



Was heißt eigentlich Grid-Computing?

Fasst man die Rechenleistung von Computern innerhalb eines Netzwerks zusammen, um nicht nur Daten auszutauschen, sondern zudem die parallele Lösung von rechenintensiven Problemen zu ermöglichen (verteiltes Rechnen), so spricht man von **Grid-Computing**. Grids sind gut skalierbar: durch Einbindung weiterer Rechner oder gar Verknüpfung mehrerer Grids kann die Leistung um Größenordnungen verändert werden. Ein solches Grid kann die Rechenkapazität und -geschwindigkeit von Supercomputern weit übertreffen.

Werden auch Datenbestände, auf welche die Grid-Nutzer zugreifen können, vernetzt bzw. für Grid-Anwendungen erschlossen, entsteht ein **Data Grid**, in Unterscheidung zum **Computing Grid**, in welchem die Verarbeitung dieser Daten stattfindet. Schließlich können auch gezielt unterschiedliche Funktionen oder Algorithmen, die jeweils in verschiedenen Systemen verfügbar sind, miteinander verknüpft und für komplexe Aufgaben simultan nutzbar gemacht werden.

Bestimmte Dienste oder Anwendungen sind oftmals integraler Bestandteil von Grids und werden durch diese den Nutzern zur Verfügung gestellt (**Application Sharing**).

Grid-Computing zielt also insbesondere auf die Zusammenführung, Auswertung und Darstellung großer Datenmengen ab, wie sie beispielsweise in der medizinischen Forschung, der Meteorologie oder Astronomie anfallen.

Einige Charakteristika von Grids erinnern an das herkömmliche **Stromnetz**. Das Grid stellt durch eine intelligente dynamische Verwaltung seiner verteilten Ressourcen Rechenleistung und Datenmaterial zur Verfügung. Dem Nutzer eines Grids ist letztlich egal, wo seine Daten genau verarbeitet werden. Er will sich möglichst einfach mit dem Grid verbinden und Rechenleistung beziehen.

Aus der Analogie zum Stromnetz drängen sich auch unmittelbar die Fragen nach geeigneten Betreibermodellen, Sicherheitsaspekten sowie Verfahren zur Messung des „Verbrauchs“ von Leistung und dessen Abrechnung auf. Insofern muss ein Grid als Gesamtkonstrukt gedacht werden: seiner „Inbetriebnahme“ muss die Beantwortung dieser Fragen ebenso vorausgehen wie die Lösung der hard- und softwaretechnischen Herausforderungen.



Das MediGRID-Konsortium

Das MediGRID-Konsortium besteht aus acht Partnern.

Sprecher: Otto Rienhoff

Universität Göttingen, Abteilung Medizinische Informatik
haegar@med.uni-goettingen.de

Modul A: Koordination

Telematikplattform für Medizinische Forschungsnetze e.V.
medigrid@tmf-ev.de
Sebastian C. Semler, Mathias Blaurock, Stefan Scholz

Modul B: Ontologie-Werkzeuge

Universität Leipzig
rahm@informatik.uni-leipzig.de
Markus Löffler, Erhard Rahm, Toralf Kirsten,
Michael Hartung, Katrin Loebe

Modul C: Ressourcenfusion

Konrad Zuse Institut, Berlin
steinke@zib.de
Thomas Steinke, Ralph Müller-Pfefferkorn, Ulrich Schwardmann

Modul D: Middleware

Fraunhofer Gesellschaft, Stuttgart
anette.weisbecker@iao.fraunhofer.de
Anette Weisbecker, Jürgen Falkner, Thilo Ernst

Modul E: Bioinformatik

Uniklinikum Schleswig-Holstein
jhampe@1med.uni-kiel.de
Jochen Hampe, Tobias A. Knoch,
Burkhard Morgenstern, M. Schroeder

Modul F: Bildverarbeitung

Charité, Universitätsmedizin Berlin
dagmar.krefting@charite.de
Dagmar Krefting, Thomas Tolxdorff, Michael Vossberg

Modul G: Klinische Forschung

Universität Marburg
canisius@staff.uni-marburg.de
Thomas Penzel, Richard Dodel, Wolfgang Oertel, Sebastian Canisius

Modul H: eScience

Universität Göttingen
usax@med.uni-goettingen.de
Ulrich Sax, Yassene Mohammed

Internetseite: <http://www.medigrid.de>

Applikationsportal: <https://portal.medigrid.de>



Vernetzte Rechnerleistung für die Biomedizinische Forschung

– ein Projekt im D-GRiD-Verbund –



Förderkennzeichen: 01AK803A - F

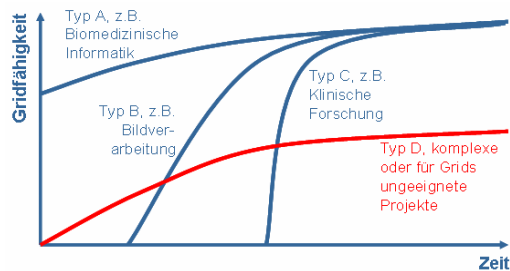


Das MediGRID-Projekt

Grid-Computing gilt inzwischen auch in der biomedizinischen Forschung als wichtiger Fortschrittsfaktor. Im Verbundprojekt MediGRID haben sich renommierte Forschungseinrichtungen in den Bereichen Medizin, Bioinformatik und Gesundheitswissenschaften als Konsortialpartner zusammengeschlossen und haben eine **Grid-Middleware-Integrationsplattform** und darauf aufsetzende eScience-Dienste für die biomedizinische Wissenschaft entwickelt. Die weitere Einbeziehung zahlreicher assoziierter Partner aus Industrie, Versorgungs- und Forschungseinrichtungen stellt das Projekt auf eine breite Interessensbasis.

Für das Projekt wurde eine modulare Aufgabenverteilung gewählt: In den vier methodischen Projekt-Modulen (Middleware, Ontologie-Werkzeuge, Ressourcenfusion und eScience) entwickelten die entsprechenden Konsortialpartner schrittweise eine Grid-Infrastruktur. Sie berücksichtigen dabei insbesondere die Anforderungen der Grid-Nutzer aus dem biomedizinischen Umfeld, welche exemplarisch in den drei anwendungsorientierten Projektmodulen (Biomedizinische Informatik, Bildverarbeitung, Klinische Forschung) erarbeitet wurden.

MediGRID soll anhand ausgewählter Anwendungsbeispiele zeigen, ob, wie und in welchen Fällen Grid-Technologie in der biomedizinischen Forschung gewinnbringend eingesetzt werden kann. Die Untersuchung der Anwendungsbeispiele wird dabei nicht parallel gestartet (Pipelining):



Typ-A-Projekte (s. Grafik) sind auf Grid-Computing vorbereitet und wurden sofort gestartet, andere können aufgrund der Erfahrungen aus den A-Projekten dann rasch gridfähig gemacht werden (Typen B und C). Projekte vom Typ D sind von vornherein zu vermeiden.

Die ersten Applikationen im Bereich Biomedizinische Informatik und Bildverarbeitung sowie Ontologie-Werkzeuge können Sie im MediGRID Applikationsportal selbst aufrufen:

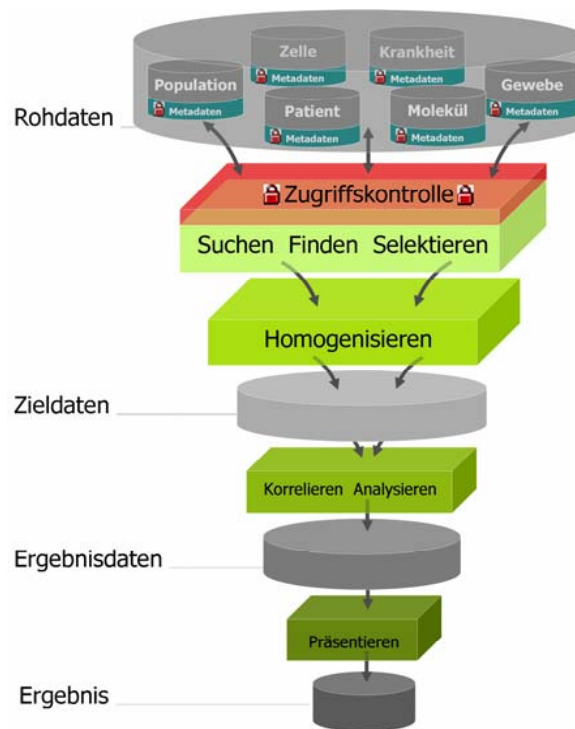
<https://portal.medigrid.de>

Zur internationalen Anbindung des Projektes findet ein intensiver Austausch mit Partnern in Frankreich, UK und den USA statt. Deshalb hat MediGRID auch internationale Workshops durchgeführt, z.B. auf dem AMIA-Kongress 2006 in Washington, DC (USA).



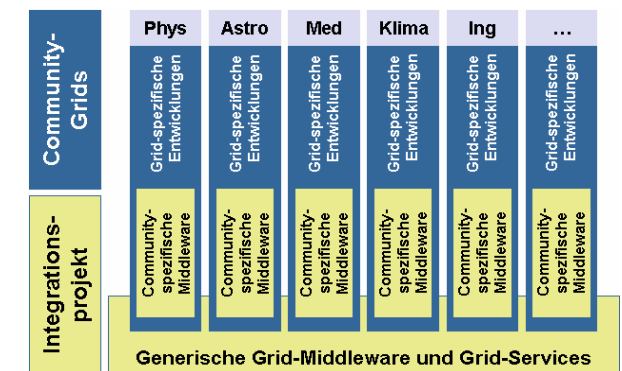
Von den Rohdaten zum Forschungsergebnis

Ein Grid, das in erster Linie für Forschungs-Communities design wurde, muss zunächst aus **heterogenen Datenquellen**, beispielsweise aus dem biomedizinischen Bereich, über entsprechende **Zugriffskontrollmechanismen** Datensätze **suchen und selektieren**. Wegen der Heterogenität der Daten ist dann ein **Homogenisierungsschritt** notwendig, bei dem mit Hilfe von Ontologien semantische Interoperabilität erreicht wird. Die homogenisierten Zieldaten können dann mit biometrischen Methoden **korreliert und ausgewertet**, schließlich verdichtet und als Ergebnisdaten **präsentiert** werden.



MediGRID im D-GRID-Verbund

Unter dem Dach der D-GRID-Initiative werden vom Bundesministerium für Bildung und Forschung (BMBF) weitere Grid-Projekte gefördert, unter anderem für die Bereiche Astronomie, Klimaforschung, Hochenergiephysik und Ingenieurwissenschaften. Durch das D-GRID-Integrationsprojekt werden die Aktivitäten der einzelnen Grid-Projekte koordiniert und Synergien genutzt, mit dem Ziel, letztlich eine gemeinsame Grid-Plattform für die deutsche Wissenschaftsgemeinde zu schaffen.



Weitere Informationen zu D-GRID und den verschiedenen Community-Projekten:

www.d-grid.de

