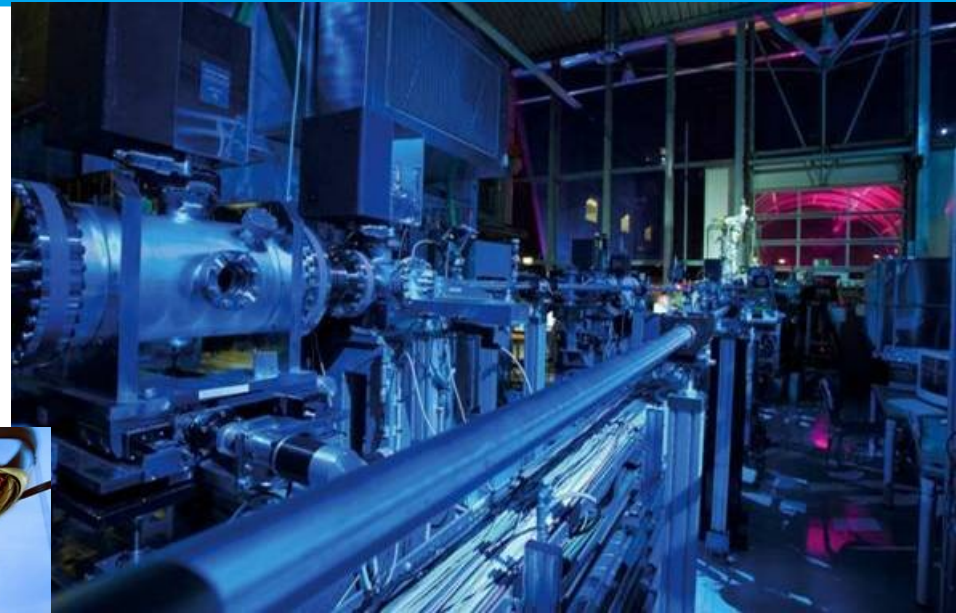


DESY Site Report.

HEPiX Spring meeting 2011



Wolfgang Friebel
HEPiX
Darmstadt, May 2011

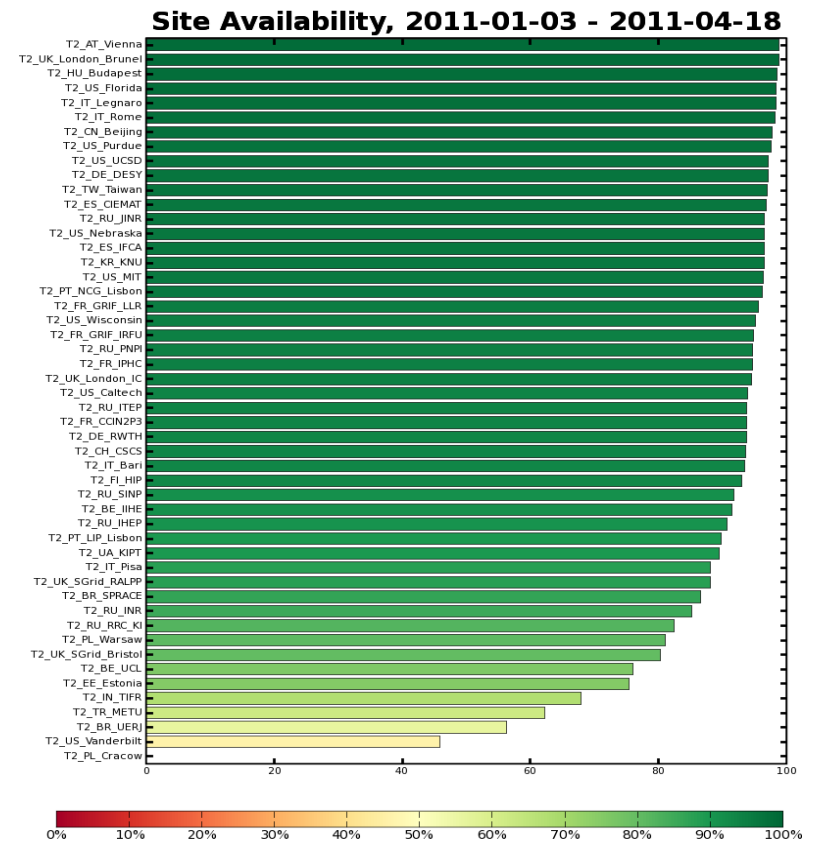
Grid activities

> Hardware

- About 5000 cores for the grid (4224HH + 784 Zn)
- About 4 PB (2.3 PB + 1.4 PB Zn + 0.3 PB NAF) accessible in Grid

> Supported VOs

- Atlas, CMS, LHCb: Tier2 centers
- Zeus, Hermes, H1: Monte Carlo simulations
- IceCube: Tier 1
- ILC, Calice
- CTA: Monte Carlo simulations
- LQCD: Lattice Data Grid
- ... and many more



Local farms

➤ National analysis facility

- 1500 -> 1620 cores, currently 540 registered users

➤ Local farm Hamburg

- 450 -> 500 cores

➤ Local farm Zeuthen

- 720 -> 936 cores

➤ Using Gridengine batch system

- UNIVA is (also) maintaining the open version of GE
- For details see <http://gridengine.org/blog/faq/>
- Talking about support contract with UNIVA
- Commercial version GE 8.0 released (compatible but enhanced open GE)



High performance Computing

➤ Long tradition with special purpose LQCD computers

- 1994: APE100 45 GFlops installed
- 1999: APEMille 550 Gflops
- 2005 APENext 2.5 Tflops, decommissioned in March
- 2010 QPACE 26 Tflops/rack, major DESY contributions installed elsewhere



➤ Conventional parallel computing

- 1152 (9*16*8) tightly connected cores 13.5 Tflops peak performance
 - QDR Infiniband interconnect using Mellanox cards
 - switches for fat tree between blade enclosures currently being installed
 - Integrated into local farm, tight MPI integration into Sun GridEngine
 - shared by several groups
 - 100 TB Lustre storage accessed over DDR IB from cluster
- Can also be accessed over Ethernet from farm and workgroup servers



High performance computing using GPUs

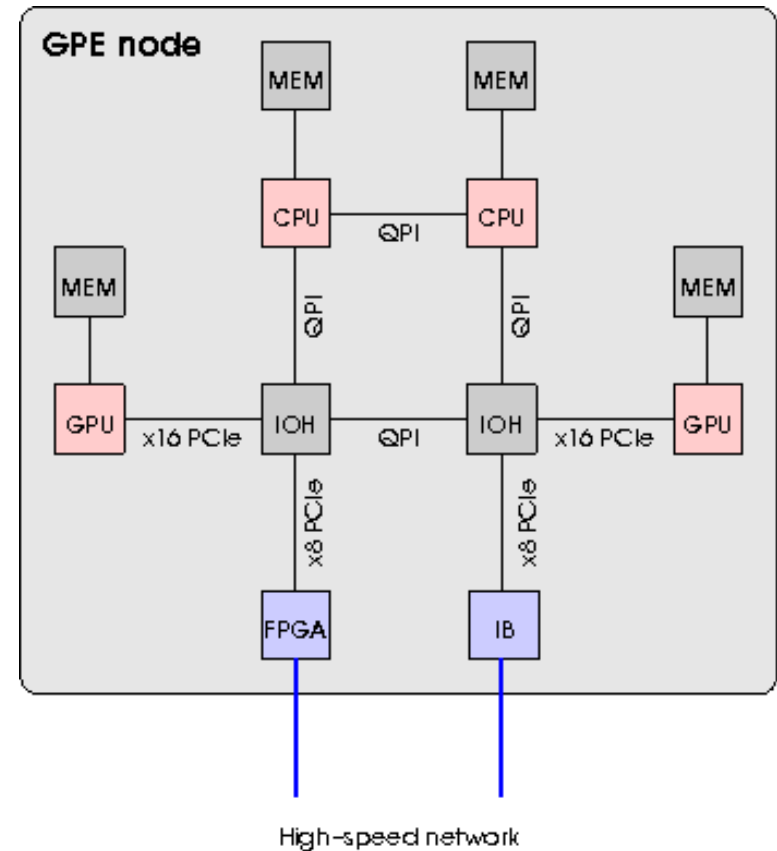
➤ GPU testbed in Zeuthen using two machines

- 2 Intel X5667 CPU Quadcore 3.07 GHz
- 2 Intel 5520 IOH Rev 22
- PCIe with 3 ports, 36 channels
- 2 NVIDIA C2050 boards
- 3GB GDDR5
- Altera FPGA development kit
- Mellanox IB HCA

➤ Peak 1TFlop single precision/machine

➤ GPUs for Photon science:

- pilot systems have been installed there as well

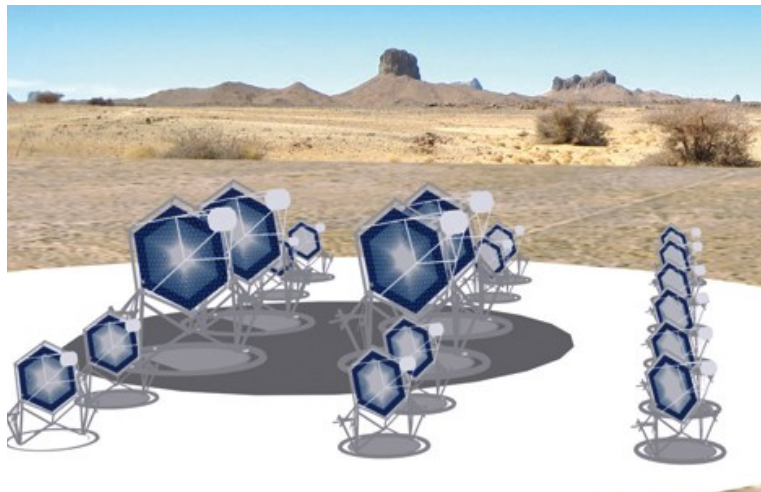


➤ IceCube

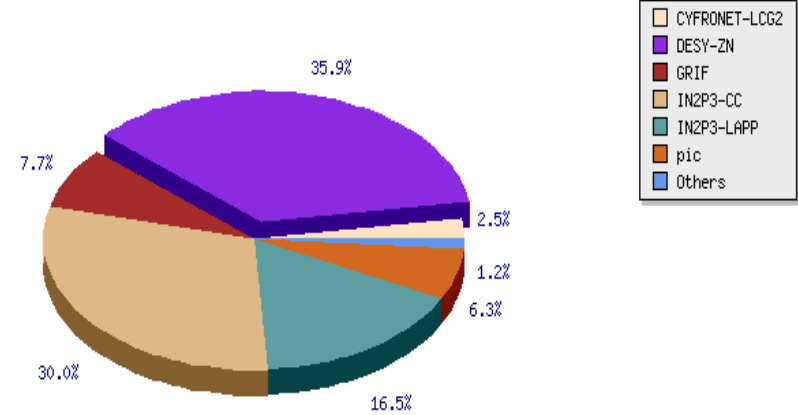
- IceCube completed Dec 18, 2010, all strings deployed
- Zeuthen is European Tier 1 for IceCube

➤ Cherenkov Telescope Array Project

- Biggest Monte Carlo Production grid site



vo.cta.in2p3.fr VO Normalised Elapsed time (HEPSPEC06) per SITE
Production Sites. October 2010 - March 2011



(C) CESGA 'EGI View': vo.cta.in2p3.fr VO / normetop-HEPSPEC06 / 2010-10-2011-3 / SITE-DATE / Production / ACCBAR-LIN 2011-04-18 19:53 UTC

- Focusing primarily on data management and support for scientific computing for the PS communities
 - common data management requirements have been collected
 - a prototype portal application for data management and data access is being developed (tightly connected to the DOOR User Portal for Photon Science at DESY)
 - a new dCache instance for PS is in preparation
 - both to be available to users in May or June
- A machine has been installed for CFEL data reconstruction and analysis
 - SGI with 12 hexacore CPUs @ 2.7 GHz and 384 GB of RAM
 - .5 PB of storage on DataDirectNetworks storage array
- Looking into identity management issues for thousands of new users
 - typically very many small VOs both at DESY, in Germany and on an European scale.
 - Many activities regarding federated identity management are afoot in these communities.



- Majority of systems at DESY running SL 5(.5) 64 bit
- SL6 deployment at DESY
 - Work on SL6 integration began in April 2010
 - By November 2010 installation and maintenance process for SL6 hosts completed
 - By now several desktops, servers and virtual machines running SL6
 - Waiting for Lustre clients on batch nodes and workgroup servers
- Major work areas for the SL6 deployment
 - openafs 1.6 RPM provided
 - Setup and maintenance of KVM based virtual machines
 - Building kernel modules using kabi (nvidia/CUDA, openafs)
 - Adaptation of the automated SL6 installation to the changed software stack, e.g. ext4 file system, rsyslog, changes in X11 configuration process, ...



Storage

- See separate talk 'The DESY Grid Lab'
- Looking for pNFS/NFS 4.1 storage solutions
 - dCache
 - IBM SONAS
 - NetApp Cluster Mode solutions
 - Plain GPFS probably too complex to get optimum performance with minimal manpower



- Zeuthen continues to use storage bricks (commodity hardware)

- ongoing AFS troubles are being investigated
 - Symptoms: short (several minutes) freeze of many clients when a directory is simultaneously accessed in read and write mode by hundreds of clients (lock up of AFS cell due to CB issues)
 - remedies suggested by the OpenAFS developers are being perused
 - partial success using their patches as well as patches from CERN
 - no definitive solution reached so far
- openafs RPMs (1.4 and 1.6) provided for Scientific Linux
- No current AFS-OSD activities (SVN repository hosted at DESY)
 - waiting for a deployable development branch w/ OSD (2.0)



Storage - Lustre

- Lustre version 1.8.5 on SL5 64 bit
- Evaluation on SL6 - not yet stable enough
 - Whamcloud's development looks promising
- Installed storage
 - Zeuthen: 385 TB, more ordered by the experiments
 - NAF: 275 TB (170 TB HH + 105 TB Zn)
 - connected with DDR Infiniband or bonded 2x1GB ethernet



Networking, IPv6 activities

- Taking part in LHCone (dedicated network to connect Tier1's and Tier2's)
 - DESY has taken the leading role for LHCone in Germany
- Cisco Firewall replaced by Juniper Networks
 - led to increase in bandwidth
- IPv6 capable HW existing (Catalyst 6500-E)
 - separate networks in HH and Zn as a first step
 - installation of several machines planned for application tests
- Tests being performed, achieved thus far:
 - IPv6 connection to the “DFN-Verein”
 - IPv6 on two networks for testing purposes
 - IPv6 enabled DNS-Server
- Focus on connectivity tests, not yet applications
 - Several decisions still to be made, e.g. address assignment by DHCPv6 or RA



Miscellaneous

- Looking for an Exchange 2003 replacement (groupware, calendaring,...)
 - Candidates: OpenExchange, Exchange2010, Oracle Beehive, Zimbra
 - No decision yet
- Migration of mail accounts from UW-IMAP to dovecot in HH underway
- Moving to WebLogic as application Server in Oracle 11g suite
 - Used internally (accelerator controls, administration)
 - move causes adaptation work by users of the application server



- Many Services virtualized with HYPER-V and XEN
- Office 2010 rollout on XP nearly completed without big problems
- Windows 7
 - Rollout will start in May
 - Problems with Samba Printing (64 bit printer drivers on Samba 3.5 and 3.6)
 - better experiences with older releases (3.3.8 and 3.3.10)
- Netapp is used for File service
 - Windows 7/Windows Server 2008 R2 SP1 compatibility issues required a NetApp filer upgrade (OnTap release 7.3.5.1), which was tedious
- Thoughts on using Forefront Endpoint Protection
 - Who has experiences?

