

CASPUR Site Report

Andrei Maslennikov
Lead - Systems

Rome, April 2006

Contents

- Update on central computers
- Batch news
- Storage news
- Projects 2006
- HEPiX services

Central computers

IBM SMP:

- Purchased a POWER-5 cluster (21 nodes, 168 p575 CPUs at 1.9 GHz and 400 GB of RAM)
- Communication subsystem: High Performance Switch (Federation)
- Plenty of problems while putting it in production:
 - 2 nodes required parts replacements
 - HPS was failing stress tests (HW issues)
 - buggy software (both AIX and cluster software issues)
- Now everything is solved, the system is running pre-production tests and is being tuned
- Old 80 POWER-4 CPUs at 1.1 GHz and 144 GB of RAM – will soon be decommissioned

HP SMP:

- One EV7 system with 32 CPUs at 1.15 GHz, RAM: 64 GB, Tru64 5.1B+
- Pretty stable

Opteron SMP:

- Getting very popular: 2 more clusters with 50 CPUs each, one with Infiniband, one with QSnet
- Most probably this area will be growing, the platform is very competitive

NEC SX-6:

- 8 CPUs, 64 GB of RAM

Batch news

- Since many years we are using SGE on all platforms, but this is now going to change
- Got impressed with PBS on Opteron clusters:
 - Better support for MPI jobs
 - Configuration is resource-based and allows for more flexibility
 - Fits very well with our new accounting scheme
 - Commercial variant (and hence support) available: PBSpro
 - May run on all our platforms
- Now evaluating PBSpro on our new PWR5 cluster

Storage news

Decommissioned IBM SANFS (StorTank) and now moving to GPFS:

- Simultaneous problems on both MDS units, metadata lost (of course, we had a backup and immediately brought the data online on NFS)
- This coincided with arrival of PWR5 cluster with DS4800 disk system (20TB)
- First benchmarks of DS4800: 750 MB/sec aggregate
- GPFS has a small overhead and may operate in the range 600-700 Mb/sec on our hw
- Already recycled all StorTank hw base (disks and machines)

Purchased 2 new powerful NFS servers (CERN disk server with R6)

- Currently under stress test, will shortly be put in production

Purchased 2 new IFT disk systems (G2422, R6)

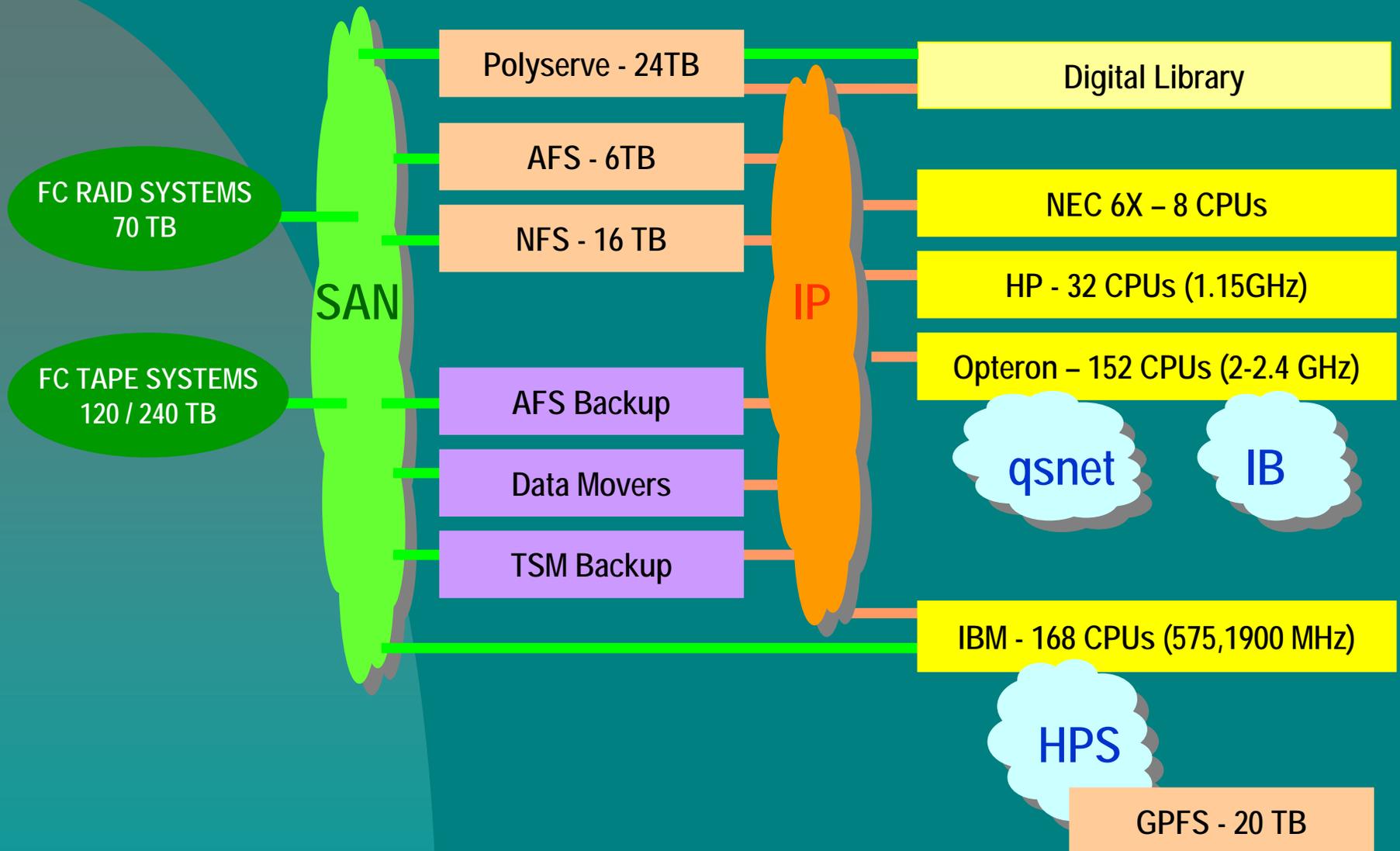
- Currently being evaluated, will replace AFS RAID-5 arrays

Tapes: some upgrades

- Replaced the remaining LTO-1 units with LTO-3
- Data migration in progress

SAN: migrated from Brocade to Qlogic, 2 new 5600 switches at 4 Gbit

CASPUR: principal resources in 2006



Some projects, 2006

Technology tracking (in collab. with CERN and other centers) – 0.5 FTE

- Just renewed the lab, tests in progress, plan to report at JLAB
- New R6 devices (disk arrays and PCI boards)
- Fast interconnects (10Gbit and IB)
- Distributed file systems (new and updated solutions):
 - GFS
 - PVFS2
 - Lustre
 - Terragrid
 - GPFS
 - StorNext
 - ... etc
- New appliances like Open-E

AFS/OSD (in collaboration with CERN, ENEA and RZ Garching) - 2.2 FTE

- Implementation of an Object Shared Device (OSD) in accordance with T10 specs
- OSD integration with AFS
- Progressing reasonably, v 1.0 in August, will be reported during Storage Day

HEPiX services

As was agreed shortly after Karlsruhe meeting:

- Put in place a new K5 domain (HEPIX.ORG) and a new AFS cell (/afs/hepik.org)
- Partial archive of past meetings (not all yet collected)
- Photo archive
- Web access to AFS areas

- This service will be integrated with the SLAB in 2006
- Access granted to all HEP institutes

- Complementary to <http://www.hepik.org/> site