

Softwareentwicklung in der Wissenschaft

Prof. Dr. Jörn Behrens, CliSAP, Numerische Methoden in den
Geowissenschaften

Alexander Njemz

8njemz@informatik.uni-hamburg.de

Universität Hamburg

9. Februar 2011

Übersicht

- 1 Einführung
- 2 Projekte
- 3 Softwareentwicklung: Die Praxis
- 4 Abschluss

Interviewführung

- Vorbereitung

- ▶ <http://www.amatos.info>
- ▶ Behrens, J., Rakowsky, N., Hiller, W., Handorf, D., Läuter, M., Pöpke, J., and Dethloff, K. (2005). amatos: Parallel adaptive mesh generator for atmospheric and oceanic simulation. *Ocean Modelling*, 10(1-2):171 – 183. The Second International Workshop on Unstructured Mesh Numerical Modelling of Coastal, Shelf and Ocean Flows
- ▶ Behrens, J. (2005). Adaptive atmospheric modeling - key techniques in grid generation, data structures, and numerical operations with applications. Thesis of Habilitation
- ▶ Materialien von der Webseite des Seminars

- Hilfsmittel

- ▶ Aufnahmegerät der des Medienkompetenzzentrums (MCC) am Regionalen Rechenzentrum

- Ablauf

- ▶ Treffen in Behrens' Büro, Dauer des Interviews etwa eine Stunde

Interviewpartner

- Seit 2009 am CliSAP
- Leiter der Junior Research Group „Numerische Methoden in den Geowissenschaften“
- 2006-2009 Leiter der Arbeitsgruppe Tsunami Modellierung am Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI)
- Schon von 1991-1998 am AWI, Promotion
- Von 1998-2006 an der Technischen Universität München, Habilitation



Quelle:

<http://www.klimacampus.de/joernbehrens.html>

Was ist das CliSAP?

Der Exzellenzcluster “Integrated **C**limate **S**ystem **A**nalysis and **P**rediction”

- Gestartet im Oktober 2007
- Gefördert von der Deutschen Forschungsgemeinschaft
- Teil des KlimaCampus an der Universität Hamburg, hieran sind 18 universitäre Institute und außeruniversitäre Partner wie das Max-Planck-Institut für Meteorologie, das Helmholtz-Zentrum Geesthacht und das Deutsche Klimarechenzentrum beteiligt

AWI

Alfred-Wegener-Institut für Polar- und Meeresforschung

- Das AWI wurde 1980 gegründet
- Fachbereiche:
 - ▶ Geowissenschaften
 - ▶ Biowissenschaften
 - ▶ Klimawissenschaften



Quelle:

<http://www.awi.de/de/institut/standorte/>

Aufgabengebiet

Wissenschaftliches Rechnen

- Wissenschaftliches Rechnen
 - ▶ „Rechnen“: damit Ergebnisse herauskommen
 - ▶ „wissenschaftlich“: weil es keine Verfahren von der Stange gibt
- Adaptive Gitter und darauf aufsetzend Lösung von partiellen Differentialgleichungen
- Geowissenschaftliche Anwendungen:
 - ▶ Ozeananwendungen
 - ▶ Atmosphärenanwendungen
- Optimierung, Parallelisierung und Lösung großer Gleichungssysteme

Tsunamis

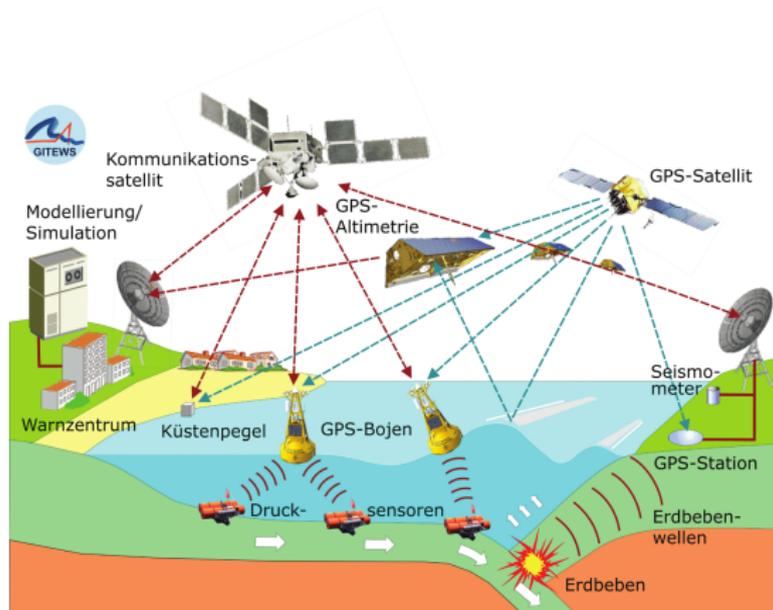
Der Tsunami im indischen Ozean vom 26. Dezember 2004



http://upload.wikimedia.org/wikipedia/commons/a/a6/2004-tsunami_balanced.jpg

Die Komponenten des Deutsch-Indonesischen Tsunami-Frühwarnsystems

German Indonesian Tsunami Early Warning System (GITEWS)



http://www.gitews.de/fileadmin/documents/content/press/GITEWS_Flyer_DE2008.pdf

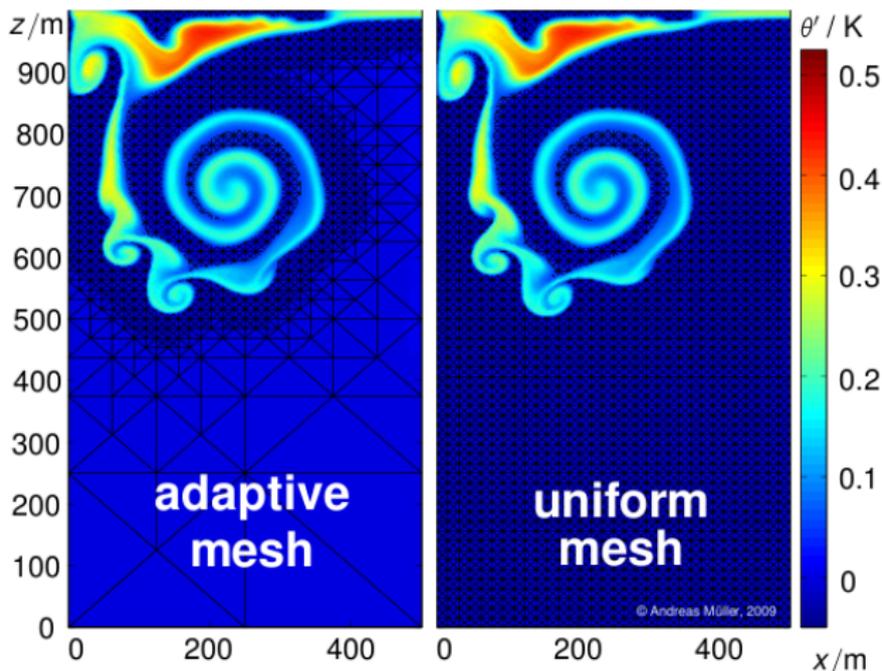
- Prof. Behrens war verantwortlich für einen maßgeblichen Teil der Modellierungs-Komponente
- Eine wesentliche Leistung von Behrens:
Die Unsicherheit in dem Vorhersageprozess zu quantifizieren
- Beteiligte Institutionen: Deutsches GeoForschungsZentrum (GFZ) Konsortialführer, Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI), Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Deutsches Zentrum für Luft- und Raumfahrt (DLR), GKSS Forschungszentrum Geesthacht, Gesellschaft für Technische Zusammenarbeit (GTZ), Konsortium Deutsche Meeresforschung (KDM), Leibniz-Institut für Meereswissenschaften (IFM-GEOMAR), United Nations University - Institute for Environment and Human Security (UNU-EHS).

amatos

Adaptive Mesh generator for **AT**mospheric and **O**ceanic **S**imulation

- Eine Funktionenbibliothek zur Generierung adaptiver Gitter
- Entwicklung begann 1997
- Softwareentwicklungskennnisse nicht vorhanden, autodidaktisch vorgegangen

Was sind adaptive Gitter?



aus: Mueller, A., Behrens, J., Giraldo, F., and Wirth, V. (2010). An adaptive discontinuous galerkin method for modeling cumulus clouds. In *Fifth European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010*. (Submitted)
[Mueller et al., 2010]

- 1 Hardware
- 2 Sprachen
- 3 Verwendete Bibliotheken und Compiler
- 4 Code Editing
- 5 Debugging
- 6 Testing und Validierung
- 7 Versionierung
- 8 Projekt Management und Bugtracking
- 9 Codequalität & Code Review
- 10 Codedokumentation
- 11 Ausblick

Hardware

- Sun-Cluster mit 8 Cores pro Knoten, IBM mit 8 (oder 16?) Prozessoren pro Knoten
- Amatos ursprünglich auf einer SGI entwickelt
- TsunAWI:
 - ▶ IBM Power6 und Power7 Architekturen mit 4-8 Cores und ca. 16 GB je Knoten
 - ▶ Intel Architekturen (SGI Altix) mit Intel Nehalem Prozessoren

Verwendete Programmiersprachen

- Fortran 90/95
- Amatos OpenMP mit parallelisiert
- Bisher nicht mit MPI parallelisiert, vielleicht in Zukunft
- Python zu lernen ist ein Ziel von Prof. Behrens

Verwendete Bibliotheken und Compiler

- NetCDF wird als generisches Datenformat verwendet (nicht Open Source aber frei verfügbar)
- Lineare Algebra Bibliotheken, freie Varianten verfügbar aber proprietäre sind aufgrund der Optimierungen vorzuziehen
- Prof. Behrens selbst benutzt auf Linux den Intel Fortran Compiler

Code Editing

- Die Wahl des Text Editors ist jedem selbst überlassen
- Prof. Behrens benutzt aktuell den XCode Editor auf dem Mac
- Integrierte Entwicklungsumgebungen werden bisher nicht genutzt

Debugging

- Kommandozeilenorientierte Debugger werden verwendet
- Nach Vergleich mit den anderen auf dem Markt, hält Prof. Behrens Totalview für den einzig brauchbaren parallelen Debugger

Testing & Validierung

- Eine Variante von Test Driven Development (TDD):
Es werden nicht zunächst für sämtliche Routinen Testfälle implementiert, aber der Code wird mit Sicht auf evtl. Testfälle geschrieben.
- Validierung bisher: Erhaltungseigenschaften, Vergleich mit analytischen Lösungen
- Rigorose Validierung als Problem und interessantes Forschungsthema

Versionierung

- Amatos zunächst “manuell” versioniert, später mit SVN
- TsunAWI wird mit SVN versioniert
- Nutzen von Versionierung so ziemlich für alles sinnvoll, z.B. auch Doktorarbeiten, das ist jedem unmittelbar klar

src/gridgen/FEM_Gridgen.F90

```

!*****
! USES:
!     MISC_globalparam, MISC_error, FEM_define, FEM_handle, FEM_create,
!     FEM_gridmanag
! LIBRARIES:
!
! REFERENCES:
!     the grid refinement strategy has been taken from
!     M.C. Rivara: "Algorithms for Refining triangular grids suitable
!     for adaptive and multigrid techniques", INT. JOU. NUM. METH. ENG.,
!     Vol. 20, pp. 745-756 (1984)
!     E. Baensch: "Local mesh refinement in 2 and 3 dimensions",
!     IMPACT COMP. SCI. ENG., Vol. 3, pp. 181-191 (1991)
! VERSION(S):
!     1. original version                j. behrens      8/96
!     2. now with edges                  j. behrens     10/96
!     3. now with time and handles, clearup_grid and
!        grid_timeduplicate added, renamed      j. behrens     10/96
!     4. grid_clean added, clearup_grid
!        modified                          j. behrens     5/97
!     5. hashing tables                  j. behrens     7/97
!     6. many changes and bug-fixes      j. behrens     8/97
!     7. grid_timetoggle moved to GRID_api    j. behrens     9/97
!     8. extracted FEM_create            j. behrens    10/97
!     9. parallel version                 j. behrens     1/98]
!    10. find_resolvpatch added          j. behrens     1/98
!    11. find_resolvpatch alg. changed    j. behrens    2/98
!    12. changed edge data struct.       j. behrens     2/98
!    13. updated for amatos version 1.0   j. behrens    11/2000
!
!*****
MODULE FEM_gridgen
  USE MISC_globalparam
  USE MISC_error
  USE FEM_define
  USE FEM_handle
  USE FEM_create
  USE FEM_dataretrieve
  USE FEM_signature
  USE GRID_utils
!--daFlo--
  USE IO_emit
  USE FEM_boundary

```

Projektmanagement und Bugtracking

- Amatos wird unter trac verwaltet
- Am Alfred-Wegener-Institut wurde GForge genutzt
- Am Clisap wird Redmine genutzt
- Ticketing je nach Anzahl der Programmierenden unterschiedlich stark genutzt

Codequalität & Code Review

- Amatos ist gut Modularisiert
- Für das TsunAWI-Project wurde eine Coding Guideline erstellt, angelehnt an die European Standards For Writing and Documenting Exchangeable Fortran 90 Code
- Code Review ist nicht institutionalisiert

Coding Guideline für TsunAWI

Draft for a TsunAWI Coding Guideline

V0.1, Lars Mentrup

V0.2, Joern Behrens, Florian Klaschka, Lars Mentrup

V0.21, Sven Harig, Florian Klaschka, Lars Mentrup

Based on „European Standards for Writing and Documenting Exchangeable Fortran 90 Code“ by MetOffice, UK. http://www.metoffice.gov.uk/research/nwp/numerical/fortran90/f90_standards.html

With little adaptations.

Documentation

- External Documentation split up into
 - Scientific documentation
 - Implementation doc (Developer's guide)
 - User's Guide
- Internal Documentation
 - „Literate Programming“ paradigm: document within the source code using ProTeX
 - Every procedure has a ProTeX header
 - Comment hierarchy
 - Section comments:
!-----
! Comment
!-----
 - Other comments
! Comment

Codedokumentation

- Der Code von amatos ist gut kommentiert
- Das Handbuch zu amatos ist auf dem Stand von 2003
- Für TsunAWI wurde ProTeX genutzt

tsunami.f90

```

!new!
!please put the new comments only from the third column!
!----- Introduction -----
! AWI Tsunami Research Unit
! Tsunami Code Initial Revision
!-----
! PROTEX Introduction section:
!BOI
! !TITLE: \emph{TsUn}AWI -- A Tsunami model on Unstructured meshes provided by Alfred Wegener Institute
!
! !AUTHORS: J.-Behrens, S.-Braune, S.-Harig, W.-Hiller, J.-Schr\^oter, J\{\sc D.-Sein, D.-Sidorenko, O.-Startseva, E.-Taguchi
!
! !AFFILIATION: Alfred Wegener Institute for Polar and Marine Research J\{\em Am Handelshafen 12 J\{\em 27570 Bremerhaven, Germany
!
! !DATE: 20.09.2006
!
! !INTRODUCTION: Introduction
! This document describes the mathematical formulation, discretization, code implementation
! of the AWI Tsunami model...
! \section{Mathematical Reasoning}
! ! The model is based to the barotropic (i.e.\ 2D) shallow water equations in velocity-depth formulation...
!EOI
!-----
!
!----- Module -----
!BOP
!
! !ROUTINE: program MAIN
!
! !INTERFACE:
!
!
!-----
!
! ***** M A I N *****
!-----
PROGRAM MAIN
  USE PARAMETERS
  USE MESH
  USE ELEMENTS
  USE SWE
  USE INITIAL_CONDITIONS
  USE OKADA_FAULT_PARAMETERS

```

Ausblick

- Amatos auf MPI portieren
- CUDA evaluieren unter Nutzung des Fortran Compilers der Portland Group
- Testing automatisieren

Äußere Umstände der SiW

- Zeitdruck ist generell nicht so ein großes Problem in den Wissenschaften, wenn man jedoch ein Paper fertigstellen will, wird auch schonmal „quick'n'dirty“ entwickelt
- Software schreiben ist in Anbetracht der Karriereoptimierung nicht sinnvoll

Kurs zur Softwareentwicklung

- Softwaretechnische Ausbildung der Wissenschaftler verbesserungswürdig
- Prof. Behrens hat einen Kurs von einer Woche angeboten

Kursplan

Course outline Scientific Programming

Time/Day	Monday	Tuesday	Wednesday	Thursday	Friday
10:00-10:45	Introduction: aim, preliminaries, course outline, registering names etc.	Recap of previous day; Intro to today's subject: How to plan software development, Software Life Cycle ("Vorgehensmodelle"), etc.	Recap of previous day; Intro to today's subject: testing and verification, code validation, automated build; what is it good for, etc.	Recap of previous day; Intro to today's subject: code management, maintenance, etc.	Recap of previous day; Intro to today's subject: documenting your code, intro to using doxygen, etc.
10:45-11:00	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break
11:00-12:30	Facts & Fallacies..., the hardware-software mapping problem, difference to scientific software development	Intro to project management, programming approaches, estimation, etc.	Sensitivity analysis, testing strategies, hierarchies of verification tests, unit tests, etc.	Introduction to debugging, gdb, ddd, et al.	Documentation: levels of documents, tools for online documentation, applying doxygen
12:30-13:30	Lunch break	Lunch break	Lunch break	practical work	Lunch break
13:30-15:00	Versioning: Introduction to SVN, etc.	Using ticketing, release plans, etc.	Intro to testing software, automated build (Make, et al)	free	Exam: I want to see your own project checked in, at least two versions, partly documented, Feedback etc.
15:00-15:15	Coffee break	Coffee break	Coffee break	free	END
15:15-16:00	Setting up an individual svn repository, checking in example code, checkin in your own project...	Practical lesson with demo.redmine.org	Practical work with make and tests	free	

OLAT-Page: <https://uhh-srv-olatweb.rz.uni-hamburg.de:443/olat/auth/repo/go?rid=148897792>

Zusammenfassung

- Nutzung von Tools und Methoden
- Verhältnis der Wissenschaftler zur Softwaretechnik

Fazit

(Subjektive Meinung des Vortragenden!)

- Prof. Behrens ist sich des Nutzens softwaretechnischer Methoden bewusst, aber es gibt noch Raum für Verbesserung.
- Fazit zum Interview



Behrens, J. (2005).

Adaptive atmospheric modeling - key techniques in grid generation, data structures, and numerical operations with applications.

Thesis of Habilitation.



Behrens, J., Rakowsky, N., Hiller, W., Handorf, D., Läuter, M., Pöpke, J., and Dethloff, K. (2005).

amatos: Parallel adaptive mesh generator for atmospheric and oceanic simulation.

Ocean Modelling, 10(1-2):171 – 183.

The Second International Workshop on Unstructured Mesh Numerical Modelling of Coastal, Shelf and Ocean Flows.



Mueller, A., Behrens, J., Giraldo, F., and Wirth, V. (2010).

An adaptive discontinuous galerkin method for modeling cumulus clouds.

In *Fifth European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010*.

(Submitted).

- amatos Webseite

<http://www.amatos.info>

Letzter Zugriff: 05.02.2011

- Gitews Flyer

http://www.gitews.de/fileadmin/documents/content/press/GITEWS_Flyer_DE2008.pdf

Letzter Zugriff: 05.02.2011

- Gitews Webseite

<http://www.gitews.de>

Letzter Zugriff: 05.02.2011