Experiences and lessons learnt from running a shared and heterogeneous university computer cluster in grids

Volker Büge\textsuperscript{1,2}, Christopher Jung\textsuperscript{1,2}, Yves Kemp\textsuperscript{2}

1: Institut für Wissenschaftliches Rechnen, Forschungszentrum Karlsruhe
2: Institut für Experimentelle Kernphysik, Universität Karlsruhe
Grid Computing in Karlsruhe

Forschungszentrum Karlsruhe:

- Founded in 1956 for nuclear research
- Focus on science and engineering
- Germany‘s Tier-1 center for all 4 LHC experiments, CDF, D0, Compass and Babar

=> Closer collaboration planned in the framework of the “Karlsruhe Institute of Technology” (from July 2006 on)

University of Karlsruhe:

- Founded in 1825 as a technical university
- Today research university with focus on science and engineering
- Tier-2/3-Prototype for CMS experiment at IEKP institute
- SAMStation for CDF at IEKP
Karlsruhe on the WLCG monitoring map
Why do we need grid computing?

- In all areas of science, the available amount of data for research grows rapidly, e.g. in High Energy Physics
- The need for larger computing power for analyses grows
- Data storage and processing power widely available at institutes which are working on the same project
  - **Collaborating groups** cope with these challenges via grid tools
  - **Opportunistic/shared use** of resources between local users and grid users
    → minimization of idle time, interception of peak loads, shared data storage and shared effort in deployment of common services
Computing in university groups

- Diverse challenges:
  - Several user groups with different applications and possibly conflicting interests
  - Computer clusters have been growing over time
    -> structure of clusters much influenced by 'historic' decisions
  - Heterogeneity: hardware, software, funding and ownership
  - Clusters have to adhere rules and guidelines set by institutes, faculties and university
  - Fluctuations in responsible team of administrators, as they are often PhD students and post-docs

→ Integration into existing grids not easy at all!
A typical example: IEKP

IEKP (Institut für Experimentelle Kernphysik) is an institute for experimental high energy physics

Local users in working groups

Supported grids:
- Sequential Access via Metadata Grid
- Worldwide LHC Computing Grid

heterogeneous environment in:
- software & hardware
- local and grid users
- access policies
- grid middleware
Resources at IEKP

- **Hardware (Linux cluster):**
  - Portal machines for the experiments (3 for CDF, 2 for CMS, 1 for AMS)
  - 5 file servers with ~20 TB of disk space
  - 27 computing nodes with 36 CPUs
  - Independent of desktop cluster

- **Human resources**
  - Grid team of 3 persons (~2 FTE)
  - Often changes in team of responsible persons
Peculiarities of IEKP Linux cluster

- PBS/Torque (Open Source) as local Batch System, using MAUI scheduler
- Institute firewall protects cluster (incl. grid components)
  - some special ports have to be open for grid services
  - no university firewall for our Linux cluster
- PCs in desktop cluster provide access to portal machines
- 1 GBit firewall connects clusters
- Software on portals and worker nodes is experiment dependent
- Linux flavor is Scientific Linux 3 (at the moment experiments do not require different flavors; this might change)
Site specific grid services @ IEKP

**SAMGrid Station**
- One portal machine per cluster
- Offers following services:
  - User interaction
  - Mass storage
  - File import, file export, file delivery to analysis programs
  - File exchange via GSIFTP
  - Analysis job activity is written to central database

**LCG services**
- Computing Element:
  - Gateway to Worker Nodes of the local batch system
  - Globus Gatekeeper
- Storage Element
  - Gateway to local disk storage
  - Globus GridFTP server, SRM interface
- Monitoring Box
- User Interface
  - Users‘ access point to the grid
Hardware Consolidation by Virtualisation (I)

Advantages:

- **Only one** single high-performance server needed for:
  - **Complete** LCG installation (including a test Worker Node)
  - Second LCG installation for testing
  - SAMGrid station

- **Easy setup** of basic OS by copying image files

- Possibility of **migration** VMs to other servers; easy **backup**

- Balanced load and **efficient use** of the host machines
  → interception of CPU peaks
Host system:

- AMD Dual Opteron with 2 GB RAM
- RAID-10 for stability and performance
- Debian Linux

Virtualization done with XEN 3.0:

- No hardware emulation at all. Guest systems access hardware indirectly via host system and virtual machine monitor. -> Only small loss of performance
- Guest-OS has to be ported to XEN (kernel patch), but not the applications!
Conclusions and Outlook

Conclusions

- University groups profit from grid computing
- Institutes‘ clusters can be successfully integrated into grids
- University groups have to deal with heterogeneity in various aspects (hardware, funding, requirements set by software, …)
- Virtualization is useful for consolidation of hardware

Outlook

- Virtualization will be used for allowing different Linux flavors on one Worker Node
- Further improvements in middleware needed