und anderswo eingesetzt werden, entweder von Wissenschaftlern oder aber auch im Business- und Bankenbereich. Hier wird EGEE genutzt, um beispielsweise neue Ölvorkommen zu orten, Marktbedingungen zu simulieren oder Versteuerungsverfahren abzubilden.

Die nächste Konferenz, „EGEE'08“ (www.eu-egee.org/egee08), wird vom 22. bis 26. September 2008 in Istanbul in der Türkei stattfinden. „Dort bietet sich Teilnehmern aus Wissenschaft und Wirtschaft eine einmalige Gelegenheit, EGEE-Communities, angeschlossene Projekte, Entwickler und Entscheidungsträger zu treffen und an der Vision eines nachhaltigen, interoperablen europäischen und weltweiten Grids mitzuwirken“, so Klaus-Peter Mickel.

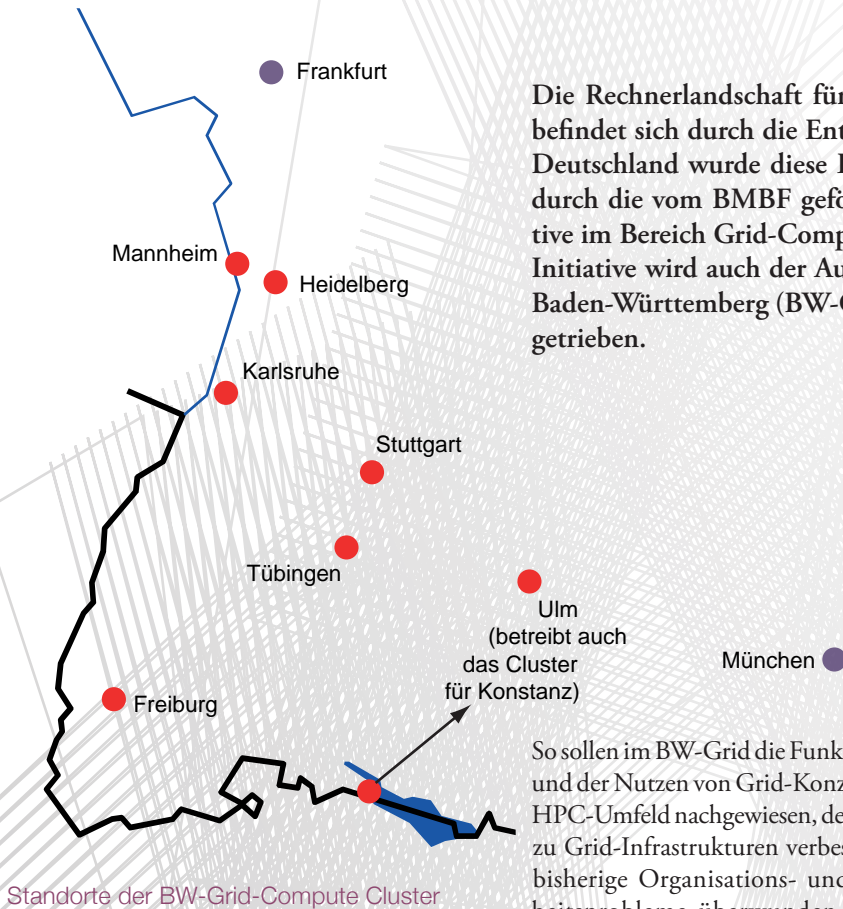
Karin Schäufele

Weitere Informationen

www.eu-egee.org
oder bei Dr. Rüdiger Berlich
German Press Office
press@eu-egee.de
Tel.: 07247/82-5678
E-Mail: Ruediger.Berlich@kit.edu
Fax: 07247/82-49 72

Baden-Württemberg Grid - **bw**GRiD

Regionale Grid-Infrastruktur für Baden-Württemberg



Standorte der BW-Grid-Compute Cluster

Die Rechnerlandschaft für technisch-wissenschaftliche Anwendungen befindet sich durch die Entstehung von Grid-Verbänden im Wandel. In Deutschland wurde diese Entwicklung in den letzten Jahren vor allem durch die vom BMBF geförderte Forschungs- und Entwicklungsinitiative im Bereich Grid-Computing „D-Grid“ geprägt. Im Rahmen dieser Initiative wird auch der Aufbau einer regionalen Grid-Infrastruktur für Baden-Württemberg (BW-Grid) an den Universitäten des Landes vorangetrieben.

So sollen im BW-Grid die Funktionalität und der Nutzen von Grid-Konzepten im HPC-Umfeld nachgewiesen, der Zugang zu Grid-Infrastrukturen verbessert und bisherige Organisations- und Sicherheitsprobleme überwunden werden. Darüber hinaus sollen neue Cluster- und Grid-Anwendungen entwickelt und bislang bestehende Lizenzprobleme gelöst werden. Beteiligt sind die Universitäten Freiburg, Heidelberg, Karlsruhe, Mannheim, Stuttgart, Tübingen und Ulm. Im Rahmen dieses Projekts streben die Universitäten zudem eine weitere Spezialisierung ihrer Rechenzentren an, um die Unterstützung unterschiedlicher Anwendergruppen zu verstärken. Jedes Rechenzentrum wird einen spezifischen Themenbereich als Schwerpunkt haben und die Unterstützung der Nutzer aus diesem Bereich koordinieren. Die Zentren werden auch einen weiteren, möglichst verwandten Schwerpunkt unterstützen, sodass lokale und verteilte Kompetenzen vernetzt werden.

Gefördert wird dieses Vorhaben durch das Bundesministerium für Bildung und Forschung (BMBF), das Ministerium für Wissenschaft, Forschung und Kunst (MWK) des Landes Baden-Württemberg

sowie die Deutsche Forschungsgemeinschaft (DFG).

Das BMBF genehmigte Ende 2007 eine einmalige Hardware-Investition zum Aufbau von Linux-Clustern an den sieben beteiligten Universitäten. Das Land finanziert gemeinsam mit den Universitäten die Mitarbeiter, die sich um das Management der Systeme und die Betreuung der Anwender kümmern. Eine landesübergreifende Speicherlösung bestehend aus lokalen Komponenten an den Cluster-Standorten und einem zentralen System am Steinbuch Centre for Computing (SCC) wird derzeit bei der DFG beantragt.

Leistungsfähige Compute Cluster für die Forschung

Das BW-Grid stellt für die wissenschaftliche Forschung an den Universitäten und Hochschulen in Baden-Württemberg sogenannte Blade-Systeme der Firma IBM zur Verfügung. Diese Systeme sind modular aufgebaut und erlauben den Aufbau großer Rechenkapazitäten bei geringem Platzbedarf. Insgesamt wurden 101 BladeCenter H beschafft. Jedes dieser BladeCenter ist ein äußerst kompaktes

Gehäuse für insgesamt 14 Blades, d. h. 14 individuelle, leistungsfähige Rechner vom Typ IBM BladeServer HS21 XM mit extrem kompakten Ausmaßen und folgender Ausstattung:

- 2 x Intel Xeon Quadcore 2.8Ghz
- 16 GB Hauptspeicher
- 4 x DDR dual port Infiniband-HCA

Diese BladeCenter sind folgendermaßen auf die sieben Standorte des BW-Grid verteilt:

- Jeweils 10 BladeCenter (140-Knoten) an den Standorten Freiburg, Heidelberg, Karlsruhe, Mannheim und Tübingen
- 20 BladeCenter (280 Knoten) am Standort Ulm (im Verbund mit Konstanz)
- 31 BladeCenter (434 Knoten) am Standort Stuttgart

Die InfiniBand-Vernetzung an den einzelnen Standorten erfolgt mit „Switchfabrics“ der Firma Voltaire vom Typ Grid Director ISR 2012 mit der derzeit höchstmöglichen Übertragungsrate von 20 Gbit/s (IB-DDR) bidirektional.

SCC koordiniert landesübergreifende Datenhaltung

Der Clusterverbund des BW-Grid soll durch eine landesübergreifende Speicherlösung ergänzt werden, die aus lokalen Komponenten an den einzelnen Cluster-Standorten und einem zentralen Datenhaltungssystem besteht. Das Gesamtsystem soll als so genannter „Federated Storage“ organisationsübergreifend verwaltet und genutzt werden. In Abstimmung mit den beteiligten Universitäten koordiniert das SCC dieses Teilprojekt.

Verteiltes Speicherkonzept noch Neuland

Natürlich ist ein solches verteiltes Speicherkonzept auch eine große Herausforderung. Zum einen ist die Zuverlässigkeit und Stabilität des globalen Dateisystems extrem wichtig, denn wenn das zentrale Dateisystem nicht funktioniert, ist eine Nutzung der Rechner im BW-Grid nur eingeschränkt möglich. Zum anderen ist das Zusammenwirken lokaler und zentraler Speicherkomponenten an verschiedenen Einrichtungen in der hier vorgeschlagenen Form noch Neuland.

Insbesondere wird eine Schulung der Nutzer im Umgang mit den neuen Systemen und den dahinter stehenden Konzepten erforderlich.

Erhöhte Effizienz durch Bündelung

Ein wesentlicher Vorteil der gemeinsamen Beschaffung einer Speicher-Lösung für alle sieben Standorte des BW-Grid ist die Kostenersparnis durch Bündelung. Zum anderen verringert sich der Gesamtaufwand für Administration und Wartung, da künftig an allen Standorten gleichartige Systeme eingesetzt werden. Zusätzlich wird durch das Zusammenwirken von globalen und lokalen Speicherbereichen mehrfache Datenhaltung vermieden, was die Benutzbarkeit vereinfacht und den Plattenspeicherbedarf verringert. Mit dem Backup- und Archivierungssystem am SCC sind zudem die Voraussetzungen für eine Datensicherung dieser gemeinsamen Datenhaltungslösung gegeben.

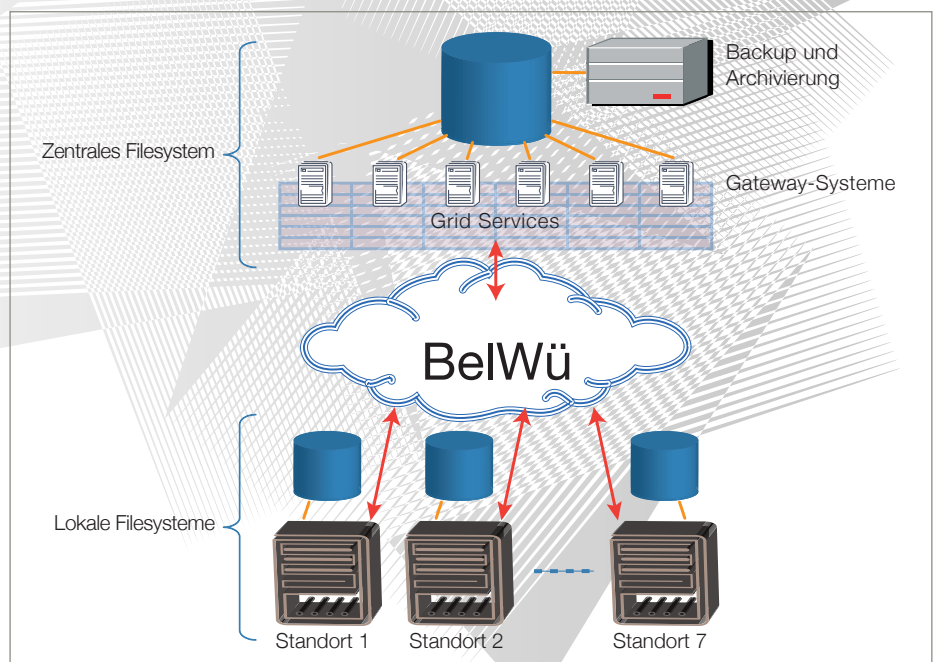
Zugang

Die Infrastruktur steht jedem Wissenschaftler der Baden-Württembergischen Universitäten und Hochschulen offen, der Mitglied der Virtuellen Organisation BW-Grid ist. Darüber hinaus können die Cluster als Teil der Infrastruktur von D-Grid auch von allen anderen „Communities“ verwendet werden. Voraussetzung zur Anmeldung im BW-Grid oder D-Grid ist der Besitz eines

persönlichen Grid-Zertifikats, einer Art elektronischem Ausweis, mit dem sich Benutzer im Grid identifizieren. Innerhalb der Virtuellen Organisation BW-Grid stehen den Benutzern nicht nur die reinen Hardware-Ressourcen zur Verfügung, sondern auch die jeweiligen Experten der Rechenzentren, die die Anwender tatkräftig unterstützen. Um den Support für die unterschiedlichen Anwendungen zu optimieren, streben die Rechenzentren eine Spezialisierung an. So konzentriert sich etwa das Know-How für CAE-Anwendungen, insbesondere Strukturmechanik und Strömungssimulationen auf das Höchstleistungsrechenzentrum Stuttgart (HLRS) und das Steinbuch Centre for Computing (SCC) in Karlsruhe.

Das SCC wird zum kommenden Wintersemester Einführungskurse zur Nutzung des BW-Grid anbieten.

Olaf Schneider



Speicherkonzept für BW-Grid

GridKa School 2008

Applikationen - Middleware - Techniken



Großes Interesse aus aller Welt: Internationale Teilnehmer der GridKa School 07.
Foto: Martin Lober

Mit der GridKa School 2008 veranstaltet das SCC vom 8. bis 12. September 2008 zum sechsten Mal eine internationale Sommerschule für Wissenschaftler und Studierende zu den Themen Grid-Computing und E-Science. Dabei ist das Themenangebot gezielt auf Anfänger, erfahrene Nutzer und Grid-Administratoren ausgerichtet.

Das Thema der diesjährigen Schule lautet „Applikationen, Middleware, Techniken“. Die angebotenen Präsentationen und praktischen Übungen beziehen sich unter anderem auf die Analyse der Daten des Teilchenbeschleunigers (Large Hadron Collider, LHC) am Forschungszentrum CERN in Genf, UNICORE (Uniform Interface to COmputing REsources) und Virtualisierung. Zusätzlich wird ein optionaler Grid

Security Workshop angeboten.

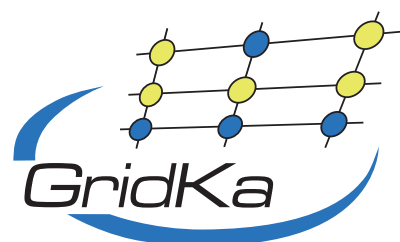
Die Hochenergiephysik-Gemeinde ist eine der wissenschaftlichen Gruppierungen, die Grid Computing intensiv nutzen und voran bringen. Daher wird Professor Rolf-Dieter Heuer, der designierte Generaldirektor des CERN, in einem öffentlichen Abendvortrag über das Großprojekt Large Hadron Collider (LHC) berichten. Keynote-Sprecher ist Fabrizio Gagliardi, Direktor Technical Computing der Microsoft Corporation. Die Firma Microsoft unterstützt die Gridka-School als Hauptsponsor; weitere Sponsoren sind die Firmen SUN und Phi-T.

Die GridKa School kooperiert 2008 u. a. mit dem vom BMBF geförderten D-Grid-Projekt SuGI (Sustainable Grid Infrastructures), der

HGF (Helmholtz-Gemeinschaft Deutscher Forschungszentren)-Allianz „Physik an der Teraskala“ sowie dem EU-Projekt EGEE (Enabling Grids for e-Science).

Karin Schäufele

Weitere Informationen
www.fzk.de/gks08 oder kontaktieren Sie
Dr. Christopher Jung
Tel. 07247/82-8606
E-Mail: gridka-school@fzk.de
Fax: 07247/82-4972



Integrated Site Security for Grids

Projekt zur modellierten Zusammenführung von Grid und Site Security erfolgreich abgeschlossen

Das SCC führte in den letzten zwei Jahren gemeinsam mit dem europäischen Forschungszentrum CERN in der Schweiz und dem Science and Technology Facilities Council (STFC) in Großbritannien das von der EU geförderte Forschungsprojekt „Integrated Site Security for Grids (ISSeG)“ durch, das die Optimierung der Sicherheit von Grid Sites zum Ziel hatte. Die Ergebnisse dieses Projekts werden nun auf der Website <http://www.isseg.eu> präsentiert.

Im ersten Projektabschnitt wurde gemeinsam mit den Projektpartnern eine Strategie entwickelt, mit der sich die Sicherheit von Grid Sites verbessern lässt. An der Umsetzung dieser Strategie und der damit verbundenen technischen Lösungen waren fast alle Abteilungen des SCC beteiligt. Es wurden Arbeiten im Bereich des Identitäts- und Ressourcen-Managements und der „Intrusion-Detection“ im Umfeld von 10Gbps durchgeführt. Weitere Schwerpunkte waren das zentrale Account-Management, die zentrale Patch-Verteilung von Software, die Definition von Firewall-Policies, die Separation sensibler Netzwerke, der Betrieb von Application Gateways sowie der Einsatz von Security Scan und Network Monitoring Tools. Auch der IT-Sicherheits- und der Datenschutzbeauftragte leisteten in Form von Administrator- und User-Schulungen ihren Beitrag zum erfolgreichen Gelingen des Projektes.

Die ISSeG-Website stellt ein Risk-Assessment-Tool sowie verschiedene Sicherheitsempfehlungen und Trainingsmaterialien zur Verfügung, mit denen die Sicherheit einer Einrichtung optimiert werden kann. Obwohl der Projektfokus auf der Verbesserung der Sicherheit von Grid Sites lag, kann fast jede an-

dere Institution, insbesondere im wissenschaftlich-technischen Umfeld, die Ergebnisse nutzen.

Auf der Website (<http://www.isseg.eu>) können einrichtungsspezifische priorisierte Listen von Bedrohungen, eine maßgeschneiderte Auswahl an Sicherheitsempfehlungen sowie technische Beschreibungen zur Umsetzung angefordert werden. Eine allgemeingültige Liste von Top-Bedrohungen und Top-Sicherheitsempfehlungen für Grid Sites können auch direkt ausgewählt werden. Trainingsmaterial für User, Systemadministratoren, Softwareentwickler und Manager sind dort ebenfalls erhältlich. Darüber hinaus können Materialien, die in die Thematik einführen, Sicherheitschecklisten und druckbare Dateien, wie Poster und Flyer, heruntergeladen werden. Diese können beispielsweise zur Erhöhung des Sicherheitsbewusstseins von Mitarbeitern verwendet werden. Weitere zahlreiche EU-Dokumente, die sich mit Sicherheitsstrategien, deren Implementierung und Risiko-Management befassen, stehen zum Download bereit. Darüber hinaus geben diese Dokumente Auskunft über die Gültigkeit und Methodik des ISSeG-Verfahrens. Bei Rückfragen stehen die Autoren dieses Artikels gerne zu Verfügung.

Ursula Epting, Bruno Hoefft, Tobias König

Das SCC stellt sich vor

In dieser Ausgabe: Die Abteilung Integration und Virtualisierung (IVI)



Foto: Privat

Dr. Marcel Kunze promovierte als Elementarteilchenphysiker an der Universität Karlsruhe und habilitierte sich 1996 mit einem Thema der Neuroinformatik an der Universität Bochum. Forschungsaufenthalte führten ihn zu internationalen Einrichtungen wie CERN und Stanford University. Seit 2002 ist er Abteilungsleiter im Forschungszentrum Karlsruhe und bearbeitet am SCC schwerpunktmäßig die Themenbereiche „Integration“ und „Virtualisierung“.

Dr. Martin Nussbaumer arbeitet als Entwicklungsleiter im Projekt KIT Integriertes InformationsManagement (KIM). Als stellvertretender Abteilungsleiter der Abteilung Integration und Virtualisierung (IVI) des SCC setzt er sich dabei neben Fragestellungen zur technischen Lösbarkeit auch mit methodischem Vorgehen auseinander. Im Rahmen seines Lehrauftrags an der Universität hält er die Vorlesung „Web Engineering“.



Foto: Privat

In vielen Anwendungsdomänen stellt die Skalierbarkeit und Verfügbarkeit von IT-Diensten eine große Herausforderung dar: Wie skalieren Anwendungen im Hinblick auf Durchsatz oder Zahl und Art der Klienten, ohne dass Änderungen am Programm oder der Konfiguration erforderlich sind? Wie können Anwendungen möglichst unterbrechungsfrei betrieben werden? Idealerweise sollte sich dabei die zugrunde liegende Infrastruktur zur Gewährleistung von Service-Garantien dynamisch und automatisch an sich ändernde Bedingungen und Anforderungen anpassen. Die Querschnittsabteilung „Integration und Virtualisierung (IVI)“ widmet sich schwerpunktmäßig diesem Themenkreis und bearbeitet so innovative Felder wie „Service-orientierte Architekturen (SOA)“ und „Cloud Computing“.

Die Anwendungslandschaften und Prozesse an der Universität und am Forschungszentrum Karlsruhe sind seit mehreren Jahrzehnten unabhängig voneinander organisch gewachsen und werden nun im Rahmen des KIT zusammengeführt. Die Abteilung IVI entwickelt die dazu nötigen Konzepte: Komplexe Anwendungslandschaften und Prozessketten profitieren von einer umfassenden SOA-Strategie, bei der sich Geschäftsprozesse über virtuelle Infrastrukturen dynamisch an Änderungen anpassen können. Ein Paradebeispiel für den integrativen Ansatz durch SOA ist das Projekt „KIT Integriertes Informationsmanagement (KIM)“, in dem die Abteilung IVI die führende Rolle spielt und Entwicklungen maßgeblich mit vorantreibt.

KIT Integriertes Informationsmanagement – SOA technische und organisatorische Grundlage

Zur Realisierung eines integrierten Informationsmanagements an Forschungseinrichtungen und Hochschulen wie auch in der Industrie setzen sich in zunehmendem Maße Service-orientierte Architekturen als technisches und organisatorisches Lösungskonzept durch. Heterogene, autonome und über den gesamten Campus verteilte Informationssysteme werden durch Dienste gekapselt und einheitlich über standardisierte Schnittstellen und Protokolle zugänglich gemacht. Durch die Integration und Komposition von Daten und Funktionalitäten verschiedenster Systeme entlang übergreifender Geschäftsprozesse können deren Effizienz und Qualität deutlich verbessert und neuartige Mehrwertdienste den verschiedenen Nutzergruppen angeboten werden. Einrichtungsübergreifende Geschäftsprozesse beherbergen allerdings auch inhärent die Gefahren von Informations- und Medienbrüchen, die den Austausch von solch prozessrelevanten Daten über einheitliche Schnittstellen erheblich erschweren.

Das Vorhaben „KIT Integriertes InformationsManagement (KIM)“ strebt eine ganzheitliche Betrachtung sämtlicher einrichtungsübergreifender Prozesse an. Dies soll durch konsequente Modellierung und Analyse, kontinuierliche Verbesserungen und eine übergreifende prozessorientierte IT-Plattform erreicht werden. Angebotene Informationen und Dienstleistungen orientieren sich an den Anforderungen der Nutzer und werden diesen über ein Informations- und Dienstleistungsportal zugänglich gemacht. Die Grundlage dafür stellt die in KIM entwickelte Referenzarchitektur für integrative SOA-Projekte (iSOA) dar, die verschiedene Integra-

tionsniveaus einführt und die jeweiligen Schichten und deren Aufgaben von einander abgrenzt (siehe Abbildung 1). Aufbauend auf der Integrationsschicht „Technische Infrastruktur“, die sich mit Betrieb, Wartung und Erbringung grundlegender Infrastrukturdienste wie Netze, Systeme und Server oder Anwendungs- und Hostsystemen befasst, werden im Folgenden die weiteren Integrationsschichten der iSOA dargestellt. Die grundlegende Idee der iSOA-Referenzarchitektur stellt die Spezifikation einer Schablone für einen Baukasten dar, mit dessen Hilfe Mehrwertdienste aus autonomen wieder verwendbaren Werkstücken zusammengesetzt werden können.

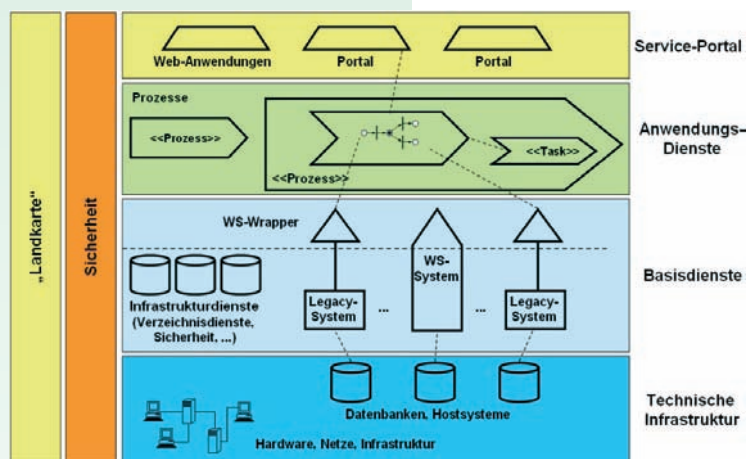


Abbildung 1: Integrierte Service-orientierte Architektur (iSOA) des KIM-Vorhabens.

Basisdienste

Als datenzentrierte, wieder verwendbare Komponenten in Form von Web Services stellen Basisdienste die grundlegenden Bausteine der iSOA dar. Als genormte, plattformunabhängige Zugangspunkte bieten sie sich idealerweise an, um die Integration von Altsystemen (Legacy) beziehungsweise deren Daten in die iSOA zu realisieren. Dazu stellen sie für eine begrenzte, semantisch stark kohäsive Menge an Geschäftsobjekten Operationen zu deren Erstellung, Abfrage und Modifikation bereit. Legt man unterschiedliche Integrationsszenarien zu Grunde, so wird als Ergebnis eine spezifische Untermenge von Basisdiensten erwartet. Das dabei verfolgte Ziel umfasst deren Wiederverwendung in mehreren, von einander unterschiedlichen Anwendungs-

szenarien, beispielsweise zur Unterstützung der Navigation über große Mengen von Informationen, zur situationsbezogenen Darstellung von Daten oder zur anwendungsübergreifenden Kosten- und Lizenzkontrolle.

So gibt es mittlerweile unter anderem Basisdienste für den Zugriff auf Prüfungsergebnisse, Veranstaltungsdaten und Personendaten. In diesen Fällen wird direkt auf die Daten in den Systemen der HIS GmbH zugegriffen. Dazu wurde ein bereits vorhandener, äußerst generischer Web Service so erweitert, dass dadurch ein gesicherter Zugriff auf zuvor spezifizierte Geschäftsobjekte ermöglicht wird. Weiterentwicklungen in Kooperation mit der Firma HIS manifestieren sich derzeit in einer strategischen und technologischen Entwicklungspartnerschaft im Projekt „KIM@HISinOne“. Bereits im Produktivbetrieb wird ein Dienst eingesetzt, der das Gebäudeverwaltungssystem mit Informationen über die Hörsaalbelegungen versorgt und damit eine deutlich effizientere Steuerung der Hörsaalklimate realisiert.

Anwendungsdienste

Basierend auf den wieder verwendbaren Bausteinen der iSOA, den Basisdiensten, verknüpfen Anwendungsdienste diese zu höheren, prozessorientierten Diensten. Diese können sowohl Geschäftsprozesse realisieren, die ausschließlich zwischen mehreren Systemen ablaufen als auch Workflows unter Einbeziehung von Benutzerinteraktion. Im Gegensatz zu Basisdiensten sind Anwendungsdienste stark funktionsorientiert ausgeprägt. Aufgrund ihrer Fundierung auf den Komponenten der Basisdienstschicht und deren loser Kopplung können sie flexibel angepasst werden. Ein Beispiel dafür ist die Erstellung eines „Transcript-of-Records“, der die Studienleistungen eines Studierenden umfasst. Konkret werden hier prüfungs-, Lehrveranstaltungs- und personenbezogene Daten aggregiert und zu einem neuen Dienst zusammengefasst.

Portale

Diese Schicht stellt den unterschiedlichen Benutzergruppen über verschiedene Formen von Benutzerschnittstellen wie Web-Anwendungen und Portalen bis hin zu Büroanwendungen, zentrale und einheitliche Zugangspunkte auf die Geschäftsprozesse und die damit verbundenen Informationen, Anwendungen und Dienste zur Verfügung. Im Rahmen des KIM-Vorhabens entstand jüngst das Studierendenportal, mit dessen Hilfe umfangreiche Studienassistentenfunktionen umgesetzt werden. Das KIT-Studierendenportal strebt die Unterstützung des gesamten Lebenszyklus eines Studiums am KIT von der Immatrikulation bis hin zur Exmatrikulation an, um dadurch die durchgängige Exzellenz der Lehre zu fördern.

Landkarte

Die Landkarte bietet einen Überblick über alle iSOA-Komponenten, ihre Zusammenhänge sowie dedizierte Ansichten der Systemlandschaft zur Unterstützung von Betrieb, Wartung und Evolution. Die zugrunde liegenden Komponenten erstrecken sich über alle Integrationsschichten. Die Grundlage für die Landkarte ist eine Menge von Modellen, welche die Beschreibung der Metainformationen über die

iSOA von der technischen Infrastruktur bis hin zu den Service-Portalen umfassen. Die mit Hilfe der Modelle beschreibbaren Systeminformationen werden während der Laufzeit über Infrastrukturdienste, die grundsätzlich Basisdienste sind, von den Komponenten nach außen zur Verfügung gestellt. Ein Landkartendienst verschaltet mehrere Infrastrukturdienste und konsolidiert die gesammelten Informationen aus diesen Diensten.

Neben den angestrebten Eigenschaften einer iSOA wie die lose Kopplung, die Autonomie von Diensten oder die Plattformunabhängigkeit bildet eine einheitliche Beschreibbarkeit einzelner Ressourcen eine weitere maßgebliche Säule. So benötigen einerseits die verschiedenen beteiligten Zielgruppen Daten über die Fähigkeiten und die Konfiguration einzelner angebotener Dienste. Über geeignete Ansichten müssen relevante Informationen derart aufbereitet sein, dass sie der Situation und den Bedürfnissen der Zielgruppe entsprechen. Andererseits werden Mechanismen benötigt, die ein Fehlverhalten beziehungsweise Ausfälle einzelner Bestandteile frühzeitig erkennen und gegebenenfalls Schritte zur Beseitigung der Betriebsbeeinträchtigung einleiten. Im Idealfall lassen sich das Finden und Verschalten einzelner Services (Entwicklungsaspekt) vom Betrieb und Hosting auf den realen Systemen und Infrastrukturkomponenten (Betriebsaspekt) physisch trennen aber im virtuellen Verbund als ganzheitliche Services anforderungsabhängig anbieten.

Virtuelle Infrastrukturen - Optimierte Services und positive Bilanzen

Ein weiteres innovatives und relevantes Thema, mit dem die Abteilung IVI sich auseinandersetzt, ist die Virtualisierung von IT-Infrastrukturen: Die Virtualisierung gibt einem Anwender oder einer Anwendung stets die gewünschte Sicht auf die Ressourcen-Landschaft, unabhängig davon, wie die tatsächliche reale Implementierung aussieht.

Der wohl durchdachte und strategisch geplante Einsatz der Virtualisierung setzt dabei große Effizienzpotenziale frei. Die Abteilung IVI erarbeitet eine übergreifende Virtualisierungsstrategie, konzipiert und plant die erforderlichen virtuellen Infrastrukturen im Hinblick auf Anwendungen, Speicher und Server. Dazu werden Standards, Verfügbarkeits- und Ausfallsicherheitsszenarien definiert und eingeführt.

Die Aufgaben im Virtualisierungsbereich umfassen die Bereiche Server-, Speicher-, Netzwerk-, Desktop- und Anwendungsvirtualisierung. Durch die Entkopplung der Services von der physischen Infrastruktur entfällt in der virtuellen Umgebung die Notwendigkeit, Wartungsfenster für die Infrastrukturpflege zu vereinbaren: Dienste können während des Betriebs beliebig migriert werden, um beispielsweise auf den Infrastrukturkomponenten Firmware-Upgrades zu fahren. Auch das Tauschen von Komponenten im Rahmen eines Technologiewechsels ist während des laufenden Betriebs möglich. Generell kann so die allgemeine Service-Verfügbarkeit deutlich verbessert werden.

Zusätzlich führt die Mehrfachnutzung von Ressourcen bei Virtualisierung von Services auch zu einer prinzipiell besseren Auslastung und damit zu einer günstigeren Energiebilanz. Als besonders vielversprechendes Beispiel soll an dieser Stelle die Desktopvirtualisierung genannt werden: Die Verwendung von so genannten Thin Clients im Zusammenspiel mit einer Server-Infrastruktur anstelle von Desktop-PCs ist in vielen Bereichen ohne spürbare Komforteinbußen für die Benutzer denkbar. Man spart besonders große Mengen Energie, wenn die benötigte Rechnerleistung dabei zentral auf einer energietechnisch günstigen Blade-Installation vorgehalten wird. Es wird in der Praxis eine Konsolidierung von 20 bis 40 virtuellen Desktops pro Blade-Einschub erreicht. Bei der Einrichtung und Durchführung von Schulungsveranstaltungen ergeben sich dadurch interessante Möglichkeiten.

Besonders im Entwicklungsbereich zahlt sich der Einsatz von Virtualisierungstechnologien aus. Ein Beispiel ist die Datenverarbeitung des Auger-Projekts, das in Argentinien einen großflächigen Detektor für Gamma-Strahlung betreibt. Die Datenerfassung umfasst vor Ort ein komplexes verteiltes System aus 23 Linux- und Windowsrechnern, das vom Institut für Prozessdatenverarbeitung und Elektronik entwickelt wurde und gewartet werden muss. Für die Software-Entwicklung sollte daher in Karlsruhe eine exakte Kopie des Systems in Malargue nachgebaut werden. Die mit IVI gemeinsam erarbeitete Lösung sieht so aus, dass das System inklusive aller Netzwerkkomponenten komplett virtualisiert wurde und nun prinzipiell in einem einzigen Server-Blade betrieben werden kann. Auf diese Weise wurden nicht nur Investitions- und Betriebskosten gesenkt, sondern Test und Entwicklung profitieren zusätzlich von fortgeschrittenen Eigenschaften wie Snapshots von virtuellen Maschinen und dynamischer Ressourcenzuteilung im Falle eines Engpasses.

Cloud Computing - Ressourcen „von der Stange“

Die Virtualisierung ermöglicht darüber hinaus durch die logische Sicht auf die üblicherweise heterogen gewachsenen Ressourcen das transparente und dynamische Management von IT-Kapazität: So verkürzt sich in der virtuellen Welt die Bereitstellungszeit von Infrastrukturkomponenten wie zum Beispiel Servern oder Datenspeichern dramatisch. Aufgaben, die in der physischen Welt unter Umständen Tage oder Wochen in Anspruch nehmen, sind hier in wenigen Minuten erledigt, indem Systeme nach Kundenanforderung schnell aus einem Ressourcen-Pool kommissioniert werden. Es können dabei aus Anwendersicht sogar mehr Ressourcen zugeteilt werden, als physisch verfügbar sind (Over-Commitment). Es muss nur sichergestellt sein, dass zum rechten Zeitpunkt genügend Ressourcen zur Verfügung stehen, um eine Anforderung in einer bestimmten Güte zu befriedigen.

Diese Ressourcen können zum einen in-house bereitgestellt werden; zum anderen entwickelt sich speziell im Virtualisierungsumfeld zurzeit ein umfangreicher Service-Markt, sodass in der nahen Zukunft Ressourcen auch nach Bedarf quasi „von der Stange“ von einem Service Provider dazu gebucht werden können. Im Rahmen des so genannten

Cloud Computing werden zu diesem Zweck weltweit verteilte Services im Internet dynamisch über virtuelle Maschinen angeboten. Das SCC richtet derzeit zusammen mit namhaften internationalen Partnern wie HP, Intel und Yahoo! als eines von sechs Exzellenz-Zentren eine globale Testumgebung ein. Die Zusammenschaltung mit den anderen Testumgebungen wird noch in diesem Jahr erste weltweite Forschungsvorhaben zum Cloud Computing ermöglichen. Die Abteilung IVI untersucht und entwickelt in mehreren Projekten, Doktor- und Diplomarbeiten die Technologien, die für das Cloud Computing bereitgestellt werden müssen.

Dynamische IT-Umgebungen durch Integration und Virtualisierung

Die Kombination der beiden Bereiche „Integration und Virtualisierung“ in einer Abteilung ist deutschlandweit einmalig und eröffnet vielfältige neue Möglichkeiten zur Steigerung der Service-Qualität und zur Entwicklung innovativer Services: Die Verknüpfung einer Service-orientierten Architektur mit virtuellen Infrastrukturen führt zu dynamischen IT-Umgebungen, in denen Dienste sich automatisch an sich ändernde Anforderungen anpassen und im laufenden Betrieb zwischen verschiedenen Ressourcen-Providern migriert werden können. Im Bereich des Service Management ergeben sich daraus fortgeschrittene Methoden zur Automatisierung und damit einhergehend zu einer Steigerung der Effizienz.

Die Forschung und Entwicklung in den Bereichen SOA, Virtualisierung und Cloud Computing ist nicht nur spannend, sondern auch äußerst industrierelevant und bietet viele interessante Aspekte für die Ausbildung von Diplomanden und Doktoranden. Insofern ergeben sich daraus gute Chancen zur Bereicherung von Lehrveranstaltungen mit diesen innovativen Themen.

Die Abteilung IVI bietet kompetente Beratung und Unterstützung nicht nur im Bereich der dienstorientierten Anwendungsentwicklung an, sondern ebenso bei der Transformation physischer und virtueller Ressourcen sowie bei Fragen des Hosting von Services.

Nationale und internationale Projekte

Die Abteilung IVI engagiert sich zur Bearbeitung der Fragestellungen in einer Vielzahl nationaler und internationaler Projekte, die in den kommenden Ausgaben der SCC-News ausführlich vorgestellt werden:

D-Grid

Zum sicheren und hoch verfügbaren Betrieb der Kerndienste wurde für das D-Grid Infrastrukturprojekt eine virtuelle Infrastruktur konzipiert, die sich über zwei Standorte erstreckt und dadurch fehlertolerant betrieben werden kann. Die hier gesammelten Erfahrungen können direkt in den Regelbetrieb der SCC-Services einfließen.

D-MON

Die Vision des D-Grid Projekts D-MON ist es, eine verteilte Monitoring-Architektur über mehrere darunter liegende, heterogene Systeme zu realisieren und dabei zu berücksichtigen, dass es mehrere Ressourcen-Provider und Virtuelle Organisationen gibt. Dabei werden die verschiedenen Komponenten für Ressourcen- und Dienstmonitoring horizontal integriert und in einem Testbett betrieben. Die Eignung industrierelevanter Protokolle wird untersucht.

Int.EU.Grid

Das Int.EU.Grid-Projekt hat zum Ziel, Infrastrukturen in einem verteilten System spontan für interaktive parallele Anwendungen nutzbar zu machen. Die Virtualisierung ermöglicht hier die dynamische Erfüllung der Anforderungen.

EUFORIA

Das EUFORIA-Projekt will Anwendungen aus der Fusionsforschung zur Nutzung dynamischer verteilter Services ertüchtigen.

g-Eclipse

Im g-Eclipse-Projekt werden Komponenten verschiedener Middleware-Architekturen auf der Basis der Eclipse-Entwicklungsumgebung in einer gemeinsamen Benutzeroberfläche integriert. Zielgruppe sind Anwender, Entwickler und Service-Betreiber. Eine aktuelle Erweiterung ermöglicht den Zugang zu Ressourcen aus der Amazon Elastic Compute Cloud.

Dr. Marcel Kunze, Dr. Martin Nussbaumer



Florian Allarding ist wissenschaftlicher Mitarbeiter am SCC. Neben seiner Arbeit im Projekt KIT Integriertes InformationsManagement (KIM) ist er in der Forschungsgruppe „Effiziente Algorithmen“ am Institut AIFB engagiert.



Dr. Ariel Garcia ist in zwei Projekte involviert: In D-Grid leistet er die gLite Middleware Unterstützung und in g-Eclipse ist er Release Engineer. Er ist ein Experte für die Virtualisierung von Ressourcen mit Xen.



Christian Baun arbeitet als Doktorand am Thema „Aufbau und Evaluation von virtualisierten Clustern unter besonderer Berücksichtigung der Eingabe/Ausgabe“. Ein Schwerpunkt ist die Arbeit mit verteilten Dateisystemen.



Patrick Freudenstein ist Doktorand am SCC und arbeitet im Projekt KIT Integriertes InformationsManagement (KIM). Im Rahmen seiner Forschung beschäftigt er sich mit der modellgetriebenen Entwicklung workflow-basierter Web-Anwendungen unter besonderer Berücksichtigung von Stakeholder-Aspekten.



Dr. Rüdiger Berlich betreut die Ausgründung „Gemfony scientific“ aus dem Forschungszentrum Karlsruhe. Deren Ziel besteht darin, verteilte Algorithmen im Bereich der parametrischen Optimierung zu entwickeln und zu kommerzialisieren. Darüber hinaus ist er für das Press-Office des EU-Projekts EGEE-III in Deutschland und der Schweiz verantwortlich.



Ahmad Hammad arbeitet im EU-Projekt EUFORIA an der Adaptierung, Optimierung und Gridifizierung wichtiger Applikationen für die Fusion Community sowie an der Bereitstellung virtualisierter Ressourcen.



Jan Buck ist wissenschaftlicher Mitarbeiter am SCC und arbeitet im Projekt KIT Integriertes InformationsManagement (KIM). Sein zweiter Schwerpunkt liegt auf der Mitarbeit in der Forschungsgruppe „IT-Management and Web Engineering“.



Marcus Hardt ist in den EU-Projekten Int. eu.grid und EUFORIA verantwortlich für die Integration von Fusionsanwendungen mit verteilten virtualisierten Ressourcen. Er arbeitet auch an der Integration von Matlab sowie an einem verteilten Algorithmus für die medizinische Bildverarbeitung. Er ist ein Experte für die Virtualisierung von Ressourcen mit Xen.



Volker Büge arbeitet als Doktorand am Thema “Measurement of the Z Boson Production with Associated Jets using virtualised Grid Resources”. Ein weiterer Schwerpunkt seiner Arbeit ist der Einsatz von Virtualisierungstechniken zur Konsolidierung der Grid-Services am WLCG Tier3-Zentrum des Instituts für Experimentelle Kernphysik sowie die Entwicklung eines auf virtuellen Maschinen basierenden Batch-Systems. Darüber hinaus arbeitet er an der Entwicklung einer Jet Kalibration basierend auf Z Boson-Ereignissen für das CMS-Experiment am LHC.



Stefan Klett übernimmt Entwicklungs- und Dokumentationsaufgaben im Zusammenhang mit der Ausgründung „Gemfony Scientific“.

Mitarbeiter der Abteilung IVI

Fotos: Privat



Stefan Link ist akademischer Mitarbeiter und Doktorand am SCC. Neben seiner Mitarbeit im Projekt KIT Integriertes InformationsManagement (KIM) liegt ein zweiter Schwerpunkt auf der Forschung im Bereich der modellgetriebenen Anwendungsentwicklung, fokussiert auf die Mensch-Maschine-Interaktion.



Dr. Mathias Stümpert koordiniert das europäische F&E-Projekt g-Eclipse. Dieses beschäftigt sich mit der Entwicklung eines generischen Frameworks für den benutzerfreundlichen Zugriff auf bestehende Grid-Infrastrukturen.



Frederic Majer ist wissenschaftlicher Mitarbeiter und Doktorand am SCC. Neben seiner Tätigkeit im Projekt KIT Integriertes InformationsManagement (KIM) liegt der Schwerpunkt seiner Forschung auf der Unterstützung des Betriebs servicebasierter Systeme unter besonderer Berücksichtigung des inhärenten Beziehungs- und Abhängigkeitsgeflechts.



Dr. Jie Tao ist als Software-Entwicklerin im g-Eclipse-Projekt beschäftigt. Sie ist auch am Aufbau einer aus virtuellen Maschinen bestehenden Umgebung für das Cloud Computing beteiligt.



Dr. Jens Milke befasst sich konzeptionell mit Virtualisierungstechniken zur Steigerung der Effizienz beim Betrieb von Ressourcen. Dazu gehört auch die Entwicklung von Ressource-Pools und Managementverfahren zur Kapazitätsplanung.



Lizhe Wang arbeitet als Doktorand an der Entwicklung von verteilten virtuellen Systemen. Dies beinhaltet die Untersuchung von virtuellen Datenspeichern, Netzwerken und Applikationen für das Cloud Computing.



Oliver Oberst ist Doktorand am SCC sowie am Institut für Experimentelle Kernphysik an der Universität Karlsruhe und beschäftigt sich aktuell mit der dynamischen Virtualisierung von Rechnerknoten des Instituts-Clusters der Universität Karlsruhe



Dr. Gevorg Poghosyan integriert Monitoringssysteme in verteilte Systeme. Darüber hinaus unterstützt er Wissenschaftler im EUFORIA-Projekt bei der Anpassung ihrer Codes an das Grid.

Mitarbeiter der Abteilung IV

Fotos: Privat

Netzwerksimulator optimiert Simulation drahtloser Netzwerke

Der Netzwerksimulator NS-2 ist der in Forschung und Wissenschaft meist genutzte und am weitesten verbreitete Simulator für drahtgebundene und drahtlose Netzwerke. In der Forschungsgruppe von Prof. Hartenstein am SCC wurde in Zusammenarbeit mit Mercedes-Benz Research & Development North America eine im Detaillierungsgrad verbesserte und strukturell vollständig überarbeitete Modellierung und Implementierung für drahtlose Netzwerke gemäß dem Standard IEEE 802.11 entwickelt, die nun zum Standardumfang der aktuellen Version 2.33 des Simulators gehört und unter der Adresse <http://hnsam.isi.edu/hnsam> frei verfügbar ist.

Die Implementierung verfolgt einen modularen Ansatz, in dem die unterschiedlichen Funktionalitäten von IEEE 802.11 klar in unterschiedliche Module aufgeteilt sind. So enthält die Medienzugriffsschicht unter anderem Module für Send- und Empfangskoordination, die Umsetzung der Backoff-Funktionalität beim Medienzugriff, die Verwaltung des tatsächlichen und virtuellen Kanalzustandes sowie die eigentlichen Send- und Empfangsvorgänge.

Auf physikalischer Schicht wurde insbesondere auf eine detaillierte und realitätsnähere Umsetzung geachtet, die es erlaubt, die Send- und Empfangscharakteristika drahtloser Komponenten detailliert zu simulieren und zu analysieren. Dies umfasst insbesondere die Unterstützung von kumuliertem Rauschen sowie der Capturing-Funktionalität.

Im aktuellen Stadium wird insbesondere direkt die Simulation von IEEE 802.11a-Netzwerken unterstützt, die die dezentrale Koordinierungsfunktion (decentralized coordination function, DCF) als Medienzugriffsverfahren nutzen. Der modulare Aufbau erlaubt jedoch die einfache Erweiterbarkeit des Simulators. So kann zum einen die Simulation weiterer IEEE 802.11-Standards ermöglicht werden. Zum anderen kann zusätzlich zu simulierende Funktionalität integriert werden, beispielsweise in den Bereichen Netzwerkmanagement und -sicherheit.

Felix Schmidt-Eisenlohr

Weitere Informationen

http://dsn.tm.uni-karlsruhe.de/Overhaul_NS-2.php.

SCC betreibt Intrusion Prevention-System mit 10 GBit/s Bandbreite

Neueste MacAfee-Plattform erstmals im Produktionsnetz

Auf dem Campus Süd des KIT wurde im April 2008 die Internetverbindung von 1Gbit/s auf 10Gbit/s ausgebaut. Damit wurde die Beschaffung eines neuen Intrusion Prevention-Systems (IPS) notwendig, da das bisher eingesetzte nicht die entsprechenden Datendurchsatz-Kapazitäten sowie die für 10 GBit/s notwendigen optischen Verbindungen bereitstellen konnte. Auf dem Campus Nord war die Internetanbindung mit einer Bandbreite von 10Gbit/s bereits realisiert, jedoch wurde bisher noch kein IPS für diese Anbindung betrieben. Um ein einheitliches Sicherheitsniveau am KIT zu erreichen, initiierte das SCC daher die gemeinsame Beschaffung eines neuen Intrusion Prevention-Systems.

Die Beschaffung war für das SCC eine Herausforderung, da es keine definierte Vorgehensweise für ein solch gemeinsames Vorhaben gab. Die Einkaufsabteilungen von Campus Nord, Süd und das SCC mussten in kürzester Zeit ein neues Procedere entwickeln. Durch die hervorragende Zusammenarbeit war es jedoch möglich, die IPS-Plattform kurzfristig zu beschaffen, so dass die Erweiterung der Anbindung im geplanten Zeitraum umgesetzt werden konnte.

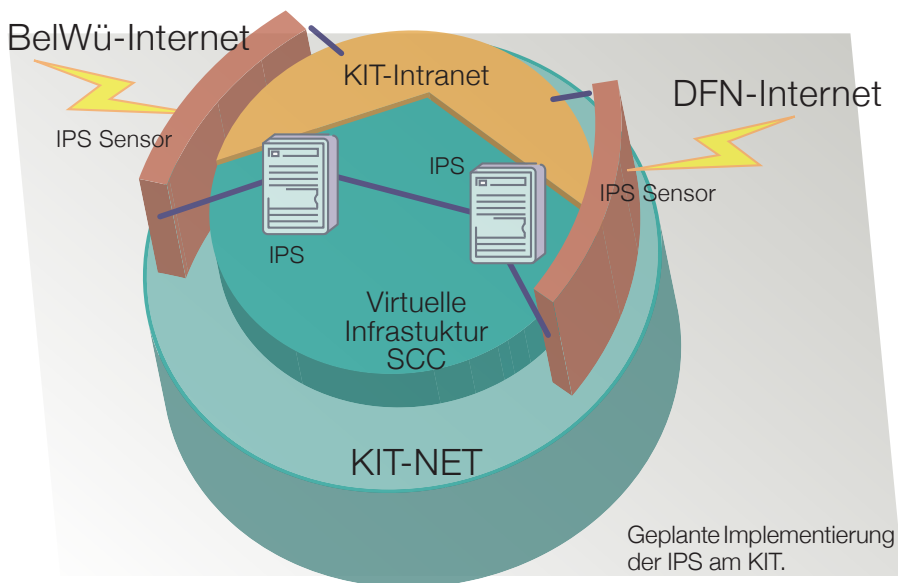
Bei dem neuen System handelt es sich um die Network Security-Plattform M-8000 (früher McAfee Intrushield) der Firma McAfee. Diese Plattform stellt einen Datendurchsatz von 10Gbit/s für die Filterung von Datenströmen bereit. Anders als die älteren Modelle bietet sie mehr Platzierungsmöglichkeiten im Netzwerk: Insgesamt stehen 12 Anschlüsse für Bandbreiten bis 10Gbit/s und 16 Anschlüsse für Bandbreiten bis 1Gbit/s zur Verfügung. Das KIT ist nicht nur der erste Kunde in Europa, sondern auch die

erste Einrichtung, die diese Plattform im Produktionsnetz einsetzt - das SCC nimmt in diesem Bereich eine Vorreiterrolle ein.

Das bisher am Campus Süd eingesetzte IPS besteht aus einer einzelnen Hardwareplattform, die als Sensor dient. Aufgrund der höheren Anforderungen im Datendurchsatz und der gestiegenen Anzahl der Anschlüsse besteht das neue IPS aus mehreren Teilsensoren. Die einzelnen Sensoren

Verwendung der virtuellen Infrastruktur auf Basis von VMWare ESX des SCC notwendig, die Abbildung zeigt die geplante Realisierung der IPS-Plattform am KIT.

Um bei Wartungsarbeiten an den Sensoren keine Unterbrechung des Netzes auszulösen, werden optische Umschalter eingesetzt. Diese optischen Umschalter erlauben das Einschleifen der IPS-Sensoren in die Anbindungen des KIT-NET in einem



sind durch optische Verbindungen mit einer Gesamtbandbreite von 20Gbit/s miteinander verbunden. Diese verfügen jeweils über redundante Stromverbindungen und zusätzliche dedizierte serielle Verbindungen.

Die Managementsysteme für dieses IPS werden aufgrund der ersten Realisierung auf dem Campus Süd redundant am SCC Süd betrieben. Mit der Inbetriebnahme der Plattform auf dem Campus Nord werden die Managementsysteme sowohl auf dem Campus Nord als auch Süd redundant aufgebaut. Die veränderten Anforderungen durch mehrere Sensorplattformen machen die Ablösung der bisherigen Hardware und die

Zeitraum von wenigen Mikrosekunden, so dass dadurch kein nennenswerter Paketverlust auftritt.

Die neue Plattform, die seit Mitte April in Betrieb ist, ersetzt die bisher auf dem Campus Süd betriebene Lösung. Auf dem Campus Nord steht die Inbetriebnahme kurz bevor. Eine zeitgleiche Aufnahme des Betriebs konnte nicht durchgeführt werden, da die IPS-Plattform auf dem Campus Nord zunächst ins Verfahrensregister aufgenommen werden musste; die dafür notwendige Begutachtung des Systems durch den Datenschutzbeauftragten ist bereits erfolgt.

Adrian Wiedemann

SCC auf der Internationalen Supercomputing Conference „ISC 2008“ in Dresden

KIT belegt vier Plätze in der „TOP500“

Die jährlich stattfindende Konferenz zählt weltweit zu den renommiertesten Veranstaltungen im Bereich „High Performance Computing“ (HPC) und ermöglicht Spitzenwissenschaftlern aus Forschung, Industrie und Wirtschaft, ihre Erfahrungen und Visionen auszutauschen. Hochkarätige Fachleute aus anerkannten Supercomputing-Zentren wie auch führende Hard- und Software-Anbieter präsentieren dort modernste Anwendungen, Architekturen und Trends.

Auch das SCC war in diesem Jahr wieder mit einem Stand vertreten und präsentierte innovative Projekte und Infrastrukturlösungen aus dem HPC-Bereich. In der renommierten Liste der weltweit leistungsfähigsten Rechnersysteme (www.top500.org) konnte das KIT vier Lösungen etablieren: Auf Platz 265 den Institutscluster der Firma Transtec (<http://www.rz.uni-karlsruhe.de/ssck/ic.php>), auf Platz 297 den Landeshöchstleistungsrechner HP XC4000 (<http://www.rz.uni-karlsruhe.de/ssck/hpxc4000.php>), auf Platz 404 die am SCC installierten 10 BladeCenter (IBM) des Baden-Württemberg Grid (www.bwGrid.de) und auf Platz 431 die x3550-Lösung im CampusGrid (www.CampusGrid.de) und GridKa (www.GridKa.de).

Frank Schmitz

Neues Buch zum Thema Netzwerk- und IT-Sicherheitsmanagement

Die Autoren Jochen Dinger und Hannes Hartenstein beleuchten Bezug zwischen theoretischen und betrieblichen Aspekten

Dem Netz kommt heutzutage eine fundamentale Rolle zu, da es einzelne Systeme verbindet und somit die Erbringung von verteilten kollaborativen IT-Diensten erlaubt. Durch geeignete Verfahren muss sichergestellt werden, dass die Netze verfügbar, leistungsfähig und sicher sind. Neben den technischen Aspekten sind dabei die strategischen, planerischen, organisatorischen und steuernden Maßnahmen des Managements von besonderer Bedeutung. Aus betrieblicher Sicht kommt den beiden Themengebieten Netzwerk- und IT-Sicherheitsmanagement daher eine elementare Rolle zu.

Prof. Dr. Hannes Hartenstein, Wissenschaftler der Forschungsgruppe „Dezentrale Systeme und Netzdienste“ sowie Mitarbeiter des SCC haben zu diesem Themenkomplex ein Buch mit dem Titel „Netzwerk- und IT-Sicherheitsmanagement: Eine Einführung“ verfasst. Dabei wird insbesondere der Bezug zwischen theoretischen Gesichtspunkten und den betrieblichen Aspekten aus der Praxis des SCC aufgezeigt.

„The network is the computer“

Das Motto der Firma SUN drückt kurz und prägnant eine seit mehreren Jahrzehnten anhaltende Entwicklung aus, die durch das Internet für jedermann erfahrbar wurde und heute durch Entwicklungen wie etwa dem ubiquitären Kommunizieren zum „Internet der Dinge“ oder zu Grid-Strukturen weiterhin fortschreitet. Dem Netz kommt demnach eine fundamentale Rolle zu, da es einzelne Systeme verbindet und somit erst die Erbringung von komplexen kollaborativen und typischerweise geschäftskritischen IT-Diensten erlaubt. Daher muss durch geeignete Techniken sichergestellt werden, dass die Netze verfügbar, leistungsfähig und sicher sind. Darüber hinaus resultieren aus der gestiegenen Bedeutung der Informationstechnik auch mehr und mehr rechtliche Implikationen, sodass die so genannte IT-Compliance auch auf dem Gebiet des Netz- und IT-Sicherheitsmanagement eine zunehmend größere Rolle einnimmt.

Praxisnahe Bezüge hilfreich in Lehrveranstaltungen

Neben theoretischen Konzepten sind in den beiden Themenbereichen sehr oft empfohlene Vorgehensweisen (Best Practices) und spezifische Einzellösungen anzutreffen. Im Rahmen von Lehrveranstaltungen ist es daher eine große Herausforderung, Studierenden den Zusammenhang zwischen Theorie und Praxis näher zu bringen.

Das SCC bietet einen großen Fundus, um die betrieblichen Aspekte den Studierenden im Bereich Netzwerkmanagement zu verdeutlichen. So sei an dieser Stelle beispielhaft die selbstentwickelte Netzwerkmanagementplattform am SCC Süd erwähnt. Mittels dieser integrierten Plattform werden unter anderem ca. 550 aktive Netzkomponenten wie Router und Switches mit insgesamt ca. 15.000 Ports und mehr als 650 virtuelle LANs (VLANs) gemanagt.

Auch im Bereich IT-Sicherheitsmanagement finden am SCC verschiedene Maßnahmen Anwendung, wie zum Beispiel ein Policy-basiertes Managementverfahren für die Einrichtung von Firewall-Regeln auf den zentralen Routern oder der Betrieb von Honeypot-Systemen. Zudem widmen sich mehrere Mitarbeiter im Projekt KIM-IDM dem Themenkomplex Identitäts- und Zugangsmanagement.

Im Rahmen von Vorlesungen und Seminaren ist das SCC bestrebt, den Studierenden die Erfahrungen und Konzepte aus den Themengebieten Netzwerk- und IT-Sicherheitsmanagement näher zu

bringen. Dabei wird in Lehrveranstaltungen vor allem die teilweise doch sehr abstrakte Theorie durch reale Fallbeispiele aus dem Betrieb des SCC angereichert und somit versucht den Studierenden ein tieferes Verständnis zu vermitteln.

Vorlesungen zum Netzwerkmanagement haben an der Universität Karlsruhe eine lange Tradition; hier haben unter anderem Prof. Dr. Jochen Seitz, jetzt Technische Universität Ilmenau, und Prof. Dr. Martin Gaedke, jetzt Technische Universität Chemnitz, Grundlagen geschaffen. Im Jahr 2004 wurde die Vorlesung „Netzwerkmanagement“ von Prof. Hartenstein übernommen und um den Themenkomplex IT-Sicherheitsmanagement erweitert. Durch die Doppelfunktion von Prof. Hartenstein als Professor des Instituts für Telematik und Direktor des SCC wurde auch die Möglichkeit eröffnet, Fallbeispiele aus dem Alltag des SCC in die Vorlesung zu integrieren, sodass in den letzten Jahren Mitarbeiter des SCC die Vorlesung durch praxisnahe Präsentationen regelmäßig bereicherten. Die Studierenden erwähnten diese regelmäßig als sehr hilfreich. Im Wintersemester 2007/08 wurde die Vorlesung außerdem durch eine begleitende Übung ergänzt, sodass mehr Raum für die praxisnahen Teile zur Verfügung steht.

„Netzwerk- und IT-Sicherheitsmanagement“ beste Wahlvorlesung

Im Rahmen der Akademischen Jahresfeier der Fakultät für Informatik im Juli 2008 wurden die Vorlesung „Netzwerk- und IT-Sicherheitsmanagement“ sowie die zugehörige Übung als bestbeurteilte Wahlvorlesung sowie bestbeurteilte

Übung zu einer Wahlvorlesung ausgezeichnet. Die Beurteilung beruht dabei auf der halbjährlichen Lehrevaluation, die in Form einer Studierendenbefragung stattfindet.

Die Vorlesung wurde von den Dozenten Prof. Hartenstein und Jochen Dinger gehalten. Außerdem trugen weitere Mitarbeiter der Forschungsgruppe von Prof. Hartenstein sowie die beiden SCC-Mitarbeiter Adrian Wiedemann und Matthias Müller durch praxisnahe Einblicke zum Gelingen der Lehrveranstaltungen bei.

„Netzwerk- und IT-Sicherheitsmanagement: Eine Einführung“

Das Buch mit dem Titel „Netzwerk- und IT-Sicherheitsmanagement: Eine Einführung“ ist im Februar diesen Jahres im Universitätsverlag Karlsruhe unter der ISBN 978-3-86644-209-2 erschienen. Es wurde von den beiden Hauptautoren Jochen Dinger und Hannes Hartenstein sowie Mitautoren aus der Forschungsgruppe DSN und dem SCC geschrieben, da sich in der Vergangenheit gezeigt hat, dass viele existierende Werke nur Teilaspekte behandeln und teilweise auch nicht mehr aktuell sind. Zudem ließ die bestehende Literatur den oben erwähnten Praxisbezug sehr oft vermissen.

Das Buch führt den Leser in die beiden Themengebiete Netzwerk- und IT-Sicherheitsmanagement ein und gibt einen Überblick über die relevanten Themenbereiche. Dabei verdeutlichen die Autoren die Inhalte durch Verknüpfung von Praxis und Theorie. Die angeführten Fallbeispiele nehmen insbesondere auf Erfahrungen und Maßnahmen des SCC Bezug. Ferner werden die Grundlagen aus historischer Sicht betrachtet und es wird auf aktuelle Arbeiten verwiesen. Die Autoren geben einerseits einen Überblick und eröffnen dem Leser andererseits durch die Referenzierung ergänzender Literatur die Möglichkeit zur Vertiefung.

Nach einem einleitenden Kapitel widmet sich der erste Teil des Buches dem Netzwerkmanagement. Hier werden zunächst allgemein sowie anhand von Fallbeispielen die Aufgaben des Netzmanagements

verdeutlicht. Danach wird abstrakt eine Managementarchitektur betrachtet, bevor konkrete Standards und Verfahren wie SNMP, RMON und NetFlow vorgestellt werden. Nach einem Überblick über die Verwaltung und Vergabe von IP-Nummernräumen und Domain-



Namen werden Managementwerkzeuge und -plattformen anhand von Beispielen des SCC vorgestellt. Zum Abschluss des Teils „Netzmanagement“ folgt eine Darstellung aktueller Entwicklungen und Trends.

Das IT-Sicherheitsmanagement bildet den zweiten Teil des Buches. Hierzu wird der IT-Sicherheitsprozess anhand des Grundschriftbuches des Bundesamtes für Sicherheit in der Informationstechnik vorgestellt. Zur effektiven Umsetzung des IT-Sicherheitsmanagements ist ein Identitäts- und Zugangsmanagement nötig, das in zwei Kapiteln betrachtet wird. Im Anschluss werden Sicherheitspatch-Management, Firewalls und Vorfallsbehandlung dargestellt. Die gestiegene Relevanz der IT spiegelt sich auch immer mehr in gesetzlichen Regelungen wieder, insofern muss eine Einhaltung im Sinne der IT-Compliance sichergestellt werden. Daher erfolgt abschließend eine Darlegung rechtlicher Aspekte durch den Mitautor Dr. iur. Oliver Raabe, der dem Zentrum für angewandte Rechtswissenschaft (ZAR) angehört.

Prof. Dr. Hannes Hartenstein wurde im Jahr 2003 auf die Professur „Dezentrale Systeme und Netzdienste“ an die Universität Karlsruhe (TH) berufen; die Professur beinhaltet die Position des stellvertretenden Leiters des Universitätsrechenzentrums. Am SCC vertritt Prof. Hartenstein im Direktorium die Bereiche IT-Sicherheit, Servicemanagement und innovative Netztechnologien. Mit dem Projekt KIM-IDM treibt Prof. Hartenstein die Einführung eines integrierten Identitätsmanagements am KIT voran. Er ist Co-Autor zahlreicher Veröffentlichungen auf den Gebieten der mobilen Netze, der virtuellen Netze und des IT-Managements.

Dipl.-Inform. Jochen Dinger ist seit 2003 als Wissenschaftlicher Mitarbeiter am Institut für Telematik der Universität Karlsruhe, Forschungsgruppe Prof. Hartenstein, tätig. Seine Forschungsschwerpunkte im Rahmen des BMBF-Projektes SESAM sowie des vom Bundesamt für Sicherheit in der Informationstechnik (BSI) beauftragten Projektes KAI sind sicherheitstechnische und telekommunikationsrechtliche Aspekte in selbstorganisierenden Netzen und Systemen. Daneben bilden Lehrveranstaltungen im Bereich Netzwerk- und IT-Sicherheitsmanagement einen weiteren Tätigkeitsschwerpunkt.

Weitere Informationen zu den beiden Themengebieten und zum Buch „Netzwerk- und IT-Sicherheitsmanagement: Eine Einführung“ finden sich auf den Webseiten des SCC sowie unter <http://dsn.tm.uni-karlsruhe.de/netman.php>.

(red)

Neuer „Masterkurs Parallele und Verteilte Systeme“

Die Autoren Gunther Bengel, Christian Baun, Marcel Kunze und Karl-Uwe Stucky haben einen neuen „Masterkurs Parallele und Verteilte Systeme“ verfasst, der 2008 im Vieweg Verlag erschienen ist und sich mit Grundlagen und Programmierung von Multicoreprozessoren, Multiprozessoren, Cluster und Grid-Computing beschäftigt.

Das Buch bietet eine systematische Darstellung des Stands der Technik und der aktuellen Entwicklungen auf dem Gebiet des Parallelen und Verteilten Rechnens. Es stellt alle relevanten Hardwarearchitekturen für Multiprozessoren und Multicoreprozessoren sowie ihre Betriebssysteme bis hin zum Google-Cluster vor. Das parallele Programmieren bildet einen Schwerpunkt des Werkes. Dazu gehören Client-Server-Modelle und Service-orientierte Architekturen sowie Programmiermodelle für unterschiedliche Speicherarchitekturen. Eine ausführliche Erläuterung von Leistungsmaßen, Parallelisierungstechniken und verteilten Algorithmen zeigt dem Programmierer die Möglichkeiten aber auch die Grenzen der Verteilung auf. Methoden der statischen und der dynamischen Rechenlastverteilung sind ebenso enthalten wie moderne Virtualisierungstechniken. Die abschließenden Kapitel über Cluster und Grid-Computing geben Einblick in die aktuellsten Themen des Gebiets und einen Ausblick auf die zukünftigen Entwicklungen.

Prof. Gunther Bengel lehrt die Gebiete Verteilte Systeme, Parallele Systeme, Cluster- und Grid-Computing und Betriebssysteme und ist Leiter des Instituts für Betriebssysteme an der Hochschule Mannheim.

MSc Christian Baun arbeitet als Wissenschaftlicher Mitarbeiter am Steinbuch Centre for Computing (SCC) im Bereich Grid-Computing.

Dr. Marcel Kunze hat sich an der Ruhr-Universität Bochum habilitiert, forscht im Bereich der parallelen und verteilten Systeme. Am Steinbuch Centre for Computing (SCC) ist er Leiter der Abteilung „Integration“ und „Virtualisierung“.

Dr. Karl-Uwe Stucky leitet am Institut für Angewandte Informatik des Forschungszentrums Karlsruhe die Arbeitsgruppe Optimierung und Ressourcenmanagement im Grid.

(red)

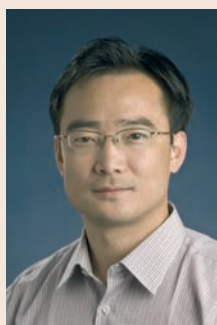
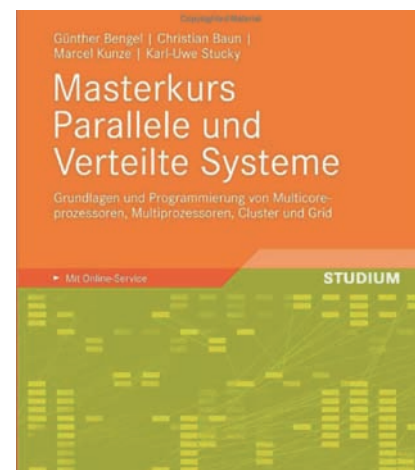


Foto: Privat

Lizhe Wang bester chinesischer Doktorand im deutschsprachigen Raum

Lizhe Wang, Doktorand des Leiters der SCC-Abteilung „Integration und Virtualisierung“ PD Dr. Marcel Kunze, hat im März 2008 als bester chinesischer Doktorand im deutschsprachigen Raum von der chinesischen Regierung die Auszeichnung „2007 Chinese government excellent Ph.D. student abroad“ erhalten. Verliehen wurde ihm diese in der chinesischen Botschaft in Berlin für seine Arbeit „Virtualisierungsumgebungen für das Grid-Computing“. Die Auszeichnung wird alljährlich vergeben und bezieht sich auf alle Forschungsbereiche.

(red)

MATLAB - Numerische Berechnungen einfach erledigen

Lizenzen für Universitätsangehörige und Studierende kostenfrei nutzbar

MATLAB hat sich an der Universität Karlsruhe in den letzten 14 Jahren zunehmend zu einem der wichtigsten numerischen Werkzeuge in vielen Bereichen der Forschung entwickelt. Die Nutzung war zuletzt so sehr gestiegen, dass es immer wieder zu Engpässen bei der Bereitstellung von Lizenzen kam. Nach längeren Diskussionen mit den Instituten und mit der Firma The MathWorks konnte im vergangenen Jahr eine so genannte TAH-Lizenz finanziert werden, die es Universitätsangehörigen und Studierenden erlaubt, das Programm beliebig oft, auch auf privaten Rechnern, zu installieren. Im ersten Jahr mussten die Lizenzkosten noch auf die Universitätseinrichtungen umgelegt werden. Seit Mai dieses Jahres übernimmt das SCC Süd die Kosten nicht nur für Mitarbeiter, sondern auch für Studierende vollständig.

TAH steht für Total Academic Headcount und bedeutet, dass die Kosten für die jeweiligen Lizenzen von der Anzahl der Beschäftigten beziehungsweise Studenten abhängt. Der wesentliche Vorteil dieser Lizenz gegenüber der üblichen Netzwerklizenz besteht darin, dass es keine obere Grenze für die Anzahl an Installationen und gleichzeitiger Nutzung von MATLAB gibt und eine Installation nicht mehr über ein Netz an einen Lizenz-Server gekoppelt sein muss. Dies garantiert eine wesentlich flexiblere und von Infrastrukturen unabhängige Handhabung des Programms.

Die TAH-Lizenzen für Universitätsangehörige und Studierende sind vom Umfang her gleich, das heißt neben MATLAB und SIMULINK stehen 54 weitere Toolboxes und Blocksets zur Verfügung.

Das Produkt sowie die notwendigen Informationen und Schlüssel kön-

nen wie bei allen Campus-Lizenzen des SCC-Süd im Softwareshop der Universität unter <https://rzunika.asknet.de/cgi-bin/home/ml=DE> angefordert werden. Berechtigt sind alle Beschäftigten und Studierenden der Universität. Die Authentifizierung geschieht über die Registrierung bei „The MathWorks“ mit der E-Mail-Adresse, die aus der Domäne *.uni-karlsruhe.de stammen muss.

Angehörige des Forschungszentrums Karlsruhe können die Lizenz leider noch nicht nutzen. Erst wenn die Universität und das Forschungszentrum im KIT eine gemeinsame Rechtsform haben, kann die Authentifizierung über „kit.edu“ erfolgen.

Das neueste Release MATLAB R2008a wurde im April 2008 ausgeliefert. Dies ist ein komplettes System aus dem MATLAB-Programm in der Version 7.6 und vielen weiteren Toolboxes. Dazu kommt mit SIMULINK eine zweite Produktlinie. Dabei handelt es sich um ein Simulationsprogramm, in dem Modelle mittels Funktionsblöcken in einer grafischen Umgebung erzeugt und analysiert werden können und das, ähnlich wie die Toolboxes, eine Reihe von Block Sets für speziellere Anwendungen enthält.

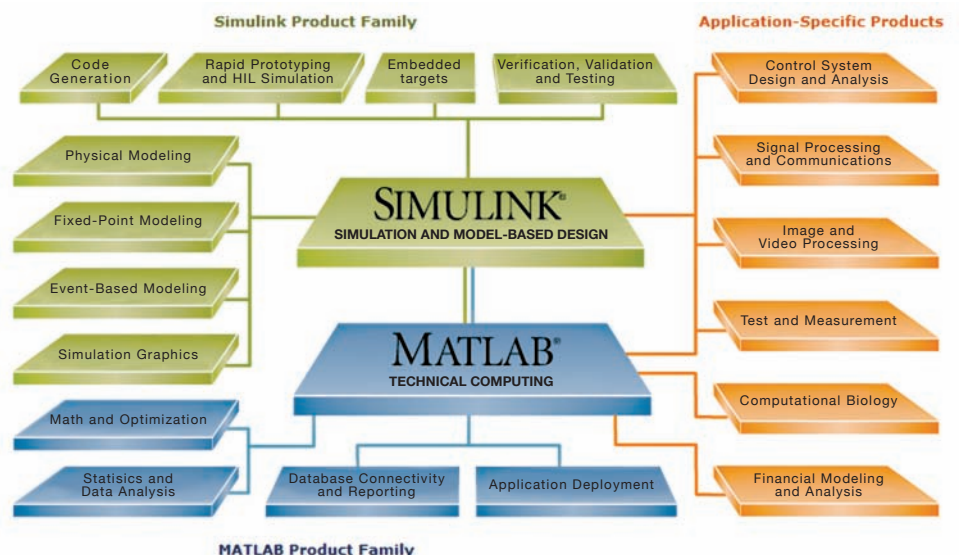
Zu jedem Produkt gibt es auf der Support-Seite der Firma „The MathWorks“ www.mathworks.com/access/helpdesk/help/helpdesk.html einen Link, unter dem man einen kompletten Dokumentationsatz bestehend aus

- Getting Started
- User Guides
- Getting Help
- Examples
- Programming Tips
- Demos

findet, der jeweils auch als PDF-Dokument vorhanden ist. Speziell für Universitäten gibt es unter www.mathworks.com/academia/ weitere Tutorien, den MATLAB Digest und Vieles mehr.

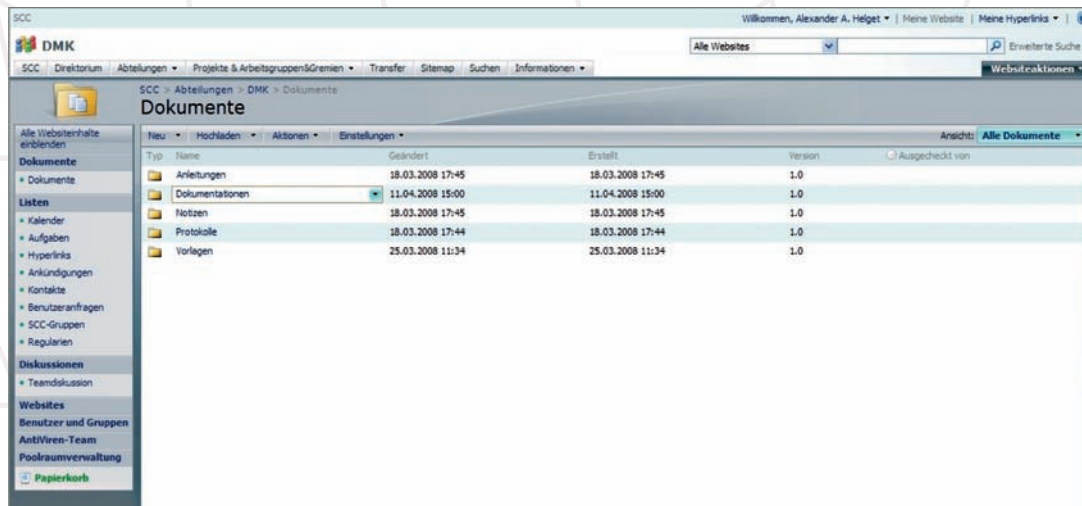
Auf der User Community-Seite www.mathworks.com/matlabcentral/ findet der Informations- und Erfahrungsaustausch unter den MATLAB-Nutzern statt. Dadurch wird es dem Neueinsteiger leicht gemacht, sich in MATLAB einzuarbeiten und von dem Wissen und den Erfahrungen anderer zu profitieren und sich mit ihnen auszutauschen.

Dr. Paul Weber



Zentrale Plattform für Team- und Projektarbeit

SCC bietet Sharepoint-Technologie innerhalb des KIT an



Teamportal: Dokumente können strukturiert mit Versionsverlauf abgelegt werden.

Um effektiv und komfortabel in Projekten, Gruppen und Abteilungen zusammenzuarbeiten, ist eine gemeinsame Plattform zum Austausch von Informationen und Dokumenten ein wichtiges und unentbehrliches Arbeitsmittel. Das SCC hat hierfür ein Teamportal auf der Basis von Microsoft Office Sharepoint Server im Einsatz.

Relevante Informationen sind zentral über eine Teamseite im Intranet zu erreichen und sofort verfügbar. Damit wird sichergestellt, dass jeder Mitarbeiter zeitnah informiert wird. Alle Kolleginnen und Kollegen können auf Dokumente, Kalender und Aufgabenlisten der jeweils aktuellen Version zugreifen. Durch die gemeinsam genutzte Plattform verringern sich Mehrfachablagen von Dokumenten und das E-Mailaufkommen.

Abteilungsinterne Strukturen lassen sich sowohl vertikal nach Abteilungen und Gruppen als auch horizontal für die abteilungsübergreifende Projektarbeit abbilden. Detaillierte Zugriffsberechtigungen grenzen interne und öffentliche Bereiche ab.

Die strukturierte Dokumentenablage erleichtert sowohl das Auffinden von Standarddokumenten

wie Anleitungen, Regelungen oder Vorlagen als auch die teamorientierte Arbeit an Berichten oder Dokumentationen. Durch eine einfache, funktionelle Versionskontrolle ist erkennbar, wer aktuell an einem Dokument arbeitet. Eine komfortable Volltextsuche ergänzt die schnellen Zugriffsmöglichkeiten. Die vertraute Benutzeroberfläche durch Office-Integration erleichtert den Einstieg für den Anwender. Die Möglichkeit, auf weitere wichtige Unternehmensanwendungen aus dem Portal heraus zuzugreifen, macht die Teamseiten für viele Anforderungen interessant.

Das SCC möchte diese Technologie künftig auch für Team- und Projektarbeiten innerhalb des KIT anbieten. Bei Interesse und weiteren Fragen können Sie sich gerne an Sabine Rodegher oder Alexander Helget wenden (mailto:sharepoint-admins@scc.kit.edu).

Sabine Rodegher

Spamfilterumstellung Campus-Nord

Neues System hat sich bewährt

Ende März wurde auf dem Campus-Nord der Spamfilter auf ein neues System umgestellt. Es handelt sich dabei um das gleiche System wie auf dem Campus-Süd. Die Umstellung hat reibungslos funktioniert, außerdem hat sich der neue Spamfilter bereits bewährt, denn es kommen deutlich weniger Spams durch als früher. Da es jedoch einige Fragen zum Umgang mit dem Junk-E-Mail-Ordner im Postfach gab, weist das SCC an dieser Stelle noch einmal auf einige Punkte hin.

Der Ordner Junk-E-Mail sollte nicht gelöscht werden, denn dann können Spams nicht mehr automatisch dorthin verschoben werden. Out-

look 2003 und höher lassen das Löschen dieses Ordners nicht zu und bieten stattdessen die Möglichkeit, diesen zu leeren.

Nachrichten, die älter als 30 Tage sind, werden automatisch aus dem Ordner Junk-E-Mail gelöscht. Der Ordner kann also im Prinzip sich selbst überlassen werden. Der Zugriff auf den Inhalt erfolgt bei Bedarf, zum Beispiel wenn eine Nachricht vermisst wird, die fälschlicherweise als Spam erkannt wurde. Das Postfach läuft auch bei einigen hundert Spams nicht über, da Spams relativ klein sind.

Darüber hinaus sollte eine Outlook-Version ab 2003 verwendet werden.

Zum einen kann der Junk-E-Mail Ordner in Outlook 2002 versehentlich gelöscht werden, zum anderen zeigt Outlook 2002 für Spams die Benachrichtigung für neue Nachrichten an. Der Workaround, die Benachrichtigung umzukonfigurieren, ist nur eine zweitklassige Lösung.

Eine kurze Anleitung zum Spamfilter steht unter http://www.fzk.de/fzk/groups/iwr/documents/internetdokument/id_062501.pdf zur Verfügung.

Gerne werden auch weitere Fragen beantwortet.

Eugenie Bantle

Anregungen erwünscht

IT-Services werden weiter optimiert

Nach der Gründung als gemeinsam getragenes Institut der Universität und des Forschungszentrums Karlsruhe möchte das SCC IT-Dienstleistungen so anbieten, dass die tägliche Arbeit der Kunden und Nutzer optimal unterstützt wird. Das SCC steht nach der Zusammenführung der beiden Organisationen vor großen Herausforderungen und hat sich vorgenommen, die Qualität der IT-Services durch standardisierte Verfahren sicherzustellen und zu verbessern. Um dies zu erreichen, wurde am SCC die Abteilung „IT-Security und Service-Management (ISM)“ eingerichtet.

Gute IT-Services lassen sich jedoch nicht „im Labor“ entwickeln, sie sind vielmehr ein Produkt, das in Zusammenarbeit mit dem Kunden entsteht. Teilen Sie dem SCC deshalb mit, wie Ihre Geschäftsvorhaben noch besser durch Lösungen aus der Informationstechnologie unterstützt werden können. Wenn Sie Wünsche oder Anregungen zu den IT-Services haben, können Sie sich jederzeit an „servicemanager@scc.kit.edu“ wenden.

Bei aktuell anstehenden Störungen, Problemen oder Standard-Service-Anforderungen (wie „neues Benutzerkonto“ oder „Passwort-Reset“)

benutzen Sie bitte die interne Telefonnummer -6060 (FZK) beziehungsweise -8000 (UniKA).

Das SCC freut sich auf eine gute und partnerschaftliche Zusammenarbeit.

Achim Grindler

Projekt „Gefahrstoffe an Schulen“ erfolgreich beendet

Die Verbindung von Theorie und Praxis ist Ziel der Lehr- und Lernmethodik von TheoPrax®. Industriepartner bieten Projektarbeit im Angebot-Auftrags-Verhältnis an Schulen und Hochschulen an, um die Wirtschaft mit der Ausbildung zu verzahnen. Das Forschungszentrum Karlsruhe hat bereits einige Projekte an Theo-Prax erteilt. Im Juli absolvierten Schüler des Markgrafen-Gymnasiums Durlach am SSC das jüngste Projekt „Gefahrstoffe an Schulen“.

Die Schüler stellten mit Unterstützung des TheoPrax-Teams und ihrer Chemielehrerin einen Projektplan auf, der die Themen sowie den Zeit- und Kostenaufwand beschreibt. Nun begann die eigentliche Projektarbeit: Es wurde eine Befragung der Besucher des regionalen Lehrerkongresses 2007 „Dialog Schule-Chemie“ durchgeführt und ausgewertet. Ergänzend dazu wurden einige Lehrer im persönlichen Gespräch befragt. Eine Recherche im Internet über Softwareprodukte zur Verwaltung von Chemikalien an Schulen rundeten die Arbeiten ab.

Anfang Juli präsentierten die Schüler dem TheoPrax-Team und den Lehrern die Ergebnisse und legten einen Abschlussbericht vor. Das Fazit war, dass es sinnvoll und wünschenswert wäre, in den Schulen eine Software einzuführen, mit deren Hilfe sich Schüler und Lehrer über die im Unterricht verwendeten Gefahrstoffe informieren können und die die Verwaltung vorhandener Stoffe erleichtert. Darüber hinaus machten sie Vorschläge für eine schultaugliche Version von „ChemA“ – ein am SCC entwickeltes Chemikalien-Management und Informationssystem, das Unternehmen bei der Verwaltung ihrer Chemikalien und Gefahrstoffe unterstützt.

Doris Heathman



Foto: Privat

Frau Parrisius, stellvertretende Leiterin des TheoPrax-Teams, überreicht die Urkunden für den erfolgreichen Abschluss des Projektes „Gefahrstoffe in Schulen“ an die Schüler Karoline Schäfer (Projektleiterin), Tobias Hochstatter und Franziska Holler (von rechts nach links).

Einführungskurs SAS

Der nächste SAS-Einführungskurs (Statistical Analysis System) findet vom 6.10. bis 10.10.2008 statt. Ziel des Kurses ist das Kennenlernen und Anwenden von SAS. Voraussetzung für die Teilnahme sind grundlegende Kenntnisse einer Programmiersprache. Die erworbenen Kenntnisse können an unterschiedlichen Rechnern eingesetzt werden. Die statistischen Grundlagen sind nicht Lehrstoff des Kurses.

Kursbeginn: Montag, 6.10.2008, 9.00 Uhr, Raum 217
 Kursende: Freitag, 10.10.2008, 17.00 Uhr
 Ort: SCC-Süd, Raum 217, Übungen im Raum 114

Programm

Montag, 6.10.2008	
9.00 - 9.45	Überblick über Statistikprogramme am Rechenzentrum, Vorstellung von SAS
10.00 - 12.00	Die interaktive Oberfläche von SAS, der SAS-Editor, Literatur
14.00 - 15.00	Aufbau von SAS-Programmen, Variablen, Daten, Dateien
15.00 - 17.00	Betreute Übung
Dienstag, 7.10.2008	
9.00 - 10.30	Eingabe von Daten in SAS
10.30 - 12.00	Betreute Übung
14.00 - 15.00	Ausgabe und einfache Auswertungen von Daten
15.00 - 17.00	Betreute Übung
Mittwoch, 8.10.2008	
9.00 - 10.15	Interaktive Dateneingabe, Maskenerstellung (SAS/FSP)
10.15 - 12.00	Betreute Übung
14.00 - 15.30	Grafische Darstellung von Daten (SAS/GRAPH) I
15.30 - 17.00	Betreute Übung
Donnerstag, 9.10.2008	
9.00 - 10.30	Grafische Darstellung von Daten (SAS/GRAPH) II
10.30 - 12.00	Betreute Übung
14.00 - 15.00	Statistische Prozeduren - Überblick und Beispiel (SAS/STAT)
15.00 - 17.00	Betreute Übung
Freitag, 10.10.2008	
9.00 - 10.15	Überblick: SAS/ETS (Zeitreihenanalyse) SAS/OR (Operations Research) SAS/IML (Interactive Matrix Language) SAS/AF (Programmierung von Menü-Oberflächen)
10.15 - 12.00	Betreute Übung
14.00 - 15.00	Zusammenfassung der Kursinhalte, Fragen, Abschlussdiskussion
ab 15.00	Betreute Übung

An den Übungen können nur Angehörige der Universität Karlsruhe teilnehmen. Die Teilnehmerzahl ist auf 20 begrenzt. Zur Anmeldung liegen vorbereitete Listen beim BIT 8000 im Foyer des SCC-Süd aus (Tel. -8000, E-Mail: bit8000@scc.kit.edu).

Dr. Klaus Braune

Weitere Informationen

zu SAS und zum SAS-Kurs unter
<http://www.rz.uni-karlsruhe.de/produkte/2657.php>.

Literatur

SAS Version 6 - Eine Einführung mit Beispielen.
 Skript zum Kurs.
 SAS-Anwenderhandbuch im Netz. Verfügbar unter
<http://www.urz.uni-heidelberg.de/statistik/sas-ah/>.

