CMS use case

D. Ciangottini for the CMS-ESCAPE team

ESCAPE - The European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement n° 824064.

Outline

- CMS plans overview
- Compute integration
- Data access
- WP2-5 contact surface

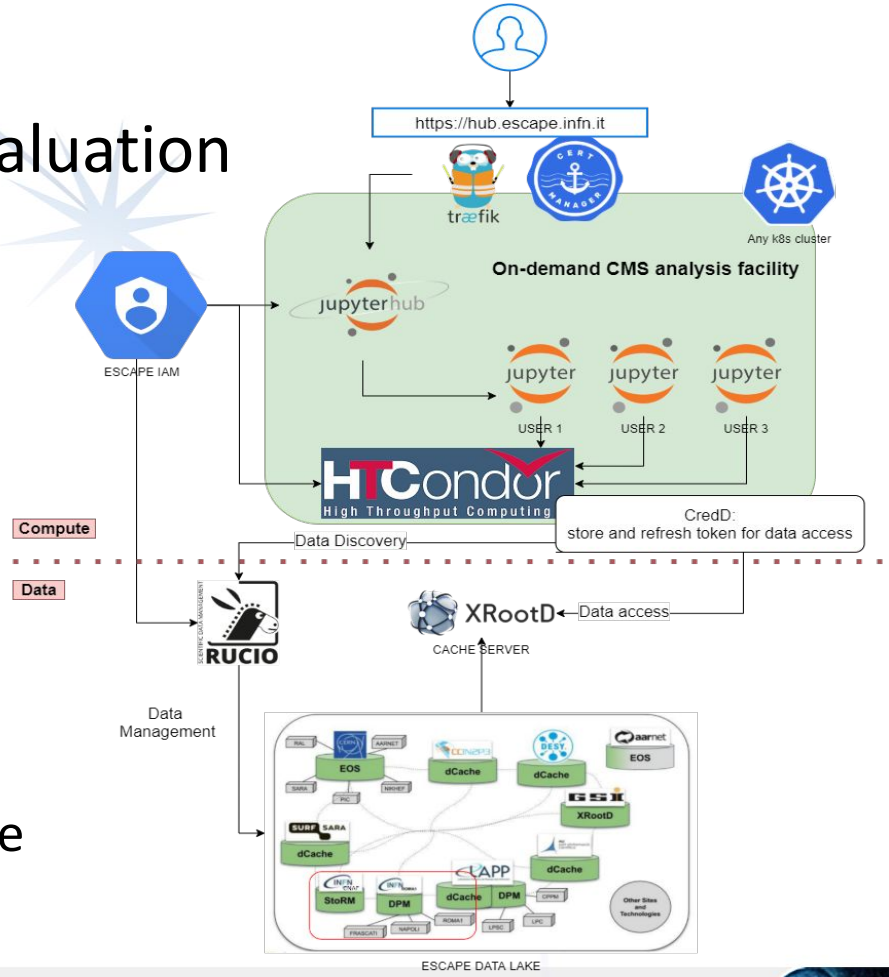
Funded by the European Union's
Horizon 2020 - Grant N° 824064

CMS ESCAPE data-lake evaluation

- Objectives:
  - On demand analysis facility to deploy on any k8s cluster
    - Stateless and reproducible
  - IAM based AuthN/Z for both data and compute resources access
    - X509 free
  - Support both interactive and batch workflows
  - Support data access for multiple kind architecture and resources

Bonus: keep it as much general as possible

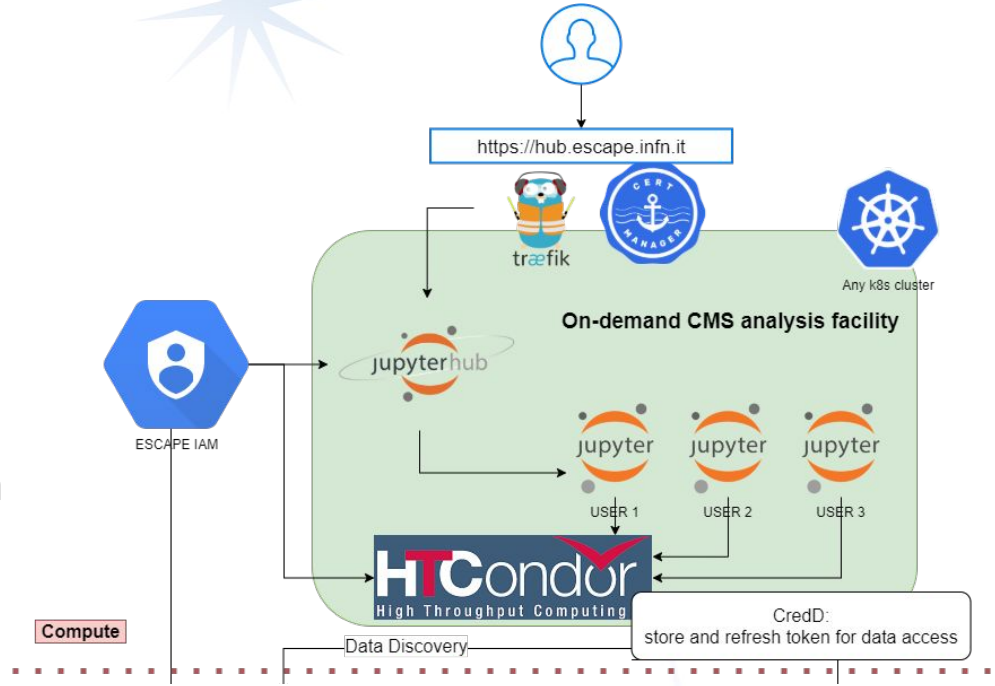
Sharable approach with other communities, e.g. in the context of the INFN cloud federation



Funded by the European Union's Horizon 2020 - Grant N° 824064

On-demand compute

- Provide software packed into Helm chart
- Multi user and possibly multi group
  - Access IAM managed
- User containers already configured to access HTCondor batch and remote data through access token

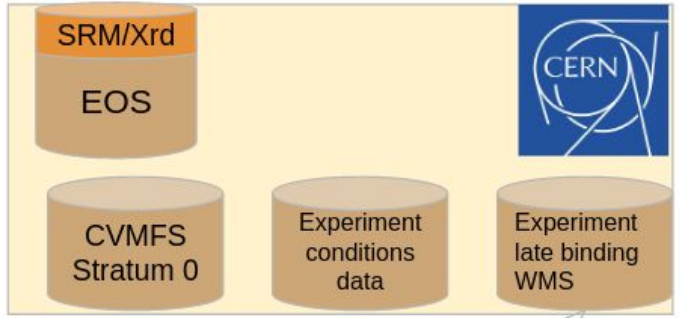


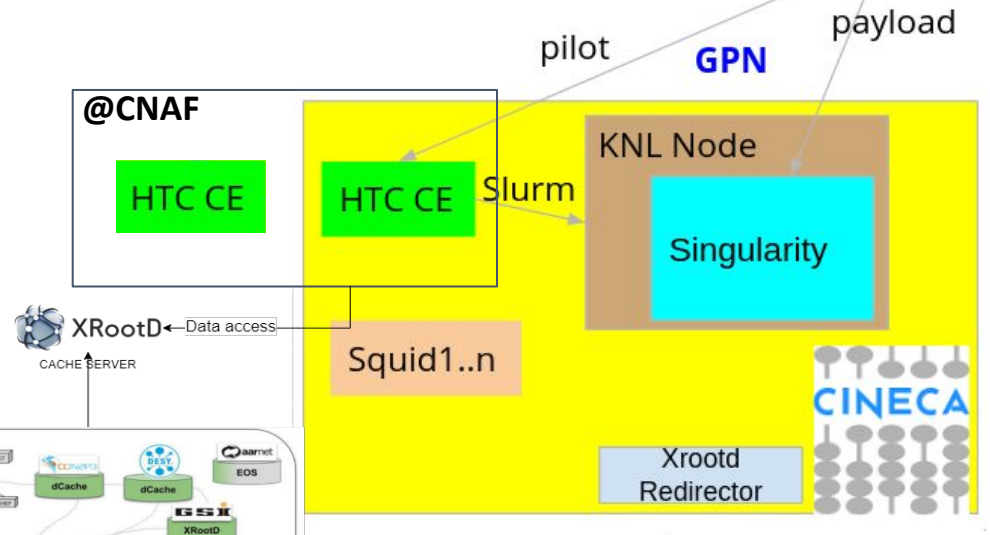
Funded by the European Union's Horizon 2020 - Grant N° 824064

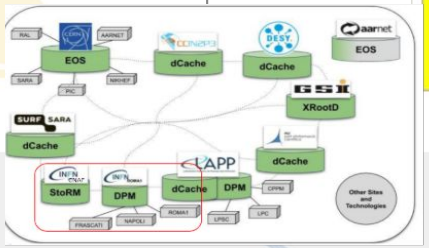
CMS-CINECA WMS integration

- SW dependencies on CVMFS
- Distributing payloads:
  - Accept / submit workloads which fit the RAM / walltime and IO bandwidth
- We went for **(site) customizable pilots** inside CMS WM
- Access to data through XCache server

*Thanks to CNAF support





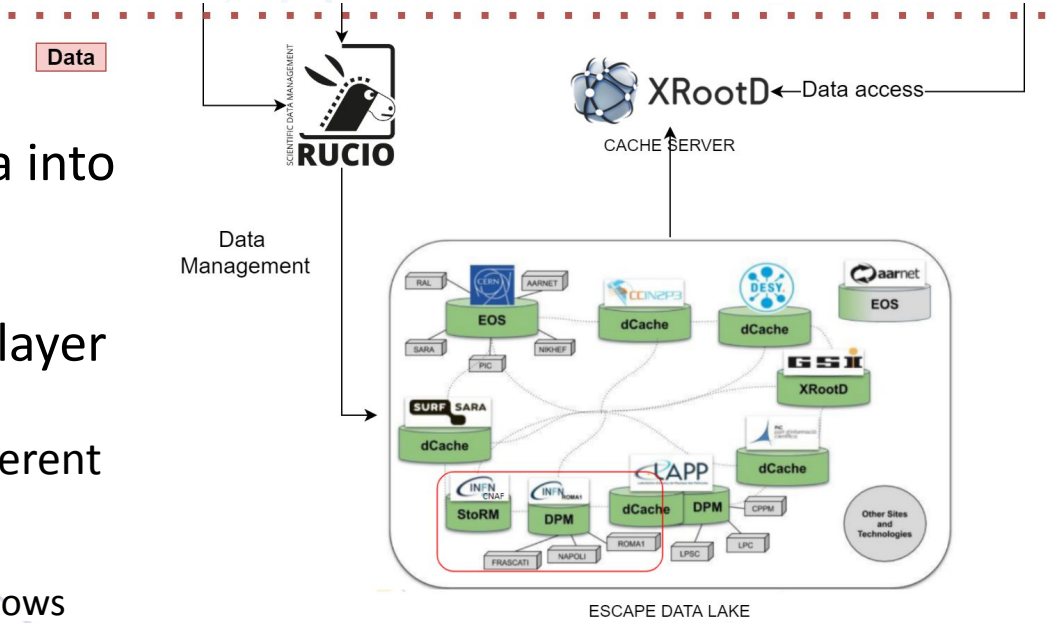


ESCAPE DATA LAKE

Funded by the European Union's Horizon 2020 - Grant N° 824064

Data access and caches

- Importing and managing data into the data lake
- Rucio as data discovery tool
- Data access through a cache layer for better performances
  - Multiple layers foreseen with different purposes:
    - Remote read latency hiding
    - Data pre-staging for high I/O workflows



Funded by the European Union's Horizon 2020 - Grant N° 824064

WP2-WP5 integration interest

- Jupyter Rucio integration (token based?)
- Multi-layered Cache system
  - Localhost level (ssd/io intensive)
  - Site level (latency hiding)
  - Region level (disk space optimization)
- Embargoed data management:
  - automatic replication + cache aware storage federation (Rucio)

Funded by the European Union's Horizon 2020 - Grant N° 824064