

DESY Site Report.

HEPiX Spring meeting 2010



Wolfgang Friebel
HEPiX
Lisbon, April 2010

Personalia

- Already since last year ...
- Helmut Dosch Director General of DESY



- Joachim Mnich Director in charge of HEP and Astroparticle Physics



- Edgar Weckert Director in charge of Photon Science



➤ Planned and recently finished facilities



- Installation of PETRA III DAQ storage system in HH computing centre
 - data from beamlines will be transported to computer center via aggregated 10 GbE links
 - annual volume expected to be 500TB to 1 PB when PETRA III is fully operational
 - experiment data will be transferred to analysis space still under specification
 - storage system is dual NetApp 3140 cluster w/ OnTap 8, 90 TB net each
 - near real time analysis feedback loop to experiments is in preparation

- Setup of a analysis farm prototype for CFEL
 - analysis jobs will take ~ 100GB main memory and many cores
 - combination of 'regular' CPU code and GPU based code (CUDA, OpenCL)
 - possibly going for stronger MPI based methods
 - eyeing new quad 8 core machines from DELL and 6 core SGI machines
 - starting pilot in May with 3 weeks worth of experimental data (~ 40TB)
 - during analysis data volume increase 3x, 50-75% is considered for long time archiving
 - this is a prototype for one research group, one beamtime, one analysis run
 - of each of those we will have many



New Astrophysics Activities

➤ Participation in the Cherenkov Telescope Array Project

- Evaluation of ACS (Alma Control Software) for the Array Observation Center
- Taking part in Monte Carlo Production



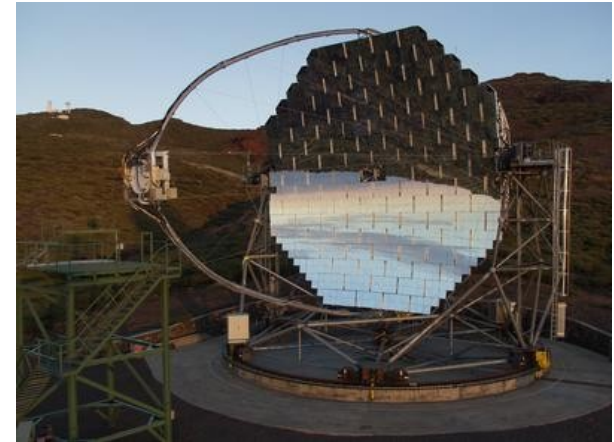
➤ Computing support for the IceCube Experiment

- Zeuthen is European IceCube Data Center
- Plans to become Disaster Recovery Center for IceCube
- IceCube is now using the (WLCG) Grid Infrastructure



➤ Analysis of MAGIC (Gamma Ray Telescope) data

- Using noticeable parts of mass storage and farm nodes



Remote control rooms for LHC

- Remote control rooms established for ATLAS and CMS



CMS remote control room in Hamburg



ATLAS remote control room in Zeuthen

Grid activities

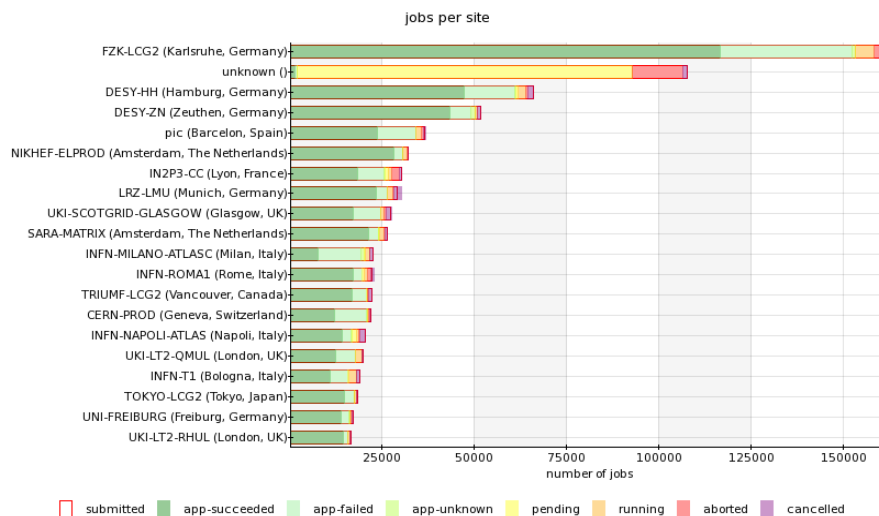
➤ Playing well within the german cloud

- More than 4900 cores for the grid (4236 HH + 672 Zn)

➤ Grid activities not restricted to LHC

- Good fraction of H1 Monte Carlo simulations running on the grid at DESY
- IceCube activities
- CTA (Cerenkov Telescope Array) simulations started
- LQCD uses the LDG (Zeuthen development) for world wide data distribution

➤ DESY is performing very well in Challenges and Cloud tests



> National analysis facility

- Upgrade from 800 -> 1500 cores, currently more than 400 registered users

> Local farm Hamburg

- About 450 cores, growing fast

> Local farm Zeuthen

- Upgrade from 600 -> 720 cores

> New parallel cluster Zeuthen

- Will extend and replace APEnext (summer 2010)
- Main usage LQCD, Astrophysics, external groups
- Installed 8 systems (6 in production), 16 blades each, 8 cores/blade (12 TFlops peak)
- Tightly connected using Infiniband
- Integrated into local farm, tight MPI integration into Sun GridEngine



High performance Computing: QPACE

➤ QPACE: LQCD optimized supercomputer

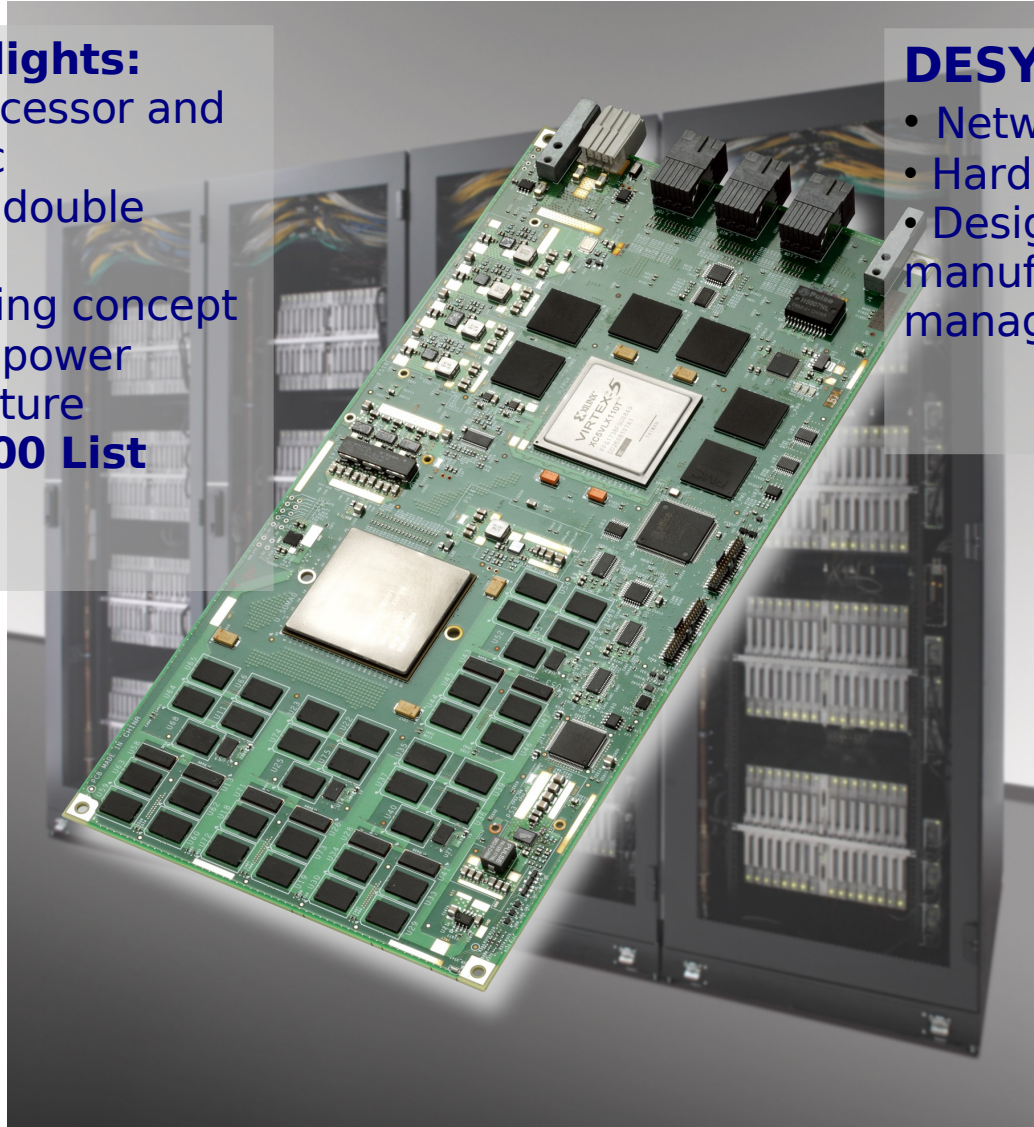
Technical highlights:

- Commodity processor and custom I/O fabric
- 26 Tflops/rack (double precision peak)
- Innovative cooling concept
- Currently most power efficient architecture
- **No 1: Green500 List**

DESY contribution

- Network design
- Hardware design
- Design lead, manufacturing management

Partners:
U Regensburg
U Wuppertal
FZ Jülich
U Ferrara
U Milano
IBM



> Installations

- National Analysis Facility: 13(9 Zn + 4 HH) servers with 100 (36 + 64) TB, clients and servers 1.8.2
- Local batch farm Zeuthen 14 servers with 181 TB, clients 1.8.0.1 (testing 1.8.2), servers 1.8.1.1 (testing 1.8.2)

> Problems: stability

- Client crashes
- clients tend to use much RAM and become slow
- not always clean reboots possible due to Lustre unmount problems
- Some bug fixes introduced new problems, new fixes + downtimes for users

> Problems: HW and OS compatibility

- HW support of some Sun hardware, maybe solved in 1.8.2 version, but got rid of Sun
- SL 5.4 kernel requires upgrade to 1.8.2 for Infiniband support
- No more SL 4 clients as 1.8.x not supported on that OS
- Each upgrade requires downtime and heavily affects users



- AFS-OSD SVN repository hosted at DESY
 - <https://svnsrv.desy.de/public/openafs-osd>
- Active Participation in development until 9/2009 (Felix Frank)
 - Test cell was operational, AFS-OSD with Lustre was demonstrated to work
- Currently development takes place mostly at Garching
 - Other sites showed interest, but no substantial contributions yet
 - DESY is lacking manpower to take part in development
- Testbed setup for AFS in Hamburg – ready for use
 - Planned for automated build and test of OpenAFS code
 - For testing and hardening OpenAFS-OSD code
- Decision concerning AFS-OSD future at DESY pending
 - Depends on commitments of other sites to support and use AFS-OSD



Windows

- Domain level is now Windows Server 2008 R2
 - Not only DC's but also most structure defining servers at this level
- Win7 rollout (64bit) in preparation for PC's and notebooks
 - Start of domain migration planned in 2010
 - Domain is technically ready for Win7
 - Group policies, admin consoles, ...
- Still evaluating new software distribution solutions (besides Netinstall)
- Integration of Windows 7 Printing into Samba service mostly done
- Hamburg continues with Netapp filer 3050 Cluster for domain storage
 - Stability, performance and cost are satisfactory
 - 38 TB capacity being upped to 48 TB in coming weeks
 - Investigating use of SAS disks (instead of FC) for further expansion
- Zeuthen uses new storage solution based on SANmelody (DataCore)
 - 2 head nodes (Dell) with 4TB each redundant storage, backbone using FC and iSCSI
 - Base for further highly reliable virtualizing solutions
 - Data migration from the old HP MSA1000 to SANmelody underway



Floor Space and Cooling

➤ Computer room RZ2 in HH now fully equipped

- 200 m² floor space, 58 water cooled cabinets
- In addition to RZ1 (700 m², 370 kW)
- currently 300kW for servers, 120 kW for cooling used

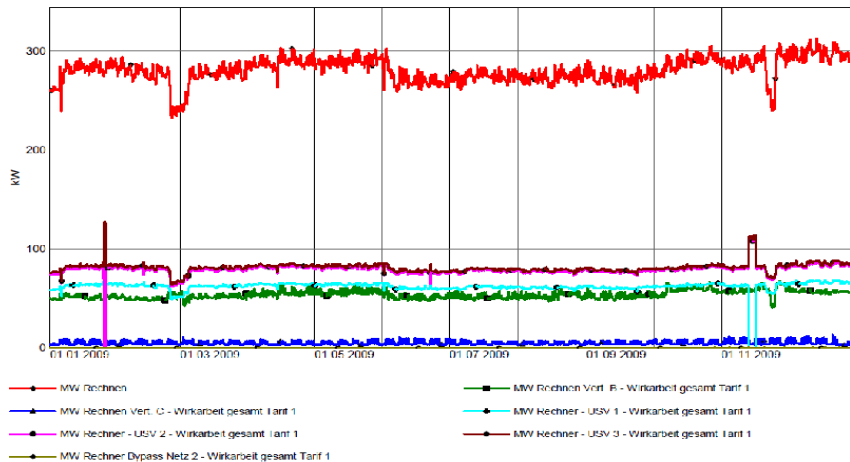
➤ New computer room is being built in ZN

- New floor space (160m²) for approx 35 cabinets on top of old one
- Need to remain within 300kW for whole computer center Zeuthen
- Several measures taken (water cooled cabinets, sandwich cooling)



FRAKO EMVIS-NET - Historische Daten (DESY)

Tageskurve Leistung von: Donnerstag, 1. Januar 2009 bis: Freitag, 1. Januar 2010



> Printing

- Cups now introduced in HH as well. Does not scale well with number of printers
- Introduction was labor intensive
- Leasing contracts for printers advantageous

> dCache

- Migration of the HERA dCache instance (from pNFS to Chimera) done
- Huge task, instance is the site with largest number of files world wide
- Migration effort mainly determined by number of files

> Subversion

- service for DESY (introduced in 2008) well accepted

> Mail

- Unix mail server: Migration of 500 mailboxes from UW-IMAP to dovecot went well
- Dovecot now in production at Zeuthen, server side filtering with sieve
- Investigating new central mail solution with Groupware functionality



> Hardware problems

- Fewer broken disks (partly due to usage of 2.5" SAS disks)
- Dying DIMMS (excessive single bit error rate, computer may operate at small fraction of nominal speed) need to be identified and exchanged, very labor intense

> Mac usage

- Increasing number of Mac users led to activities within computing groups
- 1st user meeting, collecting wishes and requirements
- Limited support for Mac users established



Thank you for your attention

Questions/Remarks ?

