

## 2. C3Grid Userworkshop

Zusammenfassung der Diskussion vom 17.02.2009 mit Nutzeranforderungen

### **Portal**

- Übersichtliche und gut funktionierende Suche
- Betonung des C3Grid als allg. Datenquelle (Nutzer laden sich ihre gewünschten Daten lokal auf ihren Rechner, weitere Rechnungen werden dann außerhalb des Grids gemacht)
- Wünsche der Nutzer
  - a) Abschätzung von Zeit für Datenbereitstellung & Größe der zu erwartenden Datei
  - a) Fortschrittsanzeige & Prozessablauf, Arbeitsschritte auf Anforderung anzeigen evtl. Fortschrittsbalken
  - b) bei Abbruch der Jobs, wo wurde der Job in der Prozesskette abgebrochen, welche Arbeitsschritte wurden abgearbeitet
  - c) Fehlermeldungen für Techniker & Anwender in Mails wenig aussagekräftig
- C3Grid Antworten
  - a) Abschätzung von Dauer & Größe theoretisch möglich, benötigt aber nach Abschicken selbst eine gewisse Zeit; Anfrage läuft durch das gesamte System (muss bis zum Datenprovider durchgereicht werden – ist im Prinzip vorhanden)
  - b) Kommunikationskanal bzw. Schnittstelle vom/zum Provider fehlt noch  
  
Information über die bei einem C3-Job abgearbeiteten Arbeitsschritte müssen Datenprovider (Staging und Preprocessing), Compute-Provider (Start und Ende der eigentlichen Analyse) liefern; Schritte in der Analyse müssen von WF-Entwicklern dokumentiert werden.
  - c) Soll verbessert werden.

## **CDO-Workflow**

- Ensembles-Funktionalität
- Wünsche der Nutzer
  - a) Individuelle Zusammenstellung der einzelnen Ensemblemembers für die Berechnung von ensmin, ensmax usw.
- C3Grid Antworten
  - a) Aufwand ist realisierbar
  - b) Workflow muss angepasst werden
  - c) bei Datenprovidern entsteht erhöhter Speicherplatzbedarf
  - d) sequentielle Abarbeitung kann zur numerischen Instabilität führen

## **Realisierung des Anwendersupports**

- Wünsche der Nutzer
  - a) Wiki oder
  - b) moderierte Mailingliste oder
  - c) Diskussionsforum
  - d) jede Komponente sollte Ansprechpartner haben

## Rechnerarchitektur & Workflows

- Wünsche der Nutzer
  - a) um die Vergleichbarkeit der Ergebnisse bei wiss. Auswertungen zu gewährleisten, sollte für die Abarbeitung von Workflows die gleiche Rechnerarchitektur verwendet werden, Reproduzierbarkeit muss gegeben sein
  - b) Steuerung der Wahl der Rechnerarchitektur im Portal ohne die Übersichtlichkeit zu verlieren
  - c) Information über die Rechnerarchitektur
  - d) Berechnung von Workflows immer auf dem selben Rechner soll angefordert werden können, d.h. Nutzer kontrolliert, wo gerechnet wird
  - e) Download von Skripten, Quellcodes der Workflows
  - f) Zuordnung der Bilder aus Workflow zum gewählten Datensatz bzw. Informationen über den Datensatz
- C3Grid Antworten
  - a) es ist nicht vorgesehen, dass gleiche Skripte auf allen unterschiedlichen Architekturen laufen. Falls Einschränkungen bzgl. der Rechnerarchitektur vom Nutzer für den WF notwendig sind, muss das beim Submittieren angegeben werden (noch zu implementieren)
  - b) Information über Rechnerarchitektur werden im C3Grid über das RIS bereitgestellt.
  - c) Berechnung der Workflows immer auf ein und demselben Rechner im Grid hat wenig Sinn, besser direkt Account beim jeweiligen Institut oder Datendownload und Berechnungen auf dem „Heimrechner“ durchführen
  - d) Workflowentwickler stellen Skripte & Quellcodes zur Verfügung
  - e) Zuordnung der Bilder über Datenbank möglich
- Stromtrack: Wintermittel nicht möglich? Gelieferte Spezifikation von FUB dahingehend überprüfen

## Eigene Workflows

- 2 Arten von Workflows:
  - 1. spezielle Workflows: Benutzer ist voll verantwortlich für den Workflow und steuert den Workflow über Parameterlisten; weiß was für Daten er benutzen darf
  - 2. allgemeine Workflows – Änderungen am Portal; C3Grid übernimmt Datenauswahl
- Wünsche der Nutzer
  - a) freie Wahl der Programmiersprache für eigene Workflows
  - b) minimaler Aufwand für das Schreiben eigener Workflows
  - c) Persistenz von eigenen Workflows im C3Grid
  - d) eigene Workflows der Arbeitsgruppe zur Verfügung stellen
- C3Grid Antworten
  - a) bisher keine Notwendigkeit zur Einschränkung der Programmiersprache
  - b) Didaktisches Beispiel für die Gegenüberstellung Original-Script vs. gridifiziertes Script soll erstellt werden
  - c) Definition einer Spezifikation für einen Container, in dem eigene Workflows abgelegt werden können und die zentral verwaltet werden
  - d) Rechteverwaltung für eigene Workflows unklar

## Benutzergruppen

- 2 Arten von Benutzergruppen:
  - 1. Benutzer, die direkten Zugang zu Klimadaten als primäre Informationsquelle (z.B. IPCC-Modelldaten) wünschen, jedoch wenig wissenschaftlichen Background auf dem Gebiet der Klimaforschung haben
  - 2. Klimaforscher, die Daten in komplexer Art und Weise flexibel auswerten möchten
- kurz & mittelfristig: Weiterentwicklung der Funktionalität für Klimaforscher (Benutzergruppe 2)
- langfristig: Erweiterung auf Nutzergruppen außerhalb der Community (Benutzergruppe 1), Beratung als Filter vor Nutzung der Daten notwendig

## Datenprovider

- Wünsche der Nutzer
  - a) Erstellung der Metadaten sollte weitgehend automatisiert werden
  - b) Temporäre Einstellung von Daten mit einem Minimum an Informationen für Metadaten – sollten nicht im Metadatenkatalog auftauchen, nicht zwangsweise an Übermittlung an einen Datenprovider gebunden
- C3Grid Antworten
  - a) Erstellung weitgehend automatisiert aus einem netcdf-header rausholen
  - b) Sichtbarkeit von Metadaten durch zentralen Metadatenkatalog:  
Frage: „Will ich die Metadaten, zu Datensätzen, die ich nicht herunterladen darf, sehen oder will ich nichts von diesen Datensätzen wissen?“

## Sonstiges

- Wünsche der Nutzer
  - a) Abbruchmöglichkeit für gestartete Jobs
  - b) Für laufend ergänzte (aktuelle) Daten im C3Grid soll es möglich sein, Datendownloads über das Portal mittels cronjobs durchzuführen.
  - c) Möglichkeit, eine Deadline für die Abarbeitung von Jobs anzugeben
- C3Grid Antworten
  - a) Abbruchmöglichkeit theoretisch gegeben, muss durchgereicht werden
  - b) je nach Priorität der Nutzer möglich
  - c) möglich, bedeutet auch, dass Lebenszeit der Ergebnisdaten an Jobausführungszeit angepasst werden muss