

PIC
port d'informació
científica

PIC: Helping to turn information into knowledge

Prof. Manuel Delfino, Director of PIC
EuroCRIS 2015 national session
Barcelona, November 9th, 2015

Our funding bodies

- PIC is the result of a coordinated funding effort by:



with additional funds from:



Unión Europea

Fondo Europeo
de Desarrollo Regional
“Una manera de hacer Europa”

Our agencies

- PIC's personnel and expenditures are handled through:

Institut de Física
d'Altes Energies



Ciemat

Centro de Investigaciones
Energéticas, Medioambientales
y Tecnológicas

with additional infrastructure help from:

UAB
Universitat Autònoma
de Barcelona



Red **IRIS**

Main PIC activities

- Particle physics
 - First level center for CERN's Large Hadron Collider (out of 13 around the world)
 - Analysis data center for Neutrino data from Japan
- Astrophysics and Cosmology
 - Custodial data center for:
 - MAGIC Cherenkov gamma-ray telescope
 - Physics of the Accelerating Universe dark energy survey
 - Supercomputer simulations of the expansion of the Universe, star evolution, aerodynamic turbulent flow, neuroimages
 - Science Data Center for ESA EUCLID satellite
- Support for smaller projects in a variety of fields, ranging from neuroimaging to electric vehicle telemetry

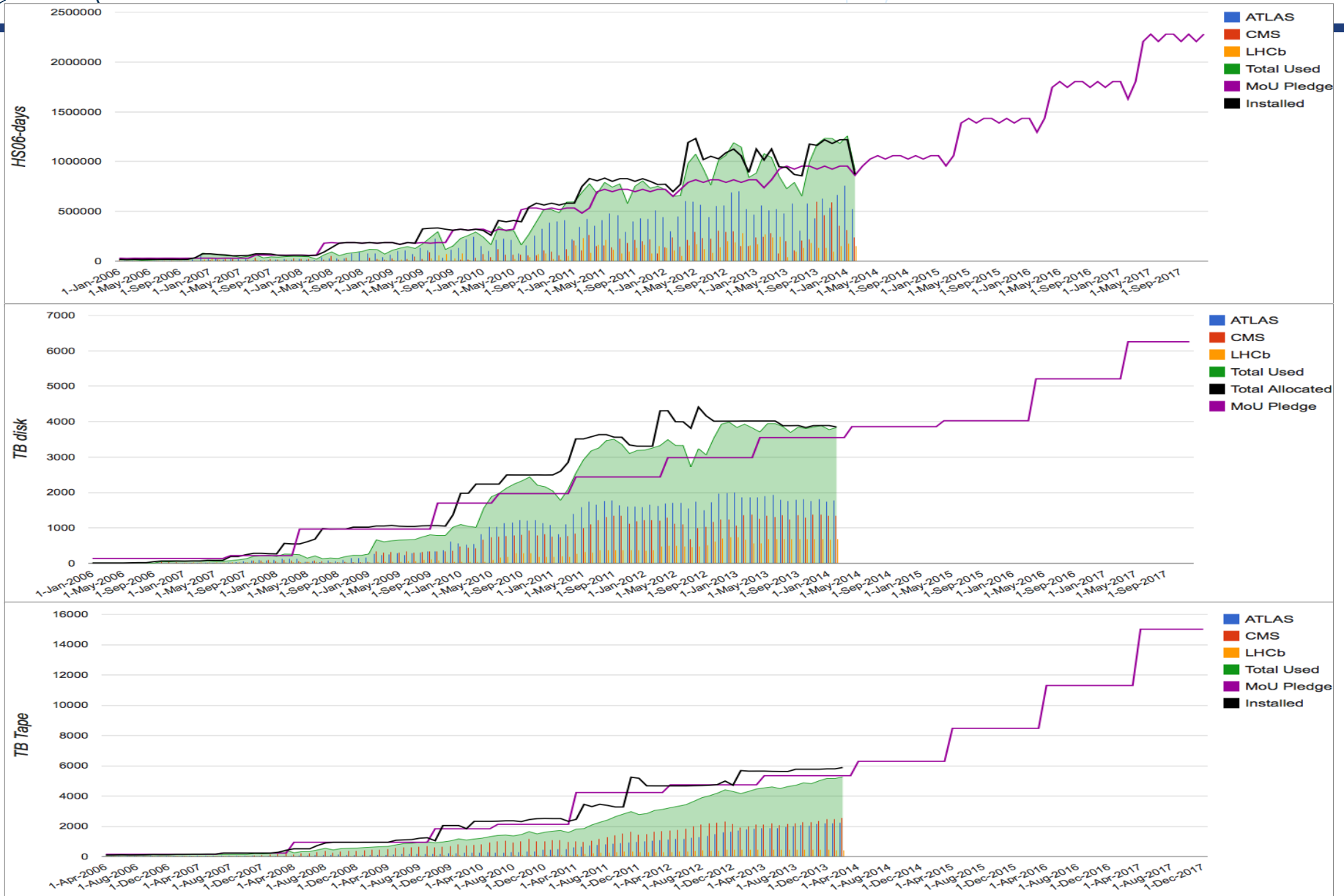
Our most important asset: the PIC team



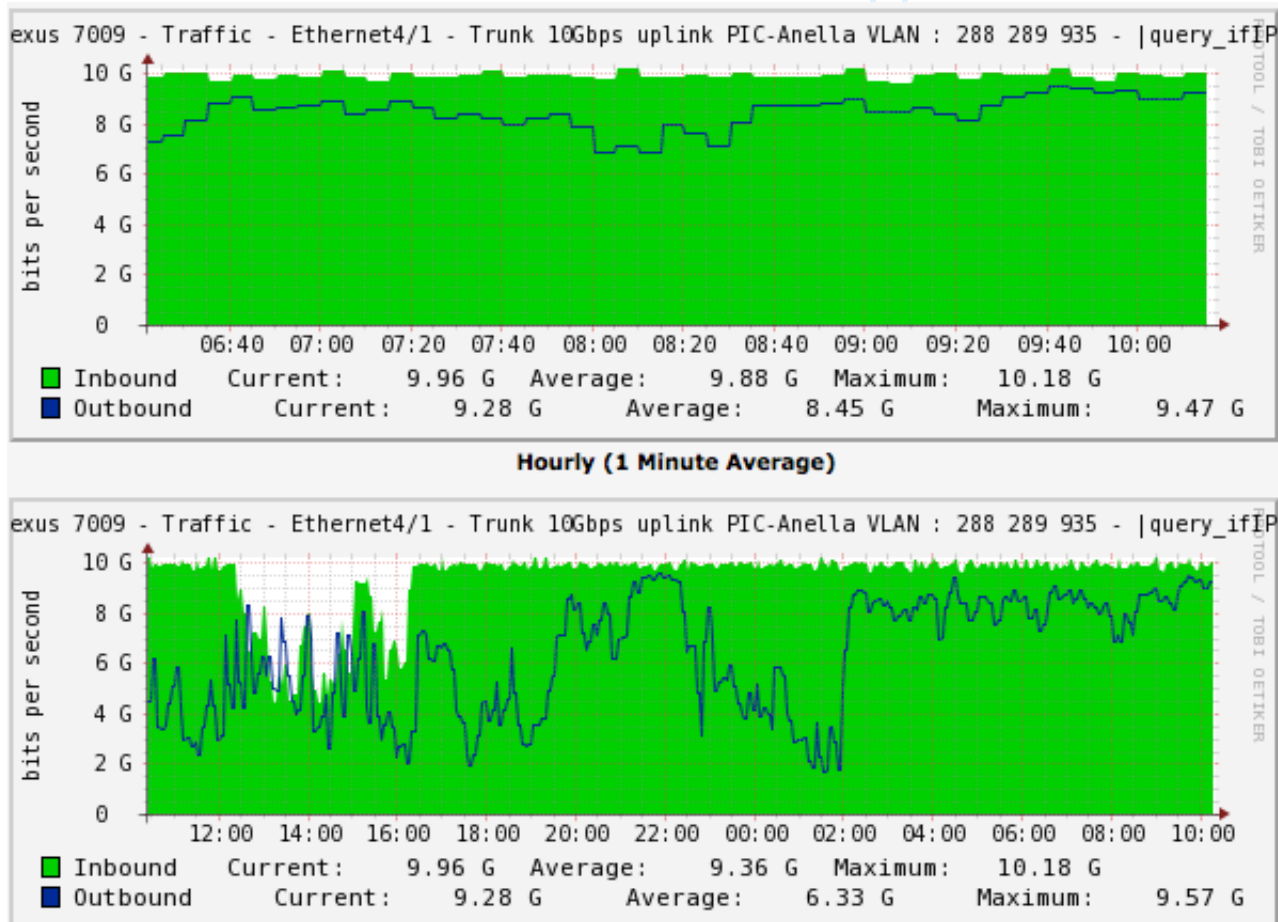
Resources at PIC

- Average growth rate over 10 years: 37% year-to-year
- About 4000 x86 cores in “farm” configuration
- Batch system can schedule “multi-core” jobs
- Over 6 Petabytes of spinning disk – NAS model
- Over 12 Petabytes of active tape – LTO4, LTO5, T10KC/D – SL 8500 library
- Glued by 10 Gbps Ethernet. Also 10 Gbps WAN link
- Major investment in automation and monitoring: Puppet, Icinga, Ganglia, gateways to infrastructure
- Virtualization platform for internal services

Resources delivered/planned for LHC



10 Gbps network flat-out



LHC computing system

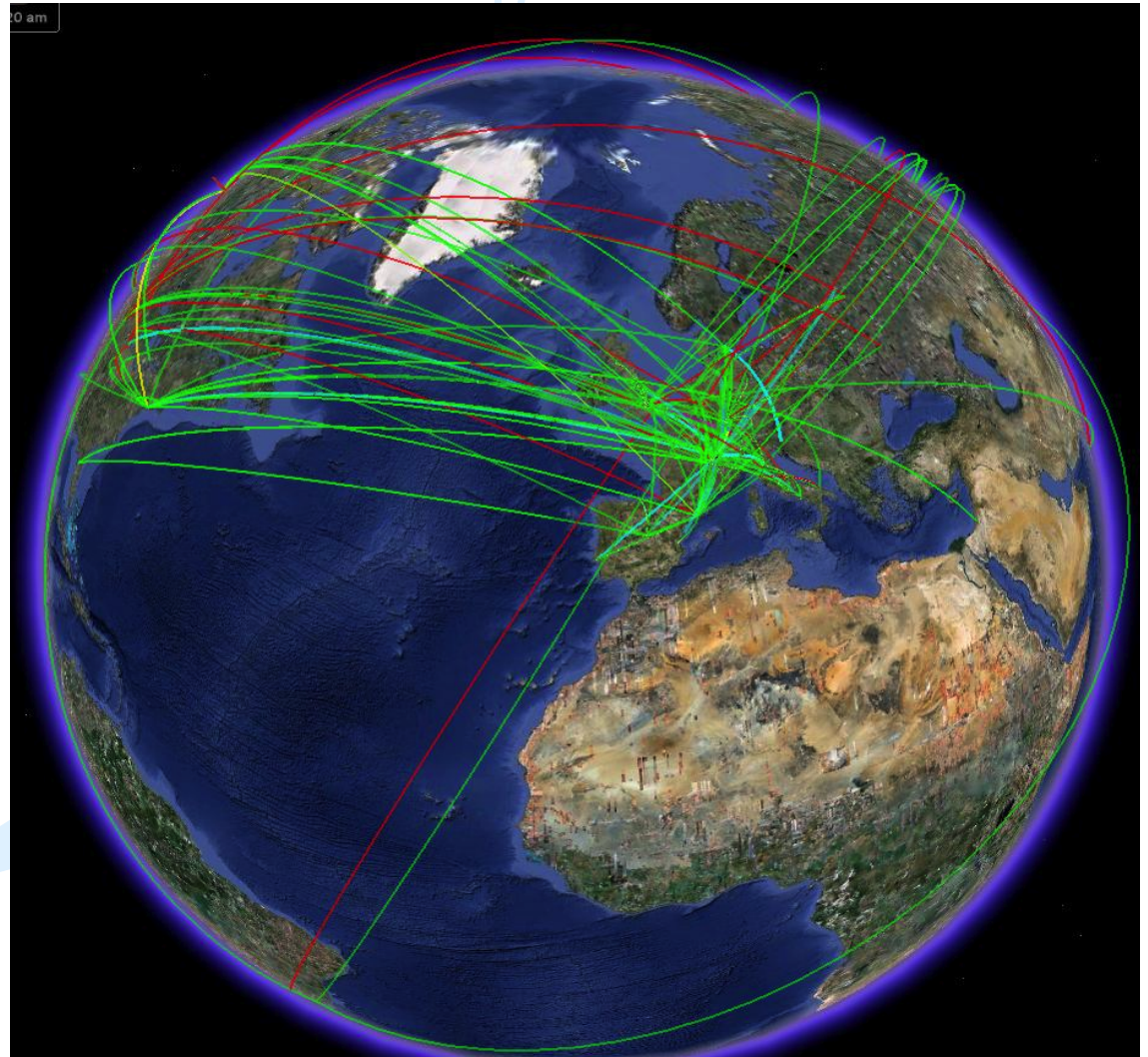
Integrated, distributed system

300k cores

10 Gbytes/s global dataflow

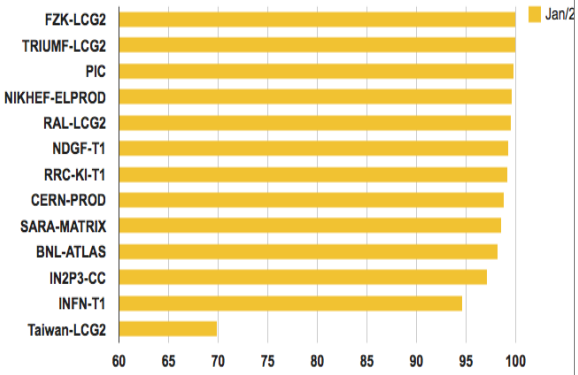
170 sites in 42 countries

Completes one job every 5s

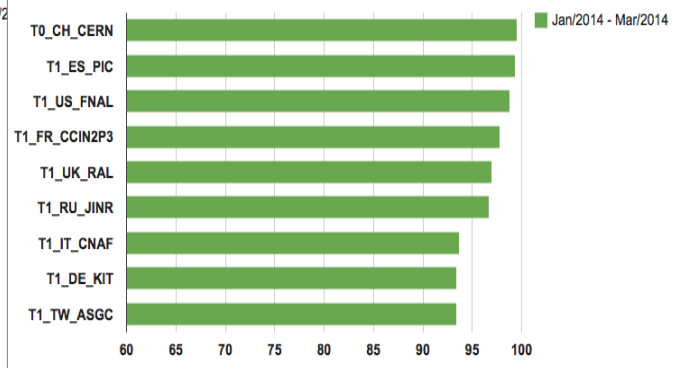


Reliability

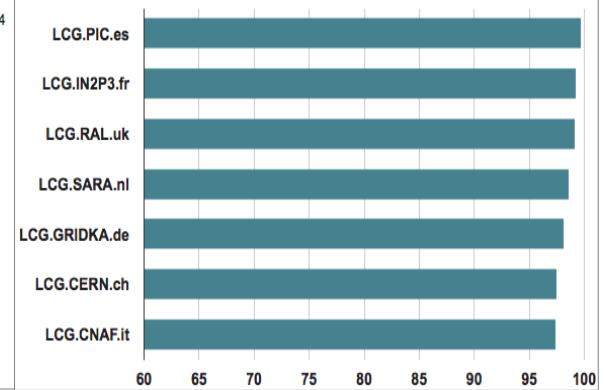
ATLAS Reliability Ranking



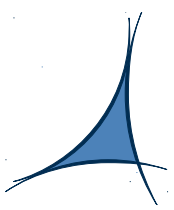
CMS Reliability Ranking



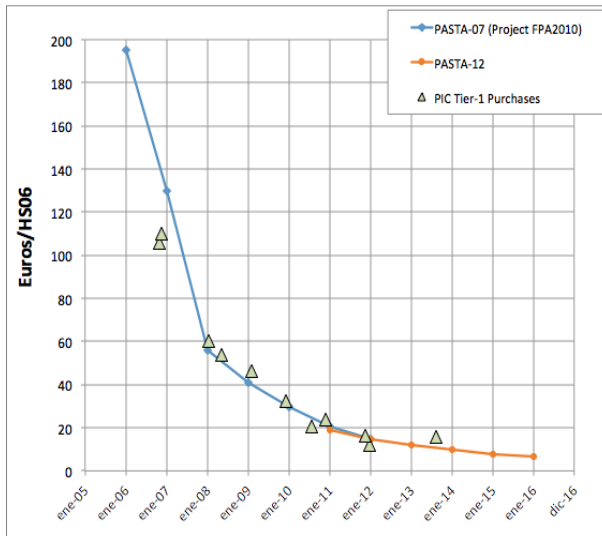
LHCb Reliability Ranking



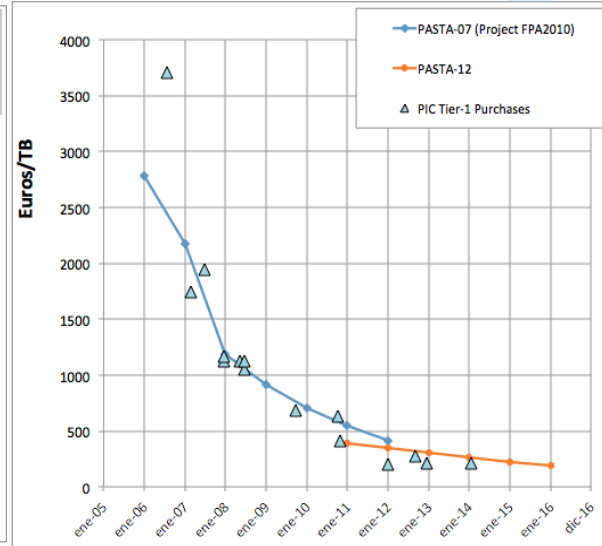
1/1/2014 → 5/4/2014



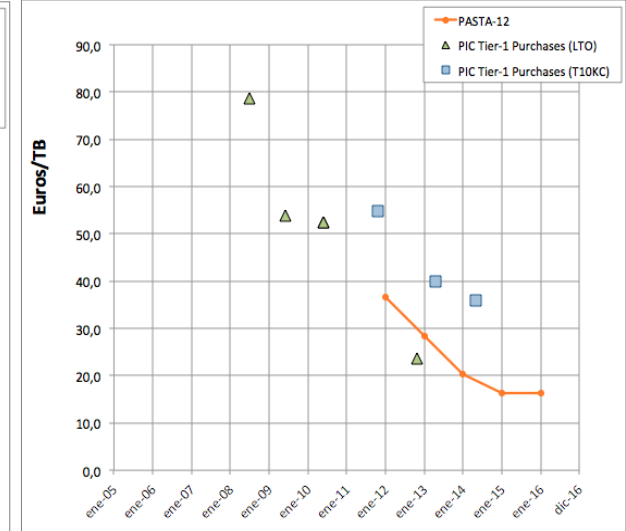
cpu



disk



tape (cartridges)



Data management issues

- Bit preservation
 - Archive data from disk to tape
 - Migrate from one tape technology to the next
 - Technology under control. Funding sources unclear.
- Context preservation
 - Need a lot of software to re-do analysis
 - “Useful” metadata often not known *a-priori*
 - Established ad-hoc standards vs. emerging
- Boundary between preserved/non-preserved
- Analysis functionality built into access portals

Example: Astronomy


 M51 r=1m

Displaying 32 of 160 Total Rows

Export Table As... MESSIER 051 (RA: 13:29:56.201, Dec: +47:13:50.02), radius: 0.01667°

	Short Name	Title
1	CADC	CADC Image Search
2	HLA	Hubble Legacy Archive

Filters: Clear Filters Edit Facets... Help...

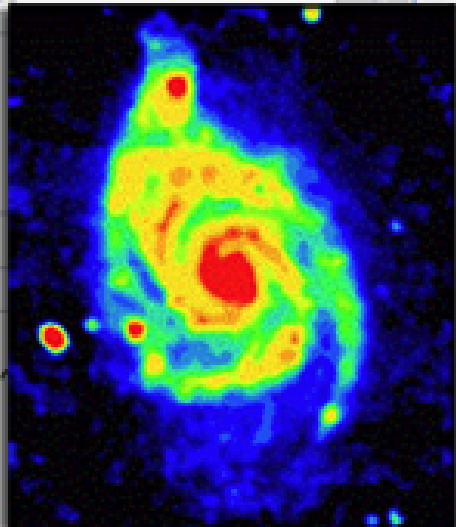
Filter All Record Fields

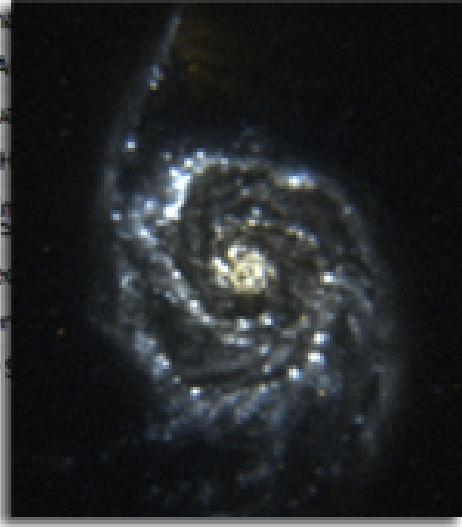
Categories:

- ☐ Catalog
- ☒ Images
- ☐ Spectra
- ☐ X-ray

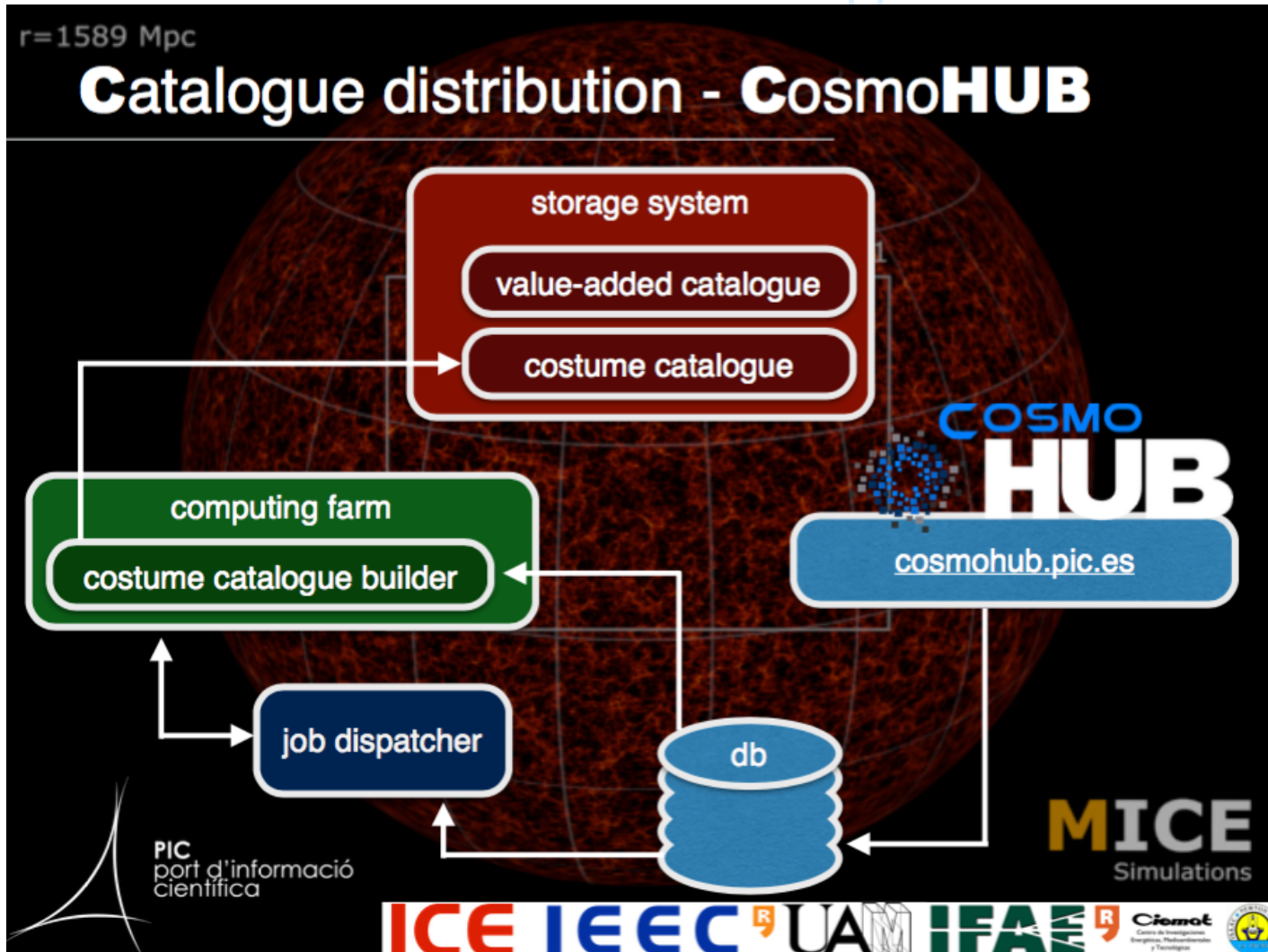
Publisher:

- ☐ Canadian Astronomical Society





Example: Cosmology

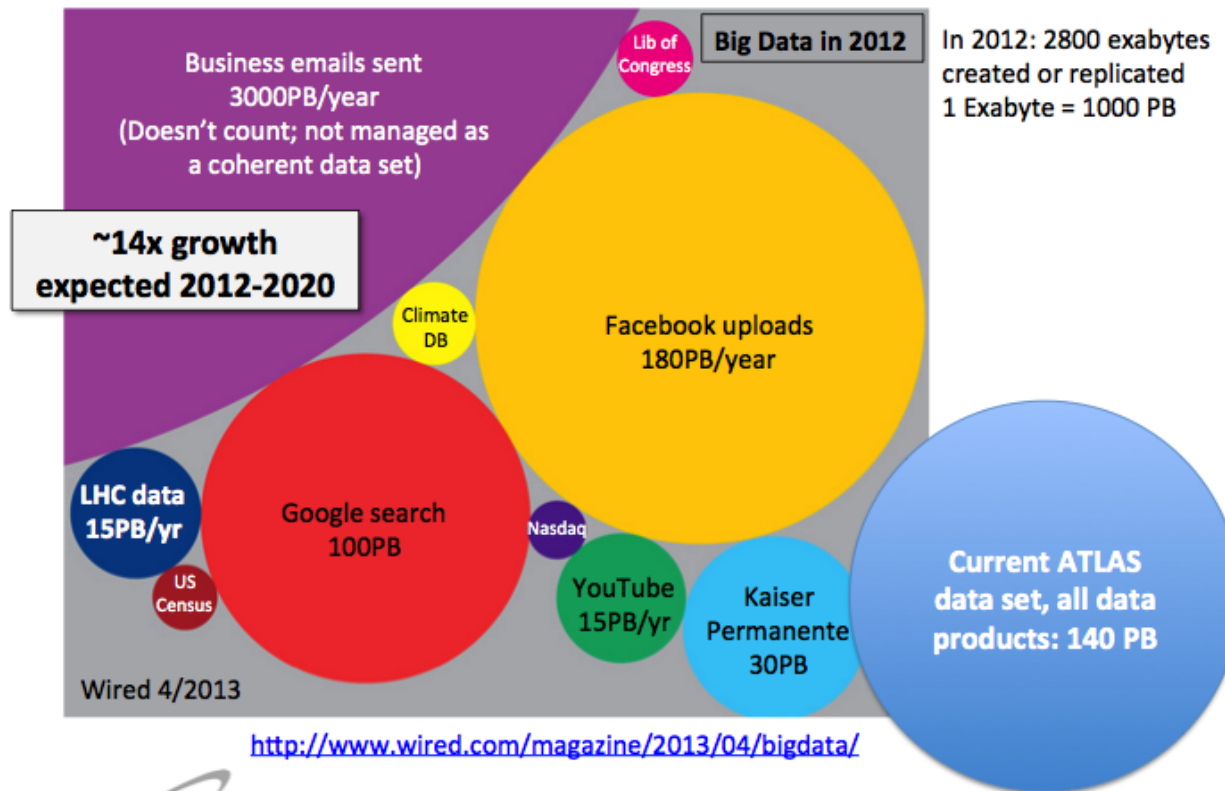


Example: Particle Physics

Data Management

Torre Wenaus; CHEP 2013

Where is LHC in Big Data Terms?



Summary

- PIC provides a Peta-scale data processing center for international projects in particle physics and astrophysics, as well as for smaller projects
- There are a lot of issues in the preservation, access and re-use of scientific data
- A center like PIC hosts, and increasingly generates on-demand, science data. Useage of these data needs to be monitored and citations should be ensured
- Very interested in standards for this