Data Management and Data Access: lessons learnt from ESCAPE

Xavier Espinal (CERN)

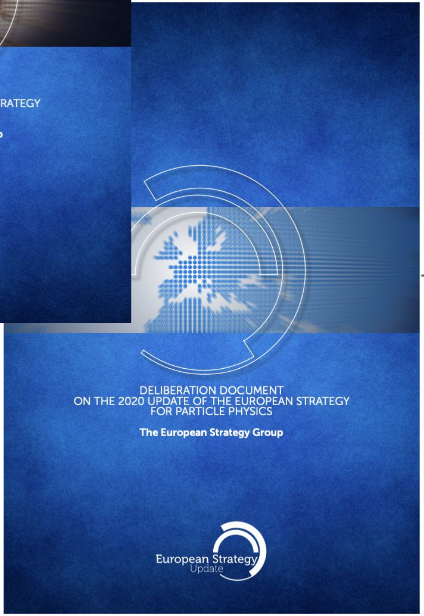
WLCG workshop, Lancaster, 7-9 November 2022

European Strategy and the ESFRI science clusters

2020 European Strategy for Particle Physics

4 Other essential scientific activities for particle physics

Full document



Full document

D. Large-scale data-intensive software and computing infrastructures are an essential ingredient to particle physics research programmes. The community faces major challenges in this area, notably with a view to the HL-LHC. As a result, the software and computing models used in particle physics research must evolve to meet the future needs of the field. ***The community must vigorously pursue common, coordinated R&D efforts in collaboration with other fields of science and industry, to develop software and computing infrastructures that exploit recent advances in information technology and data science. Further development of internal policies on open data and data preservation should be encouraged, and an adequate level of resources invested in their implementation.***

The scientific outcomes of particle physics experiments are made possible by the development of an efficient computing and software infrastructure. Computing and software are profound R&D topics in their own right and are essential to sustain and enhance particle physics research capabilities. There is a need for strong community-wide coordination for computing and software R&D activities, and for the development of common coordinating structures that will promote coherence in these activities, long-term planning and effective means of exploiting synergies with other disciplines and industry. Some recently initiated examples are the HEP Software Foundation addressing the common computing and software challenges related to particle physics, and ESCAPE (European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures) exploring the synergies in the areas of astronomy, astroparticle and accelerator-based particle physics.

H2020: Connecting ESFRI infrastructures through Cluster projects

The H2020 cluster concept was aimed at **supporting open data intensive driven science**, lead to **new insights and innovation**

The approach:

- Foster the establishment of cross-border open innovation environment
- Develop synergies between the involved ESFRI[1] RIs
- Adopt global standards and common solutions for data management favoring economy of scale

The scope:

- Open Science commitment, implement the FAIRness of scientific data
- Link the ESFRI RIs to the EOSC

[1]The **European** Strategy Forum on Research Infrastructures

ESCAPE - The European Science Cluster of Astronomy & Particle physics ESFRI Research Infrastructures

Synergies ESCAPE, WLCG and DOMA

Setting the scene 2018 - now (1/2)

- DOMA Access 09/2018 =>12/2020
  - Studies on Storage Consolidation
    - Content delivery, streaming caches, data (pre)placement
    - Event services
  - Early thoughts on “Data Lakes“ prototypes

- ESCAPE 02/19 =>Jan/23*
  - Large scale distributed computing needs for SKAO, CTAO, Vera Rubin/LSST, KM3Net, FAIR, EGO/Virgo
    - Data and Metadata Management
    - File Transfers
    - Content Delivery
    - Analysis Environments
    - Distributed computing “structure”: sites, resources and users

*An Open Collaboration Agreement signed by ESCAPE ESFRIs, starting Jan/23 is providing a formal statement to foster scientific collaboration in computing, maintaining the community together and following-up on ESCAPE work program

Setting the scene 2018 - now (2/2)

- **Large overlap** in distributed computing needs across scientific communities
- LHC's mature computing services and tools well received
- Common need for ***heterogeneous*** **computing integration**
    - Data processing: commercial clouds, HPCs, ephemeral/sporadic resources (private clouds)
    - Data storage: *register* existing Storage Elements, cloud endpoints, data placement to exploit *ephemeral* resources
    - Common and **interoperable** AAI: x509"free" scenario => token based
- Common need to orchestrate Data Lifecycles:
    - Data management from the source (inc. very remote locations) and data *preparation*: from raw to analysis ready products
    - Match data value to storage *price-tag*: QoS
- Common interest on Analysis Platforms/Frameworks, Virtual Research Environments
    - Open science, learning and citizen science but also user analysis
        - Need a key step forward to be able to scale out in number of jobs and link different types of resources (local and external)

ESCAPE

ESCAPE - Work program

Data Lake:
Build a scalable, federated, data infrastructure as the basis of open science for the ESFRI projects within ESCAPE.

Software Repository:
Repository of "scientific software" as a major component of the “data” to be curated in EOSC.

Science Platforms:
Flexible science platforms to enable the open data analysis tailored by and for each facility as well as a global one for transversal workflows.

Citizen Science:
Open gateway for citizen science on ESCAPE data archives and ESFRI community

Virtual Observatory:
Extend the VO *FAIR* standards, methods and to a broader scientific context; prepare the VO to interface the large data volumes of next facilities.

25/10/2022

The ESCAPE Data Lake

The ESCAPE Scientific-Data Lake is a policy-driven, reliable distributed data infrastructure capable of managing Exabyte-scale data sets. And able to deliver data on-demand at low latency to all types of processing facilities

Services operated by the ESCAPE partner institutes

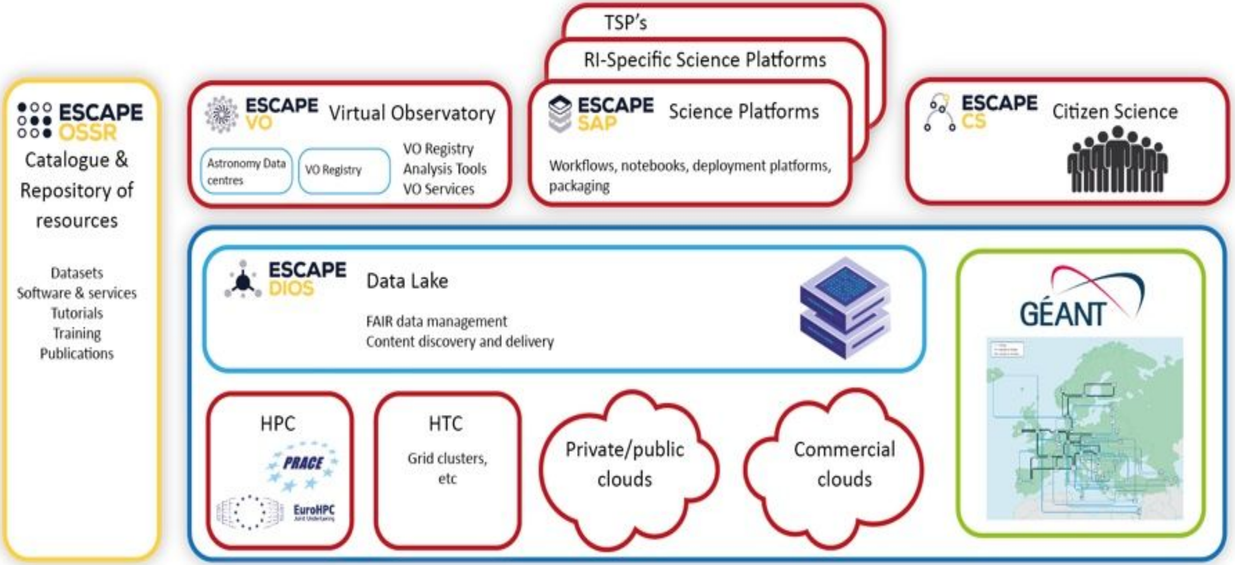
Petabyte scale storage: DESY, SURF-SARA, IN2P3-CC, CERN, IFAE-PIC, LAPP, GSI and INFN (CNAF, ROMA and Napoli)

Data management and storage orchestration (Rucio)

File transfer and data movement services (FTS)

Global Data Lake Information System (CRIC)

ESCAPE IAM: common Auth/Authz/IM (AAI)



The ESCAPE Data Lake

The ESCAPE Scientific-Data Lake is a policy-driven, reliable distributed data infrastructure capable of managing Exabyte-scale data sets. And able to deliver data on-demand at low latency to all types of processing facilities

Services operated by the ESCAPE partner institutes

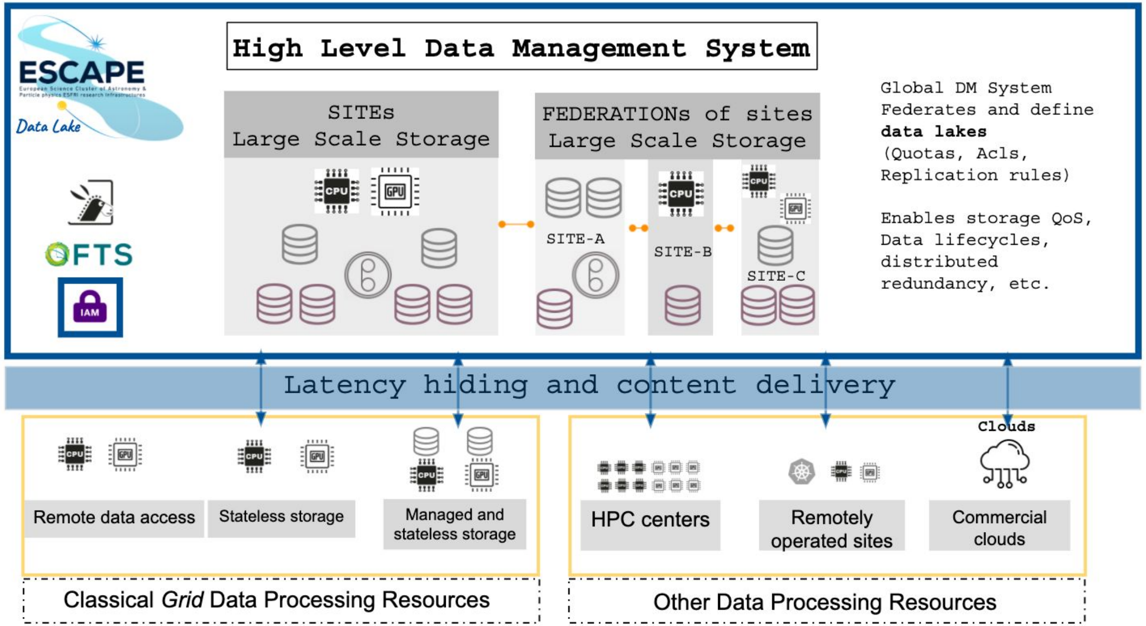
Petabyte scale storage: DESY, SURF-SARA, IN2P3-CC, CERN, IFAE-PIC, LAPP, GSI and INFN (CNAF, ROMA and Napoli)

Data management and storage orchestration (Rucio)

File transfer and data movement services (FTS)

Global Data Lake Information System (CRIC)

ESCAPE IAM: common Auth/Authz/IM (AAI)



ESCAPE DIOS Roadmap

Start 02/2019

Sep 2019

Pilot Data Lake design and implementation plan ✓

Aug 2020

Initial Data Lake implementation with at least 3 centers ✓

Nov 2020

FDR!

Pilot Data lake performance testing and analysis ✓

Jan 2021

Analysis and assessment of the pilot Data Lake ✓

Feb 2021

Nov 2021

DAC21!

Working infrastructure. Verify RI data access from compute platforms: **clouds and HPCs** ✓

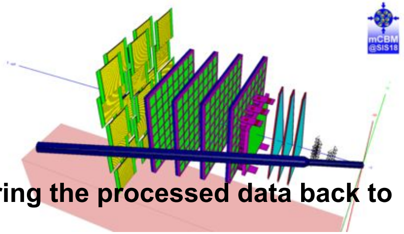
Jun 2022

Preparing post-ESCAPE scenario:
Keeping scientific community engaged
DIOS model sustainability

End 08/2022+6M

Putting the system to work (1/5)

FAIR



- **Registration of RAW data acquired by the mCBM detector on FAIR-ROOT**
- Ingestion and replication of simulated R3B data
- Ingestion and replication of simulated and digitised raw PANDA fallback data
- Particle-transport and digitisation of Monte-Carlo events
- Live ingestion of simulated data
- **Retrieval of stored RAW data from the data-lake, processing of the data and storing the processed data back to the data lake**
- Retrieve raw mCBM data from the data lake, run reconstruction on it and store the results
- Analyse simulated R3B data stored in the data lake, upload resulting histograms and bitmaps to the DL

Raw data injected, stored and preserved in the DL. Data processed by users, results are stored back in the DL.

KM3NeT

- **Ingestion of raw data from the storage at the KM3Net shore station to the Data Lake, and policy-based data replication across the Data Lake infrastructure**

Offload data from the storage buffer in the coast, replicate across sites, run data calibration, store back. Data product ready for user consumption

cta cherenkov telescope array

- Long-haul transfer and replication. CTA-RUCIO @PIC: non-deterministic (La Palma) and deterministic (PIC) RSEs
- **Data reprocessing. Primary data stored and findable in the datalake (using the CTA Rucio instance). Data is accessed and processed. New data products stored back in the Data Lake**
- Data analysis. Data access via Jupyterhub/mybinder via ESAP. Higher-level analysis products produced

Distributed data re-processing taken at remote locations

Putting the system to work (2/5)

HiLumi HL-LHC PROJECT

ATLAS EXPERIMENT

- Exercises (data production, replication and documentation) before and during the DAC21. Include the creation of datasets for real-kind final user analysis examples using current open-access datasets. ~200*10 = 2000 files uploaded in the Datalake. Two copies of such files (rules) into at least two RSE's
- **User analysis pipeline tests on experimental particle physics by using augmented open data** (http://opendata.atlas.cern/software/). Testing and validating the reading access of the samples via de Jupyter rucio extension, and running multiple analysis pipelines.

MAGIC Major Atmospheric Gamma Imaging Cerenkov Telescopes

- **Long haul raw data ingestion and replication. Data is successfully transferred from the telescope station and replicated to the Data Lake, file deleted on the telescope storage buffer.**
- Data transfer monitored. Data can be discovered using the CTA-RUCIO instance.

Data management from remote locations

of DL3 file. Validating the reading access of the samples via ltiple analysis using gammapy library.

LOFAR

- **Ingestion of LOFAR data from a remote site** to the Data Lake. Data transfer and replication into off-site storage, after successful replication delete data at the source
- **Process data in the Data Lake at an external location, combine results with other astronomical data to produce a multiwavelength image.**
- Include a read-only RSE to a location outside the data lake. Get data from there into the DL.
- Extending use cases by using larger files and leveraging several QoS, running all processing in the DLaaS, requiring a the availability of specific LOFAR software in the DLaaS.

Full-cycle scientific data management and data processing

Putting the system to work (3/5)

SKAO

- Data replication. Data in correct place in timely manner.
- **Long haul data replication. SKAO Rucio (Australia and South-Africa to UK RSEs), using the RUCIO SKA instance.**
- End-to-end proof of concept data lifecycle test, AUS/SA to northern hemisphere sites
- Data Analysis. Successfully running SKAO's science data challenge pipeline using data stored in datalake.

HiLumi HL-LHC PROJECT

CMS

- **Multi purpose Analysis Facility PoC with data access via DASK (workload orchestrator) leveraging computing at Marconi (HPC) and large batch clusters**
- Access control for embargo data, test in CNAF and DESY
- Content delivery and caching: XCache Protocol Translation: xroot internal vs http External for Data Lake data transfer. Performance comparisons for Analysis workflows

DL interface with local and heterogeneous resources, CDN and caching

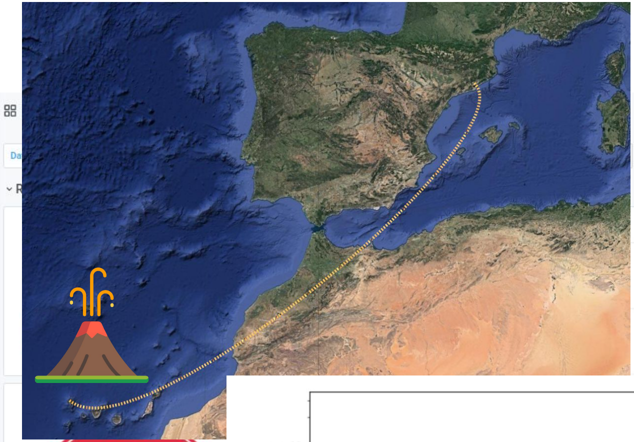
LSST

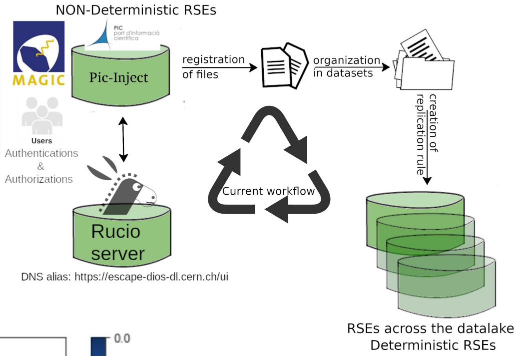
VERA C. RUBIN OBSERVATORY

- **Simulate replication of one night's worth of raw images data between two Vera C. Rubin data facilities, perform the exercise several times.** Each iteration is composed of 15TB, 800k files, ideally to be replicated in 12 hours or less
- Incorporate SLAC National Accelerator Laboratory (US) in the data replication chain (postponed)

Leverage telescope local storage data replication to fulfill daily data management cycles

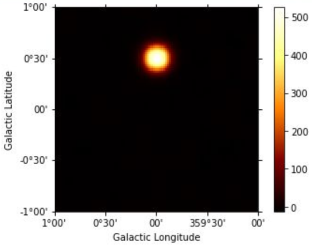
Putting the system to work (4/5)






```
[38]: # we can also compute the significance of our source
analysis.get_excess_map()
analysis.excess_map["sqrt_ts"].plot(add_cbar=True);
```

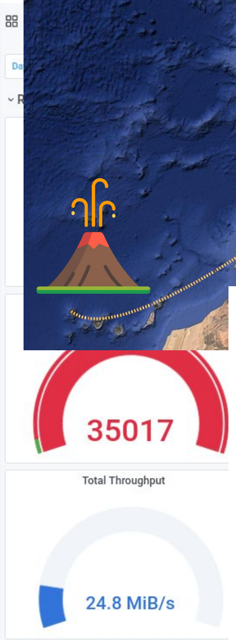
```
Computing excess maps.
Position <SkyCoord (Galactic): (l, b) in deg
    (0., 0.)> is outside valid IRF map range, using nearest IRF defined within
```

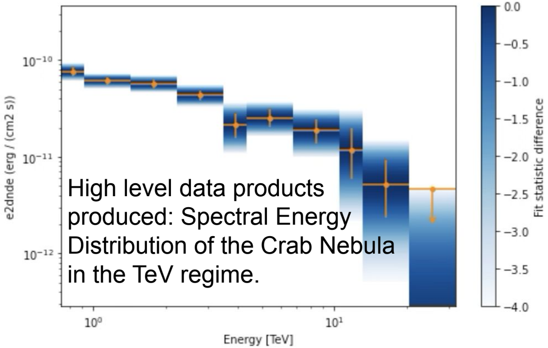


perform the fit

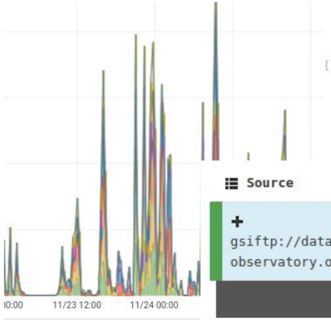
As a final step we fit the spectrum of the source, and we compare to the one we actually used for simulation

```
[43]: # let us load the model we used for the simulation
models = Models.read("./data/models/point-source-pwl.yaml")
# let us create a copy of the spectral model for later comparison
original_spectral_model = models[0].spectral_model.copy()
```

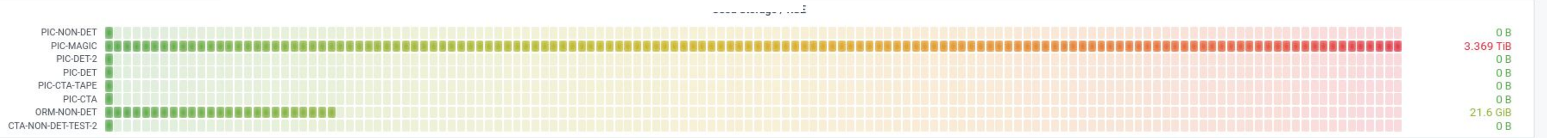




High level data products produced: Spectral Energy Distribution of the Crab Nebula in the TeV regime.



Source	Destination	VO	Submitted	Active
gsiftp://datatransfer.ctan.cta-observatory.org	gsiftp://door05.pic.es	pic01-rucio-server.pic	5589	129
			5589	129



Putting the system to work (5/5)

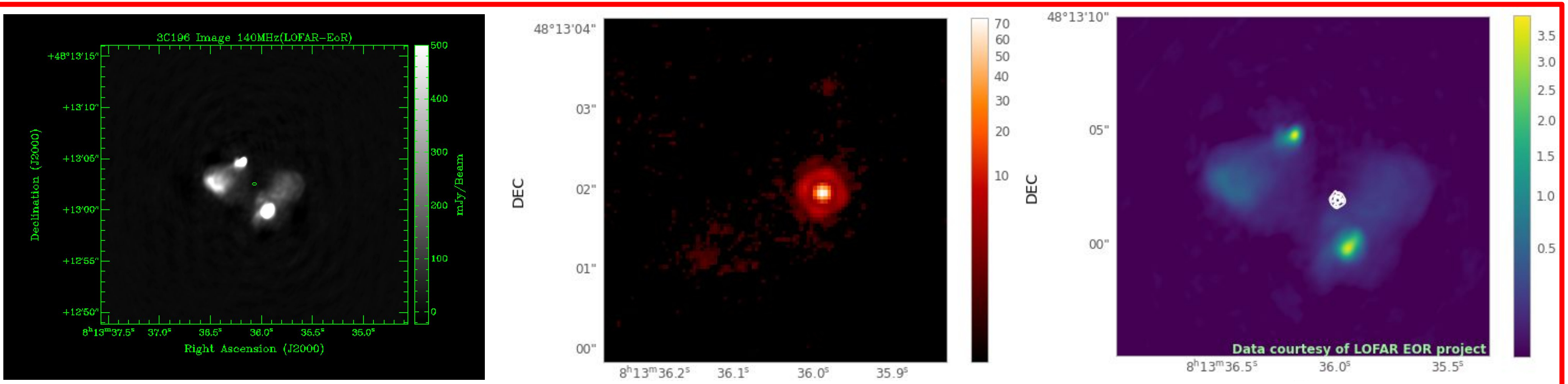
1. Data injected to the DL from **three** radio source observations in external locations

2. User in external location download the data, process and store results back to the DL

3. User interested in combining results stored with other public data to cover also visible spectrum

4. Combined optical data from the Hubble located via the **VO (WP4)**

5. Optical and radio data aggregate in via the **ESAP (WP5), combined analysis done.** Results uploaded back to the DL.



From left to right: Radio image, Optical image and the Combined image (LOFAR with optical contours)

Some items for discussion inspired by the ESCAPE experience

Aim:

- See if there is interest from the experiments?
- See if there is interest from the sites?
- See if there is interest from the software/services/tools providers?
- Could DOMA be the place to start/coordinate some of them?

For discussion: 1 - Joining Efforts

- Large computing Commonalities across sciences. Still the supporting “sites” are quite the same. Fostering homogeneity of tools, frameworks and services might lead to a more sustainable scenarios
  - Software and services: cross collaborations and synergies to enrich tools and enforce support/coding community, e.g metadata, Rucio, FTS
  - Give value to the technical expertise at the sites: foster similar technologies, industry standards that are attractive for engineers and maximise job opportunities
  - Ease integration of diverse computing infrastructures, e.g. enforcing coordination/mutual benefit of HPC integration work, eg. FENIX project-ESCAPE, LHC experiments and related HPCs.

For discussion: 2 - User Analysis “portals”

- Analysis Platforms (Environments, Facilities or Frameworks): User’s portals, visual working modes
  - Evolving frameworks controlled by the experiments, user friendly portals, all-included: software, code, data browsing, etc.
  - Interesting for learning, outreach, citizen science and some user analysis
- Step forward would be to be able to “encapsulate” jobs (data surls, code, software) to scale out
  - User test code and cuts on few events, then scale out when happy to backend computing resources, local or external.
- Facilitator for Open science and Open Data portals and Citizen science campaigns
  - Fostering reusability and reproducibility, e.g. REANA

For discussion: 3 - Focussing Resources

- Towards “Specialisation” of sites?
  - Following developments and needs of the evolving Analysis Platforms
  - Running rigid-old-style storage becoming less attractive and effective?
  - Content delivery and active/streaming caches?
  - More useful to the experiment community to have more tailored computing resources?
    - GPUs, AI/ML services, accessible also through Analysis Environments/Frameworks?

- Time to re-discuss storage consolidation ?

For discussion: 4 - Strengthen a *global* AAI collaboration

- AAI: A common trust layer is fundamental. Many activities and efforts.
  - Feeling efforts/coordination is still to siloed/scattered?
    - Token definitions?
    - Mismatch: services vs. token providers?
    - Flavors/technologies?
- The way out of x509 should not be horribly painful and what we got should not be horribly difficult…
  - Should be EASY to accommodate new communities, services and resources.

Thank you