



GitLab – zentraler Dienst
für kollaborative
Softwareprojekte

*GitLab – Core Service
for collaborative
Software Projects*

Dateneffizientes
Maschinelles Lernen
von Biomolekülen

*Data Efficient
Machine Learning
of Biomolecules*

New Work am SCC:
gemeinsam, flexibel,
agil – innovativ

*New Work at SCC:
collaborative, flexible,
agile – innovative*

Liebe Leserin, lieber Leser,

am Beginn eines Jahres richten wir den Blick noch einmal zurück auf vergangene Erfolge und Herausforderungen, auf schöne und erfüllende, auf schwierige und anstrengende Tage und Wochen. Mit dieser Ausgabe möchten wir gemeinsam mit Ihnen auf das Geleistete zurückblicken und wie dies auf Bestehendes aufbaut. In den letzten drei Jahren hat sich wegen der Corona-Pandemie so einiges in unserer Arbeitswelt getan – Stichwort Homeoffice, Virtuelle Meetings und Co. Dies führte am KIT zu der Frage, wie können wir die Vorteile aus der „alten Welt“ mit einer „New Work“ sinnvoll kombinieren, um so sogar zu mehr Zufriedenheit, Agilität und Flexibilität zu gelangen (S. 33).

Vor fast zehn Jahren hat das SCC schon eine zentrale Plattform für die kollaborative Entwicklung von Software am KIT betrieben, nach und nach verbessert, und nun, nach einem stark angewachsenen Bedarf aus den Projekten, deutlich modernisiert und ausgebaut – ein langer Weg (S. 18).

Wenn wir das Hochleistungsrechnen in Karlsruhe von seinen Anfängen bis heute betrachten, gehen wir sogar 40 Jahre zurück, denn 1983 ging an der Universität Karlsruhe der erste „Supercomputer“ in Betrieb. Seither wurden mehr als 30 Hochleistungsrechner am KIT und den beiden Vorgängereinrichtungen betrieben, was wir mit einem Festkolloquium gebührend feierten (S. 35 und Titelseite).

Manchmal lohnt es sich jedoch, noch weiter zurückzuschauen, und sich zu fragen, ob sich hinter den uralten Prozessen der Evolution auch clevere Ansätze verbergen, die uns helfen effiziente Algorithmen zur Lösung aktueller Probleme zu entwickeln. Ein neues Paper von SCC-Forschenden über dateneffizientes maschinelles Lernen von Biomolekülen bestätigt genau das (S. 24).

Und nun noch etwas in eigener Sache: Ab dem 1. Januar 2024 blicken wir mit neuem Namen – Scientific Computing Center – aber weiterhin gleichem Kürzel SCC (S. 39) in die Zukunft und wünschen Ihnen

Viel Freude beim Lesen

Martin Frank, Martin Nußbaumer, Bernhard Neumair, Achim Streit

Dear reader,

At the beginning of a new year, we can reflect on the successes and challenges of the past year, on days and weeks that were beautiful and fulfilling, and on those difficult and exhausting. In this issue, we want to look back with you at what we have realized and how it builds on what we have already achieved. Over the past three years, the coronavirus pandemic has brought about several changes in our working world - home offices, virtual meetings, etc. At KIT, this has led to the question of how we can combine the advantages of the "old world" with the "new work" in a meaningful way to reach even greater satisfaction, agility, and flexibility in our work (p. 33). Already ten years ago, the SCC was operating a central platform for collaborative software development at KIT. Gradually improved and following a sharp increase in demand from projects the platform is being modernized and expanded - a long way to go (p. 18). When looking at high-performance computing in Karlsruhe from its beginnings 40 years ago to the present day, the first "supercomputer" went into operation at the University of Karlsruhe in 1983. Since then, more than 30 supercomputers have been operated at KIT and its two predecessor institutions, which we properly celebrated with a festive colloquium (p. 35 and front page).

Sometimes however, it is worth looking back even further and asking whether the ancient processes of evolution also reveal clever approaches that help us to develop efficient algorithms to solve current problems. A new paper by SCC researchers on data-efficient machine learning of biomolecules confirms just that (p. 24).

And now on a personal note: from 1 January 2024 we will be looking into the future with a new name - Scientific Computing Center - but still with the same abbreviation SCC (p. 39) and wish you

happy reading

Martin Frank, Bernhard Neumair, Martin Nußbaumer, Achim Streit



DIENTE UND INNOVATION

- 04 Flexible Anmeldeverfahren mit SignMeUp
- 06 Neue Version des Netzdienstverwaltungssystems NETVS
- 09 Neues Helpdesk-System in der EOSC – vom Konzept zur Umsetzung
- 12 Performante Datenmigration zwischen Bandsystemen
- 14 Erneuerung des OpenVPN-Dienstes am KIT
- 16 GitLab – zentraler Dienst für kollaborative Softwareprojekte
- 18 Relaunch des KIT Intranets
- 21 Arbeitskreis Clientmanagement als Wissensmultiplikator

FORSCHUNG UND PROJEKTE

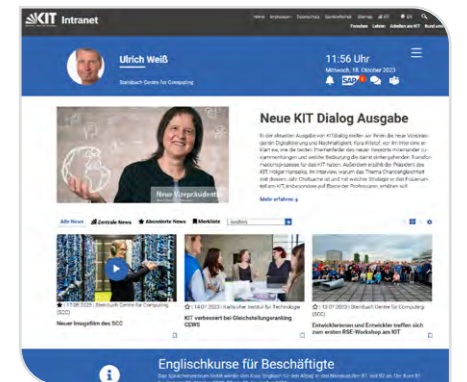
- 21 Abgeschlossene Promotionen
- 22 Kein Weg zu weit – GridKa integriert nachhaltige Computere Ressourcen aus Übersee
- 24 Dateneffizientes Maschinelles Lernen von Biomolekülen
- 26 Einfacher Zugang zu hydrologischen Daten in Deutschland unterstützt die Umweltforschung

STUDIUM UND WISSENSVERMITTLUNG

- 30 Im Zeichen der Zeit – der zweite AI-Hero Hackathon
- 31 12. HPC-Statuskonferenz der Gauß-Allianz zu Gast am KIT

VERSCHIEDENES

- 33 New Work am SCC: gemeinsam, flexibel, agil – innovativ
- 35 KIT feiert 40 Jahre Hochleistungsrechnen in Karlsruhe
- 39 KIT benennt Informationstechnologie-Zentrum um



Flexible Anmeldeverfahren mit SignMeUp

Mit SignMeUp bietet das SCC seit Herbst 2022 ein neues System zur Durchführung von Anmeldeverfahren für Veranstaltungen an. Das System ist dabei vielseitig einsetzbar und ermöglicht den Nutzenden sowohl bei der Anmeldung als auch bei der Platzvergabe einen hohen Grad an Individualisierung, um ein breites Spektrum an Anwendungsfällen abzudecken. Die Nutzung ist für alle Mitarbeitenden des KIT möglich. *Michael Meier*

Die Durchführung von Anmeldungen und Platzvergaben bei der Planung von Veranstaltungen ist eine immer wiederkehrende Tätigkeit. Für einfache Anmeldeverfahren bietet das Campus-System des KIT bereits entsprechende Tools, die sich aber für komplexere Anwendungsfälle wie z. B. die Verteilung auf Termine anhand von Bewertungen oder Losverfahren nicht eignen. Daher waren viele Organisationseinheiten gezwungen, entweder eigene Lösungen zu entwickeln oder auf Drittanbietersoftware zurückzugreifen. Mit der neuen Funktion des Campus-Plus-Portals SignMeUp stellt das SCC nun eine zweisprachige Lösung zur Verfügung, welche auch komplexere Anmelde- und Vergaberegeln unterstützt.

Vielfältige Anwendungsszenarien

SignMeUp eignet sich für eine breite Palette von möglichen Anwendungsszenarien. Dank individuell auswählbarer Anmeldeparameter sowie Platzvergabeverfahren können nicht nur die typischen Anwendungsfälle in der Lehre wie Tutorien, Seminare und Praktika abgedeckt werden, sondern auch Terminvergaben oder -vereinbarungen sowie Exkursionen oder Abstimmungen (Abbildung 1). Die Möglichkeit, den Zugriff auf Anmeldeverfahren auf „Privat“ zu schalten, erlaubt es, die Anmeldung nur einer eingeschränkten Personengruppe zugänglich zu machen.

Unterstützte Platzvergabemethoden

Die Art der Platzvergabe spielt eine wichtige Rolle für die Gestaltung eines Anmeldeverfahrens.

Derzeit werden fünf verschiedene Platzvergabeverfahren unterstützt:

- Die Manuelle Auswahl ermöglicht es, die Teilnehmenden selbst zu bestimmen.
- Über das Losverfahren erhalten alle Anmeldungen eine Losnummer, und die Plätze werden bis zur Kapazität anhand dieser Nummer vergeben.
- Das Fachsemester-Ranking ist ein Losverfahren, welches das Fachsemester in Relation zur Regelstudienzeit berücksichtigt. Anmeldungen in höheren Semestern werden bevorzugt behandelt.
- First Come, First Served vergibt die Plätze nach Eingangszeitpunkt.
- Das Verfahren Optimale Verteilung verteilt Teilnehmende anhand abgegebener Bewertungen möglichst optimal auf verschiedene Themen oder Termine, z. B. Tutorien (Abbildung 2).

Viele dieser Vergabeverfahren unterstützen ein automatisches Nachrücken innerhalb einer festgelegten Frist, falls Absagen erfolgen oder weitere Anmeldungen eingehen.

Anmeldevorgang individuell gestalten

Die Anmeldung kann flexibel an das gewünschte Platzvergabeverfahren

angepasst werden. So gibt es z.B. die Möglichkeit, über ein einfaches Formular weitere Angaben bei der Anmeldung abzufragen. Zusätzlich können verfügbare Termine oder Themen definiert werden, für welche entweder eine Auswahl oder eine Bewertung erfolgen muss. Einige der Vergabeverfahren berücksichtigen dann die Auswahl oder Bewertung bei der Platzvergabe.

Richtet sich eine Anmeldung nur an bestimmte Personengruppen oder Studiengänge, kann dies in den Einstellungen definiert werden. Eine Anmeldung ist in dem Fall nur möglich, wenn die Voraussetzungen dafür erfüllt werden.

Mit SignMeUp können auf Wunsch auch Personen, die dem KIT nicht oder noch nicht angehören, für die Anmeldung zugelassen werden.

Wenn es für eine Veranstaltung sinnvoll ist, können Lerngruppen aktiviert werden. Angemeldete Personen einer Lerngruppe erhalten immer den gleichen Status und bei einer Zusage das gleiche Thema oder den gleichen Termin.

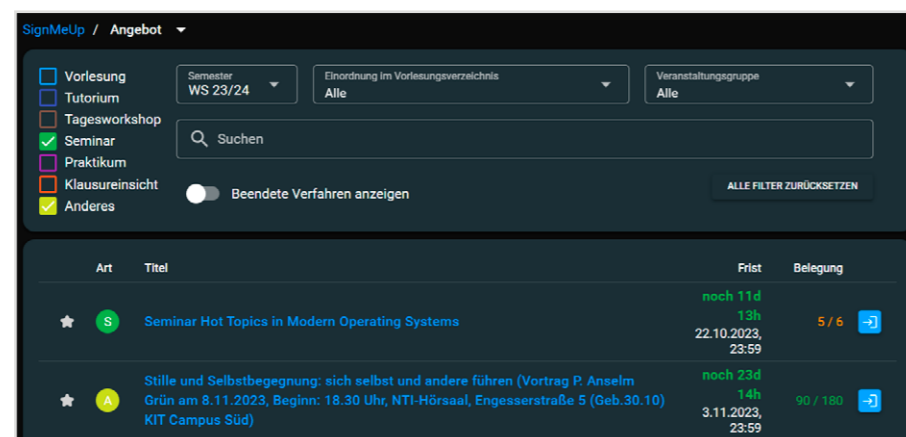


Abbildung 1: Das öffentliche Angebot in SignMeUp

Veranstaltungen gruppieren

Für Nutzende, welche mehrere Veranstaltungen anbieten, wurde in SignMeUp ein Konzept für Veranstaltungsgruppen eingeführt. Eine Gruppe kann mehrere Anmeldeverfahren, beispielsweise für eine Seminarreihe, zusammenfassen und diese so nicht nur einfacher im Angebot auffindbar machen, sondern sie erlaubt es auch, Einschränkungen für diese Gruppe zu definieren. So sind Nutzerinnen und Nutzer z. B. in der Lage, die maximale Anzahl an Anmeldungen und Teilnahmeszusagen innerhalb dieser Gruppe zu beschränken.

Eine neu hinzugekommene Funktion ist die Möglichkeit, sogenannte Boosts zuzulassen, welche es den angemeldeten Personen erlaubt, einzelne Anmeldungen stärker zu gewichten, um abhängig vom Vergabeverfahren die Chance auf einen Platz zu erhöhen.

Für die einfachere Pflege von großen Veranstaltungsangeboten bietet das Tool eine Import- und Export-Schnittstelle, über welche mittels Excel-Dateien neue Anmeldeverfahren angelegt sowie bestehende aktualisiert werden können.

Integration in die Systemlandschaft

SignMeUp ist in das Campus-Management-System des KIT integriert, so dass Nutzerinnen und Nutzer eine bestehende Veranstaltung im Vorlesungsverzeichnis des Campus-Systems mit einem Anmeldeverfahren in SignMeUp verbinden können. Existiert eine solche Verbindung, wird automatisch im Vorlesungsverzeichnis des Studierendenportals sowie bei Auszügen auf Webseiten ein Hinweis auf die Anmeldung über SignMeUp eingeblendet. Die Integration in das Campussystem liefert für die Teilnehmerlisten in SignMeUp automatisch aktuelle Informationen zu den angemeldeten Personen (z.B. Fachsemester und Studiengang bei Studierenden). Eine mühsame Abfrage dieser Angaben entfällt somit.

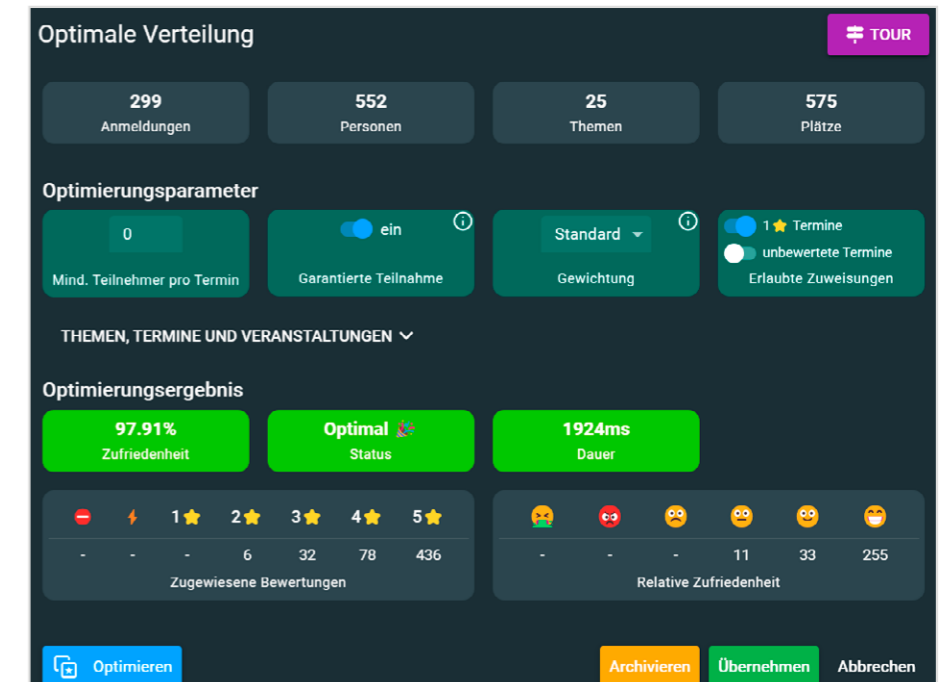


Abbildung 2: Durchführen einer optimalen Verteilung für ein Tutorium

Über einen ILIAS-Export können Teilnehmerlisten in das am KIT etablierte Lernmanagementsystem ILIAS übertragen werden, so dass an einer Lehrveranstaltung teilnehmende Personen einen direkten Zugriff auf den dazugehörigen ILIAS-Kurs erhalten. Auf Wunsch legt der Export im ILIAS-Kurs den Themen oder Terminen entsprechende Kursgruppen an und weist die Teilnehmer diesen zu.

Mit SignMeUp steht ein umfangreiches und flexibles Tool zur Abwicklung von Anmeldeverfahren aller Art zur Verfügung, das sich an den Bedürfnissen und Wünschen der Nutzerinnen und Nutzer orientiert und kontinuierlich weiterentwickelt wird.

Flexible and easy registration processes using SignMeUp.

SignMeUp is a new tool available for all KIT employees to manage registration processes. With multiple supported assignment methods like First Come First Served, Lottery or Optimal Distribution based on ratings the software covers many different use cases. Registration procedures can be further customized for target specific user groups, to request additional information during registration and to allow or disallow external users and study groups. Topics or appointments can be specified to be selected or rated. Procedure groups combine multiple registration procedures and allow the definition of group specific restrictions. Managing multiple procedures is easy with Excel import and export functions. SignMeUp integrates with the Campus Management System to support events in the course catalog and features an ILIAS export.

Neue Version des Netzdienstverwaltungssystems NETVS

Mit einer Vielzahl an Verbesserungen in der Nutzeroberfläche, einigen neuen Features und der Unterstützung von mehreren Sprachen ging die neue Version 4.0 des Netzdienstverwaltungssystems NETVS Ende September an den Start.

Alexander Kaschta

Das SCC betreibt mit dem Netzdienstverwaltungssystem (NETVS) seit Jahren ein hauseigenes Tool für den Netzbetrieb am KIT. Das NETVS steht den IT-Beauftragten und IT-Admins am KIT zur Verwaltung der vielen verschiedenen Netzwerkdienste im KIT-Netz zur Verfügung (s. SCC-News 2/2020).

Dieses Tool wird ständig weiterentwickelt, und so war es am Wochenende vom 30.9.2023 mal wieder soweit: Eine neue Version des NETVS und die NETDB-API in der Version 4.0 wurden veröffentlicht. Dieser Artikel gibt einen Überblick über die Änderungen und neuen Funktionen.

Neue Startseite

Die ersten Änderungen erkennt man schon direkt auf der NETVS-Startseite¹, denn nach dem Login sieht man nun direkt eine Kachelansicht mit einer sys-

temübergreifenden Übersicht über alle selbstverwalteten Broadcast-Domains (also der Adressräume, kurz BCDs). Diese gibt eine Zusammenfassung über alle wichtigen Informationen, und von dort kann man schnell zu den einzelnen Anwendungen (DNSVS, MacAuth usw.) der jeweiligen Broadcast-Domain navigieren. Für jede Broadcast-Domain wird nun auch die Sicherheitsstufe angezeigt, was bisher nicht der Fall war. Um die Übersicht zu vereinfachen, wurde eine Filtermöglichkeit implementiert (Abbildung 1).

Zur Authentifizierung, um überhaupt zur Startseite zu gelangen, ist in der neuen Version nun auch ein zweiter Faktor notwendig². Dies erhöht die Sicherheit der Administration der Netzdienste. Weitere Änderungen finden sich unter anderem in der linken Navigationsleiste und in der Menüleiste, wo z. B. die

Einträge Service-Accounts, NETVS Hub und IP-Tools neu dazugekommen sind. Die weiteren Änderungen verstecken sich alle in der Menüleiste. Sollte die neue Startseite nicht zusagen, gibt es über die Nutzereinstellungen in der Menüleiste nun die Option, eine beliebige NETVS-Seite als individuelle Startseite festzulegen. Somit kann man direkt beim Aufruf des NETVS in seine Administrationsumgebung springen.

Die bekannte Funktion Expertenmodus wurde um ein einheitliches Icon erweitert, welches nun jede Expertenaktion visuell hervorhebt. In der Menüleiste informiert nun ein Glockensymbol über Systembenachrichtigungen oder anstehende Wartungen. Dies ist kein Ersatz für die bisherigen Ankündigungen, sondern erweitert diese.

¹ netvs.scc.kit.edu/

² Zwei-Faktor-Authentifizierung am KIT: www.scc.kit.edu/dienste/2fa.php

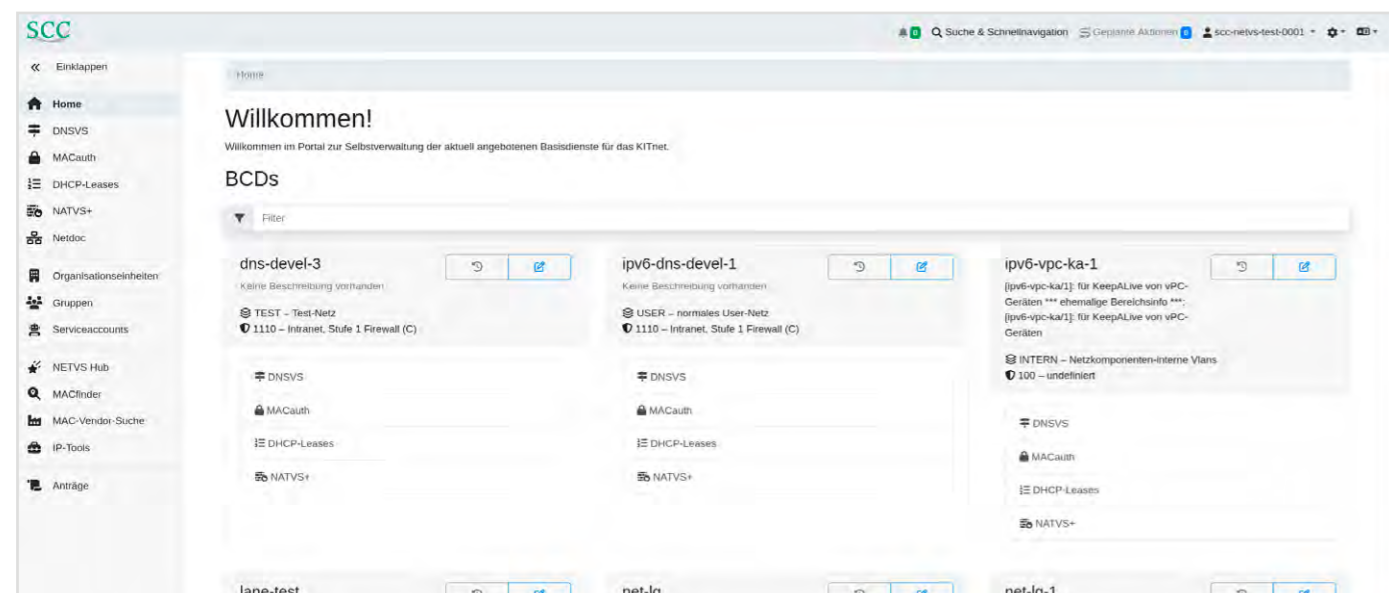


Abbildung 1: Neue Startseite des NETVS mit übersichtlichen Einstiegskacheln

Das NETVS wird mehrsprachig

Mit zunehmender Zahl an englischsprachigen Studiengängen und einer sowieso internationalen Forschungsgemeinde steigt die Anzahl englischsprachiger Nutzende am KIT, die nicht zwingend die Deutschkenntnisse haben, um das Webinterface zu nutzen. Dies ist eine Entwicklung, mit der das SCC mithalten möchte, daher unterstützt das NETVS neben Deutsch nun auch Englisch. Beim ersten Aufruf der Seite wird versucht, anhand der Browserdaten die passende Sprache auszuwählen. Auf Wunsch kann die initiale Spracheinstellung über ein Icon in der Menüleiste geändert und im Browser gespeichert werden. Mit der Umstellung auf verschiedene Sprachvarianten sind die Voraussetzungen gegeben, bei Bedarf nicht nur Englisch und Deutsch, sondern auch weitere Sprachen zu implementieren.

NETVS-Service-Accounts

Eine weitere Neuerung ist die Einführung von sogenannten NETVS-Service-Accounts. Dabei handelt es sich um nicht-personengebundene, NETVS-interne bzw. NETDB-interne Accounts, die vielseitig eingesetzt werden können. Ein Anwendungsbeispiel ist die Automatisierung der ACME-Challenge (Automated Certificate Management Environment) zur Beantragung von Zertifikaten mittels ACME4NETVS³. Diese Service-Accounts können über die Navigationsleiste eingesehen werden. OE-Betreuer können dort die Service-Accounts für ihre OE erstellen und verwalten.

Bei der Erstellung eines Service-Accounts wird nach einer zentralen Berechtigungsgruppe gefragt, welche über die Gruppenverwaltung des KIT angelegt und einer Organisationseinheit zugeordnet sein muss. Der Service-Account wird dann dieser Gruppe zugeordnet. OE-Betreuer können über diese Gruppe steuern, welche IT-Admins auf diesen Service-Account zugreifen und diesen nutzen können.

³ git.scc.kit.edu/KIT-CA/acme4netvs

Für die Nutzung der Service-Accounts wurde ein neuer Modus im NETVS implementiert, der es erlaubt, sich als dieser Account auszugeben (engl. impersonation). Wird dieser Modus aktiviert, stehen dem Service-Account im NETVS nur Funktionen zur Verfügung, für welche er berechtigt ist. Die gleiche Funktion wurde nun auch für die bereits bekannten Unterkonten implementiert.

Eine weitere Neuerung ist der NETVS Hub⁴ unter dem gleichnamigen Eintrag in der Navigationsleiste. Dieser ist eine Sammlung von anwendungsspezifischen Workflows, um komplexere Prozesse im NETVS zu vereinfachen. Jeder Workflow wird durch eine JSON-Datei definiert, welche die benötigten Eingaben angibt und die auszuführenden NETDB-Transaktionen spezifiziert. Die auf den JSON-Dateien basierenden Workflows werden von Nutzenden beigetragen und bieten

⁴ netvs.scc.kit.edu/tools/hub

so die Möglichkeit, sie auch mit anderen Nutzenden zu teilen.

Der erste verfügbare Workflow dient zur Erstellung von NETVS-Service-Accounts für ACME4NETVS. Im Vergleich zu der manuellen Erstellung eines Service-Accounts übernimmt dieser Workflow die Erstellung der Gruppe, fügt die angegebenen Mitglieder hinzu und weist auch die angegebenen FQDNs zu. Damit entfällt der Bedarf für KIT-Service-Accounts, da alle Funktionen mit den NETVS-Service-Accounts umgesetzt werden. Demnächst wird noch ein weiterer Workflow für ACME4NETVS folgen, welcher dann mit dem neu erstellten NETVS-Service-Account für ACME4NETVS ausgeführt wird, um das notwendige Token für ACME4NETVS zu generieren. Neue Workflows aus dem Kreis der Nutzenden sind herzlich willkommen, um so die Effizienz und den Wissenstransfer innerhalb der NETVS-Community weiter zu stärken.

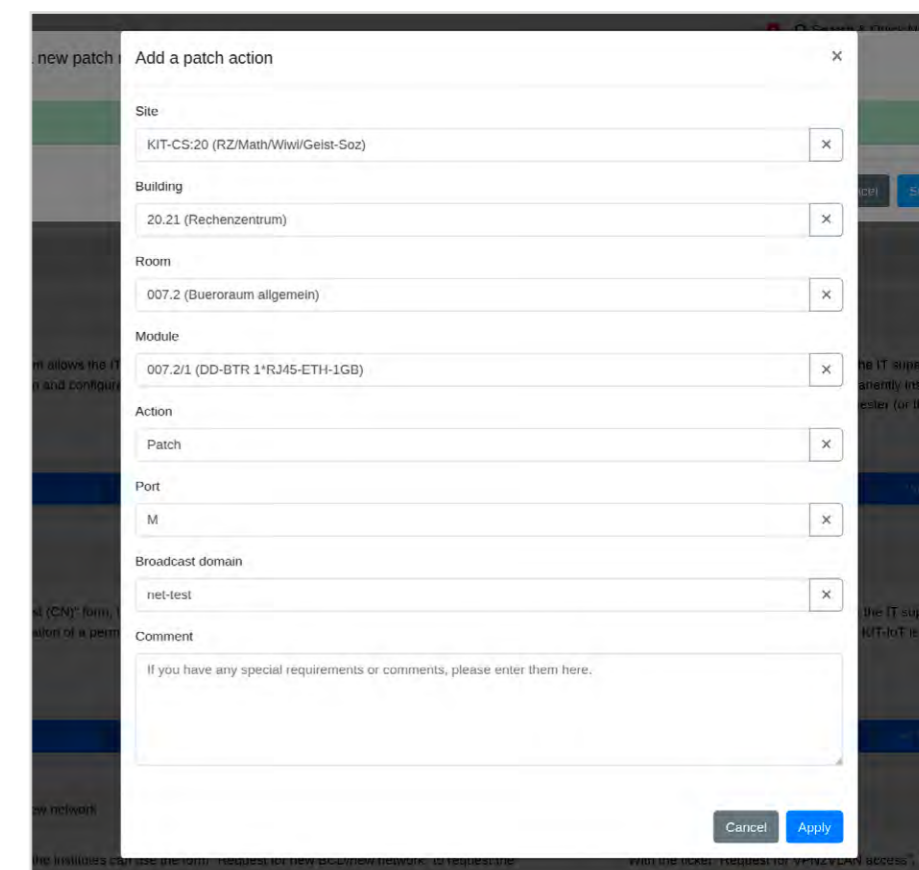


Abbildung 2: Neues Antragsformular für das Auflegen von Netzwerkanschlüssen

Patch-Request-Formular

Dieses Formular (Abbildung 2) vereinheitlicht nicht nur den Prozess für die KIT-Standorte Campus Süd, Nord, Ost und West, sondern unterstützt mit Hilfe von datenbankbasierten Eingabevorschlägen die Auswahl der richtigen Netzwerkdose. Zusätzlich lassen sich nun mehrere Patch-Anfragen in einem Antrag zusammenfassen, sodass zum Beispiel direkt ein Einsatz in der Netzwerkdose getauscht und der Zugriff auf eine neue Broadcast-Domain auf diese Dose geschaltet werden kann.

Weitere neue Funktionen

Nachfolgende Liste beschreibt weitere, interessante Funktionen:

- Mit der neuen NETVS-Version werden nun nicht BCD-bezogene Record-Typen, wie z. B. TXT- oder SSHFP-Records, nicht mehr in der BCD angezeigt. Diese Records haben im DNS-Verwaltungssystem DNSVS, ebenfalls ein Teil des NETVS, einen eigenen, neuen Reiter erhalten.
- Das Anlegen von Records sieht nun etwas anders aus, da API-seitig nichtexistierende FQDNs beim Anlegen eines Records nicht mehr automatisch angelegt werden; das übernimmt jetzt das NETVS. Dafür gibt es jetzt auch die Option, eine bereits bestehende FQDN-Beschreibung beim Anlegen eines weiteren Records mit dem FQDN zu überschreiben.
- Im DNSVS bekommen IP-Adressen mit spezieller Funktion, wie die Gateway-Adressen, Netzwerk-Adressen, Broadcast-Adressen und Router Anycast-Adressen, entsprechende Labels.
- Es gibt nun unter dem Navigationspunkt IP-Tools eine Sammlung an nützlichen Werkzeugen für die Arbeit mit IPv6. Diese können z. B. IPv6-Adressen in die komprimierte und die erweiterte Schreibweise umformen und die Größe von Subnetzen ausrechnen. Diese Werkzeugsammlung ist ein Beispiel dafür, wie aus einem Funktionswunsch der Nutzenden eine NETVS-Implementierung wurde.

Ausblick in die Zukunft

Die Entwicklung am NETVS und dessen Programmierschnittstelle (API) geht weiter, denn es gibt noch eine ganze Reihe an Issues zu lösen und Feature Requests umzusetzen. Das SCC freut sich immer über Rückmeldungen der Nutzenden, denn gerade diese helfen, die Software zu verbessern. Für den regelmäßigen Austausch zwischen Nutzenden und dem NETVS-Entwicklerteam gibt es die KIT-Matrix-Gruppe NETVS Community⁵. Des Weiteren können über den Issue Tracker⁶ Funktionswünsche eingebracht und Fehler gemeldet werden.

Ein Ausblick auf die kommenden Entwicklungen: Die Unterstützung zusätzlicher Sprachen wird weiter fortgesetzt. Die API unterstützt bereits mehrere Sprachen, jedoch müssen die Inhalte noch übersetzt werden, damit z. B. das Eventlog dann auch in Englisch verfügbar ist.

Im Hintergrund wird es noch weitere Änderungen geben. Zum einen steht der Umzug des Sourcecodes auf die neue GitLab-Instanz des KIT⁷ an und die Middleware des NETVS wird umgeschrieben. Dort soll das flask-Framework durch das fastapi-Framework ersetzt werden.

Für das NETVS steht außerdem die Implementierung der DHCP-Funktionalität an. Mit der NETDB-API 4.0 sind nun entsprechende Schnittstellen dafür eingeführt, und es wird derzeit an einer Oberfläche sowie an der neuen DHCP-Server-Architektur entwickelt.

⁵ matrix.to/#/#netvs-community:kit.edu

⁶ git.scc.kit.edu/scc-net/netvs/netvs-core/-/issues

⁷ gitlab.kit.edu/

New Version of the Network Administration System NETVS

Version 4.0 of NETVS and the accompanying NETDB-API were released on September 29th. They deliver many new features to the user interface and the application programming interface (API). One major change is the multilingual support, making the user interface available in both English and German. This support also extends to the API, but there may still be gaps as not all content has been translated yet. Other languages could also be added if needed, but there are currently no plans to add more.

The NETVS service accounts are another notable improvement in this version because they simplify the account management process through an entire organizational unit. These can be used, for example, for the certificate automation of the ACME challenge, so that a KIT service account is no longer necessary. In addition, the new NETVS Hub introduces a user-contributed collection of application-specific workflows to simplify complex processes in NETVS. The Patch-Request form has also received a major update. It is now completely integrated into the NETVS and provides suggestions from the database while typing to help selecting the right network socket.

SCC is always happy to receive feedback from users, as this helps to improve the software. Feature requests and bugs can be reported using the Issue Tracker. Furthermore, there is a KIT Matrix group NETVS Community for regular exchange between users and the NETVS development team.

Neues Helpdesk-System in der EOSC – vom Konzept zur Umsetzung

Das Helpdesk ist eine wichtige Komponente der European Open Science Cloud (EOSC), die einen zentralen Unterstützungsdienst für Nutzende über mehrere Communities und e-Infrastrukturen hinweg bietet. Im Rahmen des EU-Projekts EOSC Future, dem Nachfolgeprojekt von EOSC-Hub, schlug das KIT ein neues Konzept und eine neue Implementierung des bestehenden EOSC-Helpdesk-Systems vor, einschließlich der Integration mit anderen Helpdesk-Systemen und einer Bereitstellung als Dienst für mehrere EOSC-Communities.

Pavel Weber

Das EOSC-Helpdesk (im Folgenden auch kurz Helpdesk) ist der Einstiegspunkt und das Ticketsystem für Serviceanfragen und Problemmeldungen zu den verfügbaren Diensten innerhalb der EOSC. Der etablierte Incident- und Service-Request-Management-Prozess des innerhalb der EOSC-Service-Landschaft angewendeten IT-Service-Management-Standards FitSM¹ definiert klare Abläufe, die im Helpdesk-System umgesetzt werden und eine effiziente und strukturierte Kommunikation zwischen EOSC-Nutzenden und Anbietern von IT-Ressourcen und -Diensten ermöglichen.

Das Helpdesk bietet Nutzenden wichtige Funktionen, die mit einer Anforderungsanalyse identifiziert wurden. Zu nennen sind hier Funktionen wie Self Service, Reporting, flexibles Benachrichtigungssystem und Helpdesk-Portale für EOSC-Diensteanbieter.

Diese Funktionen tragen dazu bei, die Integrität der IT-Infrastruktur und die Qualität der bereitgestellten Dienste zu gewährleisten. Im Rahmen des Projekts EOSC Future wird das Helpdesk als verteiltes System implementiert, das für die effiziente Unterstützung der technischen Cloud-Plattform und der EOSC-Dienste der sogenannten Exchange-Ebene genutzt werden kann. Nachfolgend wird diese komplexe Cloud-Plattform vorgestellt, deren Services und Architektur nun schon fünf Jahre aufgebaut und fortentwickelt werden.

Einblick in die Cloud-Plattform EOSC

Die ersten EOSC-Schichten wurden im EU-Projekt EOSC-Hub entwickelt, worüber das SCC in den SCC-News 1/2020 ab S. 40

ausführlich berichtet hat. Das Folgeprojekt EOSC Future setzt den Aufbau der EOSC und ihrer verschiedenen Komponenten fort. Die erklärte Aufgabe von EOSC Future ist: Die Europäischen e-Infrastrukturen und Wissenschafts-Communities zusammenzubringen, um eine funktionsfähige EOSC-Plattform mit Schwerpunkt auf Technologie und Interoperabilität zu implementieren, Ressourcen zur Verfügung zu stellen sowie die Anforderungen der Nutzenden und deren Erfahrung zu berücksichtigen.

Das hier vorgestellte Helpdesk-System² ist in eine komplexe Servicestruktur eingebettet. Es unterstützt die EOSC-Kerndienste und ist zudem in die unterschiedlichen Community-Helpdesk-Systeme integriert. Abbildung 1 zeigt die Top-Level-Architektur der EOSC-Plattform.

EOSC Core (grün): Diese Ebene beinhaltet die internen Kerndienste und stellt den Betrieb der EOSC-Komponenten sicher. Auf ihre Ressourcen greifen die EOSC-Dienste der Exchange-Ebene zurück (blau). Sie umfasst zudem nicht-technische, aber für den Betrieb wichtige Koordinierungsfunktionen, wie z. B. das Onboarding von neuen Service Providern und das IT-Sicherheitsmanagement.

EOSC Exchange (blau): Diese Ebene umfasst Dienste und Datenquellen, die von Forschungsinfrastrukturen, Wissenschafts-Communities und Forschungsprojekten in die EOSC eingebracht wurden.

Die EOSC Support Activities unterstützen die EOSC-Core- und die EOSC-Exchange-Dienste und umfassen die

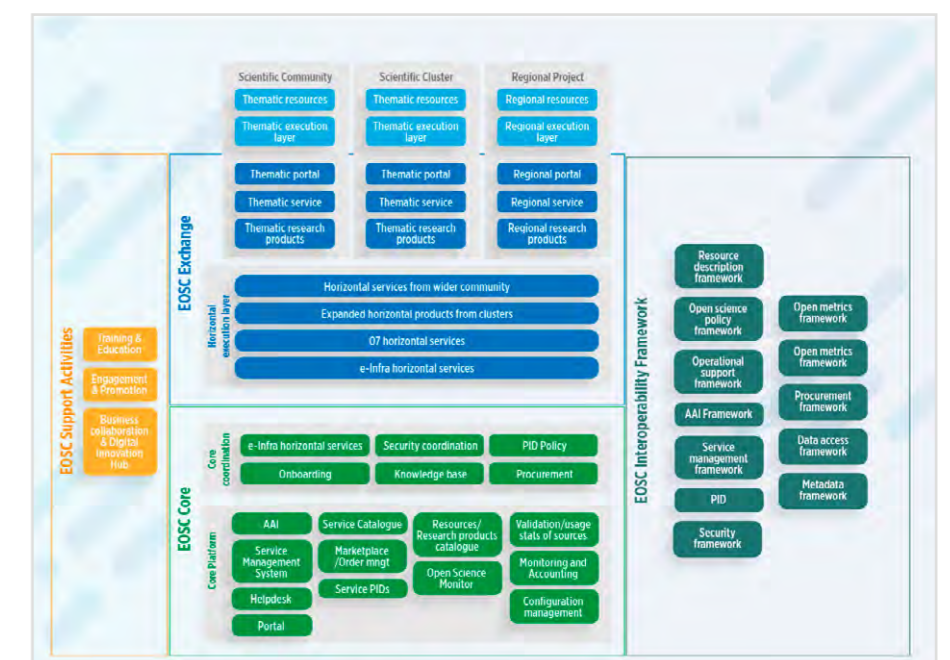


Abbildung 1: Top-Level-Architektur der EOSC

¹ www.fitsm.eu/

² eosc-helpdesk.eosc-portal.eu

Ausbildung, Verpflichtungen und andere nicht betrieblich orientierte Aktivitäten, die die EOSC-Plattform attraktiver und benutzerfreundlicher machen.

Das **EOSC Interoperability Framework (EOSC IF)** bietet beispielsweise Security- und Service-Management-Verfahren und Dienste, die etablierten Standards und Leitlinien zur Unterstützung des EOSC-Betriebs folgen.

Anforderungen an ein modernes Helpdesk-System in der verteilten EOSC-Umgebung

Das bisherige Helpdesk-System basierte auf der in die Jahre gekommene GGUS-Technologie², die im Projekt EOSC-Hub weiterentwickelt und im Projekt EOSC Future durch eine neue, flexiblere Open-Source-Lösung ersetzt wurde.

Die Sammlung und Analyse von Anforderungen aus den EOSC-Communities und dem Technical Coordination Board (TCB) des Projekts EOSC Future war entscheidend, um eine passende Helpdesk-Technologie zu finden und eine Roadmap für die Helpdesk-Erweiterung der nächsten Jahre zu entwickeln.

Das Helpdesk soll für die EOSC nach Planungen im Projekt EOSC Future als Helpdesk-as-a-Service aufgebaut werden, d. h. es soll für neue Anwender- oder Forschungsgruppen schnell eine neue Helpdesk-Oberfläche erstellt werden können. Die Inbetriebsetzung einer Helpdesk-Instanz soll mit geringem personellen Aufwand und ohne zusätzlich anfallende Kosten erfolgen. Weitere Anforderungen an die Helpdesk-Software waren:

Das Helpdesk-System

- soll über die Standardprotokolle SAML oder OAuth 2.0 an eine föderierte Authentifizierungs- und Autorisierungsinfrastruktur angebunden werden können,

² GGUS ist eine über Jahre im Grid-Umfeld eingesetzte Technologie und steht für Global Grid User Support (s. SCC-News 2/2014)

- es soll als mandantenfähiges System mit modernen Benutzeroberflächen für mehrere (mindestens zehn) Communities ausgelegt sein,
- eine Knowledge-Base-Funktion (Wissensdatenbank) für jede Community wird benötigt
- und es soll eine Open-Source-Lösung zum Einsatz kommen.

Bereitstellung des neuen Helpdesk-Systems

Das SCC hat mit mehreren Anbietern eine umfangreiche Marktanalyse durchgeführt, jeweils eine Testinstanz erworben und diese 10 bis 12 Wochen lang im SCC eingesetzt. Jede Helpdesk-Instanz wurde aus der Perspektive eines Nutzers, Unterstützers, Administrators oder Entwicklers analysiert und getestet. Untersucht wurde das Verhalten des Tools für die unterschiedlichen Zielgruppen, die Anpassungsmöglichkeiten des User Interfaces und der Dashboard-Funktionen sowie die Implementierung kundenspezifischer Workflows, sofern dies mit der jeweiligen Lösung möglich war. Die aus den verschiedenen Bewertungen kombinierten Daten sprachen für die Zammad GmbH als besten Kandidaten unter den Open-Source-Produkten.

Bisher nicht vorhandene, aber von der EOSC Community geforderte Funktionen hat die Firma Zammad dankenswerterweise in ihre Software integriert. Die Arbeiten am SCC konzentrierten sich auf

die Bereitstellung des Produkts und auf die Anpassung an die Bedürfnisse der Gemeinschaften. Bereits 2021 sind Test-, Produktions- und Entwicklungsinstanzen des Helpdesk-Systems erfolgreich am SCC installiert und in Betrieb genommen worden.

Optionen der Helpdesk-Anwendung

Das Incident und Service Request Management der EOSC bietet einen webbasierten Helpdesk-Dienst zur Vorgangsverwaltung. Daneben gibt es drei Anwendungsszenarien (Abbildung 2), die nachfolgend erläutert werden.

Helpdesk-as-a-Service:

In diesem Szenario kann das Helpdesk als Ticketsystem für EOSC-Provider und deren eingebundene Dienste angeboten werden. Jeder EOSC-Provider, der sich für diese Option entschieden hat, erhält einen eigenen Bereich für die Verwaltung der Support-Einheiten im Helpdesk-System. Dieser kann an die Anforderungen des Providers und seiner Nutzenden durch die Aktivierung von kundenspezifischen Funktionen angepasst werden.

Ticket-Umleitung:

Hier verwenden Nutzende des Providers das Webinterface des EOSC-Helpdesk als Kontaktstelle für ihre Anfragen. Alle Anfragen werden automatisch als sogenannte Tickets an eine bestimmte Organisation umgeleitet, die diese bis zur Lösung verwaltet.

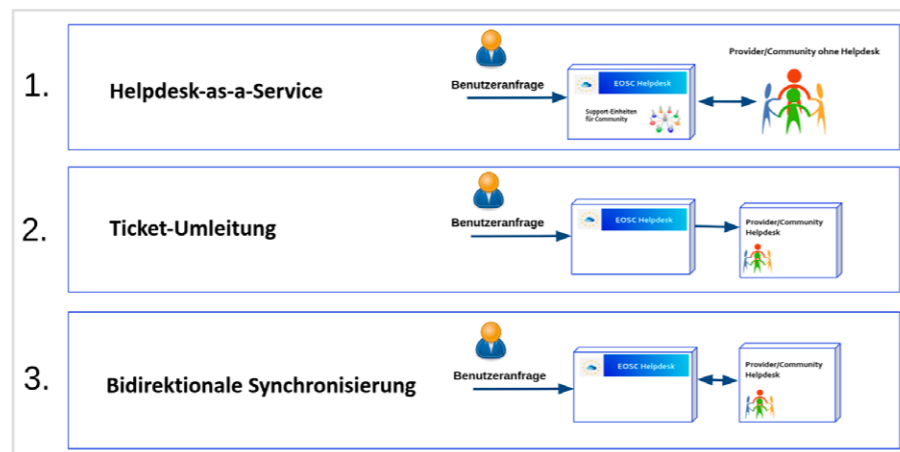


Abbildung 2: Optionen für verschiedene EOSC-Helpdesk-Anwendungsfälle



Abbildung 3: Helpdesk-Portale für unterschiedliche Communities in der EOSC

Bidirektionale Synchronisierung:

Dies entspricht der Integration mit dem Helpdesk einer Community, was eine vollständige bidirektionale Synchronisierung zwischen dem EOSC-Helpdesk und dem Helpdesk der Community voraussetzt. Diese Integration kann durch die Anwendung einer Reihe von Helpdesk-REST-APIs erreicht werden. Anbieter und Communities mit ausgereiften Helpdesk-Systemen, die eine enge Verbindung mit dem EOSC-Helpdesk wünschen, können so Tickets nahtlos in beiden Systemen verwalten.

Bereitstellung und Anwendungsszenarien

Das EOSC-Helpdesk hat seit seinem Produktionsstart im März 2022 einen bemerkenswerten Erfolg erzielt. Die Weiterentwicklung umfasste umfangreiche Anpassungen und zusätzliche Implementierungen, die auf die spezifischen Bedürfnisse verschiedener Gemeinschaften zugeschnitten wurden. Diese Verbesserungen wurden direkt als Reaktion auf Community-Anfragen vorgenommen, um sicherzustellen, dass die Funktionen des Helpdesks mit den Benutzeranforderungen übereinstimmen.

Das Helpdesk wird derzeit als Dienstleistung für mehrere EOSC-Communities angeboten. Im Rahmen des Szenarios Helpdesk-as-a-Service wird für jede Community ein eigenes Portal mit maßgeschneiderten Funktionen eingerichtet, die individuell konfiguriert sind. Abbildung 3

zeigt Beispiele von Portal-Anmeldeseiten für verschiedene Communities mit unterschiedlichen Logos und URLs.

place, verschiedenen Community-Portalen, mit SLACK und mit dem Messenger-Dienst Telegram integriert. Diese Interoperabilität

1. Helpdesk-as-a-Service	2. Umleitung	3. Synchronisierung
Production	Production	Production
Production	Production	Production
Production	Production	Beta
Production		Beta
Production		
Beta		
Beta		

Abbildung 4: EOSC-Communities, die mit dem EOSC-Helpdesk integriert sind oder das Anwendungsszenario Helpdesk-as-a-Service nutzen

Abbildung 4 zeigt die Communities, die zurzeit das EOSC-Helpdesk nutzen oder ihre eigenen Helpdesk-Systeme integriert haben.

verbessert die Anwenderfreundlichkeit und optimiert die Kommunikationskanäle, was positiv zur Effizienz der Support-Dienste beiträgt.

Die ursprünglich im Projekt EOSC Future festgelegte Roadmap für das EOSC-Helpdesk wurde vollständig umgesetzt und wie oben beschrieben mit vielen individuellen, von den EOSC-Communities angeforderten Anpassungen ergänzt.

Als Beleg für seine weitreichende Akzeptanz verzeichnet das System mehr als 700 registrierte Nutzende und wird von etwa 80 Support-Einheiten mit 70 aktiven Support-Agenten unterstützt. Es wurde nicht nur mit anderen Helpdesk-Systemen, sondern auch mit anderen EOSC Core Services wie dem EOSC-Portal, dem EOSC-Market-

New helpdesk system at EOSC – from concept to implementation
The Helpdesk is an important component of the European Open Science Cloud (EOSC). It provides a centralized support service for users across multiple communities and e-infrastructures. In the scope of the EOSC Future Project, the successor of the EOSC-Hub project, KIT proposed a new concept and implementation of the EOSC Helpdesk including its integration with other helpdesk systems and provisions the service for multiple EOSC communities.

Performante Datenmigration zwischen Bandsystemen

Die Nachfrage und Nutzung von Speichersystemen (Festplatte und Band) steigt von Jahr zu Jahr, was zu Herausforderungen bei der Unterstützung und Wartung für Datenzentren des Worldwide LHC Computing Grid (WLCG) führt. Um diese Herausforderungen zu bewältigen, werden flexiblere und leicht anpassbare Speicherlösungen erforderlich. Bandspeichersysteme sind hier nach wie vor die billigsten, zuverlässigsten und sichersten Systeme für die langfristige Speicherung großer Datenmengen. Das Ziel der verschiedenen Datenzentren ist dasselbe – ihren Nutzenden einen schnellen Zugriff sowie skalierbare und zuverlässige Speichersysteme zu bieten. Die Bereitstellung eines leistungsstarken und zuverlässigen Bandspeichersystems hat daher für das Grid Computing Centre Karlsruhe (GridKa) oberste Priorität.

Haykuhi Musheghyan, Doris Ressmann, Artur Gottmann

Seit 20 Jahren verwendet das Betriebsteam von GridKa die Software IBM Storage Protect (IBM SP), früher bekannt als Tivoli Storage Manager (TSM), als Bandspeichersystem. Um den wachsenden Anforderungen aus dem WLCG¹ gerecht zu werden, wurde das Bandspeichersystem von GridKa kürzlich von IBM SP auf High Performance Storage System (HPSS) migriert². Dies sind zwei völlig unterschiedliche Bandspeichersysteme (mit unterschiedlichem Hardware-Layout und Design). Die Lösung IBM SP verwendet eine Oracle-SL8500-Bibliothek mit Bandlaufwerken vom Typ StorageTek T1000D. Die unkomprimierte Kapazität beträgt bis zu 8,5 TB, und der unkomprimierte Durchsatz beträgt 252 MB/s. Die HPSS-Umgebung von GridKa speichert in einer Spectra-Logic-Tfinity-Bibliothek mit Bandlaufwerken vom Typ TS1160. Die unkomprimierte Kapazität dieser Laufwerke beträgt bis zu 20 TB, und der native Durchsatz beträgt 400 MB/s.

Beide Bandspeichersysteme sind unterschiedlich konfiguriert und ermöglichen die Gruppierung oder Aggregation von Daten, bevor die Daten auf Bandkassetten gesichert werden. Im Falle von IBM SP werden die Dateien direkt auf die Bandkassetten geschrieben. Im Falle von HPSS werden die Dateien zuerst auf die HPSS-Cache und erst dann auf die

Bandkassetten gespeichert. Im Setup für GridKa wird ein einzelner Datensatz³ mit nur einem Bandlaufwerk geschrieben, unabhängig von der Größe des Datensatzes. Diese Vorgehensweise verhindert, dass Datensätze auf zu viele Bandkassetten gleichzeitig verteilt werden. Durch das Schreiben eines kompletten Datensatzes auf eine einzige Bandkassette kann der Datensatz zum Lesen schneller wieder bereitgestellt werden, da die Bandkassette nur einmal eingehängt und das zeitaufwändige Spulen sowie ein erneutes Einlegen der nächsten Kassette vermieden wird. Dennoch gibt es einen Ausnahmefall für sehr große Datensätze des Teilchenphysikexperimentes ATLAS am CERN, denn diese passen nicht auf eine einzige Bandkassette. Solch ein einzelner Datensatz, der größer als 40 TB ist, wird mit mehreren Bandlaufwerken gleichzeitig geschrieben.

Ein weiterer Vorteil der neuen Speicherlösung liegt in der Möglichkeit, Aggregate zu erstellen. Dies ist eine Bündelung von verschiedenen Dateien für einen schnelleren Zugriff. Die Erstellung von Aggregaten durch das HPSS erfolgt unter Berücksichtigung der vordefinierten Einstellungen, zum Beispiel basierend auf der Verzeichnisstruktur der in HPSS zu speichernden Daten. Jedes Aggregat hat eine maximale Größe von 300 GB.

Dateien, die kleiner als 10 GB sind, werden in Aggregate aufgenommen, Dateien, die größer als 10 GB sind, werden einzeln auf Bandkassetten geschrieben. Das Schreiben von Daten aus dem HPSS-Cache auf Bandkassetten erfolgt bei GridKa alle zwölf Stunden und ist über die HPSS-Einstellungen leicht konfigurierbar.

Datenmigration von IBM SP nach HPSS

Um von einer Technologie zur anderen wechseln zu können, müssen die bereits ins alte System geschriebenen Daten migriert werden. Hierfür hat das SCC einen Ansatz entwickelt, die Daten zu restrukturieren und die Konsistenz aller Daten auf dem Band zu überprüfen. Seit Mitte 2021 werden bei GridKa mit diesem Ansatz Daten von IBM SP zu HPSS migriert. Um diesen aufwändigen Prozess zu vereinfachen, hat das GridKa-Team ein Werkzeug entwickelt, mit dem sich der Migrationsvorgang größtenteils automatisieren lässt.

Das Werkzeug migriert einen kompletten Datensatz von IBM SP auf ein unabhängiges Online-Speichersystem (hier IBM SS⁴). Danach werden die Daten in den HPSS-Cache kopiert, wobei die dafür erforderlichen Attribute angegeben werden und die Datenkonsistenz anhand der

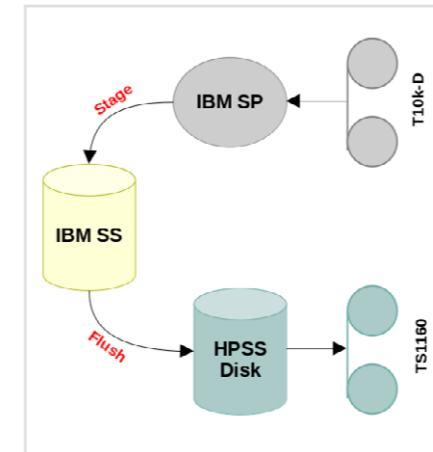


Abbildung 1: Datenmigrationsprozess bei GridKa

Prüfsumme sichergestellt wird. Wenn eine Prüfsummenabweichung auftritt, wird die Datei automatisiert aus dem HPSS-Cache entfernt, und es wird erneut versucht die Datei zu schreiben, bis der Datentransfer erfolgreich war. Erschwerend kommt hinzu, dass jedes Experiment, für das bei GridKa Daten gespeichert werden, seine eigenen, besonderen Datenstrukturen und Namensgebungen für seine Dateien hat. Diese Besonderheiten der Experimente wurden ebenfalls in dem Werkzeug berücksichtigt.

Abbildung 1 zeigt den Ablauf der Datenmigration bei GridKa. Die Datenmigration ist unabhängig von den laufenden Produktionsaktivitäten und erfolgt im Hintergrund.

Daten mehrerer Experimente wurden bereits vollständig in das neue Bandsystem migriert, und diese Experimente verwenden bereits seit einiger Zeit die aktuelle HPSS-Umgebung. Für andere Experimente ist die Migration noch nicht abgeschlossen und läuft im Hintergrund weiter.

Insgesamt hat GridKa zur Zeit 42 Millionen Dateien und 85 PB im HPSS System.

Ergebnisse

Die Abbildung 2 (oben) zeigt beispielsweise den maximalen Durchsatz von ≈ 7 GB/s beim Sichern von Dateien von einer IBM SS auf den HPSS-Cache während des Datenmigrationsverfahrens für ATLAS. Abbildung 2 (Mitte/unten) zeigt den maximalen Durchsatz von ≈ 1800 MB/s für

fünf Bandlaufwerke oder ≈ 350 MB/s pro Bandlaufwerk beim Sichern von Dateien vom HPSS-Cache auf die Bandkassetten.

Dies ist ein sehr gutes Ergebnis im Vergleich zu der früheren Lösung mit IBM SP. In der früheren Konfiguration lag der maximale Migrationsdurchsatz bei ≈ 200 MB/s pro Bandlaufwerk. Ähnlich gute Ergebnisse wurden auch bei der Migration von Daten für andere Experimente erzielt.

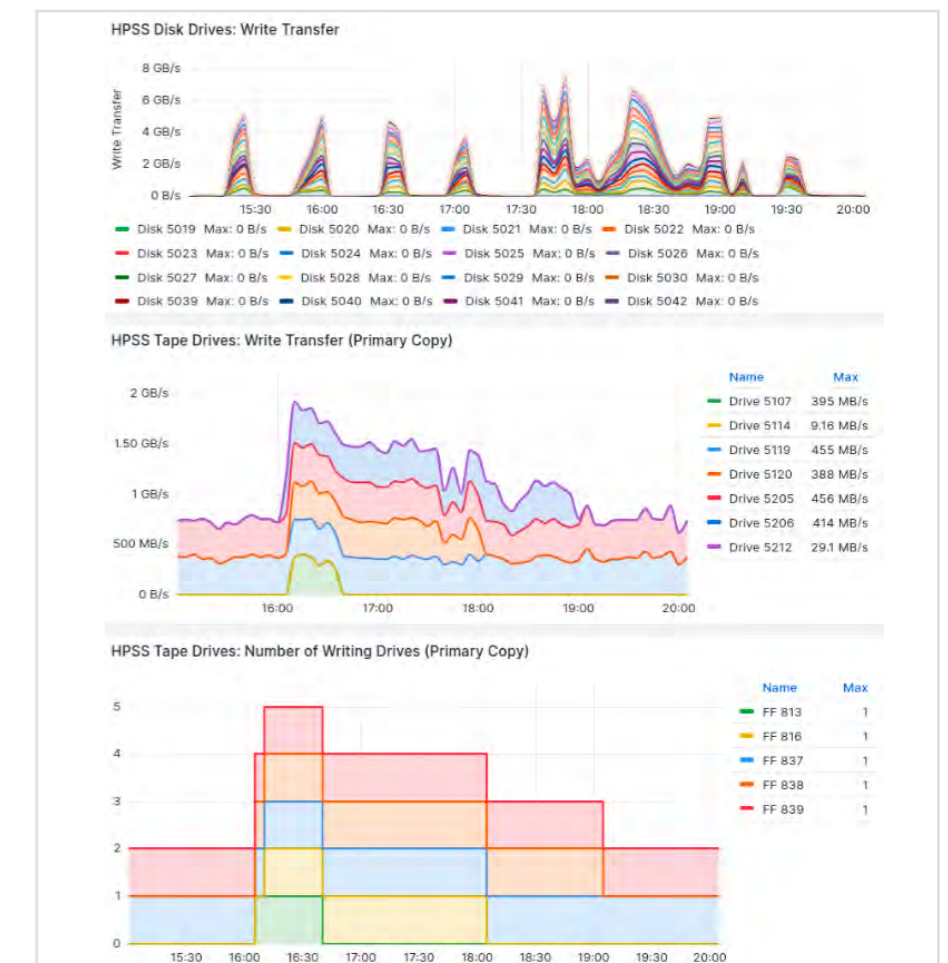


Abbildung 2: Migrationsprozedur: Datentransfer von Platte in den Cache und von Cache auf Band

Performant data migration between tape systems

The demand and usage of storage systems (disk and tape) is increasing year by year, leading to challenges in support and maintenance for Worldwide LHC Computing Grid (WLCG) data centers. To overcome these challenges, more flexible and easily customizable storage solutions are required. Tape storage systems are still the cheapest, most reliable and most secure systems for long-term storage of large amounts of data. The goal of the various data centers is the same – to provide their users with fast access and scalable and reliable storage systems. The provision of a high-performance and reliable tape storage system is therefore a top priority for the Grid Computing Center Karlsruhe (GridKa).

¹ Shiers, Jamie, The worldwide LHC computing grid (worldwide LCG), Computer physics communications 177.1-2 (2007): 219-223

² [hps-collaboration.org](https://collaboration.org)

³ Ein Datensatz ist eine Ansammlung von Dateien, die in dem selben Verzeichnis gespeichert sind und zum gleichen Datentyp gehören.

⁴ www.ibm.com/products/spectrum-scale

Erneuerung des OpenVPN-Dienstes am KIT

Mitte 2022 begannen die Planungen für die Erneuerung des OpenVPN-Dienstes am KIT, der inzwischen schon seit über 10 Jahren angeboten wird. Der Dienst hat durch die Zunahme des mobilen Arbeitens in den letzten Jahren noch einmal stark an Bedeutung gewonnen. Diesmal war auch ein Austausch der Client-Konfigurationsdateien bis Ende November 2023 unbedingt nötig. Bis die neuen Konfigurationsdateien veröffentlicht werden konnten, dauerte es insgesamt ein Jahr, was deutlich länger war als geplant. Dies war verursacht durch Änderungen in der neuen OpenVPN-Version, wobei dadurch entstandene Probleme erst analysiert und gelöst werden mussten, sowie durch das Einbringen weiterer Verbesserungen in den Dienst. Seit Juli 2023 steht nun den Nutzenden ein rundum erneuerter und optimierter VPN-Dienst zur Verfügung.

Klara Mall, Benedikt Neuffer

Planung der Erneuerung des Dienstes

Seit über 10 Jahren stellt das SCC den KIT-Angehörigen einen OpenVPN-Dienst als Lösung für den Remote-Zugriff bereit¹. Für die Nutzung des Dienstes müssen die Nutzenden auf ihren Systemen den OpenVPN-Client installieren und die OpenVPN-Konfigurationsdatei in den Client importieren (Abbildung 1). Da das SCC seinen Nutzenden die bestmögliche Komfortabilität bieten möchte, versucht es als Betreiber des Dienstes zu vermeiden, dass alle Nutzenden ihre Client-Konfigurationsdateien austauschen müssen. Letztes Jahr musste bei den Planungen für die Erneuerung des OpenVPN-Dienstes, was zunächst einmal die Erneuerung der Server-Seite bedeutet, aber auch die zwingende Erneuerung der Client-Konfigurationsdateien eingeplant werden.

Der Grund war die Abkündigung der Zertifikatsinfrastruktur des DFN (DFN-PKI) und dass somit für die OpenVPN-Server ab Dezember 2023 ein neues Zertifikat unter einer neuen Zertifizierungsstelle (Root-CA) nötig wurde. Die Root-CA wird in der Client-Konfigurationsdatei angegeben, und damit wurde der Wechsel der Konfigurationsdatei bis Dezember 2023 zwingend erforderlich. Das SCC hat sich dafür entschieden, hier eine sogenannte Community CA des DFN zu verwenden. Da eine solche CA nicht in den Betriebssystemen verankert ist, ist es für die Nutzenden am einfachsten in der Handhabung,

wenn diese direkt in die Client-Konfigurationsdateien für VPN eingebettet wird. Im Juli 2022 wurden die IT-Beauftragten des KIT über die Erneuerung des OpenVPN-Dienstes informiert und die Gründe für die damit einhergehende Erneuerung der Client-Konfigurationsdateien erläutert. Geplant war es, die neuen Konfigurationsdateien im vierten Quartal 2022 zur Verfügung zu stellen. Am Anfang standen Überlegungen, welche Optionen die neuen Client-Konfigurationsdateien enthalten sollten. Da das SCC mit der Community CA des DFN über eine CA mit 20 Jahren Gültigkeit verfügt, war es Ziel, auch den Rest der Konfigurationsdateien möglichst so zu implementieren, dass sie bestenfalls erst wieder in 20 Jahren getauscht werden müssen – falls bis dahin nicht sowieso ein anderer VPN-Dienst angeboten werden würde. Hilfreich war hierbei, dass ab der Version OpenVPN 2.4 die Datenchiffre ausgehandelt werden kann, die zur Verschlüs-

selung der Datenpakete verwendet wird. Die neuen Client-Konfigurationsdateien enthalten nun gerade so viele Optionen, wie nötig und insbesondere auch aus Sicherheitsgründen notwendig sind. Außerdem war es möglich, die neuen Konfigurationsdateien so zu realisieren, dass sie für alle Betriebssysteme gültig sind.

Probleme mit der neuen Version 2.6

Im Dezember 2022 waren die neuen OpenVPN-Server soweit installiert und befanden sich in der internen Testphase. Allerdings war das neue OpenVPN-Release 2.6 noch nicht verfügbar. Der Rollout der neuen Konfigurationsdateien sollte allerdings erst mit dem neuen Release erfolgen. OpenVPN 2.6 wurde dann Ende Januar 2023 veröffentlicht und die Server darauf aktualisiert. Im Februar konnte die erneute interne Testphase starten, bei der zunächst keine Probleme auftraten. Als

nächstes wurden die neuen Konfigurationsdateien über die interne Softwareverteilung auf den verwalteten Windows-Systemen am SCC ausgerollt. Nun meldeten sich allerdings mehrere Beschäftigte des SCC, weil ihre VPN-Verbindung nicht mehr funktionierte. Nach einer Analyse wurde schnell klar, dass es sich um ein Problem an bestimmten Internet-Anschlüssen handelte, da sich mit der neuen OpenVPN-Version ein Verhalten geändert hatte: Pakete waren plötzlich größer als sie es früher waren. Hierdurch gab es auf einmal bei viel mehr Anschlüssen als vorher Probleme mit der Maximum Transmission Unit (MTU). Für die wenigen Anschlüsse mit extrem kleiner MTU, bei denen das Problem bereits mit den vorherigen OpenVPN-Versionen aufgetreten war, wurde seit jeher eine Spezial-Konfiguration zur Verfügung gestellt, die aber bei einer nun viel größeren Anzahl von Nutzenden unbedingt vermieden werden sollte.

Um das Problem mit der MTU und die schlussendlich gefundene Lösung zu verstehen, hier ein kurzer Ausflug in die Welt der Paketvermittlung: Der VPN-Dienst macht nichts anderes als zwischen einem Computer und dem KIT-Netzwerk eine virtuelle Leitung herzustellen. In Computernetzwerken werden zu übertragende Daten in einzelne Pakete aufgeteilt, übertragen und auf der Gegenseite wieder zusammengesetzt. Beim VPN-Dienst werden Pakete verschlüsselt und in andere Pakete eingepackt. Das neue Paket ist hierdurch etwas größer als das ursprüngliche Paket. Auf dem Übertragungsweg kommt es nun vor, dass auf einem Teilpfad die maximale Paketgröße überschritten wird. Ist ein Paket zu groß, gibt es Mechanismen, den Absender hierüber zu informieren. Diese Mechanismen funktionieren allerdings leider nicht zuverlässig – Software-Bugs

und fehlerhafte Einstellungen sind hier übliche Gründe. Eine gängige Umgehungslösung ist daher, dass man von Anfang an etwas Puffer lässt und kleinere Pakete als theoretisch möglich versendet. In vorherigen OpenVPN-Versionen wurde die maximale Paketgröße unterschätzt; die neue Version berechnet diese maximale Paketgröße genauer und verschickt daher größere Pakete. Dies hat dazu geführt, dass für manche Internet-Anschlüsse die Pakete auf einmal zu groß waren. Als Umgehungslösung wurde daher die maximale Paketgröße in der Konfiguration auf den OpenVPN-Servern beschränkt. Für Nutzende an Internet-Anschlüssen, die größere Pakete zulassen, ist es dadurch zwar etwas weniger effizient, jedoch kann so die volle Funktionalität für alle Nutzenden sichergestellt werden.

Weitere Verbesserungen

Nachdem dieses Problem gelöst war, gab es noch eine weitere Idee zur Verbesserung des Dienstes. OpenVPN bietet seit längerem die Möglichkeit an, mehrere VPN-Zugänge in einer Client-Konfigurationsdatei zu listen, für die man normalerweise mehrere Konfigurationsdateien benötigen würde. In einer Konfigurationsdatei mit mehreren Zugängen werden alle Zugänge probiert und der erste, bei dem eine Verbindung erfolgreich hergestellt werden kann, genutzt. Dies ermöglicht das Zurückfallen auf schlechtere Verbindungen, wenn die beste Verbindung nicht funktioniert. Dies ist beispielsweise im Hotel vorteilhaft, wo der WLAN-Zugang für Gäste oft viele Beschränkungen aufweist. Man benötigt damit für solche Zugänge keine spezielle Konfigurationsdatei mehr. Außerdem kann auf diese Art und Weise als erstes der IPv6-Zugang angegeben werden, so dass immer, wenn vorhanden, IPv6

genutzt wird. Das Zurückfallen auf IPv4, wenn IPv6 nicht vorhanden ist, geschieht nahezu unbemerkt, da der Timeout sehr kurz gewählt wurde. Mit den neuen Konfigurationsdateien werden auf den Clients neben der aktuellen OpenVPN-Version 2.6 noch die jeweils neuesten Versionen von 2.4 und 2.5 unterstützt. Einschränkend ist zu sagen, dass die Software Viscosity für macOS und Network Manager für Linux mehrere Zugänge in einer Konfigurationsdatei noch nicht unterstützen.

Go-Live neue VPN-Zugänge

Nach einer erneuten internen Testphase im Mai wurden zunächst die OpenVPN-Server mit dem für den 10. Juni 2023 angekündigten neuen Debian Release 12 „bookworm“ aktualisiert sowie der Dienst VPN2VLAN auf die neue Betriebssystemumgebung angepasst. Erst danach konnte ein Rollout auf den verwalteten Systemen am SCC geplant werden. Nachdem die neue Installation mit den bereits vorhandenen Nutzenden eine Zeit lang ohne Probleme lief, veröffentlichte das SCC am 13. Juli 2023 die neuen Konfigurationsdateien auf den Webseiten und versendete eine Ankündigung per E-Mail. Alle Nutzenden eines vollverwalteten Endgerätes am KIT bekamen automatisch die neuen Konfigurationsdateien zeitgleich auf ihren Windows-Clients zur Verfügung gestellt. Die IT-Beauftragten wurden darin geschult, die Konfigurationsdateien für die Windows-Clients in ihrer Organisationseinheit automatisch per Group Policy Object (GPO) auszurollen. Ende November 2023 wurden die alten OpenVPN-Server abgeschaltet, das alte Zertifikat lief im Dezember aus. Bis dahin mussten alle Nutzenden des VPN-Dienstes auf die neuen Konfigurationsdateien migriert sein.

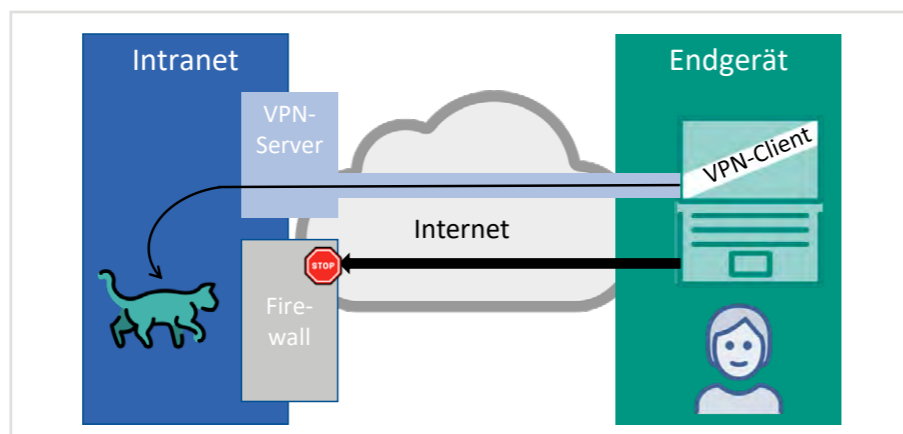


Abbildung 1: Ohne VPN-Verbindung blockiert die KIT-Firewall den Zugriff auf das Intranet (unten), während der Zugriff mit VPN möglich ist (oben)

¹ www.scc.kit.edu/dienste/vpn.php

GitLab – zentraler Dienst für kollaborative Softwareprojekte

Seit Anfang Juli 2023 stellt das SCC den zentralen GitLab-Service auf Basis der gleichnamigen Software bereit. Das neu aufgebaute System ermöglicht einen unterbrechungsfreien Betrieb für die Kollaboration in Softwareprojekten innerhalb des KIT und mit externen Mitarbeitenden. Zudem ermöglicht die komplett neue Plattform weitere Entwicklungen und den Ausbau des Dienstes.

Donghee Kang, Matthias Leander-Knoll, Klaus Scheibenberger

Die kollaborative Entwicklung von robuster und nachhaltiger Software ist heutzutage eine Kernkompetenz in der modernen Wissenschaft und Forschende sind entsprechend zu unterstützen. Hierfür hat sich das Werkzeug GitLab inzwischen fest etabliert.

Bereits seit 2014 stellt das SCC Beschäftigten des KIT GitLab in einer grundlegenden Installation mit einer Vielzahl von Funktionen bereit. Diese wurde sehr gut angenommen, wegen der zunehmenden Nutzung und der steigenden Anzahl von Projekten und Gruppen waren jedoch grundlegende Anpassungen der Plattform und der Betriebsprozesse notwendig.

Das SCC hat daher von Januar bis Juli 2023 die IT-Infrastruktur für diese zentrale Versionsverwaltung für Softwareprojekte

mit GitLab komplett neu aufgebaut. Der Dienst „GitLab am KIT“ steht seitdem allen Mitarbeitenden und Studierenden des KIT zur kollaborativen Softwareentwicklung zur Verfügung. Mit der neuen Installation können sich auch Nutzende aus anderen Einrichtungen über die Federated Login Services (FeLS) des KIT mit den gewohnten Login-Daten der Heimatinstitution anmelden und Projektgruppen beitreten.

GitLab als zentraler Dienst

In der Softwareentwicklung hat sich das Werkzeug git für die Versionsverwaltung als De-facto-Standard durchgesetzt. Darauf aufbauend bietet die Softwaresuite GitLab eine komplette Webapplikation (Abbildung 2) für Projektmanagement, Dokumentation, Bug Tracking, Issue- sowie

Build-Erstellung und Deployment inklusive Prozesse für Continuous Integration (CI). Die Einsatzszenarien am KIT sind daher unter anderem Programmierung und Projektmanagement im Kontext von Lehrveranstaltungen sowie die Softwareentwicklung in der Forschung, besonders auch das Research Software Engineering (RSE). Daneben eignet sich die Plattform zur Verwaltung textbasierter Dokumente, etwa in LaTeX oder Markdown, die Autoren häufig für Abschlussarbeiten oder Vorlesungsskripte verwenden.

Detaillierte Informationen über die Ausprägung des Dienstes und seiner Nutzungsmöglichkeiten sind in der Servicebeschreibung¹ nachzulesen.

¹ www.scc.kit.edu/dienste/gitlab.php

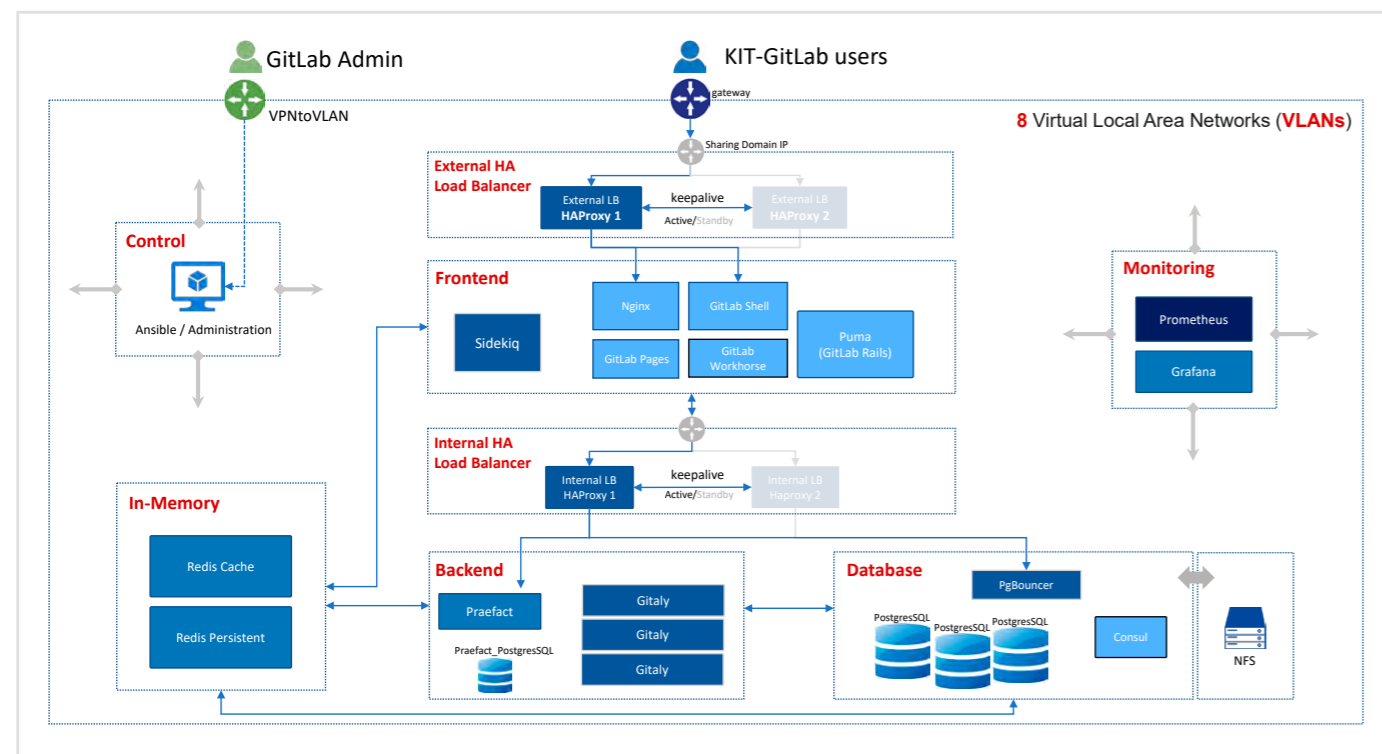


Abbildung 1: Gesamtstruktur der neu aufgebauten und bereitgestellten GitLab-Systemlandschaft

Neue Infrastruktur für den Betrieb

Seit Frühjahr 2023 hat das SCC die GitLab-Installation im Clusterbetrieb aufgebaut. Das Gesamtsystem umfasst dabei neun Kernkomponenten für den performanten und einfachen Zugriff auf die Dienste. Die Installation besteht aus insgesamt 36 virtuellen Maschinen auf Basis des freien Linux-Betriebssystems Debian (Abbildung 1).

Das GitLab-Portal

Im GitLab-Portal² sind die notwendigen Informationen zu den ersten Schritten und die zugehörige Dienst-Dokumentation für Nutzende einfach erreichbar und werden vom Betriebsteam stetig aktualisiert (Abbildung 3). Neben der umfangreich veröffentlichten Dokumentation von GitLab selbst³, erstellt und pflegt das SCC auch spezifische Informationen zur Nutzung der Software sowie den Verfahren rund um die Einbindung von externen Nutzenden und den besonderen Merkmalen des Dienstes.

Neues Anmeldeverfahren über OIDC

Im Zuge der Einführung stellte das SCC den gesamten Dienst zudem auf eine Authentifizierung mit Single-Sign-On um, so dass Nutzende sich lediglich einmal am System einloggen müssen. Die Nutzerdaten sowie IDs stellt das Identity Managementsystem des KIT (IDM) bereit.

Externe Nutzende

Durch die Anbindung an das IDM ergeben sich weitere und neue Nutzungsszenarien. So können intern und extern Kooperierende in Projekten einfach zusammenarbeiten. Projektverantwortliche legen hierfür eine Projektgruppe im föderierten Login-Service FeLS an⁴. Anschließend können sie dieser Gruppe externe Mitarbeitende hinzufügen. Durch die Zugehörigkeit zu einer Projektgruppe

² docs.gitlab.kit.edu

³ docs.gitlab.com

⁴ fels.scc.kit.edu/project/

erhalten Externe Zugriff auf GitLab am KIT und können sich einloggen. Nach dem Login lassen sich diese Nutzende in Projekte und Gruppen in GitLab ebenso hinzufügen wie interne Nutzende. Externe Nutzende erhalten die Rolle External und können damit keine eigenen Projekte und Gruppen anlegen. Zunächst sind öffentliche Gruppen und Projekte von ihnen mit den zugewiesenen Rechten nutzbar, später auch diejenigen, zu denen sie von KIT-Angehörigen hinzugefügt werden. Um Externe in GitLab einzubinden, stellt das SCC eine ausführliche Anleitung der nötigen Schritte bereit⁵.

Ausblick und weitere Entwicklungen

Die deutlich erweiterte neue Dienst-Infrastruktur sowie der Wechsel zu einem föderierten Login mit SSO ermöglicht den weiteren Ausbau. Bereits kurz nach seiner Verfügbarkeit in GitLab wurde der Large File Service (LFS) auch in der KIT-Instanz für alle Nutzende freigeschaltet. Dieser ermöglicht es, größere Datensätze und Dateien effizient zu verknüpfen und zu nutzen, ohne dass die Code-Repositoryn zu groß und schwerfällig werden. Eine erweiterte Nutzung dieses Dienstes ist durch die Anbindung per SSO und die neue globale Gruppenstruktur⁶ auch über das KIT hinaus möglich. Kooperationen lassen sich so über Organisationen hinweg einfacher auf der Plattform einrichten und pflegen, was den Dienst GitLab am KIT noch zukunftssicherer macht.

⁵ docs.gitlab.kit.edu/de/Externer_Benutzer

⁶ docs.gitlab.kit.edu/de/Organisation

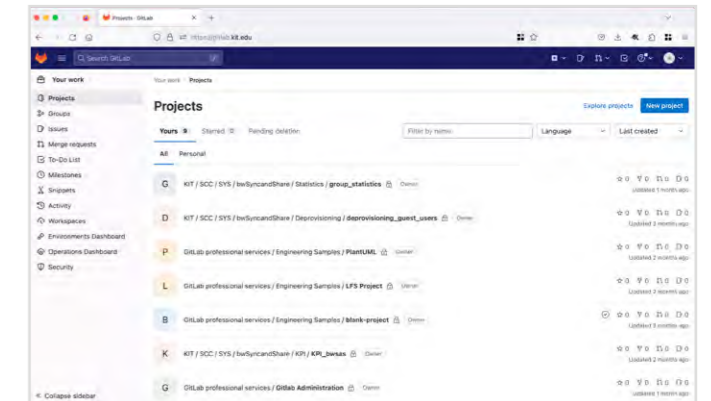


Abbildung 2: GitLab-Startseite

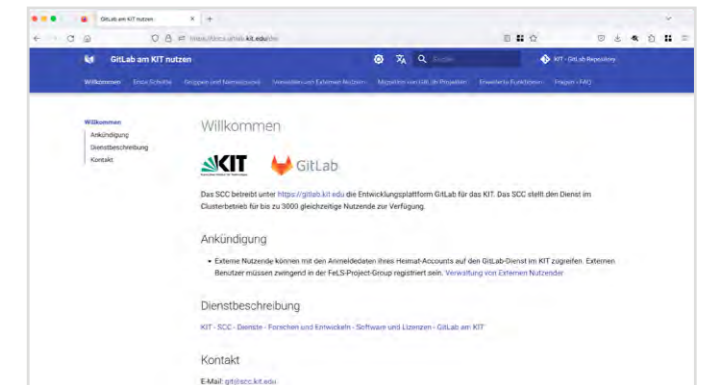


Abbildung 3: Startseite Dienstdokumentation

GitLab – central service for collaborative software projects

Since July 2023, SCC has been providing the central GitLab service based on the same-named software. The newly established system enables uninterrupted operation for collaboration in software projects within KIT and with external employees. In addition, the completely new platform enables further developments and expansion of the service.

Relaunch des KIT-Intranets

Das KIT hat seit Ende November einen umfassenden Relaunch seines Intranets durchgeführt. Das neue „Portal für alles“ erleichtert mit einem personalisierten Newsfeed die Auffindbarkeit von Informationen und steigert mit individuell konfigurierbaren Funktionen die Nutzerfreundlichkeit. Im Fokus stehen eine effizientere Integration und ein verbessertes Intranet-Erlebnis für Mitarbeitende und Studierende. Das SCC hat dabei die technische Umsetzung durchgeführt und die redaktionellen Abläufe durch ein hohes Maß an integrierten Workflows vereinfacht.

Ulrich Weiß, Larissa Beinhorn, Michael Philipp

Mit dem neuen Intranet wird die interne Kommunikation optimiert und Mitarbeitende, Studierende sowie Verwaltungsangestellte zielgerichteter und persönlicher angesprochen. Das neue Portal bündelt Neuigkeiten, wichtige Informationen sowie Zugänge zu Diensten und bietet

Nutzenden eine individuelle Konfiguration eines eigenen Dashboards: das ist der zentrale Bereich auf der Startseite, in dem sie ihre meistgenutzten Anwendungen und Links verwalten. Das Dashboard liefert u. a. Benachrichtigungen über anstehende Aufgaben aus den SAP-Systemen,

dem KIT-Genehmigungsportal und zukünftig auch von weiteren Diensten.

Zu den bedeutenden Neuerungen gehört ein personalisierter Newsfeed auf der Startseite, der Informationen nach Zielgruppen differenziert anzeigt, was die

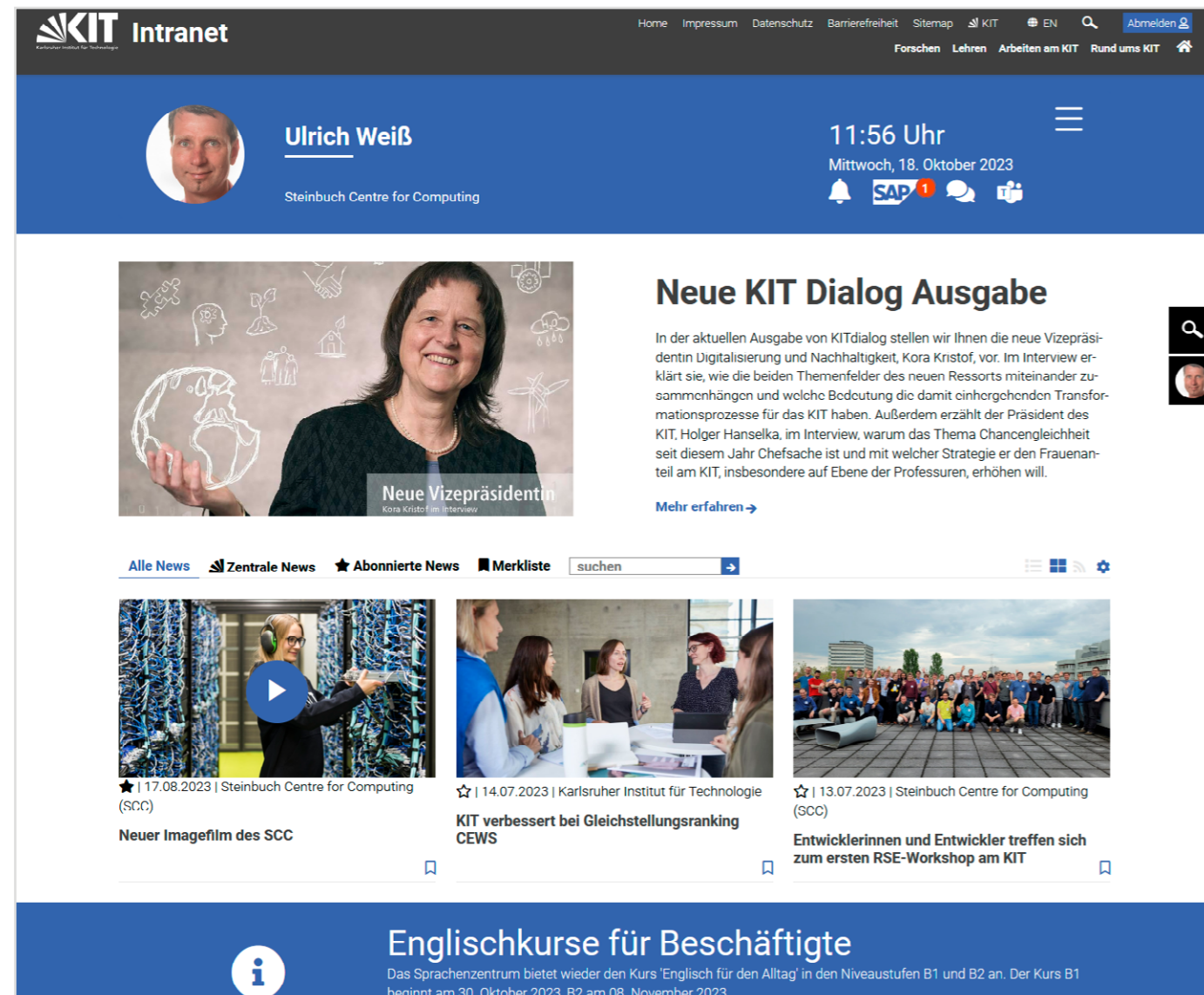


Abbildung 1: Startseite des Intranets mit Newsfeed

Auffindbarkeit relevanter Inhalte erheblich erleichtert (Abbildung 1).

Die Navigationsstruktur des Intranets ist zielgruppenspezifisch bzw. nutzendenzentriert und thematisch strukturiert, um die Auffindbarkeit von Inhalten zu verbessern, insbesondere für neue Mitarbeitende und Studierende. Selbstverständlich kann die Sprache im neuen Portal auch auf Englisch umgeschaltet werden.

Die Umsetzung des Projekts erfolgte gemeinsam mit der Gesamtkommunikation des KIT. Die enge Zusammenarbeit mit anderen KIT-Einrichtungen stellte sicher, dass vielfältige Perspektiven und Dienste in die Neugestaltung einfließen.

Der „Portal für alles“-Ansatz beinhaltet sowohl relevante Inhalte für Mitarbeitende und Studierende direkt auf den Intranet-Webseiten, als auch eine News-Funktion für auf persönliche Interessen zugeschnittene Kurzinformationen und Links zu weiteren Detailinformationen. Alle Dienstleistungseinheiten (DE) können dabei Inhalte, die sie in den eigenen Webauftritten pflegen, automatisiert ins thematisch zentrierte Intranet „einblenden“ bzw. synchronisieren lassen.

Umsetzung im Redaktionssystem

Dabei übernimmt die Intranetredaktion in Absprache mit den DEs die dezentralen Inhalte ins Intranet und macht sie über die normale Navigation zugänglich. Der Vorteil dieses Vorgehens liegt darin, dass die DEs auch weiterhin nur ihre eigenen Webauftritte im zentralen Redaktionssystem des KIT pflegen. Die zentrale Intranetredaktion kümmert sich lediglich um die korrekte Positionierung und unterstützt bei der zielgruppengerechten Aufbereitung.

Persönlicher Newsfeed

Neben der kompletten Übernahme von Inhalten ins Intranet besteht für alle Web-Redakteure die Möglichkeit, kurze Teaser im zielgruppen- und themenorientierten

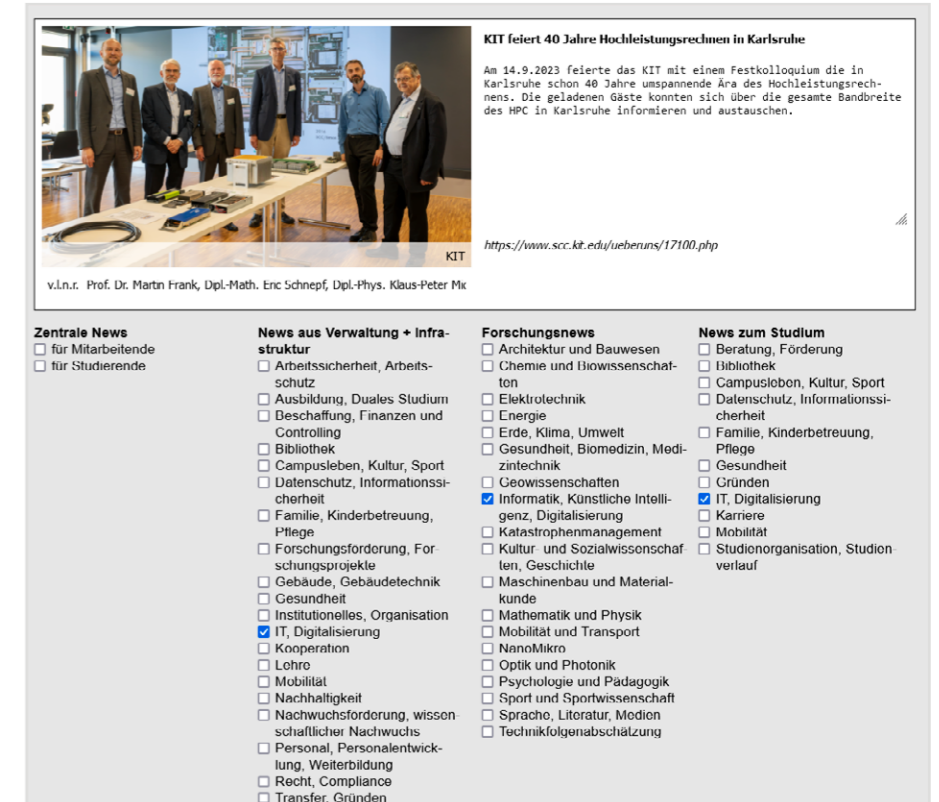


Abbildung 2: Inhaltserstellung im Redaktionssystem

Newsfeed auf der Startseite zu erstellen. Diese neue Funktion bietet für alle Redakteure im CMS die Möglichkeit, News-Teaser mit Bild und Verlinkung zur Webseite im zielgruppen- und themenorientierten Newsfeed auf der Startseite zu veröffentlichen. Dem Teaser werden dann noch die notwendigen Schlagworte für die

Zielgruppe(n) und Themengebiet(e) zugeordnet (Abbildung 2), so dass Besucher des Intranets, die diese Themen abonniert haben, die für sie relevanten Beiträge innerhalb des Newsfeeds erhalten.

Beiträge im Newsfeed des neuen Intranets sind chronologisch absteigend

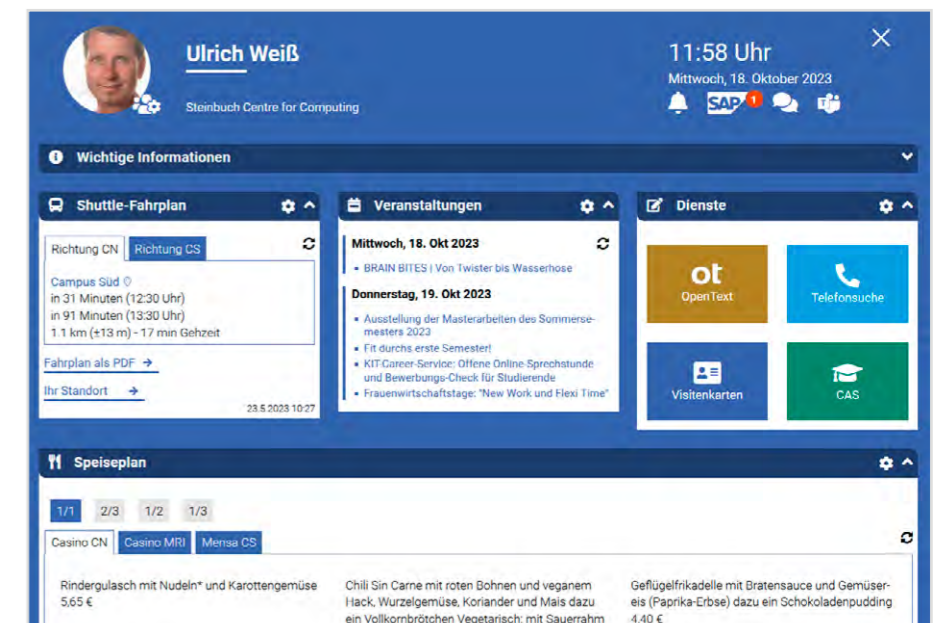


Abbildung 3: Das Dashboard – ein individuell konfigurierbarer persönlicher Bereich



Abbildung 4: Konfiguration des persönlichen Bereichs

sortiert. Die thematischen Interessen können von den Nutzenden individuell konfiguriert bzw. ausgewählt werden und führen zu einem persönlichen Newsfeed, der alle Bereiche des KIT abdeckt.

Persönliches Dashboard

Neben den statischen Seiten bzw. der normalen Navigation und dem Newsfeed lässt sich auf der Intranet-Startseite nun auch ein persönlicher Bereich, das Dashboard (Abbildung 3), individuell und personenbezogen konfigurieren (Abbildung 4). Der persönliche Bereich wird über das Burger-Menü ein- und ausgeblendet und kann mit individuellen Widgets belegt werden. Derzeit bietet das KIT die Widgets Wichtige Informationen, Shuttle-Fahrplan in Abhängigkeit vom eigenen Standort, Veranstaltungshinweise, Speisepläne, Links zu häufig genutzten Diensten sowie eine individuell zusammenstellbare Linksammlung an.

Anmeldung erforderlich

Das neue Intranet ist über alle gängigen Geräte zugänglich und unterstützt insbesondere die Nutzung von mobilen Endgeräten. Der Zugang ist allerdings erst nach Anmeldung mit dem KIT-Account möglich, damit die o. g. Angebote auf

die persönlichen Vorlieben zugeschnitten werden können.

Verbesserte Suche

Neben der persönlichen Startseite und den statischen, für Zielgruppen aufbereiteten Webseiten gibt es eine überarbeitete Suchfunktion. Diese ist speziell auf Informationen ausgelegt, die im Intranet und auf den Webseiten der Einrichtungen angeboten werden, um schnell zu den einzelnen Angeboten und Portalen zu gelangen.

Das Intranet-Modernisierungsprojekt zielt darauf ab, die Bedürfnisse der Nutzenden besser zu erfüllen und die interne Kommunikation effizienter zu gestalten. Die hier vorgestellten Änderungen verbessern nicht nur die Auffindbarkeit von Informationen, sondern steigern auch die Nutzungsfreundlichkeit und Interaktion insgesamt. Mit der hier vorgestellten benutzerfreundlichen Plattform für Information und Interaktion unterstreicht das KIT sein Engagement, mit den sich wandelnden Anforderungen an die interne Kommunikation Schritt zu halten.

The new KIT Intranet

The new intranet site is designed to foster internal communication at KIT among employees and students. It now offers a personalized user dashboard to manage applications and receive notifications about tasks and services.

A standout feature is the personalized news feed, making relevant information easily accessible. The intranet's navigation is user-centric and thematically structured for improved content access. The project was executed in collaboration with various KIT units, allowing automated content delivery to central intranet pages. Web authors can create teaser content and personalize their areas. For access to the intranet and enhanced interaction login using the KIT account is required. The relaunch reflects the commitment of KIT to invest in the evolving needs for internal communication.

Arbeitskreis Clientmanagement als Wissensmultiplikator

Im November 2022 wurde der Arbeitskreis Clientmanagement am KIT ins Leben gerufen. Es entstand eine gut besuchte Wissens- und Austauschplattform für die Themen zur Verwaltung der Arbeitsplatzendgeräte des KIT, an der sich IT-Beauftragte und -Admins sowie der technische Clientsupport am SCC beteiligen.

Andreas Lorenz

Der allergrößte Teil der Beschäftigten am KIT nutzt für die tägliche Arbeit Desktop-Computer, Laptops oder Tablets. So kommen schnell Tausende Arbeitsplatz-Endgeräte – sogenannte Clients – zusammen, die in das Netzwerk des KIT integriert sind und zentrale sowie dezentrale Services nutzen. Neben der schieren Menge ist die Vielfalt dieser Geräte und die dafür anfallenden Aufgaben, gerade im wissenschaftlich-technischen Umfeld, enorm groß, während für die Verwaltung erfolgreich umgesetzte Standardisierungsmaßnahmen einfache und zentrale IT-Beschaffungs-, Verwaltungs-, und Supportprozesse ermöglichen. Um diesen unterschiedlichen Ausprägungen und Anforderungen begegnen zu können, hat das SCC vier Betriebsmodelle

definiert. Dabei behält es den sicheren und zuverlässigen Betrieb der gesamten IT-Infrastruktur im Blick (SCC-News 2/2021, S. 9). Es liegt in der Natur der Sache, dass bei einem so komplexen, vielseitigen und individuellen IT-Umfeld ein hoher Unterstützungsaufwand anfällt. Mit der Einführung eines Konzepts zur partnerschaftlichen Zusammenarbeit hat das SCC schon vor Jahren die Beschäftigten und IT-Beauftragten/Admins sowie den IT-Service des SCC zusammengebracht, mit dem Ziel, die IT-Infrastruktur geregelter, sicherer und zukunftsfähiger zu betreiben. Damit und mit den vier Betriebsmodellen entsteht eine solide Basis für einen auf die Bedürfnisse der Verwaltung und Technik sowie Forschung und Lehre zugeschnittenen IT-Endgerätebetrieb.

Eine dritte Säule verstärkt nun diese Basis zusätzlich: Der vom SCC geleitete Arbeitskreis Clientmanagement bringt seit November 2022 IT-Beauftragte und IT-Admins des KIT sowie Personen des technischen Clientsupport des SCC zusammen. Bisher fanden sieben Online-Treffen mit zwischen 50 und 60 Teilnehmenden via MS Teams statt. Die Themen reichen aktuell von der klassischen Arbeitsplatzbetreuung über den Support für Mac-Systeme zur Softwareverteilung und den Windows 11 Support. Dieses neue, unkomplizierte Format wurde sehr gut angenommen, so dass mehr und mehr Themen aus der dezentralen IT zu den für die Arbeit des SCC wichtigen Themen hinzukommen.

FORSCHUNG UND PROJEKTE

Abgeschlossene Promotionen

Zahlreiche Menschen, die am SCC in verschiedenen Wissenschaftsdisziplinen forschen, streben eine Promotion an. Seit Erscheinen der SCC-News im Winter 2021 konnten acht Wissenschaftlerinnen und Wissenschaftler ihre Dissertation erfolgreich am KIT oder anderen Universitäten abschließen.

Pratik Nayak, Informatik, 23. November 2023, Synchronization free algorithms for exascale and beyond

Gayatri Čaklović, Mathematik, 28. Juni 2023, ParaDiag and Collocation Methods: Theory and Implementation

Steffen Schotthöfer, Mathematik, 3. Mai 2023, Synergies between Numerical Methods for Kinetic Equations and Neural Networks

Pia Stammer, Mathematik, 22. März 2023, Uncertainty quantification and numerical methods in charged particle radiation therapy

Germaine Götzelmann, Gesellschafts- & Geschichtswissenschaften, TU Darmstadt, 20. Oktober 2022, Computergestützte Analyse von Illustrationswiederverwendungen in deutschsprachigen Drucken des 16. Jahrhunderts. Am Beispiel Frankfurt am Main

Arthur Voroin, Physik, 13. Mai 2022, Application and Optimization of Contact-Guided Replica Exchange Molecular Dynamics

Fathia Idiris, Chemie und Biowissenschaften, 27. April 2022, Multiscale Molecular Dynamics Simulations of Histidine Kinase Activity

Sarah Schönbrodt, Mathematik, 23. Februar 2022, Optimierungsprobleme in der mathematischen Modellierung: Grundlegende Aspekte und Chancen aus Sicht der Mathematikdidaktik – herausgestellt an aktuellen Problemen aus der Forschung zu künstlicher Intelligenz und erneuerbaren Energien

Kein Weg zu weit – GridKa integriert nachhaltige Compute-Ressourcen aus Übersee

Der nachhaltige Betrieb von Rechenzentren ist auch eine Frage des Standorts, denn deren Energieversorgung und Kühlung bauen stark auf lokale Infrastruktur, Wetter- und Umweltverhältnisse auf. Da sich ein bestehendes Rechenzentrum aber nicht einfach verlegen lässt, benötigt es moderne Ansätze, um Rechenaufgaben flexibel auf andere Standorte übertragen zu können. Im Rahmen eines Pilotprojekts wurden dafür nachhaltig betriebene Rechenressourcen aus mehr als 8.500 km Entfernung in die Compute Farm des Grid Computing Centre Karlsruhe (GridKa) eingegliedert.

Max Fischer, Manuel Giffels

Um die Idee der Integration von nachhaltig betriebenen Computern ins eigene Rechenzentrum umzusetzen, hat sich ein Team aus Forschenden am SCC nicht nur in Europa, sondern weltweit umgesehen und geeignete Projektpartner gesucht.

Unser Partner: Lancium Compute

Lancium ist ein US-amerikanisches Unternehmen, das in Texas angesiedelt ist. Das Geschäftsfeld von Lancium umfasst verschiedene Aktivitäten im Bereich Strom-Management. Es betreibt unter anderem große Rechenzentren in Regionen mit einem Überschuss an erneuerbaren Energien, mit dem Ziel, damit das US-Stromnetz zu stabilisieren und den Ausbau von Wind- und Solarenergie zu fördern. Unter

dem Namen Lancium Compute fungierte die Firma bis vor kurzem als nachhaltiger Platform-as-a-Service-Anbieter. Lancium verwendete ausgemusterte Hardware von großen US-Technologieunternehmen wieder und betrieb diese CO₂-neutral mit erneuerbaren Energien.

Machbarkeitsstudie

Im Rahmen einer Machbarkeitsstudie bekam das SCC Ende 2022 die Möglichkeit, einen kleinen Teil der Lancium-Compute-Ressourcen zu nutzen. Da GridKa allerdings nicht selbst rechnet, sondern für die Wissenschaft Rechenressourcen zur Verfügung stellt, musste hierfür ein Teil der Lancium-Compute-Ressourcen in den GridKa-Rechencluster eingebunden werden.

Das SCC war dank vorheriger Entwicklungen in Kooperation mit Partnern aus der Hochenergiephysik hierfür gut aufgestellt: Mit der Software COBaID/TARDIS (SCC-News 1/2021, S. 5) gibt es die Technologie, um je nach Bedarf und Verfügbarkeit dynamisch Rechner in ein bestehendes Cluster einzubinden. Weitere Entwicklungsarbeiten in der Software waren notwendig, damit diese mit der Lancium-Compute-Plattform kommunizieren konnte. Anfang 2023 konnten dann erstmals Lancium-Ressourcen den Wissenschaftlerinnen und Wissenschaftlern über GridKa zur Verfügung gestellt werden (Abbildung 1).

Im Laufe der Machbarkeitsstudie wurden den Experimenten ATLAS und CMS am

CERN so insgesamt 700.000 Compute-Core-Stunden (Abbildung 2) an Rechenleistung für Simulationsrechnungen zur Verfügung gestellt. Dabei haben die Nutzenden nicht einmal gemerkt, dass ihre Rechenjobs nicht in Karlsruhe, sondern 8.500 km entfernt ausgeführt wurden.

Im Rahmen eines Skalierungstests wurden im April dieses Jahres bis zu 3.000 Lancium-Compute-Rechenkerne in das weltweite LHC Computing Grid integriert, was etwa 7 % der Rechenleistung des aktuellen GridKa-Clusters entspricht (Abbildung 3). Dies demonstriert eindrucksvoll, dass die Nutzung nachhaltiger Rechenressourcen einen signifikanten Beitrag zum wissenschaftlichen Rechnen leisten kann.

Zukunft für die Zukunft

Leider wurde dieses Platform-as-a-Service-Angebot Ende April 2023 von Lancium überraschend eingestellt, so dass eine bereits geplante Beschaffung von nachhaltigen Compute-Ressourcen nicht mehr zustande kam. Doch sowohl mit der Technologie als auch der Expertise ist das GridKa und das SCC nun bereit, auch mit anderen Betreibern nachhaltiges wissenschaftliches Rechnen umzusetzen.

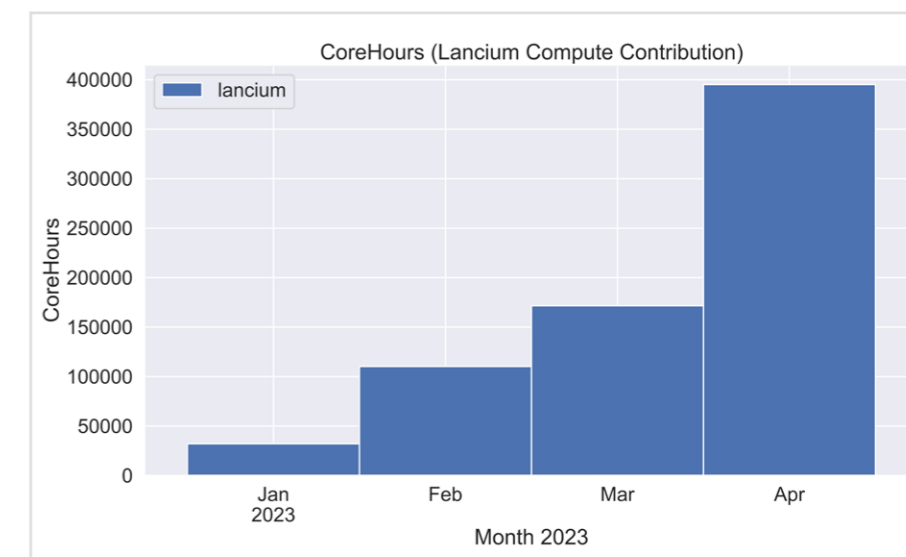


Abbildung 2: In GridKa zur Verfügung gestellte Lancium-Compute-Ressourcen (Kumulierte monatliche Core-Stunden)

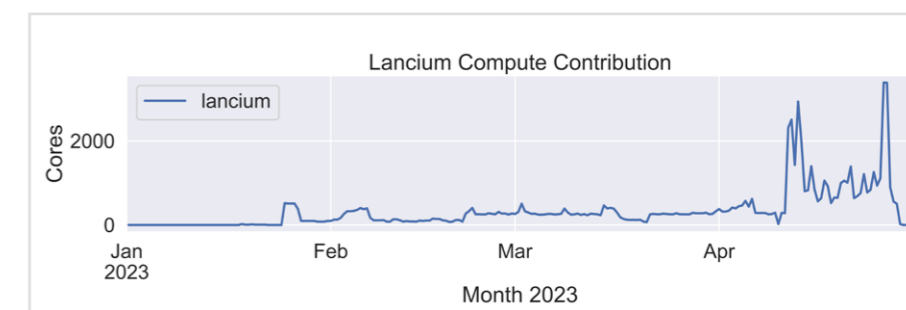


Abbildung 3: In GridKa zur Verfügung gestellte Lancium-Compute-Ressourcen (Anzahl Cores)

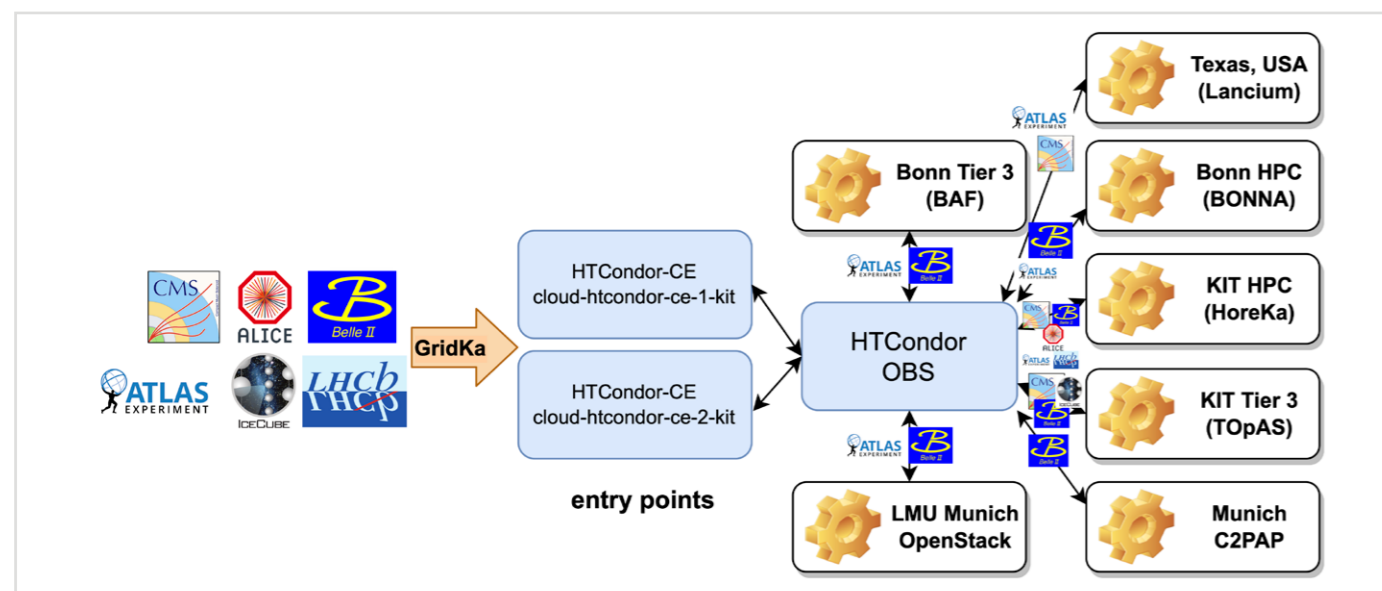


Abbildung 1: Setup des GridKa Overlay Batch Systems (OBS) und kooperierender Ressourcenbetreiber

No distance too far – GridKa integrates sustainable compute resources from overseas

The sustainable operation of data centers depends also on its location: energy supply and cooling rely heavily on local infrastructure, weather, and the environment. Since data centers cannot be easily relocated, instead, sustainably operated computing resources from more than 8500 km away were integrated into the GridKa compute farm in a pilot project together with our partner Lancium. Lancium reused decommissioned hardware from major US technology companies and operated it in a CO₂-neutral manner using renewable energies.

The integration was possible thanks to previous development of the COBaID/TARDIS software, which supports dynamic integration or remote compute resources depending on availability and demand.

In total 700,000 core hours and in a scale test up to 3000 cores were integrated into the Worldwide LHC Computing Grid, which corresponds to about 7% of the current GridKa cluster.

With both technology and expertise, GridKa showed that it is ready to integrate sustainable computing from remote. The GridKa research and development team already has new, challenging ideas to look forward to.

Dateneffizientes Maschinelles Lernen von Biomolekülen

Leben wird auf der zellulären Ebene von verschiedenen Biomolekülen bestimmt. Sie stellen die Maschinerie von lebenden Organismen dar und spielen eine entscheidende Rolle für das Funktionieren jeder Zelle. Um ihre Funktion und die damit zusammenhängende Struktur zu studieren, wird zunehmend auch maschinelles Lernen verwendet. Mitarbeitende der Forschungsgruppe Multiscale Biomolecular Simulation und des Helmholtz-AI-Teams im SCC haben in Kooperation mit dem Forschungszentrum Jülich und dem Deutschen Zentrum für Luft- und Raumfahrt (DLR) nun eine Methode vorgeschlagen, wie moderne und klassische Methoden des tiefen maschinellen Lernens kombiniert werden können, um auch in datenarmen Szenarien Modelle bauen zu können.

Oskar Taubert, Alexander Schug

Biologische Makromoleküle

Verschiedene Biomoleküle bilden als kleinste Bausteine die molekulare Maschinerie aller lebenden Organismen. Sie übernehmen entscheidende Funktionen in jeder Zelle wie die Speicherung, das Auslesen und die Interpretation von genetischen Informationen, das Zusammenspiel zahlreicher chemischer Reaktionen oder das Erkennen und Interpretieren von Signalen oder Gefahren für die Zellen. Zu diesen Biomolekülen gehören Proteine, Ribonukleinsäuren (RNA), Kohlenhydrate und Lipide, die auf komplexe Weise miteinander interagieren, um die Prozesse des Lebens zu ermöglichen.

Eines der wichtigsten Biomoleküle ist die schon erwähnte RNA. Etwas vereinfacht gesagt ist RNA eine einzelsträngige Kopie von Desoxyribonukleinsäure (DNA), die viele verschiedene Aufgaben übernimmt. Eine wichtige Aufgabe ist der Transport der genetischen Information, die

die DNA trägt. Einige Viren speichern ihre genetische Information auf RNA-Molekülen (so ist etwa das Corona-Virus ein RNA-basiertes Virus, und auch viele der Corona-Vakzine beruhen auf RNA-Technologien). Einige RNA-Moleküle können sich, wie etwa auch Proteine, zu komplexen dreidimensionalen Strukturen falten. Diese 3D-Struktur ist wiederum von entscheidender Bedeutung für das Verständnis der Funktion der RNA und auch für die Entwicklung von Behandlungen für Krankheiten, die durch RNA-Funktionsstörungen verursacht werden.

Strukturbestimmung

Die experimentelle Bestimmung von solchen 3D-Strukturen von RNA-Molekülen ist sehr aufwändig – und hier kommt das maschinelle Lernen (ML) ins Spiel. ML hat viele Bereiche revolutioniert, darunter auch die Biologie, indem es leistungsstarke Werkzeuge für die Analyse komplexer Daten und das Erstellen von Vorhersagen

bereitstellt. In den letzten Jahren hat sich Deep Learning (DL) als besonders vielversprechender Ansatz für die Analyse biologischer Daten erwiesen, einschließlich der Vorhersage von biomolekularen Strukturen. DL ist eine Form des maschinellen Lernens, bei der künstliche neuronale Netze verwendet werden, um aus Daten zu lernen. Diese Netze bestehen aus zahlreichen Schichten miteinander verbundener Knoten, die Informationen verarbeiten. Durch das Training dieser Netze mit großen Datensätzen können DL-Algorithmen lernen, Muster zu erkennen und genaue Vorhersagen zu treffen. Für Proteine wurde die Aufgabe der Strukturvorhersage vor Kurzem durch AlphaFold2, ein tiefes neuronales Netz, mit beeindruckenden Ergebnissen gelöst. AlphaFold2 nutzt die umfangreichen Datenbanken für Proteininformationen, um direkt dreidimensionale Strukturen vorherzusagen. Für RNA ist eine direkte Übertragung aber nicht möglich, da für RNA nur Bruchteile der Trainingsdaten

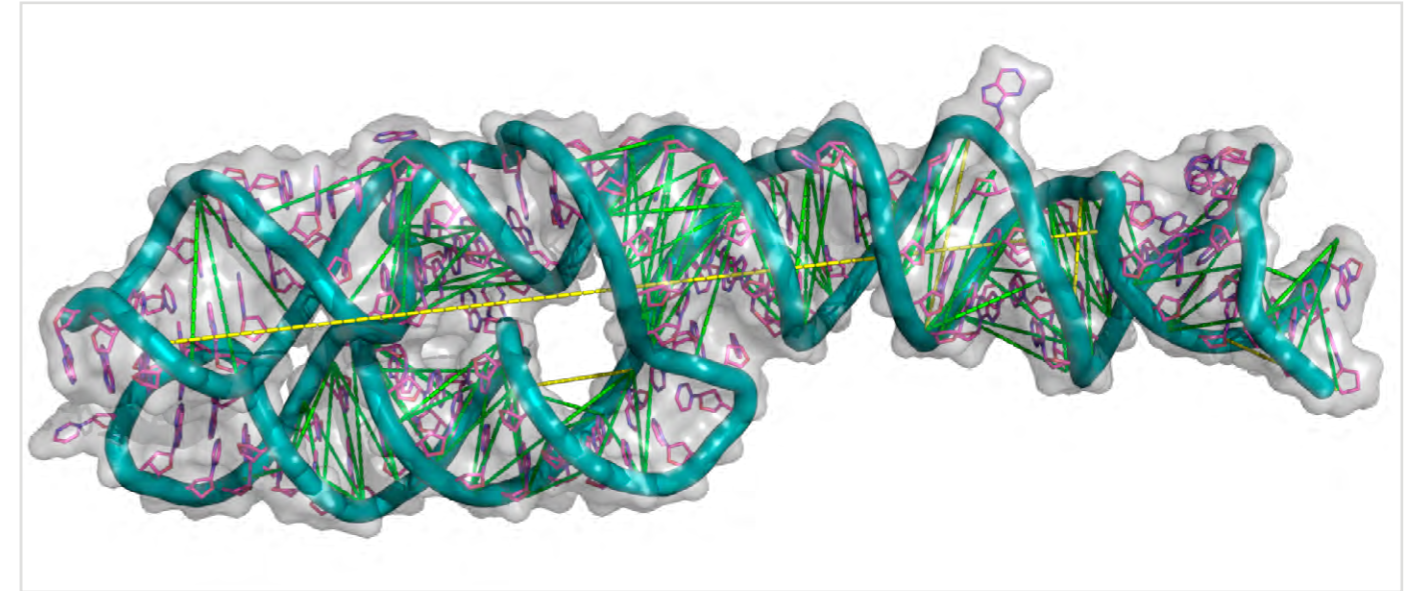


Abbildung 2: Beispielsystem mit korrekten Vorhersagen in grün und falschen Vorhersagen in gelb, Hauptstrang in türkis und Basen in violett

zur Verfügung stehen: So gibt es auch für RNA umfangreiche Sequenzdaten, also Abfolgen der einzelnen RNA-Bausteine, dargestellt durch lange Buchstabenketten. Allerdings gibt es nur etwa 200 bekannte Strukturen für RNA, die sich deutlich voneinander unterscheiden.

Das BARNACLE Modell

Ein neues Modell, BARNACLE genannt, verwendet nun einen DL-Ansatz, um räumliche Nachbarschaften zwischen RNA-Bausteinen (sog. Nukleotide, einzelne Buchstaben) vorherzusagen.

Man kann sich das wie ein Lego-Modell vorstellen: Wenn an einer Stelle einzelne Legosteine ausgetauscht werden, müssen die Steine in der Umgebung sich anpassen, damit das gesamte Bauwerk noch zusammenpasst. Das BARNACLE-Modell nutzt diese Idee für RNA: Nukleotide, die in RNA räumlich nahe beieinander liegen, mutieren während der Evolution auch mit größerer Wahrscheinlichkeit gemeinsam. Und genau nach diesen entstehenden Mutationsmustern sucht das Modell. Um das Modell trainieren zu können, wird auf eine Kombination aus selbstüberwachtem Vortraining auf vielen Sequenzdaten (Abbildung 1) und einer effizienten Nutzung der wenigen Strukturdaten

zurückgegriffen. Beim selbstüberwachten Vortraining wird das Modell auf einem großen Datensatz mit Sequenzdaten trainiert, um allgemeine Merkmale der Daten zu lernen. Dazu wird das Eingabedatum verfälscht, und das Modell muss das Original wiederherstellen. Dabei lernt das Modell die Regeln in den Daten. Dieses vortrainierte Modell wird dann auf dem kleineren Datensatz mit Strukturdaten feinabgestimmt, um spezifische Vorhersagen zu treffen.

Nachdem das Modell feinabgestimmt ist, wird ein weiteres Modell trainiert, das aus den internen Repräsentationen des Netzwerks die eigentliche Vorhersage produziert. Dieses Modell liest also in gewissem Sinne „die Gedanken“ des ersten Modells. Der Vorteil dieses Vorgehens ist, dass das zweite Modell kein tiefes neuronales Netz sein muss, sondern in diesem Fall ein sogenannter Random Forest ist, der effizienter von den wenigen Daten lernt. BARNACLE zeigte eine signifikante Verbesserung gegenüber etablierten klassischen statistischen Ansätzen, aber auch anderen neuronalen Netzen (Abbildung 2). Es zeigt sich auch, dass diese Herangehensweise auf verwandte Aufgaben mit ähnlichen Datenbeschränkungen übertragbar ist.

Die Arbeit wurde als „RNA Contact Prediction by Data Efficient Deep Learning“ in Communications Biology veröffentlicht¹.

¹ doi.org/10.1038/s42003-023-05244-9

Data Efficient Machine Learning of Biomolecules

Life on the cellular level is governed by different types of biomolecules. These building blocks represent the machinery of living organisms and play crucial roles for the functioning of each cell. To study their structure and the function tightly coupled to it, the use of machine learning is becoming increasingly established. The publication “RNA contact prediction by data efficient deep learning” in the journal Communications Biology by the junior research group Multiscale Biomolecular Simulation and Helmholtz AI Energy at SCC in cooperation with JSC and DLR presents a method to make structural predictions in data sparse scenarios combining self-supervised machine learning and classical machine learning techniques.

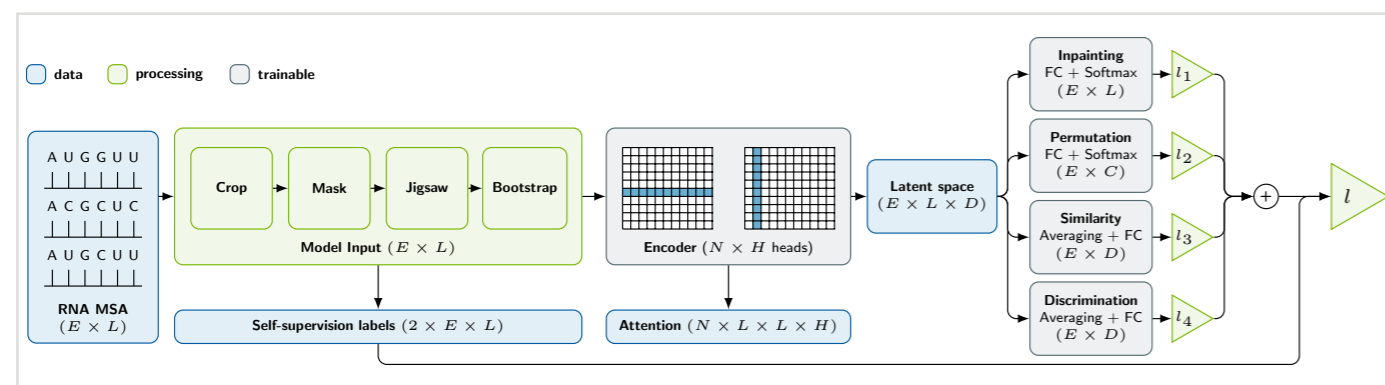


Abbildung 1: BARNACLE: selbstüberwachtes Training



Einfacher Zugang zu hydrologischen Daten in Deutschland unterstützt die Umweltforschung

Die Arbeitsgruppe um das Projekt V-FOR-WaTer entwickelt ein Webportal, in dem Umweltwissenschaftlerinnen und -wissenschaftler auf hydrologische Daten aus unterschiedlichen Quellen einfach und schnell zugreifen, aber auch eigene Daten teilen können. Zusätzlich arbeiten die Forschenden des KIT an Schnittstellen zu bestehenden Datenportalen, damit bereits veröffentlichte Datensätze mit einbezogen werden können. Mithilfe integrierter Werkzeuge sind diese Daten leicht skalierbar und in einem einheitlichen Format verfügbar. Außerdem stehen im Portal Analysewerkzeuge für Auswertungen und Unsicherheitsabschätzungen zur Verfügung. Ein leichtes Abspeichern und Teilen der Analyseschritte und der verwendeten Werkzeuge macht die Berechnungen reproduzierbar, und durch die Anbindung an vorhandene Datenrepositorien soll die Publikation eigener Daten vereinfacht werden. Damit bietet V-FOR-WaTer und die Weiterentwicklung des Portals im Rahmen des neuen DFG-geförderten Projekts ISABEL, Forschenden aus den Wasser- und Umweltwissenschaften Zugriff auf Daten von z. B. Landesämtern oder Hochschulprojekten und zu standardisierten Werkzeugen für reproduzierbare Datenanalyse-schritte. Das Entwicklungsteam am SCC, Elnaz Azmi und Marcus Strobl, erweitert beständig die Funktionen und arbeitet dazu zusammen mit dem Institut für Wasser und Gewässerentwicklung – Hydrologie (IWG) am KIT und Hydrologen in Deutschland.

Jos van Wezel im Gespräch mit Elnaz Azmi und Marcus Strobl

Liebe Elnaz, lieber Marcus, bitte stellt euch vor und erzählt uns etwas über euren Hintergrund und eure Rolle im Projekt.

Elnaz: Ich erstelle meine Doktorarbeit in Informatik und begann am SCC schon im Jahr 2016. Ich war von Beginn an im Projekt eingebunden, anfangs als Entwicklerin und wegen meiner Promotionsarbeit heute eher als Beraterin. Meine Dissertation hängt aber nur zu einem kleinen Teil mit V-FOR-WaTer¹ zusammen. Allerdings habe ich einige Anregungen für sie im Projekt gefunden. Vor allem aufgrund der engen Zusammenarbeit mit dem Institut für Wasser und Gewässerentwicklung – Hydrologie (IWG) begann meine Forschung mit der Entwicklung einer Methode zur Optimierung eines hydrologischen Modells.

Marcus: Ich bin promovierter Geowissenschaftler und habe nach der Promotion noch zwei Jahre als Postdoc gearbeitet. In dieser Zeit hatte ich das Gefühl, mit meinen Forschungsmethoden einen „Hammer“ in der Hand zu haben, für den ich

ständig nach „Nägeln“ suchen musste, was ich als unbefriedigend empfand. Mit dem Ziel, mich mehr in die IT und Softwareentwicklung zu entwickeln, arbeite ich nun für das Projekt V-FOR-WaTer. Es ist sehr reizvoll, eine Forschungsumgebung zur Sammlung und Analyse von Daten für Hydrologen aufzubauen. Eine Forschungsumgebung, wie ich sie als Geowissenschaftler selbst gerne gehabt hätte. Ich bin von Anfang an zusammen mit Elnaz im Projekt dabei.

Wie hat das Projekt angefangen?

Das Projekt begann als Landesprojekt, gefördert vom Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg, ursprünglich für zwei Jahre, und war angesetzt, um eben diese Weboberfläche zu entwickeln (s. SCC-News 2/2016). Es gab ein paar Ideen hier aus dem Hause, wie man es anpacken könnte, wobei wir schnell sahen, dass dies nicht ganz den Ansprüchen der Hydrologen entspricht, dass es doch ein bisschen mehr braucht, damit die Anwendung am Ende von der hydrologischen Gemeinschaft angenommen und genutzt wird.

Daher lag von Anfang an der Fokus auf der Entwicklung eines für die Community bedienerfreundlichen Webportals, so, wie es heute verfügbar ist (Abbildung 1). Ursprünglich kam die Idee zu dem Projekt von Erwin Zehe aus der Hydrologie, der an einem vorherigen größeren Projekt beteiligt war, wo die Daten aufbereitet und geteilt werden sollten. Nur das funktionierte nicht so befriedigend. Es gibt zwar schon ein ähnliches System, das war allerdings nur eingeschränkt nutzbar. Da hätte man, salopp gesagt, „die Wissenschaft an die Organisation der Daten und die Datenbank anpassen müssen“, weshalb dies nicht in Frage kam. Die Motivation kam also im Grunde von den Daten. Mit unseren Kollegen aus der Hydrologie, Sibylle Häbler und Mirko Mälicke, kamen wir zu dem Schluss, dass es die bessere Lösung wäre, etwas Eigenes zu entwickeln.

Mit wem arbeitet ihr im Projekt zusammen?

Wie schon gesagt, arbeiten wir zusammen mit dem IWG am Campus Süd. Anfangs zu zweit, heute sind wir auf acht Leute im Projekt angewachsen. Drei

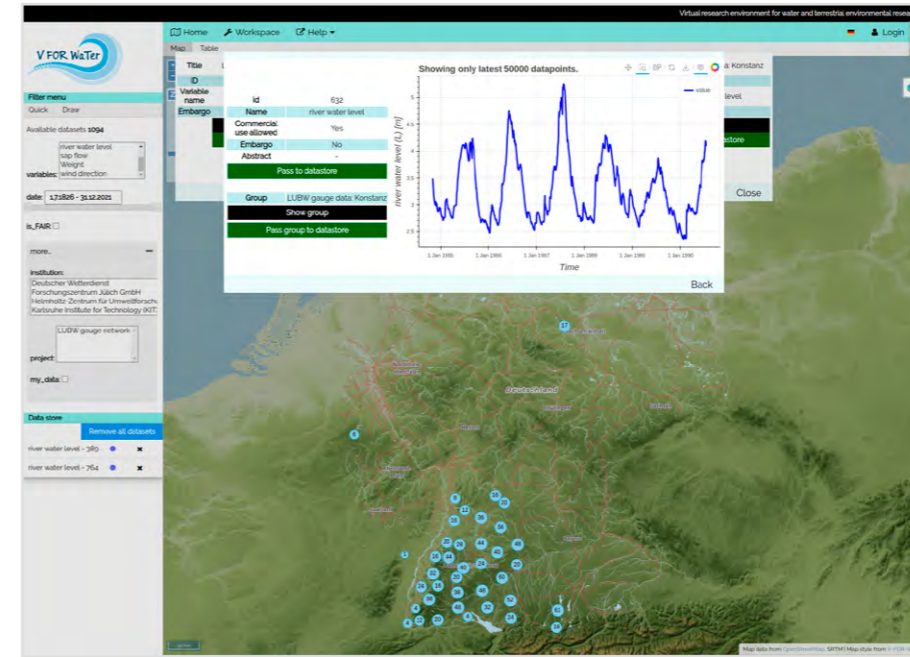


Abbildung 1: Das Portal von V-FOR-WaTer

Personen hier am SCC und fünf weitere am IWG (Abbildung 2). Wir haben eine klar gegliederte Arbeitsaufteilung in den zwei Instituten. Am SCC entwickeln wir das Webportal, halten die Services am Laufen und kümmern uns um die Online-Dienste. Die Datenbank und die Tools, die im Portal zur Verfügung stehen, werden von den Hydrologen am IWG entwickelt. Wenn das Portal produktiv wird, sind wir aber auch an der Beteiligung der hydrologischen Gemeinschaft interessiert, die dann das Portal mit eigenen Tools erweitern kann.

Wie wurde das Portal in den vergangenen sechs Jahren aufgebaut?

Das Ursprungsprodukt wird ständig weiterentwickelt, und es gibt immer noch neue Ideen für weitere Funktionen. Wir erkannten schnell, dass es wichtig ist, einen gewissen Ausbau und Vielfalt der Datensätze zu erreichen, um relevant zu werden und überhaupt an Daten und letztendlich Nutzende zu kommen. Eine der Hauptfragen war stets „Welche Daten können wir anbieten?“ und davon abgeleitet „Woher bekommen wir die Daten?“ Daten von Universitätsprojekten sind relativ einfach zu bekommen, damit haben wir auch angefangen zu entwickeln. Aber wir wollen eben auch Daten anbieten, die generell

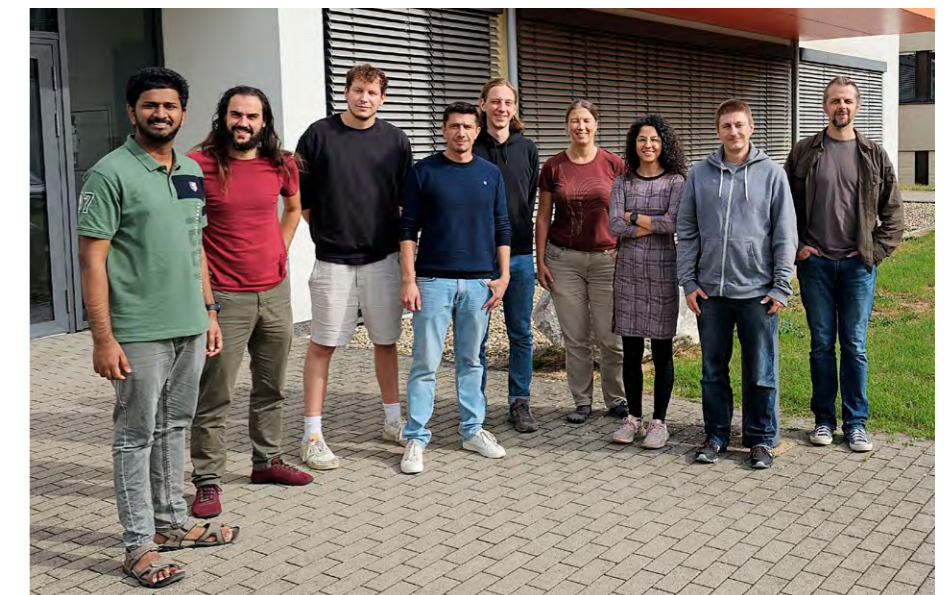


Abbildung 2: Das Projektteam V-FOR-WaTer – Ashish Manoj Jaseetha, Mirko Mälicke, Balazs Bischof, Carlos Alberto Zuleta Salmon, Alexander Dolich, Sibylle Hassler, Elnaz Azmi, Jörg Meyer und Marcus Strobl (v. l. n. r.)

sich lohnt, mit uns zusammenzuarbeiten. Wir hatten also reichlich Pläne, aber diese Ideen zu entwickeln und umzusetzen, brauchte viel Zeit. Und langsam wird das Projekt sogar zu groß, als dass wir es in unserem kleinen Team stemmen könnten. Trotzdem, wir entwickeln weiter, und unser Produkt ist in seinen Grundzügen nutzbar. Um das Projekt besser reproduzieren und um unsere Dienste besser verwalten zu können, haben wir angefangen, die unterschiedlichen Module in Docker-Container zu verteilen (Abbildung 3). Mit unserem kürzlich bewilligten DFG-Antrag können wir neue Personen einstellen, um sowohl am IWG als auch am SCC weitere Neuerungen einzubauen und für Nutzende attraktiver zu werden.

Welche Daten fließen in das V-FOR-WaTer-Portal ein?

Die Daten sind sehr divers. Eigentlich alle Daten, die in den hydrologischen Wissenschaften gebraucht werden. Es gibt verschiedene Zeitreihen wie zum Beispiel Daten von Flusspegeln, Abflussdaten, Luftdruck und -temperatur, Windrichtung und -geschwindigkeit, Niederschlag und Evaporation, Saftflussdaten von Bäumen, Luftfeuchte, und dazu digitale Höhenmodelle als Rasterdaten und Vektordaten zum Beispiel von Flusseinzugsgebieten.

¹ www.vforwater.de/english/index.php

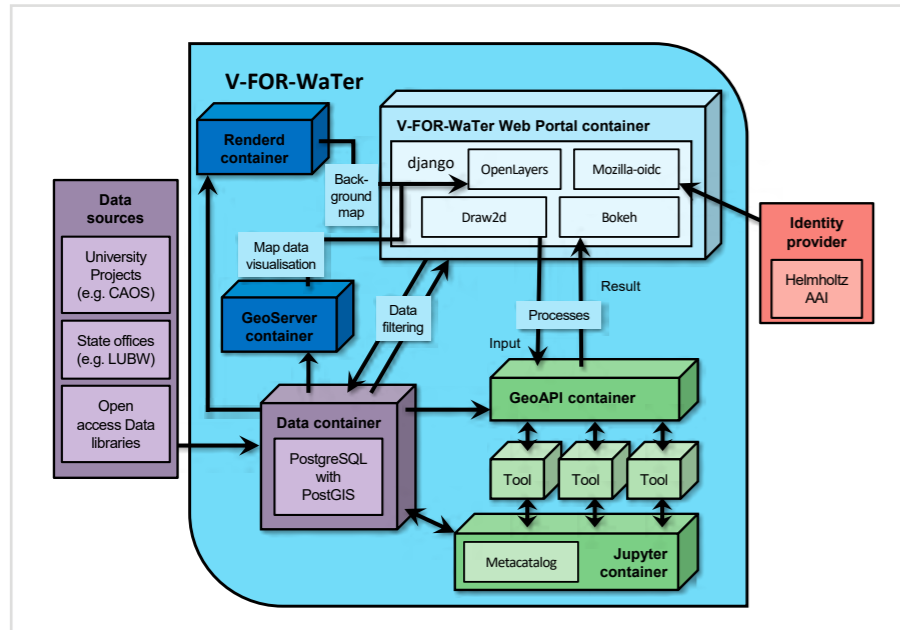


Abbildung 3: Architektur des Webportals V-FOR-WaTer mit allen Komponenten in separaten Docker-Containern

Insgesamt also eine sehr umfangreiche Datenbasis.

Werden weitere Daten in V-FOR-WaTer aufgenommen?

Im Portal werden sämtliche Daten, die etwas mit Wasser zu tun haben, zusammengebracht und ähnlich wie in einem Geoinformationssystem dargestellt. Das heißt, wir sehen die Daten auf einer Karte, können dort gezielt eine Region auswählen und zum Beispiel betrachten, wie Niederschläge oder Feuchtigkeit sich dort über die Zeit entwickeln. Die Flut an neuen Daten ist auch der Grund, dass das Projekt immer weiterwächst. Das Erweitern mit bereits bekannten Daten, wie z. B. Abflussdaten von neuen Standorten, kann ohne Probleme erfolgen, weil diese immer ähnlich erhoben werden und keine neuen Datenbanktabellen und keine weitere Visualisierung dafür erzeugt werden müssen. Spannend wird es allerdings mit völlig neuen Daten, die weder in der Datenbank noch am Frontend mit den bisher vorhandenen Diagrammen dargestellt werden können. Aktuell arbeiten wir daran, Regen-Radar-Daten mit aufzunehmen und stoßen damit an technische Grenzen der verwendeten Postgres-Datenbank; und die Visualisierung dieser Daten ist auch noch einmal eine Herausforderung.

Werden Daten anderer Anbieter über V-FOR-WaTer zugänglich sein?

Erstmal werden Daten angeboten, die wir selbst in der Datenbank gespeichert haben. Daten aus Universitätsprojekten sind oftmals zwar nicht so umfangreich, aber diese im Portal zu integrieren ist ein großer Fortschritt gegenüber der beliebten Praxis vieler Wissenschaftler, die Daten irgendwo als CSV-Data abzuspeichern. Im Portal verknüpfen wir die Daten mit umfangreichen Metadaten, um den Inhalt möglichst nutzungsfreundlich darstellen zu können und insbesondere, um den Nutzenden der Tools Hinweise geben zu können, welche Daten sie miteinander verknüpfen können. Das heißt, das Portal bietet viele Informationen, die Hydrologen üblicherweise nicht zu ihren Daten bereitstellen. Um von externen Datenbanken, also von externen Datenportalen, etwas anbieten zu können, müssen wir noch die entsprechenden Schnittstellen entwickeln. Derzeit sind wir mit den GFZ Data Services² im Gespräch, wie wir auf deren Daten zugreifen können. In einer Fallstudie ist angestrebt, dass deren Daten in V-FOR-WaTer eingelesen und im Portal dargestellt werden können, aber nicht permanent gespeichert werden. Das gleiche streben wir auch mit Daten

² dataservices.gfz-potsdam.de

der LUBW an. Damit die LUBW mitspielt, muss V-FOR-WaTer allerdings erst eine ausreichend große Datensammlung bzw. Nutzerzahl erreichen. Daneben sind wir auch noch dabei, bestimmte Daten des offenen Datenportals des Deutschen Wetterdiensts (DWD) einzupflegen. Es gibt also zum einen eine eigene Datenbank, in der Daten gesammelt werden, und zusätzlich möchten wir über V-FOR-WaTer den Zugriff auf externe Daten ermöglichen.

Sind Daten immer frei verfügbar?

Prinzipiell ja, allerdings gibt es Daten, die nicht jeder sehen darf und die wir nur auf Anfrage freigeben können. Das gilt sowohl für die Mess- als auch für die Metadaten. Dafür verwendet das Portal Nutzerprofile, womit der Zugriff eingeschränkt wird. Umweltwissenschaftler teilen mit anderen oftmals ihre Daten erst, wenn sie selbst eine Publikation auf Basis ihrer Daten veröffentlicht haben. Damit sie dennoch unser Webportal nutzen und ihre Daten hochladen, können sie ein bis zu zweijähriges Embargo beantragen. Während dieser Zeit sind die Daten nur für sie beziehungsweise für Personen, denen sie den Zugriff erlauben, erreichbar. Nach Ablauf des Embargos, oder nach der Publikation, werden die Daten für jedermann frei verfügbar sein. Im Fall der LUBW-Daten ist es wichtig, einen Teil der Metadaten zu schützen. Zum Beispiel muss der Zugriff auf die Daten der Position von Brunnen aus Sicherheitsgründen eingeschränkt werden, um keine Verunreinigung des Trinkwassers durch Kriminelle zu begünstigen. Für die LUBW kann es dennoch interessant sein, Daten mit Restriktion über unser Portal gezielt auszugeben, da darüber vereinfacht bestimmte Datensätze für bestimmte Nutzergruppen freigegeben werden können. Das ist gegenüber der LUBW eines unserer „Verkaufsargumente“, wobei wir auf solch eingeschränkte Daten auch keinen vollen Zugriff bekommen, sondern nur entsprechend der Nutzeranfragen Daten weiterreichen werden.

Wer ist eure Zielgruppe?

Die Zielgruppe wird in erster Linie die Wissenschaft sein. Daneben besteht auch Interesse von Ingenieurbüros, wenn über das Portal zentralisiert und vereinfacht Daten von Landesanstalten und -ämtern erreichbar sind. Aber wir wollen ja mehr anbieten als nur Daten und stellen auch Werkzeuge zur Prozessierung dieser zur Verfügung. Solche Features sind wahrscheinlich nur für Forschende von Interesse. Um diese zu erreichen, machen wir Werbung auf Konferenzen und Workshops. Dabei erfahren wir auch mehr über die Wünsche der Wissenschaftlerinnen und Wissenschaftler, sodass wir das Portal im engen Austausch mit ihnen und nach deren Interessen entwickeln. Dadurch maximieren wir die Akzeptanz, so dass diese das Portal nutzen und Daten darin einstellen. Das heißt, es werden aktiv Daten gesammelt, und das regionale Netz breitet sich langsam aus. Angefangen mit Daten aus Karlsruhe über ganz Baden-Württemberg, und momentan auch aus dem Harz, aus Garmisch und dem Elsass. Es gibt hier keine Begrenzung, und die erweiterte Datenbasis bedeutet, dass das Portal für einen erweiterten Kreis interessant wird. Die Idee dahinter ist, dass man auf diese Weise ein selbstlaufendes System produziert, was mit Daten gefüllt wird, Daten ausgibt und Werkzeuge zur Verfügung stellt, und sich damit quasi selbst trägt. Vor allem auch deshalb, weil Analysen mit den Daten und Datenverarbeitungstools zusammen in unserem Portal komplett reproduzierbar werden. Aber letztendlich ist das Portal offen für jedermann und kann von allen benutzt werden, die mit den Daten etwas anfangen können. Dabei orientieren wir uns an offenen Standards. Wir können die Daten z. B. im ISO19115-Format exportieren, damit sie eben auch mit Daten von anderen Quellen kombinierbar sind.

Wie versucht ihr, die Datenbasis zu erweitern?

Neben der genannten Werbung durch Präsentationen bei Workshops und Konferenzen gibt es die Idee, mit kleinen

Sensoren Werbung zu machen, die den Saftfluss in Bäumen messen und deren Daten dann direkt ins Portal übertragen werden können. Das wäre auch ein Testlauf ob und wie wir Daten von Citizen Science³ mit aufnehmen können. Die Hydrologen basteln allerdings noch an der Technik dafür. Die Sensoren würden wir dann erstmal auf Konferenzen verteilen. Daten von Saftflüssen sind bereits im Portal. Das ist hydrologisch von Interesse, weil diese als ein Indikator der natürlichen Verdunstung herangezogen werden können. Verdunstung findet nicht nur über den Erdboden statt, sondern eben auch über Pflanzen oder Bäume. Diese Daten fließen dann wiederum in größere Modelle ein.

Was sind jetzt die nächsten Schritte, und welche Highlights können wir noch erwarten?

Das nächste Highlight ist ein Hackathon, den wir für Anfang 2024 anvisiert haben. Auf dem Hackathon können Studierende weitere Tools entwickeln bzw. die vorhandenen Tools testen und zu deren Verbesserung beitragen. Entsprechend ist geplant, dass bis Ende des Jahres die Prozessierungstools im Portal flüssig benutzt werden können. Wenn wir weiter produktiv am Portal arbeiten, gewinnen wir weitere Anwender und schaffen mehr Nutzungszufriedenheit. Das Hauptziel im jetzigen DFG-Projekt ist es daher, das System in den nächsten zwei Jahren von der Prototyp- in die Produktionsphase zu bringen. Zugleich werden weiterhin Daten und Werkzeuge gesucht und eingebracht. Und wir stellen das Portal weiterhin bei Kongressen wie zum Beispiel auf der Jahrestagung der European Geoscience Union (EGU) in Wien vor. Zudem versuchen wir, uns mit der NFDI4earth⁴ zu verknüpfen.

Vielen Dank Elnaz und Marcus für dieses Gespräch.

³ Eine Kooperation von Forschenden mit Bürgerinnen und Bürgern. Mit vielfältigen Projekten stärkt das BMBF Citizen Science in Deutschland und deren nachhaltige Verankerung im Wissenschaftssystem.

⁴ www.nfdi4earth.de/

Easy access to hydrological data in Germany supports environmental research

A working group of the V-FOR-WaTer project is developing a web portal in which environmental scientists can easily and quickly access hydrological data from different sources, but also share their own data. In addition, KIT researchers are working on interfaces to existing data portals so that already published data sets can be included. With the help of integrated tools, these data are easily scalable and available in a uniform format. In addition, analysis tools for evaluations and uncertainty assessments are available in the portal. Easy saving and sharing of the analysis steps and the tools used makes the calculations reproducible. Repositories connected in the portal simplify publication of data of its users. Thus, V-FOR-WaTer and the further development of the portal within the framework of the new DFG-funded project ISABEL, offers researchers from the water and environmental sciences access to data from e.g. state offices or university projects and to standardized tools for reproducible data analysis steps. In close collaboration with the Institute for Water and River Basin Management at KIT and the hydrological community in Germany the development team at SCC continuously expands the functionality of the portal.

Im Zeichen der Zeit – der zweite AI-Hero Hackathon

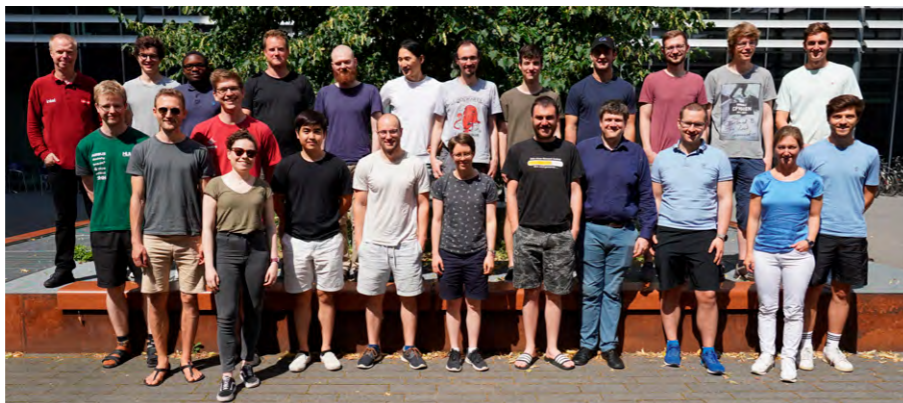
Am zweiten AI-Hero Hackathon on Energy Efficient AI, der vom 19. bis 21. Juni 2023 stattfand, nahmen 28 Wissenschaftlerinnen und Wissenschaftler aus ganz Deutschland teil. Sie arbeiteten in Teams mit Modellen und Ansätzen der Künstlichen Intelligenz und des Machine Learnings an beispielhaften Forschungsfragen und vertieften so ihr Wissen im energieeffizienten Umgang mit diesen Methoden.

René Caspart, Charlotte Debus, Markus Götz

Im Juni 2023 fand zum zweiten Mal der AI-Hero Hackathon on Energy Efficient AI statt. Dieser Hackathon wird von Teams der Helmholtz-Plattformen Helmholtz AI, Helmholtz Imaging und Helmholtz Metadata Collaboration (HMC) in Kooperation mit dem Nationalen Hochleistungsrechenzentrum am KIT (NHR@KIT) und der KIT-Graduiertenschule Computational and Data Science (KCDs) organisiert. 28 Wissenschaftlerinnen und Wissenschaftler von Forschungseinrichtungen der Helmholtz Gemeinschaft und Universitäten aus ganz Deutschland nahmen an dem Hackathon teil, der nach einer ersten, rein virtuellen Veranstaltung in 2022 dieses Jahr nun am Deutschen Krebsforschungszentrum (DKFZ) in Heidelberg in Präsenz stattfand. In Teams von jeweils drei Personen arbeiteten die Teilnehmenden mit Unterstützung der Tutoren an je einem von zwei Anwendungsfällen aus den Forschungsbereichen Energie und Gesundheit.

Ein Hackathon mit zukunftsweisender Zielsetzung

Dabei verfolgte der Hackathon nicht nur das klassische Ziel der Performance-Verbesserung von Workflows, sondern zielte darauf ab, die Teilnehmenden für den Energieverbrauch moderner KI-Workflows zu sensibilisieren und Möglichkeiten zur energieeffizienten Nutzung von KI-Modellen zu erlernen und testen zu können. Für das Training und die Inferenz ihrer Modelle verwendeten die Teilnehmenden die Rechenkapazitäten des Hochleistungsrechners Karlsruhe (HoreKa), der besonders geeignet ist, den Energieverbrauch von Workloads zu messen und diesen Wissenschaftlerinnen und Wissenschaft-



Teilnehmende und Mentoren des zweiten AI-Hero Hackathons

lern zur Verfügung zu stellen. Nils Smeds von der Firma Lenovo, Hersteller des HoreKa, gab mit seiner Keynote persönliche Einblicke in die Entwicklungen im HPC-Bereich und die Bedeutung von Energieeffizienz für KI aus Sicht eines Herstellers von HPC-Systemen.

Keine einfache Aufgabe

Die Aufgabe, die die Teilnehmenden im Rahmen des Hackathons angingen, entpuppte sich für viele schnell als eine anspruchsvolle Balance: Initiale Recherche, Überlegungen und Datensatzverständnis mussten gegen das Austesten verschiedener Modelle abgewogen werden; Aufwändige Modelle und Ansätze, die eine höhere Genauigkeit versprechen, gegen den dafür nötigen höheren Energieverbrauch. Im Rahmen dieser Abwägungen ergaben sich für die Teams jeweils andere Schwerpunkte, die mit unterschiedlichsten Ansätzen bearbeitet wurden, so dass es am Ende acht teils sehr verschiedene Herangehensweisen und Ergebnisse gab, die alle Teilnehmenden zum Abschluss in einem Pitch vorstellten.

Second AI-Hero Hackathon on Energy Efficient AI

At the second iteration of the AI-Hero Hackathon on Energy Efficient AI, 28 participants tackled the challenge of energy efficient AI and ML models. The Hackathon took place from 19th–21st of June 2023 as an in-person event at Deutsches Krebsforschungszentrum (DKFZ) in Heidelberg. The participants worked as teams of three persons each on one of two AI use cases from the fields of energy and health, tackling the fine line between achieving the optimal performance while at the same time being as energy efficient as possible.

12. HPC-Statuskonferenz der Gauß-Allianz zu Gast am KIT

Vom 31. Mai bis 2. Juni 2023 war das SCC Gastgeber der HPC-Statuskonferenz der Gauß-Allianz, die zum 12. Mal Forschenden eine Plattform zum interdisziplinären Austausch im Bereich des Hoch- und Höchstleistungsrechnens in Deutschland sowie zur Netzbildung bot. In diesem Jahr widmete sich die Konferenz vor allem den Themen Nachhaltiges Rechnen und Künstliche Intelligenz (KI).

Jasmin Hörter



In Zeiten von Energiekrisen, Klimawandel und dem Beginn der Exascale-Ära stehen HPC-Zentren vor besonderen Herausforderungen. Um den Wettlauf zu immer größeren und schnelleren Systemen mit der Ressourcenknappheit in Einklang zu bringen, bedarf es kluger Lösungen, sowohl für die Planung, Beschaffung und den Betrieb der Hardware als auch im Software-Bereich.

Dieses Spannungsfeld nahm auch Vizepräsidentin Professor Kora Kristof (Foto unten) in ihrem Grußwort zu Beginn der Konferenz auf. Sie betonte dabei die Bedeutung des SCC nicht nur als HPC-Betreiber für das KIT, sondern auch seine Rolle im Landesnetzwerk bwHPC und dem Verbund für Nationales Hochleistungsrechnen (NHR-Verbund).



Im Anschluss durfte das SCC Dr. Stefan Schenk (Foto rechts) von der BASF begrüßen, der die Keynote zu HPC-Anwendungen in der Industrie hielt. Simulationen sind heute aus keinem Wirtschaftsfeld mehr wegzudenken und kommen von der

Medizinforschung bis zur Waschmittelentwicklung überall zum Einsatz.

Danach stellten sich die vom Bundesministerium für Bildung und Forschung (BMBF) geförderten GreenHPC-Projekte (Energieeffizientes High Performance Computing) vor. Neben Forschung im Bereich energieeffizienter Klimasimulationen wurden auch Projekte wie ESN4NW¹ und WindHPC vorgestellt. Sie haben sich zum Ziel gesetzt, Rechenzentren direkt dort anzusiedeln, wo Energie produziert wird – beispielsweise in Windrädern. Die massiven Türme bieten am Boden ausreichend Schutz und Lagerfläche für Server, die „grüne Energie“ wird direkt eingespeist, und für die nötige Kühlung wird die Höhe des Mastes ausgenutzt. In einem Kreislauf wird Wasser im Turm nach oben gepumpt, kühlt dort ab und wird erneut zur Kühlung verwendet. Ein Besuch am KIT Campus Nord mit einer Führung durch den Hochleistungsrechner Karlsruhe (HoreKa) rundete das Programm des ersten Tages ab.

Nicht nur im Hardware- sondern auch im Software-Bereich bereitet man sich auf die Exascale-Ära vor. Die sogenannten SCALEX-Projekte werden ebenfalls vom BMBF gefördert und haben sich zum Ziel gesetzt, skalierbare und innovative Methoden zu entwickeln, um die kommenden Systeme sowohl zeit- als auch energieeffizient

¹ Energieoptimierte Supercomputer-Netzwerke durch die Nutzung von Windenergie

zient nutzen zu können. Das SCC ist hier mit dem Projekt EXASIM vertreten, in dem Exascale-fähige Softwarewerkzeuge zur Strömungssimulation entwickelt. Hier wird die Open Source Software OpenFOAM über eine spezielle Softwareebene mit der Softwarebibliothek Ginkgo verknüpft und so auf GPUs anwendbar gemacht. Dies verkürzt die Rechenzeit und spart somit Energie.



Mit aufregenden neuen Tools wie ChatGPT, FourCastNet oder Mindjourney herrscht im Bereich der Künstlichen Intelligenz Goldgräberstimmung. Getrieben von immer mehr neuen Anwendungsfeldern wächst die KI-Forschungslandschaft rasant. Der damit verbundene, stetig wachsende Bedarf an Rechenressourcen ist wiederum für Rechenzentrumsbetreiber relevant, und so wird die Zusammenarbeit der HPC-Community mit den KI-Forschenden immer wichtiger.

In einer Podiumsdiskussion mit Betreibern und KI-Fachleuten verschiedener Bereiche wurde ein Dialog angestoßen, um den



Die Gauß-Allianz ist ein gemeinnütziger Verein, der das Wissenschaftsthema „High Performance Computing“ (HPC) als eigenständige strategische Forschungsaktivität fördert und für eine verbesserte internationale Sichtbarkeit der deutschen Forschungsanstrengungen auf diesem Gebiet sorgt. Der Fokus liegt dabei auf der Erforschung und Entwicklung von Strategien zur Verbesserung der Effizienz, Anwendbarkeit und vereinfachten Nutzbarkeit des Hoch- und Höchstleistungsrechnens.

großen Themen Energiekrise und Klimawandel gemeinsam entgegenzutreten.

In der Keynote zum Thema „KI im Zeitalter des Exascale Computing“ griff Charlotte Debus, Leiterin der Forschungsgruppe Robust and Efficient AI am SCC, die großen Themen der vergangenen Konferenztage nochmals auf und gab den Teilnehmenden

eine klare Botschaft mit auf den Weg: Aktuell ist das Leben für HPC- und KI-Experten (noch) gut, doch mit dem immer größeren Bedarf an Rechenzeit und den ungewissen Energiekosten der Zukunft gilt es, sich heute schon fit für die Zukunft zu machen. Das Schlusswort des Vorsitzenden der Gauß-Allianz, Wolfgang Nagel, rundete die Konferenz ab.



12th HPC Status Conference of the Gauss Alliance at KIT

From May 31 to June 2, 2023, SCC hosted the HPC Status Conference of the Gauss Alliance, which for the 12th time offered researchers a platform for interdisciplinary exchange in the field of high performance and supercomputing in Germany and for networking. This year, the conference focused primarily on the topics of sustainable computing and artificial intelligence (AI).

New Work am SCC: gemeinsam, flexibel, agil – innovativ

Flexibles, agiles Arbeiten ist Teil der Aufgabenkompetenz im SCC, mit der sich die Mitarbeiterinnen und Mitarbeiter den stetig wandelnden Anforderungen der Digitalisierung stellen. So traf am SCC der Aufruf für einen Projektantrag zum Leitprojekt New Work am KIT auf offene Ohren. Im Verlauf des Projekts wurde ein Drei-Säulen-Modell für die individuellen Bedürfnisse entwickelt, flexible Arbeitsplätze mit Raumbuchungssystem eingerichtet sowie selbstorganisierte Teamarbeit in einem kollaborativen und ortsunabhängigen, hybriden Arbeitsmodell eingeführt.

Karin Rische

Serendipität: nicht gesucht und doch gefunden

Der Projektantrag wurde im SCC von einer freiwilligen Gruppe Mitarbeitender formuliert, aus der dann auch die koordinierende Projektsteuergruppe entstand. Mit Direktoren, Abteilungsleitenden und Mitarbeitenden sowohl aus dem Service als auch aus dem Forschungsbereich, waren alle Ebenen und Aufgabenfelder des SCC vertreten. Die Arbeit in der Steuergruppe gestaltete sich entlang holokratischer Prinzipien¹ sehr flexibel. Insbesondere die integrative Entscheidungsfindung, bei der eine

die Erfahrung machen, dass Anregungen und Kritik direkt in das Projekt einfließen konnten.

pen schnell und unbürokratisch zur gemeinsamen Arbeit in einem Raum treffen. Schon am nächsten Tag finden sich neue



Abbildung 1: Die unterschiedlichen Arbeitsplatzmodelle am SCC

„Wir begreifen New Work nicht nur als Neuausrichtung mobilen Arbeitens, sondern auch als Ermöglichung und Hinwendung zu agilen Arbeitsformen entlang holokratischer Prinzipien.“
 SCC-Projektantrag 2022

eventuelle Kurskorrektur und Anpassung Teil der Steuerung ist, ermöglichte die kreative Bearbeitung von Fragen, die zu Beginn des Prozesses noch nicht absehbar waren (Serendipitätsprinzip). In einer wöchentlichen Gesprächsrunde, an der alle Mitarbeitenden des SCC teilnehmen konnten, wurden Neuerungen vorgestellt, ihre Praxistauglichkeit diskutiert und notwendige Anpassungen vorgenommen. Diese offene Kommunikation, flankiert durch regelmäßige Newsletter per Mail und zwei Umfragen, trug im Projekt New Work sehr zur Akzeptanz und Mitwirkung bei. Dabei konnten alle

¹ de.wikipedia.org/wiki/Holokratie

Drei Säulen der individuellen Arbeitsplatzgestaltung

Das Raumkonzept des SCC wird durch drei Säulen gebildet, die gleichwertig nebeneinanderstehen: fester Arbeitsplatz, Bürogemeinschaft und Vollflex (Abbildung 1). Obwohl ein fester Arbeitsplatz, der der klassischen Arbeitsumgebung entspricht, keine Fläche für gemeinschaftliche Nutzung freigibt, hat er aufgrund der Tätigkeit oder aus persönlichen Gründen seine Berechtigung. In Bürogemeinschaften finden sich Teams oder Abteilungen zusammen. Sie nutzen gemeinsam einen Pool von Arbeitsplätzen, wobei es mehr Mitarbeitende als Plätze gibt. Die Absprache, wer wann welchen Platz belegt oder wie z. B. Technik oder Stauraum genutzt wird, organisiert die Bürogemeinschaft selbst. Wer sich für Vollflex entscheidet, bucht sich für jeden Arbeitstag am SCC flexibel über das neu geschaffene Raumbuchungstool einen freien Platz. Dabei können sich Arbeitsgrup-

pen schnell und unbürokratisch zur gemeinsamen Arbeit in einem Raum treffen. Schon am nächsten Tag finden sich neue Konstellationen zusammen. Sie arbeiten in den sogenannten Flex-Offices, die durch die Flächenfreigabe der „Vollflexer“ und Bürogemeinschaften entstanden sind. Diese können auch von den Mitgliedern der Bürogemeinschaften gebucht werden. Darüber hinaus konnten neue Raumkonzepte für agiles Arbeiten in diversen Kreativräumen (Abbildung 2) eingerichtet werden, die allen Mitarbeitenden zur Verfügung stehen. „Wir haben bisher sehr gute Erfahrungen gemacht, weil genau die Wirksamkeit der Personen gestärkt wird und in Verbindung mit Team-Chartas für die individuellen Bedürfnisse quasi wörtlich ‚das richtige Modell zur rechten Zeit am rechten Ort‘ gewählt werden kann“, so Martin Nußbaumer, Direktor des SCC.

Von Räumen zu Flächen

Um das Flächenmanagement der flexibleren Nutzung anzupassen, setzt das SCC einen Abteilungsflächenquotienten ein.

Leitprojekt New Work

Aufbauend auf den Erfahrungen der Corona-Lockdowns startete im September 2021 das KIT-Leitprojekt New Work, um neue Ansätze für die Zukunft der Arbeit am KIT zu erproben. Fünf Pilotprojekte testeten verschiedene Ideen und ganzheitliche Lösungen. Ob hybrides Arbeiten und Führen, neue Raumkonzepte, Plattformen und Werkzeuge: Der Kern ist dabei immer, die Perspektive der Beschäftigten in den Vordergrund zu rücken.

Weitere Informationen:

www.sts.kit.edu/projekt-new-work.php
KITdialog 01/2023, S. 13



Abbildung 2: Der Kreativraum bietet Platz und Inspiration für neue Formen der Zusammenarbeit

Dazu wird die von einer Abteilung genutzte Fläche in Bezug zur Mitarbeitendenzahl in einen Vergleichswert umgerechnet, der auch den Anteil der Abteilung am Drei-Säulen-Modell berücksichtigt. Die Diskussion über Flächenbedarfe wurde damit versachlicht und ein flexibles Instrument für die Raumvergabe geschaffen, das z. B. auch auf den personellen Aufwuchs einer Abteilung reagieren kann. In Bewerbungsgesprächen wird inzwischen gezielt nach flexiblen Arbeitszeit-/ort-Modellen gefragt. Das SCC hat sich durch das Drei-Säulen-Modell einen Standortvorteil bei der Personalakquise geschaffen. Martin Nußbaumer ist begeistert: „Die Mitarbeitenden erfahren sich mit dieser Art mobilen Arbeitens als selbstwirksam und können ihre ‚Work-Life-Balance‘ verbessern. Wir steigern damit unsere Attraktivität als Arbeitgeber auf einem hart umkämpften Arbeitsmarkt“.

Der Umbau der freigewordenen klassischen Arbeitsplätze in Flex-Offices und Kreativräume wurde im laufenden Betrieb realisiert. Der Service des SCC läuft bei gleicher Qualität auch im Drei-Säulen-Modell reibungslos weiter. Die nötigen technischen Lösungen wie das Raumbuchungssystem, die Technik für hybride und virtuelle Konferenzen sowie ein buchungsabhängiges Schließsystem für die Flex-Offices konnte erfolgreich evaluiert und implementiert werden.

Das KIT-Leitprojekt New Work wird zurzeit in ein Folgeprojekt übergeleitet, in dessen Rahmen das SCC sein erfolgreiches Konzept weiterbetreiben und weiterentwickeln wird.

**New Work at the SCC:
collaborative, flexible,
agile – innovative**

Flexible, agile working is part of the job competence at SCC, where employees face the constantly changing requirements of digitalization. The call for a project application for the New Work flagship project at KIT therefore met with an open ear at SCC. Over the course of the project, a three-pillar model was developed to meet individual needs, flexible workstations with a room booking system were set up and self-organized teamwork was implemented in a collaborative and location-independent hybrid working model.

KIT feiert 40 Jahre Hochleistungsrechnen in Karlsruhe

Am 14. September 2023 feierte das SCC mit einem Festkolloquium die in Karlsruhe schon 40 Jahre umfassende Ära des Hochleistungsrechnens. Die geladenen Gäste aus Forschungspolitik und -management, Rechenzentrumsplanung, -bau und -betrieb sowie Wissenschaftlerinnen und Wissenschaftler, ehemalige wie auch aktive, konnten sich bei Vorträgen, Paneldiskussion sowie Ausstellung und Führungen aus erster Hand über die gesamte Bandbreite von 40 Jahren Hochleistungsrechnen in Karlsruhe informieren. Es gab viel Raum, Erfolge und Herausforderungen Revue passieren zu lassen, ein wenig über die Zukunft zu philosophieren, und natürlich gemeinsam zu feiern und interessante Geschichten auszutauschen.

Achim Grindler

Grußworte und Einführung

Der für HPC zuständige Referatsleiter im Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK), Peter Castellaz, hob in seinen Grußworten die Landesstrategie bwHPC heraus, mit der das SCC nicht nur wichtige inhaltliche Aspekte umgesetzt, sondern sich in einer intensiv gelebten Kooperationskultur federführend engagiert hat.



„Neben den HPC-spezifischen Ressourcen und dem zugehörigen Methodenwissen hat sich das KIT besonders mit seiner Expertise im Bereich Daten und Forschungssoftware erfolgreich eingebracht“,
so Peter Castellaz.

Lobende Worte fand er zum landesweiten föderierten Identitätsmanagement bwIDM, mit dem das SCC gemeinsam mit anderen Landeseinrichtungen entscheidende Grundvoraussetzungen für kooperativ erbrachte Dienste – auch über das HPC

hinaus – erarbeitet hat. Um neben technologischen Entwicklungen auch die Themen Forschungssoftware und Nachhaltigkeit zu akzentuieren, wird eine umfassende Landesstrategie bis zum Jahr 2032 erarbeitet, ließ Peter Castellaz wissen.

Die KIT-Vizepräsidentin für Digitalisierung und Nachhaltigkeit, Kora Kristof, zeigte sich in ihren Grußworten beeindruckt von der Community, die sich über lange Zeit im HPC-Umfeld entwickelte, vom Gauss Centre for Supercomputing (GCS) angefangen über die unterschiedlichen Zentren der Ebenen Tier 2 (Gauß-Allianz) und Tier 3, national und in Baden-Württemberg.



„Was im Bereich des HPC vom KIT und anderen Einrichtungen auf Landesebene entwickelt wurde, hat auch die Entwicklungen national mitgeprägt, und dafür gilt ein besonderer Dank“,
stellt Kora Kristof fest.

Zudem habe das KIT das Hochleistungsrechnen mit dem Thema Energieeffizienz verbunden und damit herausragende Erfolge erzielt. Hier sprechen der Deutsche Rechenzentrumspreis 2017 für ForHLR II und Platz 13 auf der internationalen Bestenliste der energieeffizientesten Rechner für HoreKa für sich. „Und zukünftig werden uns auch im HPC-Umfeld noch viele interessante Themen, die Nachhaltigkeit prägen, beschäftigen – dies betrifft nachhaltige Gebäude und Infrastrukturen, den Einsatz nachhaltiger Materialien und Ressourcenschonung sowie Aspekte der nachhaltigen Software“, prognostizierte Kora Kristof.

In der Überleitung zu den Fachvorträgen charakterisierte Martin Frank, Direktor des SCC, das HPC-Geschäft als eine Mischung aus etwas sehr Dynamischem und etwas Konservativem. „Das Dynamische ist an der Entwicklung von HPC-Systemen in den letzten 40 Jahren sehr deutlich zu erkennen, das Konservative steckt beispielsweise in der Bearbeitung sehr komplexer Prozesse wie Beschaffung und dem sicheren Betrieb der Infrastruktur“, konkretisierte Martin Frank in seinem Grußwort und wusste, dass Erfahrung und Innovation die wichtigen Pole im HPC-Geschäft sind, die das SCC im Nationalen Hochleistungsrechnen zu einer wichtigen Größe machen.



„All das sorgt dafür, dass Wissenschaftlerinnen und Wissenschaftler in ihrer Forschung mit Hochleistungsrechnern und Forschungssoftware bestmöglich unterstützt werden“,
so Martin Frank.

Anfänge des HPC in Karlsruhe

Eric Schnepf führte im ersten Fachvortrag in die Anfänge der Hochleistungsrechner in Karlsruhe ein und spannte einen Bogen über die Entwicklungen bis hin zur Gegenwart. Er machte seine ersten IT-Erfahrungen in den 70er-Jahren an der Universität Karlsruhe mit Algol-Programmen, die er mittels eines Siemens T100 Fernschreibers auf Lochstreifen erstellte und auf der Zuse Z 23 zum Ablauf brachte. In den 80ern konnte er sich neben den Universalrechnern auch mit Vektorrechnern vertraut machen, auf denen man Anwendungen genauer und schneller berechnen konnte. Den Beginn der HPC-Ära in Karlsruhe datierte Eric Schnepf auf Mai 1983, wo ein Landesvektorrechner, eine Cyber 205, nach vorherigen Tests an einer gleichartigen Maschine der Uni Bochum im Universitätsrechenzentrum in Karlsruhe installiert und betrieben wurde. Die Anwendungsbetreuung erfolgte durch ein Team aus Uni Karlsruhe und dem Forschungszentrum Karlsruhe (FZK). „Die Beschaffung kam nur zustande, weil eine große Community sehr gut zusammenarbeitete: Uni, FZK und Industriepartner“ bekräftigte Eric Schnepf in seinem Vortrag. Neben vielen interessanten technischen Ausflügen in

die Welt der damals in der Uni installierten Rechnersysteme, führte Eric Schnepf auch Anwendungsbeispiele – beispielsweise aus der Klimaforschung – und ging auf die zwischen Siemens-Nixdorf und der Uni geschaffenen Kooperation ODIN ein, was für Optimale Datenmodelle und Algorithmen für Ingenieur- und Naturwissenschaften auf Hochleistungsrechnern steht. Ein Meilenstein war die erste TOP500-Liste der Supercomputer, die im Jahr 1993 erschien. Das Original-Handout zeigt die Deutschlandliste mit den beiden erstplatzierten Systemen vom Typ S600/20 aus Karlsruhe und Aachen. Seinen Vortrag rundete Eric Schnepf mit einer Übersichtsfolie der wichtigsten HPC-Systeme der letzten 40 Jahre in Karlsruhe ab und ordnete diese in die Grenzlinien der TOP500 (Platz 1 .. Platz 500) ein.



„Von der Cyber 205 (1983) bis HoreKa (2023) sind das acht Zehnerpotenzen Leistungssteigerung, also im Schnitt alle 10 Jahre Faktor 100 Beschleunigung. Ich denke, das kann sich sehen lassen“,
wertschätzte Eric Schnepf die Entwicklung im HPC in Karlsruhe.

Kooperation im Virtuellen Rechenzentrum

Klaus-Peter Mickel, Physiker und ehemaliger Direktor des SCC, war schon Ende der 60er-Jahre als Programmierer für die IBM-Maschinen am Rechenzentrum des FZK tätig und erlebte und gestaltete ebenfalls die Entwicklungen der Hochleistungsre-

chenanlagen in Karlsruhe von Anfang an mit. Als er 1970 eine Stelle am Rechenzentrum Karlsruhe annahm, übernahm er die Betreuung der Uni-Mitarbeitenden, die die Maschinen des FZK benutzen wollten. Nach verschiedenen beruflichen Stationen hat Klaus-Peter Mickel 1996 dann die Leitung des Rechenzentrums am FZK angetreten. In seinem Rückblick auf die Jahre zwischen 1966 und 1996 schilderte Klaus-Peter Mickel die intensive Kooperation der Computerfachleute in Uni und FZK, die schließlich, ab 1996, zur Planung einer ausgeklügelten technischen und organisatorischen Kooperation beider wissenschaftlichen Rechenzentren führte – das Virtuelle Rechenzentrum Karlsruhe wurde gegründet. Virtuell, ja, denn es kam nämlich nicht, wie ursprünglich überlegt, zu einem gemeinsamen Rechenzentrum beider Institutionen an einem Standort, sondern zu einem Verbund mit einer juristisch abgesicherten Kooperationsvereinbarung.



„Wir gucken uns ab, was in der Industrie als Virtuelles Unternehmen bezeichnet wird, und bauen in Karlsruhe ein Virtuelles Rechenzentrum“,
so schilderte Klaus-Peter Mickel die Idee des damaligen Leiters des Uni-Rechenzentrums Professor Schreiner.

Es gab ein gemeinsames Leitungsgremium und unterschiedliche Architekturschwerpunkte auf beiden Seiten, jeweils mit gegenseitig eingebrachten Ressourcen. Die Uni legte ihren Schwerpunkt auf die massiv parallelen Rechner und das FZK auf die damals sehr leistungsfähigen



Vektorrechner. Eine eigens dafür gelegte Datenleitung verband über 10 km Luftlinie beide Rechenzentren und erreichte die damals respektable Geschwindigkeit von 155 Megabit pro Sekunde. Durch die Einrichtung des Virtuellen Rechenzentrums wurden viele positive Effekte erzielt. Neben einer hohen Effizienz, weil nicht beide Seiten beide Architekturen zu betreuen hatten, entstand ein großer Nutzen für die Anwendergruppen, weil ihnen beide Architekturen zur Verfügung standen.

Die Welt der Clustersysteme

Rudolf Lohner gab in seinem Vortrag einen intensiven Einblick in die Entstehung des Rechenzentrums der Uni und die damit verbundene Entwicklung sowie den Betrieb der massiv parallelen Rechner in Karlsruhe, der sogenannten Rechencluster. Rudolf Lohner war 20 Jahre Mitarbeiter bei den Mathematikprofessoren Alefeld und Kulisch, die er als Wegbereiter der ersten Stunden und Gründer des Uni-Rechenzentrums würdigte. Er ist dann in den 2000er Jahren vom Mathematikinstitut ans Rechenzentrum der Uni Karlsruhe gewechselt und war bis zum

Ende seiner aktiven Tätigkeit Experte für Energieeffizienz in Hochleistungsrechnern am SCC. Mitte der 90er haben sich zunehmend die massiv parallelen Rechencluster etabliert und auch am Rechenzentrum der Uni wurden solche Systeme über die letzten Jahre, bis heute, für die Forschung betrieben.



„Rechner, Rechnerinfrastruktur und -verwaltung werden ins Rechenzentrum ausgelagert. So können sich die Institute ihren eigentlichen Aufgaben widmen – Forschung und Entwicklung“,
sagte Rudolf Lohner zum erfolgreichen Konzept des Institutsclusters.

Er stellte in einem kurzweiligen Streifzug nicht nur die verschiedenen Clustersysteme vor, sondern auch einige wichtige Projekte und Anwendungsszenarien. Das Spektrum geht von der mathematischen Simulation von Segelboot-eigenschaften für den America's Cup, über die Erstellung von präzisen Wettervorhersagen bis hin zur Entwicklung institutseigener Clustermanagementsysteme. Für den heute am SCC betriebenen Hochleistungsrechner Karlsruhe (HoreKa) und dessen Vorgängersysteme ForHLR I und ForHLR II entwarf Rudolf Lohner gemeinsam mit dem HPC-Team das äußerst effiziente Energie- und Betriebskonzept. Der dazugehörige Neubau wurde 2015 fertiggestellt und beherbergt am Campus Nord des KIT deutschlandweit nutzbare HPC-Systeme. Vor ein paar Monaten wurden die notwendigen baulichen und technischen Vorbereitungen abgeschlossen, um das Gebäude auch für zukünftige, noch viel leistungsfähigere Rechensysteme mit bis zu zwei Megawatt Leistungsaufnahme fitzumachen.

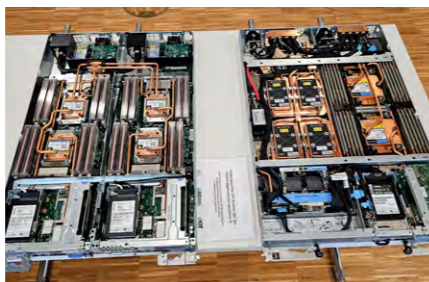


Foto: Martin Frank, Eric Schnepf, Klaus-Peter Mickel, Peter Castellaz, Simon Raffener und Rudolf Lohner (v. l. n. r.) bestaunen im Ausstellungsraum Rechnertechnik der vergangenen Jahrzehnte.

Danksagung

Im Anschluss an die spannenden Fachvorträge, die die gesamten 40 Jahre HPC in Karlsruhe aus verschiedenen Blickwinkeln beleuchteten, konnten die Gäste an Führungen durch die HPC-Infrastruktur teilnehmen sowie in einem eigens dafür eingerichteten Ausstellungsraum Rechenzentrumsexponate der letzten 40 Jahre bewundern.

Das SCC bedankt sich beim Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg, dem Präsidium des KIT, bei allen Beteiligten, die zur Veranstaltung 40 Jahre HPC in Karlsruhe beigetragen haben, sowie dem Organisationsteam dieses Festkolloquiums und natürlich bei allen seinen Gästen.



KIT celebrates 40 Years of High-Performance Computing in Karlsruhe

On 14th September 2023, SCC celebrated 40 years of high-performance computing at the university of Karlsruhe and KIT with a festive colloquium. The invited guests from research policy and management, data center planning, construction and operation, as well as scientists, both former and active, were able to learn first-hand about the entire range of high-performance computers in Karlsruhe during lectures, panel discussions, as well as an exhibition and guided tours. There was plenty of room to review successes and challenges, to philosophize a bit about the future, and of course to celebrate together and share interesting stories.

Fotos: Markus Breig

KIT benennt Informationstechnologie-Zentrum um

Distanzierung vom Namensgeber Karl Steinbuch – „Scientific Computing Center“ als neue Bezeichnung

Karl Steinbuch gilt als Mitbegründer der Informatik. Das Informationstechnologie-Zentrum des Karlsruher Instituts für Technologie (KIT) trug seit 2008 den Namen „Steinbuch Centre for Computing“. Neue Erkenntnisse zeigen nun, dass sich Karl Steinbuch, der in Karlsruhe forschte und lehrte, bereits als junger Mensch während der NS-Zeit mit nicht zu billigen Kriegshandlungen identifizierte. Unabhängig von seinen fachlichen Leistungen distanziert sich das KIT aufgrund der neuen Forschungsergebnisse von Steinbuch und wird daher das Informationstechnologie-Zentrum des KIT in „Scientific Computing Center“ umbenennen.

Nachdem Karl Steinbuch in seiner wissenschaftlich kreativen Phase eine gemäßigte politische Ausrichtung gezeigt hatte, wandte er sich nach der Emeritierung rechtsextremen Positionen zu. Da bereits 2017 zum 100. Geburtstag von Karl Steinbuch bekannt war, dass seine Ansichten mit zunehmendem Alter von der politischen Mitte abrückten, er sich nach seiner Emeritierung dem Rechtsextremismus zuwandte und in hohem Alter regelmäßig Artikel in der rechtsextremen NPD nahestehenden Zeitschriften veröffentlichte, verband das Präsidium des KIT die Anerkennung für die fachliche Leistung von Karl Steinbuch bereits 2017 ausdrücklich mit dem Hinweis, dass das KIT die politischen Ansichten, die Steinbuch in hohem Alter vertrat, nicht teile.

Neue Erkenntnisse über Karl Steinbuch

Mittlerweile liegen neue Erkenntnisse zu Steinbuchs Biografie vor, die zeigen, dass sich Karl Steinbuch bereits als junger Mensch während der NS-Zeit mit nicht zu billigen Kriegshandlungen identifizierte¹.

Namensänderung ab 1. Januar 2024

Das KIT hat daher eine Namensänderung des Informationstechnologie-Zentrums des KIT beschlossen. Die Entscheidung wurde in den Gremien des KIT bestätigt. Das KIT setzt die Namensänderung in „Scientific Computing Center“ ab 1. Januar 2024 um.

¹ (Anton F. Guhl: Kurskorrekturen eines Technokraten – Die politische Rechtswendung des Nachrichtentechnikers und Zukunftsforschers Karl Steinbuch nach 1970; Technikgeschichte Bd. 87 (2020) H. 4, S. 315-334)

IMPRESSUM

SCC news
Magazin des Scientific Computing Center

Herausgeber
Prof. Dr. Oliver Kraft
In Vertretung des Präsidenten des KIT
Karlsruher Institut für Technologie (KIT)
Kaiserstraße 12
76131 Karlsruhe

Anschrift
Scientific Computing Center (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 (verantwortlich),
Karin Rische, Andreas Ley
E-Mail: redaktion@scc.kit.edu

Gestaltung, Satz und Layout
Heike Gerstner, Nicole Gross, Elias Kobel
AServ – CrossMedia
Karlsruher Institut für Technologie (KIT)
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen

Titelfoto
Eric Schnepf, Klaus-Peter Mickel, Simon Raffener und Rudolf Lohner (v. l. n. r.) im Ausstellungsraum des SCC.
(Foto: Markus Breig, KIT)

Fotos
SCC, KIT

Druck
AServ – CrossMedia
Karlsruher Institut für Technologie (KIT)

Erscheinungstermin dieser Ausgabe
Januar 2024

www.scc.kit.edu/publikationen/scc-news

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



Karlsruher Institut für Technologie (KIT)
Scientific Computing Center (SCC)

ISSN: 1866-4954

www.scc.kit.edu
www.scc.kit.edu/twitter
contact@scc.kit.edu