

# Das CrossGrid-Projekt

M. Kunze, IWR

## Grid-Anwendungen

Die moderne Forschung und Entwicklung bedient sich der Expertise einer großen Zahl von Wissenschaftlern, die mit unterschiedlichem Know-how zusammenarbeiten müssen. Diese interdisziplinäre Zusammenarbeit stützt sich dabei in der Regel auf den Austausch von Informationen in weltweit verteilten Organisationen. Grid-Computing ist ein Mittel der Wahl um diesen Prozess zu unterstützen: Grids bilden eine Infrastrukturbasis zur generischen Unterstützung verteilter Anwendungen z.B. in der Hochenergiephysik, der Medizin und der Umweltforschung [1].

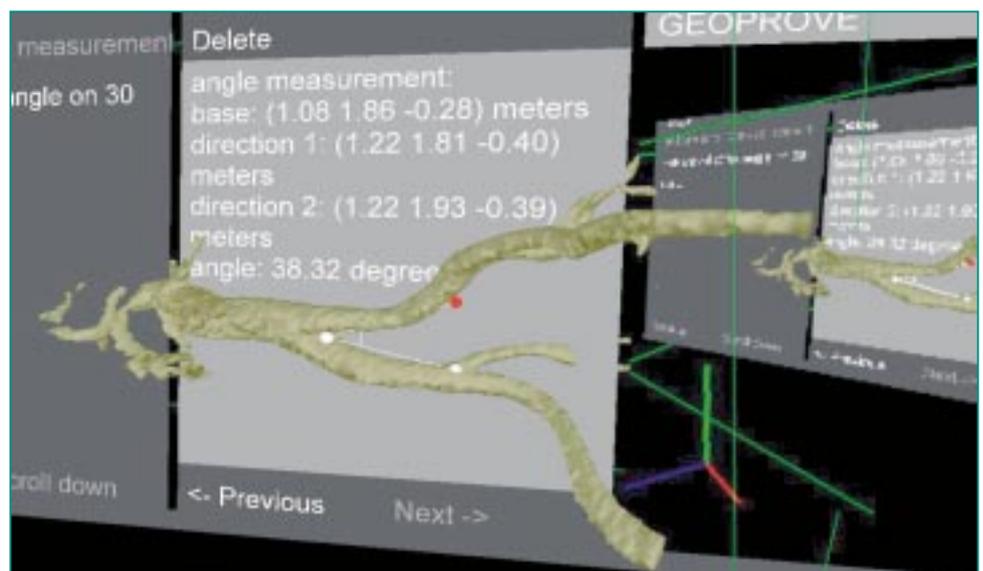
Das von der EU geförderte CrossGrid Konsortium<sup>1)</sup> umfasst 21 Partner aus 11 Ländern und beschäftigt sich mit interaktiven Grid-basierten Anwendungen aus den genannten Bereichen [2]. Die Anwendungen sind dadurch gekennzeichnet, dass eine Person in einer Verarbeitungsschleife permanent mit dem verteilten System wechselwirkt: Charakteristisch dabei ist, dass eine schnelle Antwort vom Computersystem als Reaktion auf eine datenintensive Eingabe garantiert sein muss. Beispiele für diesen Typ von Anwendungen sind:

- Virtuelle chirurgische Modelle zur Operationsplanung und Simulation,
- Entscheidungsunterstützungssysteme für Flutkatastrophen,
- Verteilte Datenanalyse in der Hochenergiephysik,

- Kombination von Luftverschmutzungsmodellen mit Wettervorhersagen.

Exemplarisch soll im Folgenden ein Beispiel aus dem medizinischen Bereich angeführt werden: Gefäßerkrankungen stellen ein großes Gesundheitsproblem dar und die Behandlung erfordert oftmals einen chirurgischen Eingriff. Dieser beinhaltet das Legen eines Bypass um die betroffenen Arterien zur Verbesserung des Blutflusses. Die Operationsplanung stützt sich dabei auf 3D-Aufnahmen, die üblicherweise mit Computertomographen gewonnen werden. Die hier zum Einsatz kommenden Informationssysteme profitieren erheblich durch Kombination mit Werkzeugen zur Simulation und Visualisierung. Diese Herangehensweise erfordert aber im Augenblick der Behandlung schnellen Zugriff auf große Rechnerleistung, die aus

Kostengründen in der Regel vor Ort nicht installiert ist. Grid-Computing bietet hier einen eleganten Ansatz zur Problemlösung, indem verteilte Rechenressourcen dynamisch zugeschaltet werden können. Im CrossGrid-Projekt wurde ein solches prototypisches Grid-System zur Unterstützung virtueller Modelle in der Operationsplanung und Simulation entwickelt. Das System unterstützt den Chirurgen beim Entscheidungsprozess über verschiedene Operationsmethoden, indem der Blutfluss simuliert und vorhergesagt wird auf der Basis von vermessener Arteriengeometrie sowie von gespeichertem Expertenwissen über die verschiedenen Behandlungsformen. Interaktive 3D-Visualisierung wird eingesetzt, um den Chirurgen in klarer und intuitiver Weise zu leiten (siehe Abb. 1). Verschiedene Operationspläne können so vor der eigentlichen Operation im



**Abb. 1:** Aus computertomografischen Daten gewonnenes 3D-Modell einer Arterie. In einer Virtual Reality-Umgebung kann der Chirurg die Operation planen (Messung der Länge eines Bypass) und durch Simulation den Blutfluss vorhersagen.

<sup>1)</sup> Vertrag IST-2001-32243, Laufzeit 2002-2005, Fördersumme ca. 2.3 Millionen Euro.

Hinblick auf ihre Effizienz und Machbarkeit interaktiv überprüft werden.

Der transparente und sichere Zugriff auf verteilte Ressourcen wird dabei durch einen Resourcebroker organisiert. Die CrossGrid-Infrastruktur erlaubt in diesem Zusammenhang die nahtlose Einbindung von medizinischen Scannern und Daten, der Visualisierungsumgebung sowie die Vermittlung und die Bereitstellung der zur Simulation des Blutflusses nötigen Ressourcen in örtlich verschiedenen administrativen Rechenzentrumsdomänen.

### Grid-Entwicklungs- umgebung

Nicht nur für die Anwender, sondern auch für Programmentwickler ist einer der kritischen Aspekte des Grid die inhärente Komplexität. Die Entwicklung, die Fehlersuche und die Optimierung von verteilten Anwendungen auf dem Grid ist schwierig und zeitaufwändig aufgrund der sich dynamisch ändernden heterogenen Infrastruktur aus Rechnern, Speichern und Netzen. Um die genannten Anwendungsfelder effizient auf einem Grid implementieren zu können, werden daher nicht nur neue Werkzeuge zur syntaktischen Überprüfung paralleler Programme, sondern auch Methoden zur Leistungsvorhersage und Leistungsbewertung benötigt. Ferner werden Hilfsmittel gebraucht zur Ermittlung der Anwendungsleistung sowie zum interaktiven Management der zur

Durchführung der Aufgaben benötigten Ressourcen im Grid.

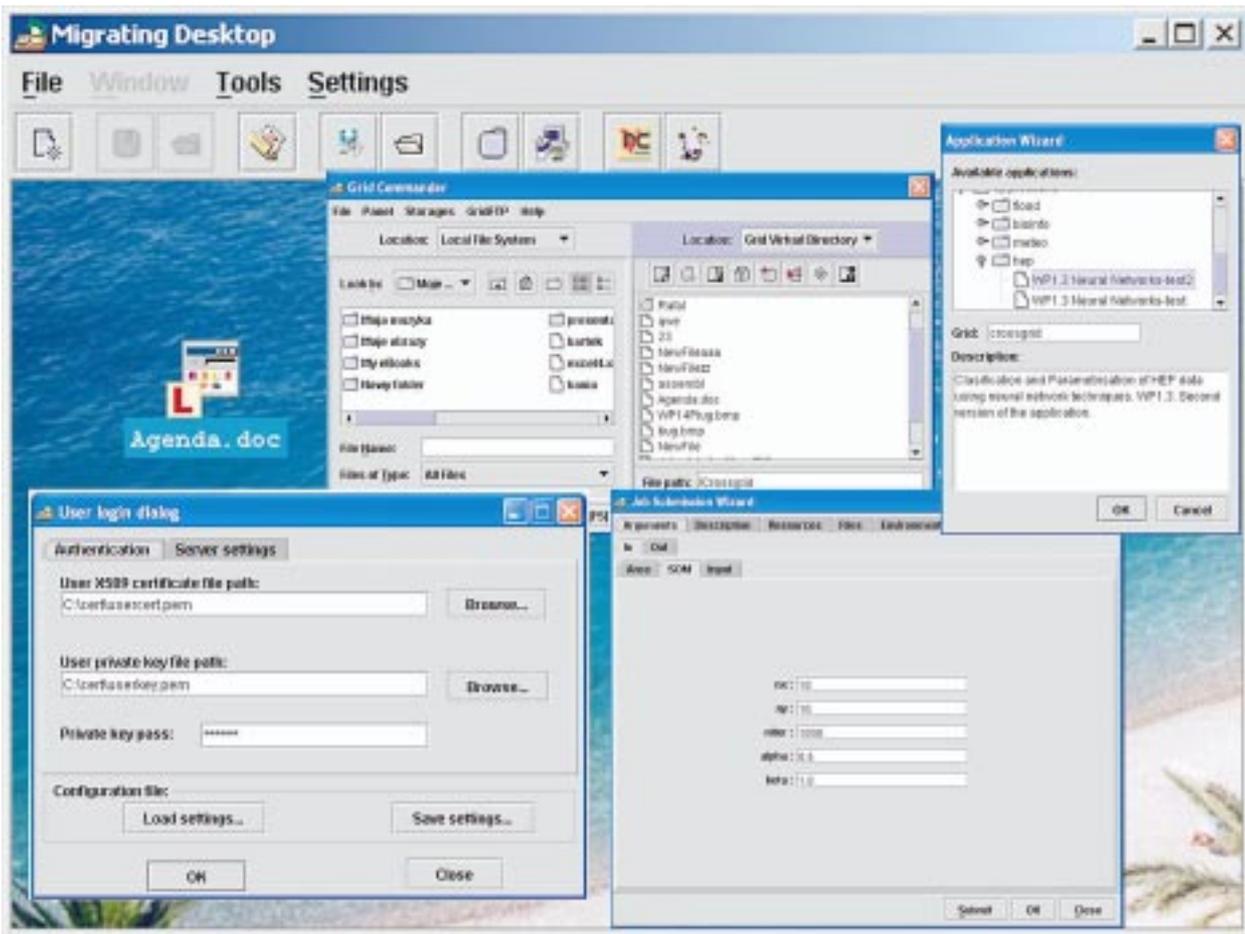
Es ist ein weiteres Ziel des CrossGrid-Projekts, solche Monitoring-Werkzeuge zu entwickeln und in die Grid-Umgebung zu integrieren, um damit die Erstellung und die Optimierung von verteilten, parallel auf mehreren Rechnern ausgeführten Programmen zu erleichtern. Um solche Parallelapplikationen bei reproduzierbaren Ergebnissen portabel zu gestalten, wurde ein spezielles Verifikationstool für MPI-Programme erstellt (MARMOT). Die automatische Überprüfung, ob Benutzeranwendungen mit dem MPI<sup>2)</sup>-Standard übereinstimmen, verringert die Fehleranfälligkeit schon im Vorfeld der Implementierung. Es werden dabei Programmiermodelle unterstützt, die sowohl rechenzeitintensiv sind als auch datenintensiv. Um das Grid auch für im Hinblick auf Ressourcenanforderungen anspruchsvolle interaktive Anwendungen tauglich zu machen, benötigen die Entwickler darüber hinaus Werkzeuge, um die Leistung ihrer Programme überwachen und analysieren zu können. In diesem Bereich wurden entsprechende Hilfsmittel entwickelt zum Aufspüren von Engpässen, sowie entsprechende Benchmark-Programme entworfen, die sensitiv im Hinblick auf Datentransfer, Synchronisation und Latenzzeit sind. Zusammen mit den Monitoring-Tools kann ein verteiltes System so mit nahezu beliebiger Granularität (Ort und Zeit) bis auf den Prozess-Level untersucht und optimiert werden.

### Grid-Portale

Es ist eine der Grid-Visionen, dass Benutzer ihre Anwendungen auf dem Grid in einer einfachen und transparenten Weise ausführen können, ohne die genauen Details der Grid-Infrastruktur zu kennen. Internetportale mit dynamisch zusammengestellten Informationen und Verweisen helfen, den Zugang zur Grid-Welt zu vereinfachen: Der Zugriff über Internetportale sowie die Plattform- und Ortsunabhängigkeit spielen eine entscheidende Rolle für die Akzeptanz der neuen Technologie seitens der Endanwender. Entsprechend einer Umfrage über die Anforderungen von Gridnutzern im Rahmen des ENACTS-Projekts [3] würde eine größere Zahl von Anwendern Grid-Services in Anspruch nehmen, wenn der Zugriff einfacher wäre. Portaltechnologien, die etwa im Grid Portal Development Toolkit (GPDK) propagiert werden, zielen in diese Richtung [4].

CrossGrid entwickelt aus diesem Grund benutzerfreundliche Portale und mobile Arbeitsumgebungen. Eine der Portaltechniken, die in CrossGrid implementiert wird, ist „Roaming“, d.h. der mobile persönliche Zugriff auf Ressourcen. Ein Grid-Service (Portal-Server) führt Buch über alle Interaktionen des Benutzers, der sich auf seinem Endgerät (Portal-Client) zunächst über eine Eingabemaske authentifizieren muss. Die Benutzerschnittstelle ist dabei gängigen Desktop-Oberflächen nachempfunden.

<sup>2)</sup> MPI = Message Passing Interface



**Abb. 2: Mobile Benutzerschnittstelle für CrossGrid-Applikationen.** Durch enge Anlehnung an das Erscheinungsbild gängiger Desktop-Oberflächen finden sich auch Neueinsteiger schnell zurecht. Zu sehen sind neben Verknüpfungen zu Dateien im Grid (Agenda.doc, oben links) auch Fenster zur Anmeldung (User Login Dialog, unten links) und Job-Abgabe (Job Submission Wizard, unten rechts) sowie zur Navigation (Grid Commander, oben Mitte) und Erstellung von Benutzerschnittstellen (Application Wizard, oben rechts).

den (siehe Abb. 2), um die Einarbeitungszeit zu verkürzen. Es ist nicht nur möglich, Grid-Ressourcen interaktiv zu verwalten (Grid Commander), sondern es können Dateien im Grid auch so verarbeitet werden, als ob sie lokal gespeichert wären (Virtual Directory Browser). Für Grid-Applikationen können durch einen sog. „Application Wizard“ auf einfache Art und Weise grafische Schnittstellen zur Eingabe und

Ausgabe von Informationen gebaut werden, die Applikation kann durch einen „Job Submission Wizard“ gestartet werden. Der Zustand einer Sitzung ist persistent, d.h. ein Benutzer kann sich jederzeit abmelden und seine Arbeit an einem anderen Ort zu einer späteren Zeit auf einer anderen Plattform wieder aufnehmen, wobei die Oberfläche identisch restauriert wird.

### Internationales Testbett

Organisatorische Probleme sowie Fragen der Leistung und Sicherheit im Netzwerk können nur in einem realen Umfeld gelöst werden. Im CrossGrid Projekt wird daher eine internationale Testumgebung aufgebaut und gepflegt, um die Integration und Optimierung der verteilten Anwendungen und Werkzeuge zu ermöglichen. Die Zielsetzungen

der Testbettorganisation im Projekt sind:

- Bereitstellung verteilter Ressourcen,
- Integration nationaler Grids,
- Integration der Softwarepakete,
- Koordination der Software-Release Versionen,
- Sicherstellung der Verträglichkeit mit Grid-Standards.

Das Testbett wird an 10 unterschiedlichen europäischen Standorten verteilt betrieben. Ein wesentlicher Beitrag wird in diesem Bereich vom Forschungszentrum Karlsruhe geleistet: Neben der Bereitstellung von Ressourcen im Grid (Rechnerelemente, Speicherelemente) wird hier auch der

zentrale Server zur Verwaltung der Software-Releases betrieben. In regelmäßigen Abständen wird mittels einer speziell entwickelten Infrastruktur aus den Eingaben der Partner in automatisierter Weise ein Softwarepaket gebaut und in Testprozeduren auf Konsistenz und Fehlerfreiheit überprüft. Auf diese Weise ist sichergestellt, dass ein Software-Release eine gewisse Robustheit und Reife besitzt, bevor er den Server verlässt und in der Kollaboration sowie auf dem Testbett verteilt wird.

### Zusammenfassung

Das CrossGrid-Projekt beschäftigt sich mit interaktiven Anwendungen im Grid. Interaktivität im Grid ist ein durchaus neues,

schwieriges und spannendes Thema, das hier erstmals erfolgreich aufgegriffen wird. Neben prototypischen Anwendungen aus dem Bereich der Hochenergiephysik, der Medizin und der Umweltforschung werden auch Spezialwerkzeuge bereitgestellt zur Programmentwicklung, sowie Portaltechniken, die den Umgang mit dem Grid erleichtern und damit die Akzeptanz der neuen Technik verbessern. Das Forschungszentrum Karlsruhe erbringt wesentliche Beiträge im Bereich der Grid-Infrastruktur und der Grid-Services.

### Literatur

[1] M. Kunze, H.Marten, *Grid-Computing und e-Science.*, Nachrichten 4/2002

[2] *The CrossGrid Collaboration*, IST-2001-32243, <http://www.crossgrid.org/>

[3] <http://www.enacts.org/>

[4] <http://doesciencegrid.org/projects/GPDK/>