Massive data management (using Big Data technologies)

ESCAPE WP2/WP5 Workshop (Amsterdam 1st - 3rd of July 2019)

J. Rico, M. Delfino, G. Merino, J. Delgado on behalf of IFAE-PIC

IFAE 25 years 1991-2016 Institut de Física d'Altes Energies
EXCELENCIA SEVERO OCHOA
BIST Barcelona Institute of Science and Technology
PIC port d'informació científica
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
EXCELENCIA MARÍA DE MAEZTU

PIC
port d'informació
científica

Outline

- Port d'Informació Científica (PIC) overview
- Motivation for a Big Data platform - CosmoHub
- IFAE-PIC proposal for the WP5

Port d'Informació Científica (PIC)

- Founded in 2003:
  - Collaboration between IFAE and CIEMAT

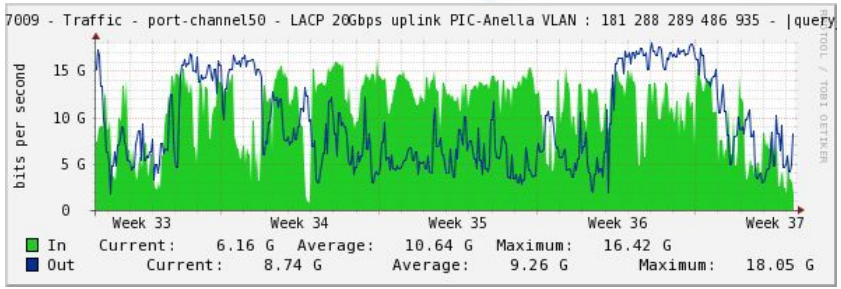
- Supported projects

We are in UAB campus

- Spanish Tier-1 Large Hadron Collider (LHC - WLCG)
- MAGIC Data Center
- EUCLID SDC-Spain

PIC - Network

- External network
  - 20 Gbps Wide Area Network (WAN)
    - 49,4 PB of data transferred last year (in and out)
  - 1 x 10 Gbps optical paths to CERN
  - 1 x 10 Gbps to Observatorio del Roque de los Muchachos (ORM - MAGIC Control House)
    - 250TB/year of data transferred to PIC from ORM for MAGIC
- Internal network
  - 240x 10 Gbps ports wire speed in core router
- 100 Gbps WAN connection expected by 2021

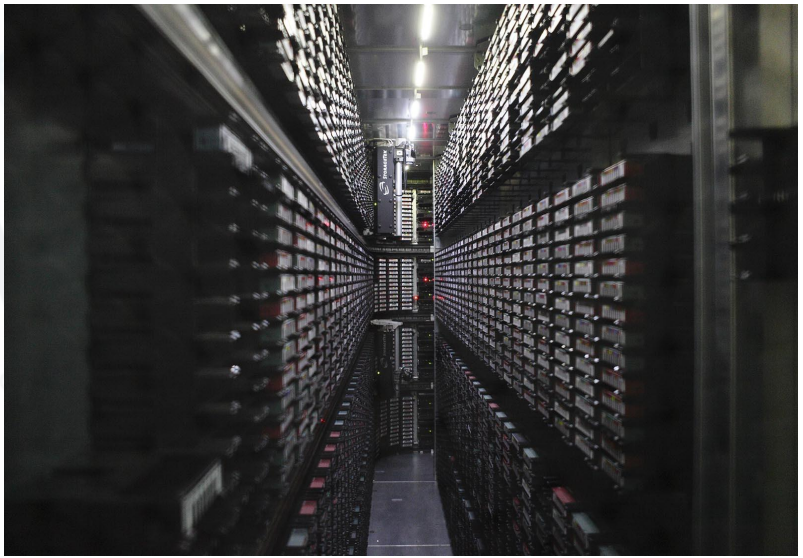


PIC - Disk Storage

- Distributed Hierarchical File System (dCache):
  - 50 nodes
  - 11 PB
  - Interfaces (how one can access the storage)
    - NFS
    - WebDAV
    - gridFTP
    - Xrootd
  - Integrated with the tape system

PIC - Tape Storage

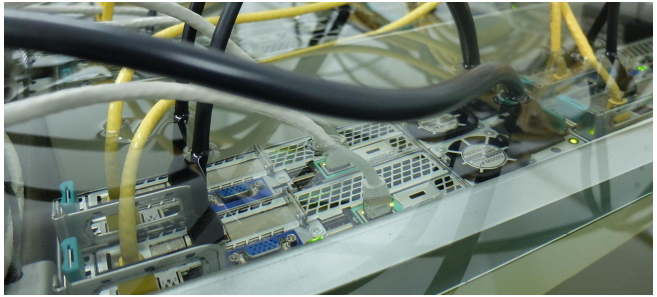
- Hardware - Oracle Tape Library SL8500
  - Library capacity (max. 6630 slots)
    - 5010 occupied slots
    - Current capacity 25 PB
  - Drive capacity (max. 64 drives)
    - 18 drives (4 LTO5, 8 T10KC , 6 T10KD)
  - Cartridge capacity
    - LTO-5: 1.5 TB
    - T10KC: 5.4 TB
    - T10KD: 8.4 TB
- Second tape library IBM under installation and test
- Software - Enstore (developed by Fermilab)



PIC
port d'informació
científica

PIC - Computing resources

- Resources
  - 8426 cores computing farm
    - ~ 11,000,000 jobs / year
  - 448 cores in Hadoop
  - User interfaces and other grid elements
- Job types:
  - Single / Multi core
  - MPI
- Batch job management:
  - PBS
  - HTCondor
  - YARN (Hadoop)
- Liquid immersion cooling for nodes (PUE=1.1)



Cosmology Projects

150TB raw
~30 users

30PB images
~1500 users

Galaxy Catalogs: our main roles

- **Generation** and **validation**
- **Exploration** and interactive **visualization**
- Data **Distribution**

PIC
port d'informació
científica

Main challenge: distribute and explore large datasets

After storing, reducing and calibrating the observed images one of the main outputs is a galaxy catalog, which contains their positions along with hundreds of other properties such as the luminosity, colour, morphology, etc

Project	Date	volume/night	Total Volume	Number of Objects (catalog)
SDSS	2000-now	variable	116 TiB	$2x10^6$
MICE GS	2013	NA	42 TiB	$5x10^8$
DES	2013-2018	2.5 TiB	2 PiB	$4x10^8$
GAIA	2014-2019	40 GiB	1 PiB	1.1×10^9
Euclid	2020-2025	100 GiB	580 TiB	1.5×10^9
LSST	2022-2032	15 TiB	50 PiB	$1x10^{10}$

PIC
port d'informació
científica

MICECat2: How we did it back in 2015?

MICECat2: $5x10^8$ number of objects

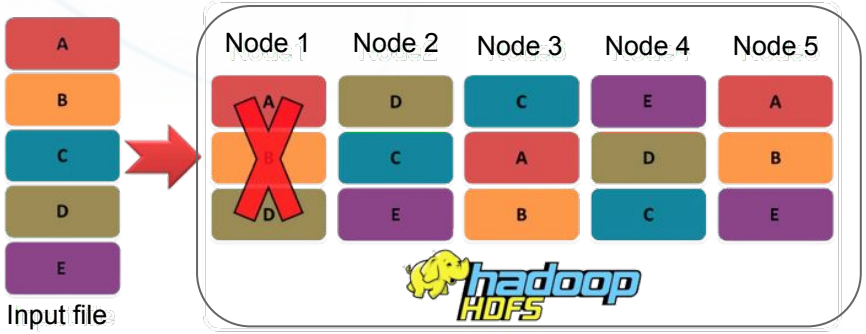
- Data located in different personal computers
  - Many data transfers
- Different people coding the same thing
  - No software version control
- It took some months to create it
- We finally ingested it into our distribution platform (relational database). **It took 26 hours!**

Big Data Platform

- Based on Hadoop (Hortonworks HDP 2.6)
  - Open source Big Data Platform
  - Distributed storage and processing
    - Runs on top of commodity computer clusters
    - Scalable from dozens up to thousands of nodes
      - Performance scales with HW
    - Fault tolerant
      - Simple machines working together - no single point of failure

PIC Hadoop platform

- 16 nodes + 3 head + UI
- 448 total cores
- (192 + 128) TiB HDD



PIC
port d'informació
científica

SciPIC: Scientific pipeline @ PIC

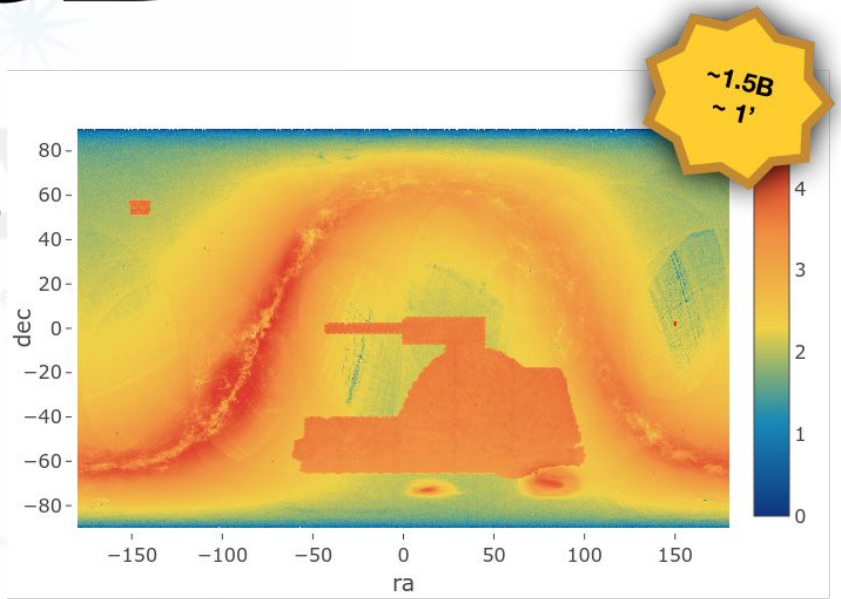
- Set of python code / algorithms to generate synthetic galaxy catalogs using Dark Matter simulations
- Runs on top of the PIC Big Data platform using Spark
- Software version control (git)
- Stored in CosmoHub ready to be explored and distributed

	input		output		time	time (+WL)
	objects	size	objects	size		
octant	5 x 10^9	0.7 TiB	7.4 x 10^9	3.2 TiB	< 2 h *	< 10 h *
full sky	4 x 10^{10}	5.5 TiB	6 x 10^{10}	25.5 TiB	< 16 h *	< 80 h *

PIC
port d'informació
científica

COSMO HUB

- Based on **Apache Hive** (ORC file format)
- Distribution
  - Query time range: seconds to minutes
    - 85% in < 3 min.
- Exploration (Visualization)
  - Unlimited time
  - Full dataset plots (over all raws)
    - May use sampling
    - 1D histogram & 2D heatmap



https://cosmohub.pic.es

GammaHUB: contribution of IFAE-PIC to ESCAPE WP5

- Will contain a substrate of the Hadoop platform
- Data ingestion:
  - Based on DL3 format files from the archive of several gamma-ray observatories (i.e HESS, MAGIC, CTA, other ESFRIS...?)

    MAGIC Is collecting ~10^8 entries per year

  - DL3 refers to a reduced level of data, describing gamma-ray candidates
    - Fields describing the the primary particle (energy, arrival direction…)
    - Fields describing primary interaction (time, height…)
    - Fields describing the telescope status (pointing direction, pointing IRFs…)

  - Apart from the event list, DL3 should contain the IRFs describing the events

GammaHUB: contribution of IFAE-PIC to ESCAPE WP5

- Main services on the platform
  - Web interface for interactive/on-line exploration and visualization of the selected datasets. Some examples:
    - Full sky skymap with the arrival direction of all events
    - Histogram counting the number of events coming from a given direction
    - Compute the significance of detection of a gamma-ray source in a given position of the sky
    - Plot the (unfolded) energy spectrum of gamma-rays coming from a given direction in a given interval of time
    - Other tasks of general interest to be identified…
  - Provide standard basic final analysis products, such as spectra and lightcurves
  - User-friendly (of course), guided for expert and non-expert users

GammaHUB: contribution of IFAE-PIC to ESCAPE WP5

- Main services on the platform
  - Web-based analysis on demand service based on Python Notebooks (JupyterLab)
    - Accessing to the DL3 archive
    - Load python packages for gamma ray analysis
    - Linked to Hadoop engines
  - For expert users

PIC
port d'informació
científica

Summary and conclusions

- Data management is crucial to:
  - Deliver scientific results successfully & on time
  - Engage community and boost outreach
- PIC has lots of expertise in data management in different science cases:
  - Particle physics, Astrophysics, Cosmology...
  - Development, integration and operation
  - Small and large collaboration projects
- PIC explores and develops on new technologies:
  - Big Data (Hadoop, Hive, Spark)
  - File formats (FITS, ASDF, Parquet, Arrow)
  - Machine Learning (over GPU)

Take the expertise from
Big Data technologies (CosmoHub)

Applied to Gamma-ray
astronomy

Massive data management (using Big Data technologies)

ESCAPE WP2/WP5 Workshop (Amsterdam 1st - 3rd of July 2019)

J. Rico, M. Delfino, G. Merino, J. Delgado on behalf of IFAE-PIC

IFAE 25 years 1991-2016 Institut de Física d'Altes Energies
EXCELENCIA SEVERO OCHOA
BIST Barcelona Institute of Science and Technology
PIC port d'informació científica
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD
Ciemat Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas
EXCELENCIA MARÍA DE MAEZTU