



ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

Building exabyte-scale federated storage for the ESFRI communities

European HDF5 Users Group (HUG) meeting 2021

<https://www.hdfgroup.org/hug/europeanhug21/>

Paul Millar (on behalf of WP2) <paul.millar@desy.de>



ESCAPE in a nutshell

ESCAPE - <https://projectescape.eu/> - convenes a large scientific community

- **31** partners (including 2 SMEs), representing:
 - **7** ESFRI projects & landmarks: CTA, ELT, EST, FAIR, HL-LHC, KM3NeT, SKA
 - **2** pan-European International Organizations: CERN, ESO.
 - **4** supporting ERA-NET initiatives: HEP (CERN), NuPECC, ASTRONET, APPEC
 - **1** involved initiative/infrastructure: EURO-VO
 - **2** European research infrastructures: EGO and JIV-ERIC
- Budget: **16 M€**
- Started: **2019-02-01**
- Ends: **2023-01-31**
- Coordinator: **CNRS**

Details courtesy of Kay Graf





MAX-PLANCK-GESELLSCHAFT



Heidelberg Institute for Theoretical Studies



THE UNIVERSITY of EDINBURGH



Details courtesy of Kay Graf



ESCAPE Partner RIs

Radio



JIVE-VLBI

SKA

Visible light



ELT

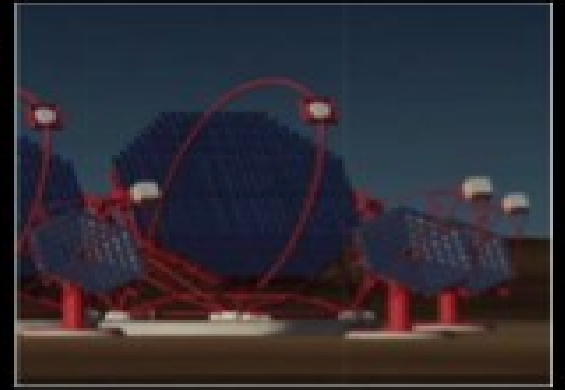


ESO



EST

Gamma rays

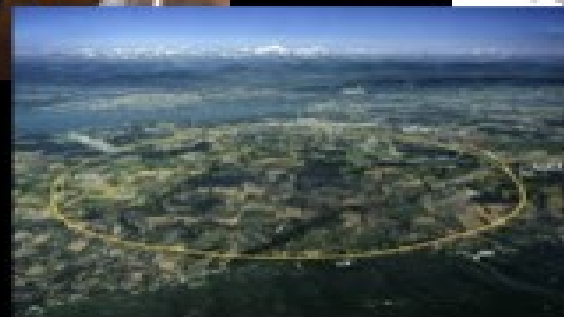


CTA

Accelerator-based Particle Physics

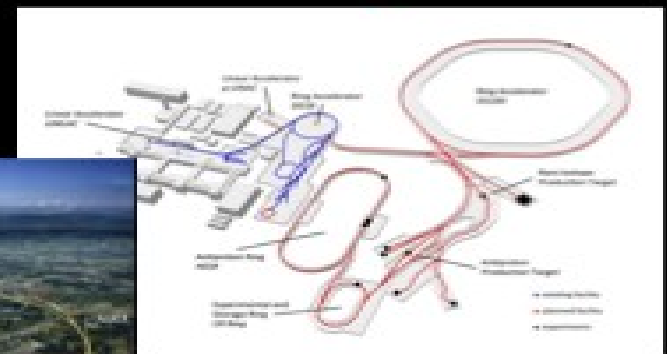


HL-LHC



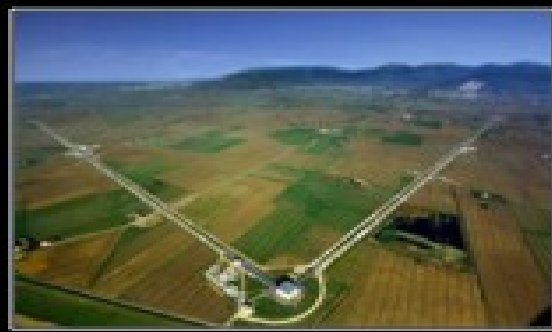
CERN

Accelerator-based Nuclear Physics



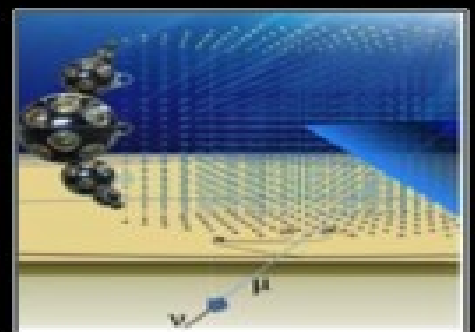
FAIR

Gravitational Waves

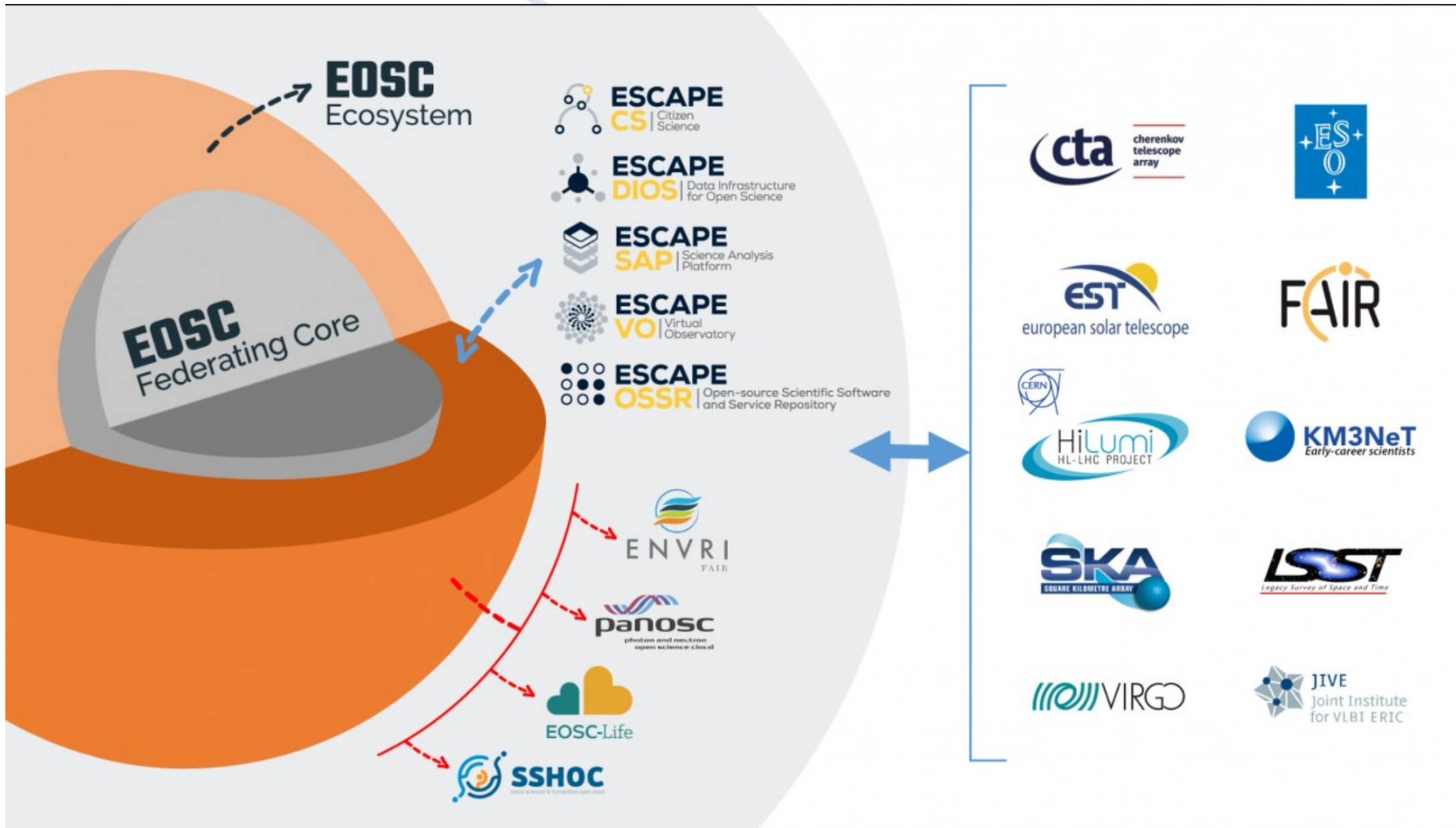


EGO-VIRGO

Cosmic-rays Neutrinos



KM3NeT



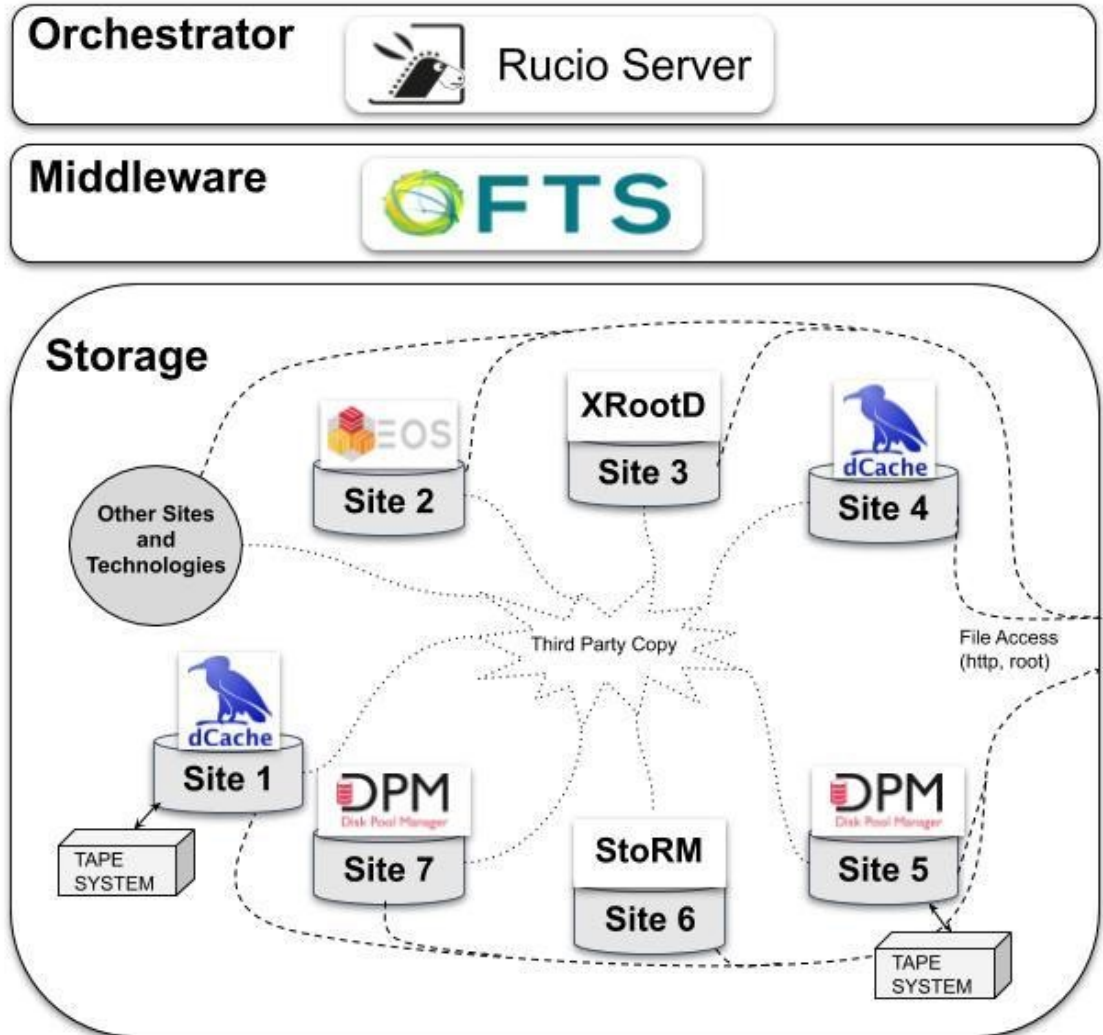
DIOS (“WP2”) overview

- Data Infrastructure for Open Science (DIOS): a federated data infrastructure for open access that can scale up to multi-Exabyte capacity.
- Based on the Data Lake concept, which is distilled from decades of production experience in WLCG (LHC / CERN).
- The work is split into five “tasks”:
 - Data lake infrastructure and federated services,
 - Data lake orchestration service,
 - Integration with Compute Services,
 - Networking,
 - Authentication & Authorisation Mechanisms.



DataLake overview

- DataLake concept involves **abstracting** data locality.
- Data is placed according to rules in **Rucio**.
- **FTS** reliably transfers data.
- **Caches** may be deployed near computing resources.
- Jobs can read data “directly” from DataLake (perhaps through a cache) or download all data locally.
- Output may be written into the DataLake



How HDF5 is being used?

- **KM3NeT**: "mostly" for publication-level data. Underlying data format is ROOT, but for the highest level we are aiming for HDF5 + others ([further details](#)).
- **CTA**: has made no decision on which format will be used. CTA is investigating HDF5, which is a clear contender.
- **MAGIC**: doesn't use HDF5; however, the same people are involved with a project where HDF5 bioimaging data from ICFO (The Institute of Photonic Sciences) is written into Rucio.
- **SKA**: is using it experimentally in a few cases from things like a file format for an i/o performance test for distributed degridding to some input data for simulations (not extensively but there are examples). Future plans: no decision has been made yet.
- **EGO-Virgo**: GWOSC (open science center) is using HDF5 for storing public-access data, otherwise a custom format is used.
- **LOFAR**: currently uses HDF5 along with other formats. Have written [a paper](#) on HDF5 as a standard approach for radio astronomy.



What might be of interest?

- HDF5 libraries are currently POSIX (local filesystem storage) only.
 - Must download complete dataset onto local storage.
- Is there interest in HDF5 libraries supporting remote IO?
 - Allow scientific workflows to take advantage of Data Lake with few changes.
 - Remote reading
 - Some work-flows read only part of a file: downloading the complete file is wasteful
 - Replica discovery – learning where data is available
 - Read-ahead: fetch data while processing chunks (latency hiding)
 - Exploring different caching options.
 - Remote writing (simplify creating HDF5 file within the data-lake)
- Metadata handling: Is there scope for pushing metadata into Rucio
 - This would avoid downloading data that turns out not to be useful.



Conclusions

- **ESCAPE** is a EU-funded project to build a blueprint for next-generation, cross-disciplinary computing and storage resources to scientific communities.
- **The Data Lake** is a key concept for managing federated storage at multiple sites.
- Various ESCAPE scientific communities are **adopting or investigating HDF5** as a file format.
- Although the Data Lake is largely file-format agnostic, there is **potential for collaboration**.



Thanks for listening!



Backup slides



WP2 Testbed

