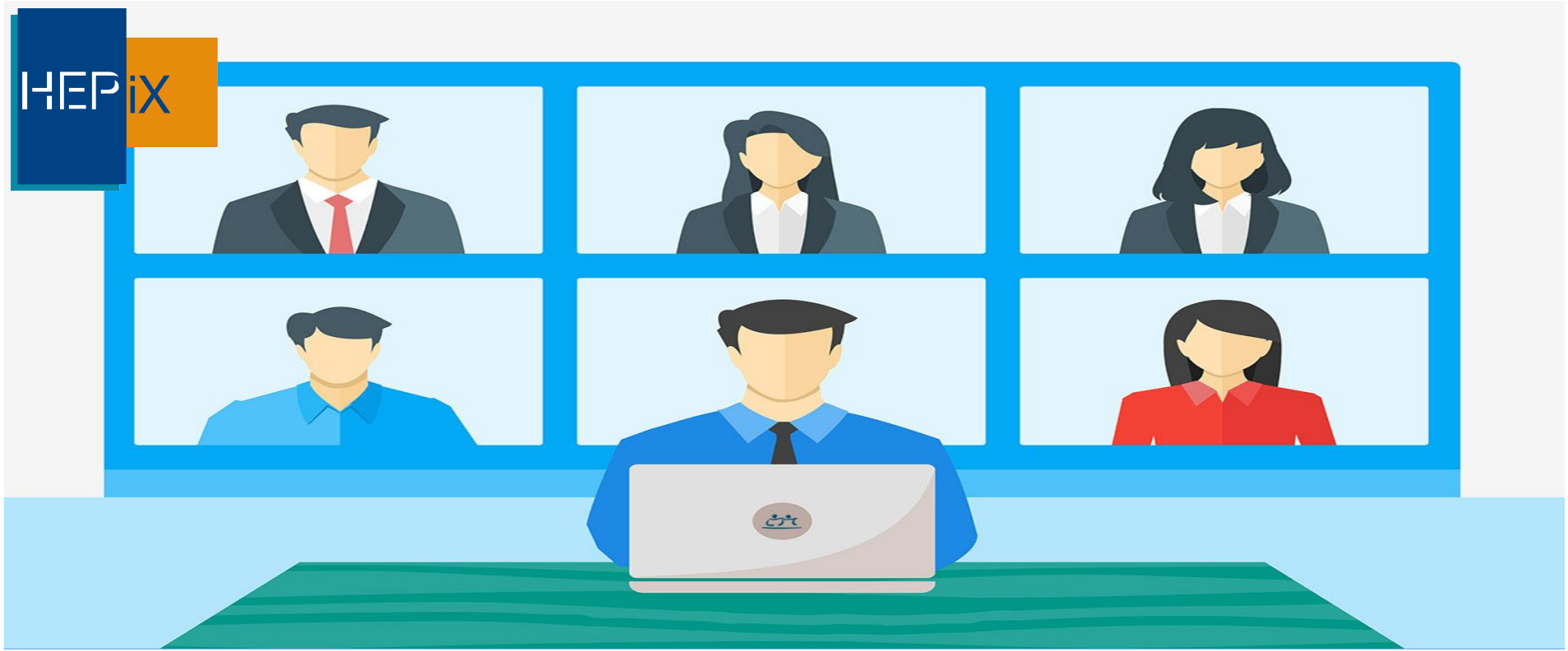


PIC Report - J. Flix [on behalf of PIC team]

HEPiX Spring 2021 / Virtual

15-19 March 2021

PIC
port d'informació
científica

Institut de Física d'Altes Energies
Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

PIC in numbers

March 2021

CPU: 110 kHS06
Disk: 10.1 PB
Tape: 33.1 PB

Spanish WLCG Tier-1 centre → **~80% of resources**

→ Provides ~5% of Tier1 data processing of CERN's LHC detectors ATLAS, CMS and LHCb

¼ of the Spanish ATLAS Tier-2 and **a Tier-3 ATLAS data analysis facility** → **~10% of resources**

T2K [neutrinos], **MAGIC** and **CTA** [gamma-ray astronomy], **PAU** and **EUCLID** [cosmology], **VIP** [instrumentation], opportunistic access to **LIGO/VIRGO** and **DUNE**, among others...

PIC farm updates

321 compute nodes (8056 slots), under **HTCondor v.8.8.12**

→ 84% of compute nodes dual-stack; old hardware still in IPv4 (won't be migrated)

2x HTCondor-CE v3.4.3-1
2x ARC-CE v.6.10.1 (used by ATLAS and LHCb as HPC gateways - see later)

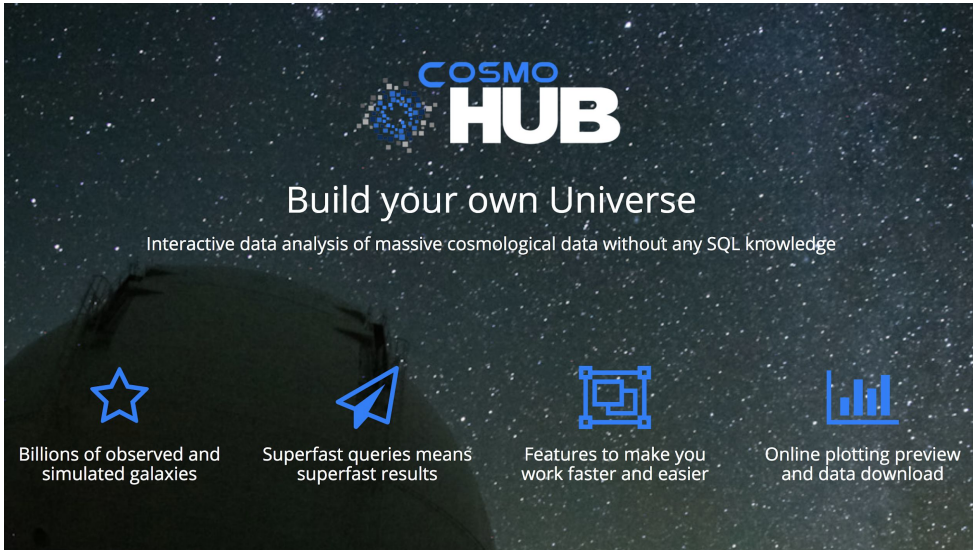
Continued tests at low scale with CMS workloads on **AWS** and **CloudSigma**

A portal is available to select different profiles (CPU, mem, GPU) to **spawn Jupyter Notebooks to the PIC farm**

10 GPUs available: 2 in use for farm jobs (VIRGO/LIGO users) and 8 for JupyterHub

Support to new groups in **Quantum Informatics** and **BioInformatics**

Upgrade/new Hadoop cluster @CosmoHub



New Hadoop cluster: 12 nodes AMD Threadripper 1920x, 128 GB RAM, 12 x 3 TB SATA HDD hot-swap, 2x1 TB NVMe SSD i 2x10-GBASE-T LAN

Hadoop upgrade from HDP 2.6.5 to HDP 3.1.4 (+ Hive 3.1.0 and Spark 2.3.2)

Cone Search (sky position and an angular distance, defining a cone on the sky) integrated

GammaHub (ESCAPE WP5)

GammaHub is intended to be an interactive science analysis platform by itself, taking previous experiences from https://cosmohub.pic.es, ingest datasets from multi-instrument astronomical gamma-ray experiments like MAGIC, HESS, VERITAS and CTA among others

Web interface

- Provided by CosmoHub https://cosmohub02.pic.es
- User friendly data search
- Expert mode for complex queries
- Subsets generator and download
- A web interface for fast visualization/histogramming
- Ingested data from MAGIC, HESS, VERITAS, CTA
- Cone Search implementation for data search in the sky
- Instrument filter by name

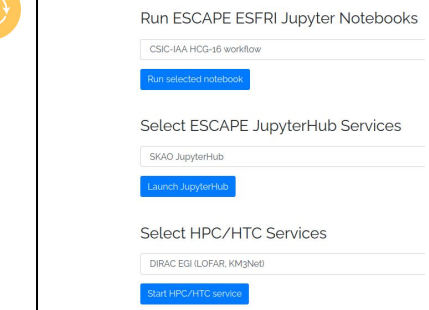
Data Products

- Compute the significance of detection of a gamma-ray source
- Source spectra calculation

- Data Lake (WP2) integration with RUCIO
- REANA (Reusable Analysis) Deploying

Python Notebooks

- Jupyter service at PIC (PIC Account + Kerberos User) https://jupyter.pic.es
- Integrated with a link in ESAP https://sdc.astron.nl/esap-gui/interactive



- Connected to Hadoop
- Virtual Environment and Kernel

GammaHub (sketch)

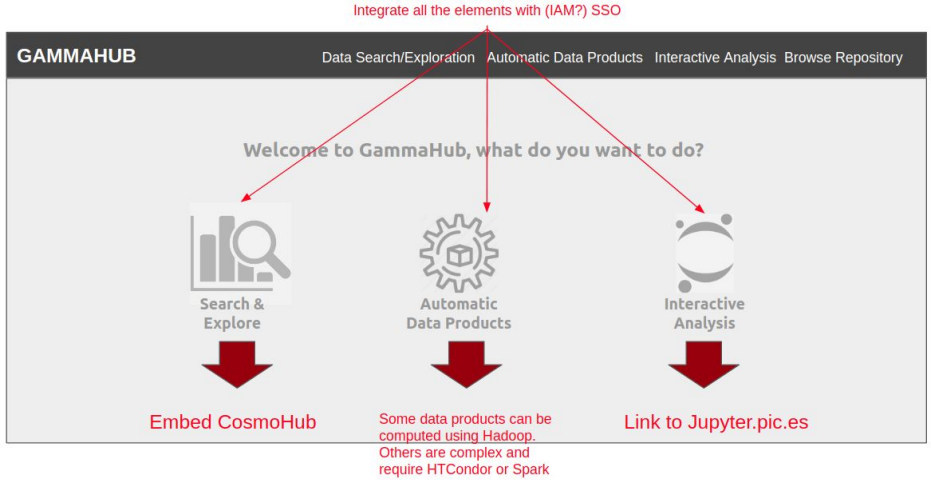
Design of a web interface to integrate all components

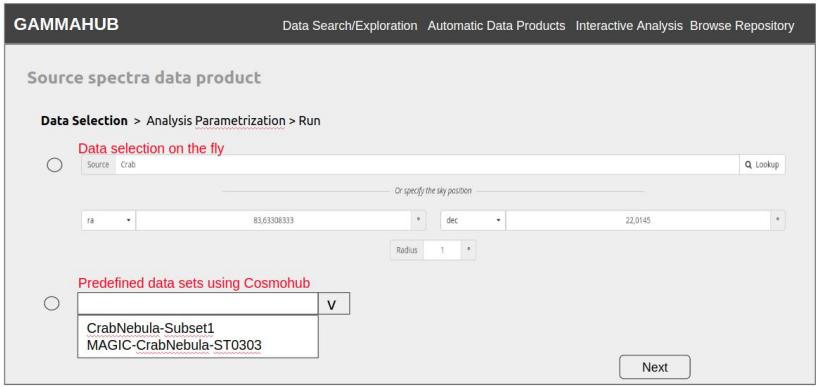
Single Sign On integration to allow working with cosmohub engines and jupyter.pic.es

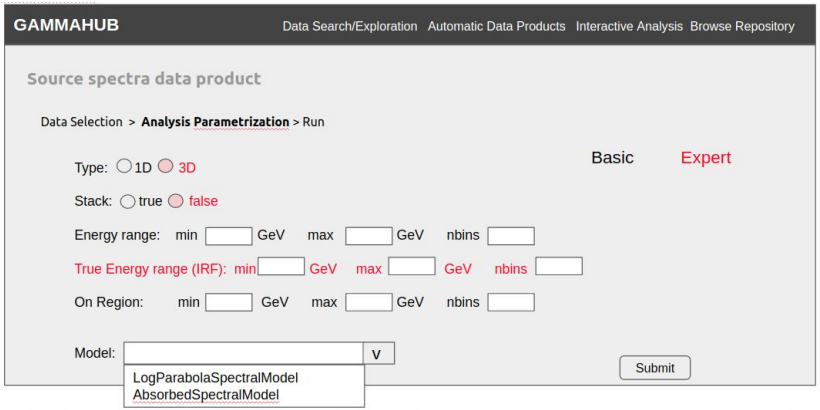
Interface to parametrize the data products workflows

Data staging using RUCIO

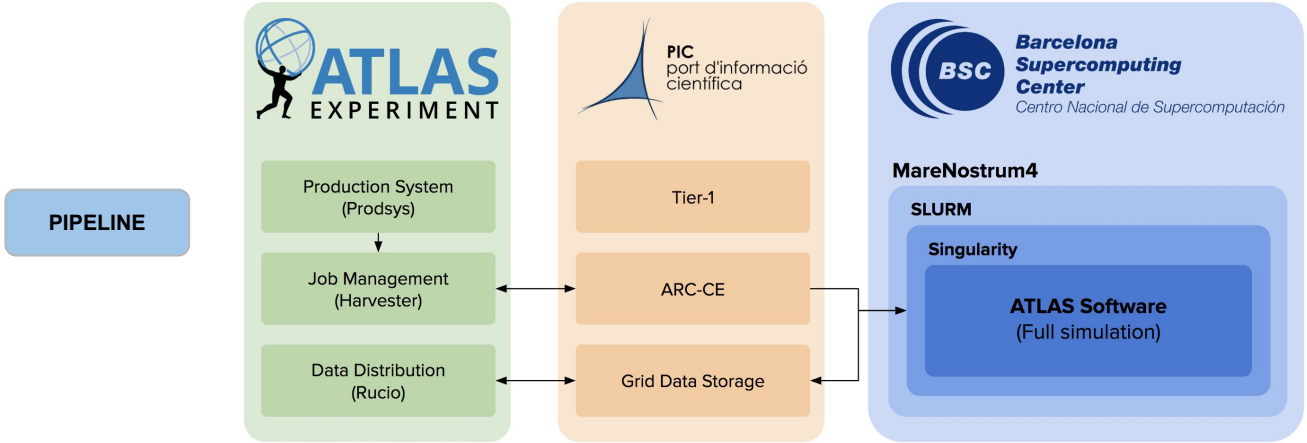
Integration in ESCAPE Science Analysis Platform (ESAP)







Use of the BSC by ATLAS PIC Tier-1



Using two **ARC-CEs** at PIC to interconnect MareNostrum and ATLAS production system

Only simulation workflow validated - singularity containers, pre-placed at MareNostrum GPFS

MareNostrum accepts only SSH protocol for job submission and data transfer

Use of the BSC by ATLAS PIC Tier-1

Submitting **ATLAS** payloads to BSC from PIC Tier-1 since 2018, <u>in production since 2019</u>

Stats for 2020: 10 million hours approved in 2020 (100% used) and one request of 4.5 million hours approved [1 hour = 16.75 HS06-hours] → 53 millions of events simulated

LHCb testing similar technical implementations in the same grant

All ATLAS Tier-2s in Spain send jobs to MareNostrum

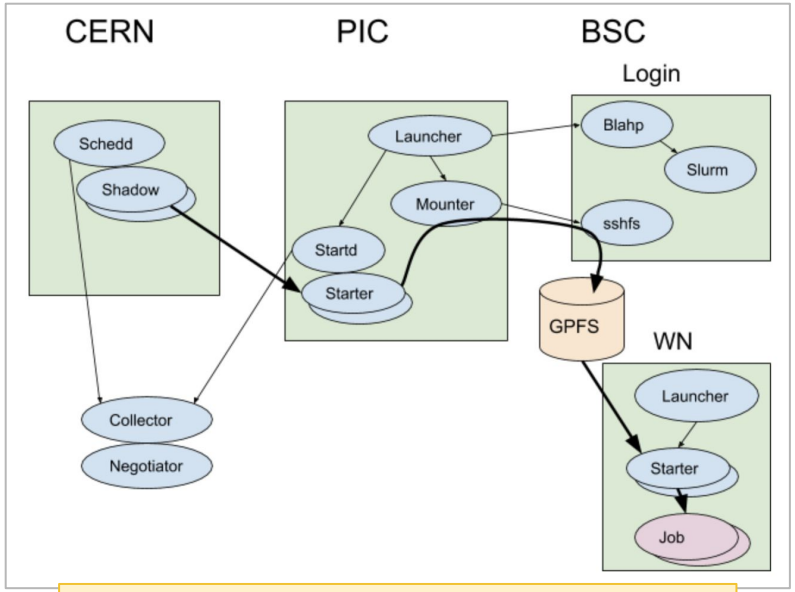
2020 figures

- ~80% of simulation jobs ran at BSC
- Overall 15% CPU beyond pledges have been obtained

→ **Draft paper submitted to CHEP2021 (link)**

Use of the BSC by CMS PIC Tier-1

PIC and HTCondor team collaboration to **use a shared FS as control path for HTCondor**



Setup that interconnects all of the HTCondor daemons for the CMS Global Pool, PIC Tier-1 center and the BSC

→ **Draft paper submitted to CHEP2021 (link)**

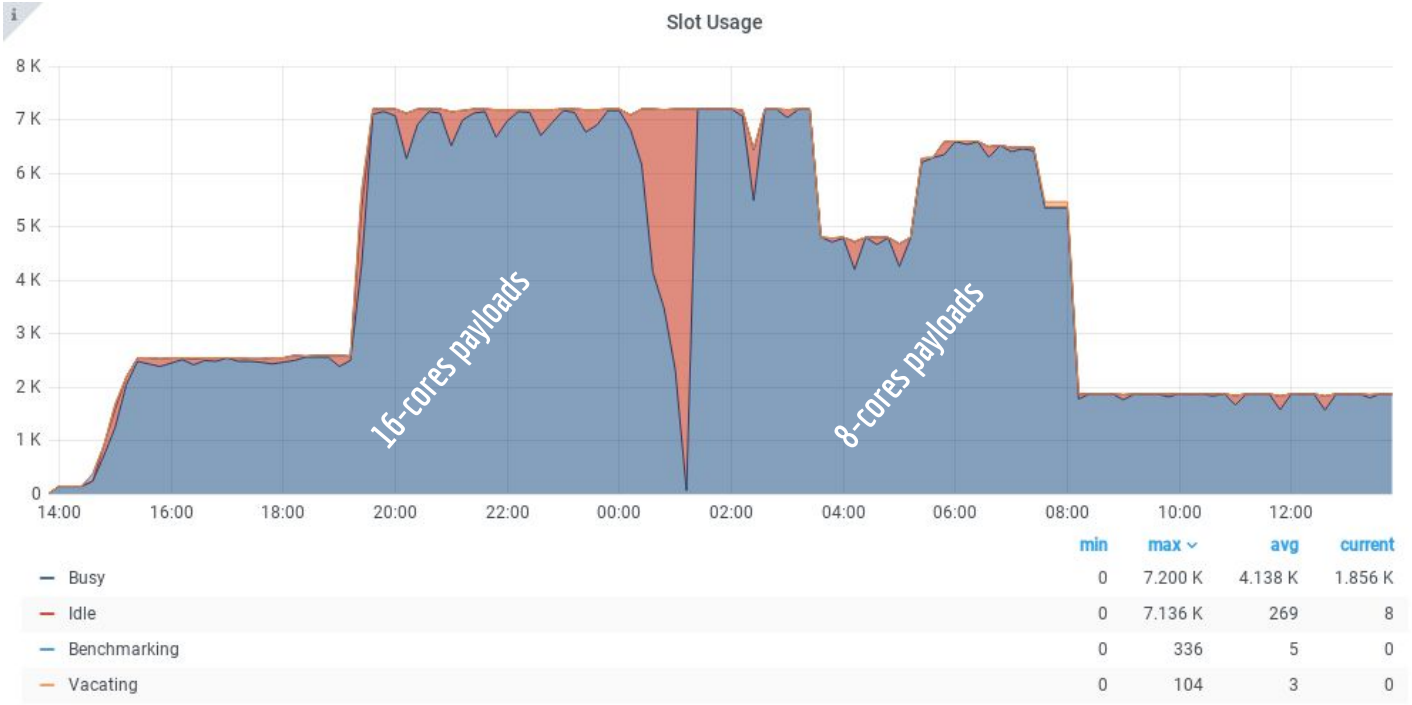
Current status

- An HTCondor-bridge has been deployed at PIC to interact with BSC execute nodes through the login node, mounting the shared FS through **sshfs** and sending jobs to the Slurm scheduler via **ssh**
- Ran a self-contained payloads which do not require external connectivity connected to the CMS global pool (application packaged inside a **Singularity container**, and **conditions data** read at run time dumped into a **sql file**, no I/O)
- **CMS Software modified** to accept sql files for conditions data at runtime
- Allocation of 1Mhours (Oct-Dec. 2020) consumed at 40%, due to **scalability Issues and saturation** at the bridge capacity. New grant of 6 Mhours approved

Next steps

- Load balanced bridges at PIC to absorb peaks in allocated capacity
- Testing cvmfs_preload to bring CMS software into BSC storage
- Developing a data management service to get output files from BSC

Use of the BSC by CMS PIC Tier-1



Scale tests: running singularity images for CMS simulation on ~7k slots (running in 48 cores machines, tuning payload core usages to maximize global CPU efficiency), plugged into the CMS Global Pool (test instance) through PIC HTCondor infrastructure, using the shared FS at BSC

PIC storage updates

~10 PB running on **dCache 5.2.35**

→ dCache pools in dual-stack
→ TPC enabled for HTTPs and XRootD and token authentication (PIC in DOMA testbeds)

StashCache deployed on K8s (OSG repo) for Virgo/Ligo

→ 3.2 TB - 90% occupancy
→ Migrated to XRootD 5.0

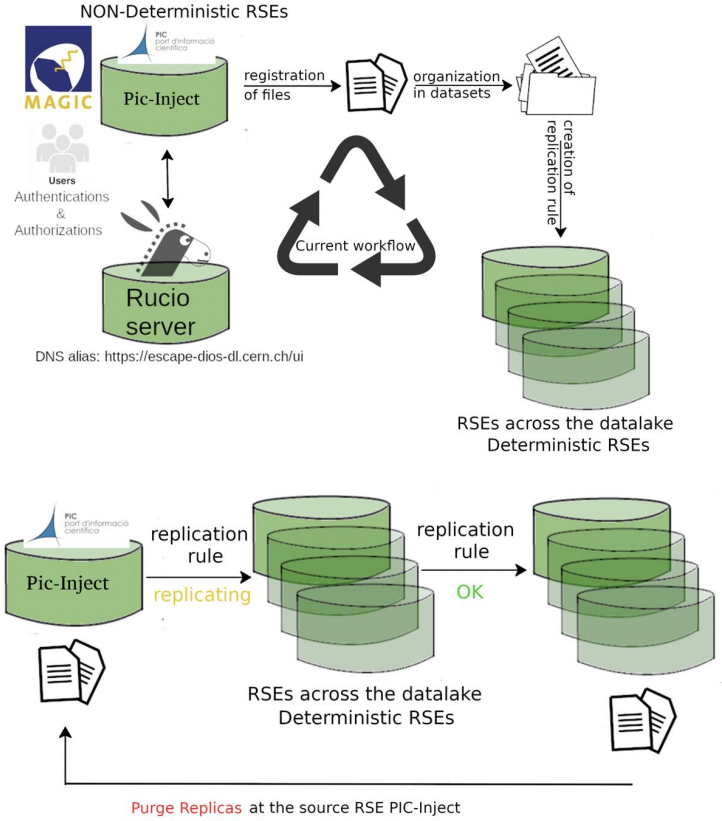
xCache deployed (OSG repo) for the CMS experiment

→ 4TBx36, 16 cores L5630 (HT enabled), 48 GB RAM, 10 Gbps - ~75% occupancy
→ Migrated to XRootD 5.0
→ Currently at low scale. Monitors and proper setting being deployed and/or investigated

PIC data injector (ESCAPE)

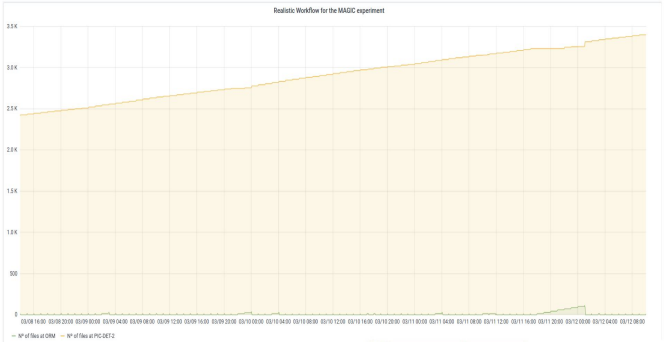
Aim: develop solutions to handle the large datasets produced by Gamma ray telescopes, adopting Rucio to stream files from the telescopes to a Data Lake for permanent storage and access

Using **PIC and Cherenkov telescopes in Canary islands** as the testbed (MAGIC and CTA)



Currently testing

Currently testing the automatic workflow from the non-deterministic RSE configuration at the site to a deterministic RSE at PIC (1 month of successful tests so far!)



Expansion of the new tape library

IBM TS4500

SL8500

IBM TS4500: 2 frames (L55+D55) + 8 LT08 drives

→ 4.8 PB capacity installed with cartridges LT07 M8

→ 750 TB capacity installed with cartridges LT08

This library is <u>expected to grow</u> to host future data

→ It will host new data and data migrated from SL8500 library (ongoing)

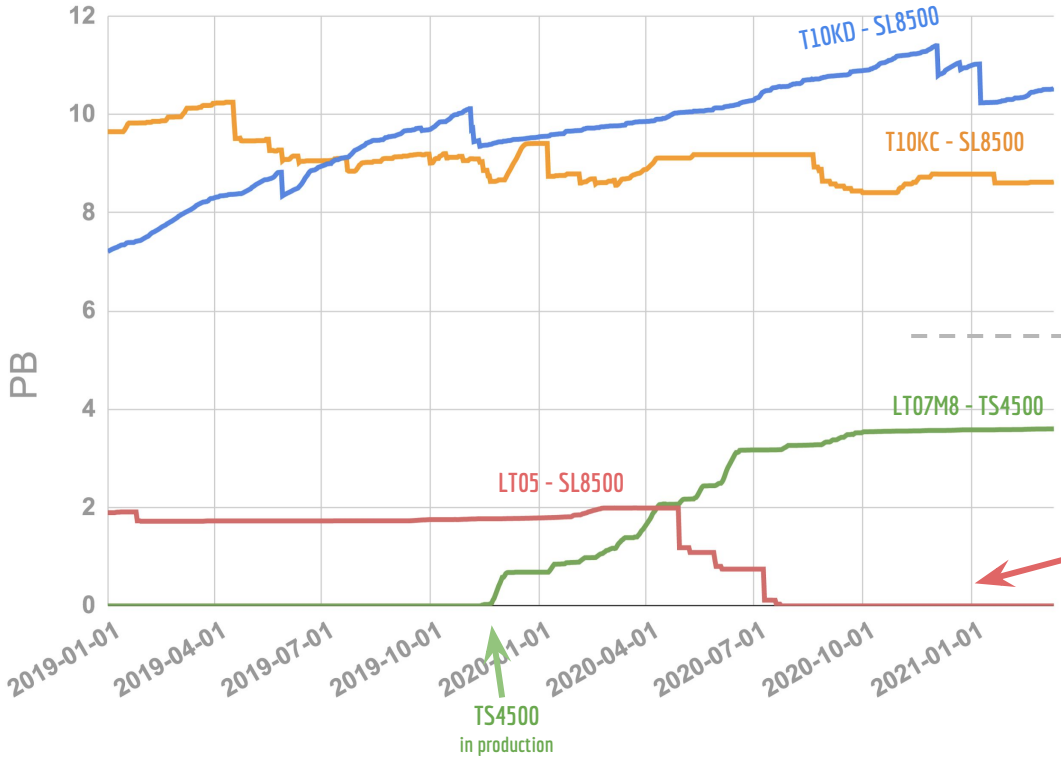
→ Dedicated drives, frames and cartridges installed to handle this

All new **CMS**, **LHCb** and **MAGIC** data go to the IBM

PIC currently runs **Enstore 6.3.4-2** (CentOS7)

Data migrations to TS4500

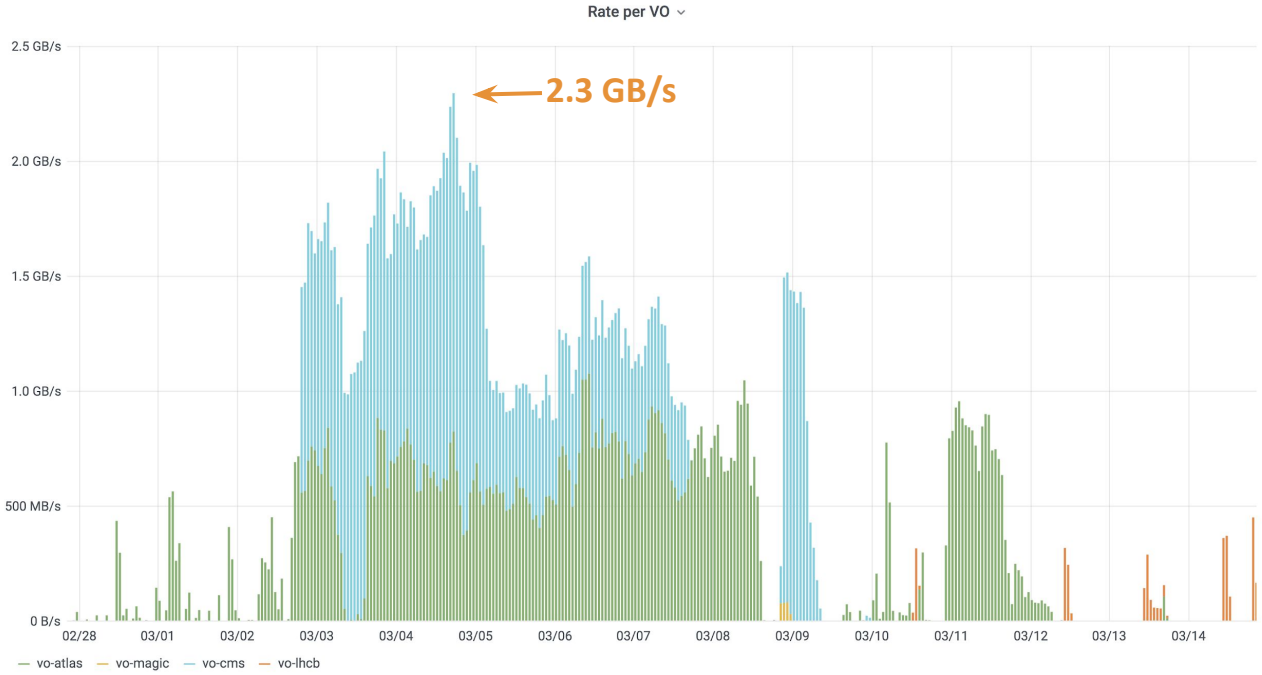
Used space by WLCG



Migrating T10KC data from CMS to LT07M8 technology
So far, 1.4 PB migrated

Fully migrated all LHCb data from LT05 (~2PB) to LT07M8
LT05 technology deprecated at PIC

Data migrations to TS4500



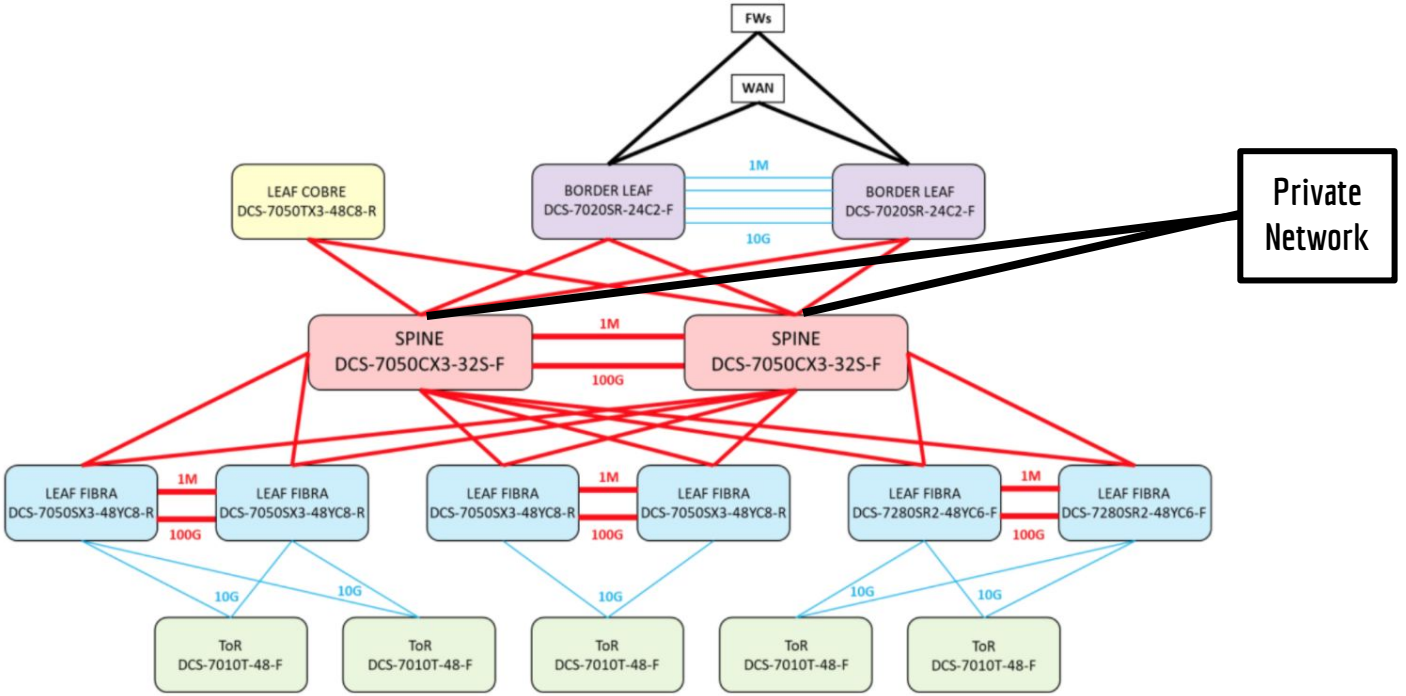
Currently, massive read tests from tape @PIC for both ATLAS and CMS

Network Upgrade @ PIC: 2x100 Gbps

Current 10Gbps core network (NEXUS 7009) being upgraded to 2x100 Gbps (ARISTA)

- **2x DCS-7050CX3-32S-F:** 32 ports 100GbE QSFP28 and 2 ports SFP+
  - Interconnection and routing equipment for private networks
- **4x DCS-7050SX3-48YC8-R:** 48 ports 10/25GbE SFP28 and 8 ports 100GbE QSFP
  - To connect equipments with 10 Gbps ports
- **2x DCS-7280SR2-48YC6-F:** 48 ports 10/25GbE SFP28 and 6 ports 100GbE QSFP
  - To connect storage equipments with 10/25 Gbps
- **1x DCS-7050TX3-48C8-R:** 48 ports 10GBaseT and 8 ports 100GbE QSFP
  - To connect IT equipments with 10 Gbps copper ports
- **5x DCS-7010T-48-F:** 48 ports 1GBaseT and 4 ports SFP+ 1/10GbE
  - ToR function
- **2x DCS-7020SR-24C2-F:** 24 ports 10GbE SFP+ and 2 ports 100GbE QSFP28
  - To connect Firewall and DMZ equipments

Network Upgrade @ PIC: 2x100 Gbps



Keeping us busy for the next months - all elements expected in place before Summer 2021

Thanks!
Questions?

Credits to: E. Acción, V. Acin, C. Acosta, A. Bruzzese, J. Carretero, J. Casals, R. Cruz, M. Delfino, J. Delgado, J. Flix, G. Merino, C. Neissner, A. Pacheco, C. Pérez, A. Pérez-Calero, E. Planas, M.C. Porto, B. Rodríguez, P. Tallada, F. Torradeflot

www.pic.es