GRID activities in Wuppertal

Taking advantage of GRID software now

D0RACE Workshop
Fermilab 02/14/2002

Christian Schmitt
Wuppertal University

Overview

- Why GRID ?
- Computing power
- Software tools
- Current status
- Future plans
- Summary

Why GRID in Wuppertal ?

- Current and future experiments need huge amount of CPU power
  - MC production
  - Data reconstruction
  - Analysis jobs
- we need a Grid framework for own software development
- Combine needs of different experiments in one Software

Where to get the CPU power from ?

- Current desktop machines are very fast (~1 GHz) and have enough memory (256 Mb)
- They are only used on working hours for daily work
- There are more of them than you might think: Wuppertal: ≈30 (³ 800 MHz)
- **connect all 5 German institutes and**
- **Karlsruhe as regional analysis center**

Software tools

- Condor(-G)
  - High Throughput Computing environment: uses also PC which are „normal desktop machines“ – CPU power which is lost otherwise.
  - Can manage very large collections of distributively owned workstations
  - Standard (non Grid) jobs independant of Linux environment (like FermiLinux, SuSE, etc)
  - No need for user account on the machines, no shared file system → easy to administrate

Software tools

- SAM
  - Data transfer from Fermilab to the different SAM stations and vice versa
- Globus
  - Authentification of users
  - Data transfer from nearest SAM station to the desktop PC which is running the job

Current status

- Condor running on 14 (future: >30) PCs
  - different Linux versions
  - different memory/CPU resources
  - different network throughput
  - some Grid enabled, some not – no problem!
- Achived: have a powerful queueing system running which uses **all** possible CPUs **and** have a Grid environment for software development
- Example: before → 2 mio AMANDA MC/day

  now → 30 mio /day

  because of intelligent use of idle PCs !!!

Current status

- Software running in the Condor-Pool:
  - D0GSTAR
    - executable, dynamic libraries and rcp database put together in a „Sandbox“ which is then submitted
  - CORSIKA (MC generator for AMANDA)
- Successful: first extension to other institute:

  GRID connection with Aachen
- Started programming a jobmanager which takes care of Analysis chains, e.g.:
  Pythia→Geant→Reconstruction →Analysis

Future plans

- Write „Sandbox“ for all main D0-Programs:
  - D0Reco, RecoAnalyze, ...
- Use SAM to get the files from/to the local stations
- Connect all the 5 German institutes
- Setup of Karlsruhe as regional analysis center

Summary

- Successful usage of desktop PCs to run batch jobs
- Data transfer done using GRID technology (Globus)
- Jobs submission over institute boundaries
- Join all 5 German institutes and Karlsruhe
- Take advantage of existing GRID software/technology **now**