

Prague WLCG Tier-2 Report

- **Jan Kandrát**

- M. Lokajíček
- J. Chudoba
- T. Kouba
- J. Švec
- L. Fiala
- J. Horký
- M. Eliáš

Institute of Physics of the AS CR, v.v.i.



History: 2002



- 34 HP LP1000
 - 2x 1.3 GHz Pentium 3
 - 1 GB RAM
- First 1 TB of disks
- “Terminal room”

History: 2004

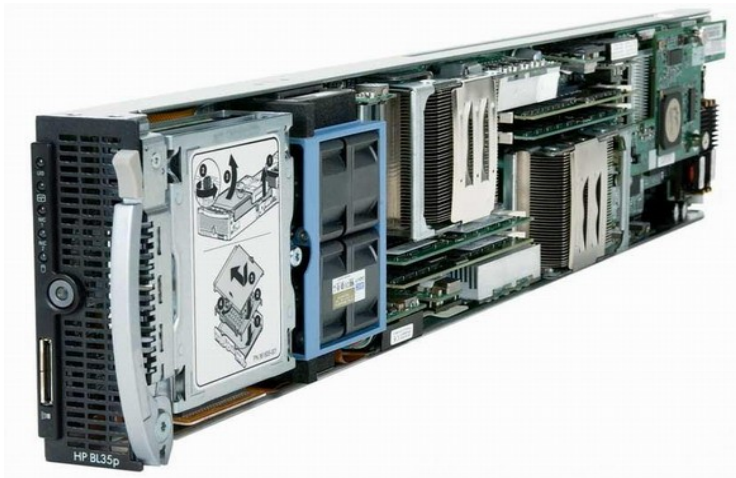


- A real data center
- 200 kVA UPS
 - 380 kVA diesel
- 2x 56 kW CRACs
- 67 HP DL140
 - 3.06 GHz Prescotts
- 10 TB disk storage

History: 2006 - 2007

- 2006

- 36 HP BL35p
- 6 HP BL20p



- 2007

- 9 HP BL460c
- 12 HP BL465c



History: 2006 – 2007 (cont.)

- 3 HP DL360 for Xen
 - FC infrastructure
- HP EVA
 - 70+ TB usable capacity, FATA drives
 - Disk images for Xen machines
 - Rest used as DPM
 - Warranty ends in Dec 2010
 - A new box cheaper than warranty extension

History: 2008

- 84x IBM iDataPlex node x340
 - 2x Xeon E5440 => 8 cores
- 20x Altix XE 310 twins (40 hosts)
 - 2x Xeon E5420 => 8 cores
- 3x Overland Ultamus 4800 (48TB raw each) SAN
- Tape library NEO 8000
- 2x VIA based NFS storage
- First decommissioning of computing nodes
 - 2002's HP LP1000r

History: 2009

- 65x IBM iDataPlex x360 nodes
 - 2x Xeon E5520 with HT => 16 Cores
- 9x Altix XE 340 twins (18 hosts)
 - 2x Xeon E5520 without HT => 8 cores
- All water cooled
- 3x Nexsan SataBeast (84TB raw each) SAN
- 3x Atom-based storage nodes
 - Wrong idea, WD15EADS-00P8B0 are just weird

Cooling & Power Problems

- High UPS utilization
- CRACs unevenly loaded
- Simple fans work surprisingly well
- New HW on unprotected power distribution



Cooling Problems - Long Term Solution

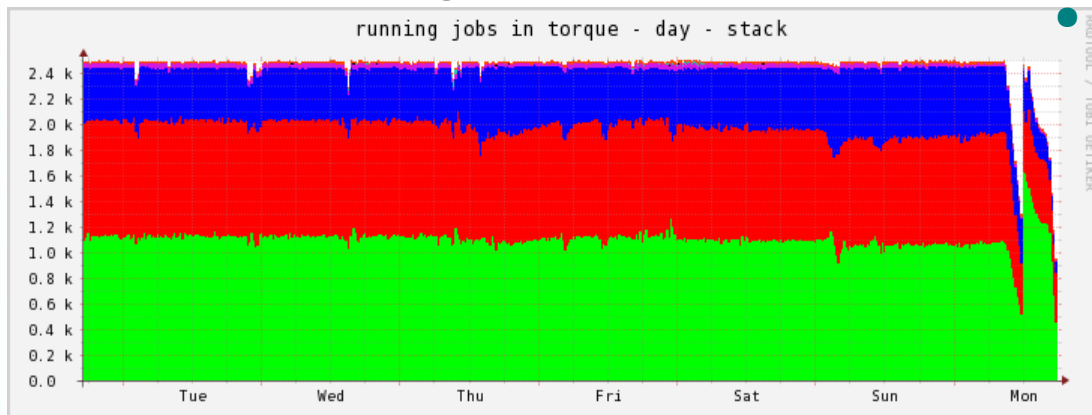
- 2009: New water cooling infrastructure
 - STULZ CLO 781A 2x 88 kW



- Additional ceiling for hot-cold air separation

Current WLCG numbers

- Atlas (T2 in DE cloud)
- Alice (T2)
- FNAL's D0
- Other EGEE Vos
 - Auger, Star
- 336 WNs
- 2630 cores (approx. 20.5k HEPSpec, 5M SI2K)
- Torque & Maui
- 520 TB on disk servers (DPM or NFSv3)



Network facilities

- External connectivity delivered by CESNET, a GEANT stakeholder
 - 1Gb to public Internet
 - 1Gb to FZK (Karlsruhe)
 - 10Gb connection demuxed to several dedicated 1Gb lines
 - FNAL, BNL, ASGC, Czech Tier-3s,...
- 10G internal network
 - Old machines still on 1G switches → LACP

2010: Current Procurements

- Estimates:
 - 1.5 PB for DPM
 - 4000 HEPSpec
- A “reliable disk array”
 - HP EVA replacement
- One more 10G Ethernet switch
- Additional HA servers etc.

Power Issues

- How to deal with a 100 A fuse blowing out



Power Issues (cont.)

- 30th Jun 2009 12:30: a 100 A fuse blown out
 - 17:00: fuse replaced, services up & running
- 7th Dec 2009 20:26:38: the same fuse blown out again
 - 21:26: we've noticed and went back on site to investigate
 - 22:30: problem understood, fixing stuff
 - ~ midnight: time to go home, critical services up, some machines turned off

Power Issues (cont.)

- 00:17:01: second 100 A fuse blown out
 - ...three minutes after the support went home
 - HP blade PSUs remained running on 2 phases
- 01:20: we noticed it's down again → downtime, controlled shutdown
- Most WNs turned off
- Fuses etc. replaced the next day



Power Issues: Lessons Learned

- Don't expect to run 96 A over a 100 A fuse indefinitely
- Act if the wires are getting warm on touch
- Beware of 3-phase PSUs
- Make sure that dual-PSU servers are on different phases
- Distribute the load evenly

HW Inventory

- In-house web-based inventory inconvenient to use
- Students from Charles University developing a CLI tool
 - IOS-like interface – we like CLI
 - Versioning, flat text dumps
 - Config for services generated from this DB
- More information in CHEP'10 paper

Q & A

Questions?