



DESY Site Report

HEPIX Spring 2006

CASPUR, Rome

Dirk Jahnke-Zumbusch

compiled by Peter van der Reest

DESY – IT

- Communications: X-Win, IP Telephony
- User Support: DOOR and Indico
- Systems: apeNEXT, Disk Experiences, SLD, Quattor
- Operations: second computing centre
- Services: WebOffice, AFS Backup, Oracle, Enterprise Storage
- GRID: HERA, ILDG, ILC, EGEE, D-GRID, LCG: CMS & ATLAS

- Wide Area Network
 - since Jan. 1st 2006 DESY-HH is X-WiN core node (X-WiN = German 10 Gb/s research network)
 - planned 10 Gb/s private link to FZK, 1 Gb/s link to DESY-ZN
- Telephony
 - DESY is fully migrating to a VoIP system for office use no more investments in classical phone exchange systems
 - only for emergency use and non-networked areas classical phones will remain available

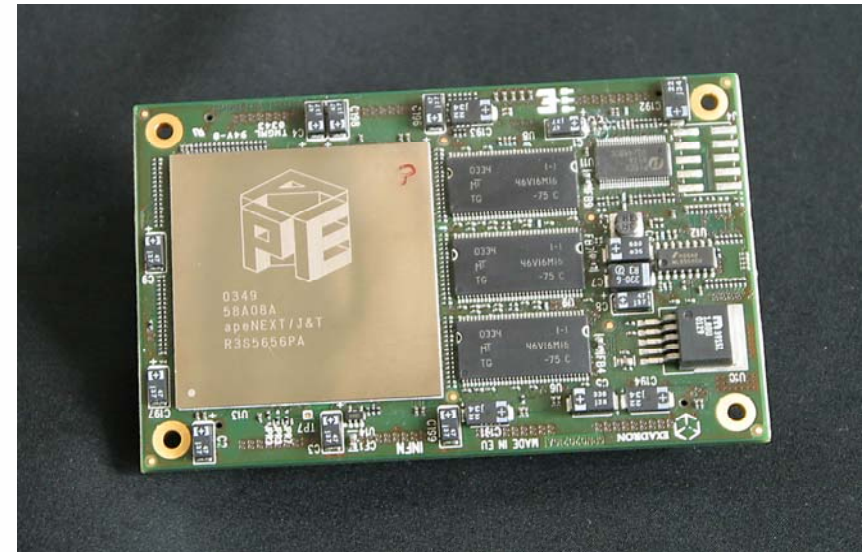
■ DOOR

- the „**DESY Online Office for Research with photons**“ has been established in close cooperation with HASYLAB
- based on DUO from Paul Scherrer Institute it implements workflows
- used for project proposal & review, beam time application & allocation as well as project reporting

■ INDICO

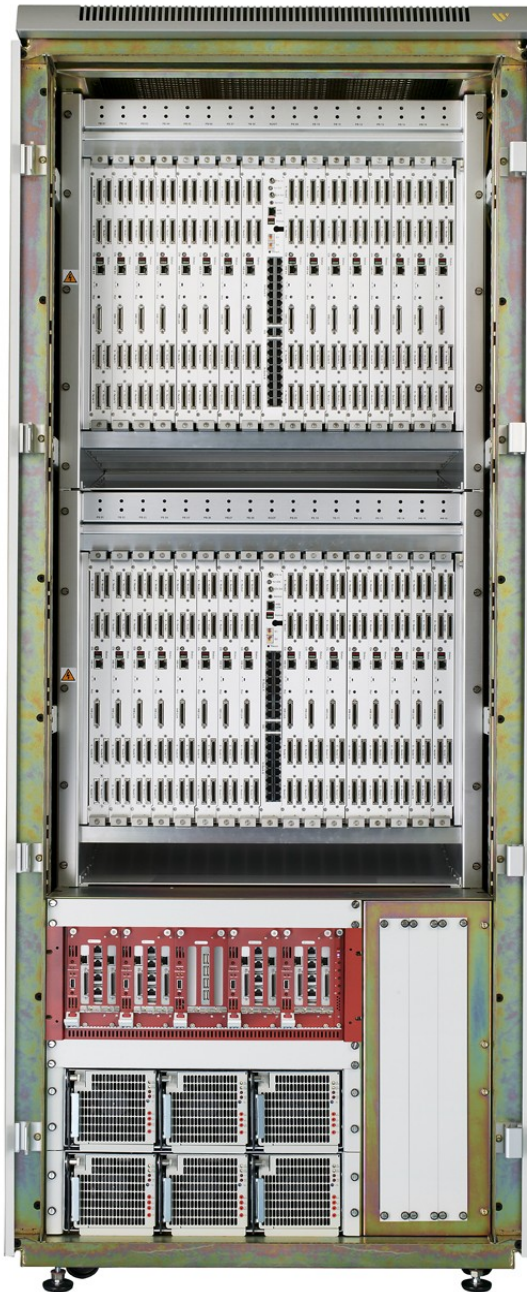
- the CERN Integrated Digital Conferencing system has been introduced at DESY as a centrally supported service
- after first experiences we will try to provide for DESY specific needs
- in cooperation and exchange with CERN

- Project goal: Build custom designed *TFlops-computer with 50% sustained performance* for key lattice gauge theory applications
- Collaboration: INFN (I), DESY (D), U Orsay (F)
- Processor features:
 - System-on-chip design
 - Large memory bandwidth
 - Large bandwidth, very low latency network
- System features:
 - High density
 - Low power-consumption



contributed by D. Pleiter, ZN-DV

apeNEXT (2)



- Currently deploying 25 racks
- Status at DESY
 - 3 racks installed(1 rack = 512 CPU = 0.6 TFlops)
 - Tested by long (3-20 h) physics application programs running in replication mode
 - Stable operation
(mean time between interventions: >1 weeks)
 - Soon: transition to user operation
- Application kernel performance
 - 30-55% for assembly routines
 - 30-45% for high level languages (to be improved)

- after having dealt with a multitude of SATA-Disk related problems, we moved to using 'higher quality' SATA units
 - server edition & nearline are commonly used descriptions
- in one case we now found the same silent data corruptions as before (data returned is not data written/requested)
- maybe the quality of factory testing is not as good as manufacturers want us to believe...
- ...on the other hand, we only found it in one of 16 400GB disks...

for details, please contact stephan.wiesand@desy.de

- SLD3 has been adopted as replacement for SuSE based DL5
- SL4 is currently recommended for notebooks
- and heavily used for services e.g. dCache, web, LDAP
- after the announcement that LCG is going for SL4, we aim for a complete SLD4, both in serverland as well as on the desktop
- we would welcome adaptation of Quattor methods for desktop configuration and management by the community (of which we are an active member, so ...)
- ... DESY-IT has offered to host the next Quattor workshop in October 2006

Operations: 2nd computing center



HAMBURG · ZEUTHEN

- due to a thermal upper limit on the number of systems we can host in the main computer center
- and the need to expand for our various Grid activities, a new center is being built
 - in the basement of the current building
 - using water-cooled racks
 - laid out as a lights-out facility
- there should be room for more than 1000 boxes
- once we reach that number, we will have to start looking at expanding the UPS system to allow further growth

- WebOffice was founded three years ago as joint venture IT/PR
- selected, installed and operates the ZOPE ZMS content management system
- architected a high availability server environment
- provides service for user community: documentation, training, planning and skeleton implementations, as well as implementing functionality for new requirements
- has implemented some 40 active sites and is working on some 40 in the pipeline
- is happy to discuss their work with others: web-office@desy.de

- DESY-ZN-DV has developed a backup scheme for AFS
 - allows for full and incremental backups in configurable frequency
 - provides end user operations by means of arcx delegated authorisation
 - both full volume and single file restore is possible while respecting ACL
 - in production, disk caching of backup files results in high recovery speeds
- system is operational at DESY in Zeuthen
 - backup since April 2005, recovery procedures since March 2006
- dependant on file-based backup mechanisms only, therefore easily adapted to your backup system

- Oracle
 - IT is moving towards Oracle 10g RAC – high availability RDBMS
 - migrating old users that are able and willing from FORMS to IAS apps
 - urging new users not to use FORMS

- SAN & enterprise storage
 - expanding fabric based storage for rapid response to ad-hoc requirements
 - while separating areas where there is no need for sharing fabrics
Windows, Unix core services, TSM backup and physics mass storage
are in four distinct fibre channel fabrics



- Grid activities were initially driven by the demand for resources for MonteCarlo production of the HERA experiments H1 and ZEUS
- The International Lattice Data Grid (ILDG) and the International Linear Collider Community (ILC) joined the Grid activities
- DESY participates in EGEE (since 2004) and D-GRID (since 2005)
- DESY sets up LCG Tier-2 centres for ATLAS (in federation w/ U Freiburg) and CMS (in federation w/ RWTH Aachen)
- DESY participated in the Service Challenges SC3 and prepares for SC4 and the Pre-Production Service PPS for LHC
- DESY operates a complete Grid infrastructure incl. all services

Grid @ DESY (2)



HAMBURG · ZEUTHEN

VOs hosted at DESY:

- Global: hone, ilc, zeus (registration via LCG registrar system)
- Regional: calice, dcms, ildg
- Local: baikal, desy, herab, hermes, icecube

VOs supported at DESY: (Tier-2)

- Global: atlas, cms, dteam
- Regional: dech

H1 Experiment at HERA (hone)

- desy, uni-dortmund, cscs, gridpp, bham, ucl, lancs, ox, marseille, cyf-kr, saske

ILC Community (ilc, calice)

- desy, uni-freiburg, bham, ox, ral, ic, qmul, lal, ed
- contact to SLAC and FermiLab

ZEUS Experiment at HERA (zeus)

- desy, uni-dortmund, gridpp, bham, ucl, lancs, ox, marseille, cyf-kr, saske, infn, utoronto, uam, scotgrid, weizmann, scai, bris, tau, ed

Grid @ DESY (3)



HAMBURG · ZEUTHEN

- Tier2 activities are shared between Hamburg and Zeuthen

- the certified DESY-ZN site is operational from April 1st, 2006
 - running with two CE: Torque and SGE based
 - using an adapted job manager lcgsgc.pm based on lcgpbs.pm
 - the systems are installed by RPM and YAIM only, no fabric management system in use

- the DESY-HH site uses Quattor for fabric management
- both sites are providing Tier2 services for CMS and ATLAS



¿ Questions ?

peter.van.der.reest@desy.de
dirk.jahnke@desy.de

Links



HAMBURG · ZEUTHEN

- <http://door.desy.de>
- <http://indico.desy.de>
- <http://www-zeuthen.desy.de/ape>
- <http://scientificlinux.desy.de>
- <http://wof.desy.de>
- <http://grid.desy.de>

DESY

AFS Backup

- backup of AFS volumes by creating snapshots using 'vos backup'
- frequency and level of backup determined by config file
- currently 3 levels of backup (full, level1, incremental) implemented
- converting .backup volume into ordinary file using 'vos dump'
- backup of the file
(backup date in the name and coded within file)

contributed by Wolfgang Friebel, ZN-DV

AFS Recovery



HAMBURG · ZEUTHEN

- command line interface accessible to all users (arcx client)
- recovery of whole volumes by name, get mounted in ~/Recover

or

- recovery of individual files into original location (with .R appended)
- all access rights to the volumes or files are strictly respected

contributed by Wolfgang Friebe, ZN-DV

■ Benefits

- backup frequency can be tailored to individual needs (volume patterns)
- recover speed is high by caching up to 1.5 TB of recent data
- Even faster when recovering individual files by decoding the AFS dumpfile structures

■ Status:

- New backup procedure in place since 04/05
- New restore procedure introduced in 03/06
- currently dependent on Legato Networker (backup, retrieve and index commands), easily adoptable to other backup systems

Grid @ DESY (4)



HAMBURG · ZEUTHEN

Quattor (SL 3.05 for all nodes; complete installation for WNs)

LCG-2_6_0, Yaim (for all service nodes)

Central VO Services: (*unique per VO*)

VO (LDAP) [grid-vo.desy.de]

VOMS [grid-voms.desy.de]

Catalogue Services: RLS / LFC [grid-cat/2.desy.de]

Distributed VO Services: (*mandatory per VO*)

Resource Broker (RB) [grid-rb0/2.desy.de]

Information Index (BDII) [grid-bdii.desy.de]

Proxy (PXY) [grid-pxy.desy.de]

contributed by A. Gellrich, HH-IT

Grid @ DESY (5)



HAMBURG · ZEUTHEN

Site Resources: [DESY-HH] (*multiple*)

GIIS: ldap://grid-giis.desy.de:2170/mds-vo-name=DESY-HH,o=grid

CE: 83 * XEON/2.8-3.06 GHz [grid-ce0.desy.de]

CE: 166 * Opteron/2.4 GHz [t2-ce0.desy.de]

CE: 34 * XEON/1GHz (ZEUS) [zeus-ce.desy.de]

SE: 75 GB; dCache-based w/ access to the full DESY data space

contributed by A. Gellrich, HH-IT