



## Collaborative Climate Community Data and Processing Grid (C3-Grid)

### Anbindung der Compute-Ressourcen-Provider an das C3-Grid (Schnittstelle F)

---

Arbeitspaket:	AP7
Autor:	Volker Winkelmann
Version:	März 2008
Veröffentlichungsdatum:	März 2008
AP-Koordination:	
Partner:	
Ansprechpartner:	Volker Winkelmann
Email:	Winkelmann@Uni-Koeln.DE

---

*Projekt im Rahmen des D-Grid  
Gefördert vom BMBF*

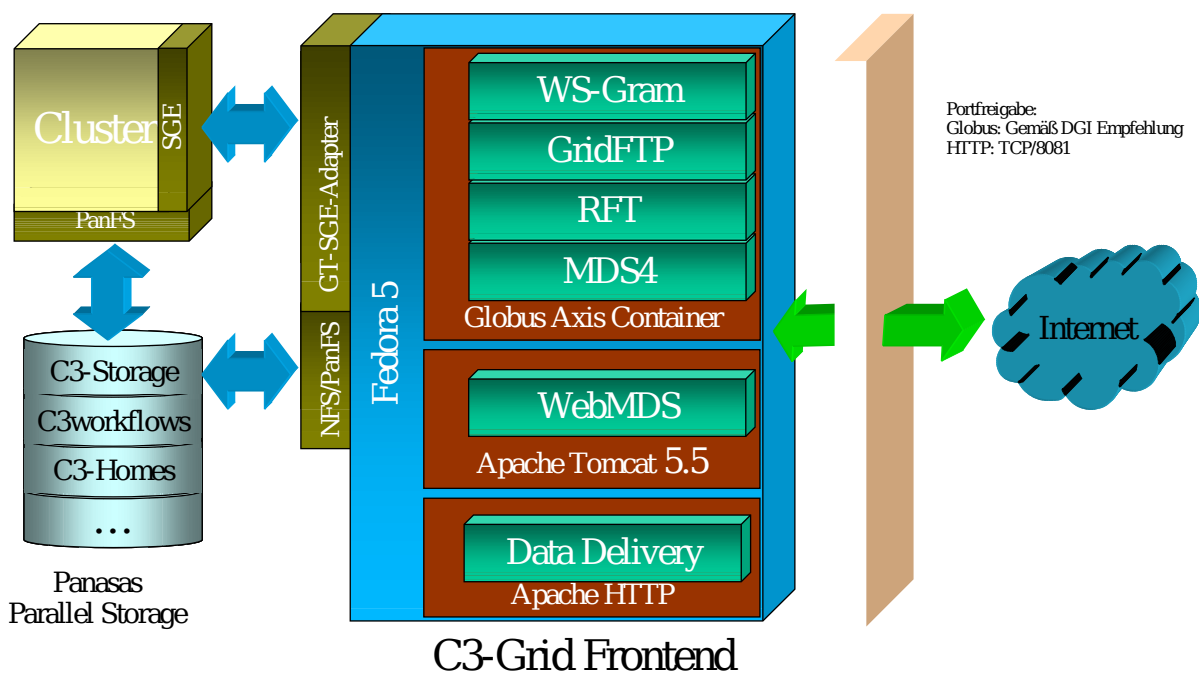


# 1 Anbindung der Compute-Ressourcen-Provider an C3-Grid

## (Schnittstelle F)

Flankierend zum Aufbau der C3-Grid-Middleware mussten die Compute-Ressource-Provider die notwendige Infrastruktur für das Processing der C3-Grid-Workflows bereitstellen. Die Aufgaben, die eine solche Infrastruktur übernehmen muss, sind die Bereitstellung von Workflows, der Empfang von Aufträgen aus dem C3-Grid zur Abarbeitung von Workflows, der Empfang von Eingabedaten aus dem C3-Grid sowie die Auslieferung der berechneten Ergebnisse eines Workflows zurück an das Portal.

Für die Kopplung von Compute-Ressourcen an das C3-Grid ist bei jedem Provider ein Grid-Frontend erforderlich, welches gridseitig Daten und Aufträge annimmt und an die vom Provider zur Verfügung gestellten Compute-Ressourcen, typischerweise HPC-Cluster, übermittelt, wobei konzeptionell Grid-Frontend und Compute-Ressourcen im C3-Grid hardwaremäßig getrennt sind, siehe Abbildung 1.



**Abbildung 1: C3-Grid-Frontend und Compute-Cluster an der Universität zu Köln**

Abbildung 2: C3-Grid-Frontend an der Universität zu Köln

Da im C3-Grid Globus Toolkit 4 (GT 4) als Grid-Middleware eingesetzt wird, musste jeder Compute-Provider zunächst auf einem eigenen Frontend-Rechner GT 4 mit den Services WS-GRAM, MDS, WebMDS, RFT installieren und konfigurieren. Da die Anbindung der WS-GRAM-Komponente vom Batch-System der eingesetzten Compute-Ressource abhängt, war jeweils ein spezieller WS-GRAM-Batch-Adapter erforderlich, der an die lokalen Rahmenbedingungen angepasst werden musste.

Jedes Grid-Frontend wird mit einem Workspace ausgestattet, der den Empfang von Daten und die Auslieferung von Workflow-Ergebnissen an das Portal ermöglicht. Für den Datenaustausch zeigte es sich zweckmäßig, einen geschalteten Datenbereich zwischen der Compute-Ressource und dem Grid-Frontend einzurichten: es werden unnötige Datentransfers im lokalen Netz des Providers vermieden, und die Ergebnisse sind aus dem C3-Grid-Portal über einen HTTP-Server auf dem Grid-Frontend leicht abrufbar. Für die Sicherheitsaspekte der Lösung wurde ein Konzeptpapier entwickelt.

Im C3-Grid werden zu bearbeitende Workflows vom C3-User nicht vollständig auf eine Compute-Ressource übertragen, sondern es wird erwartet, dass diese auf der Compute-Ressource bereits vorinstalliert sind. Daher mussten die Provider die notwendigen C3-Grid-Workflows auf den Compute-Ressourcen vorinstallieren und auf dem jeweiligen Grid-Frontend einen zusätzlichen C3-Info-Serviceprovider bereitstellen, der Auskunft über die vorhandenen Workflows, Softwareumgebungen oder Bibliotheken etc. gibt.

Um die Portabilität der Workflows zwischen den vorkommenden unterschiedlichen Compute-Ressource-Plattformen zu erleichtern und lokale Software-Abhängigkeiten zu umgehen, wurde ein Konzept zur Modularisierung der Workflows und seiner einzelnen Tasks unter Verwendung der sogenannten Environment-Modules (-> vgl. Deliverable) auf den Compute-Ressourcen entwickelt und umgesetzt.