

Abschlußbericht

zum DFN-Forschungsprojekt GRIKSL:

Development of Grid Based Simulation and Visualization Techniques

Auftragsnummer:	TK 602 – AN 200	
Beteiligte Einrichtungen:	Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Golm (AEI) Konrad-Zuse-Zentrum für Informationstechnik, Berlin (ZIB)	
Autoren:	Ralf Kähler (ZIB)	kaehler@zib.de
	Thomas Radke (AEI)	tradke@aei.mpg.de
Projektverantwortliche:	Hans-Christian Hege (ZIB)	hege@zib.de
	Ed Seidel (AEI)	eseidel@aei.mpg.de
Projektzeitraum:	1. April 2002 – 31. März 2004	

1 Übersicht

1.1 Aufgabestellung

Ziel des GRIKSL-Projekts war es, adäquate Grid-Werkzeuge und Technologien für existierende Simulationsprogramme aus dem Bereich des High-Performance-Computings zu entwickeln, um deren Anwendern die Nutzung bereits vorhandener, jedoch innerhalb eines virtuellen *Computing Grids* geografisch verteilter Rechenressourcen zu erschließen. Der inhaltliche Schwerpunkt von GRIKSL lag dabei im Entwurf und der Implementierung von gridbasierten, standard-orientierten Methoden für die effiziente Speicherung und die transparente Verwaltung von verteilten, hochaufgelösten hierarchischen Datensätzen sowie deren Visualisierung mittels Datenfernzugriffstechniken.

1.2 Voraussetzungen

Besonderer Wert bei der Realisierung von GRIKSL wurde auf den engen Kontakt zu bestehenden Wissenschaftlergruppen mit realen Anwenderprogrammen gelegt, insbesondere zu den Physikern aus der von Ed Seidel geleiteten Gruppe “Numerische Relativität” am AEI mit ihrer parallelen Simulationsumgebung Cactus. Dies erlaubte, die praktische Relevanz und Verwendbarkeit der entwickelten Grid-Softwarekomponenten zu erproben und sicherzustellen.

1.3 Planung und Ablauf

Zur Bearbeitung der Gesamt-Aufgabenstellung an den beteiligten Einrichtungen AEI und ZIB wurde das Projekt organisatorisch in fünf Arbeitspakete WP-I bis WP-V unterteilt (WP steht für Work Package):

- WP-I:** Dynamische Techniken und Werkzeuge des Grid-Computings
- WP-II:** Datenbeschreibung und -management
- WP-III:** Effizientes I/O in Grid-Umgebungen
- WP-IV:** Grid-fähige Visualisierungstechniken und -werkzeuge
- WP-V:** Verwaltung, Koordination, Support

Diese Arbeitspakete wurden von den aus Projektmitteln finanzierten drei wissenschaftlichen Mitarbeitern (BAT-IIa) bearbeitet, welche jeweils zur Hälfte am AEI und am ZIB angesiedelt waren. Verwaltungstechnisch fungierte das AEI als Hauptauftragnehmer für das GRIKSL-Projekt gegenüber dem DFN-Verein, das ZIB schloß einen Untervertrag mit dem AEI ab.

Die technische Infrastruktur, in die das GRIKSL-Projekt eingebettet war, bildete auf Hardwareebene ein am AEI installierter Hochleistungsrechner (Linux-Cluster mit 128 Knoten) als Produktionsmaschine sowie weitere europa- und weltweit verteilte Rechenressourcen als Testumgebung (GRIDLAB-Testbett mit bis zu 10 beteiligten Institutionen, Allokationen bei Supercomputing Zentren in den USA); für die Visualisierung standen moderne Grafikmaschinen am AEI und am ZIB zur Verfügung.

1.4 Wissenschaftlicher und technischer Stand

Als existierende Softwarekomponenten wurde von seiten des ZIB das wissenschaftliche Visualisierungspaket AMIRA (<http://amira.zib.de>) eingebracht; CACTUS (<http://www.cactuscode.org>), ein am AEI in internationaler Zusammenarbeit entwickeltes modulares Programmpaket zur numerischen Simulation von Problemen aus Einsteins allgemeiner Relativitätstheorie sowie verwandten astrophysikalischen Forschungsrichtungen, diente als Beispielapplikation aus einem realen wissenschaftlichen Anwendungsfeld.

Als wissenschaftliche Ausgangsbasis für GRIKSL wurde sehr stark auf Erfahrungen aus dem ebenfalls vom DFN-Verein geförderten Vorgängerprojekt TIKSL *“Tele-Immersion – Kollision Schwarzer Löcher”* zurückgegriffen. Dieses im Zeitraum von 1999-2001 laufende Projekt befaßte sich mit der Entwicklung von grundlegenden Netzwerkfähigkeiten für große, in Netzwerkumgebungen auf Supercomputern laufende Simulationen aus dem Bereich der numerischen Relativität. Wichtige Ergebnisse aus TIKSL flossen dabei direkt in die Forschungsarbeiten von GRIKSL ein, so die Erweiterung der standardisierten HDF5-I/O-Bibliothek um effiziente Möglichkeiten zum entfernten Datenzugriff unter Ausnutzung allgemeiner und sicherer Grid-I/O-Techniken, der Entwurf und die Entwicklung von Kontroll- und Steuermechanismen, welche Anwendern die direkte Interaktion mit laufenden verteilten Simulationen ermöglichen, sowie die Erweiterung vorhandener Visualisierungswerkzeuge um geeignete Techniken, die sowohl Echtzeitdaten von der laufenden Simulation für einen entfernten Beobachter visualisieren als auch Rohdaten aus entfernten Dateien darstellen können.

1.5 Zusammenarbeit mit anderen Einrichtungen und Projekten

Weiter gepflegt wurde die bereits bestehende langjährige Zusammenarbeit von AEI und ZIB mit Wissenschaftsgruppen aus den USA, so z.B. dem HDF5-Entwicklerteam (<http://hdf.ncsa.uiuc.edu/HDF5>) am National Center for Supercomputing Applications (NCSA) in Urbana-Champaign und dem Globus-Toolkit-Entwicklerteam (<http://www.globus.org>) am Argonne National Laboratory (ANL) unter Leitung von Ian Foster. HDF5 als standardisierte I/O-Bibliothek und Globus als etablierte Grid Middleware Software stellen wichtige Basis-Komponenten der im Rahmen von GRIKSL entstandenden Grid-Lösungen dar.

Während der Laufzeit von GRIKSL entstand eine sehr enge Kooperation mit dem zeitgleich laufenden EU-Projekt GRIDLAB (<http://www.gridlab.org>), welches sich der Entwicklung eines flexiblen, generischen und modularen Grid Application Toolkits (GAT) sowie weiterer Grid Services widmet.

Desweiteren beteiligten sich Mitarbeiter des GRIKSL - Projektes an den Aktivitäten zur Gründung des Global Grid Forums (<http://www.ggf.org>) und leisteten im Rahmen von WP-II wesentliche Beiträge zur Standardisierung von Techniken zu allgemeiner Datenbeschreibung und Datenmanagement in Grid-Umgebungen.

- GRIDLAB-Projekt, <http://www.gridlab.org>
Leibniz-Rechenzentrum München (LRZ), <http://www.lrz-muenchen.de>
Rechenzentrum Garching der Max-Planck-Gesellschaft (RZG), <http://www.rzg.mpg.de>

In enger Abstimmung mit dem GRIDLAB-Projekt sowie dem LRZ und RZG wurde ein europaweites Grid-Computing-Testbett mit verschiedenen Globus-Services etabliert. Besonders hervorgehoben werden soll hier die in Zusammenarbeit zwischen dem LRZ-Systemadministrator Helmut Heller und die im GRIKSL - Projekt erfolgte erstmalige Portierung der Globus-Software und GridFTP-Komponenten auf einen Hitachi-Vektorrechner.

- Globus-Projekt, <http://www.globus.org>

Im Rahmen der Installation des GRIDLAB-Testbetts setzte sich auch die langjährige Kooperation mit dem Globus-Team fort. Mit dem GridFTP-Treiber gehörte GRIKSL zu den Alphatestern von Komponenten des Globus 2.0 Toolkits. Die durch ein von GRIKSL entwickeltes Testprogramm festgestellten Fehler in der Implementierung der GridFTP-Server-Komponente von Globus wurden behoben und sind in der neuesten Globus-Release 2.2 behoben.

Desweiteren wurde bei einem Treffen mit den Globus-Entwicklern über zukünftige Erweiterungen des GridFTP-API diskutiert, die insbesondere für die in GRIKSL geplante Implementierung eines Datenservers relevant sein können.

- HDF5-Entwicklungsteam, <http://hdf.ncsa.uiuc.edu/HDF5>

Zur HDF5-Gruppe bestand reger Kontakt per email zu technischen Fragen bezüglich einer effizienten Implementierung des GridFTP-Treibers, insbesondere der Optimierung von Metadatenzugriffen und des Datencaching bei entfernten Dateizugriffen.

Es wurden ebenso Fragen einer zukünftigen Integration des GridFTP-Treibers in die offizielle HDF5-Release angesprochen, um diesen Treiber für jedermann verfügbar zu machen.

- Cactus Entwicklungsteam, <http://www.cactuscode.org>
International Relativity Group, <http://jean-luc.ncsa.uiuc.edu>

Diese Wissenschaftler, insbesondere die Gravitationsphysiker am AEI, bildeten die wichtigste Anwender-Zielgruppe für die im Rahmen des GRIKSL-Projekts erzielten Forschungsergebnisse.

Zum einen konnte sich GRIKSL durch die Verwendung des Cactus-Codes an konkreten, praxisrelevanten Simulationsproblemen aus dem Bereich des Grid Computings orientieren und sich den jeweiligen Anwendungserfordernissen anpassen. Dies betrifft insbesondere die Verwaltung, Analyse und Visualisierung von entfernten Cactus-Daten.

Andererseits profitierte GRIKSL unmittelbar von den Physikern, indem diese die im Rahmen des Projektes entwickelten Softwaretools testeten und entsprechendes Feedback gaben. Konkret trifft dies auf die Entwicklung des GridFTP-Treibers zu, der im OPENDX-Visualisierungs-Toolkit integriert und auf allen Arbeitsplatzmaschinen am AEI installiert wurde. Diese Lösung wird von den Physikern zunehmend genutzt, um die Ergebnisse aus ihren am LRZ bzw. RZG durchgeführten Cactus-Simulationen zu visualisieren.

2 Detaillierte Darstellung

2.1 Erzielte Forschungsergebnisse

2.1.1 Themenbereich: Dynamische Techniken und Werkzeuge des Grid-Computings

Dieses Arbeitspaket wurde in enger Zusammenarbeit mit GRIDLAB bearbeitet, um die sich aus den Zielstellungen beider Projekte ergebenden Synergieeffekte maximal auszunutzen und fernere Ziele zu erreichen.

Ein erster Meilenstein hierbei war die Planung und Installierung eines Grid-Computing-Testbetts als Grundlage für die Durchführung aller weiteren Vorhaben. Das von GRIKSL initiierte Testbett beinhaltete zunächst nur die am AEI und ZIB verfügbaren Supercomputer und Grafikrechner, wurde dann jedoch in Absprache mit GRIDLAB auf weitere Rechenressourcen in Deutschland (Rechenzentrum Garching, Leibniz-Rechenzentrum München, Fraunhofer-Rechenzentrum) und europaweit (Polen, Tschechien, Ungarn, Italien, Niederlande, Großbritannien) ausgedehnt und mit GRIDLAB gemeinsam genutzt und gewartet.

Dabei flossen in den vom GRIDLAB-Projekt in einem *TestBed Requirements Document* definierten Standard auch spezifische Testbett-Anforderungen von GRIKSL ein, so z.B. die Installation notwendiger Basis-Software wie CACTUS und HDF5 zur Durchführung von Testsimulationen sowie die Aktivierung von GridFtp-Services zum entfernten Datenzugriff.

Die in WP-I durchgeführten Arbeiten zu Entwurf und Implementierung eines dynamischen Grid-Frameworks für CACTUS gingen vollständig in gleichgelagerte Arbeiten von GRIDLAB ein. Als wichtigstes Ergebnis entstand hieraus das *Grid Application Toolkit* GAT (<http://www.gridlab.org/WorkPackages/wp-1/index.html>), ein standardisiertes API zur flexiblen und einheitlichen Interaktion von Grid-Applikationen mit entfernten Grid Services innerhalb einer dynamischen Grid-Umgebung.

Darauf aufsetzend wurde ein CACTUS-spezifisches CGAT entwickelt (<http://www.gridlab.org/WorkPackages/wp-2/index.html>), welches unter Ausnutzung verschiedener Grid Services gridbasierte Funktionen für beliebige CACTUS-Simulationen – wie z.B. die Erkundung und Reservierung freier Grid-Ressourcen, das automatische Starten neuer Jobs auf diesen, die Registrierung eines Jobs bei Grid-Informationsdiensten, die Replizierung seiner Ausgabedateien –

zur Verfügung stellt.

2.1.2 Themenbereich: Datenbeschreibung und -management

Bei der Entwicklung von allgemeinen Datenmodellen und -repräsentationen wurde deutlich, dass eine enge Zusammenarbeit mit Anwendern der verschiedenen Anwendungsbereiche notwendig und eine Akzeptanz nur zu erreichen ist, wenn die diversen, von historisch gewachsenen Datenmodellen und -formaten geprägten Vorstellungen berücksichtigt werden. Dies macht langwierige Abstimmungsprozesse notwendig. Vom GriKSL-Projekt wurde daher eine Lösung in einem breiteren und internationalen Rahmen, dem Global-Grid-Forum (GGF), angestrebt. Entsprechende Initiativen zur Gründung von zwei GGF-Arbeitgruppen wurden gestartet: eine zum Thema Datenbeschreibung 'data format definition language' (DFDL, GGF7, Mrz 2003, Tokio) und 'data models' (DaMo, GGF8, Juni 2003, Seattle).

Das Ziel der DFDL-WG ist die Schaffung einer XML-Beschreibung für Datenformate, die Informationen über Daten auf Dateistrukturerebene enthält. Dadurch sollen existierende Datenformate durch allgemeineren Zugriffsmechanismen handhabbarer werden. Durch die DAMO-WG sollten auf einer höherliegenden Abstraktionsebene existierende Datenmodelle im Grid-Kontext klassifiziert werden. In dem Abstimmungsprozess ließ sich die DAMO-WG jedoch nicht etablieren. Ihre Ziele werden nun, zumindest teilweise, in der Application Metadata Working Group (AMD-WG) verfolgt, die Standards (XML-Schemata) für Anwendungs-Metadaten definiert.

Da die Bemühungen auf internationaler Ebene erst längerfristig zu praktisch nutzbaren Ergebnissen führen, wurde die Implementierung der Datenmodelle im Projekt von diesen Entwicklungen entkoppelt. Es wurde eine schlanke Abstraktionsschicht geschaffen, die es erlaubt, verschiedene existierende Datenmodelle auf das in der Visualisierungssoftware Amira verwendeten Datenmodell abzubilden. Insbesondere wird diese Abstraktion für die im Projekt verwendeten Cactus-Datentypen verwendet. Die implementierte Lösung stellt eine Abstraktion der Verteilung der Daten sowohl über verschiedene Dateien, als auch über verschiedene entfernte Quellen zur Verfügung. Zur Zeit beschränkt sich die Implementierung auf Datenmodelle, die in HDF5 (weit verbreitetes general-purpose-Datenformat für wissenschaftliche Daten mit zugehöriger Bibliothek, s. <http://hdf.ncsa.uiuc.edu/HDF5>) formuliert sind.

Zur Unterstützung des Datenmanagements wurde die Klassenstruktur von Amira erweitert. Die hiermit erreichte Flexibilität in der Unterstützung verschiedener Datenmodelle machte eine Erweiterung von Cactus zur Beschreibung der verwendeten Datenmodelle überflüssig. Das Design von Amira erlaubt damit die Unterstützung verschiedenster Applikationsdaten ohne besondere Vorbereitungen auf Anwendungsseite. Mit der Integration der Grid-Informationen und Datenverwaltungsdienste (u.a. Replika-Management) in Amira und Cactus ist das System komplett.

Im Projekt wurden weiterhin die Arbeiten zum effizienten Zugriff auf entfernte Daten fortgeführt. In Zusammenarbeit mit dem WP-8 (Data Management) des EU-Projektes GridLab wurden die vorhandenen Mechanismen zum Zugriff auf entfernte Daten im HDF5-Format auf den Zugriff auf beliebige regulär strukturierte Binärdaten übertragen. Diese neuen Techniken nutzen ebenfalls die Möglichkeiten des verwendeten Grid-FTP-Servers, insbesondere, eigene Plugins zur Bearbeitung von spezifischen Requests auf dem Server zu installieren. Die Requestsprache ist eine einfache, mächtige Sprache (FALLS) zur Beschreibung von hierarchischen regulären Mustern, die auf die Binärdaten angewandt wird. Das Client-API für diese Zugriffe wurde dem POSIX-API für Datei-Zugriffe nachempfunden, und erlaubt eine einfache Implementierung von entsprechenden Remote-Readern. Amira wurde für einige wesentliche Datenformate um die Unterstützung dieses Remote-Zugriffs erweitert (über den ursprünglichen Projektvorschlag hinausgehend).

2.1.3 Themenbereich: Effizientes I/O in Grid-Umgebungen

Eine wichtige Anforderung an die Ein-/Ausgabe von großen Datenmengen in parallelen Supercomputing-Applikationen ist eine höchstmögliche Effizienz der Speicherung und des anschließenden Wiedereinlesens der erzeugten Datensätze zwecks Visualisierung. In einer üblicherweise heterogenen Grid-Umgebung kommt die Notwendigkeit der Standardisierung erschwerend hinzu: Daten werden auf verschiedensten Plattformen mit unterschiedlichen Betriebssystemen generiert und ggf. auf anderen Maschinen visualisiert.

Ein Schwerpunkt in WP-III war es deshalb, eine I/O-Infrastruktur für CACTUS zu entwickeln, die diesen Anforderungen Rechnung trägt. Zu diesem Zweck wurde die am NCSA entstandene HDF5-Bibliothek ausgewählt, um damit effiziente I/O-Methoden zur parallelen Ein-/Ausgabe von hierarchischen Datensätzen in CACTUS zu implementieren. Die Ausgabe erfolgt dabei standardmäßig gleichzeitig von jedem Prozessor aus in sogenannten chunks (eine Datei je Prozessor). Für herkömmliche Visualisierungstools wurde ein separates Rekombinierungstool entwickelt, welches die einzelnen chunks wiederum zu einer gemeinsamen Datei zusammenfügt. Zugleich wurde dieser HDF5-Rekombinierer in die von GRIKSL benutzten Visualisierungstools AMIRA und OPENDX direkt als Modul integriert.

Desweiteren wurden in WP-III geeignete Methoden für den Zugriff auf entfernt gespeicherte großvolumige hierarchische Cactus-Datensätze entwickelt, welche dann in Visualisierungstools wie AMIRA und OPENDX zur *remote visualization* eingesetzt werden können.

Dazu wurde – in Kooperation mit dem HDF5-Entwicklerteam am NCSA – die Standard-HDF5-Bibliothek um einen sogenannten *Virtual File Driver* erweitert, der das partielle Lesen und Schreiben von HDF5-Dateien auf entfernten GridFTP-Servern erlaubt. Dieser HDF5-I/O-Treiber basiert auf dem Globus-Toolkit und nutzt das von Globus bereitgestellte GridFTP-API zum direkten Zugriff auf entfernte Dateien. Dabei wird implizit die Globus-Sicherheits-Infrastruktur (GSI) zur Nutzeridentifikation und Prüfung von Zugriffsrechten auf im Grid verteilte Daten verwendet (single sign-on). Der Zugriff auf einen einzelnen Datensatz kann selektiv über sogenannte *hyper slab reads* erfolgen, d.h. es werden nur diejenigen Rohdaten aus einem N-dimensionalen physischen Datensatz gelesen, die für die vom Nutzer für die jeweilige Visualisierung (z.B. anhand der gewählten Zoom- bzw. Downsampling-Parameter) erforderlich sind. Für den Datenaustausch selbst besteht zudem die Möglichkeit, über mehrere, gleichzeitig geöffnete Socketverbindungen Daten zwischen Grid-Applikationen parallel zu übertragen.

Insbesondere für das Lesen von Metadaten aus entfernten HDF5-Dateien mit sehr vielen zeitabhängigen Datensätzen erwies sich das vom Globus-GridFTP-Server verwendete Kommunikationsprotokoll als zu ineffizient, da die Vielzahl der hierbei ausgeführten einzelnen low-level I/O-Operationen die Latenzzeiten zum Lesen der eigentlichen Nutzerdaten stark ansteigen ließ. Als Lösung wurde deshalb in Zusammenarbeit mit dem Globus-Entwicklerteam ein eigener Ansatz für einen optimierten GridFtp-Server verfolgt. Dieser basiert auf dem vom Globus-Team als Alphaversion bereitgestellten SFTP-Server, welcher im Rahmen von WP-III durch einen HDF5 Plug-in erweitert wurde, um somit direkt HDF5-Operationen auf Serverseite zu unterstützen.

Die gesamte im Rahmen von WP-III entwickelte HDF5-Funktionalität ist in den jeweiligen Software-Distributionen von CACTUS (HDF5-basierte I/O-Methoden zur parallelen Ausgabe von Simulationsdaten) sowie den Visualisierungspaketen AMIRA und OPENDX (HDF5-Dateneingabemodule für den transparenten Zugriff sowohl auf lokale als auch entfernt gespeicherte hierarchische HDF5-Datensätze) enthalten.

2.1.4 Themenbereich: Grid-fähige Visualisierungstechniken und -werkzeuge

Im Rahmen des WP-IV wurde das am ZIB entwickelte 3D-Visualisierungspaket AMIRA zunächst auf Grundlage des oben erwähnten Grid-FTP-Treibers mit einem Reader für den Zugriff auf entfernte uniforme Datensätze versehen.

Ein für große und entfernte Datensätze umgestaltetes AMIRA-I/O-Interface lieferte die Basis für einen transparenten Netzzugriff der Visualisierungsmodule. Dieser Mechanismus befreite von der bisherigen Limitierung, den Reader nur über spezielle Schnittstellen angesteuert zu können. Dieser Zugriffsmechanismus diente als Basis zur *progressiven* Visualisierung großer und/oder hierarchischer Datensätze.

Mithilfe der übergebenen Metadaten wird hierzu eine hierarchische Baumstruktur (in diesem Fall ein Octree) generiert, deren Wurzelknoten das gesamte Datenvolumen in grober Auflösung überdeckt. Jeder innerer Knoten wird rekursiv in jeweils acht Unterbereiche höherer Auflösung unterteilt, bis die ursprüngliche Auflösung der Daten erreicht ist.

Die Liste der zu lesenden Knoten wird vom Reader in einem eigenen Thread abgearbeitet, so dass die parallel operierenden Visualisierungsklassen nicht blockiert werden. Sobald die Daten eines Knotens bereitgestellt werden, wird die Verwaltungsklasse davon in Kenntnis gesetzt, dass neue, bzw. höher aufgelöste Daten vorliegen.

Die im Rahmen von WP-IV entwickelten Visualisierungsmodule operieren direkt auf diesen hierarchischen Daten, um so die Darstellung der Daten zu beschleunigen, bzw. im Fall sehr großer Daten, überhaupt erst zu ermöglichen.

Neben Verfahren zur Darstellung von Daten in ebenen 2D-Schnitten, wurde ein Algorithmus zur Isoflächendarstellung von skalaren Feldern auf strukturierten AMR (Adaptive Mesh Refinement) Gittern entwickelt. Das Standardverfahren für komforme Hexaedergitter, der sogenannte 'Marching Cubes'-Algorithmus, liefert bei Anwendung auf lokal verfeinerte Hexaedergitter Artefakte an den Grenzen zwischen Bereichen verschiedener Auflösungsstufe. Unser Algorithmus vermeidet dies, und ist insbesondere anwendbar auf Gitter, bei denen sich benachbarte Zellen um beliebig viele Hierarchiestufen unterscheiden können, oder bei denen die Datenwerte an den hängenden Knoten beliebige Werte annehmen können.

Ferner wurden hardware- und software-basierte Algorithmen zum direkten Volume Rendering von AMR-Daten entwickelt. Hierzu wird das Datenvolumen in getrennte Bereiche konstanter Auflösungsstufen unterteilt, welche in einer blickpunktabhängigen Reihenfolge bearbeitet werden.

Desweiteren wurde eine Datenklasse für vektorwertige Felder auf adaptiv verfeinerten Gittern entworfen und implementiert. Durch Einhaltung der Schnittstellen für den Zugriff auf Vektorfelder in Amira können bereits vorhandene leistungsfähige Visualisierungsverfahren, wie zum Beispiel beleuchtete Stromlinien oder LIC-Flächen, auch für diesen Gittertyp genutzt werden.

Ferner wurde ein Verfahren entwickelt, das die zeitliche Interpolation von strukturierten AMR-Daten ermöglicht. Derartige Interpolationsverfahren werden in der Visualisierung unter anderem dann benötigt, wenn sich der zeitliche Abstand der von der Simulationsumgebung gespeicherten Zeitschritte für eine glatte Animation der Daten als zu groß erweist. Im Falle entfernter Daten ergibt sich ferner die Möglichkeit, die Zahl der zu übertragenden Zeitschritte deutlich zu reduzieren und die nicht übertragenen Zeitschritte während der Visualisierungsphase 'on-the-fly' zu approximieren.

Weiterhin wurde in Amira ein Modul zur einfachen Integration der Grid-Sicherheitsdienste (GSI) implementiert. Dieses Modul erlaubt dem Nutzer die Verwaltung und Initialisierung seiner Zertifikate, ohne dazu die Umgebung verlassen zu müssen. GSI wird von allen im Projekt verwendeten Diensten (GridFTP, Erweiterungen des GridFTP, Replica-Management, GIIS) genutzt.

Alle hier beschriebenen Visualisierungsroutinen wurden an die Bedürfnisse der am Albert-Einstein-Institut arbeitenden Wissenschaftler angepasst, was insbesondere Änderungen in der verwendeten I/O-Schicht zum Zugriff auf die im HDF5-Format abgespeicherten AMR-Daten erforderte.

2.1.5 Verwaltung, Koordination, Support

Die in diesem Arbeitspaket bearbeiteten Aufgaben umfassten die Projektkoordination, interne Verwaltungsdienste (Einrichtung eines CVS-Servers) sowie die öffentliche Präsentation von GRIKSL im Internet auf einem eigenen Webserver (<http://www.griksl.org>). Dieser Server bedient Webseiten mit Informationen über die Zielstellung des GRIKSL-Projekts, erzielte Ergebnisse sowie eine Liste der gegebenen Präsentationen und Vorträge sowie der erfolgten Veröffentlichungen. Der Inhalt dieser Webseiten wird über die Laufzeit von GRIKSL hinaus weiter gepflegt.

2.2 Verwertbarkeit

Für OPENDX stehen die HDF5-Erweiterungen als separate Importmodule innerhalb eines *OpenDXutils* Pakets zur Verfügung. Der zugehörige Quellcode ist frei erhältlich; Dokumentation dazu ist unter der Webseite <http://www.cactuscode.org/VizTools/OpenDX.html> abrufbar.

Eine OPENDX-Installation mit integriertem CARPET-Importmodul ist auf allen Arbeitsplatzmaschinen am AEI eingerichtet und von den Wissenschaftlern dort direkt zur Visualisierung ihrer CACTUS-Simulationsdaten nutzbar. Dokumentation und Quelltext dieses SFTPD-HDF5-Servers sind unter der URL <http://www.cactuscode.org/VizTools/SFTPD-HDF5.html> erhältlich.

Eine Dokumentation zum GridFTP-Treibers ist auf der Webseite <http://www.cactuscode.org/VizTools/GridFtpVFD-HDF5.html> abrufbar, der zugehörige Quellcode ist als *open source* frei verfügbar.

Ferner wurden alle im Kapitel 2.1.4 beschriebenen Visualisierungsroutinen in die Visualisierungssoftware AMIRA integriert und werden inzwischen von zahlreichen Wissenschaftlern nicht nur an AEI und ZIB verwendet.

2.3 Fortschritt auf dem Gebiet bei anderen Stellen

Während der Laufzeit des Projektes ist das Interesse an allgemeinen Datenmodellen und -repräsentationen innerhalb des Global-Grid-Forums stark gewachsen. Wie bereits in Kapitel 2.1.2 erwähnt, wurden dort in den vergangenen Jahren entsprechende Arbeitsgruppen mit dem Ziel der Beschreibung binärer Datenformate gegründet.

Insbesondere geht es dabei um die Schaffung einer XML-Beschreibung für Datenformate, die Informationen über Daten auf Dateiestructurebene enthält. Ferner sollen existierende Datenformate durch allgemeinere Zugriffsmechanismen handhabbarer werden. In diesem Rahmen wurden ferner verschiedene konkurrierende Datenmodelle miteinander verglichen.

An diesen Diskussionen beteiligten wir uns weiterhin aktiv und erwarten auch für die Zukunft, dass hieraus Prozesse erwachsen, welche die oben genannten Probleme beseitigen und an deren Ende schließlich eine weithin akzeptierte, für viele Anwendungsbereiche attraktive Lösung steht.

2.4 Erfolgte und geplante Veröffentlichungen

2.4.1 Präsentationen und Vorträge

- H.-C. Hege:
Visual Data Analysis: Applications in Biomedicine
PARA'02 6th Int. Conf. on Applied Parallel Computing (invited talk)
Espoo, Finland; 18. Juni 2002
- H.-C. Hege (zusammen mit H. Schirmacher und D. Stalling):
Visualization and Supercomputing: From Atoms to Galaxies
ISC2002 17th International Supercomputer Conference (invited talk)
Heidelberg; 22. Juni 2002
- H.-C. Hege:
Virtual Laboratories for Science, Engineering and Medicine
Japan Korea Computer Graphics Conference 2002 (JKCGC2002) (invited talk)
Kanazawa, Japan; 23. August 2002
- H.-C. Hege:
Virtual Laboratories for Physical Sciences
Research Institute for Information Science and Engineering, Hiroshima University
Hiroshima, Japan; 27. August 2002
- H.-C. Hege:
Virtual Laboratories in Geo-Sciences
Earth Simulator Center (ESC)
Yokohama, Japan; 30. August 2002
- H.-C. Hege:
Data Visualization in Science: from Atoms to Galaxy Clusters
IAMG 2002 Annual Conference of the International Association for Mathematical Geology
(keynote talk)
Berlin; 17. September 2002
- R. Kähler, A. Merzky, T. Radke:
Demonstration erster in GRIKSL erreichter Ergebnisse mittels verschiedener Computersimulationen und Visualisierung von Simulationsdaten auf entfernten Ressourcen
iGrid 2002, 23.-26. September 2002, Amsterdam.
<http://www.igrid2002.org/>
- Thomas Radke:
Remote Visualization with OpenDX
Präsentation zum *Numerical Relativity Group Meeting*
AEI; 30. Oktober 2002
<http://www.griksl.org/Pics/OpenDX.ps.gz>
- H.-C. Hege (mit 3D-Demo durch Andre Merzky):
Supercomputing und Visualisierung Einweihung des HLRN-Rechners, Hannover, Deutscher Pavillon der Expo 2000 Hannover; 13. November 2002

- Ed Seidel, Denis Pollney, Thomas Radke:
Remote Visualization Demonstrations with OpenDX
SuperComputing 2002 Conference in Baltimore, USA; 16.-22. November 2002
<http://sc2002.aei.mpg.de/demos/RemoteViz.html>
- Thomas Radke:
The GriKSL Project - An Overview
Präsentation zum *Gravitational Wave / Numerical Relativity Group Meeting*
AEI; 2. Dezember 2002
http://www.griksl.org/Pics/GriKSL_Overview.ps.gz
- Ralf Kähler, Andre Merzky:
Remote hierarchical Visualization with Amira
Live-Präsentation auf dem *Numerical Relativity Group Seminar*
AEI; 19. Dezember 2002
- H.-C. Hege:
Visualization in Engineering: Methods and Applications
Summer School ‘Jacques Louis Lions’ – Multidisciplinary Methods for Analysis, Optimization and Control of Complex Systems, MACSI-net
Montecatini, Italien; 20. März 2003
- E. Seidel, A. Merzky, R. Kähler, B. Ullmer
Live-Präsentation von GRIKSL im Rahmen der DFN-Future-Talk-Veranstaltung: Kollision schwarzer Löcher - Grid-basierte Simulationen und Visualisierungen
CeBIT 2003, Hannover
<http://www.futureparc.cebit.de/>.
- H.-C. Hege
Scientific Visualization: from Atoms to Galaxies
Tagung Géométrie et Visualisation
Maubeuge, Frankreich; 17. Juni 2003
- H.-C. Hege
Effective 3D Visualization of Complex Data
2003 Computer Graphics and Visualization Festa (Keynote Lecture)
Nikkei Science, Tokyo, Japan; 4. Oktober 2003
- H.-C. Hege
Visual Analysis of Fluid Flow and Engineering Data
EU-Projektplanungstreffen ‘Virtual Basin’
Amsterdam; 24.11.2004
- Thomas Radke: “The Visualization Package Chombo Viz”, “The OpenDX Toolkit”, “Remote HDF5 via GridFTP”
Ralf Kähler: “The 3D-Visualization System Amira”, “Visualization of Adaptive Mesh Refinement Data”, “Remote Visualization in Grid Environments”
Präsentationen auf dem *Numerical Relativity Mesh Refinement Visualization Meeting*
26. bis 28. Februar 2004, Universität Innsbruck
<http://astro.uibk.ac.at/visworkshop2004/index.html>

- R. Kähler, A. Merzky, A. Hutanu:
Live-Präsentation von GRIKSL im Rahmen der DFN-Future-Talk-Veranstaltung: GRID-basierte Visualisierungen großer medizinischer Datensätze
CeBIT 2004, Hannover
<http://www.futureparc.cebit.de/>.

2.4.2 Veröffentlichungen

- H.-C. Hege, A. Merzky:
GriKSL – Immersive Überwachung und Steuerung von Simulationen auf entfernten Supercomputern.
DFN-Mitteilungen, Heft 59
<http://www.dfn.de/uploaded/heft59.pdf>
- H.-C. Hege, R. Kähler, E. Seidel, T. Radke:
GriKSL zeigt Live-Demos auf der iGrid2002.
DFN-Mitteilungen, Heft 60, November 2002
<http://www.dfn.de/uploaded/heft60.pdf>
- R. W. Moore, A. Merzky:
Persistent Archives Basic Components.
Technical Report (GGF Recommendation Draft),
Global Grid Forum, 2003
http://www.ggf.org/documents/Drafts/default_b.htm
- Thomas Radke:
GriKSL – Simulating Black Hole Collisions on the Grid
Informationsblatt als Handout für GRIKSL-Präsentationen und Konferenzen
<http://www.griksl.org/flyer.html>
- A. Hutanu, A. Merzky, R. Kaehler, H.-C. Hege, B. Ullmer, T. Radke, E. Seidel
Progressive Retrieval and Hierarchical Visualization of Large Remote Data
Proceedings of the 2003 Workshop on Adaptive Grid Middleware, p. 60-72
<http://www.zib.de/PaperWeb/abstracts/ZR-03-40/>
- R. Kähler, H.-C. Hege,
Visualization of Time-Dependent Adaptive Mesh Refinement Data
ZIB-Preprint 03-16, 2003
<http://www.zib.de/PaperWeb/abstracts/ZR-03-16/>
- S. Prohaska, A. Hutanu, R. Kähler, H.-C. Hege,
Interactive Exploration of Large, Remote Micro-CT Scans
erscheint in den Proceedings der IEEE Visualization 2004 Konferenz,
10-15 Oktober, Austin (USA)
<http://vis.computer.org/vis2004/>

3 Erfolgskontrollbericht

3.1 Wissenschaftlich-Technische Ergebnisse

Im GRIKSL-Projekt wurden erfolgreich Grid-Werkzeuge und Technologien für existierende Simulationsprogramme aus dem Bereich des High-Performance-Computings entwickelt.

Insbesondere wurden effiziente Methoden für die Speicherung und den Zugriff entfernter hochaufgelöster, sowohl uniformer als auch hierarchischer Daten, entworfen und implementiert. Für eine detaillierte Darstellung der erzielten wissenschaftlichen Ergebnisse sei auf die Unterkapitel 2.1.1 bis 2.1.4 verwiesen.

3.2 Erfindungen und Schutzrechtsanmeldungen

Zur Zeit sind keine Schutzrechtsanmeldungen von im Rahmen des GRIKSL-Projektes entwickelter Algorithmen geplant.

3.3 Wirtschaftliche Erfolgsaussichten

In dem Projekt wurde eine Themenstellung verfolgt, die auch von wirtschaftlicher Relevanz ist. Die tatsächliche Nutzbarkeit der erzielten Resultate wird derzeit von der Indeed - Visual Concepts GmbH (<http://www.indeed3d.com/>) geprüft. Die Indeed - Visual Concepts GmbH ist eine – mit Unterstützung des BMBF – aus dem ZIB hervorgegangene Spinoff-Firma.

3.4 Wissenschaftliche und technische Erfolgsaussichten

Die im GRIKSL-Projekt geschaffenen Lösungen werden schon jetzt von einem großen Anwenderkreis zur effizienten Visualisierung entfernter Daten genutzt. Insbesondere aufgrund der Offenlegung von Teilen des im Projektes entwickeltem Quellcodes, erwarten wir, dass sich der Nutzerkreis in Zukunft weiter vergrößern wird.

3.5 Wissenschaftliche und wirtschaftliche Anschlussfähigkeit

Eine weiterführende Phase des Projektes ist zur Zeit nicht geplant.

3.6 Arbeiten, die zu keiner Lösung führten

Wie bereits in Kapitel 2.1.2 angedeutet, stellte sich die Entwicklung eines umfassenden und allgemeinen Modells zur Beschreibung wissenschaftlicher Daten als für den zeitlichen Rahmen des Projektes zu langfristige heraus, und konnte daher nur in Teilbereichen realisiert werden.

An den, unter anderem von Mitgliedern des GRIKSL- Projekt ins Leben gerufenen Arbeitsgruppen innerhalb des Global-Grid-Forums (GGF) beteiligten wir uns weiterhin aktiv und erwarten in der näheren Zukunft Ergebnisse, welche die oben genannten Probleme beseitigen und die schließlich eine weithin akzeptierte, für viele Anwendungsbereiche attraktive Lösung bieten werden.

3.7 Präsentationsmöglichkeiten für mögliche Nutzer

Wie in Kapitel 2.2 beschrieben, wurden die im Projekt entwickelten Verfahren auf zahlreichen internationalen Konferenzen potentiellen Anwenderkreisen vorgestellt.

3.8 Einhaltung der Kosten-, und Zeitplanung

Der Budget- und Zeitplan konnte eingehalten werden.