

# SCC *news*

Steinbuch Centre for Computing

Ausgabe 1 | 2017



**Deutscher Rechenzentrumspreis – Erster Platz für Energieeffizienz**  
*Deutscher Rechenzentrumspreis – First Place for Energy Efficiency*

**Simulation von turbulenter Verbrennung**  
*Simulation of Turbulent Combustion*

**Neuer Self-Service – Freigabe persönlicher Daten**  
*New Self-Service – Conferring Access to Personal Data*

*Liebe Leserinnen und Leser,*

der Sommer in Karlsruhe beschert uns wieder einmal Temperaturen von über 35 Grad Celsius. Vielleicht wird auch dieses Jahr in unserer Region ein Hitzerekord aufgestellt. Apropos Hitze: Bei diesem Wetter kann man sich erst recht nicht vorstellen, dass der am SCC betriebene Hochleistungsrechner mit heißem Wasser gekühlt wird. Unter anderem wegen seines ausgeklügelten Kühlkonzepts erhält der Neubau des Forschungshochleistungsrechners ForHLR II Ende April den ersten Platz beim Deutschen Rechenzentrumspreis in der Kategorie „Neu gebaute energie- und ressourceneffiziente Rechenzentren“ (Seite 28). Die Titelseite dieser Ausgabe zeigt einen Teil der Kühltechnik des ForHLR II. Den Preis stolz in Händen hält Professor Rudolf Lohner (Energie-Effizientes Computing) – rechts im Bild.



„Heiß her“ geht es auch in einer Forschungsarbeit zur Simulation von turbulenten Verbrennungsvorgängen. Aber nicht nur die Temperaturen spielen hier eine Rolle sondern insbesondere die Geräusche, der sogenannte Verbrennungslärm. Mittels geeigneter Computersimulationen kann herausgefunden werden, wie man diesen Lärm deutlich verringern kann. Die Berechnungen sind dabei so komplex, dass dafür ein Supercomputer, wie etwa der ForHLR am KIT, benötigt wird (Seite 18).



Auch beim Schutz von persönlichen Daten gibt es immer wieder mal hitzige Diskussionen und es braucht „kühle Köpfe“ für die Entwicklung von geeigneten IT-Lösungen und -Diensten. Diese sollen den Anforderungen an die Zusammenarbeit in einem internationalen Umfeld gerecht werden, die internen Arbeitsabläufe nicht behindern und darüber hinaus die Selbstbestimmungsrechte der Mitarbeiter des KIT nicht aus dem Auge verlieren. Mit einem neuen IT-Self-Service leistet das SCC dazu einen zentralen Beitrag (Seite 9).

Viel Freude beim Lesen.

*Bernhard Neumair, Achim Streit*

*Dear reader,*

The summer in Karlsruhe once again granted us with temperatures above 35 degrees Celsius. Perhaps this year again there will be a heat record put up in our region. In regards to heat: With this kind of weather, it is hard to imagine that the SCC powered high performance computer is being cooled with hot water. Amongst other things due to its sophisticated cooling concept the new building of the HPC system ForHLR II receives at the end of April the first place with the German Data Centre Prize in the category "newly-built energy and resource efficient data centres" (page 28). The title page of this issue features a part of the cooling technology of the ForHLR II. Professor Rudolf Lohner (to the right of the picture) proudly displays the prize.

It also heats up in a research work to simulate turbulent combustion processes. Not only do the temperatures play a role in this but especially the noises, the so-called combustion noise. Through the use of appropriate computer simulation, one can find out how to significantly reduce this noise. The calculations for that are so complex that it calls for the use of a Supercomputer like the ForHLR at KIT (page 18).

In regards to the protection of personal data, there are always some heated discussions and cooler heads are needed for the development of appropriate IT solutions and services. These have to justify the demands for collaboration within an international environment, do not hinder internal workflow and beyond that does not lose sight of the individual desires of the KIT employees. With a new IT-Self-Service, the SCC makes a central contribution (page 9).

Enjoy reading.

*Bernhard Neumair, Achim Streit*

# Inhaltsverzeichnis

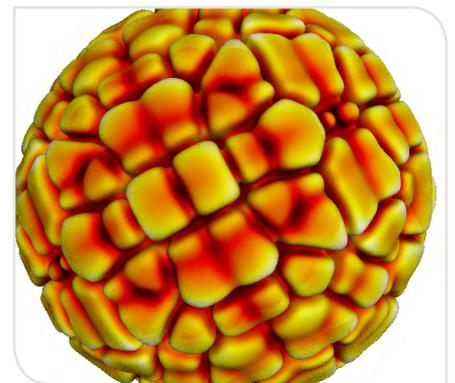
## DIENTE UND INNOVATION

- 04 KITnet wird zur High-Speed-Autobahn
- 06 Der zentrale Virenschutzdienst am KIT
- 08 Neues Onlinespeichersystem für GridKa
- 09 Neuer Self-Service – Freigabe persönlicher Daten
- 12 Nutzertreffen zu Hochleistungsrechnen und datenintensiven Wissenschaften



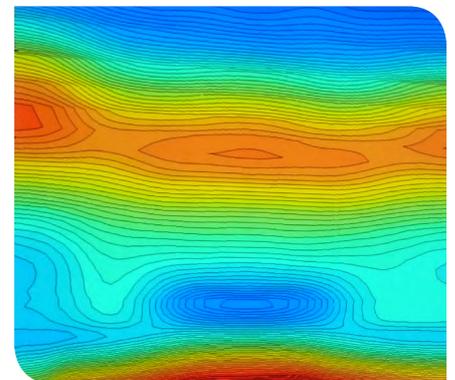
## FORSCHUNG UND PROJEKTE

- 13 Proteinstruktur mittels Big Data vorhergesagt
- 14 Architektur für die standortübergreifende Betriebsüberwachung und -analyse in bwCloud
- 18 Numerische Simulation turbulenter Verbrennung auf Hochleistungsrechnern des SCC
- 21 bwITsec – eine kooperative IT-Sicherheitsstruktur entwickelt sich
- 22 EU Projects: Contributions from SCC at a Glance



## STUDIUM UND WISSENSVERMITTLUNG

- 24 In vivo – In vitro – IN SILICO
- 26 Simulierte Welten: Schülerinnen und Schüler präsentieren ihre Projekte



## VERSCHIEDENES

- 27 Neue Helmholtz-Nachwuchsforschergruppe am SCC
- 28 Supercomputer des KIT in jeder Hinsicht hervorragend
- 30 Tag der offenen Tür am KIT – Impressionen aus dem SCC
- 31 Impressum & Kontakte



## KITnet wird zur High-Speed-Autobahn

Leistungsfähigere Server- und Client-Systeme bewirken steigende Auslastungszahlen im KITnet. Um auch in der Zukunft den steigenden Anforderungen Rechnung zu tragen, plant die Netzabteilung des SCC nach der Hochrüstung der Internetverbindung nun die nächste Geschwindigkeitsdekade im Intranet des KITnet.

*Reinhard Strebler*

### Das KITnet

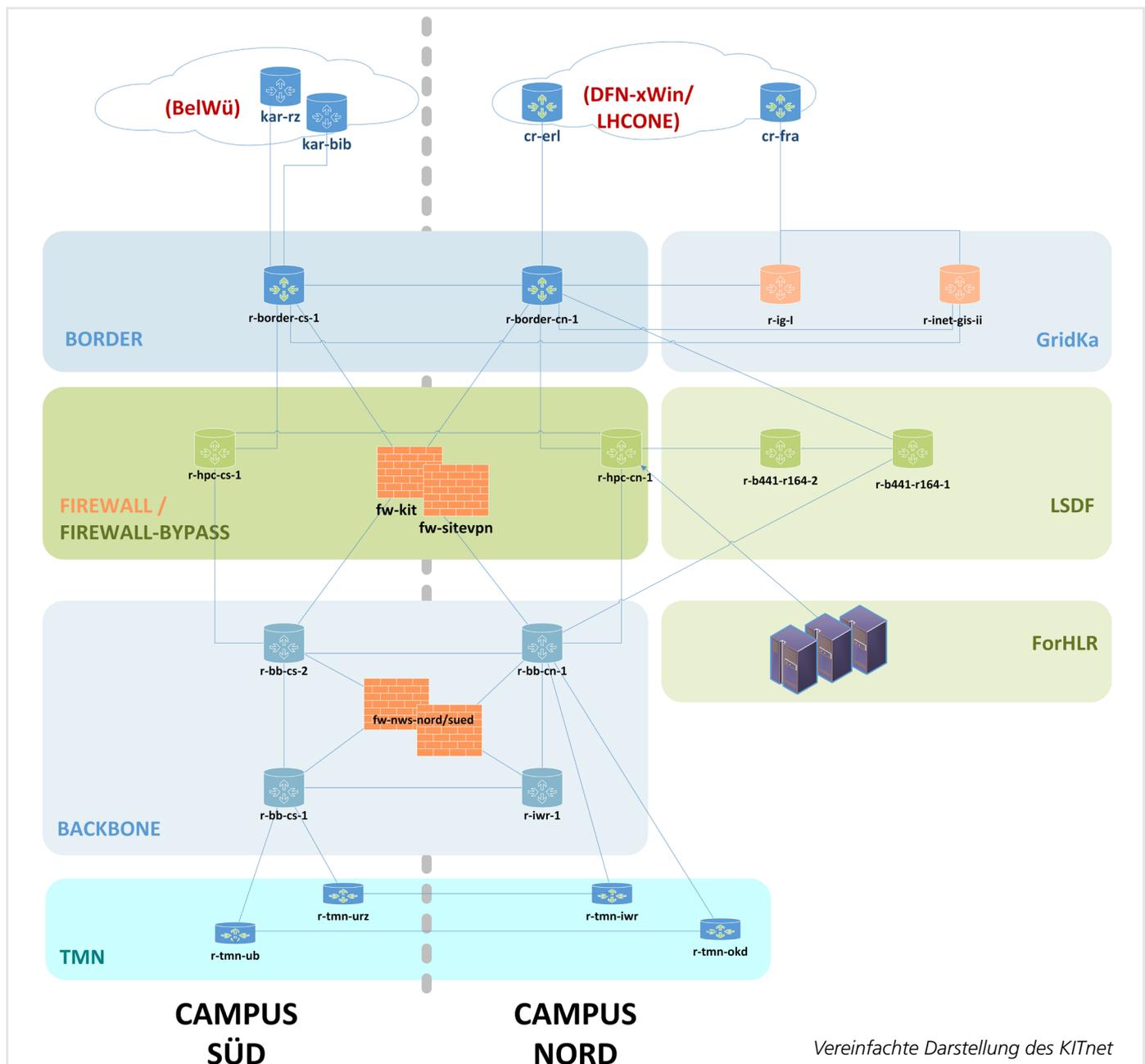
Das SCC betreibt das Datennetz des KIT (KITnet). Dieses besteht im Wesentlichen aus dem Festnetz und dem drahtlosen Netz (WLAN). Das Festnetz lässt sich ein-

teilen in den Border-Bereich, die zentrale Firewall mit Intrusion Prevention System (IPS) und das Backbone.

Der Border-Bereich ist das Bindeglied zwischen dem KITnet und dem Internet. Hier sind zwei große Router-Systeme mit

hoher Leistungsfähigkeit platziert, die die weltweiten Routing-Tabellen speichern und verarbeiten können.

Die zentrale Firewall regelt, auf welche IT-Ressourcen des KIT von außerhalb zugegriffen werden darf und überwacht dabei



auch den aus Protokoll-Sicht geregelten Zugriff. Das IPS unterbindet Angriffe auf die freigeschalteten Ressourcen im KITnet durch eine Analyse des Datenstroms.

Das Backbone verbindet die Netze der Institute mit den zentralen Servern des SCC und weiterer Dienstleister am KIT. Im Backbone befinden sich neben den vier Backbone-Routern die dezentralen Firewalls zur besonderen Absicherung von Netzbereichen gegenüber dem restlichen Backbone.

Zusätzlich betreibt das SCC spezielle Router-Systeme zur Anbindung des Grid Computing Centre Karlsruhe (GridKa), der Hochleistungsrechner-Systeme (HPC) und der Large Scale Data Facility (LSDF). In den Gebäuden des KIT befinden sich über 2200 Ethernet-Switches in 500 Datenverteilern zur Versorgung der Einrichtungen des KIT.

Das WLAN versorgt die Gebäude und Freiflächen des KIT und deckt auch umliegende Flächen mit hoher Aufenthaltswahrscheinlichkeit von KIT-Mitarbeitern und KIT-Studierenden ab. In den Gebäuden des KIT werden insgesamt 1700 WLAN-Accesspoints betrieben.

### Die Ausgangslage

Bis Anfang 2016 verfügte das KIT über vier Internet-Uplinks mit einer Bandbreite von je 10 Gbit/s und im Backbone überwiegend gebündelte 10 Gbit/s Verbindungen. Bereits Anfang 2015 wurde offensichtlich, dass auf Grund gesteigener Auslastungszahlen im KITnet eine Erhöhung der Gesamtleistung des Netzes erforderlich wird. Die erste Priorität lag im Border-Bereich. Da die vorhandenen Systeme nicht hochrüstbar waren, wurde zuerst ein Anforderungskatalog für neu zu beschaffende Border-Router erstellt. Nach einer umfangreichen Marktanalyse wurden ab Ende 2015 zwei neue Hochleistungsrouter (Brocade MLXe) beschafft und zuerst am Campus Nord der vorhandene alte Border-Router ersetzt. Seit Mitte 2016 verfügt das Netz am Campus

Nord damit über eine doppelte 100 Gbit/s Anbindung ans Internet über das Netz des DFN (Deutsches Forschungsnetz). Ende 2016 war der zweite Border-Router für Campus Süd verfügbar. Dieser wurde Anfang 2017 in Betrieb genommen. In Kürze wird auch BelWü (Baden-Württembergs extended LAN, Hochschulnetz Baden-Württemberg) dem KIT 100 Gbit/s-Anschlüsse anbieten können. Sämtliche mit dieser Hochrüstung verbundenen Umbauarbeiten im Border-Bereich konnten ohne Ausfälle und größere Wartungsfenster durchgeführt werden.

### Entwicklung im Backbone

Nach der Hochrüstung des Border-Bereichs des KITnet liegt der Fokus auf der Hochrüstung des Backbones. Hier finden grundsätzliche Überlegungen bezüglich der neuen Architektur und der einzusetzenden Technologie statt. Bei den Produkten, die seit Ende 2009 im Backbone des KITnet in Betrieb sind, werden VLANs (Virtuelle LANs) direkt in der KITnet-Infrastruktur implementiert. Das bringt Skalierungsprobleme mit sich, besonders im Zusammenhang mit Layer-2-Redundanzprotokollen. Es gab damals große Hoffnungen, dass virtualisierte Backbone-Router (die Zusammenschaltung zweier Router zu einer logischen Einheit durch Nutzung proprietärer Technologien) die Verfügbarkeit des Netzes drastisch steigern würden und ein kontinuierlicher Netzbetrieb auch bei Durchführung von Software-Updates gewährleistet sein würde. Beides trat nicht im erwarteten Umfang ein.

Aus den Erkenntnissen beim Betrieb eines großen Campus-Netzwerks heraus, plant das SCC nicht nur die verfügbare Bandbreite zu verzehnfachen. Die neuen Planungen werden mit dem Ziel vorangetrieben, das KITnet mit neuen, inzwischen verfügbaren Technologien zu modernisieren. Dieses neue KITnet soll neben einer deutlich höheren Durchsatzleistung optimierte Verfahren zur Automation des Netzbetriebs, Wartbarkeit, Skalierung, Ausfallsicherheit und möglichst eine

Integration von Netz- und Systemvirtualisierung bieten. Für den Ausbau des KIT-Backbones werden standardisierte Virtualisierungstechnologien wie MPLS/VPLS, VXLAN und EVPN genauer betrachtet und auf einen möglichen Gewinn für die genannten Ziele hin bewertet.

Im ersten Schritt werden Gespräche mit einer Vielzahl möglicher Hersteller von Backbone-Netzkomponenten geführt. Im zweiten Schritt werden die Gespräche mit den Herstellern weitergeführt, die Produktlinien vorgestellt haben, mit denen die Randbedingungen des KITnet erfüllbar sind und die eine gute Zukunftsträchtigkeit versprechen. Nach der positiven Evaluation einer Teststellung wird dann die finale Entscheidung für die einzusetzenden Produkte fallen. Es ist beabsichtigt, einen neuen Backbone parallel zum Bestehenden aufzubauen, um zeitnah die vier Backbone-Router-Systeme ersetzen zu können. Dabei ist geplant, die bestehenden Gebäude- und Gebietsanbindungen sukzessive auf den neuen Backbone umzuschwenken. Hier besteht dann die Möglichkeit, zusätzliche Anbindungen redundant auf 10 Gbit/s hochzurüsten, was aktuell wegen der limitierten 10 Gbit/s-Schnittstellen an den Backbone-Routern nicht möglich ist.

### Weitere Schritte

Als abschließende Maßnahme muss noch die zentrale Firewall zwischen KITnet-Border-Bereich und Backbone erneuert werden, damit es hier keinen Flaschenhals bei der Kommunikation mit dem Internet gibt. Auch hier werden in Gesprächen mit den infrage kommenden Herstellern Lösungsmöglichkeiten ausgelotet, um zu einer Produktfindung zu kommen. Der Beginn der Gespräche ist für Ende 2017 geplant. Die Inbetriebnahme der neuen Hochleistungsfirewall wird für Ende 2018 angestrebt.

## Der zentrale Virenschutzdienst am KIT

Die am KIT eingesetzten Computersysteme werden unter anderem durch Schadprogramme bedroht. Eine der am KIT verwendeten Schutzmaßnahmen gegen solche Bedrohungen ist der vom SCC betriebene Virenschutzdienst. Er bietet zusätzlich für IT-Administratoren eine zentrale Verwaltungsplattform für die Virenschutzkomponenten auf den Systemen. Im Rahmen des Schulungsangebotes bietet das SCC für die Administratoren der Institute und Einrichtungen einen Workshop an, der das Konzept des zentralen Virenschutzdienstes erläutert, in die Nutzung einführt und diese demonstriert.

*Helmut Gündert, Ulrich Stadie*

Im IT-Sicherheitskonzept für das KIT wird als Schutz vor Schadprogrammen der zentrale Virenschutzdienst zur Umsetzung eines funktionierenden Virenschutzkonzeptes empfohlen. Die Institute und Einrichtungen am KIT können dieses Angebot nutzen oder selbst festlegen, wie sie einen angemessenen Schutz vor Schadprogrammen in ihrem Verantwortungsbereich umsetzen.

### Hohes Gefährdungspotential durch Schadprogramme

Schadprogramme sind eine Gefahr für den ordnungsgemäßen Betrieb von IT-Systemen. Sie wirken auf unterschiedliche Weise und können zu erheblichen Beeinträchtigungen von Funktionalitäten führen. Durch Schadprogramme können Daten manipuliert, ausspioniert oder entwendet werden.

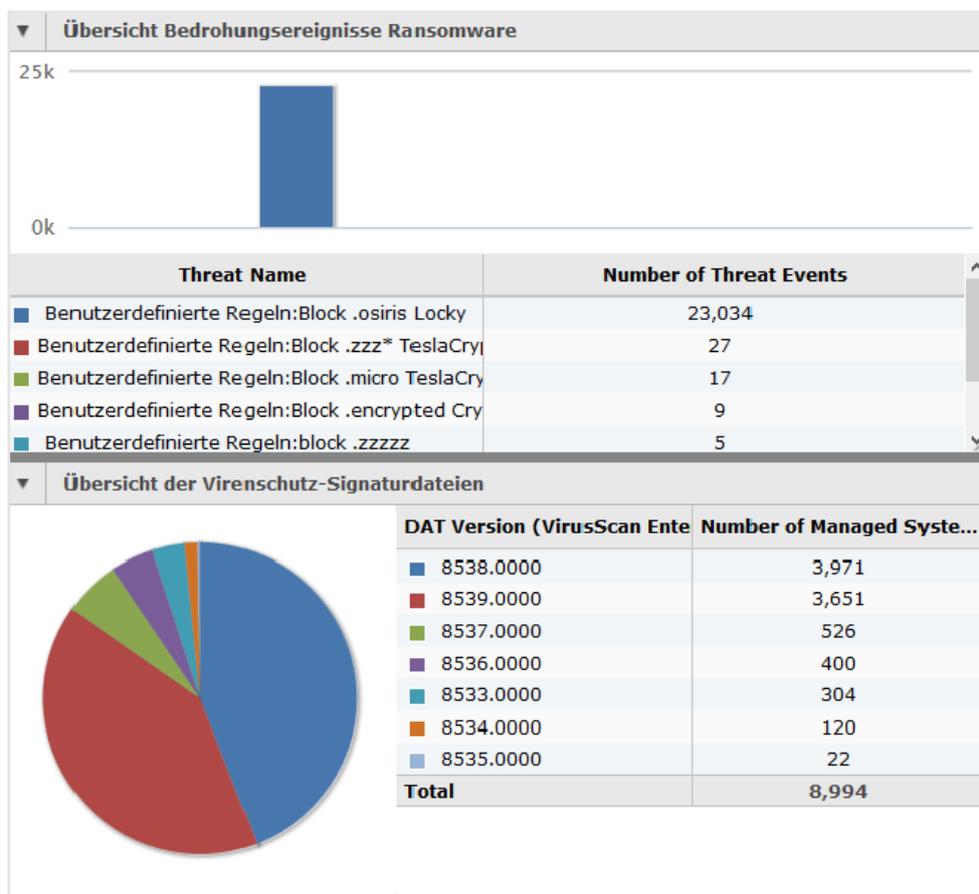
Bedroht sind derzeit vorrangig alle IT-Systeme mit Windows-Betriebssystemen sowie solche mit Anwendungsprogrammen, deren Dateien zum Beispiel durch Makro-Viren infiziert werden können. Schadprogramme können grundsätzlich auch bei Verwendung anderer Betriebssysteme oder Anwendungsprogramme auftreten. Dies gilt zum Beispiel auch für Unix/Linux-Systeme und Mac-OS-Systeme.

### Effizienzsteigerung durch zentrale Verwaltung

Das SCC bietet Organisationseinheiten am KIT (im folgenden OE genannt) die Möglichkeit einer zentralen Verwaltung der Antivirenlösungen. Mit dieser zentralen Lösung können die OE ihren Sicherheitsstandard erhöhen und auf einem stabilen Niveau halten.

Folgende Aufgaben lassen sich mit Hilfe der zentralen Verwaltung weitestgehend automatisieren:

- Verteilung und Konfiguration der Antivirensoftware
- Verteilung von Antiviren-Updates und Programm-Patches für die eingesetzte Software von McAfee
- Entdeckung von nicht-verwalteten Systemen
- Automatische Benachrichtigung der IT-Administratoren bei auffälligen Systemen
- Berichtswesen und Statistiken



Das SCC stellt den OE am KIT hierbei die Plattform des zentralen Virenschutzdienstes zur Verfügung. Die teilnehmenden OE des KIT melden dem SCC die für ihren Virenschutz zuständigen Administratoren. Diese erhalten Zugriff auf die Verwaltungskonsole der zentralen Virenschutzplattform (siehe Abbildung). Über diese Konsole können die Administratoren nur die Systeme ihrer eigenen OE verwalten. Auch bei Teilnahme am zentralen Virenschutzdienst bleiben die Administratoren der OE für die Sicherheit der Computersysteme ihrer OE verantwortlich. Dies betrifft insbesondere die Aktualität und Funktion der Virenschutzsoftware auf ihren Computersystemen. Mit Hilfe der Verwaltungskonsole können die Systeme regelmäßig überprüft und gegebenenfalls entsprechende Maßnahmen ergriffen werden.

### Schulung und Informationsaustausch notwendig

Das SCC führt zweimal jährlich einen Workshop zum zentralen Virenschutz und zur Verwendung der Verwaltungskonsole durch. Im Rahmen dieser Workshops werden neue Versionen der verwendeten Antivirenprogramme vorgestellt, und die Möglichkeiten der Verwaltungskonsole anhand von Demonstrationen erklärt. Des Weiteren können sich die Teilnehmer in den Workshops austauschen und konkrete Fragen zur Nutzung des Virenschutzdienstes stellen.

### Weitere Informationen

und die Anmelde­möglichkeit zum zentralen Virenschutz stehen unter dem folgenden Link zur Verfügung: <https://www.scc.kit.edu/dienste/virenschutz.php>

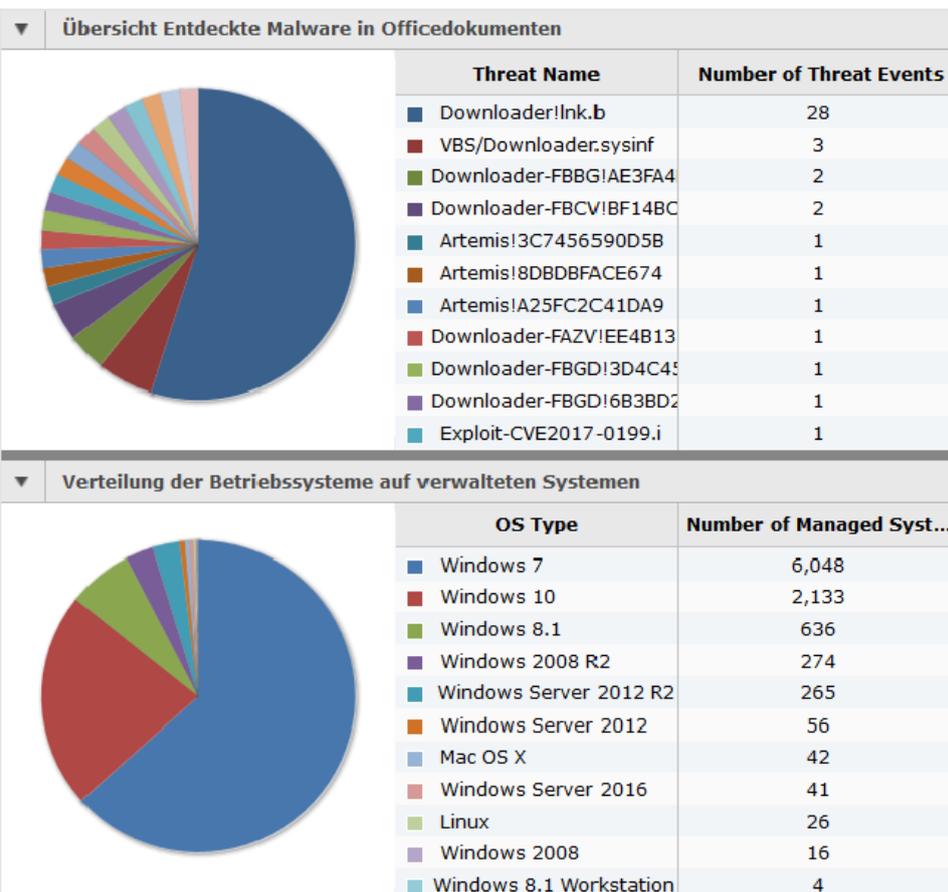


Abb.: Konsole der zentralen Virenschutzplattform, über die Administratoren Systeme in ihrer OE verwalten.

### Central Antivirus Service at KIT

Malicious software programs are a constant threat to the computer systems deployed at the KIT. The SCC operates for the KIT the antivirus service for the central administration of the antivirus software on the computer systems. To enable the administrators of participating institutes to employ the antivirus service effectively, the SCC offers an antivirus workshop, explaining the concept of the central antivirus service, demonstrating the advantages of a centralized antivirus management solution and the handling of typical scenarios.



Teil des neuen Onlinespeichersystems, von dem besonders die Hochenergiephysik-Experimente am CERN profitieren.

## Neues Onlinespeichersystem für GridKa

Im April 2017 wurde ein neues Onlinespeichersystem mit einer Kapazität von 20 Petabyte in Betrieb genommen. GridKa erhält damit zusätzlich eine Leistungssteigerung sowie eine einfache Erweiterbarkeit für zukünftig wachsende Anforderungen.

Andreas Petzold

Im April 2017 wurde ein neues 20 Petabyte Onlinespeichersystem am deutschen Tier-1 Daten- und Analysezentrum GridKa des Worldwide LHC Computing Grid (WLCG) am KIT in Betrieb genommen. Das von der Firma NEC gelieferte GxFS System besteht aus 14 Speichersystemen mit 3500 Festplatten und Solid State Disks, welche über zwei redundante Infiniband-Fabrics mit 44 Fileservern verbunden sind. Die Anbindung der Fileserver an den GridKa Netzwerk-Backbone ist mit 40 Gbit/s Ethernetverbindungen realisiert, die in mehreren 100 Gbit/s Uplinks gebündelt werden.

Insgesamt erreicht das System einen kombinierten Schreib-Lese-Durchsatz von 70 GByte/s. IBM Spectrum Scale™ wird als Software-Defined-Storage-Layer genutzt, um eine transparente Skalierbarkeit von Kapazität und Leistung zu ermöglichen. Zurzeit werden 13,5 Petabyte Daten der vier LHC-Experimente ALICE, ATLAS, CMS, LHCb sowie weiterer Teilchen- und Astro teilchenexperimente wie Belle II, Compass und Auger, ohne Betriebsunterbrechung innerhalb von GridKa auf das neue Speichersystem migriert.

Mit dieser Kapazitäts- und Leistungssteigerung sowie der einfachen Erweiterbarkeit des neuen Onlinespeichersystems können auch die weiter wachsenden Anforderungen, insbesondere der LHC-Experimente, in den kommenden Jahren erfüllt werden. Erweiterungen von insgesamt ca. 8 Petabyte sind schon bis zum Ende des Jahres 2017 vorgesehen.

### New Online Storage System for GridKa

In April 2017, a new online storage system with a capacity of 20 petabytes has been taken into production at GridKa. The new storage system achieves a combined read and write throughput of 70 GByte/s. Currently, 13.5PB of data, belonging to the LHC and several other particle and astro-particle physics experiments, are being migrated to the new storage system. The migration is fully transparent to the users and is happening while regular workloads of the experiments continue to run at GridKa. The increased capacity and performance as well as the simplicity of adding possible future expansions, enable GridKa to satisfy the steadily increasing resource and performance requirements of the LHC experiments over the coming years.

## Neuer Self-Service – Freigabe persönlicher Daten

Das SCC hat im Mai 2017 eine Neuimplementierung der Datenschutzeinstellungen für die persönlichen Daten in Betrieb genommen. Damit stehen allen KIT-Mitarbeitern und -Mitarbeiterinnen weitreichende Möglichkeiten zur Freigabe ihrer persönlichen Daten zur Verfügung. Die Freigabeeinstellungen werden über das SCC-Self-Service-Portal festgelegt.

Ulrich Weiß

### Wer pflegt die Daten?

Im SCC-Self-Service-Portal [my.scc.kit.edu](http://my.scc.kit.edu) gibt es seit kurzem den Abschnitt Meine Daten/Datenschutz, in dem Mitarbeiterinnen und Mitarbeiter des KIT die Sichtbarkeit ihrer persönlichen Daten selbst festlegen können. Einige Daten wie z.B. Vorname, Nachname, Abteilung und OE-Zugehörigkeit werden aus den zentralen Personalsystemen übernommen. Änderungen können nicht selbst durchgeführt werden, sondern müssen mit der Dienstleistungseinheit Personalservice (PSE) abgestimmt werden. Das KIT-Konto ist fest mit diesen Daten verknüpft und kann daher nicht verändert werden.

Alle weiteren Daten, wie z.B. der Standort, die anzuzeigende E-Mailadresse oder die Telefonnummer, können ausschließlich vom jeweiligen Mitarbeiter selbst gepflegt und für die einzelnen Verwendungen freigegeben werden. So kann zum Beispiel eine international arbeitende Forscherin bestimmte Daten ins Internet freigeben, während ein Sachbearbeiter im Einkauf sie nur im Intranet freigeben hat. Die Angaben sind freiwillig, und es ist zulässig auf die Angabe von Daten zu verzichten.

### Was passiert mit den Daten?

Die angegebenen Daten werden an verschiedene Systeme übermittelt und dort gemäß deren Freigabe verwendet. Derzeit gibt es drei Anwendungsbereiche, in denen persönliche Daten verwendet werden dürfen:

Das weiter unten beschriebene neue **Personenverzeichnis (KV)**, das eine um-

fassende Suche nach Personen, Telefonnummern oder E-Mailadressen ermöglicht. Darüber hinaus fließen persönliche Daten in das **Globale Adressbuch (GAB)** ein, einem geschützten zentralen Verzeichnisdienst, der im Active-Directory, Outlook, Exchange und ggf. für Kontakte auf Mobiltelefonen verwendet wird. Die **KIT-Internetauftritte**, die mit dem zentralen Redaktionssystem gepflegt werden, verwenden ebenfalls die freigegebenen Daten.

Während die freigegebenen Daten im KV und GAB sofort zur Verfügung stehen, ist die Personenseite im Internetauftritt einmalig dem KIT-Account zuzuordnen. Dies geschieht im Allgemeinen durch die

verantwortlichen Redakteure der Einrichtung, die Zugriff auf die Personendetails im OpenText Web Site Management haben. Die Redakteure können entweder einzelne Personen oder komfortabel alle vorhandenen Personenseiten auf die neue Funktionalität umstellen.

### Welche Freigaben gibt es?

Im SCC-Self-Service-Portal sind drei Freigabestufen für persönliche Daten einstellbar. Diese bezeichnen die Zugänglichkeit von persönlichen Daten in Abhängigkeit von der Herkunft/Berechtigung des Anfragenden:

		Freigabestufe <sup>1</sup>			Verwendet für <sup>2</sup>
		Öffentlich	KIT-Netz	Geschützt	
Nachname	Weiß	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Vorname	Ulrich	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Akad. Titel		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	KV, Web
Organisationseinheit	SCC	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Abteilung	SCC-DEI	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	KV, Web
Tätigkeit / Funktion	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	KV, Web
<b>Kommunikation</b>					
		Freigabestufe <sup>1</sup>			Verwendet für <sup>2</sup>
		Öffentlich	KIT-Netz	Geschützt	
Angezeigte E-Mail	<input type="text" value="weiss@kit.edu"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	KV, GAB, Web
URL für Homepage	<input type="text" value="https://www.scc.kit.edu/personen/ulrich.we"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Standort	<input type="text" value="Campus Süd"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB
Gebäude	<input type="text" value="20.21"/> <a href="#">Campus Plan</a>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Raum	<input type="text" value="006.2"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Telefon	<input type="text" value="+49 721 608-4 4868"/> (4 Ziffern)	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web
Telefon 2	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	KV, GAB, Web
Telefon mobil	<input type="text"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	KV, GAB, Web
Fax	<input type="text" value="+49 721 32550"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	KV, GAB, Web

Abb. 1: Freigabe der persönlichen Daten mittels spezieller Freigabestufen

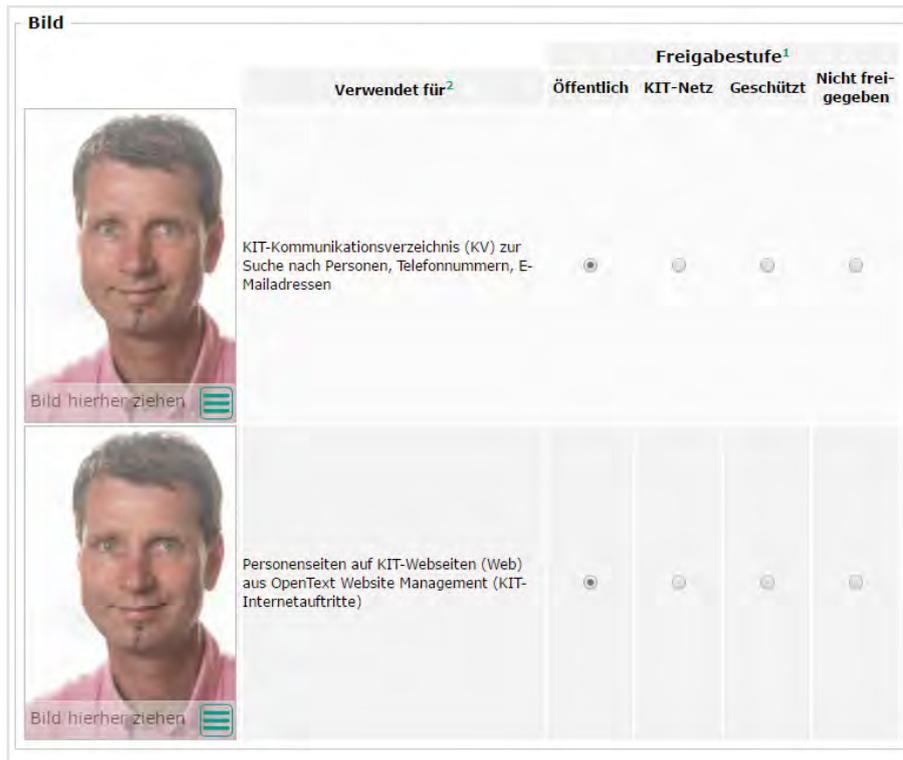


Abb. 2: Freigabe der Bilder je nach Verwendungszweck

**Geschützt** bedeutet, dass das Attribut nur für eine geschlossene Benutzergruppe nach Login/Authentifizierung mit KIT-Account sichtbar ist. Hierzu zählen

Webseiten bzw. -auftritte, die nur nach Login mit KIT-Account zugänglich sind, z.B. das Intranet-Portal des KIT (intra.kit.edu). Die zweite Freigabestufe, **KIT-Netz**,

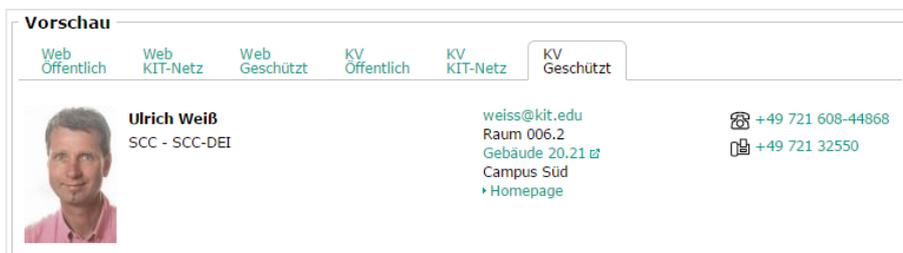


Abb. 3: Vorschau der Daten nach erfolgten Datenschutzeinstellungen



liefert eine Sichtbarkeit des Attributs nur für Anwender, welche die Information aus dem KIT-Netzwerk (IP-Netz) heraus abrufen. Am weitreichendsten ist die Stufe **Öffentlich**, womit das Attribut weltweit auf zugänglichen Seiten, wie z.B. die meisten OE-Webauftritte oder www.kit.edu, sichtbar ist.

**Wo kann man persönliche Daten recherchieren?**

Im Intranet (**intranet.kit.edu**) ist die Suche nach persönlichen Daten direkt über die Startseite (Telefonsuche) verlinkt und kann in einer einfachen oder Detailvariante aufgerufen werden. In der Detailsuche kann gezielt nach Name, Vorname, Organisationseinheit, E-Mail, Telefon oder Account gesucht werden, während die einfache Suche eine Suche über alle Felder ermöglicht. Die Suche ist auch direkt unter der URL **intranet.kit.edu/kv-suche** aufrufbar. Im Gegensatz zum bisherigen Kommunikationsverzeichnis ist auch ein eingeschränkter Zugang ohne Login aus dem KIT-Netz möglich. Dabei werden jedoch nur die Informationen angezeigt, die für diese Freigabestufe von der jeweiligen Person freigegeben worden sind. Zusätzliche Informationen sind nach einem Login verfügbar, das mittels Single-Sign-On ohne wiederholte manuelle Authentifizierung auskommt.

**Geht's mobil noch komfortabler?**

Seit kurzem ist ein auf mobile Geräte, wie Smartphones, optimiertes Portal online: **KIT-Mobile**, erreichbar unter **m.kit.edu**. Dabei handelt es sich um ein Angebot des SCC, das den Zugriff auf lokalisierte und personalisierte Inhalte mobil und leicht zu bedienend anbietet. Dieses neue Portal bietet je nach Personengruppe

Abb. 4: Personensuche auf intranet.kit.edu, Detailinformationen in Abhängigkeit der Freigabestufen und des Status des Abrufen

individuelle Zugänge: Studierende haben nach Login einen direkten Zugang zum Campus Management inkl. Stundenplan und Notenspiegel. Mitarbeiterinnen und Mitarbeiter bekommen anstelle dessen direkten Zugang zum SCC-Self-Service-Portal inkl. Konto Informationen, E-Mailweiterleitungen und Datenschutzeinstellungen.

Für alle Zielgruppen gemeinsam stehen drei Informationen an oberster Stelle: Wann fährt der nächste KIT-Shuttle? Ist

der Zugriff auf den Standort im Mobilgerät aktiviert, werden, abhängig vom eigenen Standort, Wege und Zeiten zur nächsten Haltestelle sowie die nächsten beiden Shuttle-Abfahrtszeiten angezeigt.

Außerdem zeigt das Portal die tagesaktuellen Angebote der Mensa sowie der Casinos des Campus Nord und des Max Rubner-Instituts an. Das dritte Highlight von KIT-Mobile ist die Personensuche, die auf o.g. Freigaben basiert und für den mobilen Bedarf optimiert ist. Ein

so gefundener Kontakt kann per Klick auf dessen Telefonnummer angerufen werden.

Zur Abrundung sind aktuelle Veranstaltungen, die auf der KIT-Homepage veröffentlichten News, sowie, je nach Login-Status bzw. Herkunft, die des KIT-Intranet und aktuell laufende Vorlesungen über [m.kit.edu](http://m.kit.edu) verfügbar. Der Campusplan kann zur Orientierung auf dem Campus sowie zur Suche von Einrichtungen verwendet werden.

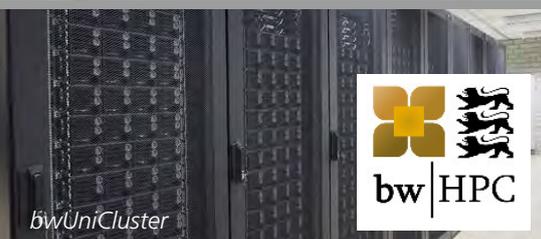
### New self-service – release of personal data

The new SCC self-service portal enables all KIT employees to determine the visibility and release of their personal information within many KIT directories. Moreover, a new mobile portal has been launched.

Abb. 5: Neues mobiles KIT-Portal [m.kit.edu](http://m.kit.edu)



Forschungshochleistungsrechner ForHLR



bwUniCluster



## Data Intensive Computing (DIC)

Large Scale  
Data Facility  
(LSDF)

Smart Data Innovation Lab (SDIL)

# Nutzertreffen zu Hochleistungsrechnen und datenintensiven Wissenschaften

Eine neue Veranstaltungsreihe erleichtert den informellen Austausch zwischen Nutzern und Betreibern von Hochleistungsrechnern und Forschungsdateninfrastrukturen.

*Dr. Andreas Heiss, Dr. Olaf Schneider*

Das SCC bietet für Wissenschaftlerinnen und Wissenschaftler eine Reihe von Dienstleistungen und Infrastrukturen (siehe Bildleiste) im Bereich des Hochleistungsrechnens und der datenintensiven Wissenschaften. Für die Weiterentwicklung dieser Dienste und Infrastrukturen ist ein enger Austausch mit den Nutzern unerlässlich. Deswegen wurde eine neue Veranstaltungsreihe ins Leben gerufen, bei der sich Betreiber und Nutzer in einem informellen Rahmen treffen können.

HPC-Investitionen informiert. Anschließend gab es Präsentationen zur interaktiven Remote-Visualisierung mit einem am SCC entwickelten Tool und zum Stand des LSDF-Upgrades. Ausführlich diskutiert wurde über die Möglichkeiten zur Speicherung und Weiterverarbeitung von Simulationsergebnissen. Zum Abschluss wurde auf laufende und kommende Umfragen verwiesen, mit denen die künftigen Anforderungen der Nutzer erfasst werden sollen.

Für den 24. November 2016 hatte das SCC zum ersten Nutzertreffen HPC+DIC eingeladen. Zirka 50 Teilnehmer aus über 20 Instituten versammelten sich im SCC-Hörsaal am Campus Nord, um nach einer Kurzvorstellung des SCC etwas über das engere Zusammenrücken von HPC-Systemen und der Large Scale Data Facility (LSDF), den neuen Dienst bwDataArchiv, die mittelfristige HPC-Roadmap und zum anstehenden Aus- und Umbau der LSDF zu erfahren. Zu den Präsentationen und in der abschließenden Diskussionsrunde wurden zahlreiche Fragen geklärt und Wünsche an das SCC geäußert.

Ein weiteres Treffen ist in der zweiten Jahreshälfte 2017 geplant. Dann wird es voraussichtlich um die Ergebnisse der Umfragen und die erreichten Zwischenstände beim LSDF-Ausbau und der Beantragung der nächsten HPC-Systeme gehen.

Als Informationsverteiler wurde eine Mailingliste eingerichtet (siehe Weitere Informationen). Auf dem Listserver kann man sich nach Anmeldung auch die bei den Treffen gezeigten Präsentationsfolien ansehen.

Dem allgemeinen Wunsch entsprechend, dass diese erfolgreiche Veranstaltung keine Eintagsfliege bleiben möge, wurde für 26. April 2017 ein zweites Treffen organisiert – dieses Mal am Campus Süd. Die ca. 15 Teilnehmer konnten sich am gerade errungenen Deutschen Rechenzentrumspreis für den ForHLR erfreuen (siehe Seite 28) und wurden über die Inbetriebnahme der bwUniCluster-Erweiterung sowie anstehende Anträge für

### User meetings on high performance computing and data-intensive sciences

The interexchange and cooperation between operators and users is essential for the further development of the services and infrastructures for high performance and data-intensive computing at the KIT. The SCC has therefore launched a new series of events for a mutual exchange of information and discussion.

#### Weitere Informationen

Mailingliste [scc-hpc-dic@lists.kit.edu](mailto:scc-hpc-dic@lists.kit.edu)

Anmeldung über [www.lists.kit.edu/sympa/info/scc-hpc-dic](http://www.lists.kit.edu/sympa/info/scc-hpc-dic)

Dokumente unter [www.lists.kit.edu/sympa/d\\_read/scc-hpc-dic/](http://www.lists.kit.edu/sympa/d_read/scc-hpc-dic/)

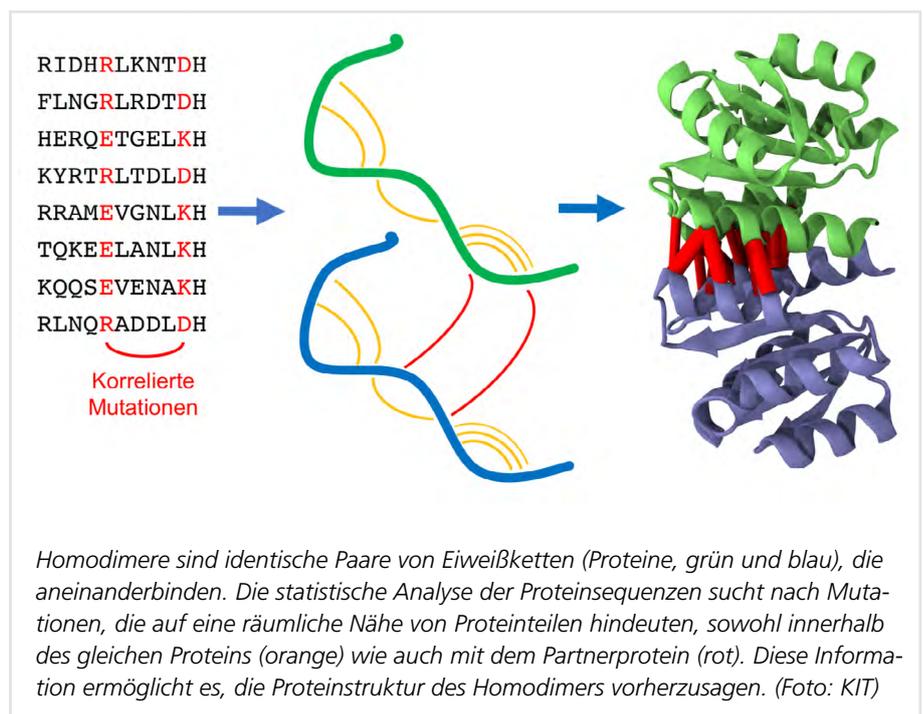
## Proteinstruktur mittels Big Data vorhergesagt

Ohne Proteine geht im Körper nichts – sie sind die molekularen Alleskönner in unseren Zellen. Arbeiten sie nicht richtig, kann das schwere Krankheiten auslösen wie etwa Alzheimer. Um Methoden zu entwickeln, nicht funktionierende Proteine zu reparieren, muss man deren Struktur kennen. Mit einem Big-Data-Ansatz haben Forscher des Karlsruher Instituts für Technologie (KIT) nun eine Methode entwickelt, mit der sie Proteinstrukturen vorhersagen können.

Dr. Felix Mescoli<sup>1</sup>

Wie die Forscher in der Zeitschrift *Proceedings of the National Academy of Sciences of the United States of America* (PNAS) berichten, ist es ihnen gelungen, auf Basis statistischer Analysen auch komplizierteste Proteinstrukturen unabhängig vom Experiment vorherzusagen. Diese experimentell zu bestimmen ist sehr aufwändig und der Erfolg nicht garantiert. Proteine bilden die Grundlage des Lebens. Als Strukturproteine sind sie am Aufbau von Gewebe wie Nägeln oder Haare beteiligt. Andere Proteine arbeiten als Muskeln, steuern den Stoffwechsel und die Immunabwehr oder übernehmen den Sauerstofftransport in den roten Blutkörperchen.

Die Grundstruktur von Proteinen mit bestimmten Funktionen ähnelt sich auch bei sehr verschiedenen Organismen. „Ob Mensch, Maus, Walfisch oder Bakterium, die Natur erfindet Proteine für verschiedene Lebewesen nicht immer neu, sondern variiert sie lediglich durch evolutionäre Mutation und Selektion“, so Dr. Alexander Schug vom Steinbuch Centre for Computing (SCC). Solche Mutationen können beim Auslesen der Erbinformationen, aus denen die Proteine zusammengesetzt sind, leicht identifiziert werden. Treten sie nun paarweise auf, liegen die beteiligten Proteinabschnitte meist räumlich nahe beieinander. Diese einzelnen Informationen vieler räumlich benachbarter Abschnitte können mit einem Computer wie ein großes Puzzle zu einer genauen Vorhersage der dreidimensionalen Struktur zusammengesetzt werden. Denn: „Um die Funktion eines Proteins wirklich



im Detail zu verstehen und eventuell auch zu beeinflussen, muss man den Ort jedes einzelnen Atoms kennen“, sagt Schug.

Der gelernte Physiker nutzt in seiner Arbeit einen interdisziplinären Ansatz mit Methoden und Ressourcen der Informatik und Biochemie. Mit Hilfe von Hochleistungsrechnern hat er die frei verfügbaren Erbinformationen tausender verschiedener Organismen von Bakterien bis hin zum Menschen nach in Verbindung stehenden Mutationen durchsucht. „Durch die Kombination von modernster Technik und einem wahren Schatz an Datensätzen konnten wir knapp 2000 verschiedene Proteine untersuchen – das ist eine

völlig neue Größenordnung im Vergleich zu bisherigen Studien“, sagt Schug. Dies unterstreicht eindrucksvoll die Leistungsfähigkeit dieser Methodik und verspricht großes Potenzial für breite Anwendungen von der Molekularbiologie bis hin zur Medizin. Auch wenn die gegenwärtige Arbeit Grundlagenforschung sei, wie Schug betont, können die Ergebnisse in Zukunft beispielsweise in neue Behandlungsmethoden von Krankheiten einfließen.

<sup>1</sup> Autor der zugehörigen KIT-Pressinformation

## Architektur für die standortübergreifende Betriebsüberwachung und -analyse in bwCloud

Das im Dezember 2016 abgeschlossene Landesprojekt bwCloud: Standortübergreifende Servervirtualisierung hat einen Prototyp für eine Infrastructure as a Service (IaaS) Science-Cloud – bwCloud für wissenschaftliche Einrichtungen in Baden-Württemberg konzipiert und aufgebaut. Das seit Januar 2017 begonnene Folgeprojekt bwCloud SCOPE hat zum Ziel, die bwCloud in einen Landesdienst zu überführen. „SCOPE“ steht dabei für die drei Anwendungsbereiche „SCience“, „OPerations“ und „EDucation“.

Neben strukturellen gibt es betriebliche und administrative Aufgaben und Fragen, die beim Übergang vom Prototyp zum Landesdienst beantwortet und umgesetzt werden müssen. Dieser Artikel beschreibt die Betriebsüberwachung und -analyse (Operations Analytics) der Cloud. Es wird die Architektur dargestellt, die zur dafür notwendigen Erfassung von Betriebsdaten implementiert wird.

*Oleg Dulov, Klaus Scheibenberger*

### Einführung

Während der zweijährigen Projektlaufzeit von bwCloud: Standortübergreifende Servervirtualisierung wurde ein Konzept für den Aufbau und Betrieb einer verteilten Science-Cloud auf Basis von OpenStack<sup>1</sup> entwickelt. Dazu wurde eine Testumgebung aufgebaut, die vier Betriebsstandorte umfasst. (s. Abb. 1).

Das Folgeprojekt „bwCloud SCOPE“ lässt sich in zwei Phasen aufteilen: In der ersten Phase findet der Übergang von einer Prototyp- zu einer Service-Infrastruktur statt. In dieser Phase – bis Ende 2017 – wird die neue Hardware-Infrastruktur installiert. In der zweiten Phase soll die bwCloud als Landesdienst für die Hochschulen in Baden-Württemberg bereitgestellt werden.

Abb. 2 zeigt die Struktur der bwCloud mit ihren regionalen und zentralen Komponenten. „Zentral“ ist hier so zu verstehen, dass solche Komponenten nicht notwendigerweise an allen Betriebsstandorten installiert und betrieben werden müssen, aber von allen genutzt werden. Beispiele für zentrale Komponenten sind die OpenStack-Komponente „Keystone“ für die Authentifizierung und Autorisierung, die am Betriebsstandort Freiburg installiert ist, oder die OpenStack-Benchmarking-Komponente Rally in Karlsruhe, die Testszenarien mit einer simulierten Nutzerlast in der Cloud ermöglicht.

In das Hardware-Layout fließen die Erfahrungen und Erkenntnisse ein, die mit dem Prototyp gewonnen wurden. Grundlage soll eine homogene Architektur für alle Betriebsstandorte bilden. Innerhalb eines Betriebsstandortes – in OpenStack als Region bezeichnet – wird die neue Hardware in zwei Cluster aufgeteilt sein: Ein Compute-Cluster und ein Storage-Cluster auf Basis

von Ceph<sup>3</sup> (je Standort ca. 30 Compute- und Storage-Cluster-Knoten). Angestrebt ist ein wachsender Aufbau von Ressourcen, die mindestens 1000 virtuelle Maschinen pro Region gleichzeitig betreiben können.

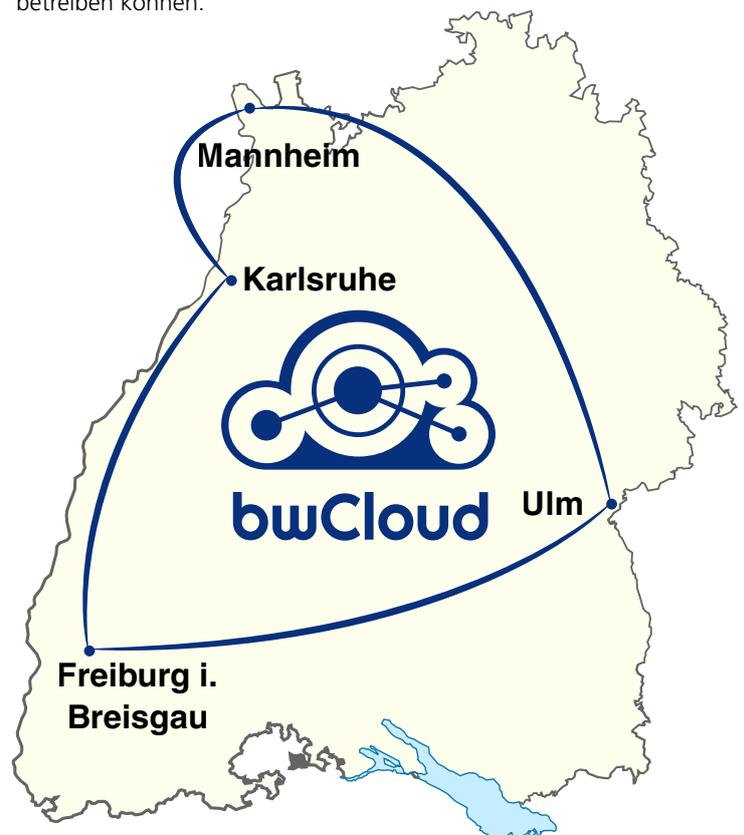


Abb. 1: Betriebsstandorte der aktuellen bwCloud. Die Betreiber sind neben dem SCC die Rechenzentren der Universitäten Mannheim, Freiburg und Ulm.

<sup>1</sup> [www.openstack.org/](http://www.openstack.org/)

<sup>2</sup> Eine ausführlichere Beschreibung wurde bereits in den SCC-News 1/2015 gegeben ([www.scc.kit.edu/downloads/oko/SCC-news01\\_2015.pdf](http://www.scc.kit.edu/downloads/oko/SCC-news01_2015.pdf)).

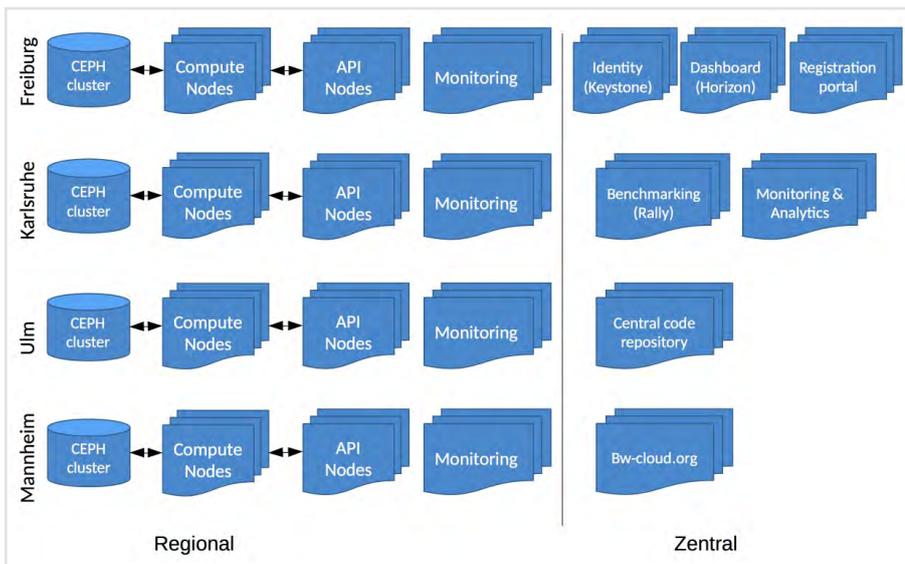


Abb. 2: bwCloud-Architektur: Regionale und zentrale Komponenten.

### Betriebsüberwachung und -analyse

Die Aufgaben der über die Betriebsstandorte verteilten bwCloud-Betriebsgruppe lassen sich logisch in zwei Bereiche aufteilen: Die Administration des Dienstbetriebs und der technischen Systemkonfigurationen und die Betriebsüberwachung und -analyse (engl. „Operations Analytics“). Für das zweite Aufgabenfeld bildet die Erfassung der Betriebsdaten die Grundlage und wird nachfolgend detaillierter betrachtet und dargestellt.

Jeder Betreiber eines IT-Dienstes benötigt Betriebsdaten von den Ressourcen, die in den Dienst eingebunden sind. Hierbei handelt es sich entweder um Echtzeitdaten, wie durchschnittliche Datentransferrate oder CPU-Load, oder um über längere Zeitintervalle erstellte Statistiken. Diese Daten helfen den aktuellen technischen Zustand zu beurteilen, z. B. ob Störungen vorliegen oder um Sicherheitsvorfälle zu erkennen. Desweiteren lassen sich aus diesen Daten z. B. Trends in der Ressourcennutzung des Dienstes ableiten, um rechtzeitig Erweiterungen anzustoßen und dadurch Performanz- oder Kapazitätsengpässe zu vermeiden, oder um Optimierungsmöglichkeiten in der Systemkonfiguration zu erkennen.

Gegenüber einfachen Dienststrukturen stellt der Landesdienst bwCloud hinsichtlich seiner technischen Architektur ein massiv verteiltes System dar, dessen Komponenten räumlich verteilt sind (s. Abb. 2) und das zudem von einem ebenfalls räumlich verteilt aufgestellten Team betrieben wird. Des Weiteren beeinflussen die für die Nutzer bereitgestellten Systeme das Verhalten dieser Dienstinfrastruktur ebenfalls massiv. Alleine schon die hohe

Anzahl von Komponenten, aber auch die der virtuellen Maschinen (VM), bedingen ein großes Volumen an Betriebsdaten.

Diese Randbedingungen erfordern eine strukturierte Umgebung zur Erfassung und Bereitstellung von Betriebsdaten, damit das Betriebsteam gemeinsam, auch standortübergreifend, darauf aufsetzende Analysen zur Störungseingrenzung, Optimierung und Planung effizient durchführen kann. Dafür werden u. a. effiziente Such- und Korrelationsmöglichkeiten für die gesammelten Daten benötigt. Den Cloud-Nutzern sollen ebenfalls Informationen zu ihren Ressourcen bereitgestellt werden, die sie für den Betrieb ihrer Dienste (z. B. Webserver) nutzen.

In den folgenden Abschnitten werden die für die Betriebsüberwachung und -analyse

geplant und für das prototypische Szenario entwickelten, grundlegenden Strukturen skizziert.

### Metriken und Ereignisse

Die technische Überwachung und Analyse eines jeden IT-basierten Dienstes umfasst Metriken- und Ereignisanalyse. Die zugrunde liegenden Betriebsdaten für diese Analysen lassen sich folgendermaßen gruppieren:

1. Netzwerkmetriken: Transaktionsdaten über ein Netzwerk oder Netzwerke.
2. Systemdaten: Das sind zum Einen Metriken aus den Hostsystemen (Host CPU, Load, usw.) und aus den OpenStack-Instanzen (virtuelle Maschinen CPU, Load, usw.), zum Anderen z. B. Log-Einträge.
3. Synthetische Daten: Durch Tools generierte Diagnose- und Benchmarking-Tests, um Konfigurationsfehler oder das Verhalten in einem System zu identifizieren und zu analysieren.

Generell sind die Grundlagen für jedes Monitoring- und Analyse-System: i) Daten sammeln und speichern, ii) Alarmierung beim Überschreiten von Grenzwerten und iii) die graphische Präsentation.

### Erfassung und Analyse von Metriken für das Echtzeit-Monitoring

Die Architektur (s. Abb. 3) für die Erfassung von Metriken innerhalb der bwCloud besteht aus verteilten Zeitreihen-Datenbanken

<sup>3</sup> ceph.com/

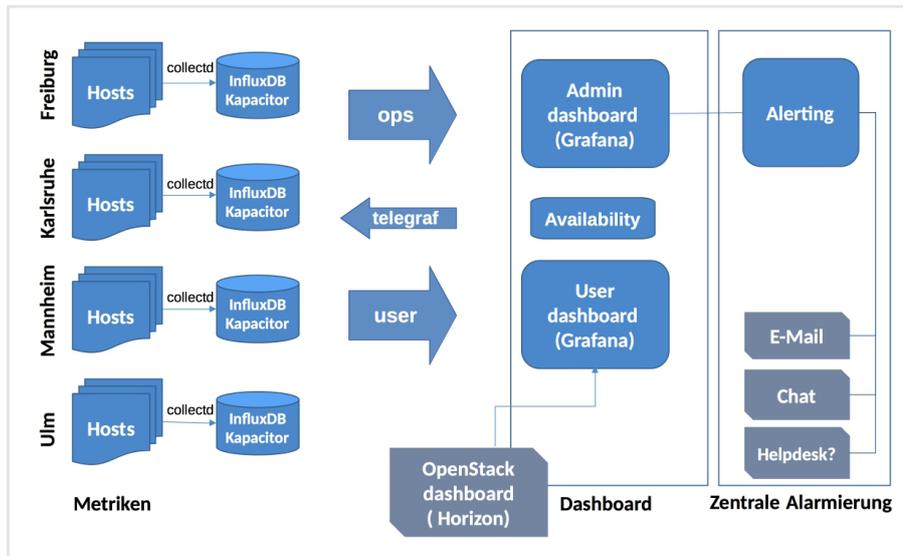


Abb. 3: Architektur der bwCloud-Metrikerfassung – links regional, rechts zentral.

(engl. „time series databases“) auf der Basis von InfluxDB<sup>4</sup> zusammen mit Kapacitor<sup>5</sup> – einem Processing- und Alerting-Werkzeug. Die Performance-Metriken der Systeme werden von Collectd<sup>6</sup> geliefert. Die Verfügbarkeits-Metriken des Dienstes werden von Telegraf<sup>7</sup> ausgeliefert und im Zeitreihenformat gespeichert. Für die Präsentation der Daten gibt es zwei Grafana<sup>8</sup>-Instanzen, um Dashboards für Administrations- und Benutzer-Daten getrennt bereitzustellen (Abb. 5).

Typische Analyseaufgaben, die anhand von Metriken durchführbar sind, wären beispielsweise dynamisch erzeugte Ranglisten von VMs. Im Fall von bw-Cloud eventuell eine Top-Ten-Liste der VMs mit hoher Ressourcen-Nutzung. Ebenso sind Trendanalysen sinnvoll, welche die Kapazitätsplanung und Entscheidungen über eine Migration von Ressourcen unterstützen.

Mittels Metrik-Analysen können Anomalien, die sich durch seltene oder ungewöhnliche Datenpunkte zeigen, erkannt werden. Durch Schätzen der Beziehung zwischen Metriken über längere Zeitintervalle (Regressionsmethode), können Prognosen erstellt werden, mit denen Abweichungen im Systemverhalten erkannt werden können und auf die so proaktiv reagiert werden kann.

### Erfassung und Analyse von Ereignissen

Die Architektur (s. Abb. 4) für die Erfassung von Ereignissen innerhalb der bwCloud besteht aus einem ELK-Stack<sup>9</sup> Cluster (ElasticSearch, Logstash und Kibana) in Verbindung mit Kafka<sup>10</sup>, dem Message Queue Cluster. Die Ereignisse aus Systemen und Anwendungen werden über Logstash aus den Regionen direkt in die Message Queue übertragen. Ein Teil davon wird dann direkt in der Elastic-search Datenbank gespeichert, andere werden vorher analysiert. Für die Suche und Präsentation der Daten wird Kibana verwendet. Das Alerting erfolgt durch Elastalert<sup>11</sup> verknüpft mit Produkten wie Alerta<sup>12</sup>.

Eine typische Analyse von Ereignissen ist die von Einträgen in Log-Dateien, um bspw. systemkritische Ereignisse oder Vorgänge, aber auch sicherheitskritische Aktivitäten zu erkennen. Um den Status und die korrekte Funktion der bwCloud-Installation zu prüfen, wird derzeit die OpenStack-Komponente Rally eingesetzt. Diese generiert durch zyklische Tests Daten, um beispielsweise die ordnungsgemäße Installation und Funktionalität der OpenStack-Umgebungen sicherzustellen. Neben den reinen Analysen der Funktion, gibt es aber auch sicherheitsorientierte Fragestellungen. Diese können durch Sicherheitsanalysen erfolgen, die auf regelmäßig erstellten OpenSCAP -Berichten für



Abb. 5: Dashboards für Administrations- und Benutzer-Daten

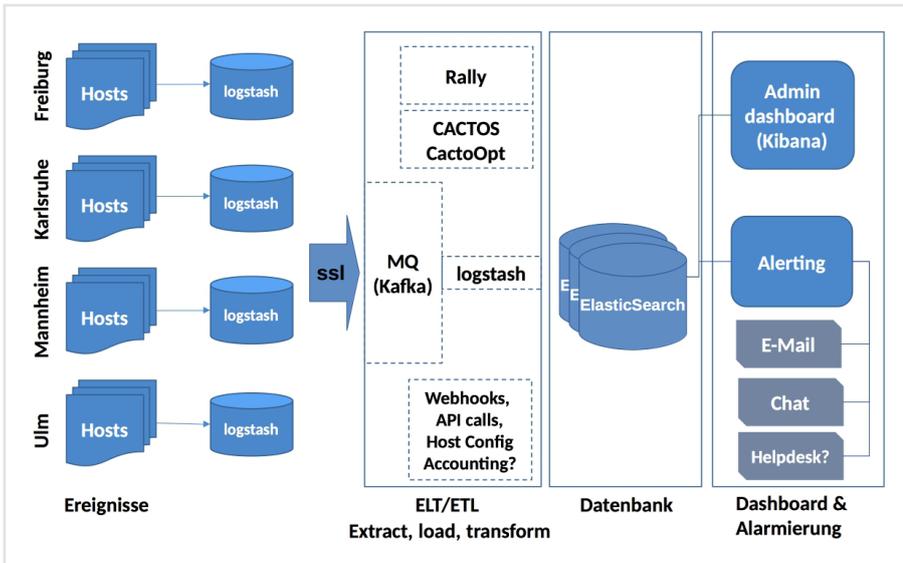


Abb. 4: bwCloud Eventprocessing

jeden bwCloud-Host basieren. Zu diesem Zweck werden zwei Richtlinien für bwCloud vergeben: Das Red Hat Standard System Security Profile als Grundlage für die Systemsicherheit und das Red Hat Firmenprofil für zertifizierte Cloud Provider (RH CCP) für die OpenStack-Zertifizierung. Um über die Systemänderungen zu informieren, wird das Advanced Intrusion Detection (AIDE) Tool in das bwCloud-System integriert.

Eine weitere Quelle für Analysen können anonymisierte Aktivitätsdaten von Nutzern sein. Diese ermöglichen Statistiken für die Klassifikationen der Benutzer, sowie für die Erkennung von nutzerbezogenen Anomalien.

**Fazit**

Das bwCloud-Projekt befindet sich derzeit in der Phase der Überführung der prototypischen Implementierung in einen Landesdienst. Es gibt viele Aspekte, die beim Übergang zu beachten und umzusetzen sind. In diesem Artikel wurde die „Betriebsüberwachung und -analyse der bwCloud“ basierend auf definierten Datenquellen hervorgehoben. Die dargestellten Technologien, Strukturen und Verfahren sind als Grundlage für die Überwachung und automatisierten Auswertungen von Betriebsdaten notwendig, um die Betriebsgruppe für Entscheidungsfindungen im Betrieb der bwCloud effizient zu unterstützen.

Darstellungen zu anderen Schwerpunkten werden in weiteren Ausgaben der SCC-News folgen.

<sup>4</sup>[www.influxdata.com/influxdb/](http://www.influxdata.com/influxdb/)

<sup>5</sup>[www.influxdata.com/kapacitor/](http://www.influxdata.com/kapacitor/)

<sup>6</sup>[collectd.org/](http://collectd.org/)

<sup>7</sup>[www.influxdata.com/telegraf/](http://www.influxdata.com/telegraf/)

<sup>8</sup>[grafana.com/](http://grafana.com/)

<sup>9</sup>[www.elastic.co/](http://www.elastic.co/)

<sup>10</sup>[kafka.apache.org/](http://kafka.apache.org/)

<sup>11</sup>[www.open-scap.org/](http://www.open-scap.org/)

<sup>12</sup>[static.open-scap.org/ssg-guides/ssg-centos7-guide-standard.html](http://static.open-scap.org/ssg-guides/ssg-centos7-guide-standard.html)

<sup>13</sup>[static.open-scap.org/ssg-guides/ssg-centos7-guide-rht-ccp.html](http://static.open-scap.org/ssg-guides/ssg-centos7-guide-rht-ccp.html)

<sup>14</sup>[aide.sourceforge.net/](http://aide.sourceforge.net/)

**bwCloud – architecture for cross-site operations analysis**

The bwCloud project is currently in transition from a prototype state to a Cloud service. There are many aspects and points to change and extend, in this article only one is highlighted: The architecture to store IT operations data within a defined structure and data sources as the baseline for bwCloud operations analytics. This enables an efficient monitoring and log processing to support decisions of the bwCloud operations team.

## Numerische Simulation turbulenter Verbrennung auf Hochleistungsrechnern des SCC

Simulationen von Verbrennungsprozessen beruhen auf komplexen mathematischen Modellen, die die grundlegenden physikalisch-chemischen Prozesse beschreiben. Die dafür etablierte Methode der Computational Fluid Dynamics stellt sehr hohe Anforderungen an die Ressourcen der verwendeten Computer wegen der komplexen Chemie während der Verbrennung, die Hunderte von chemischen Stoffen und Tausende von Reaktionen umfassen kann. Deshalb nutzt der Bereich Verbrennungstechnik am Engler-Bunte-Institut für seine aufwändigen Simulationen den am SCC betriebenen Forschungshochleistungsrechner ForHLR II sowie die enge Kooperation mit dem SimLab Energie am SCC.

Dr.-Ing. Feichi Zhang<sup>1</sup>, Thorsten Zirwes

### Verbrennungssimulationen am Engler-Bunte-Institut

Der Teilbereich Verbrennungstechnik am Engler-Bunte-Institut (EBI-VBT) der Fakultät für Chemieingenieurwesen und Verfahrenstechnik beschäftigt sich mit Forschungsthemen, deren Ziel es ist, zukünftige Verbrennungssysteme wie Strahltriebwerke oder Gasturbinen effizienter und sicherer zu machen und Schadstoffemissionen zu reduzieren. Zudem widmet sich das Institut den Fragenstellungen der Energiewende. Beispielsweise der Frage, wie Energie aus Biomasse und anderen erneuerbaren Energieträgern

gewonnen und gespeichert werden kann. Für diese Aufgabenstellungen haben sich neben den diagnostischen Messmethoden die numerischen Simulationen (Computational Fluid Dynamics – CFD) als ein wichtiges Werkzeug etabliert, sowohl in der Wissenschaft als auch in der Industrie. Für Simulationen von Verbrennungsprozessen werden die grundlegenden physikalisch-chemischen Prozesse durch komplexe mathematische Modelle beschrieben und danach mit Hilfe eines Computers berechnet. Als Ergebnisse solcher numerischen Simulationen werden physikalische Größen wie zum Beispiel Strömungsgeschwindigkeiten, Temperatur-

ren, Druck oder Konzentrationsfelder von unterschiedlichen Stoffen im gesamten Rechengebiet zur Verfügung gestellt und für die Analyse des Brennverhaltens verwendet.

### Berechnung benötigt Supercomputer

Das gesamte Rechengebiet wird in hunderte Millionen oder sogar Milliarden von sehr kleinen Teilvolumina, sog. Finite Volumes (FV), unterteilt. Zudem werden für jedes FV eine große Anzahl von gekoppelten, nichtlinearen, partiellen Differentialgleichungen anhand von numerischen Methoden gelöst. Deshalb ist der

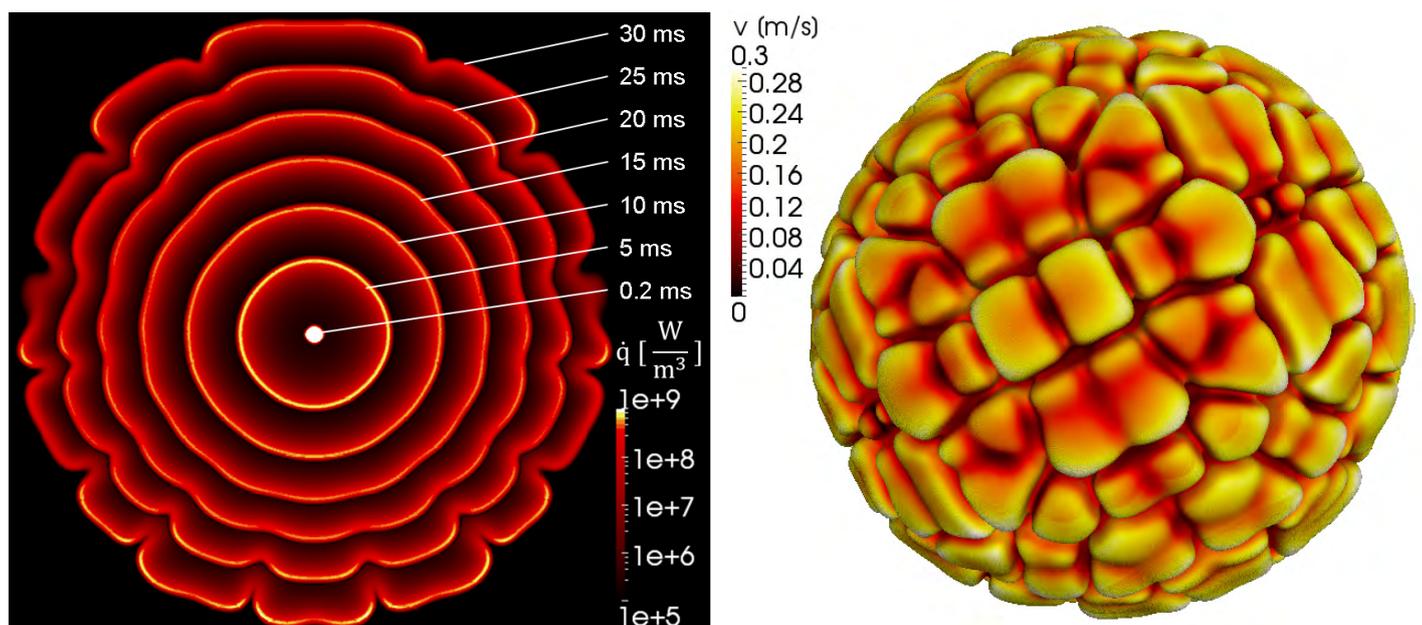


Abb. 1: 3D direkte numerische Simulation der sphärischen Flammenausbreitung: zeitlicher Verlauf der Flammenfront (links) und Snapshot von einer instabilen Flammenoberfläche (rechts).

<sup>1</sup> Engler-Bunte-Institut, Lehrstuhl für Verbrennungstechnik

Rechenaufwand für die Berechnung der turbulenten Verbrennung extrem groß und nur mit Hilfe eines modernen Supercomputers, wie z.B. dem bwUniCluster oder dem ForHLR des KIT, zu bewältigen. Der Lehrstuhl für Verbrennungstechnik am KIT beschäftigt sich seit vielen Jahren mit der CFD Berechnung von Verbrennungsvorgängen. Die Komplexität der angewandten Methoden passt sich an die Leistungssteigerung der Hochleistungsrechnersysteme am SCC/KIT (HC3 → XC2 → UC1 → ForHLR I → ForHLR II) und deutschlandweit (z.B. im HLRS Stuttgart und JSC Jülich) an. Diese Entwicklung ermöglicht und gewährleistet die detaillierteste Analyse der fundamentalen Mechanismen von Verbrennungsvorgängen und stellt die Grundbausteine für eine konkurrenzfähige Forschungsarbeit in der Öffentlichkeit, z.B. bei Publikationen, dar.

### Vorhersage von Lärm durch Verbrennungsvorgänge

Im Rahmen eines von der DFG geförderten Projektes namens „Verbrennungslärm“, können beispielsweise anhand von direkten numerischen Simulationen der durch turbulente Verbrennung induzierte Lärm vorhergesagt werden. Dieser Lärm gehört analog zu anderen Schadstoffemissionen wie  $\text{CO}_2$  oder  $\text{NO}_x$  zu den negativen Einflüssen aus Verbrennungsprozessen und ist ein wichtiges Kriterium für die Entwicklung von zukünftigen Verbrennungssystemen wie Gasturbinen oder Flugtriebwerken. Ziel der Forschergruppe um Dr. Feichi Zhang ist es deshalb, ein tiefgreifendes Verständnis der Entstehungsmechanismen von Verbrennungslärm zu gewinnen und eine wissenschaftliche Basis aufzubereiten, die die Kontrolle und Reduktion von Verbrennungslärm ermöglicht. Eine generische, mit Erdgas getriebene Jetflamme wird dafür unter Variation von Betriebsbedingungen wie Vorwärmung und Strömungsimpulse experimentell an der TU Berlin (Prof. Paschereit) untersucht und von Partnern an der TU Darmstadt (Prof. Janicka) mit gängigen numerischen Methoden berechnet. Der Fokus am EBI-VBT des KIT (Prof.

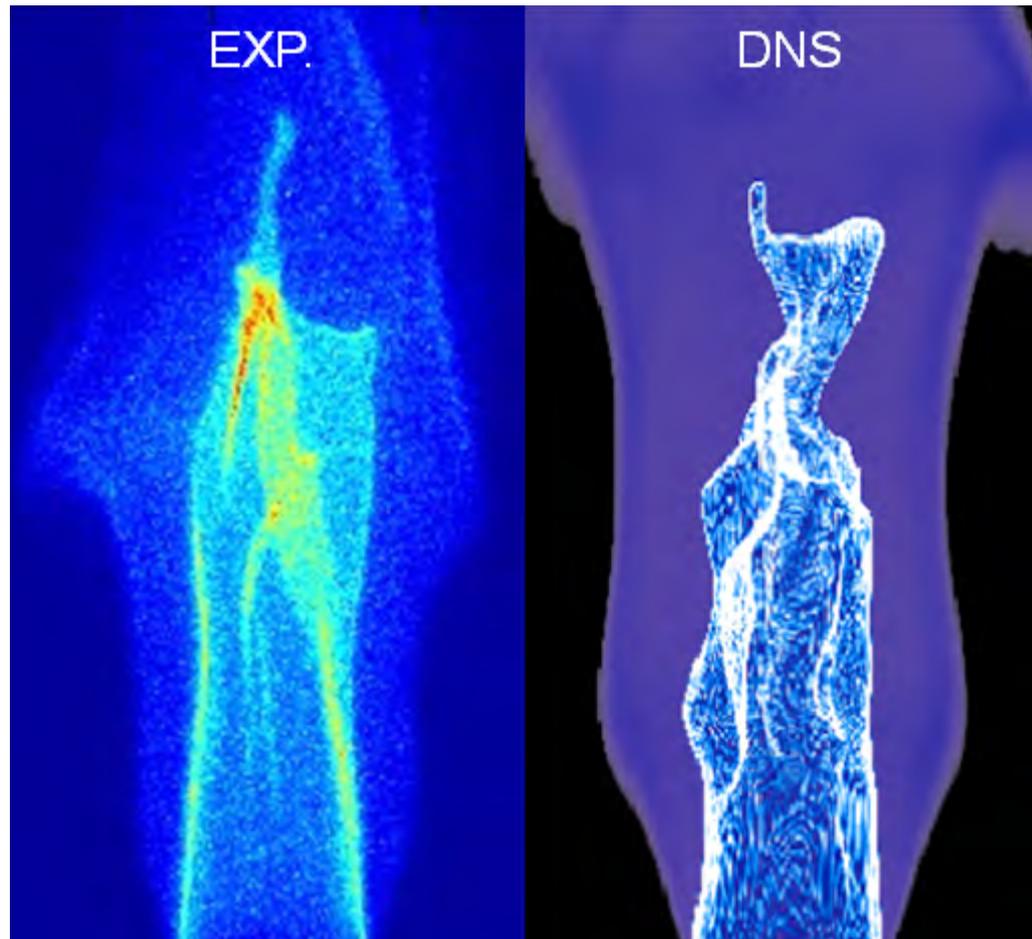


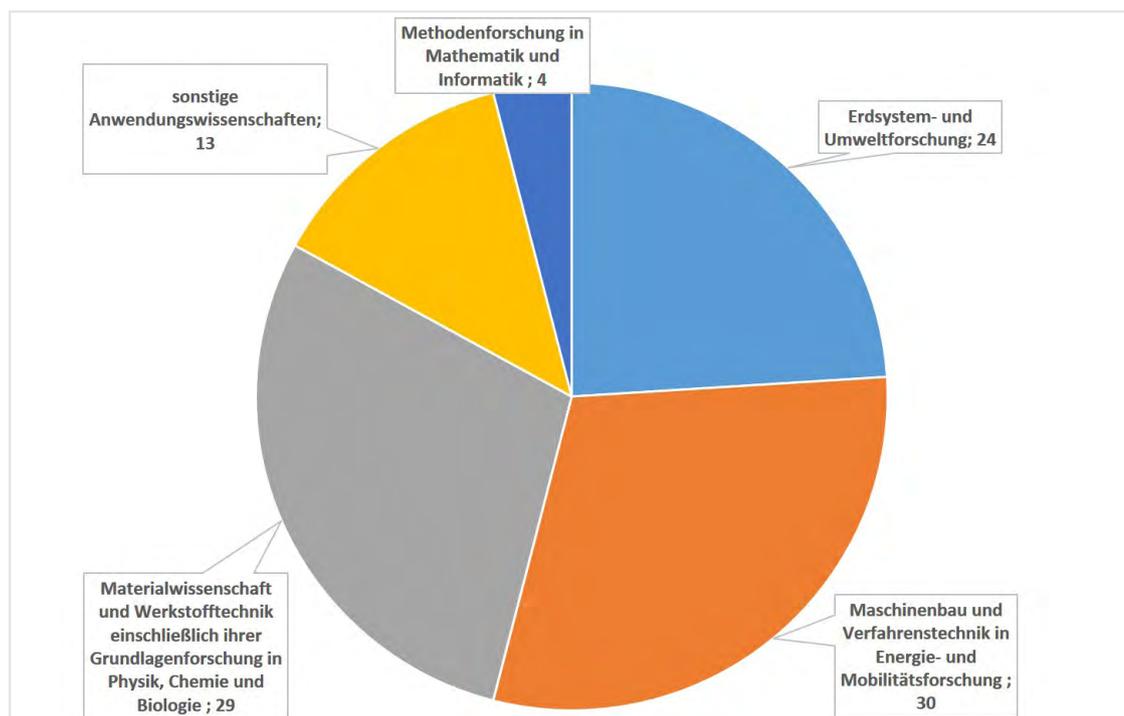
Abb. 2: Vergleich der Chemilumineszenz der Flamme aus der Messung (links) und der Simulation (rechts).

Bockhorn) liegt in der genauen Berechnung des akustischen Quellterms. Hier kommt die direkte numerische Simulation (DNS), die die aufwändigste Simulationemethode darstellt, unter Verwendung von detaillierten chemischen Reaktionen und physikalischen Transportdaten zum Einsatz. Die detailliert berechneten Lärmquellen werden dann an Kollegen der RWTH Aachen (Prof. Schröder) weitergegeben. Mit diesen Daten kann dort die Schallausbreitung in das Fernfeld anhand von Computational Aero-Acoustics Codes weiter analysiert werden.

### Umsetzung der Simulation

Um dieses Ziel zu erreichen, wurde am EBI-VBT zuerst ein Programm für die direkte numerische Simulation von chemisch reagierenden Strömungen entwickelt, das eine hohe Genauigkeit und sehr

gute Skalierbarkeit für das parallele Rechnen auf Hochleistungsrechnern gewährleistet. Abbildung 1 zeigt die Ergebnisse einer DNS der sphärischen Ausbreitung einer Wasserstoff/Luft-Flamme in einem Bombenbehälter zur Validierung des entwickelten DNS-Lösers. Das Setup wird häufig für die Bestimmung der laminaren und turbulenten Brenngeschwindigkeit verwendet, welche wichtige Parameter für die Auslegung und Modellierung von Verbrennungssystemen sind. Auf dem linken Bild ist die zeitliche Ausbreitung der Flammenfront durch Konturen der Wärmefreisetzungsrates dargestellt. Die Flamme breitet sich am Anfang mit einer sphärischen Form aus und zerbricht danach wegen thermo-diffusiver Instabilität. Ein Snapshot der 3D-Struktur nach dem Zerfall der sphärischen Flammenfront ist auf der rechten Seite dargestellt. Für die DNS werden 144 Millionen finite Volu-



Die 100 Rechenzeitprojekte auf dem ForHLR verteilen sich auf unterschiedliche Wissenschaftsgebiete (Diagramm). Das im Artikel vorgestellte Projekt gehört zu Maschinenbau und Verfahrenstechnik in Energie- und Mobilitätsforschung.

mina und detaillierte Reaktionsmechanismen verwendet.

Die Rechnung wurde parallel auf 8192 Prozessoren des CRAY-XE6 Clusters am High-Performance Computing Centre Stuttgart (HLRS) durchgeführt. Dafür wur-

den ca. 2,5 Millionen CPU-Kern-Stunden benötigt. Der DNS-Löser wurde auch mit einer sehr guten Skalierbarkeit auf den HPC Clustern ForHLR I und ForHLR II des SCC am KIT nachgewiesen.

Abbildung 2 zeigt einen Vergleich der

system für die Verbrennung des Methan/Luft-Gemischs verwendet. Die Rechnung wurde auf 12.000 Prozessorkernen auf dem Hazel-Hen Cluster am HLRS durchgeführt und benötigte ca. 3 Millionen CPU-Kern-Stunden. Ferner wird ein DNS-Datensatz mit zeitlich und räumlich hochaufgelösten Strömungs- und Reaktionsfeldern in der Größenordnung von 20 Terabyte gespeichert und an die Aachener Kollegen für die Detailanalyse des Entstehungsmechanismus von Verbrennungslärm weitergegeben. Die Analyse der Daten sowie die Visualisierung erfolgte dann auf dem ForHLR II.

Diese erfolgreiche Forschungsarbeit wurde dank der langjährigen guten Zusammenarbeit des Instituts EBI-VBT mit den Wissenschaftlern des SimLabs Energie vom SCC und des HLRS und der an diesen Rechenzentren zur Verfügung gestellten Rechenleistung und Beratung ermöglicht. Thorsten Zirwes, Co-Autor des Artikels, ist ein gemeinsamer Doktorand des SimLabs Energie und des Engler-Bunte-Instituts, Bereich Verbrennungstechnik.

Chemilumineszenz der zuvor erwähnten Erdgas/Luft-Flamme, berechnet mit DNS (rechts) und gemessen an der TU Berlin (links). Das Bild zeigt die Projektion der Chemilumineszenz von elektronisch angeregten Hydroxyl-Radikalen durch die Flamme senkrecht zur Bildebene. Die Flammenlänge und Konzentrationsverteilungen der verschiedenen Verbrennungsprodukte aus der DNS zeigen eine sehr gute Übereinstimmung mit der Messung. Für diese DNS wurden 111 Millionen FV und ein komplexes Reaktions-

### Direct Numerical Simulation of Turbulent Combustion

Optimization of combustion devices requires understanding of the underlying chemo-physical, thermodynamical and fluid-mechanical phenomena. DNS is an extremely valuable research tool, which has already been extensively used to study the fundamental mechanisms during turbulent combustion. It resolves the full range of length and time scales in a turbulent reactive flow. The chemical reactions are so complex that their detailed calculation may take more than 90% of the total computing time. However, it provides a detailed insight into the underlying physics and chemistry by resolving the internal turbulent flame structure and the interaction of the flame with the flow. The in-depth understanding of these behaviors is of central importance for designing future combustion systems and developing reliable simplified modeling concepts. The Engler-Bunte-Institute/Division for Combustion Technology works since 2011 on analyzing combustion phenomena with the DNS method with the aid of Tetra to Peta-scale supercomputers, optimizing performance of the DNS code and developing tools for the evaluation of very big data sets from the simulation results. The expertise in this field at the Engler-Bunte-Institute promoted cooperation with many other industrial and university partners. Thorsten Zirwes, the coauthor, has a joint-PhD fellowship between EBI-VBT and SCC. This is a result of the strong cooperation of SimLab Energy at SCC and EBI-VBT.

## bwITsec – eine kooperative IT-Sicherheitsstruktur entwickelt sich

Im Projekt bwITsec arbeiten die neun Landesuniversitäten gemeinsam an der Weiterentwicklung ihrer IT-Sicherheitsstrukturen. Die Dokumentationsphase des Projekts wurde erfolgreich abgeschlossen und die Ergebnisse von der Projektleitung ausgewertet. Die Auswertung zeigt auf, welche IT-Sicherheitslösungen von den Projektpartnern zu entwickeln sind. Die Projektleitung entwirft derzeit ein landesweites Rahmenkonzept für die IT-Sicherheit.

Dr. Uwe Neumann

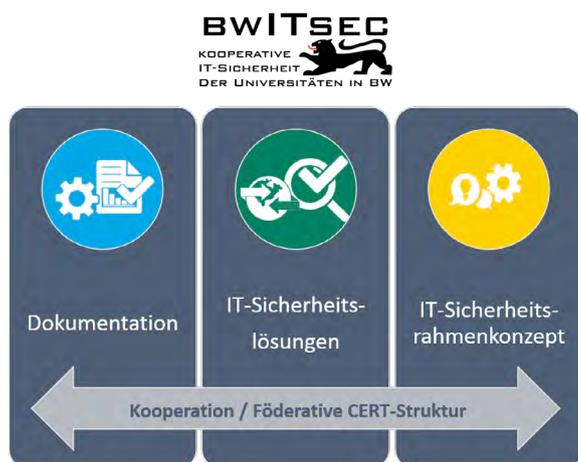


Abb. 1: Module des Projekts bwITsec

Seit dem Projektstart von bwITsec im Januar 2016 sind inzwischen fast eineinhalb Jahre vergangen. Zeit für einen Blick auf das, was inzwischen geschah.

Zu Beginn des Projekts war relativ wenig über den Stand der IT-Sicherheit an den teilnehmenden Universitäten bekannt, was unter anderem daran lag, dass es keine einheitliche Metrik gab, mit der dieser ermittelt werden konnte. Aus diesem Grund wurde in der Vorlaufphase durch die Projektleitung (Universität Stuttgart und Karlsruher Institut für Technologie) ein Fragenkatalog entwickelt, der eine einheitliche Bestandsaufnahme der jeweils vorhandenen technischen und organisatorischen IT-Sicherheitsmaßnahmen ermöglichte. Anhand dieses Fragenkatalogs haben alle Universitäten ihre Umgebungen dokumentiert und der Projektleitung zur Auswertung übergeben. Der so gewonnene Gesamtüberblick über Zustand und Baustellen zur IT-Sicherheit floss in die zweite Hauptaktivität des Projekts ein – der Entwicklung von IT-Sicherheitslösungen in Form von IT-Sicherheitsdiensten oder -prozessen.

An den insgesamt dreizehn IT-Sicherheitslösungen arbeiten die Projektpartner in Teams und stellen die Ergebnisse allen Projektteilnehmern zur Verfügung. In Karlsruhe wird derzeit an den Themen Mobilgerätesicherheit, Schwachstellenmanagement und Zertifikatsdienste gearbeitet. Bei der Mobilgerätesicherheit geht es um die Entwicklung eines Self-Services für die sichere Konfiguration

von Smartphones. Das Schwachstellenmanagement etabliert einen Dienst, der Betreibern von IT-Systemen an den Universitäten beim Auffinden von bekannten Schwachstellen oder unsicheren Konfigurationen ihrer Systeme helfen soll. Im Bereich der Zertifikatsdienste wurde bereits ein neues und verbessertes Verfahren für die Veröffentlichung von S/MIME-Zertifikaten entwickelt und in Betrieb genommen. Eine verbesserte Suchfunktion für Zertifikate, der neuen und alten Zertifizierungsstelle, steht kurz vor der Inbetriebnahme.

Weitere Themen sind unter anderem Sensibilisierung für IT-Sicherheitsthemen, Zweifaktorauthentifizierung, Ansprechpartnerdatenbank, Firewallmanagement und Vorfallsmanagement.

Der Informationsaustausch unter den Projektpartnern, auch teamübergreifend, ist rege, wozu neben regelmäßigen Projekttreffen auch die für das Projekt eingeführte Kooperationsplattform auf der Basis von GitLab und Chaträume beitragen.

In einer dritten Hauptaktivität (s. auch Abb. 1) wird derzeit ein landesweites Rahmenkonzept für die IT-Sicherheit für die Landesuniversitäten entworfen. Das Rahmenkonzept dient als Schablone für eine Sicherheitsorganisation mit den entsprechenden Rollen und Prozessen und integriert dabei die bereits vorhandenen Strukturen. Dennoch entwirft das vorgeschlagene Konzept konkrete Vorgaben, die in jeweils individuell angepassten Zeitplänen umgesetzt werden können. Ein wesentlicher Bestandteil des Rahmenkonzepts wird außerdem der Entwurf einer föderativen CERT-Struktur sein, die eine flächendeckende Vorfallsbearbeitung nach festgelegten Standards ermöglicht.

Der bisherige Bewilligungszeitraum für das Projekt bwITsec endet zum 31.12.2017. Aufgrund der erkannten Handlungsfelder sieht das Projekt den Bedarf für die weitere Entwicklung und Umsetzung der erarbeiteten Konzepte und Lösungen. Das bereits beantragte Folgeprojekt bearbeitet unter anderem diese Themen.

### **bwITsec – a Cooperative IT Security Structure is Evolving**

In the bwITsec project, the nine state universities are working together on the further development of the IT security structure of the universities of Baden-Württemberg. The documentation phase of the bwITsec project was successfully completed and the project management team evaluated the results. Based on the findings, the development of IT security solutions were developed by the project partners in teams. A statewide outlined IT security concept is currently under preparation by the project management.

## EU Projects: Contributions from SCC at a Glance

SCC participates in several projects funded by the European Commission in Horizon2020. This article provides an overview about the projects themselves and about the specific tasks that SCC is involved in.

*Arsen Hayrapetyan, Dr. Marcus Hardt*



**INDIGO-DataCloud** develops an open source data and computing platform targeted

at scientific communities, deployable on multiple hardware and provisioned over hybrid, private or public e-infrastructures. By filling existing gaps in Platform as a Service (PaaS) and Software as a Service (SaaS) levels, INDIGO - DataCloud will help developers, resource providers, e-infrastructures and scientific communities to overcome current challenges in Cloud Computing, storage and network areas.

The project unites 26 partner institutions from Germany, France, Italy, Netherlands, Spain, Portugal, Poland, Croatia, Czech Republic, Switzerland and UK.

SCC has been contributing to the definition of unified Quality-of-Service (QoS) and Data-Lifecycle (DLC) for data storage and access technologies in collaboration with the Research Data Alliance (RDA) [1], as well as to the implementation of an interface for Quality-of-Service management based on the Cloud Data Management Interface (CDMI) developed by the Storage Networking Industry Association (SNIA) [2]. The interface has been adopted to interact with the storage systems provided by the project partners, where KIT has developed a module to access and modify Storage QoS and DLC on an IBM High Performance Storage System (HPSS). During the course of the project basic QoS terms have been defined, the QoS interface has been successfully implemented and deployed over multiple storage sites, and the QoS management capabilities have been successfully de-

monstrated. The demonstrator is available at [cdmi-web.data.kit.edu](http://cdmi-web.data.kit.edu).

Another remarkable achievement of SCC is the implementation of the INDIGO-DataCloud Token Translation Service – **WaTTS**. **WaTTS** ([watts.data.kit.edu](http://watts.data.kit.edu)) creates authentication and authorisation on credentials for services that do not natively support OpenID Connect. Such services are for example ssh, S3 storage, OpenNebula or X.509 certificates. The OpenID Connect relying party (client) library developed at SCC is a certified implementation by the OpenID Connect Foundation. It is currently the only one implemented in Erlang.

More information on the project: [www.indigo-datacloud.eu](http://www.indigo-datacloud.eu)



**Human Brain Project**

Technologies (FET) Flagship Project which strives to accelerate the fields of neuroscience, computing and brain-related medicine. This acceleration will be achieved by a strategic alignment of scientific research programmes in fundamental neuroscience, advanced simulation and multi-scale modelling with the construction of an enabling Research Infrastructure. More than 100 institutions from 19 countries participate in the project.

SCC is involved in the development of this e-infrastructure by providing cloud services to HBP users. More specifically, SCC hosts and maintains an S3-compatible, WOS-based object storage with a

capacity of 480 Terabyte, complementing the storage provided locally by the HPC centres in HBP. SCC also offers cloud computing resources based on open-source technologies, such as OpenStack and Docker containers. Seamless and secure access to these services and all other resources in HBP is ensured through an OpenID Connect based single-sign-on system and the UNICORE grid middleware [3].

More information on the project: [www.humanbrainproject.eu/en/](http://www.humanbrainproject.eu/en/)



**Authentication and Authorisation for Research and Collaboration, AARC** is a project, which brings together 20 different partners from among National Research and Education Networks (NRENs) organisations, e-infrastructures service providers and libraries. AARC promotes federated access and works together with national identify federations, research infrastructures, e-infrastructures and libraries to identify building blocks and policy best practices needed to implement interoperable authentication and authorisation infrastructures (AAls).

SCC has contributed to the AARC Blueprint Architecture (BPA) [4] design. The BPA strives to provide a set of interoperable architectural building blocks for software architects and technical decision makers, who are designing and implementing access management solutions for international research collaborations. SCC also contributed to the AARC project with the design and development of three pilot implementations that enable

end users to use SSH access with OpenID Connect (OIDC) based login [5], obtain Identifier-Only Trust Assurance (IOTA) certificates using OIDC based authentication [6], and access non-web resources using LDAP-Façade [7].

The SSH pilot leverages the use of the WaTTS Token Translation Service developed by KIT in the context of the INDIGO - DataCloud project. WaTTS was developed to address the users' need to access services that cannot directly utilise federated access and require that the users use security tokens, such as SSH keys, X.509 certificates, S3 access tokens etc. In this pilot, WaTTS was integrated with the EGI CheckIn service, so that users can access WaTTS using their EGI accounts, while authenticating either at their home organisations or using their social IDs. With WaTTS, users are able to manage the SSH access to a number of trusted VMs from a single point in a secure and user-friendly manner. Although in this pilot WaTTS is integrated with the EGI CheckIn service, the solution is not limited to EGI, and can be used at any community or infrastructure service that has a need to "bridge" between different technologies, and can be run as a standalone "plug-and-play" solution. The only requirement is that the community or infrastructure service should support OIDC services.

The IOTA certificates pilot implements a plugin for the WaTTS service, which allows an end user to login to WaTTS using an OIDC provider service and obtain an IOTA-compliant digital certificate from Rcauth.eu online certification authority.

The non-web resources access pilot aims at providing access to non-web resources via Pluggable Authentication Modules (PAM) (e.g. for access to sftp, SSH, console) for users by exploiting the existing AAls, without the need to modify existing software or to obtain user certificates. The implementation is based on LDAP-Facade, a solution for access to non-web resources (e.g. available via SSH protocol) implemented by KIT.

In addition, SCC participates in the development of accounting and data protection policies [8], as well as in user training activities [9]. Here we have analysed the new European General Data Protection Regulation (GDPR, "Datenschutzgrundverordnung") with respect how these new policies can be used to support scientific use-cases.

More information on the project:  
[aarc-project.eu/](http://aarc-project.eu/)



SCC also contributed to the elaboration of the proposal for Attribute Level of Assurance (LoA) of B2ACCESS group of the EUDAT

project. The draft document represents a major effort to define the format, means of communication, and processing of LoAs of authentication and authorisation attributes exchanged between components of a federated AAI system. The use of LoAs allows Service Providers to offer fine-grained, customised services to the end user, based on the "strength" of the authentication method applied by the authenticating Identity Provider.

More information on the project:  
[www.eudat.eu/services/b2access](http://www.eudat.eu/services/b2access)

### SCC europaweit tätig

Das Steinbuch Centre for Computing ist in unterschiedlichen, von der Europäischen Kommission geförderten, Projekten tätig. Ein Schwerpunkt der Arbeiten liegt hier auf der Entwicklung von skalierbaren, sicheren IT-Förderungen und deren Einsatz in der Praxis, um Wissenschaftlern aus europäischen Forschungs- und Lehreinrichtungen einen einfachen und sicheren Zugriff auf e-Infrastrukturen, Daten- und Kollaborationsdienste zu ermöglichen. Weiterhin bringt das SCC seine Expertise in der Bereitstellung und dem Betrieb von virtuellen Prototyp-Infrastrukturen ein. Für die international notwendigen Standards im Datenmanagement konzipiert, entwickelt und erweitert das SCC wichtige Grundlagen.

[1] [www.rd-alliance.org/](http://www.rd-alliance.org/), last checked: 10.07.2017

[2] [www.snia.org/](http://www.snia.org/), last checked: 10.07.2017

[3] [www.unicore.eu/](http://www.unicore.eu/), last checked: 10.07.2017

[4] [aarc-project.eu/blueprint-architecture/](http://aarc-project.eu/blueprint-architecture/), last checked: 10.07.2017

[5] [wiki.geant.org/display/AARC/WaTTS+SSH+plugin+-+SSH+access+using+OIDC+login](http://wiki.geant.org/display/AARC/WaTTS+SSH+plugin+-+SSH+access+using+OIDC+login), last checked: 10.07.2017

[6] [wiki.geant.org/display/AARC/WaTTS+Rcauth+plugin+-+Obtaining+IOTA+proxy+certificates+with+OIDC](http://wiki.geant.org/display/AARC/WaTTS+Rcauth+plugin+-+Obtaining+IOTA+proxy+certificates+with+OIDC), last checked: 10.07.2017

[7] [wiki.geant.org/display/AARC/LDAP+Facade+CLI+pilot](http://wiki.geant.org/display/AARC/LDAP+Facade+CLI+pilot), last checked: 10.07.2017

[8] [aarc-project.eu/aarc-infoshare/](http://aarc-project.eu/aarc-infoshare/), last checked: 10.07.2017

[9] [wiki.geant.org/display/AARC/Accounting+and+Data+Protection](http://wiki.geant.org/display/AARC/Accounting+and+Data+Protection), last checked: 10.07.2017



## In vivo – In vitro – IN SILICO

Ende 2016 machten Schülerinnen und Schüler der Begabten AG des Albertus-Magnus-Gymnasiums Ettlingen eine Exkursion ans Steinbuch Centre for Computing, um mehr über den Einsatz von Computersimulationen in der Biomolekülforschung zu erfahren.

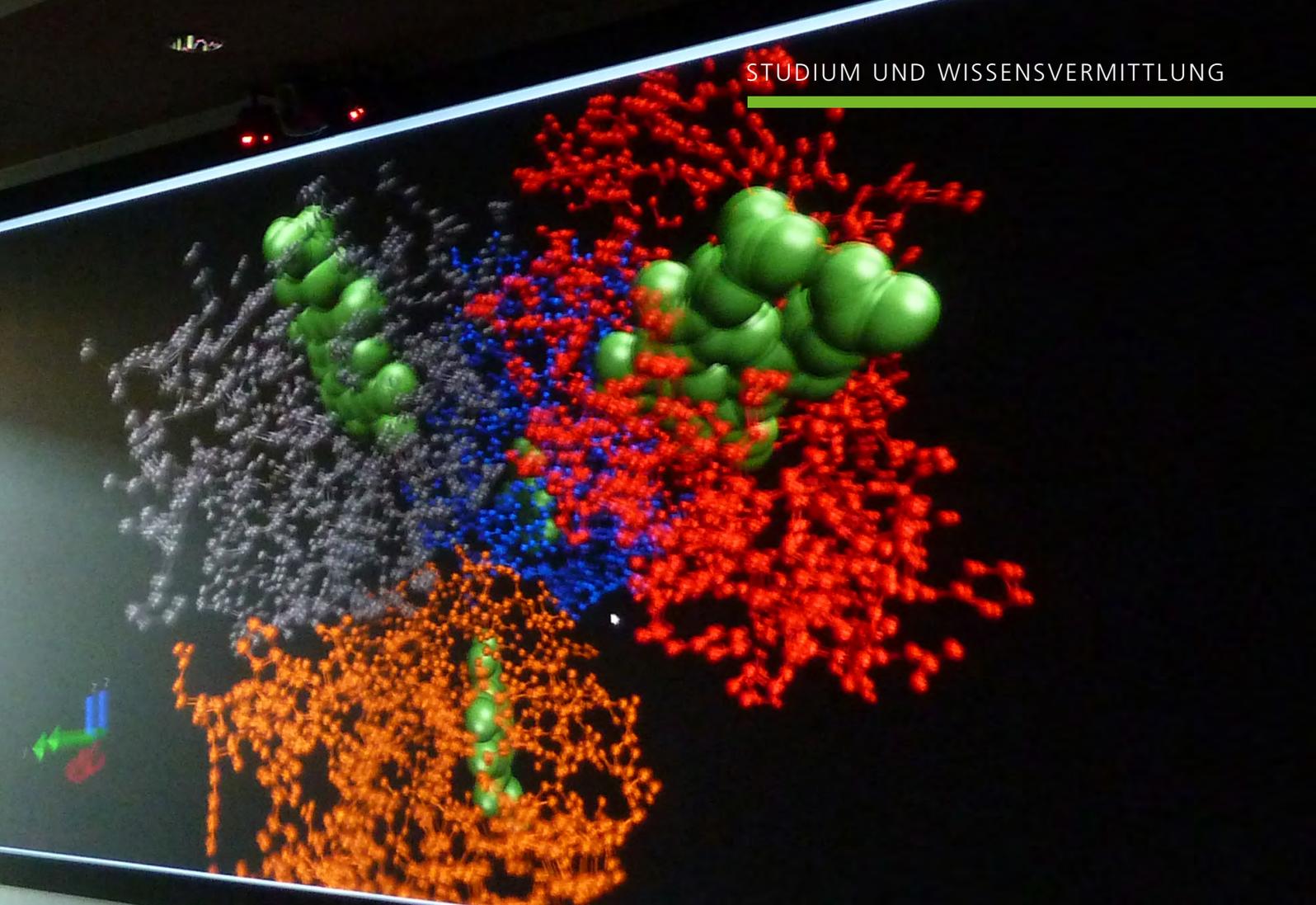
*Oliver Abel'*

Die Begriffe *in vivo*, *in vitro* und *in silico* beschreiben die Methodik des Erkenntnisgewinns in den Biowissenschaften. *In vivo* meint Beobachtungen am lebenden Organismus und *In-vitro*-Verfahren simulieren biochemische Prozesse unter Laborbedingungen. Beide Forschungsmethoden lassen jedoch viele Fragen offen, da nicht alles, was man gerne wissen möchte, sich auch im Detail beobachten lässt. Deshalb versucht man Erkenntnislücken mit Computersimulationen zu schließen. Verfügbare Daten werden *in silico*, also im Computer, mit den eigenen Hypothesen verknüpft. Im intensiven Vergleich mit der *In-vitro*- bzw. *In-vivo*-Situation kann *in silico* immer weiter ergänzt und verfeinert werden.

Schülerinnen und Schüler der Begabten AG des Albertus-Magnus-Gymnasiums Ettlingen nutzten am 20. Dezember 2016 die Chance, einem *In-silico*-Forscher bei der Arbeit zuzusehen, indem sie Alexander Schugs Arbeitsgruppe am SCC auf dem KIT Campus Nord besuchten. Der Physiker Schug, der sich übrigens für seine eigene Schulzeit auch so eine Begabten AG gewünscht

hätte, erstellt zusammen mit seinem Team Computersimulationen für Wissenschaftler, die sich für biochemische und biophysikalische Prozesse interessieren.

Wir wurden sehr herzlich empfangen und bekamen in einem inspirierenden Vortrag einen Überblick über die eigenen Forschungsschwerpunkte vermittelt. Besonders intensiv befasst sich die Arbeitsgruppe Multiscale Biomelecular Simulation mit der Entstehung der 3-dimensionalen Struktur von Eiweißstoffen, den Proteinen. Ein Protein besteht aus einer mehr oder weniger langen Molekülkette, die sich 3-dimensional faltet, um Funktionen wie z.B. Muskelbewegung, Erkennen eines Hormons, Sauerstofftransport im Blut und viele andere wahrnehmen zu können. Mit 3D-Brillen ausgestattet, durften auch wir Proteine in der großflächigen Projektion des Visualisierungslabors erleben. Es wurde sehr eindrucksvoll klar, Proteinfaltungen sind komplex und laufen äußerst schnell in einer mikroskopisch kleinen Umgebung ab. Das Erstellen einer umfangreichen Simulation dazu



erfordert enorme Rechenkapazitäten. Diese stellt im KIT u.a. der Supercomputer ForHLR II zur Verfügung, welchen wir ebenfalls besichtigen durften.

Nach dem Mittagessen durften die Schülerinnen und Schüler mit Unterstützung von Doktoranden der Arbeitsgruppe selbst am PC Proteine falten. Fast hätte man sie vom Campus Nord vertreiben müssen, als unser Exkursionstag zu Ende ging. Ob das virtuelle Austüfteln von möglichen Proteinfaltungen die Faszination für Pokemon Go & Co überbieten kann? Dabei sollte man nicht vergessen, dass die Entwicklung von Computerspielen zu einer immer leistungsfähigeren Hardware geführt haben, wovon nicht zuletzt die In-silico-Forschung profitiert.

Auch bleibt auf diesem Gebiet noch sehr viel zu tun für die angehende junge Forschergeneration. So sind zum Beispiel die detaillierten Wirkmechanismen einer Vielzahl von Medikamenten (auch von Aspirin) nicht im Detail verstanden. Ebenso im Detail ungeklärt ist die Frage, wie sich Alkohol auf Nervenzellen auswirkt. Aber so ist Forschung: je mehr man herausfindet, umso mehr Fragen wirft man auf. Die Schülerinnen und Schüler bekamen (vielleicht in nicht ganz uneigennütziger Hoffnung) gesagt: „Diese Fragen klärt dann ihr!“

Für unsere Schüler ergaben sich viele wichtige und nachhaltige Eindrücke: Motivation, Inspiration, Funktion eines großen Cam-

pus, interdisziplinäre Kooperation, Internationalität der Forschergemeinschaft, Verlauf einer Forscherkarriere, Kommunikation unter Wissenschaftlern, Hightech und nicht zuletzt Offenheit und Respekt im gegenseitigen Umgang. Dafür wollen wir uns sehr herzlich bei Herrn Dr. Schug und seiner Arbeitsgruppe bedanken.



<sup>1</sup> Biologielehrer am Albertus-Magnus-Gymnasium

<sup>2</sup> [www.scc.kit.edu/ueberuns/jrg-mbs.php](http://www.scc.kit.edu/ueberuns/jrg-mbs.php)

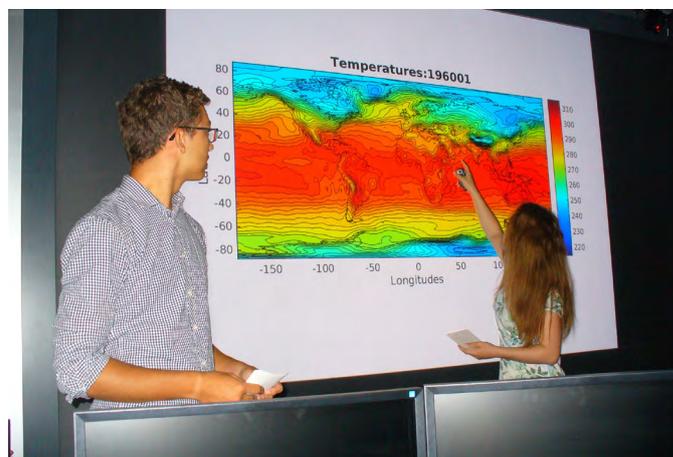
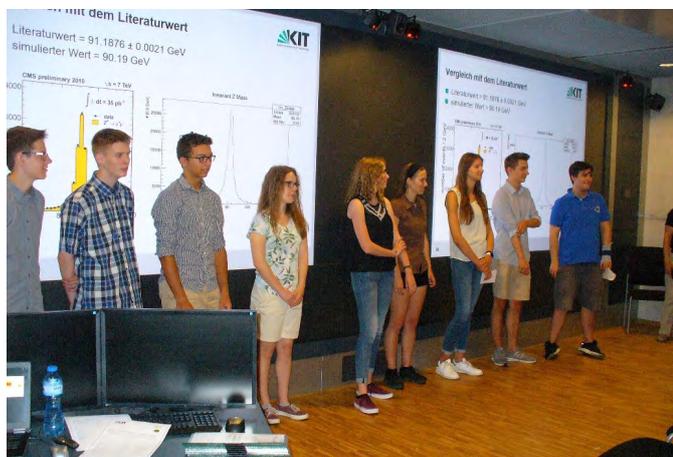
## Simulierte Welten: Schülerinnen und Schüler präsentieren ihre Projekte

Im öffentlichen Schülerkolloquium präsentierten Stipendiatinnen und Stipendiaten des Projekts Simulierte Welten die Ergebnisse ihrer Simulationsprojekte. Die Schülerinnen und Schüler kamen dieses Jahr aus Gymnasien in Karlsruhe und Ettlingen. Die Veranstaltung fand am 24. Juni 2017, dem Tag der offenen Tür des KIT am Campus Nord, statt.

*Dr.-Ing. Jordan Denev*

Im vergangenen Schuljahr 2016/2017 wurden zum fünften Mal Schülerinnen und Schüler der 11ten Klasse Karlsruher Gymnasien von Wissenschaftlerinnen und Wissenschaftlern des SCC in die spannende Welt der Simulationen und des Supercomputings eingeführt. Dies erfolgte im Rahmen des vom Wissenschaftsministerium in Baden-Württemberg finanzierten Projekts „Simulierte Welten“. Ein Ziel des Projekts ist es, Simulationen in das Ausbildungssystem der Schulen zu integrieren. An den diesjährigen Stipendien haben vier Schülerinnen und fünf Schüler aus unterschiedlichen Gymnasien in Karlsruhe und Ettlingen teilgenommen. Sie bekamen Zugang zu den Supercomputern des SCC und konnten zusammen mit ihren Betreuern Simulationen ausführen, deren Ergebnisse analysieren und Modellannahmen prüfen. Der Themenbereich war sehr breit und umfasste Atmo-

sphärenwissenschaften, Ingenieurwissenschaften, Materialwissenschaften und die Hochenergiephysik. Die Betreuer waren in diesem Jahr die Leiter der Simulation Laboratories (SimLabs) am SCC – Dr. Kirner (Klima und Umwelt), Dr. Piccioni Koch (Energy), Dr. Kondov (NanoMikro) – sowie Dr. Giffels aus der Abteilung Scientific Data Management, für die Hochenergiephysik. Die Themen wurden in regulären Treffen in kleinen Gruppen konsequent durchgearbeitet. Die „jungen Forscher“ bearbeiteten auch Aufgaben selbstständig zu Hause, die dann in den Treffen besprochen wurden. Die neu gewonnenen Erkenntnisse haben die Schüler und Schülerinnen in Präsentationen zusammengefasst und ihren Betreuern, Lehrern, Eltern und Freunden während des Schülerkolloquiums Simulierte Welten am Tag der offenen Tür des KIT vorgestellt.



## Neue Helmholtz-Nachwuchsforschergruppe am SCC

In einem Wettbewerbsverfahren für die Förderung von Nachwuchsforschergruppen in der Helmholtz-Gemeinschaft hat Dr. Hartwig Anzt erfolgreich eine Nachwuchsgruppe zum Thema „Fixed-Point Methods for Numerics at Exascale (FiNE)“ eingeworben und leitet diese seit 01.05.2017 am Steinbuch Centre for Computing. Zuvor war Dr. Anzt vier Jahre lang Mitarbeiter in Prof. Jack Dongarras Innovative Computing Lab an der University of Tennessee (USA). Dort entwickelte er unter anderem die MAGMA-sparse Software, die zum Lösen dünnbesetzter Gleichungssysteme eingesetzt wird. Sein Forschungsgebiet umfasst Numerische Lineare Algebra, Vorkonditionierer, GPU-Computing und Energie-effizientes Rechnen.

Dr. Hartwig Anzt

Will man Numerische Verfahren effizient auf Computerarchitekturen umsetzen, muss man stets mit der rasanten Hardware-Entwicklung Schritt halten und die Algorithmen – falls notwendig – auf die neuen und zukünftigen Hardwarearchitekturen anpassen. Der derzeitige Trend zeigt, dass kommende Computersysteme mit einer deutlich höheren Anzahl von Prozessoren ausgestattet sein werden als bisher und sich so bereits auf einem einzelnen Rechenknoten eine große Anzahl an Recheneinheiten (Cores) häufen werden. Einen wesentlichen Beitrag zu diesem Trend liefern Hardware Beschleuniger wie Multicoreprozessoren Intel Xeon Phi oder Grafikprozessoren (GPU). Diese ermöglichen Hunderte oder sogar Tausende parallel laufender Rechenabläufe, sog. Threads. Sie erfordern aber auch eine Umformulierung der numerischen Verfahren: Ziel ist es, die Knoten gleichmäßig auszulasten – die einzelnen Fließkomma-Rechenoperationen (FLOP) werden dabei immer günstiger, die Kommunikation zwischen den einzelnen Prozessoren und die Synchronisation aller Instanzen untereinander wird andererseits immer teurer.

Dabei stößt das traditionelle Bulk-Synchronous Programmiermodell (BSP) an seine Grenzen. Auch die Verwendung DAG-basierter Algorithmen\*), bei denen die parallele Abarbeitung von Tasks und deren Abhängigkeiten durch Laufzeit-Scheduling verwaltet werden, erfordert ausreichende Parallelität sowie eine übersichtliche Anzahl von zu verwalten den Tasks.

Mit zunehmender Hardware-Parallelität bekommen Aspekte wie Reproduzierbarkeit, Bit-Flips und Hardwareausfall eine neue Bedeutung. Bereits heute spielen Bit-Flips und Hardwareausfall bei Hochleistungsrechnern eine Rolle und wenn man die Häufigkeit des Auftretens von Fehlern extrapoliert, werden Checkpoint-and-Restart Techniken bald an ihre Grenze kommen: Das Speichern oder Wiedereinlesen eines Checkpoints wird zu lange dauern, so dass ein erneuter Bit-Flip den Algorithmus wieder unterbrechen könnte.

In seiner Forschung untersucht Dr. Hartwig Anzt mit seinem Team Algorithmen, die für diese Herausforderungen attraktive Eigenschaften besitzen: Skalierbarkeit, reduzierte Kommunikation- und Synchronisation und eine gewisse Resilienz gegenüber Fehlern.

Die Nachwuchsforschungsgruppe „Fixed-Point Methods for Numerics at Exascale“ fokussiert sich auf den vielversprechenden Ansatz, numerische Verfahren als Fixpunkt-Iteration zu formulieren. Die zugrunde liegende Idee basiert auf der mathematischen Theorie zu asynchronen Iterationsverfahren, die bei geeigneten Voraussetzungen asymptotisch konvergieren.

\*) In DAG-basierte Algorithmen lassen sich die Datenabhängigkeiten zwischen den einzelnen Arbeitspakete (Tasks) in Form eines gerichteten Abhängigkeitsgraphen („Directed Acyclic Graph“, DAG) darstellen.



Hartwig Anzt (Foto: Privat)

### New Helmholtz Young Investigator Group at SCC

With the research topic “Fixed-Point Methods for Numerics at Exascale (FiNE),” Dr. Hartwig Anzt successfully applied for research funding at the Helmholtz Foundation. In May, Dr. Anzt started his Helmholtz Young Investigator Group at the Steinbuch Centre for Computing. Previously, he was part of Prof. Jack Dongarra’s Innovative Computing Lab at the University of Tennessee (USA). His research interests are in designing numerical methods for high-performance computing on multi- and manycore architectures. He is coauthor of more than 50 journal and conference papers, and contributed to several open source linear algebra packages.

### Weitere Informationen

[www.scc.kit.edu/ueberuns/jrg-fine.php](http://www.scc.kit.edu/ueberuns/jrg-fine.php)

## Supercomputer des KIT in jeder Hinsicht hervorragend

Der neue Supercomputer des Karlsruher Instituts für Technologie (KIT) ist nicht nur superschnell, sondern auch supersparsam: Beim Deutschen Rechenzentrumspreis 2017 belegte der im vergangenen Jahr in Betrieb gegangene Forschungshochleistungsrechner ForHLR II nun den ersten Platz in der Kategorie „Neu gebaute energie- und ressourceneffiziente Rechenzentren“. Mit mehr als 24.000 Rechenkernen verfügt er im neu gebauten Rechenzentrum über ein besonders energiesparendes Kühlsystem. Er kostete 26 Millionen Euro.

*Dr. Felix Mescoli*

Federführend gefördert hat das 26-Millionen-Projekt das Land Baden-Württemberg gemeinsam mit dem Bund. Die Hälfte der bereitgestellten Mittel kam aus dem Landeshaushalt: „ForHLR II spielt eine wichtige Rolle in unserer Landesstrategie für das Höchstleistungsrechnen. Dass der Rechner auch ein Öko ist, freut uns natürlich umso mehr. Das ist ein weiterer Beleg dafür, dass Leistung und Ressourceneffizienz Hand in Hand gehen können“, sagt Baden-Württembergs Wissenschaftsministerin Theresia Bauer.

„Wer in der internationalen Spitzenforschung ganz vorne mit dabei sein will, braucht höchste Rechenleistung und Speicherkapazität“, sagt der Präsident des KIT, Professor Holger Hanselka.

„Dass hohe Rechenkapazität und neueste Visualisierungstechnik für modernste Simulationsmethoden dabei mit enorm niedrigem Energieverbrauch einhergehen, folgt der Strategie des KIT und ist Ausdruck unserer Stärke in den gesellschaftlich relevanten Bedarfsfeldern.“

„Ohne die Unterstützung aus dem Ministerium für Wissenschaft, Forschung und Kunst und auch durch die Ministerin selbst, wäre es nicht in dieser Form möglich gewesen, die Energieeffizienz zu einem besonderen Schwerpunkt des Projektes zu machen“, sagt Professor Bernhard Neumair, Direktor des Steinbuch Centre for Computing (SCC). So aber füge sich der Betrieb des Hochleistungsrechners perfekt in das auf Kraft-Wärme-Kälte-Kopplung basierende Gesamtkonzept für die Energie-Versorgung des KIT.

Wissenschaft erzeugt heute stark wachsende Datenmengen, die nicht nur verarbeitet und gespeichert, sondern auch visualisiert werden müssen. Mit dem ForHLR II steht Forschern aus ganz Deutschland ein Petaflop-System mit mehr als 1.170 Knoten, über 24.000 Rechenkernen und 74 Terabyte Hauptspeicher zur Verfügung. Ein Petaflop entspricht einer Billion Rechenoperationen pro Sekunde. Damit ist der ForHLR II noch zweieinhalb Mal leistungstärker als der ForHLR I, der seit 2014 am KIT in Betrieb ist.

Das neue Rechenzentrumsgebäude verfügt durch eine Warmwasserkühlung mit bis zu 45 Grad über neueste Kühltechnologie für einen besonders energieeffizienten Betrieb: Diese nutzt in der kalten Jahreszeit die Abwärme des Systems zur Heizung der Bürogebäude und sorgt ganzjährig für eine zuverlässige Kühlung aller heißen Systemkomponenten. Damit kommt das System praktisch ohne



*Ausgezeichnete Energieeffizienz: Projektkoordinator Professor Rudolf Lohner nimmt den Rechenzentrumspreis entgegen. (Foto: marconing GmbH)*



Die glücklichen Gewinner des Deutschen Rechenzentrumspreises 2017. Professor Rudolf Lohner hält den Preis in Händen, daneben die Direktoren des SCC, Professor Bernhard Neumair (links) und Professor Achim Streit (rechts).

energieintensive zusätzliche Kältemaschinen aus. Für diejenigen Komponenten, die noch eine klassische Kaltluftkühlung benötigen, wird das im Aufbau befindliche Fernkältenetz am Campus Nord des KIT über eine Kraft-Wärme-Kälte-Kopplung eine noch wirtschaftlichere und umweltfreundlichere Lösung bieten.

Die Konzeption eines umweltschonenden Kühlsystems sei in einer der wärmsten Gegenden Deutschlands eine besondere Herausforderung gewesen, sagt Professor Rudolf Lohner vom SCC, der das Projekt vom Start bis zur Verwirklichung begleitete und in der Endphase koordinierte. Sogenannte Nassrückkühler, bei denen Wasser auf der Oberfläche von Kühlkörpern verdunstet und so deren Inhalt kühlt, kamen laut Lohner wegen des hohen Wartungsaufwandes bei Pollenflug und der Anfälligkeit für Besiedelung durch Bakterien nicht infrage. „Also mussten wir auf Trockenrückkühler zurückgreifen.“ Deren geringere Kühlleistung habe durch eine besondere Größe der Konstruktion kompensiert werden können.

<sup>1</sup> Autor der zugehörigen KIT-Pressinformation

Viele Fragestellungen, die bei Planung und Bau eines so komplexen Projekts wie des Rechenzentrums auftreten, haben nur Dank sehr enger Zusammenarbeit mit den Partnern und Abteilungen innerhalb des KIT bewältigt werden können, betont Lohner. „Die höchste Auszeichnung für dieses gelungene Vorhaben durch ein bundesweites Fachgremium ist ein Beweis für das erfolgreiche Zusammenwachsen des KIT und die dadurch entstandenen Synergien“, findet er.

Der Deutsche Rechenzentrumspreis wird alljährlich während Deutschlands größtem Rechenzentrums-Kongress future thinking verliehen, der in diesem Jahr am 25. und 26. April in Darmstadt stattgefunden hat.

#### Weitere Informationen

ForHLR II in den SCC-News 1/2016

[www.youtube.com/watch?v=UKGOVchpM5E](https://www.youtube.com/watch?v=UKGOVchpM5E)

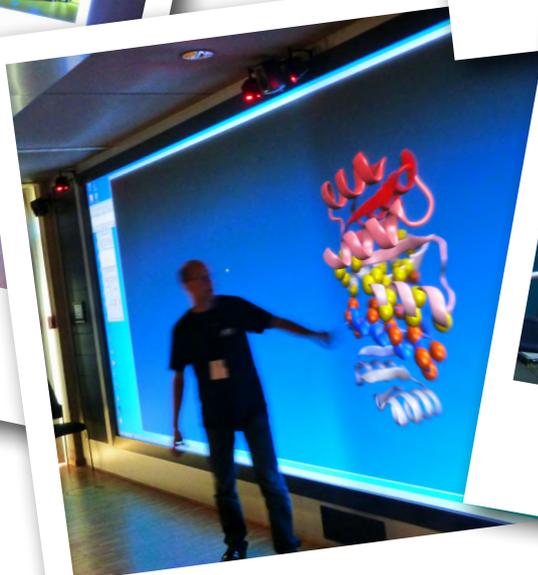
[www.scc.kit.edu/dienste/forhlr.php](http://www.scc.kit.edu/dienste/forhlr.php)

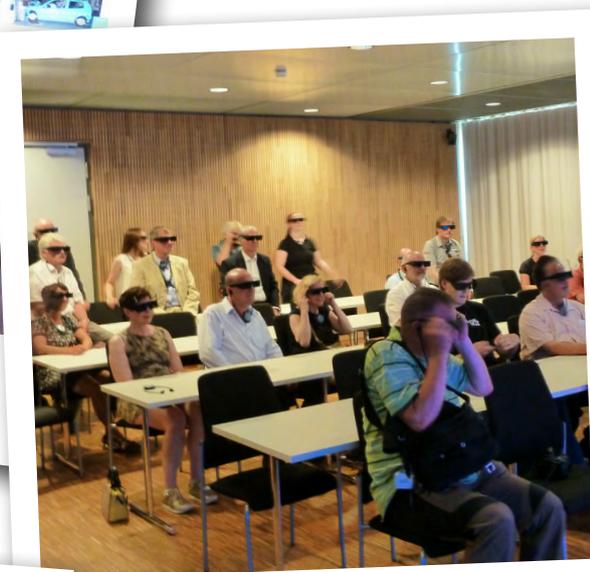
[www.future-thinking.de/news/drzp-gala-journal-2017/](http://www.future-thinking.de/news/drzp-gala-journal-2017/)

## Tag der offenen Tür am KIT – Impressionen aus dem SCC

Rund 35.000 Besucherinnen und Besucher waren am 24. Juni 2017 am KIT Campus Nord begeistert von Forschung, Lehre und Innovation zum Anfassen. Das SCC öffnete seine Tore sowohl zum preisgekrönten Forschungshochleistungsrechner als auch zur Big Data-Infrastruktur rund um das Forschungsgrößgerät GridKA und das Smart Data Labor. 250 Gäste nahmen an den Vorträgen und Führungen teil. An Programmierstationen konnten Kinder und Jugendliche ihre Talente im Umgang mit Software testen oder nach selbst erstellten 3-D-Bauplänen Lego-Autos zusammensetzen. Das Kunstwerk „Wachsender Stein“ am ForHLR-Neubau, Teil der Kinder-Forscher-Rallye, fand ebenfalls großes Interesse. Für das SCC eine wunderbare Gelegenheit, der Öffentlichkeit die faszinierende Welt der Computerwissenschaften nahezubringen.

*(Text und Fotos: K. Schäufele, A. Grindler)*





## IMPRESSUM

**SCC news**  
Magazin des Steinbuch Centre  
for Computing

**Herausgeber**  
Präsident Professor Dr.-Ing.  
Holger Hanselka  
Karlsruher Institut für Technologie (KIT)  
Kaiserstraße 12  
76131 Karlsruhe

**Anschrift**  
Steinbuch Centre for Computing (SCC)  
Karlsruher Institut für Technologie (KIT)  
Redaktion SCC-News  
Zirkel 2  
76131 Karlsruhe  
oder:  
Hermann-von-Helmholtz-Platz 1  
76344 Eggenstein-Leopoldshafen  
Fax: +49 721 608-24972

**Redaktion**  
Achim Grindler (ag, verantwortlich)  
E-Mail: [redaktion@scc.kit.edu](mailto:redaktion@scc.kit.edu)

**Gestaltung, Satz und Layout**  
Heike Gerstner  
AServ – Crossmedia – Grafik  
Karlsruher Institut für Technologie (KIT)  
Hermann-von-Helmholtz-Platz 1  
76344 Eggenstein-Leopoldshafen

**Titelfoto**  
Irina Westermann

**Fotos**  
Markus Breig, Jordan Denev,  
Andreas Drollinger, Achim Grindler,  
Karin Schäufele

**Druck**  
Systemedia GmbH, 75449 Wurmberg

**Erscheinungstermin dieser Ausgabe**  
August 2017

[www.scc.kit.edu/publikationen/scc-news](http://www.scc.kit.edu/publikationen/scc-news)

Der Nachdruck und die elektronische  
Weiterverwendung sowie die Weiter-  
gabe von Texten und Bildern, auch  
von Teilen, sind nur mit ausdrücklicher  
Genehmigung der Redaktion gestattet.



Karlsruher Institut für Technologie (KIT)  
Steinbuch Centre for Computing (SCC)

ISSN: 1866-4954

[www.scc.kit.edu](http://www.scc.kit.edu)  
[contact@scc.kit.edu](mailto:contact@scc.kit.edu)