

Steinbuch Centre for Computing

# NEWS

SCC

Innovative Plattform für  
Smart Data-Forschung

Innovative platform for smart data research

KIT betreibt zentralen Hochleistungsrechner  
der Landesuniversitäten

KIT runs central high performance computing  
system of state universities

SCC betreibt größte deutsche  
Hochschul-Cloud

SCC operates largest German university cloud



# INHALT

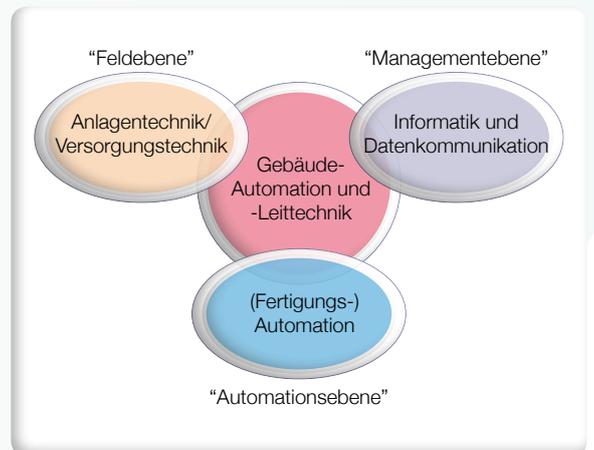
- 4 Innovative Plattform für Smart Data-Forschung
- 6 KIT betreibt zentralen Hochleistungsrechner der Landesuniversitäten
- 8 „bwUniCluster“: Leistung und Nutzungsmöglichkeiten
- 11 SCC betreibt größte deutsche Hochschul-Cloud
- 12 Landesprojekt „bwDataArchiv“  
SCC erweitert digitales Archiv für Langzeitspeicherung von Forschungsdaten
- 15 100 Gigabit/s zwischen KIT und Denver
- 16 Innovation Award für SCC-Website-Management
- 18 FM und SCC entwickeln Software-Verteilungskonzept zur Gebäudeautomation
- 23 Juristische Simulationsstudie erforscht Beweiswert elektronischer Laborbücher
- 26 KIT-WLANs bald auch in der Stadt
- 27 BW-Notebook 3 für alle Hochschulen in Baden-Württemberg
- 28 IHK-Technologiefabrik zeichnet da-cons GmbH aus
- 29 Studie zu Europas Software-Zentren  
Software-Cluster im Vergleich



8



12



18



23

Titelbild: Dr. Holger Obermaier

## EDITORIAL

Liebe Leserinnen und Leser,

diese Ausgabe der SCC-News ist die letzte, die unter der Regie von Ursula Scheller und John Atkinson erscheint. Beide Mitarbeiter verlassen das SCC Ende April 2014. Herr Atkinson wird nach rund 15-jähriger Tätigkeit am SCC in den Ruhestand gehen und Frau Scheller möchte sich nach 25 Jahren SCC-Zugehörigkeit nun auf private Projekte konzentrieren. Dies bedeutet, dass die nächste Ausgabe der SCC-News unter neuen Vorzeichen und gegebenenfalls in einem anderen Format erscheinen wird. Frau Scheller und Herr Atkinson haben über viele Jahre dem SCC und davor dem Universitätsrechenzentrum ein „Gesicht“ gegeben – dafür möchten wir uns sehr herzlich bedanken.

Im vorliegenden Heft geht es insbesondere um landesweite Projekte. So hat das SCC im Januar den „bwUniCluster“, das erste zentrale Hochleistungsrechner-System der baden-württembergischen Universitäten und Hochschulen für eine flächendeckende HPC-Grundversorgung in Betrieb genommen. Hintergrund ist die immer weiter zunehmende Bedeutung des wissenschaftlichen Rechnens in Forschung, Lehre und Studium. Wissenschaftler und Studierende im Land sollen auf der Grundlage modernster Infrastrukturen bestmöglich unterstützt werden.

Ebenfalls im Januar hat das SCC den Online-Speicherdienst „bwSync&Share“ zum einfachen Synchronisieren und Teilen von Dokumenten gestartet. Der landesweite Dienst steht rund 450.000 Mitgliedern aller Hochschuleinrichtungen in Baden-Württemberg kostenfrei zur Verfügung und bildet derzeit die größte deutsche Hochschul-Cloud.

Ein weiteres Landesvorhaben, das ebenso wie die beiden oben genannten Projekte vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg gefördert wird, ist „bwDataArchiv“. Im Rahmen dieses Projekts wird das SCC als Archiv-Standort für Baden-Württemberg seine technische Infrastruktur für die langfristige Archivierung von Forschungs- und Bibliotheksdaten ausbauen.

Ein Einblick in weitere spannende Projekte und Entwicklungen am SCC erwartet Sie auf den nächsten Seiten. Wir wünschen Ihnen viel Vergnügen bei der Lektüre!

Dear readers,

this issue of SCC News is the last one being published under the direction of Ursula Scheller and John Atkinson, as both will leave SCC at the end of April, 2014. This means that the next issue of SCC News might appear under different conditions and in a new format. Over many years, Mrs. Scheller and Mr. Atkinson have given SCC, and previously the University Computing Centre, a face – for this we would like to thank them most cordially.

Topics in the present issue are focusing on statewide projects. In January, SCC put into operation the “bwUniCluster”, the first central HPC system of the Baden-Wuerttemberg universities. Background is the growing importance of scientific computing in research, teaching and studies. The state’s scientists and students are to be supported on the basis of the latest infrastructures in the best possible way.

Also in January, SCC started the online storage service “bwSync&Share” for easy synchronisation and sharing of documents. The statewide service is made available free of charge for around 450,000 members of all state universities and is the largest German university cloud.

Another state project being funded, like the above mentioned, by the Ministry of Science, Research and the Arts Baden-Wuerttemberg is “bwDataArchiv”. In the framework of this project SCC as the state’s archive location will extend its technical infrastructure for long-term archiving of research and library data.

The present issue provides for English speaking readers brief summaries of further projects and current developments at SCC. We hope that you will enjoy reading it!

Hannes Hartenstein, Bernhard Neumair, Achim Streit



Prof. Dr. Hannes Hartenstein



Prof. Dr. Bernhard Neumair



Prof. Dr. Achim Streit

### IMPRESSUM

April 2014

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

76131 Karlsruhe bzw.

Hermann-von-Helmholtz-Platz 1

76344 Eggenstein-Leopoldshafen

Fax: 0721/32550

<http://www.scc.kit.edu/publikationen/scc-news>

Redaktion:

Ursula Scheller (verantwortlich)

Telefon: 0721/608-44865

E-Mail: [ursula.scheller@kit.edu](mailto: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.

# Innovative Plattform für Smart Data-Forschung

**Führende Vertreter aus Industrie, Wissenschaft und Politik haben Anfang Januar am KIT den Startschuss für die Einrichtung der Forschungsplattform Smart Data Innovation Lab (SDIL) gegeben. Mit der Bereitstellung einer Höchstleistungsinfrastruktur sowie Daten aus Industrieprozessen werden wichtige Voraussetzungen für die Spitzenforschung im Bereich Big Data geschaffen. Durch die enge Zusammenarbeit aller Beteiligten können neueste Forschungserkenntnisse effizient an die Industrie weitergegeben werden und zu entscheidenden Wettbewerbsvorteilen für europäische Unternehmen beitragen. Die Projekte werden sich zunächst auf die strategischen Forschungsfelder Industrie 4.0, Energiewende, Smart Cities und Personalisierte Medizin beziehen.**

Das SDIL, das in Zusammenarbeit zwischen Industrie und Forschung konzipiert wurde, wird am KIT betrieben. Neben dem KIT zählen Konzerne wie Bayer, Bosch, Microsoft Deutschland, SAP, Siemens und die Software AG ebenso zu den Gründungspartnern wie das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI), die Fraunhofer-Gesellschaft und das Forschungszentrum Jülich. Darüber hinaus unterstützen bereits heute mehr als 20 weitere Firmen und Institutionen das SDIL. Dazu zählen Unternehmen wie Infineon, Trumpf und Volkswagen sowie der Branchenverband Bitkom und die Deutsche Gesellschaft für Informatik (GI).

Das SDIL steht weiteren interessierten Vertretern aus Wissenschaft und Wirtschaft zur Mitarbeit offen. Ein besonderer Fokus des SDIL liegt auch auf der Förderung von kleinen Unternehmen, die von den Kontakten zu etablierten Anbietern profitieren und völlig neue Lösungen und Dienstleistungen im Umfeld von Big Data anbieten können.

„Die digitalen Datenmengen wachsen in unserer Gesellschaft rasant. Wir brauchen neue Instrumente, um sie zu managen und als Wissensquellen nutzbar zu machen“, erklärte Professorin Johanna Wanka, Bundesministerin für Bildung und Forschung und Co-Vorsitzende der Arbeitsgruppe „Bildung und Forschung für die digitale Zukunft“ des Nationalen IT-Gipfels. So sei das Smart Data Innovation Lab und die dort der Forschung verfügbar gemachten Daten eine optimale Ergänzung, um die Forschung auf diesem Gebiet weiter voranzutreiben.

„Wir freuen uns, unsere Kompetenz in das Smart Data Innovation Lab einzubringen und mit der deutschen Wirtschaft zu teilen“, betonte der Präsident des KIT, Prof. Dr. Holger Hanselka. Das KIT betreibe Infrastrukturen für große Datenmengen wie das Datenzentrum GridKa als Teil eines weltweit verteilten Netzwerkes für das europäische Teilchenbeschleunigerzentrum CERN. Seit Jahren befasse sich die Informatik am KIT mit Analysemethoden, Auswertelgorithmen und Datensicherheit. „Mit SDIL werden KIT und seine Industriepartner aus Big Data tatsächlich auch Smart Data machen.“

„Durch den Fokus auf Themen wie Industrie 4.0 oder personalisierte Medizin gewinnen wir Erkenntnisse, die nicht nur zur Lösung wirtschaftlicher, sondern auch gesellschaftlich relevanter Herausforderungen beitragen“, sagte Jim Hagemann Snabe, Co-CEO SAP AG. „Das erreichen wir auch durch eine Erweiterung des Netzwerkes

in Richtung kleiner Unternehmen. Innovative Ideen kommen vor allem von diesen jungen Unternehmen und es ist unsere Aufgabe, sie aktiv zu unterstützen.“

Die Arbeitsgruppen für die aktuellen vier Forschungsfelder werden von jeweils einem Vertreter aus Wissenschaft sowie Industrie geleitet. Die Arbeitsgruppen entscheiden gemeinschaftlich über die Ausgestaltung und Vergabe der Ressourcen des SDIL für Forschungsprojekte. Weitere Forschungsschwerpunkte sind im Laufe der Zeit geplant.

Die Wissenschaftler erhalten im Rahmen abgegrenzter Projekte Zugang zu Daten aus Industrieprozessen, die auf der Plattform sicher gespeichert sind. Die Analyse der Daten, beispielsweise die Beschreibung und Strukturierung von spezifischen Datensätzen oder auch das Auffinden von Anomalien, erfolgt in enger Zusammenarbeit mit dem jeweiligen Unternehmenspartner, wodurch ein schneller Wissens- und Technologietransfer möglich wird. Die beteiligten Forschungsinstitute werden außerdem allgemeingültige Werkzeuge und Methoden zur Datenanalyse entwickeln, die über die Plattform allen Teilnehmern des Smart Data Innovation Lab zur Verfügung gestellt werden können.

Das SDIL fügt sich auch in die Baden-Württemberg-Initiative „Forward IT“ ein, die das Zusammenwirken von Wirtschaft und Wissenschaft unter anderem auf den Feldern Industrie 4.0, Unternehmenssoftware, Mobilität und IT-Sicherheit stärken will.

Das Konzept für die Forschungsplattform wurde in der Arbeitsgruppe „Bildung und Forschung für die digitale Zukunft“ des Nationalen IT-Gipfels der Bundesregierung entwickelt. Die Arbeitsgruppe konzentriert sich auf die Aus- und Weiterbildung von Fachkräften sowie Forschung für die digitale Zukunft und verwirklichte in dem Zusammenhang bereits zwei erfolgreiche Initiativen: Software Campus Netzwerk (2011) und Academy Cube (2012).

(red)

## Weitere Informationen

<http://www.sdil.de/>

## Innovative platform for smart data research

At the beginning of January, leading representatives of industry, science, and politics kicked off the establishment of the Smart Data Innovation Lab (SDIL). This high performance infrastructure and the supply of data from industrial processes are major prerequisites for excellent research related to big data. Close cooperation of all parties will ensure efficient transfer of latest research findings to industry and decisive competitive advantages of European enterprises. First projects will focus on the strategic research areas of industry 4.0, energy turnaround, smart cities, and personalised medicine.

The SDIL research platform was designed by research in cooperation with industry and will be operated at KIT. Apart from KIT, large corporations like Bayer, Bosch, Microsoft Deutschland, SAP, Siemens, and Software AG as well as the German Research Center for Artificial Intelligence (DFKI), the Fraunhofer-Gesellschaft, and Forschungszentrum Jülich are among the founding partners. In addition, more than 20 other firms and institutions support SDIL. Among them are companies like Infineon, Trumpf, and Volkswagen, the ICT industry association Bitkom, and the German Informatics Society (Gesellschaft für Informatik, GI). SDIL is open for participation of other interested representatives of science and industry. A particular focus of SDIL lies on the support of small enterprises that will profit from the contacts to established suppliers and, hence, may offer entirely novel solutions and services in the area of big data.

“Digital data volumes are growing rapidly in our society. We need new instruments to manage them and to use them as sources of knowledge,” Professor Johanna Wanka said, Federal Minister of Education and Research and Co-chairperson of the working group “Education and Research for the Digital Future” of the National IT Summit. She pointed out that the Smart Data Innovation Lab and the data supplied to research are optimally suited for advancing studies in this area.

“We are happy to contribute our competence to the Smart Data Innovation Lab and to share it with German industry,” underlined the President of KIT, Prof. Dr. Holger Hanselka. He emphasized that KIT already operates infrastructure facilities for big data, such as the GridKa data centre that is part of the worldwide grid of the European particle accelerator centre CERN. For many years, computer science at KIT has dealt with analysis methods, evaluation algorithms, and data security. “With SDIL, KIT and its industry partners will really turn big data into smart data.”

“As the focus lies on topics like industry 4.0 or personalised medicine, we will obtain findings that will help



Mit der Analyse riesiger Datenmengen hat sich das KIT bereits einen Namen gemacht. Das am SCC angesiedelte Tier1-Zentrum „GridKa“ speichert und analysiert Daten aus den Teilchenbeschleuniger-Experimenten am CERN. Foto: SCC

manage not only economic, but also societal challenges,” Jim Hagemann Snabe, Co-CEO of SAP AG, says. “For this purpose, we also open our network for smaller enterprises. It is mainly these small young enterprises that create innovative ideas and it is our task to support them actively.”

The working groups of the current four research areas are headed by a representative of science and a representative of industry each. The groups jointly decide on the type and allocation of SDIL resources for research projects. Additional research topics are planned to be defined in the future.

Within the framework of specific projects, the scientists will be granted access to data from industrial processes that are stored securely on the platform. Analysis of the data, such as description and structuring of specific data sets or finding of anomalies, will be accomplished in close cooperation with the company partner, thus ensuring rapid knowledge and technology transfer. Moreover, the research institutes involved will develop universal tools and methods for data analysis. Via the platform, they will then be made available to all partners of the Smart Data Innovation Lab.

The SDIL is also part of the Baden-Württemberg initiative “Forward IT” to strengthen cooperation of industry and science in the areas of industry 4.0, enterprise software, mobility, and IT security.

The concept of the research platform was developed by the working group “Education and Research for the Digital Future” of the federal government’s National IT Summit. The working group concentrates on the education and qualification of qualified staff and research for the digital future. In this connection, two successful initiatives have already been implemented, namely, Software Campus Network (2011) and Academy Cube (2012).

(red)

## KIT betreibt zentralen Hochleistungsrechner der Landesuniversitäten

**Der „bwUniCluster“, das erste zentrale System für eine flächendeckende Grundversorgung der baden-württembergischen Universitäten und Hochschulen mit Hochleistungsrechnerkapazität, wurde am 27. Januar am KIT in Betrieb genommen. Der Cluster wird von den Landesuniversitäten und dem MWK hinsichtlich Investitions- und Betriebskosten gemeinsam getragen und ist Teil des baden-württembergischen Landeskonzepts für das Hoch- und Höchstleistungsrechnen „bwHPC“ in Forschung und Lehre.**

Das Landeskonzept „bwHPC“, das von der Deutschen Forschungsgemeinschaft (DFG) als innovativ und überregional wegweisend begutachtet und zur Förderung empfohlen wurde, gilt deutschlandweit als vorbildlich. „bwHPC“ umfasst die Förderung des Hoch- und Höchstleistungsrechnens (High Performance Computing, HPC) mit unterschiedlichen Leistungsklassen und Kompetenzzentren auf allen Ebenen: vom Ausbau des Stuttgarter Höchstleistungsrechenzentrums über einen breiten Zugang der Landeshochschulen zu den Hochleistungssystemen des KIT bis hin zu speziellen Rechen-Clustern an Universitäten und Forschungseinrichtungen des Landes. Der „bwUniCluster“ am SCC des KIT bildet dabei das zentrale Einstiegs- und Grundversorgungssystem für die Wissenschaftler in Baden-Württemberg.

„Um der zunehmenden Bedeutung des wissenschaftlichen Rechnens in Forschung, Lehre und Studium begegnen zu können, sind wissenschaftliche Institutionen auf modernste Infrastrukturen angewiesen. Die Universitäten und weitere Einrichtungen des Landes haben daher gemeinschaftlich Konzepte entwickelt, die auf dem Prinzip einer kooperativen Bereitstellung und Nutzung von Ressourcen und Diensten basieren, um so Wissenschaftler und Studierende im Land bestmöglich unterstützen zu

können. Der bwUniCluster als Teil dieser Konzeption ersetzt bzw. ergänzt die lokalen Hochleistungsrechner-Ressourcen an den Landesuniversitäten und ermöglicht so ein arbeitsteiliges und effizienteres Vorgehen“, erklärte Dr. Holger Marten, Abteilungsleiter für Wissenschaftliches Rechnen und Simulation am SCC.

Das massiv parallele System mit einer theoretischen Spitzenleistung von 176 TFLOPS, das sind 176 Billionen Rechenoperationen in der Sekunde, steht auch weiteren, bislang nicht beteiligten Landeshochschulen sowie für Kooperationen mit der Industrie gegen Kostenbeteiligung zur Verfügung. Auf der Grundlage des vom SCC koordinierten, innovativen Identitätsmanagement-Projekts „bwIDM“ bietet es den Wissenschaftlern unabhängig vom Standort einen höchst einfachen und sicheren Zugang wie bei einem System vor Ort. Zudem fördert das Land ein „bwHPC“-Begleitprojekt, das die Wissenschaftler bei der Nutzung der Hochleistungsrechnensysteme unterstützt. Die wissenschaftsadaquate Nutzung und Fortentwicklung der Systeme wird durch einen landesweiten Ausschuss gesteuert, in den alle beteiligten Hochschulen Wissenschaftsvertreter entsendet haben.

Ursula Scheller



Der „bwUniCluster“ am KIT stellt zentrale Hochleistungsrechnerkapazität für die Universitäten und Hochschulen in Baden-Württemberg bereit. Foto: Dr. Holger Obermaier

## KIT runs central high performance computing system of state universities

On January 27, 2014, the high performance computing system “bwUniCluster” took up operation at KIT. It is the first centralised compute cluster supplying the Baden-Wuerttemberg universities with basic universal high performance computing capacities. The cluster being financed by the Ministry of Science, Research and the Arts Baden-Wuerttemberg and the state universities is part of the Baden-Wuerttemberg state concept “bwHPC” for high performance computing in research and education.

The “bwHPC” state concept was found to be of innovative character and supraregional importance and recommended for funding by the German Research Foundation (DFG). It is considered a model concept for Germany. “bwHPC” covers the promotion of high performance computing (HPC) with different performance classes and competence centres at all levels. Within the framework of this concept, it is planned to extend the high performance computing centre in Stuttgart, to grant broad user groups of the state universities access to the high performance computing systems at KIT, and to establish special, dedicated computing clusters at universities and research institutions of the state. In this context the “bwUniCluster” at SCC represents the central HPC supply and entry-level system for scientists in Baden-Wuerttemberg.

The massively parallel system with a theoretical peak performance of 176 teraflops, corresponding to 176 trillion floating point operations per second, can also be accessed by state universities that are not yet involved in the program and by cooperative ventures with industry against payment. Under the “bwIDM” innovative identity management project coordinated by SCC, scientists are granted very easy and secure access independently of their location. Moreover, the state funds a “bwHPC” support project for the use of high performance computing systems by the scientists. Science-adequate use and further development of the systems is controlled by a state committee consisting of representatives of all universities involved.

(red)

“To take account of the increasing significance of scientific computing in research and education, scientific institutions are dependent on latest infrastructure facilities. For this reason, universities and other institutions in Baden-Wuerttemberg have jointly developed concepts for the cooperative supply and use of resources and services to support the state’s scientists and students in the best possible way. The “bwUniCluster” is part of this state concept and replaces or complements local high performance computing resources of state universities. So work can be shared and efficiency is enhanced,” explained Dr. Holger Marten, Head of the Scientific Computing and Simulation Department of SCC.



Baden-Württemberg

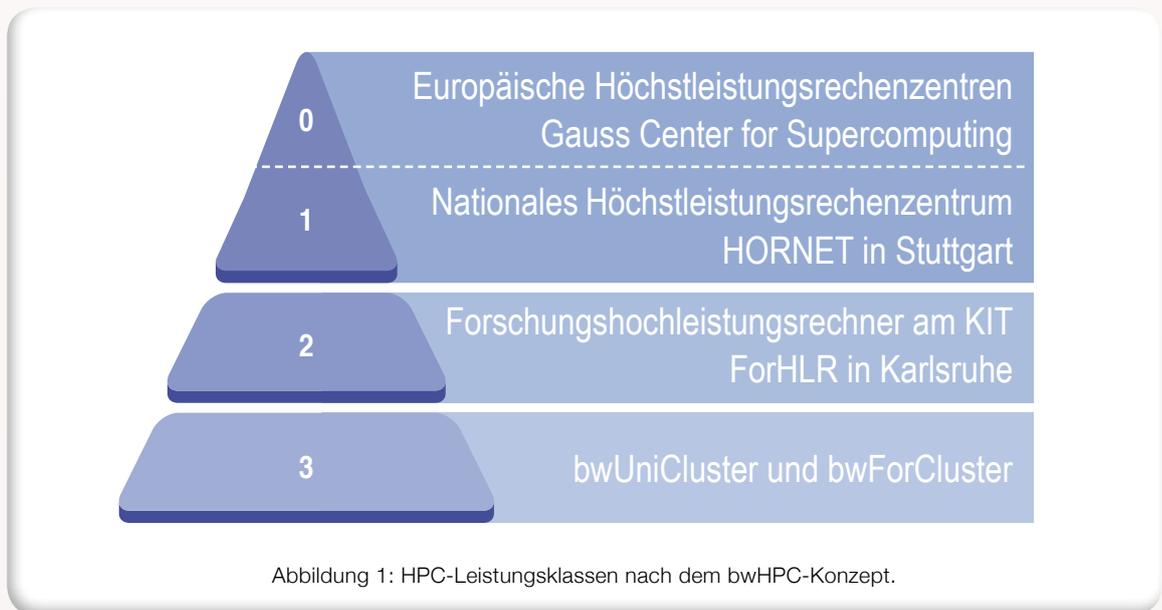
# „bwUniCluster“: Leistung und Nutzungsmöglichkeiten

Das baden-württembergische Landeskonzept „bwHPC“ für das Hoch- und Höchstleistungsrechnen (High Performance Computing, HPC) in Forschung und Lehre (s. Seite 6) sieht die Schaffung eines durchlässigen HPC-Ökosystems mit unterschiedlichen Leistungsklassen (s. Abbildung 1) für unterschiedliche Anwendungsfälle vor, die von Berechnungen auf Einzelprozessoren über „multithreaded“-Rechnungen bis zu MPI-Programmen auf sehr vielen Prozessoren reichen können. In diesem Zusammenhang werden insbesondere die HPC-Systeme der Leistungsklasse 3 neu geordnet. Dabei steht der „bwUniCluster“ in Karlsruhe für die allgemeine Fachbereichsversorgung zur Verfügung, während die vier Forschungsrechencluster („bwForCluster“) in Mannheim/Heidelberg, Freiburg, Ulm und Tübingen bis 2016 speziell auf Anwendungsprogramme aus den Forschungsbereichen Wirtschafts- und Sozialwissenschaften, Neurowissenschaften, Mikrosystemtechnik, theoretische Chemie, Physik und Biologie zugeschnitten sein werden.

Alle fünf Cluster sehen ein föderatives Betriebskonzept vor, das durch ein ebenfalls föderatives Projekt zur Unterstützung der Nutzer, „bwHPC-C5“, begleitet wird (s. [www.bwhpc-c5.de](http://www.bwhpc-c5.de)). Ziel dieses Projekts, das eine Weiterentwicklung der im Grid-Umfeld erlangten Erfahrungen darstellt, ist es, den Nutzern einen möglichst einfachen Zugang zu den Clustern sowie eine anwendungsbezogene, föderative Beratung anzubieten.

Rechenknoten mit 32 bzw. 8 Cores und 32 bzw. 18 GB Hauptspeicher pro Core.

Beide Cluster besitzen ein InfiniBand Interconnect; auch die Dateisysteme (außer den lokalen Platten) sind über InfiniBand angeschlossen. Die Nutzer sehen auf beiden Rechnern dasselbe HOME-Verzeichnis.



Um den Beschäftigten des KIT eine erste Einschätzung der Nutzbarkeit und Leistungsfähigkeit des „bwUniCluster“ zu ermöglichen, wird nachfolgend ein Vergleich zu dem von sehr vielen Mitarbeitern genutzten KIT-Parallelrechner HP XC3000 (HC3) gezogen.

Zu den Gemeinsamkeiten zählen folgende Punkte:

- Auf beiden Clustern sind Intel Xeon-Prozessoren verbaut; der „bwUniCluster“ und die „HC3“ besitzen „dünne“ Rechenknoten mit 16 bzw. 8 Cores und 4 bzw. 3 GB Hauptspeicher pro Core sowie „fette“

Sie laufen beide unter Linux und verfügen über ein ähnliches Betriebskonzept. Auf beiden Rechnern ist die Managementsoftware KITE installiert. Die Middleware ist nahezu identisch und wird auf dem „bwUniCluster“ und der „HC3“ zusammen mit der Anwendungssoftware über Module zur Verfügung gestellt.

Der Zugang zum „bwUniCluster“ wurde gegenüber dem „bwGRiD“-Rechner stark vereinfacht und entspricht nun dem Zugangsverfahren bei der „HC3“.

Wesentliche Unterschiede sind:

- Der „bwUniCluster“ ist ein auf den Durchsatz von Jobs, die auf einer kleineren Anzahl von Prozessoren laufen, ausgerichtetes Grund- und Lehrversorgungssystem, das nicht die anspruchsvollen Produktionsrechnungen der Fachbereiche Wirtschafts- und Sozialwissenschaften, Neurowissenschaften, Mikrosystemtechnik, theoretische Chemie, sowie Teilbereiche der Physik und Biologie abdecken wird. Dafür werden die zukünftigen „bwForCluster“ zur Verfügung stehen.
- Der „bwUniCluster“ besitzt mehr Rechenknoten als die „HC3“ und ist aufgrund der neueren Prozessorgeneration („Sandy Bridge“) wesentlich leistungsfähiger als die „HC3“; so haben die „dünnen“ Knoten eine mehr als 4-fach höhere theoretische Leistung im Vergleich zu den Knoten der „HC3“ mit der Architektur „Gainestown“, die „fetten“ Knoten sogar eine 8-fach höhere Leistung.
- Bei der Middleware gibt es für die Nutzer zwei wesentliche Unterschiede:
  - › Die „HC3“ nutzt eine am SCC entwickelte Nutzerschnittstelle zum Batchsystem Slurm, der „bwUniCluster“ verwendet dagegen MOAB und bietet damit einen leichteren Umstieg auf die thematisch orientierten „ForCluster“ sowie auf das künftige Tier2-System „ForHLR“, die ebenfalls mit MOAB betrieben werden sollen.
  - › Die Handhabung der Software Module Files orientiert sich an der von den „bwGRiD“-Clustern und ist durch die Kategorisierung nach Fachbereichen und Themen differenzierter als auf der „HC3“ gestaltet.
- Der Nutzer-Support wird gemeinsam mit den anderen baden-württembergischen Universitäten und Hochschulen erbracht. Anfragen werden KIT-übergreifend durch entsprechende Experten aus dem Land bearbeitet.

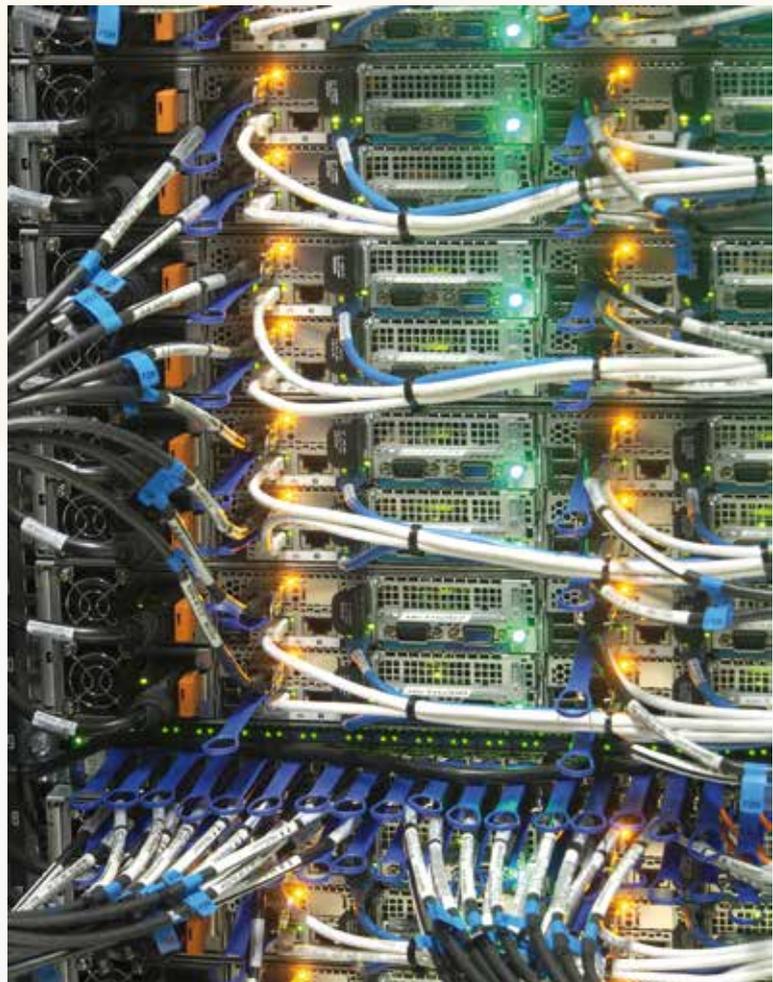
### Konfiguration des „bwUniCluster“

Die neue Maschine besteht aus Serverknoten der Firma Dell vom Typ PowerEdge C6220 bzw. PowerEdge R820. Die Dell-Knoten vom Typ PowerEdge C6220, die innerhalb eines Gehäuses von zwei Höheneinheiten jeweils vier unabhängige Server enthalten, sind preislich günstiger als Blades, haben aber eine höhere Packungsdichte als klassische Rack-Serverknoten. Beide Serverknotentypen beinhalten Intel Xeon „Sandy Bridge“-Prozessoren mit acht Rechenkernen mit einer L3-Cachegröße von 20 MB pro Sockel. Die „dünnen“ Knoten vom Typ PowerEdge C6220 haben zwei Sockel und somit 16 Cores und eine Taktfrequenz von 2,6 GHz, während die „fetten“ Knoten vom Typ PowerEdge R820 vier Sockel und somit 32 Cores und eine Taktfrequenz von 2,4 GHz besitzen.

Das Kommunikationsnetzwerk zwischen den Knoten des Clusters ist als Infiniband 4X FDR der Firma Mellanox realisiert. Diese Interconnect-Technologie besitzt die enorme Bandbreite von 56 Gbit pro Sekunde (jeweils Senden und Empfangen gleichzeitig) bei einer gleichzeitig sehr geringen Latenz von ungefähr einer Mikrosekunde.

Die Gesamtkonfiguration des neuen Clusters sieht wie folgt aus:

- 2 Login-Knoten basierend auf der „Sandy Bridge“-Architektur mit jeweils 16 Rechenkernen mit einer theoretischen Spitzenleistung von 332,8 GFLOP pro Sekunde und 64 GB Hauptspeicher pro Knoten.
- 512 „dünn“ Rechenknoten basierend auf der „Sandy Bridge“-Architektur mit jeweils 16 Rechenkernen mit einer theoretischen Spitzenleistung von 332,8 GFLOP pro Sekunde und 64 GB Hauptspeicher pro Knoten.
- 8 „fette“ Rechenknoten basierend auf der „Sandy Bridge“-Architektur mit jeweils 32 Rechenkernen mit einer theoretischen Spitzenleistung von 614,4 GFLOP pro Sekunde und 1 TB Hauptspeicher pro Knoten.



Der „bwUniCluster“ verfügt über mehr als 500 Rechenknoten und eine theoretische Spitzenleistung von 176 TFLOPS. Foto: Dr. Holger Obermaier

- InfiniBand 4X FDR Interconnect mit ConnectX-3 Single Port FDR HCA.

Die acht „fetten“ Rechenknoten mit je 1 TB Hauptspeicher sind für sehr hauptspeicherintensive Shared-Memory-Anwendungen bestens geeignet (wie zum Beispiel für CAE-Anwendungen) und für diesen Zweck zusätzlich noch jeweils mit einer lokalen Plattenkapazität von nahezu 4 TB in schneller RAID-Konfiguration ausgestattet. Insgesamt zehn weitere Server sind für die Infrastruktur und das Management des Clusters vorhanden.

Für sehr schnellen File-I/O, insbesondere von parallelen Jobs, verfügt das System über parallele Lustre-File-Systeme von ungefähr 850 TB Kapazität und einem Gesamtdurchsatz von bis zu 16 GB/s. In diesem File-System werden temporäre und sehr große Scratch-Daten abgelegt.

Die permanent verfügbaren HOME-Verzeichnisse liegen in parallelen Lustre-File-Systemen, an die sowohl der KIT-Parallelrechner HP XC3000 (HC3) als auch der Instituts-Cluster II (IC2) angeschlossen sind. Die Konfiguration des Infiniband-Netzwerks ermöglicht den Anwendern eine gemeinsame Nutzung des HOME-Verzeichnisses von allen drei Clustern aus.

### Software

Die Softwareausstattung des „bwUniCluster“ wird sehr variabel gestaltet und wird sich an den Bedürfnissen der landesweiten Grundversorgung und Lehraktivitäten des Hochleistungsrechnens orientieren.

Gegenwärtig stehen auf dem „bwUniCluster“ aus lizenzrechtlichen Gründen nur wenige CAE-Softwarepakete (CAE steht für Computer Aided Engineering) zur Verfügung. Namentlich sind das die CAE-Pakete OpenFOAM und ANSYS CFX, wobei auf ANSYS vorläufig nur von KIT-Angehörigen zugegriffen werden darf. Beide Pakete sind dem Bereich Strömungsdynamik zuzuordnen.

Weitere installierte Programmpakete sind das mathematische Programmpaket MATLAB und die Statistiksoftware R. Es stehen auch verschiedene Unterprogramm-bibliotheken wie MKL (Math Kernel Library) von Intel, GSL (GNU Scientific Library) und FFTW zur Lösung der Fast Fourier Transformation bereit.

Gegenwärtige und zukünftige Software-Installationen der Fachbereiche Wirtschafts- und Sozialwissenschaften, Neurowissenschaften, Mikrosystemtechnik, theoretische Chemie, Physik und Biologie werden mit der Inbetriebnahme der „bwForCluster“ auf diese Cluster umziehen.

Zum Übersetzen und Linken eigener Programmpakete stehen Compiler von Intel und GNU für die Programmiersprachen C/C++, Fortran95 (Fortran2008) und Java zur Verfügung. MPI-Programme können OpenMPI oder Intel MPI nutzen. Hinzu kommen Werkzeuge zum Debuggen und zur grafischen Analyse eigener Programme.

Hartmut Häfner

### Weitere Informationen

<http://www.scc.kit.edu/dienste/9237.php>

Benutzersupport

<http://www.bwhpc-c5.de>

Handbücher

<http://www.bwhpc-c5.de/wiki>

## „bwUniCluster“: performance and usage

The Baden-Wuerttemberg concept “bwHPC” for high performance computing in research and teaching provides for the creation of a permeable HPC ecosystem with different performance classes (see figure 1) for different use cases, that can range from calculations on individual processors over “multithreaded” computations to MPI programs on many processors. In this context, the HPC systems of tier 3 are rearranged. The “bwUniCluster” in Karlsruhe is available for the general supply of the university departments, while the four research computing clusters (“bwForClusters”) in Mannheim/Heidelberg, Freiburg, Ulm and Tübingen will be tailored specifically to applications from the research fields of economics and social sciences, neurosciences, microsystems technology, theoretical chemistry, physics, and biology. All five clusters provide a federated operating concept, which is accompanied by a project for the support of the users (see [www.bwhpc-c5.de](http://www.bwhpc-c5.de)). This project, which represents a further development of the experience gathered with the grid environment, offers a simple access policy and an application-related, federated support to the users.



Baden-Württemberg

# SCC betreibt größte deutsche Hochschul-Cloud

**Das SCC hat im Januar den Online-Speicherdienst „bwSync&Share“ zum einfachen Synchronisieren und Teilen von Dokumenten in Betrieb genommen. Der landesweite Dienst, der vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg getragen wird, steht rund 350.000 Studierenden und 100.000 Beschäftigten aller Hochschuleinrichtungen in Baden-Württemberg zur kostenfreien Nutzung zur Verfügung und bildet die größte deutsche Hochschul-Cloud.**

Um in der aktuellen Forschungslandschaft bestehen zu können, müssen sich Wissenschaftler zunehmend einrichtungübergreifend organisieren und kooperieren. Mit dieser Entwicklung steigt auch der Bedarf an geeigneten Werkzeugen, die eine solche Zusammenarbeit unterstützen. Die gemeinsame Bearbeitung und der einfache, sichere und flexible Austausch von Dokumenten – auch über Hochschulgrenzen hinweg – ist daher eine wichtige Voraussetzung.

Mit „bwSync&Share“ wurde am SCC in Zusammenarbeit mit der Firma Power Folder ein Speicherdienst entwickelt, der den Austausch von Daten zwischen verschiedenen Benutzern, Computern und mobilen Endgeräten ermöglicht. Forschungsergebnisse, Hausarbeiten, Bilder oder Videos können gespeichert, geteilt und synchronisiert werden. Der Zugriff erfolgt dabei über Desktop-Systeme, Smartphones, Tablets oder plattformübergreifend über eine Web-Schnittstelle und bietet dem Nutzer so größtmöglichen Komfort. In der Startphase stehen für „bwSync&Share“ in Karlsruhe zunächst 500 Terabyte bereit. Der Speicherplatz pro Nutzer ist auf 10 Gigabyte beschränkt.

„Im Gegensatz zu anderen bekannten Angeboten wie ‚Dropbox‘ oder ‚Google Drive‘ werden die Daten nicht an ausländischen Serverstandorten, sondern an der Large Scale Data Facility des KIT und damit im deutschen Rechtsraum gespeichert. Beschäftigte und Studierende der baden-württembergischen Universitäten und Hochschulen müssen sich also nicht mehr externen Cloud-Diensten anvertrauen, sondern können über diesen landesweiten Dienst einfach und sicher kooperieren“, erklärt Nico Schlitter, Projektleiter im SCC.

Die Kommunikation zwischen den Endgeräten der Nutzer und den zentralen Speichersystemen erfolgt verschlüsselt. Der Datenzugriff ist beschränkt auf den die Daten bereitstellenden Nutzer und vom ihm autorisierte weitere Nutzer.

Die Nutzung von „bwSync&Share“ ist für Mitglieder der Landesuniversitäten und -hochschulen kostenfrei, auch Angehörige anderer Einrichtungen in Baden-Württemberg können diesen Dienst in Anspruch nehmen – ohne jedoch über einen persönlichen Speicherplatz zu verfügen.

Ursula Scheller

## Zugang zum Dienst und weitere Informationen

<https://www.scc.kit.edu/dienste/bwSyncAndShare.php>

## SCC operates largest German university cloud

In January SCC put into operation the online storage service "bwSync&Share" for easy synchronisation and sharing of documents. The state-wide service, funded by the Ministry of Science, Research and the Arts Baden-Wuerttemberg, is made available free of charge to the around 450,000 students and employees of all state universities and is the largest German university cloud.

# Landesprojekt „bwDataArchiv“

## SCC erweitert digitales Archiv für Langzeitspeicherung von Forschungsdaten

Das SCC ist mit der Large Scale Data Facility (LSDF<sup>1</sup>), dem GridKa und der KIT-internen Tivoli Storage Manager-Infrastruktur (TSM) schon viele Jahre für die langfristige Aufbewahrung wissenschaftlicher Daten aus Experimenten und Messungen zuständig, auch wenn die Analyse bereits abgeschlossen ist oder regelmäßige Zugriffe nicht länger erforderlich sind. Die Menge dieser Daten wird in nächster Zeit stark ansteigen, unter anderem aufgrund der Rolle des KIT als Archiv-Standort für Baden-Württemberg. Innerhalb des kürzlich angelaufenen und vom Ministerium für Wissenschaft Forschung und Kunst Baden-Württemberg geförderten Landesprojekts „bwDataArchiv“ wird das SCC seine technische Infrastruktur für die langfristige Archivierung von Forschungs- und Bibliotheksdaten, wie etwa von Daten der Großrechner des Höchstleistungsrechenzentrums Stuttgart (HLRS), weiter ausbauen. Im Rahmen verschiedener Projekte sollen Technologien entwickelt und eingesetzt werden, die eine strukturierte, verlässliche und sichere Langzeitspeicherung von mehr als ein Exabyte ermöglichen werden.

### Bewahren ist Pflichtsache

Die Aufbewahrung digitaler Daten ist nicht nur aus rechtlichen Gründen wichtig, sondern verpflichtet auch schon aufgrund historischer Werte von Daten. Aufbewahrungspflichten sind bezüglich des Anwendungsfalls sowie der Art und des Inhalts der Dokumente geregelt. Die deutsche Rechtsordnung schreibt kein allgemeines „Aufbewahrungsgesetz“ vor, aus dem einheitliche Fristen und Anforderungen entnommen werden können. Sie werden stattdessen in verschiedenen spezialgesetzlichen Regelungen definiert. Bestehen gesetzliche Aufbewahrungspflichten schon für Papierdokumente, gelten diese auch für die entsprechenden elektronischen Dokumente. Wie für die Aufbewahrungspflicht gibt es keine allgemeine Regelung zu Aufbewahrungsform und -dauer, diese Kriterien sind ebenfalls vom Anwendungsfall und speziellen Regelungen abhängig.

Forschungsdaten bilden eine Grundlage des wissenschaftlichen Erkenntnisgewinns und werden daher allein schon wegen der Nachnutzung aufbewahrt. Die Sicherung der Daten kann jedoch auch zur Prüfung früherer Ergebnisse herangezogen werden. Um eine hohe Qualität von Forschung und Ausbildung des wissenschaftlichen Nachwuchses zu sichern und Betrug und Fälschung im Wissenschaftsbetrieb vorzubeugen, sollen Forschungsdaten in nachvollziehbarer und nachhaltiger Weise gesichert werden. Die Deutsche Forschungsgemeinschaft (DFG) sieht in ihren Vorschlägen zur Sicherung der guten wissenschaftlichen Praxis vor, dass Primärdaten als Grundlagen für Veröffentlichungen auf haltbaren und gesicherten Trägern für zehn Jahre aufbewahrt werden sollen.

Neben der Nachnutzbarkeit von Daten und (gesetzlichen) Regelungen zur Aufbewahrung, ist der Erhalt von nicht reproduzierbaren Daten oder sehr kostbaren Daten von Bedeutung. Zum Beispiel existieren in „The Language Archive“ des Max-Planck-Instituts für Psycholinguistik, Tonaufzeichnungen von Sprachen, die nicht

mehr aktiv gesprochen werden. Ein Verlust dieser Daten wäre ein kultureller Schaden. Dagegen sind die astronomischen Bilder der „Sloan Digital Sky Survey“ oder die Messdaten des Large Hadron Collider am CERN schon aus Kostengründen (zum Beispiel durch kostspielige Projekte, die Verwendung von Satelliten und den Einsatz teurer Instrumente) unersetzlich.

### Langzeitarchivierung als technische und organisatorische Herausforderung

Die digitale Langzeitarchivierung beschäftigt sich insbesondere mit der Frage, wie digitale Daten über Technologiewechsel hinweg erhalten bleiben können. Mit Daten ist jegliche Art von Information gemeint, die als Sequenz von Nullen und Einsen dargestellt werden kann. Die Aufgabe besteht daher zum einen darin, diese Sequenz und damit die Datenobjekte physisch zu erhalten. Dies wird als „bit stream preservation“ bezeichnet. Zum anderen gehört zur Interpretation der gespeicherten Informationen eine Spezifikation, die durch eine entsprechende Software unterstützt wird. Das bwFLA<sup>2</sup>-Projekt der Universität Freiburg hat dazu grundlegende Services implementiert. Zu guter Letzt muss die interpretierte Information so dargestellt werden, dass sie auch für den Menschen verständlich ist. Die digitale Langzeitarchivierung umfasst daher drei Ebenen. Das Aufbewahren der Daten wird heutzutage häufig durch magnetische Datenträger, wie zum Beispiel eine Festplatte, realisiert. Für den Erhalt der logischen Information wird entsprechende Software, wie zum Beispiel MSWord 1996, verwendet und muss daher auch aufbewahrt werden bzw. zugänglich sein. Neben der rein technischen Ausrichtung sind organisatorische Prozesse ein wichtiger Teil der Archivierung. Dazu gehören Verfahren, wie Einlagerung der Daten in Archivsysteme (Data Ingest), Lesen der Daten (Retrieval), die Wartung der Daten (Curation) und die Administration der Systeme. Das Zusammenspiel technischer und organisatorischer Prozesse ist in Abbildung 1 dargestellt.

<sup>1</sup> <https://wiki.lsd.f.kit.edu>

<sup>2</sup> <http://bw-fla.uni-freiburg.de/>

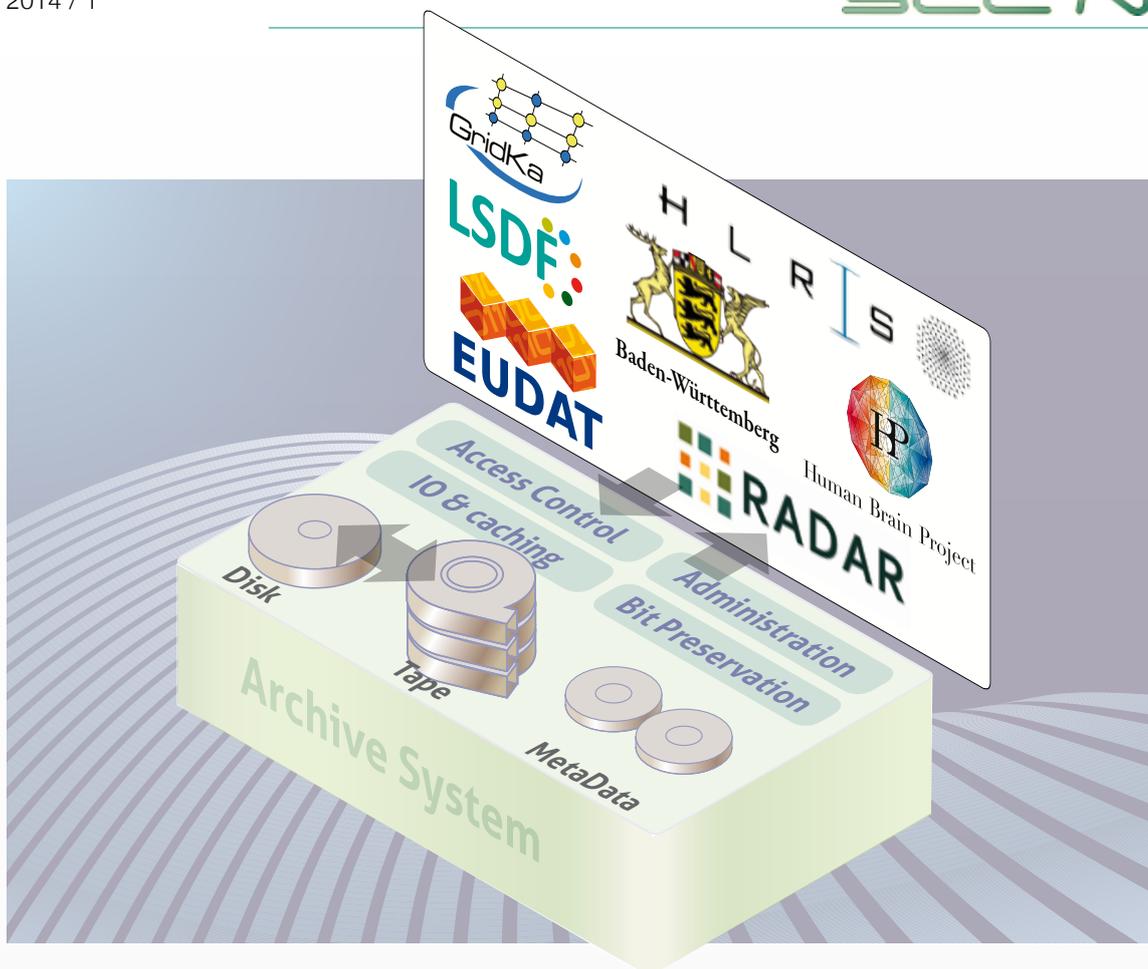


Abbildung 1: Zahlreiche Projekte benutzen das Archiv am SCC.

### Archive und Archivierungsprojekte am SCC

Über 20 Petabyte Daten werden am SCC (Stand Januar 2014) für eine Vielzahl wissenschaftlicher Disziplinen und öffentlicher Einrichtungen auf magnetischen Bändern sicher aufbewahrt. Die gespeicherte Datenmenge am KIT steigt jedes Jahr um zwei bis drei Petabyte und wird auch in Zukunft weiter stark zunehmen. Grund dafür sind beispielsweise genauere Messungen, ein bezahlbarer Zugang zu Instrumenten mit hoher Auflösung und nicht zuletzt gesellschaftliche Anforderungen, beispielsweise bei Daten, die mit öffentlichen Mitteln erhoben wurden. Dazu kommen zunehmend Daten aus Einrichtungen wie Bibliotheken, Museen und öffentlichen Archiven. Derzeit werden am KIT neben einem Teil der LHC-Daten im GridKa, unterschiedliche Datensätze von KIT-Instituten und der Universität Heidelberg, Daten der Badischen Landesbibliothek, des Staatsarchivs Stuttgart, des Bibliotheksservice-Zentrums Baden-Württemberg (BSZ) und des Heidelberger Instituts für Theoretische Studien gGmbH (HITS) archiviert.

Darüber hinaus bietet das SCC Support für Daten-Management und -analyse von KIT-Forschungsdaten sowie für verschiedene Landes-, nationale und internationale Projekte an. So werden in der LSDF, im Rahmen des Helmholtz Portfolios LSDMA<sup>3</sup>, in Kooperation mit verschiedenen Communities und im nationalen Projekt „Research Data Repositorium“ (RADAR<sup>4</sup>) Technologien und Workflows zur verlässlichen Langzeitspeicherung entwickelt und eingesetzt.

Im Rahmen des Projekts RADAR soll eine generische Infrastruktur für die Archivierung und Publikation von Forschungsdaten etabliert werden. Der Schwerpunkt wird auf Forschungsdaten aus Bereichen mit kleineren Datenmengen (small sciences) gelegt, da in diesen Forschungsdaten-Infrastrukturen meist noch fehlen. Die Nachhaltigkeit und Vertrauenswürdigkeit des aufzubauenden Services wird durch eine angestrebte Zertifizierung und einen sich selbst tragenden Betrieb mit einer Kombination aus Einmalzahlungen und institutionellen Angeboten gewährleistet. Das SCC arbeitet im RADAR-Projekt mit dem Leibniz-Institut für Informationsinfrastruktur (FIZ), der Technischen Informationsbibliothek (TIB) und verschiedenen Universitäten als wissenschaftliche Partner zusammen und ist unter anderem bei der Entwicklung von System-Schnittstellen involviert.

Mit der Bewilligung des Landesprojekts „bwDataArchiv“ Anfang 2014 wird die technologische Basis und Infrastruktur für einen Service zur verlässlichen Langzeitarchivierung von wissenschaftlichen und kulturell relevanten Daten für Archive, Bibliotheken und Universitäten des Landes Baden-Württemberg sowie für Partner in bundesweiten und internationalen Projekten ausgebaut. In Kooperation mit dem Höchstleistungsrechenzentrum Stuttgart (HLRS) ist das SCC im Rahmen des Projekts federführend für den Ausbau des digitalen Archivs am KIT zuständig. Ziel ist es, unterschiedlichen Nutzergruppen die Möglichkeit zu bieten, ihre Daten einfach und sicher zu übergeben und die Daten für einen festzulegenden Zeitraum zu archivieren. Im „bwDataArchiv“-

<sup>3</sup> <https://helmholtz-lsdma.de>

<sup>4</sup> Siehe SCC-News 01/2013

Projekt werden etablierte und zuverlässige Techniken zur digitalen Archivierung ausgewählt und nach dem Aufbau in den Betrieb überführt. Die Entwicklung des neuen Services soll bis 2015 soweit fortgeschritten sein, dass Daten vom HLRS aufgenommen und verlässlich archiviert werden können. Im ersten Schritt wird die Beschaffung der erforderlichen Hard- und Software durchgeführt und der initiale Service für das HLRS bereitgestellt. Anhand erster Erfahrungen wird der Dienst im Projekt für weitere Nutzer etabliert und im Hinblick auf Integrität, Nutzbarkeit und Verfügbarkeit im zweiten Schritt weiterentwickelt.

### Logischer und technischer Aufbau des SCC-Archivsystems

Die technische Realisierung der Archiv-Langzeitspeicher erfolgt über einen Massenspeicher, der zum Großteil aus Magnetbändern besteht – einer bewährten Technologie mit einer im Vergleich zu Festplatten sehr viel geringeren Fehlerrate. Große Rechenzentren setzen stark auf diese verlässliche und langlebige Speichertechnologie, die darüber hinaus stetig weiterentwickelt wird<sup>5</sup>. Zudem sind die Anschaffungs- und Betriebskosten von Band-Technologien im Petabyte-Bereich bis zu einem Zehnfachen kostengünstiger als rotierende Platten. Der Nachteil längerer Zugriffszeiten<sup>6</sup> bei den linearen Medien wird durch Vorteile der Langlebigkeit und geringen Kosten für Strom und Kühlung ausgeglichen.

Neben der reinen Datenspeicherung besteht ein Archivspeicher auch aus einer Datenbank, in der die Metadaten gespeichert werden, um die archivierten Daten wieder auffinden zu können. Die Bereitstellung und der Betrieb einer zuverlässigen Datenbank ist ein wesentlicher Bestandteil eines Archivierungssystems. Nach dem „Data Ingest“ liefert der Archiv-Service eine Referenz auf die Daten zurück. Diese Referenz muss mit weiteren administrativen Metadaten in einem zentralen Katalog gesichert werden. Während der Aufbewahrungszeit ist eine regelmäßige Überprüfung der Datenintegrität notwendig. Diese kann aufgrund diverser Fehlerquellen in der Hard- und Software oder durch physikalische Effekte verletzt werden. Das Bandmanagement-System kann mit speziellen Tools die Konsistenz und die Pflege eines Archivs unterstützen. Das Vertrauen in ein Archiv steht im direkten Zusammenhang mit der Zuverlässigkeit der Datenspeicherung. Diese wird durch Service Level Agreements (SLA) definiert, die entsprechende technische Maßnahmen und Prozeduren für ein bestimmtes Niveau vorschreiben.

Zur Verwaltung des Archivspeichers in Kombination mit einer geeigneten Metadaten-Verwaltung und der Bereitstellung von Funktionen, zum Beispiel um die Integrität zu gewährleisten, muss ein allumfassendes System eingesetzt werden. Das SCC hat sich für die Installation und den Betrieb des *High Performance Storage System* (HPSS)<sup>7</sup> der Firma IBM entschieden. Das HPSS ermöglicht das Speichern von Millionen von Dateien bis in den Exabyte-Bereich. Platten und Bandspeicher werden in einem virtuellen File-System zu einem hoch performanten Storage Management-System zusammengeführt, das au-

tomatisch Daten von der Platte und Bändern migriert. Zu den Benutzern von HPSS gehören in Deutschland neben dem HLRS das Deutsche Klimarechenzentrum (DKRZ), die Max-Planck-Gesellschaft (MPG) und der Deutsche Wetterdienst (DWD). Auch in den Vereinigten Staaten ist HPSS in mehreren großen Forschungs- und HPC-Einrichtungen im Einsatz.

Im Projekt „bwDataArchiv“ werden Schnittstellen und Archivspeicher von HPSS für die Entwicklungen in LSDMA, dem RADAR-Projekt und weitere Anwendungen, die eine Plattform für Langzeitspeicher benötigen, bereitgestellt. In den kommenden Jahren wird das SCC das Band-System einrichten und ausbauen, um Simulationsergebnisse aus dem HLRS über LHC-Mesdaten des CERN bis hin zu den Lustre-Dateisystemen am SCC zu speichern.

Jan Potthoff, Jos van Wezel



Abbildung 2: Magnetband-Archiv.  
Foto: Wikimedia Commons, Hannes Grobe

## State project “bwDataArchiv”

SCC extends digital archive for long term storage of research data

With the LSDF and GridKa facilities SCC has gained many years of experience in long term storage of massive amounts of data for multiple scientific disciplines. Data is a valuable good that must be made available for reference and for future research in digital archives. In the recently started state project “bwDataArchiv”, SCC and the High Performance Computing Center Stuttgart (HLRS) will jointly develop a trustworthy archive service for scientific data and data from projects which ran on the HLRS supercomputer. In the course of the project, funded by the State of Baden-Wuerttemberg, SCC will install and use HPSS for reliably managing hundreds of PB data on cost-effective tape based storage systems. In collaboration with researchers of the LSDMA initiative of the Helmholtz Association and the RADAR project, funded by the Federal Ministry of Education and Research, “bwDataArchiv” works on technologies and interfaces for archive access and bit preservation.

<sup>5</sup> <http://www.zurich.ibm.com/news/10/storage.html>

<sup>6</sup> Beim Einsatz moderner Hardware entstehen Latenzen von ein bis zwei Minuten.

<sup>7</sup> <http://www.hpss-collaboration.org/>

# 100 Gigabit/s zwischen KIT und Denver

**Forscher in Europa und Nordamerika werden mithilfe der 100 Gigabit/s-Technologie künftig in den datenintensiven Wissenschaften noch besser zusammen arbeiten können. Auf der internationalen Supercomputing-Konferenz SC13, die Mitte November letzten Jahres in Denver, USA, stattfand, demonstrierte das SCC diese Technologie erfolgreich. Eine Teilstrecke stellte dabei der Verein zur Förderung eines Deutschen Forschungsnetzes (DFN-Verein) bereit.**

Die transatlantische 100-Gbit/s-Netzwerkverbindung mit dem Namen „Advanced North Atlantic 100G Pilot project (ANA-100G)“ von Amsterdam nach New York wird von einem Konsortium aus sechs internationalen Wissenschaftsnetzen für zwölf Monate zur Verfügung gestellt. Die Verbindung soll die Entwicklung fortgeschrittener Netzwerktechnologien fördern und auch datenintensive High-End-Projekte wie die Experimente am LHC in Genf, am Fusionsreaktor ITER in Frankreich und weitere internationale Programme voranbringen. Auf der internationalen Supercomputing-Konferenz SC13 Mitte November letzten Jahres demonstrierte das SCC erfolgreich den Austausch von LHC-Experimentdaten über diese Hochgeschwindigkeitsverbindung. Der DFN-Verein stellte dabei eine 100 Gigabit/s-Strecke vom KIT nach Amsterdam bereit.

„Das Worldwide LHC Computing Grid hat in den ersten Jahren der LHC-Datennahme seine Leistungsfähigkeit bei der Datenauswertung bewiesen. Der 100 Gigabit/s-Transatlantik-Netzwerktest ist ein wichtiger Schritt für den Ausbau dieser verteilten Computing-Umgebung im Hinblick auf die sehr viel höheren Datenmengen der nächsten LHC-Datennahme und künftiger Experimente“, so Prof. Dr. Christian Zeitnitz, stellvertretender Vorsitzender des Komitees für Elementarteilchenphysik (KET).

Neben der Übertragung der LHC-Daten trat das SCC in einem Show Case auch den Beweis an, dass derzeit verfügbare Rechnersysteme die Bandbreite von 100 Gigabit/s tatsächlich füllen und die Daten in Echtzeit verarbeiten

werden. „Dies ist angesichts der Länge der Transatlantik-Verbindung sehr wohl eine technische Herausforderung“, erklärte Prof. Dr. Bernhard Neumair, Geschäftsführender Direktor des SCC.

An der Hochgeschwindigkeitsverbindung sind die Wissenschaftsnetze Internet2 (USA), NORDUnet (Nordeuropäische Länder), ESnet (U.S. Department of Energy), SURFnet (Niederlande), CANARIE (Kanada) und GÉANT (Europa) beteiligt. Die Einrichtung dieser Strecke verdeutlicht einmal mehr, dass Spitzenforschung heute zunehmend auf der Basis großskaliger wissenschaftlicher Daten stattfindet, deren Management und Analyse der Schlüssel zu neuen Erkenntnissen und Innovationen sind. Komplexe Projekte werden dabei in internationalen Kooperationen durchgeführt, bei denen die Forscher einen schnellen und sicheren Zugriff auf die rund um den Globus verteilten Daten erwarten.

Auch Dr. Takanori Hara vom japanischen Forschungszentrum für Teilchenphysik KEK und Computing-Koordinator des Belle II-Experiments, das künftig die Eigenschaften von Bottom Quarks mit höchster Präzision messen wird, gratulierte dem SCC zur erfolgreichen Demonstration der 100 Gigabit/s-Technologie. Dies sei ein Meilenstein auf dem Weg zu einer verteilten Computing-Infrastruktur für das Belle II-Experiment, bei der das KIT mit dem am SCC angesiedelten Grid Computing Centre Karlsruhe (GridKa) eine entscheidende Rolle spielen wird.

Dr. Andreas Heiss, Ursula Scheller, Reinhard Strebler

## 100 Gigabits/s between KIT and Denver

Thanks to the 100 Gigabits/s technology researchers in Europe and North America will be able to cooperate more effectively in data-intensive sciences in the future. At the International Supercomputing Conference SC13 taking place in mid-November of last year in Denver, SCC successfully demonstrated the exchange of LHC experiment data via the high-speed connection between Karlsruhe and Denver. The German National Research and Education Network, DFN, provided a 100 gigabits/s line from KIT to Amsterdam.

The transatlantic 100 Gbits/s network connection named “Advanced North Atlantic 100G Pilot project (ANA-100G)” from Amsterdam to New York will be made available by a consortium of six international science networks for a period of twelve months. The connection is to foster the development of advanced network technologies and to support data-intensive high-end projects, such as the experiments at the LHC in Geneva, at the ITER fusion reactor in France, and in other international programs. Apart from the transmission of LHC data, SCC also succeeded in demonstrating in a showcase that currently available computer systems really fill the bandwidth of 100 gigabits/s and the data are processed in real time. “In view of the length of the transatlantic connection, this certainly is a technical challenge,” Professor Dr. Bernhard Neumair, Managing Director of SCC, emphasized.

# Innovation Award für SCC-Website-Management

**Die Firma OpenText hat das SCC für seine aktuellen Entwicklungen im Bereich Web-Content Management und speziell für die Integration einer semantischen Analyse mit einem Innovation Award im Bereich Innovative Discovery ausgezeichnet. SCC-Mitarbeiter Ulrich Weiß, der mit seinem Team für das OpenText Website Management am KIT verantwortlich ist, hat auf der globalen Kundenkonferenz „Enterprise World 2013“ der Firma im November 2013 in Orlando die Entwicklungen des SCC präsentiert und nahm den Preis entgegen.**

Das SCC betreibt am KIT eine der weltweit größten Installationen des OpenText Website Management-Systems zur Pflege der KIT-Webseiten. Seit Ende 2013 wird Besuchern auf den Internetseiten des KIT eine neue Suche angeboten. Diese bietet die Möglichkeit, die Treffermengen individuell nach verschiedenen Facetten einzuschränken und schnell die gewünschten Informationen im KIT-Internetauftritt, der über 200.000 Seiten enthält, zu finden. Die Suche basiert auf einer semantischen Analyse, deren Möglichkeiten nun auch Redakteuren im Web-Content-Management-System des KIT zur Verfügung stehen. Mit automatischen Zusammenfassungen sowie der Erkennung von Fachbegriffen, Personen und Einrichtungen erhält der Redakteur nützliche Tools, um ein übergreifendes KIT-weites Internetangebot zu gestalten.

Die Suchmaske ist auf allen KIT-Seiten im linken oberen Bereich zu finden und führt eine Volltextsuche auf allen KIT-Internetseiten durch. Nach Eingabe eines Begriffs wird zunächst der Internetauftritt durchsucht, auf dem sich der Suchende gerade befindet. Nutzt ein Besucher die deutschen Webseiten, so wird die Suche zunächst nur in deutschen Dokumenten ausgeführt. Parallel dazu wird der Suchbegriff übersetzt und es wird eine zweite Suche im englischen Internetauftritt angestoßen.

Über Verfeinerungen bzw. Facetten lässt sich die Treffermenge einschränken, so dass sich der Suchende sukzessive der gewünschten Information nähern kann. Derzeit stehen die Facetten *Personen*, *Einrichtungen*, *Themen* und *Dokumenttypen* zur Verfügung, anhand derer die Treffermenge eingegrenzt werden kann.

## Intranet und Internet

Ein großer Pluspunkt der neuen Suche ist die Berücksichtigung der Intranet-Dokumente, die entweder als KIT-intern klassifiziert oder unter [intranet.kit.edu](http://intranet.kit.edu) angeboten werden. Die Suche betrachtet die Herkunft der Suchenden anhand ihrer IP-Adresse und liefert für KIT-Mitglieder alle vorhandenen Treffer, während sie bei externen Besuchern neben der eingeschränkten Trefferanzeige noch den Hinweis auf weitere nur im KIT verfügbare Dokumente liefert.

Die Suche über die Internetauftritte der KIT-Einrichtungen war seit langem nur mit begrenzter Funktionalität vorhanden. Mit Einführung dieser neuen Technologie bietet das SCC nun jedem Anwender einen individuellen Zugang zu den angebotenen Informationen. Die zugrunde liegende Anwendung erlaubt die automatische Aufbereitung der Informationen ohne Mitwirkung von Autoren und zentraler Internetredaktion.

## Optimierungen für Redakteure

Die SCC-Erweiterungen unterstützen die Redakteure der KIT-Webseiten auch bei „lästigen“ Alltagsaktivitäten, die häufig vernachlässigt werden, so zum Beispiel bei der Pflege von HTML-Metadaten wie Meta-Keywords und Meta-Description, die ebenfalls zur Suchmaschinenoptimierung dienen.

Auch eine Implementierung zur einfachen Erstellung von Teasern, beispielsweise auf der Homepage, wurde freigeschaltet. Dazu werden Inhaltsseiten automatisch analysiert und Elemente für den Teaser extrahiert. Während Überschrift und Bild direkt entnommen werden können, wird für den Teaser-Text eine Zusammenfassung des Inhalts generiert. Diese extrahierten Informationen dienen als Vorlage für den Teaser, den die Redakteure dann überarbeiten bzw. auf eine sinnvolle Länge kürzen können. Die anschließende Erstellung und Einbindung des Teasers erfolgt wiederum automatisch.

Darüber hinaus wurde für die Generierung von HTML-Metainformationen bereits ein Vorschlag implementiert, der im Redaktionssystem noch freigeschaltet wird. Als nächster Schritt wird die Erstellung eines Vorschlags für ähnliche Inhalte innerhalb anderer KIT-Auftritte folgen. Auch diese werden anhand der semantischen Analyse aus den KIT-Internetauftritten extrahiert und dem Redakteur als Vorschlag für weiterführende Links angeboten. Die Auswahl der final publizierten Links liegt damit beim Redakteur der Seite.

Mithilfe der semantischen Navigation lassen sich ganze Inhaltsseiten zu beliebigen Themen dynamisch generieren. Dabei können sowohl Webseiten als auch Dokumente oder benutzergenerierte Inhalte aus unterschiedlichen Social Media-Kanälen berücksichtigt werden.

Der Einsatz der semantischen Analyse (OpenText Content Analytics) bietet sich nicht nur für Webseiten an. Ein Fernziel könnte beispielsweise darin bestehen, eine eigene Taxonomie für die KIT-Inhalte zu erarbeiten, mit deren Hilfe alle wissenschaftlichen Publikationen klassifiziert werden könnten. Damit ließen sich automatisch ähnliche Projekte erkennen und Wissenschaftler innerhalb des KIT finden, die in ähnlichen Bereichen forschen.

Ulrich Weiß



## Innovation award for SCC's website management

SCC has won the OpenText Innovation Award for its current developments in the field of web content management and especially for the integration of a semantic analysis. SCC employee Ulrich Weiß, being responsible with his team for the OpenText Website Management at KIT, was awarded the prize in mid-November 2013 in Orlando at the company's international customer conference.

The new solution provides all stakeholders of KIT – from website visitors to intranet users and content management editors – with a personalised, user-friendly and collaborative platform to support their daily work and improve efficiency. Previously labour-intensive manual tasks have been replaced by an automated solution that assigns metadata and supports entity extraction by generating teaser texts for new pages, saving time and reducing error. Visitors are given personalised access to highly relevant information, facilitated by faceted search and related hits – resulting in a more satisfying end-user experience.

Vice President Jon Hunter und CEO Mark Barronechea von der Firma Open Text überreichen SCC-Mitarbeiter Ulrich Weiß Mitte November letzten Jahres den „OpenText Innovation Award“ für aktuelle Entwicklungen im Bereich Web-Content Management. Foto: Open Text



Abbildung 1: Suche mit Facetten zur sukzessiven Verfeinerung von Suchtreffern.



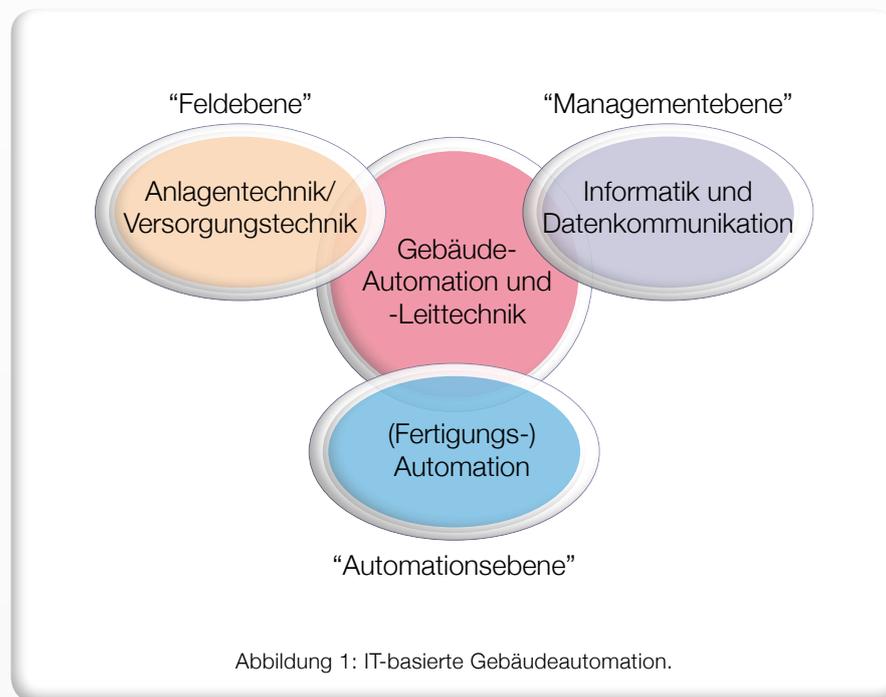
Abbildung 2: Automatische Teaser-Erstellung für jede Seite.



Abbildung 3: Vorschlag zur Erstellung eines Teasers.

# FM und SCC entwickeln Software-Verteilungskonzept zur Gebäudeautomation

Energiewende und Klimaschutz sind erklärte Ziele der Bundesrepublik. Eine Schlüsselrolle spielt dabei der konsequente Einsatz intelligenter Gebäudeautomation (GA). Die Organisationseinheit (OE) Facility Management (FM) betreut am KIT die gesamte Gebäudetechnik und ist entsprechend für die erforderliche Gebäudeautomation verantwortlich. Da die IT-Technologie auch in diesem Bereich einen immer wichtigeren Beitrag leistet, ist für einen reibungslosen GA-Betrieb u.a. eine sichere und effiziente Nutzung der entsprechenden Anwendungen durch Mitarbeiter und externe Partner notwendig. Für diese Anforderung hat FM gemeinsam mit dem SCC eine Lösung für eine effiziente Software-Verteilung erarbeitet und umgesetzt.



Der konsequente Einsatz intelligenter Gebäudeautomation ermöglicht die Steuerung unverzichtbarer Gebäudetechnik wie zum Beispiel Heizungs-, Klima- und Lüftungsanlagen, um Energie effizienter zu nutzen, die Wirtschaftlichkeit zu erhöhen, die Umwelt zu entlasten und die Lebensqualität zu steigern. Gleichzeitig bildet die Gebäudeautomation die Schnittstelle zwischen Gebäuden und zukünftigen intelligenten Stromnetzen (Smart Grid).

Als Gebäudeautomation wird die Gesamtheit von Überwachungs-, Steuer-, Regel- und Optimierungseinrichtungen in Gebäuden im Rahmen des technischen Facility-Managements bezeichnet. Ihre Hauptaufgabe ist die Regelung, Steuerung, Kontrolle und Überwachung der betrieblichen Anlagen, die optimale Betriebsführung und die sich daraus ergebende Sicherstellung eines effizienten Energieeinsatzes. Dabei sind die verschiedenen Gewerke (Heizung, Klima, etc.) übergreifend in die Gebäudeautomation zu integrieren.

Die Gebäudeautomation lässt sich entsprechend der Abbildung 1 funktional in die drei Bereiche *Feldebene*, *Automationsebene* und *Managementebene* unterteilen.

In der Automationsebene sind u.a. die Automatisierungsstationen (AS) in den Gebäuden und in der Managementebene die Management-Bedieneinheit (MBE) zentrale Elemente der GA und maßgeblich für die Steuerungsfähigkeiten und damit den effizienten Einsatz der Energie verantwortlich. Die IT-Technologie bildet heute das verbindende Element zwischen allen Ebenen.

Für ein vollständiges Anlagenmanagement sind auch noch Brandmeldeanlagen, Videotechnik, Zutrittssysteme, Einbruchmeldesysteme und elektroakustische Anlagen in die Betrachtung mit einzubeziehen.

### Kopplung von Betriebs- und Prozessebene

Ein zentrales Element der GA zur Kopplung mit anderen, außerhalb der GA angesiedelten Prozessen ist die MBE. Hier fließen Daten, zum einen aus der Feld- und Automationsebene, aber beispielsweise auch aus der OE-übergreifenden Vorlesungsplanung und der damit verbundenen Hörsaalbelegung oder der Anbindung der Zähleinheiten für das Energie-Management zusammen. Aus dem Abgleich von Soll-/Ist-Zuständen

können dann entsprechende Regelungsaktivitäten in der Feldebene angestoßen werden.

Durch die zunehmende Standardisierung von Protokollen und Schnittstellen in der GA können Aktivitäten zunehmend auch direkt aus Anwendungen, zum Beispiel für das Energie-Management, angestoßen werden. *OLE for Process Control* (OPC; OLE = *Object Linking and Embedding*), *Building Automation and Control Networks* (BACnet) und *Web Services*, sind Beispiele für solche Schnittstellen/Protokolle.

Aufgaben der Management-Bedieneinheit (MBE) sind folgende:

**Alarm/Event-Behandlung**

Im Betrieb der Anlagen wird es immer wieder vorkommen, dass es Zustände innerhalb der Anlagen gibt, die ein definiertes Alarm-Handling erfordern, das über die MBE erfolgen kann. Die Alarmer werden in den AS/MBE-Systemen definiert. Wird ein Alarm aktiv, kann dieser an die MBE gesendet werden. Die MBE nimmt den Alarm entgegen und speichert diesen im Klartext. In der MBE kann der Alarm nun weiter verarbeitet werden. Die weitere Verarbeitung geschieht durch Filter und deren Reaktionen. In der Regel wird der Alarm zur Folge haben, dass eine SMS (Fax, E-Mail, SNMP) an die Rufbereitschaft oder den Notdienst gesendet wird.

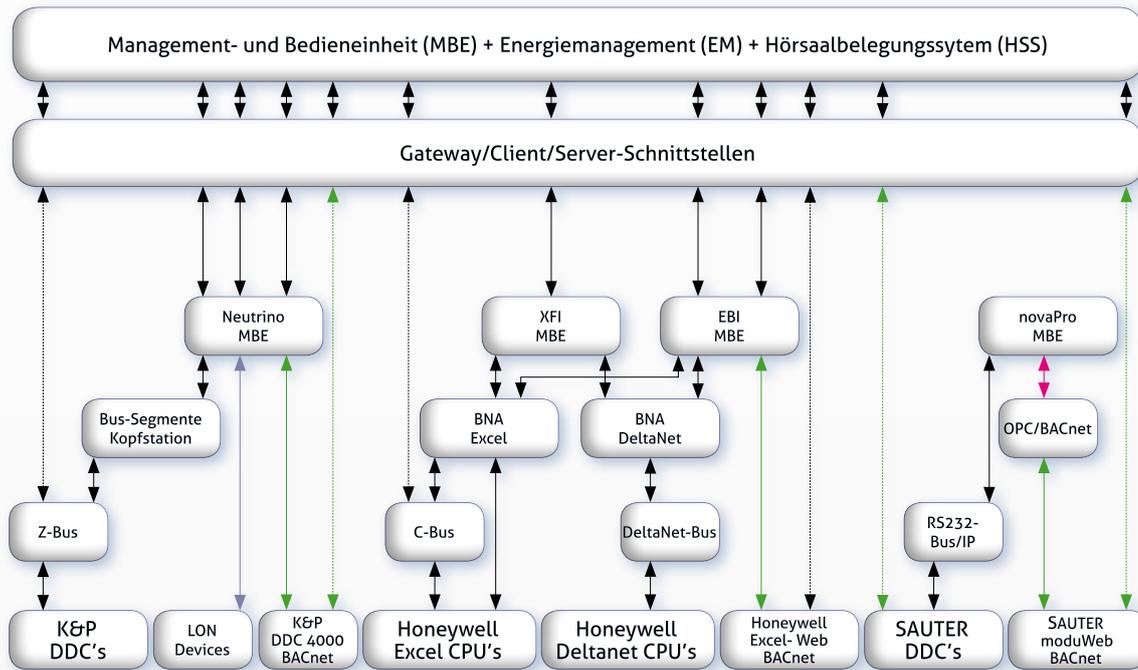


Abbildung 2: Kommunikationsstrukturen der Gebäudeautomation.

**Visualisierung und Bedienung der Anlagen**

Die Anlagen und deren Prozesswerte werden für den Nutzer visualisiert. Die Darstellung der Anlagen entspricht dem technischen Anlagenverständnis der Nutzer. Durch die grafische Darstellung als Anlagenschema und dem Einblenden der Prozesswerte wie zum Beispiel Temperaturen, Zeitprogramme, CO<sub>2</sub>-Werte oder Ventilstellungen an den jeweiligen Objekten werden die Anlagenzustände, Messwerte, Parameter, Ist-, Soll- und Stellwerte auf einen Blick erfassbar.

**Sammeln und Speichern von Messdaten**

Trenddiagramme und historisierte Daten (zum Beispiel Temperaturverläufe oder Betriebszustände) sind unverzichtbare Grundlagen für die Fehlersuche, Betriebsanalyse und Anlagenoptimierung. Die Dokumentation und Analyse des Anlagenbetriebs erfordert vom Betreiber das Speichern der wichtigsten Prozessdaten. Das MBE-System erhält in frei definierbaren Intervallen die in den Automatisierungsstationen (AS) vorgehaltenen Trenddaten und speichert diese in einer Datenbank ab. Die Daten können dann grafisch und tabellarisch abgerufen oder exportiert werden.

**Gebäudeautomation am KIT**

FM betreut die gesamte Gebäudetechnik am KIT. Dazu gehören die Anlagen zur Wärme-, Kälte-, Klima- und Stromversorgung und weitere zur Raum- und Gebäudeautomatisierung. Tabelle 1 liefert dazu Daten und Fakten.

<b>Anzahl der GA-Nutzer: ca. 50 Personen</b>	
<b>Gesamtanzahl der Datenpunkte: ca. 75.000</b>	
<b>Anzahl/Art der Sensoren/Aktoren:</b>	
Digital Input	30.959
Digital Output	2.288
Analog Input	13.503
Analog Output	6.464
<b>Gesamt</b>	<b>53.214</b>

Tabelle 1: GA-Daten und Fakten am KIT.

FM ist entsprechend seiner Aufgaben für die notwendige Gebäudeautomation verantwortlich und möchte für einen reibungslosen GA-Betrieb, den Mitarbeitern und externen Partnern die Nutzung der notwendigen Anwendungen und den Zugriff darauf in einer möglichst effizienten Weise, auch bezüglich des erforderlichen Bereitstellungsaufwands, ermöglichen.

Die Installation und Konfiguration der FM-Arbeitsplatzrechner an den internen und externen GA-Arbeitsplätzen hatten in der Vergangenheit immer wieder erhebliche Kosten und Aufwände erzeugt. Ein sicherer Zugang von außerhalb des KIT konnte überhaupt nicht ermöglicht werden. Für die kosteneffiziente und zukunftsorientierte Bereitstellung und den Zugang zu den GA-Anwendungen wurden daher verschiedene Lösungsansätze geprüft, u.a. der Aufbau einer eigenständig betriebenen Lösung mit eigenem Betrieb von dezentralen Soft- und Hardware-Komponenten. Diese wurde jedoch aus Kostengründen und unzureichender Verfügbarkeit nicht weiter verfolgt, statt dessen wurde ein Ansatz entwickelt, der bestehende Standard-Dienste des SCC nutzt und die weitestgehend eigenständige Verwaltung des Systems durch die Verantwortlichen von FM vorsieht.

Die Aufgabenstellung bestand darin, die Bereitstellung und den damit verbundenen Betrieb der Applikationen, die im Rahmen der Gebäudeautomation Informationen liefern und Wartungsaufgaben unterstützen, neu zu gestalten. Zur Evaluierung der Anforderungen und Umsetzung wurde vom Testing IT-Systems & Support-Team (TIS-Team), einer Gruppe von Studierenden in der Abteilung Systeme und Server des SCC, ein Projekt initiiert. Nach Erörterung verschiedener alternativer Lösungen wurde ein Testsystem erstellt und dieses unter Mitarbeit von FM und deren Partnern getestet.

Danach erfolgte eine Umsetzung der Lösung in ein Produktivszenario. Dabei stellte FM folgende Anforderungen an die Bereitstellung, den Zugang und die Wartung der GA-Applikationen:

- strikte Trennung zwischen Benutzer- und Management-Netz
- GA-Anwendungen:
  - › zentral installiert und gewartet
- Zugriff von innerhalb und außerhalb des KIT-Netzes
- Zugriff von KIT-Mitarbeitern und externen Partnerfirmen
  - › Zugriffsberechtigungen lesen oder schreiben
- Backup des Basissystems und aller Applikationen/Einstellungen
- Disaster Recovery
- hohe Verfügbarkeit und Ausfallsicherheit
- geringe Kosten

Die GA-Nutzer verwenden verschiedene Anwendungen (u.a. die MBE), um im Rahmen des Regelbetriebs aber auch bei Störungen einzugreifen. Abhängig von ihrer Rolle sollen entsprechende Sichten und Zugriffsmöglichkeiten auf diese SW-Werkzeuge bereitgestellt werden. Im Wesentlichen lassen sich die zwei folgenden Gruppen definieren:

- Die interne Benutzergruppe besteht aus FM-Mitarbeitern, der FM/SCC-Rufbereitschaft, Instituten und der Service-Leitstelle.
- Externe Nutzer sind Firmen (GA-Lieferanten) und das Amt für Vermögen und Bau Baden-Württemberg, Karlsruhe (VBA). Diese Gruppe soll für die Projektierung, Wartungsarbeiten, Qualitätskontrolle und Diagnosearbeiten auf GA-Systeme und Anwendungen zugreifen können.

Zu diesen Gruppen müssen in den Systemen korrespondierende Rollen und Zugriffsrechte definiert werden können. Die interaktiven Dienste wie das Energie-Management und das Raum- und Hörsaalbelegungssystem sind über eine OPC-Kopplung an die GA angebunden.

#### Technischer Lösungsansatz

Der gemeinsam erarbeitete Lösungsansatz basiert im Wesentlichen auf zwei technischen Komponenten: einem so genannten *Remote Desktop Services Server* (RDS-Server) einem Produkt der Firma Microsoft, als Frontend-Server und dem sicheren Zugang zu diesem von außerhalb des KIT über VPN (*Virtual Private Network*). Für die Unterstützung organisatorischer Abläufe, wie zum Beispiel die Zuordnung von Berechtigungen für GA-Nutzer oder -Gruppen, wurde das Szenario in die bestehenden Management-Abläufe integriert, die das SCC für das KIT betreibt.

#### Anwendungen und Systeme

Die Software wird zentral auf einem RDS-Server bereitgestellt und als Remote-Applikation (*RemoteApp*) allen berechtigten Nutzern zur Verfügung gestellt. *RemoteApps* sind Applikationen, welche auf dem RDS-Server installiert werden und per Remote Desktop Protocol (RDP) freigegeben werden können. Durch den RDS-Server ist es möglich, Benutzer und das Management-Netz strikt zu trennen. Einziger Verbindungspunkt ist die jeweilige Applikation, die vom Benutzer aufgerufen wird und dann Zugriff auf das GA-Management-Netz hat. Der Server ist außerdem Mitglied des KIT-Active-Directory (KIT-AD), über welches sich die Zugangsberechtigungen setzen lassen. Somit wird die Software zentral verwaltet und kann von jeweils berechtigten Personen genutzt werden.

Die benötigten Server-Ressourcen wurden vollständig in der Virtualisierungsumgebung des SCC basierend auf *vSphere/VMware* (ESX-Cluster) virtualisiert. Das hat nicht nur den Vorteil einer Kostenersparnis bezüglich der Hardware, sondern bietet auch flexible Skalierungsmöglichkeiten. So können die Ressourcen bei einer stufenweisen

Migration der Altsysteme problemlos angepasst werden. Hohe Verfügbarkeit wird durch das Personal und die Infrastruktur (Hardware, redundante Strom- und Kühlsysteme etc.) des SCC sichergestellt. Fachwissen über die Virtualisierung muss so nicht zusätzlich in der Organisationseinheit aufgebaut oder extern eingekauft werden. Eine Insellösung wird somit vermieden und stattdessen das Know-how des SCC verwendet.

### Zugang

Der RDP-Zugriff auf den RDS-Server ist nur aus bestimmten Netzen innerhalb des KIT zugelassen (s. Abbildung 3). So können die Mitarbeiter von FM direkt auf die *Remote Desktop Services* zugreifen und GA-Applikationen

starten darf, kann somit der IT-Beauftragte eigenständig bestimmen und verwalten.

Zu den weiteren genutzten Diensten zählen auch die Zertifizierungsstelle des KIT (KIT-CA), die eine SSL-verschlüsselte Kommunikation mit dem RDS-Server ermöglicht, sowie der zentrale McAfee-Server für den Virenschutz. Das Zusammenspiel der verschiedenen Mechanismen ist in Abbildung 4 dargestellt.

In Planung ist die Verlagerung der kompletten IT-Infrastruktur von der Organisationseinheit FM an das SCC. So sollen nach der Umsetzung des Frontend-Servers auch die Backend-Server (Datenbank-, Management-Server) in der Infrastruktur des SCC betrieben werden. Ebenso

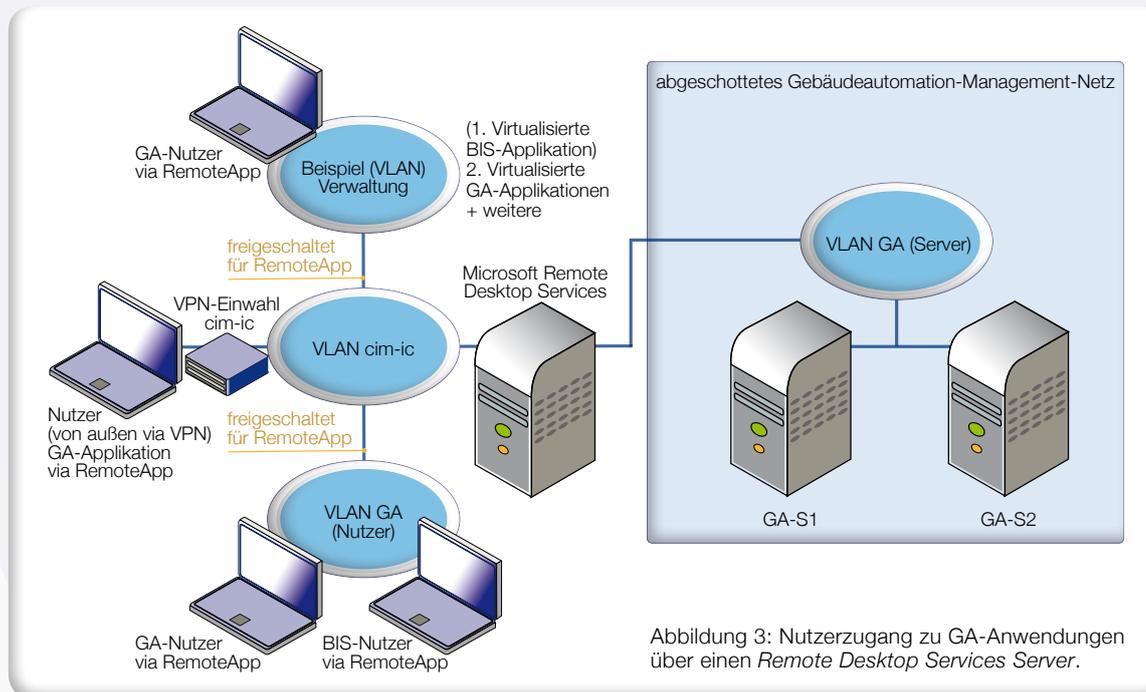


Abbildung 3: Nutzerzugang zu GA-Anwendungen über einen *Remote Desktop Services* Server.

per *RemoteApp* ausführen, um im GA-Management-Netz Aufgaben durchzuführen. Der Zugang außerhalb des KIT-Netzes, zum Beispiel für die KIT-Rufbereitschaft oder externe Wartungsfirmen, erfolgt dabei über einen speziellen VPN-Zugang: nach der Authentifizierung mit einem KIT-Account auf einer speziellen VPN-Webseite kann die *RemoteApp* direkt über diese gestartet werden. Die Nutzerberechtigungen werden dabei über Gruppen gesteuert.

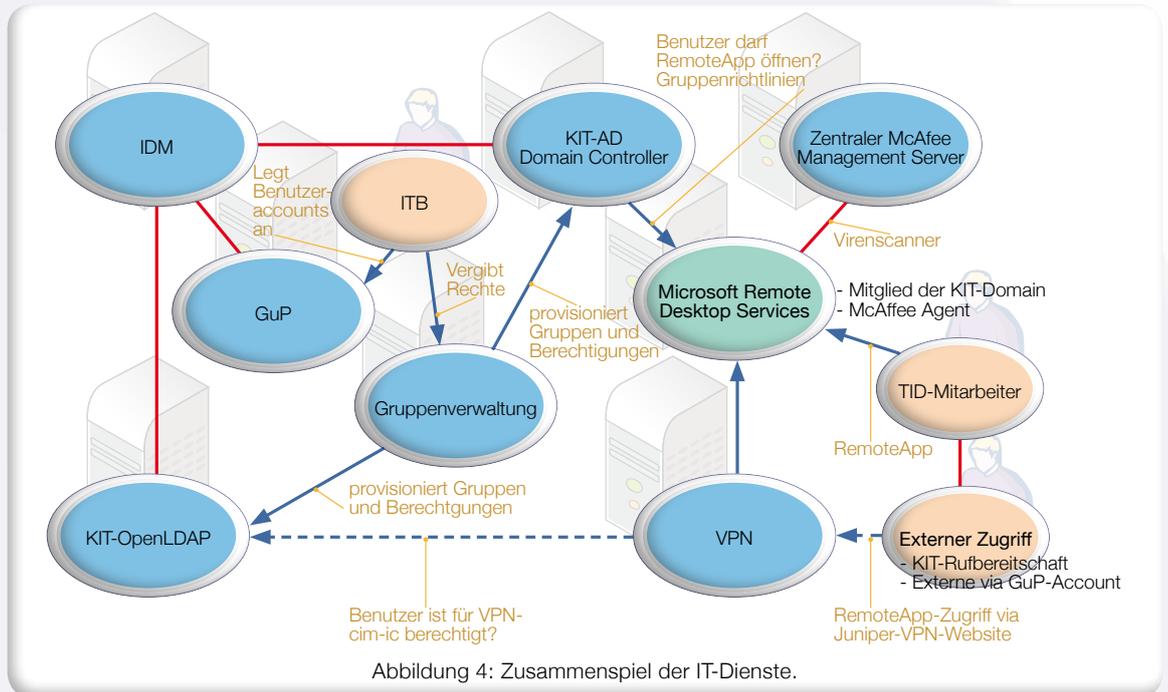
### Management

Die Migration der „alten“ Welt vom Forschungszentrum Karlsruhe und von der Universität Karlsruhe hin zum KIT führte zwangsläufig zur Einführung vieler neuer IT-Workflows und -Dienste. Ein Beispiel dafür ist die Gäste- und Partnerverwaltung, über die der IT-Beauftragte von FM Accounts für externe Partner selbständig anlegen und verwalten kann. Des Weiteren kann er über die Gruppenverwaltung Gruppen anlegen und Benutzer seiner OE den Gruppen zuweisen. Wer sich auf der speziellen VPN-Webseite anmeldet oder die *RemoteApp*

ist mittelfristig die Verlagerung der Management-Ebene von einer FM-„Insellösung“ hin zur zentralen SCC-Infrastruktur vorgesehen.

### Kostenersparnis durch Standard-Technologien

Eine Gebäudeautomation ohne eine entsprechend ausgerichtete IT-Infrastruktur ist heute nicht mehr denkbar. Durch die konsequente Nutzung von IT-Standard-Technologien können bei der Bereitstellung, Fehlerdiagnose und Pflege der Gebäudeautomation erhebliche Kosten eingespart werden. Durch die Nutzung einer sicheren VPN-Remote-Schnittstelle konnten beispielsweise die Kosten für die Störungsbehebung erheblich reduziert werden: unnötige Kosten für An- und Abfahrten und Service-Pauschalen wurden auf ein Minimum reduziert. Bei ca. 900 Rufbereitschaftalarman pro Jahr ist nur bei ca. einem Drittel der Störungen ein Vor-Ort-Einsatz notwendig.



Für die Bereitstellung und den Betrieb der GA-Anwendungen benutzt FM etablierte Standardverfahren und die Infrastruktur des SCC. Eine größtmöglich selbstständige Verwaltung des Servers wurde durch die Nutzung von IT-Diensten (Gäste- und Partnerverwaltung, Gruppenverwaltung, *Microsoft AD*, *KIT-OpenLDAP*, *VPN-Juniper* etc.) erreicht.

Sebastian Hämpel, Denis Künzel,  
Heinz Scharfenberger, Klaus Scheibenberger

## FM and SCC develop software distribution concept for building automation

Energy transition and climate protection are declared goals of the Federal Republic of Germany. A key role plays the consequent application of intelligent building automation. The organisational unit *Facility Management* (FM) services the complete building facilities at KIT and is respectively responsible for the necessary building automation. IT plays an increasingly important role also in this area. Providing necessary applications to staff and external partners in a secure and efficient manner is therefore a must, but also the effort required concerning this matter has to be taken into account.

To meet these requirements FM developed in cooperation with SCC a software distribution concept for building automation.



Rechtsanwalt J. Keuthen, Richter P. Schmid und Rechtsanwalt T. Lösche (v.l.n.r.) erörtern am simulierten Amtsgericht einen Fall zu Dokumentationspflichten in einem privaten Laborbetrieb. Foto: Paul C. Johannes

## Juristische Simulationsstudie erforscht Beweiswert elektronischer Laborbücher

**Im Rahmen des SCC-Projekts „Beweissicheres elektronisches Laborbuch“ (BeLab) wurde Mitte Januar in Kassel eine juristische Simulationsstudie zu Fragen der Archivierung und Aufbewahrung elektronischer Forschungsdokumentationen und deren Beweisbarkeit durchgeführt. Die Studie, die von den Kooperationspartnern an der Universität Kassel geleitet wurde, zeigte insbesondere, dass die im BeLab verwendeten Sicherheitsmechanismen, wie elektronische Signaturen und Zeitstempel, Systemprotokolle, Versionierung und automatisierte Datenprüfungen für Anwälte und Richter grundsätzlich überzeugend sind.**

Bei der Übergabe von Daten an ein digitales Archiv müssen diese auf eventuelle Fehler oder Unstimmigkeiten, die die Qualität der Daten einschränken, überprüft werden. Das BeLab-System ist ein Ergebnis des gleichnamigen Projekts, das die Aufbereitung von Forschungsdaten für eine beweiswerterhaltende Langzeitarchivierung verfolgt (s. Potthoff/Johannes, SCC News 2013/2). Dabei wurde eine generische Schnittstelle konzipiert, um möglichst viele Fachrichtungen abdecken zu können. Kern des Systems bilden die Daten-/Metadatenprüfungen und die anschließende Klassifizierung. In der Prüfung werden die Daten auf Vollständigkeit, Beweiskraft und Tauglichkeit zur Langzeitarchivierung untersucht. Die Klassifizierung gibt dann darüber Auskunft, mit welcher Beweiskraft die Daten im Archiv eingelagert werden. Werden beispielsweise unvollständige Datenreihen übergeben, wird dies dem Benutzer

bekannt gegeben, der dann entsprechend reagieren kann. Der Bericht des BeLab-Systems wird abschließend signiert, um wiederum darlegen zu können, mit welchen Ergebnissen die Daten übernommen wurden.

Der entwickelte Prototyp wurde sowohl aus technischer als auch juristischer Sicht evaluiert. Um zu testen, inwieweit das Konzept in der juristischen Praxis überzeugt, führten die Projektpartner eine juristische Simulationsstudie durch, in der Richter, Rechtsanwälte und Sachverständige die prototypischen Lösungen auf den prozessualen Prüfstand stellten. Die Simulation war notwendig, da zwar immer mehr Forschungseinrichtungen und Labore von der papiernen zur elektronischen Forschungsdokumentation übergehen, aber elektronische Laborbücher hinsichtlich ihres Beweiswerts bislang vor Gericht noch nicht überprüft

worden sind. Daher gibt es in der bisherigen Rechtspraxis hinsichtlich des Beweiswerts wenig Erfahrung. Das vom SCC entwickelte prototypische BeLab-System kann nur beschränkt real-juristisch erprobt werden. Um in einer solchen Situation dennoch praktische Erfahrungen zu gewinnen und die erforderliche Bewertungssicherheit zu erlangen, entwickelte die Projektgruppe verfassungsverträgliche Technikgestaltung (provet) der Universität Kassel eine juristische Simulationsstudie, in der unter Einbeziehung von Experten zukünftige Fallgestaltungen unter möglichst realen Bedingungen mit Prototypentechnik „durchgespielt“ werden. Die Methode der Simulationsstudie hat sich als geeignetes Mittel erwiesen, um praktische Erfahrungen im Umgang mit künftiger Technik und den damit zusammenhängenden Rechtsfragen zu gewinnen. Sie kann genutzt werden, um technische Gestaltungen und Rechtsnormen zu überprüfen und zu korrigieren, bevor Fehlentwicklungen irreversible Schäden verursachen können.

Die Simulationsstudie des Projekts BeLab verfolgte das Ziel, den prozessualen Beweiswert elektronischer Laborbücher zu bestimmen. Dementsprechend wurden die vom Projekt entwickelten Konzepte und Prototypen in acht Gerichtsfällen mit fiktiven Fallgestaltungen von zwei Richtern, vier Rechtsanwälten und zwei technischen Sachverständigen praktisch überprüft. Die Anwälte erhielten die Aussagen und Beweismittel ihrer Mandanten und reichten ihre Schriftsätze bei den Richtern ein. Die Vorbereitung der mündlichen Verhandlungen dauerte etwa zwei Monate, die mündlichen Verhandlungen einen Tag. Die Richter, Anwälte und Sachverständigen waren in ihren Prozesshandlungen frei und nur den Regeln ihrer Profession und den Prozessordnungen unterworfen. Die Fälle waren echten Fällen nachgebildet und betrafen das Zivil- und Verwaltungsrecht. Sie waren hinsichtlich Rollen, Behauptungs- und Beweislast unterschiedlich. In allen Gerichtsverfahren waren elektronische Laborbücher und das BeLab-System beweisheblich und damit von streitentscheidender Bedeutung.

Zum Beispiel wurde in einem der Fälle dem Kläger die Doktorwürde von seiner Universität entzogen. Diese zweifelte die Richtigkeit und Vollständigkeit der Daten im elektronischen Laborbuch des Klägers an. Die eingesetzte Prüfungskommission befand das elektronische Laborbuch als nicht „sicher“ genug und monierte explizit, dass die

Aufzeichnungen nicht auf Papier geführt wurden. Der Wissenschaftler hatte jedoch das BeLab-System zur Sicherung seiner Daten eingesetzt.

In einem anderen Fall wurden die elektronische Dokumentation und die Langzeitarchivierung allgemein stärker unter die Lupe genommen. Im simulierten Rechtsstreit stellte die fiktive Freie Forschungsgemeinschaft e.V. (FFG) einem Wissenschaftler Personal- und Sachmittel zur Verfügung. Die FFG prüfte dessen Abschlussbericht und die übersandten Daten und kam zu dem Schluss, er habe unwissenschaftlich gearbeitet. Daher will sie ihm den Rest der bewilligten Fördermittel nicht zahlen. Der Wissenschaftler nutzt das BeLab-System und klagt auf dem Zivilgerichtsweg.

Durch die konstruierten Rechtsstreitigkeiten sollten u. a. die folgenden Fragen adressiert werden: Genügt die Verwendung der eLab/BeLab-Kombination der arbeitsrechtlichen Verpflichtung und grundsätzlich den Anforderungen der guten wissenschaftlichen Praxis, Best Practices und Good Laboratory Practice? Ist das eLab/BeLab genauso sicher, wenn nicht sogar beweissicherer, wie ein normales elektronisches Laborbuch? Können BeLab-Daten dazu benutzt werden, strafrechtlich relevante Manipulationen am eLab zu beweisen? Kann mit den Daten aus einem eLab die Urheberschaft von Laborbucheintragungen bewiesen werden? Ist die Plausibilitätskontrolle des BeLab ein Indiz bei der Beweiswürdigung?

Erkenntnisgewinn konnten sowohl aus der Auswertung der von den Rechtsanwälten im schriftlichen Vorverfahren eingereichten Schriftsätze und der Urteile der Richter gewonnen werden, als auch aus der Beobachtung der Simulationsteilnehmer während der mündlichen Verhandlungen. Als Beweismittel dienten die im Rahmen des Projekts entwickelten Schnittstellen, welche die vom Fall abhängigen Daten und Auswertungen zur Verfügung stellten. Auf diesem Wege sollte die Plausibilität der Ansprüche entschieden werden können. Zusätzlich wurden Gutachten von Sachverständigen eingeholt, die die Plausibilität und Sicherheit des Systems darlegen sollten. Abschließend wurden die Gerichtsverhandlungen Mitte Januar 2014 in Kassel durchgeführt. Sie wurden in zwei nachgestalteten Gerichtssälen durch die Richter und dem Fall zugewiesene Anwälte abgehalten. Zusätzlich wurden externe Sachverständige und Zeugen, wie der fiktive Dienstanbieter von

BeLab, geladen und gebeten, zu den Fällen Stellung zu nehmen bzw. in den Zeugenaussagen die Arbeitsweise des Systems darzustellen. Hier wurde intensiv die Vorgehensweise zur Sicherung der Integrität und Authentizität der Daten, aber auch die allgemeine Vorgehensweise in der Langzeitarchivierung erfragt.

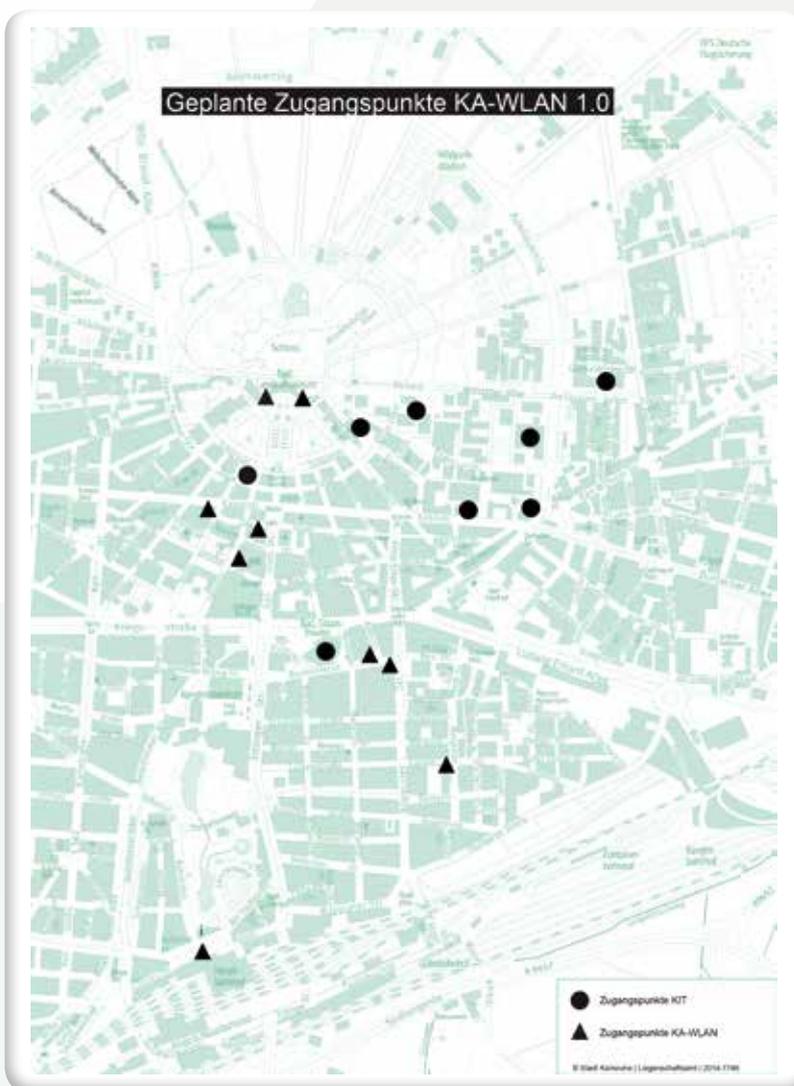
In den simulierten Fällen über elektronische Laborbücher konnten Erkenntnisse zu ihrer Eignung zur Beweisführung und zur Anwendbarkeit der §§ 371, 371a ZPO auf sie gewonnen werden. Im Allgemeinen wurde die Eignung des BeLab-Systems als Beweismittel bestätigt. Dagegen kann sowohl aus den Urteilen der Richter als auch dem Handeln der Rechtsanwälte abgeleitet werden, dass ungesicherte elektronische Forschungsdokumentationen mit erheblichem Misstrauen betrachtet werden, da sie besonders manipulationsanfällig sind. Die Studie zeigte diesbezüglich insbesondere, dass die im BeLab verwendeten Sicherheitsmechanismen, wie zum Beispiel elektronische Signaturen und Zeitstempel, Systemprotokolle, Versionierung und automatisierte Datenprüfungen für Anwälte und Richter grundsätzlich überzeugend sind. Sie sollten jedoch besonders nachvollziehbar und transparent gestaltet werden und daher Algorithmen und Parameter soweit wie möglich offen legen. Gezeigt hat sich auch, dass eine möglichst frühzeitige beweiskräftige Sicherung elektronischer Roh- und Primärdaten von entscheidender Bedeutung für ihre Eignung als Beweismittel ist.

Jan Potthoff, Paul C. Johannes

### Judicial simulation study investigates probative value of electronic laboratory notebooks

The BeLab-System has been designed to secure the probative value of electronic laboratory notebooks (ELN) concerning legal disputes. In mid-January a judicial simulation study was conducted as a proof of concept towards the completion of the project. To this end, several fictional legal cases concerning ELN and BeLab were developed. They were modelled after real cases and reflected possible areas of legal matter concerning ELN, e.g. copyrights, health and food safety inspections and rules of good scientific practice and scientific misconduct. These fictional cases were then tried by real lawyers before real judges. As in real cases arguments were exchanged in writing before hearings were held and as in real proceedings, witnesses and experts were invited to give their opinions and to present more details about the internal processes of the BeLab system. At the end the study showed that the safety mechanisms being used by BeLab are basically convincing to lawyers and judges.

## KIT-WLANs bald auch in der Stadt



In Karlsruhe wird es auf ausgewählten Plätzen in der Innenstadt bald einen kostenfreien Zugang zum Internet geben. Der Gemeinderat hat einstimmig beschlossen, dass bis Mai an mehreren zentralen Stellen so genannte „Hotspots“ eingerichtet werden sollen, an denen Bewohner und Besucher mit Laptops oder Mobiltelefonen kostenlos im Internet surfen können. Mit dem so genannten KA-WLAN will die Stadt insbesondere ihrer Bedeutung als Wissenschaftsstandort gerecht werden.

Im Rahmen des Projekts wurde auch eine Zusammenarbeit mit dem KIT vereinbart, das bereits seit mehr als zehn Jahren eigene Hotspots auf dem Campusgelände betreibt. Auf diese Weise soll die technische Infrastruktur gegenseitig genutzt werden. Mitglieder des KIT werden künftig also genau so wie auf dem Campus im Stadtgebiet WLAN-Verbindungen aufbauen können, da alle auf dem Campus vorhandenen WLANs (SSIDs) wie „vpn/web/belwue“, „eduroam“ und „wkit-802.1x“ abgestrahlt werden. Nach aktueller Planung sollen zunächst am Bahnhofplatz, Friedrichsplatz, Marktplatz, Stephansplatz sowie im Schlossgarten entsprechende Hotspots eingerichtet werden.

Im Gegenzug soll an exponierten Stellen auf dem Campus, wie den Hochhausdächern des KIT, Mensa, Bibliothek und International Department sowie am Staatstheater und an der Musikhochschule für die WLAN-Nutzer der Stadt ein zusätzliches WLAN (SSID) „KA-WLAN“ installiert werden.

Wilhelm Fries

### KIT's WLANs soon in the city centre

KIT and the City of Karlsruhe intend to share their wireless infrastructure, so that KIT members in future will be able to access the wireless infrastructure of the City of Karlsruhe and vice versa. In the first step hotspots are planned to be installed at the Bahnhofplatz, Friedrichsplatz, Marktplatz, Stephansplatz as well as in the palace garden.

## BW-Notebook 3 für alle Hochschulen in Baden-Württemberg

Im Auftrag des Landes Baden-Württemberg hat die Universität Freiburg handelsübliche PC-Systeme zum Einsatz an den Hochschulen des Landes ausgeschrieben: Auftragsgegenstand war die Lieferung von Standard-Notebooks für wissenschaftliche Anwendungen und Lehrbetrieb sowohl für die Uni Freiburg als auch für die bezugsberechtigten Hochschulen des Landes Baden-Württemberg. Den Zuschlag der EU-weiten Ausschreibung hat das Systemhaus Bechtle AG erhalten.

Als Ergebnis wird der Lenovo ThinkPad T440s für 729 Euro inklusive Mehrwertsteuer angeboten. Das Angebot richtet sich ausschließlich an die bezugsberechtigten Hochschulen des Landes Baden-Württemberg. Es gilt sowohl für Neu- als auch Ersatzbeschaffungen, wobei

ein Ersatz in der Regel erst nach mindestens vier bis fünf Jahren in Betracht kommt. Der BW-Notebook 3 darf nur zu dienstlichen Zwecken bestellt werden. Privatbestellungen sind zwingend ausgeschlossen.

(red)

### Weitere Informationen

<http://www.bw-pc.uni-freiburg.de/bwnb>



### BW-Notebook 3 for all universities in Baden- Wuerttemberg

On behalf of the State of Baden-Wuerttemberg the University of Freiburg has invited tenders for notebooks intended to be used at the state's universities. As a result the Lenovo ThinkPad T440s is offered for 729 Euro including value-added tax.

## IHK-Technologiefabrik zeichnet da-cons GmbH aus

**Im Rahmen des Wettbewerbs „Innovative Ideen 2013 gesucht“ hat die IHK-Technologiefabrik in Karlsruhe im Dezember 2013 die da-cons GmbH, eine Ausgründung des KIT, ausgezeichnet. Mit dem „da-cons-system“ belegte sie den mit 2000 Euro dotierten dritten Platz. Die Firma hat sich in Zusammenarbeit mit dem SCC und dem Buchmann Institute for Molecular Life Sciences der Goethe-Universität Frankfurt auf die neuesten Technologien in der automatischen Auswertung, Archivierung und Visualisierung großer Bilddaten spezialisiert und bietet dieses Know-how als Dienstleistung an.**



Im Rahmen des Wettbewerbs „Innovative Ideen 2013 gesucht“ der IHK-Technologiefabrik in Karlsruhe belegte die da-cons GmbH im Dezember 2013 den mit 2000 Euro dotierten dritten Platz. Foto: IHK-Technologiefabrik Karlsruhe

Moderne Mikroskope, wie sie beispielsweise in der Biotechnologie eingesetzt werden, machen kleinste Details sichtbar. Bei der hochaufgelösten Aufnahme von 3D-Objekten wie Zellen oder Organen oder von dynamischen Prozessen wie Bewegungen und Entwicklungen im Raum entstehen enorme Datenmengen. Das „da-cons-system“ ermöglicht eine effiziente und individuelle Auswertung der Rohdaten, die unter anderem das Erkennen und Vermessen von Größen, Volumen, Formen und Abständen umfasst. Dadurch lässt sich die Qualität der Untersuchungen verbessern, etwa bei der Erforschung von Krankheiten oder der Entwicklung von Arzneimitteln oder Kosmetika. In Kooperation mit Spezialisten aus Biotechnologie und Datenakquisition entwickelt da-cons Komplettlösungen zur Probenaufnahme, Aufbereitung, Analyse und Visualisierung in Form von Fotos und Videos.

Im Rahmen des Innovationswettbewerbs suchte die IHK-Technologiefabrik technische Ideen, Lösungen und Konzepte, die einen hohen Innovationsgrad und Kundennutzen aufweisen. Vier Unternehmen wurden insgesamt ausgezeichnet. Der mit 5000 Euro dotierte erste Platz ging an die Firma ColourfulBit für die Software „Kinetic Space“, die einen kostengünstigen Zugang zu dreidimensionaler Gestenerkennung und -analysetechnologie ermöglicht. Den zweiten Platz mit einem Preisgeld von 3.000 Euro erreichte das Unternehmen autoVimation mit der Entwicklung von peltier-klimatisierten Spezialkammergehäusen für extreme Bedingungen. Einen Sonder-

preis in Höhe von 1.000 Euro, gestiftet durch den Beirat der Technologiefabrik, erhielt das Konzept des „Water Harvester“ für eine kostengünstige Lösung zur sauberen Trinkwassergewinnung.

Die Entscheidung über die Gewinner wurde durch eine achtköpfige Wettbewerbsjury, bestehend aus dem Beirat der Technologiefabrik, getroffen. Die Preisübergabe an die Sieger erfolgte durch die Beiratsmitglieder und Juroren Hans Hubschneider und Robert W. Huber. Bei der Wettbewerbsausschreibung konnten sich Unternehmer, Gründer, Forscher, Erfinder und Studierende bewerben. Der Wettbewerb wurde unterstützt durch das KIT, die Hochschule Karlsruhe - Technik und Wirtschaft und durch die Duale Hochschule Baden-Württemberg Karlsruhe mit dem Studiengang Unternehmertum.

(red)

### Weitere Informationen

<http://www.da-cons.de>

### da-cons GmbH awarded

Within the framework of an innovation competition in December 2013 the IHK-Technologiefabrik in Karlsruhe awarded the da-cons company, a spin-off of KIT, with the third prize of 2000 Euro. The company, which cooperates closely with SCC, is specialised in services for the analysis, visualisation, and archiving of large multidimensional image data. The system developed by da-cons allows for an efficient and individual evaluation of raw data, including the recognition and measurement of sizes, volumes, shapes, and distances. Quality of research can be improved, for instance, when diseases are studied or drugs or cosmetics are developed. In cooperation with specialists in biotechnology and data acquisition, da-cons develops complete solutions for sample uptake, processing, analysis, and visualization in the form of photos and videos.

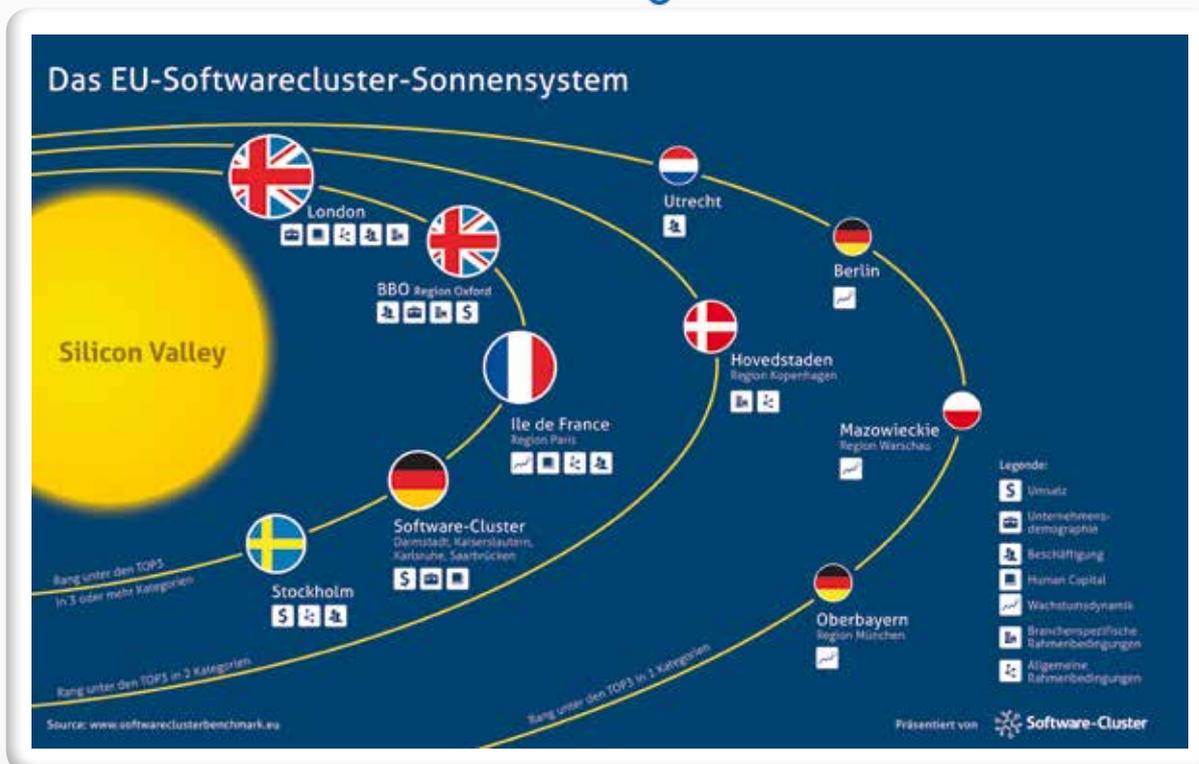
## Studie zu Europas Software-Zentren Software-Cluster im Vergleich

Anders als in den USA, wo das Silicon Valley der wichtigste Ort ist, an dem Software-Innovationen entstehen, die dann weltweit vertrieben werden, sind die Ballungszentren der europäischen Softwareindustrie über den ganzen Kontinent verteilt. Was sind die Stärken und Schwächen der europäischen Software-Cluster, wo stehen sie im Vergleich zum Silicon Valley und welche Maßnahmen können ergriffen werden, um ihre Wettbewerbsfähigkeit zu verbessern? Um Antworten auf diese Fragen zu erhalten, hat der deutsche Software-Cluster, an dem auch das KIT beteiligt ist, die Studie „EU Softwarecluster-Benchmark 2013“ beauftragt.

Europäische Software-Cluster sind Individuen: Sie haben deutlich unterschiedliche Profile in Bezug auf Wachstum, Forschungs- und Entwicklungsaktivitäten, Spezialisierung und wirtschaftlichen Erfolg. Dies ist ein Ergebnis der Studie, die vom Fraunhofer-Institut für System- und Innovationsforschung (ISI) in Karlsruhe durchgeführt wurde. Sie untersucht detailliert 15 europäische Ballungszentren der Software-Industrie und hebt die Identitäten dieser europäischen Silicon Valleys hervor. Dabei wird eine Aufteilung in zwei Arten von Clustern deutlich: einerseits solche, die groß sind, und andererseits solche, die „hungrig“ sind und dynamisch wachsen. Das Silicon Valley in den USA kombiniert beide Attribute: es ist groß und wächst weiterhin. Während einige der europäischen Cluster in bestimmten Bereichen international konkurrenzfähig sind, kommt in der Gesamtsicht keines von ihnen auch nur in die Nähe des Silicon Valley.

Die Studie analysiert die europäischen Software-Cluster anhand von sieben Kategorien: Beschäftigung, Umsatz, Wachstumsdynamik, Humankapital, Unternehmensdemographie, allgemeine und branchenspezifische Rahmenbedingungen. Von den 15 untersuchten Software-Clustern erreichen nur fünf Bestnoten in mehreren Kategorien: London, Berkshire, Buckinghamshire und Oxfordshire (BBO), Île-de-France, Stockholm sowie der deutsche Software-Cluster. Während der Software-Cluster den höchsten Umsatz erreicht, erzielt London Bestnoten in Bezug auf Humankapital und allgemeine Rahmenbedingungen und ist Europas globaler Knotenpunkt für IT Services und Venture Capital. Île-de-France ist ein noch immer wachsender Software-Riese mit einer starken F&E-Infrastruktur, und Stockholm gelingt es, einen ausgewogenen hoch spezialisierten Cluster mit einem starken Umsatz zu verbinden. Cluster,

## Software-Cluster





# Software-Cluster

die ansonsten im Vergleich weiter hinten liegen, können in einzelnen Aspekten andere Cluster übertreffen: Die Region Kopenhagen (Hovedstaden) verfügt über besonders gute branchenspezifische Rahmenbedingungen, die Region Warschau (Masowien) zeichnet sich durch ihre Wachstumsdynamik aus.

Folgende Softwarecluster wurden in der Studie untersucht: Berlin, Oberbayern, der deutsche Software-Cluster rund um Darmstadt, Kaiserslautern, Karlsruhe, Saarbrücken und Walldorf, London, Berkshire, Buckinghamshire und Oxfordshire, Île de France, Region Kopenhagen, Region Helsinki, Stockholm, Madrid, Lombardei, Utrecht, Region Budapest Region, Region Warschau, Prag. Die Ergebnisse der Studie sollen helfen, Maßnahmen zur Verbesserung der Wettbewerbsfähigkeit der europäischen Softwarecluster zu identifizieren. Vertreter mehrerer Cluster haben in diesem Zusammenhang weitere Schritte skizziert, die Politik und Industrie ins Auge fassen sollten. Sie kamen zu der Einschätzung, dass ein stärkeres Lernen der europäischen Softwarecluster voneinander helfen könnte, gute Praktiken zu identifizieren, die aufgrund des gemeinsamen politischen und kulturellen Kontextes der EU möglicherweise einfacher zu implementieren sind als US-Praktiken. Obwohl die Kluft zum Silicon Valley immer noch sehr groß ist, sollten die europäischen Softwarecluster individuelle Wachstumspfade entwickeln, die beim lokalen und regionalen Kontext ansetzen, beispielsweise bei der Industriestruktur der jeweiligen Region.

(red)

## Weitere Informationen

<http://www.softwareclusterbenchmark.eu>

## Study on Europe's software centres

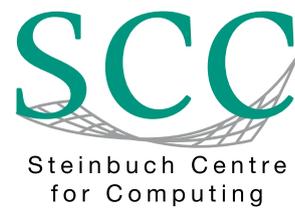
### Software clusters compared

Other than in the US where the Silicon Valley is the main place to create software innovations and distribute them worldwide, the hotbeds of the European software industry are dispersed all over the continent. What are the individual strengths and weaknesses of the European software clusters, where do they stand in comparison to the Silicon Valley and which measures could be taken to improve their competitiveness? To answer these questions, the German Software-Cluster, in which also KIT is involved, commissioned the study "EU Software Cluster Benchmark 2013" conducted by the Fraunhofer Institute for Systems and Innovation Research (ISI) in Karlsruhe.

## Further information

<http://www.softwareclusterbenchmark.eu>





Karlsruhe Institute of Technology (KIT)  
Steinbuch Centre for Computing (SCC)  
76131 Karlsruhe  
Tel: 0721/608-43754 oder -25601  
E-Mail: [contact@scc.kit.edu](mailto:contact@scc.kit.edu)