

PIC
port d'informació
científica

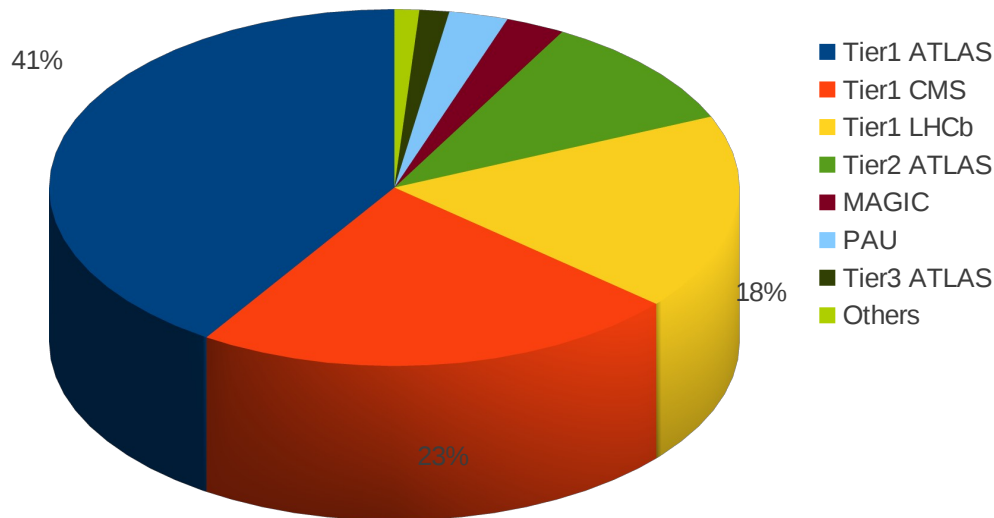
PIC Site Report Spring 2012

Gerard.Bernabeu@pic.es

- PIC is maintained through a collaboration agreement between CIEMAT (Ministerio de Economía y Competitividad), DECO (Generalitat de Catalunya), Institut de Física d'Altes Energies (IFAE) and UAB
- Data Processing support for 14 production VOs in HEP and Astroparticles, to more than a dozen Neuroimaging projects through a Web portal, and to a few technology projects, ranging from electric vehicle monitoring to simulations for a new PET device.

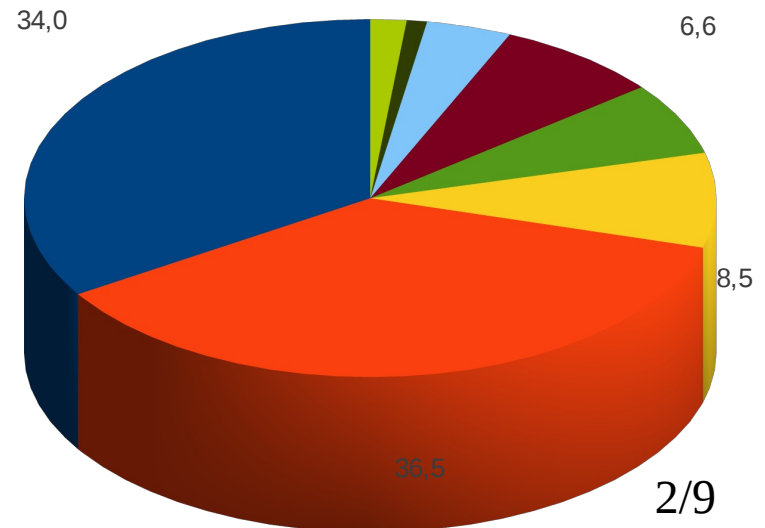
Batch System Fair Share Target

1% 1% 3% 3%
10%



Storage Capacity

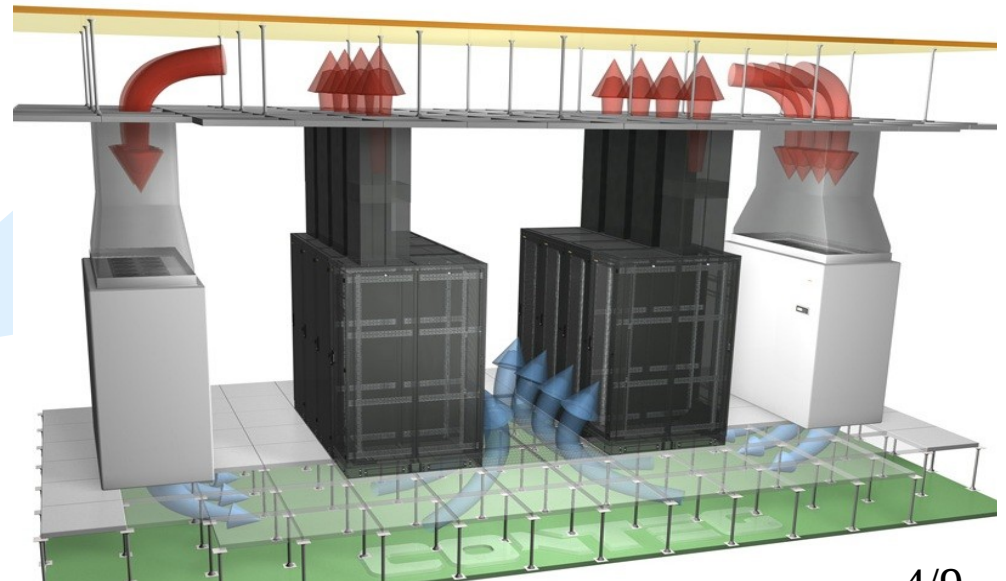
1,7 0,9 4,0
7,7



- **Services+Infrastructure teams consist of 10 people.**
 - Infrastructure (power, cooling, cabling, etc) engineer
 - Hardware and Infrastructure support technician
 - Disk Storage engineer and Production Coordinator*
 - Tape Storage engineer*
 - Virtualization and GRID engineer*
 - Network engineer*
 - Computing and GRID engineer
 - Databases engineer
 - 2 vacant positions
- **24/7 support based on weekly on call shifts by 4 trained engineers*.**

Infrastructure

- The previously reported infrastructure expansion, a module based on an AST-Modular Smart Shelter, is now full.
 - 80 KW IT, 1.55 PUE, 25 m²
 - serving 33456 HS06.
- Experimenting with chimney cooling racks in the main, traditional, machine room



Monitoring & Automation

- Moved from a centralized Nagios to 3 distributed Icinga with a joint view by Thruk
- Basic server configuration (DNS, users, bonding, RAID tools, etc) and ad-hoc tunning automated with Puppet
- Most services already deployed and maintained with Puppet.
 - Moving to parametrized classes
- Next is to focus in automating (multi-vendor) network equipment configuration

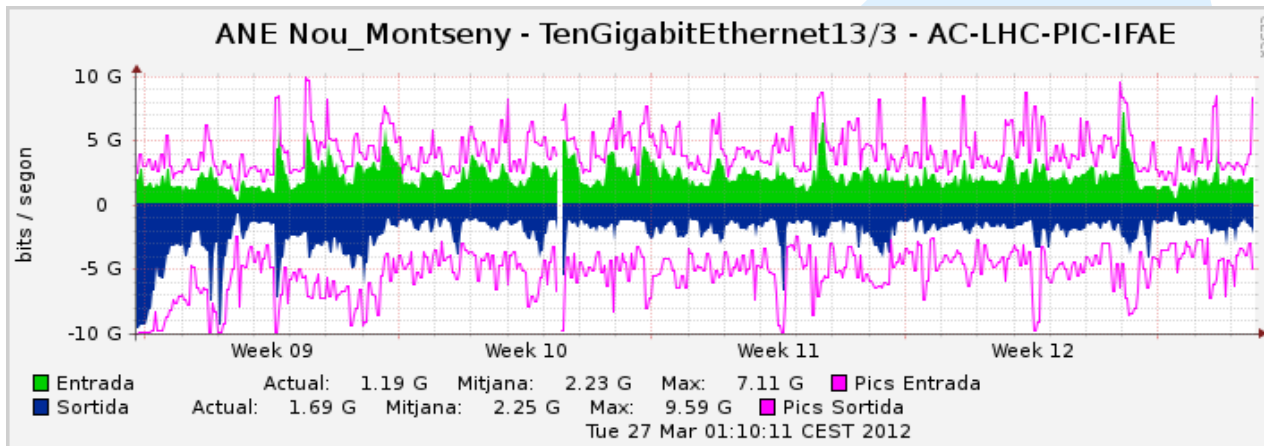
General use resources

Virtualization system

- OVM cluster running +100 virtual servers in different instances
- +90% of test servers are virtual, more than 50 production servers too.
- Looking forward to expand with a new, independent system

Network

- Ran out of 10GE slots => New Nexus 7009 for up to 336 wirespeed 10GE.
- WAN is sometimes reaching limits with 1*10GE+2*1GE. Ready to upgrade to 2*10GE when necessary.



Computing

New

- C6100 dual-twins -> 15% more cost effective.
But replacing RAM in 1 server requires shutdown of all 4 in the same chassis.

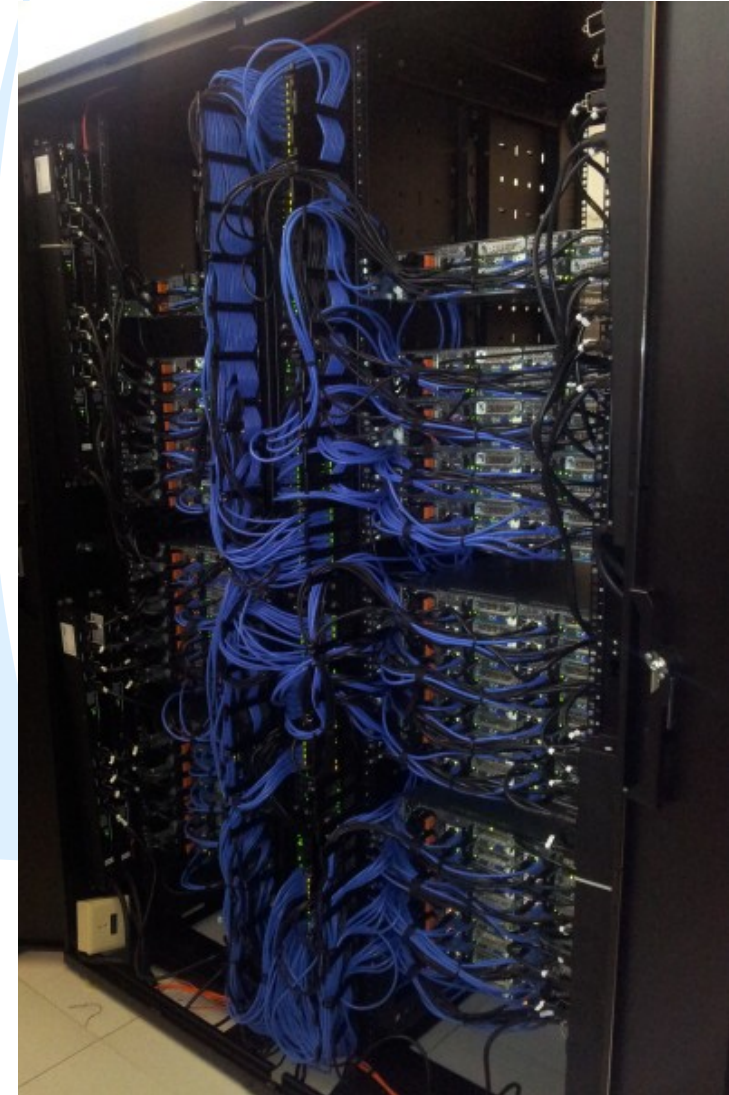
Many many cables - both power and network.
Bought two (75cm wide) SIEMON racks with cable handling accessories

Total

- 4260 Intel cores
- 50362 HS06

Next steps

- Trying to run a single PBS instance supporting more than 12k jobs -> new HW & multi-core jobs!
- Looking forward at a more scalable batch system (SLURM, SGE)



Tape Storage - FNAL Enstore

New

- 48% increase with 3,3PB in 10TKC tapes (6 drives)
- Tape server consolidation (4 drives per new server)
 - 7,5KW to 1,2KW
- Upgraded to Enstore2
 - Support for T10kC max capacity, enhanced scheduler

Total

- 6,8 PB
 - LTO3 (1446 tapes ReadOnly, 5 drives)
 - LTO4 (3611 tapes, 18 drives)
 - LTO5 (920 tapes, 4 drives)
 - T10kC (377 tapes, 6 drives)

Next steps

- PNFS to Chimera
- Small files Enstore's new feature



Disk Storage - dCache



New

- 40% increase with 1,6PB in 8 servers
- SuperMicro 4U servers with additional 4U "backpack" connected via SAS expander with Adaptec 6445 controllers -> many RAID6 issues (even data corruption)

Performance is good (see graph) but we're still debugging the corruption issue with Adaptec

- MAGIC using dCache's HTTP/WebDav door in production

Total

- 5,7 PB on SATA disks (1-3TB/disk)

Next Steps

- PNFS to Chimera
- NFS4.1

