

# C3-Grid – Gridtechnologie in der Klimaforschung



U.Ulbrich<sup>1</sup>, W.Hiller<sup>1</sup>, B. Fritsch<sup>1</sup>, R. Budich<sup>2</sup>, M. Stockhause<sup>2</sup>, I. Kirchner<sup>3</sup>, H. Kupfer<sup>3</sup>, C. Kurz<sup>4</sup>, S. Kindermann<sup>5</sup>, K. Ronneberger<sup>5</sup>

<sup>1</sup> Alfred Wegener Institut für Polar- und Meeresforschung (Projekt Koordination) <sup>2</sup> Max Planck Institut für Meteorologie Hamburg/Leibniz Institut für Meereswissenschaften Kiel <sup>3</sup> Institut für Meteorologie der Freien Universität Berlin <sup>4</sup> Deutsches Zentrum für Luft- und Raumfahrt DLR <sup>5</sup> Deutsches Klimarechenzentrum DKRZ



C3-Grid ist ein Community-Projekt der D-GRID Initiative, die im September 2006 mit dem GRIDtoday Editor's Choice Award ausgezeichnet wurde.

## Von der Nutzeranfrage zu verwertbaren Ergebnissen

In der Erdwissenschaft ist der Datenvergleich von besonderer Bedeutung. Die Daten aus Beobachtungen oder Modellstudien sind verteilt auf viele Datenarchive und variieren stark in ihrer Zugänglichkeit und Qualität. Gleichzeitig wächst das Datenvolumen kontinuierlich an. Deshalb ist für die Forscher der Weg bis zum Datenzugriff häufig beschwerlich und ineffizient.

Das C3-Grid entwickelt Lösungen für einen einfachen Zugriff auf inhomogene Datenarchive wie auch Prozessressourcen. Dafür werden eine einheitliche Datenbeschreibung sowie eine neue Infrastruktur für die vorhandenen Speicher- und Prozessmöglichkeiten in der deutschen

Klimagemeinschaft unter Nutzung von Grid-Technologien aufgebaut. Diese werden dem Nutzer den Zugriff, die Manipulation und die Verwaltung von Erdsystemdaten möglich machen: Suche, Auffinden, Sammeln, effektives Auswählen, Volumenreduktion und Übertragung.

Die Anwender können über das C3-Grid die Datenprozessfunktionalität steuern, für die es eine grundlegende und höherwertige Services anbietet. Nutzerspezifische Processingkomponenten können flexibel arrangiert werden, um einen komplexen nutzerspezifischen Workflow zu formen. Dieser wird von den Nutzern in einem Portal getriggert und kontrolliert, wofür sie nur einen Standard-Webbrowser als lokale Anwendung auf ihren PCs sowie ein Nutzerzertifikat benötigen.

## » Wo befinden sich Basisdaten für die Klimaforschung? «

### Wissenschaftliche Workflows im Grid

Alle im C3-Grid verfügbaren Daten werden durch ein einheitliches Metadatenschema beschrieben. Grundlage bilden die ISO-Richtlinien ISO 19115/19139 und die Climate and Forecast (CF) Standardnamen. Über ein Portal kann der Nutzer dann die Daten suchen und auf sie zugreifen.

Neben der Datensuche können die Nutzer vom Portal aus ihre spezifischen Workflows zusammenstellen, anstoßen und kontrollieren sowie ihre Ergebnisse abrufen. Die Authentifizierung und Autorisierung für das C3-Grid ist derzeit zertifikatsbasiert.

Die Griddienste übernehmen das Processing der nutzerspezifischen Workflows (Abb. 7): Das Scheduling entscheidet nach verfügbaren Ressourcen und Tools, bei welchem Ressourcenanbieter welcher Teil des Workflows ausgeführt wird und sorgt dort zusammen mit dem Datenmanagement für die Bereitstellung der dazu benötigten Daten.

### Beispiele für wissenschaftliche Workflows

Hier sollen zwei wissenschaftliche Workflows beispielhaft vorgestellt werden:

- Diagnose Tools
- Chemische Wettervorhersage

#### Beispiel: Diagnose Tools

Eine wissenschaftliche Aufgabe, die auf standardisierten Berechnungen basiert, ist die Berechnung eines Mittels aus einem Ensemble von Klimamodell-Rechnungen, einschließlich des mit steigendem Treibhausgas-antrieb verbundenen Klimasignals (Abb. 4a und 4b). Standard-Berechnungen für die Klima-Diagnose werden über die oben genannte Basis-Funktionalität hinaus im C3-Grid implementiert. Als Beispiele für solche komplexen Größen dienen die sogenannten Stormtracks als Maß für die synoptisch-skalierte Variabilität, die Baroklinität, Feuchtelüsse und CAPE (convective available potential energy). Diese Berechnung dieser Größen für die verfügbaren Basisdatensätze wird ebenfalls über das Portal angefordert (Abb. 2). Die Auswahl der für die Berechnung erforderlichen Parameter ist dabei Teil der Funktionalität.

Prototypisch wurde diese Berechnung des Stormtracks für das C3-Grid implementiert. Sie basiert auf verschiedenen Prozessschritten (Abb. 3):

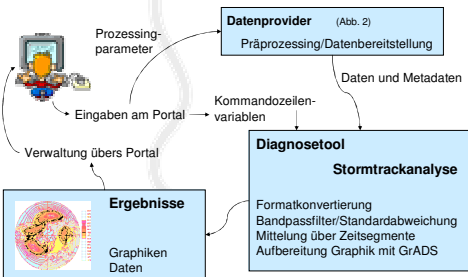


Abb. 3: Nutzerszenario für die Stormtrackanalyse als Diagnose-Tool

### Typisches Nutzerszenario (Abb. 1)

1. Datensätze im Portal suchen
2. Aufbereitung der Datensätze von verschiedenen Datenprovidern gemäß der Nutzeranfrage (Präprocessing/Datenbereitstellung; Abb. 2). Die vom Datenanbieter angebotenen Services beinhalten eine Datenreduktion durch Auswahl einzelner Klimavariablen sowie durch räumliches und zeitliches Ausschneiden.
3. Zusammenführung der Daten von verschiedenen Datenanbietern
4. Nutzerspezifische Prozessschritte (z.B. Vergleiche, Diagnosen, Statistiken usw.) unter Verwendung von grid-enabled Tools wie z.B. CDOs.
5. Anzeigen der Ergebnisse im Portal und evtl. Herunterladen durch den Nutzer oder weitere Bearbeitung über das C3-Grid.

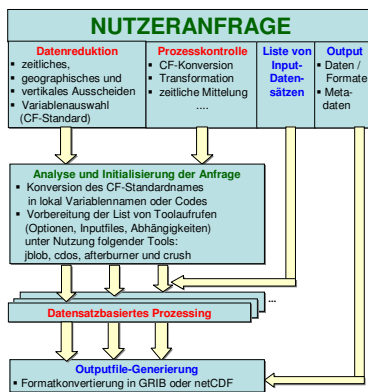


Abb. 2: Schema für das Präprocessing / Datenbereitstellung

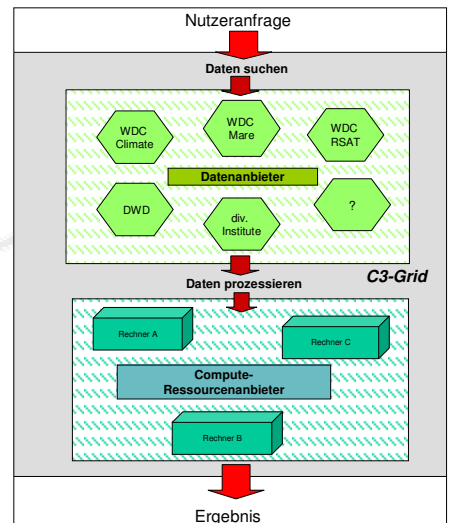


Abb. 1: C3-Grid als Black-Box-Schema



Abb. 5: DLR-Messkampfenflugzeug HALO. Mit freundlicher Genehmigung von Gulfstream.

## » Wie bekommt der Wissenschaftler Zugriff auf Klimadatensätze? «

### Chemische Wettervorhersage

Die Planung und Durchführung von Flugzeug gestützten Messkampagnen (z.B. mit dem neuen Atmosphärenforschung-Flugzeug HALO, Abb.5) sollen mit Hilfe meteorologischer und chemischer Prognosen sowie durch Analyse von Klima-Chemie-Simulationen vor Ort unterstützt werden. Dabei steht dem Experimentator typischerweise nur eine Netzanbindung geringer Bandbreite zur Verfügung. Das C3-Grid ermöglicht dem Experimentator in dieser Situation die Steuerung des Prognosesystems bis hin zur Auswertung über das Portal. Die Prognosen des meteorologischen und chemischen Zustandes der Atmosphäre werden dann vor Ort benutzt, um die Flugroute im Hinblick auf möglichst günstige Messorte zu optimieren.

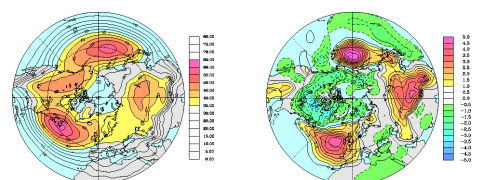
Die Vorhersage chemischer Spurenstoffverteilungen ("chemical weather forecast") erfordert sowohl hohe Rechenkapazität als auch die Kombination verteilter Ressourcen.

Das Chemie-Prognosemodell (hier: ECHAM5/MESSY, Abb.6) wird auf einem nationalen Höchstleistungsrechenzentrum betrieben. Die notwendigen Randwerte wie Vorhersagedaten eines Wettervorhersagezentrums (hier ECMWF) müssen von extern eingespielt werden. Emissionskataster werden aus weltweit verteilten Datenzentren abgerufen (z.B. EDGAR) und lokal vorprozessiert. Die so gewonnenen Chemie-Vorhersagen werden numerisch und graphisch aufgearbeitet und in komprimierter Form dem externen Experimentator im C3-Grid Portal zur Verfügung gestellt.

Über das Portal hat der Experimentator nach einer ersten Bewertung nicht nur die Möglichkeit Messflugrouten zu optimieren (z.B. Vergleiche mit klimatologischen Daten, um die Repräsentativität der Messsituation zu bewerten). Er kann auch erste Ergebnisse der Messflüge im C3-Grid zur Verfügung stellen, damit sie an verschiedenen Standorten verfolgt werden können.

## » Wie werden vergleichbare und reproduzierbare Ergebnisse erzeugt? «

Abb. 4a: MSLP-Stormtrack, IPCC-Ensemble-Mittel aus 5 Klimamodellen Kontroll-Zeitraum 1961-2000, Einheit: 10<sup>1</sup>hPa



Modelljahren 2081-2100 (Szenario A1b) und Kontroll-Zeitraum, Einheit: 10<sup>1</sup>hPa.

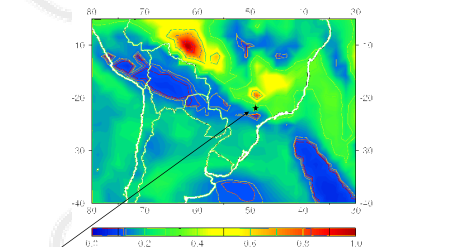


Abb. 6: Beispiel: 3-Tage-Vorhersage von Stickstoffoxiden. Mischungsverhältnis bei 200 hPa über Südamerika. Der Stern zeigt den Ort der Messstelle an.

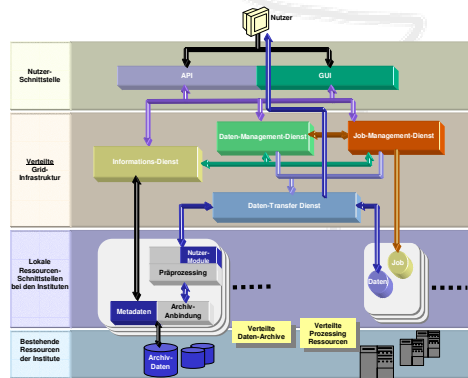


Abb. 7: C3-Grid Architektur

## » Wie kann der Einsatz verfügbarer Ressourcen optimiert werden? «

**Wissenschaftliche Anwender:** Max-Planck-Institut für Meteorologie Hamburg (MPI-M), Leibniz-Institut für Meereswissenschaften an der Universität Kiel (IFM-GEOMAR), Universität zu Köln, Freie Universität Berlin (FUB), Potsdam-Institut für Klimafolgenforschung (PIK), Deutsches Zentrum für Luft- und Raumfahrt (DLR), Alfred-Wegener-Institute für Polar- und Meeresforschung (AWI), GKSS-Forschungszentrum

**Informatik-Partner:** Universität Dortmund, Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB)

**Datenanbieter:** Weltklimazentren: WDC Climate, WDC Mare, WDC RSAT, Deutscher Wetterdienst (DWD), Deutsches Klimarechenzentrum (DKRZ), Alfred-Wegener-Institute für Polar- und Meeresforschung (AWI), Leibniz-Institut für Meereswissenschaften an der Universität Kiel (IFM-GEOMAR)

**Assoziierte Partner:** Universität Hannover, Universität Bonn, Forschungszentrum Karlsruhe (FZK), Sun Microsystems Inc., NEC Cooperation, Brockmann Consult

**Internationale Kooperationen**  
Enabling Grids for E-science - EGEE (<http://www.eu-egee.org/>)  
Nerc Data Grid NDG (<http://ndg.badc.rl.ac.uk/>)  
Earth System Grid - ESG (<http://www.earthsystemgrid.org/>)

### Kontakt

Alfred-Wegener-Institut für Polar- und Meeresforschung  
Bernadette Fritsch  
Postfach 12 01 61, 27515 Bremerhaven  
E-Mail: [fritsch@awi-bremerhaven.de](mailto:fritsch@awi-bremerhaven.de),  
Internet: <http://www.c3-grid.de>

GEORDET VOM

