

Wie bändigen wir die wissenschaftliche Datenflut? - Auf dem Weg zu internationalen Forschungsdateninfrastrukturen

**Prof. Dr. Dr. Thomas Lippert,
Direktor des Jülich Supercomputer Center, FZ Jülich**

**10. November 2016 um 17.15 Uhr
Seminarraum 9222, Gebäude E3, Informatik-Zentrum Hörn RWTH**

Abstract

Informations- und Kommunikationstechnologien haben sich zum tragenden Fundament der Wissenschaften des 21. Jahrhunderts entwickelt. Die disziplinäre Forschung widmet sich zunehmend komplexer werdenden Problemen wie z.B. dem Materials Genome, der Erdsystemforschung, der Bioökonomie, der nachhaltigen Energieerzeugung oder der Dekodierung des menschlichen Gehirns. Die Datenanalyse wird zum Schlüssel mit strategischer Bedeutung für die Forschung und auch die Wirtschaft treibt die Digitalisierung mit großen Schritten voran - Stichwort Industrie 4.0.

All dies geht einher mit einer wachsenden Heterogenität und einem exponentiell steigenden Umfang von erzeugten und zu verarbeitenden Daten, der Notwendigkeit des "data enrichment" für intelligente Suchverfahren, wie auch der unabdingbaren Unterstützung von Strukturen für die Zusammenarbeit. Schon längst wurden Big Data und cloud-basierte Software- und Plattform-Dienste (Software as a service - SaaS, Platform as a Service - PaaS) zu einem beherrschenden Thema für Hoch- und Höchstleistungsrechenzentren, deren Aufstellung zunehmend hin zu einer "Infrastructure as a Service" (IaaS) orientiert ist. Das Jülich Supercomputing Center ist auf internationaler Ebene, in Europa und in Deutschland aktiv, die genannten Herausforderungen mit föderierten Service-Infrastrukturen für die wissenschaftliche Datenanalyse anzunehmen.

Der Vortrag wird, ausgehend von Herausforderungen und Lösungsansätze im Bereich der HPC basierten Scientific Big Data Analytics, einerseits die Pläne der Helmholtz-Gemeinschaft für den Aufbau einer "Helmholtz Data Federation" beleuchten, andererseits näher auf die IT-Service-Architektur des europäischen Human Brain Flagship Projektes eingehen, die mit Hilfe einer cloud-service-basierten Kollaborations- und Entwicklungsplattform - dem COLLAB - SaaS und PaaS verwirklicht und den web-basierten Upstream- und Downstream-Zugang zu Meta-Datenbanken für die Speicherung und Verarbeitung und kollaborative Nutzung sehr heterogener wie auch sehr großer Datenmengen ermöglicht. Das HBP ist der Pilotanwender für eine europaweite IaaS namens FENIX - die Kurzform für "Federated ENGINE for Information eXchange" -, ein Projekt der im HBP vertretenen Supercomputing-Zentren aus Barcelona (BSC), Bologna (CINECA), und Lugano (CSCS). Damit wird das HBP, welches das Ziel unterstützt, Hirnforschung, medizinische Erkenntnisse und "brain-inspired" Informationstechnologie voranzubringen, auch zum Vorreiter einer pan-europäischen Forschungsdateninfrastruktur.



Kontakt: Prof. Nagl, Tel. AC 8021350, nagl@informatik.rwth-aachen.de