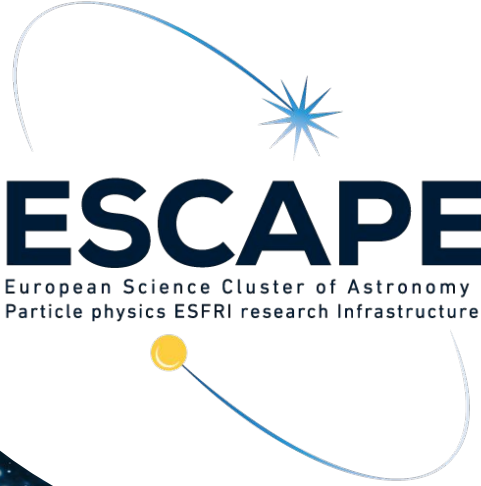


The logo graphic for ESCAPE features a stylized blue starburst at the top, a blue orbital line curving around it, and a small yellow circle at the bottom left of the orbit.

ESCAPE

European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

ESCAPE DIOS: A Data Lake for Open Science

Xavier Espinal (CERN) - ESCAPE WP2 lead

ESCAPE GA morning, 29 September 2021



Introduction

- For the first time flagship **ESFRIs** are **collaborating together since the real beginning** to implement the standards of data FAIRness and Open Access in a common scientific data management infrastructure
- **Breaking historic sociological barriers** among different scientific communities in the search for commonalities in their computing models and potential synergies
- Making use of existing dataset from precursor experiments in radio-astronomy and astro-particle physics to **validate the next generation Open Science models**



The DIOS Data Infrastructure for Open Science (DIOS)

- The DIOS Data Infrastructure for Open Science (DIOS) WP aims at **delivering a prototype of the Data Lake** concept, a common storage infrastructure that:
 - Provides **global data management** orchestration capabilities: rules, policies and storage quality of service
 - Enables **Open Science** for multi-disciplinary RIs, and by construction implementing **FAIR** principles and fostering **data openness**
 - Addressing **Exabyte-scale** needs of large science projects but also show adaptability to **modest** storage and data management needs
 - Enhance the use and the potential of **Content Delivery Networks**



The DIOS Data Infrastructure for Open Science (DIOS)

- The main Data Lake building blocks are deployed and **operated by the ESCAPE partner institutes**
 - Storage services
 - Petabyte scale capacity available deployed at **DESY, SURF-SARA, IN2P3-CC, CERN, IFAE-PIC, LAPP, GSI and INFN(CNAF, ROMA and Napoli)**
 - Data management and orchestration service: **Rucio** (from ATLAS/LHC)
 - Instances running now at **CERN, SKAO, IN2P3 and PIC**
 - File transfer and data movement: **FTS** (from WLCG/CERN)
 - Instances running at **CERN and PIC** and making use of different protocols: **http, gridftp and xrootd (and swift for HPCs)**
 - Content delivery and latency hiding: **XCache** (from XRootD/SLAC)
 - Deployments at **GSI, CERN, LAPP and IN2P3**
 - Global Information System: **CRIC** (from ATLAS/LHC) deployed at **CERN**
 - Auth/Authz/IM (AAI): Indigo **IAM** (from INDIGO-DATACLOUD, AARC/AARC2 and FIM4R projects). The ESCAPE IAM instance is deployed and operated in **INFN/CNAF**

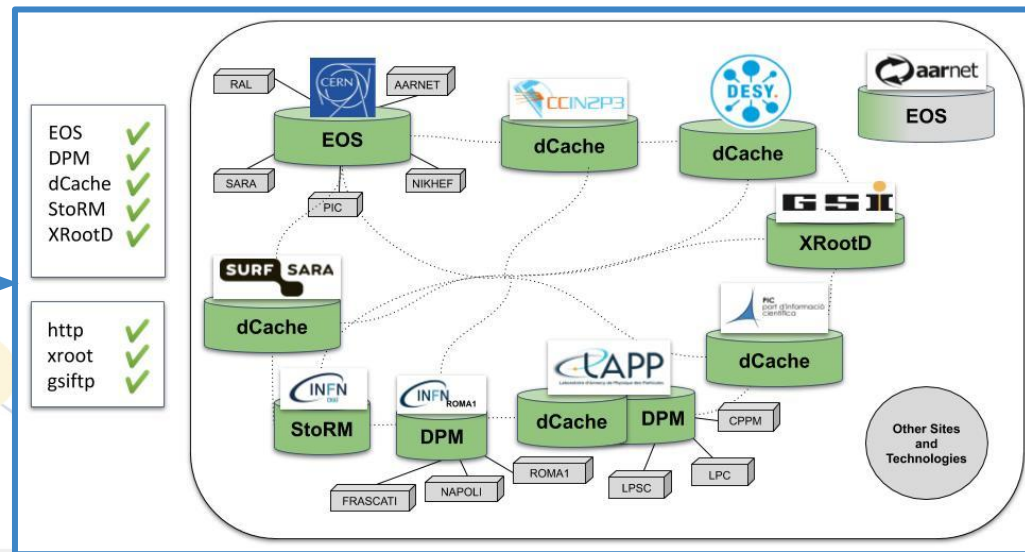
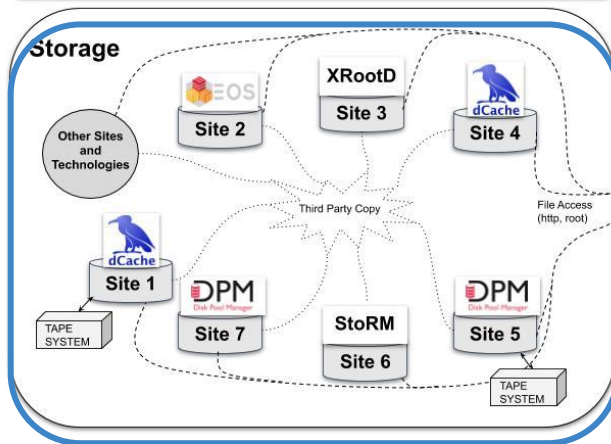


Achievements (1/6) - A functional Data Lake pilot

- Pilot datalake with **10** storage endpoints **functional**
 - INFN-CNAF, INFN-ROMA, INFN-Napoli, DESY, SURF-SARA, IN2P3-CC, CERN, IFAE-PIC, LAPP and GSI
- The high level Data Lake orchestration layer is **consolidated**

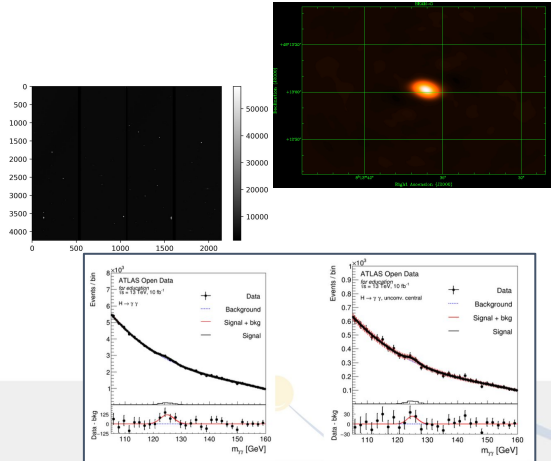
Orchestrator Rucio Server

Middleware FTS

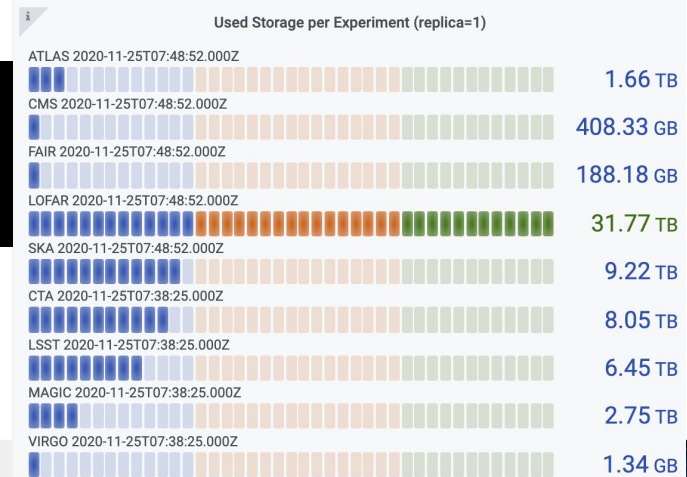


Achievements (2/6) - **SCIENCE!**

- **Strong** involvement from the **RIs and experiments**:
 - Full Dress Rehearsal exercise Nov-2020. Participation and strong involvement from: **LOFAR, CTA, MAGIC, FAIR, SKA, Vera Rubin Observatory (LSST), ATLAS, CMS** and **EGO/VIRGO**
- Data from flagship physics experiments of different disciplines **together on a common data infrastructure**:
 - Astro-particle
 - Radio-astronomy
 - Gravitational Waves
 - Cosmology
 - Particle Physics



DIDs per Experiment (replica=1)			
Experiment	Number of DIDs	Number of files	Number of datasets
LOFAR	24.9 K	24.9 K	4
FAIR	185	183	2
CMS	401	398	3
MAGIC	23.8 K	5.857 K	17.9 K
ATLAS	7.441 K	7.051 K	390
LSST	350 K	350 K	8
CTA	566 K	564 K	1.630 K
SKA	3.379 Mil	3.342 Mil	37.7 K
VIRGO	15.6 K	15.6 K	1



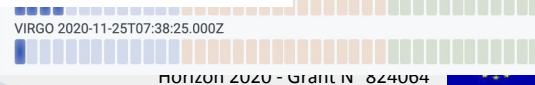
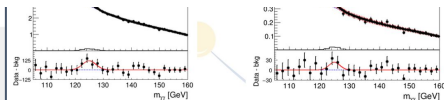
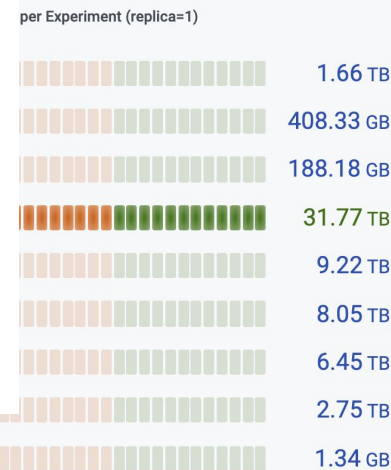
Achievements (2/6) - **SCIENCE!**

- **Strong** involvement from the **RIs and experiments**:
 - Full Dress Rehearsal exercise Nov-2020 Participation

DIDs per Experiment (replica=1)			
Experiment	Number of DIDs	Number of files	Number of datasets
LOFAR	24.9 K	24.9 K	4
FAIR	185	183	2
CMS	401	398	3

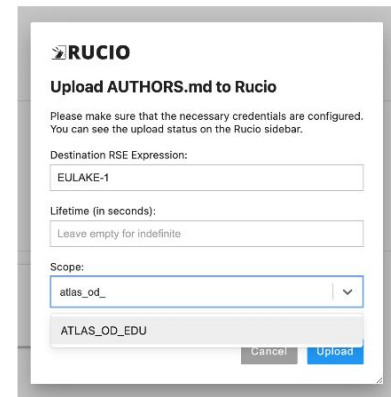
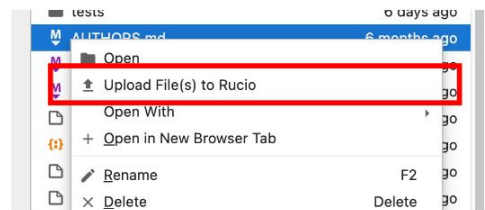
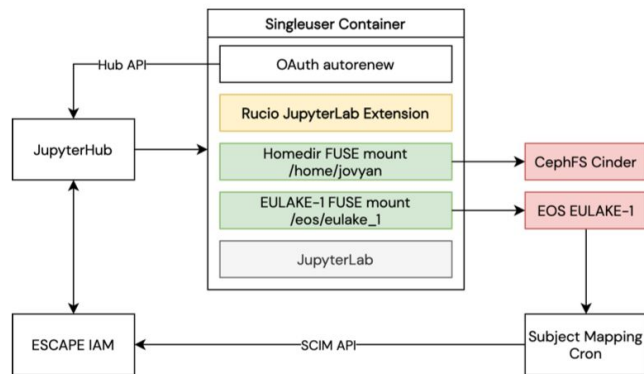
But we do not live from the past...

Full Scale Exercise in one month:
Data and Analysis Challenge (DAC21)



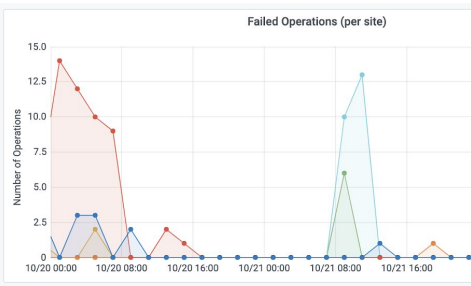
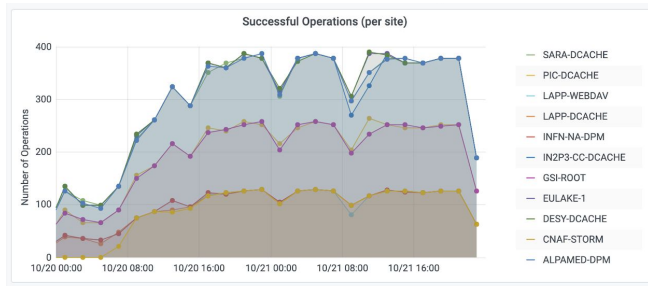
Achievements (3/6) - Enabling Data Processing and Analysis

- Step further notebook integration providing a **Data Lake as a Service (DLaaS)** platform. Demo presented yesterday at the Extended Discussion Day [\[link\]](#)
- Deployed in Kubernetes, using Zero-to-JupyterHub Helm chart (<https://escape-notebook.cern.ch>)
- OAuth authentication using ESCAPE IAM.
 - X509 and user/pass are still supported
- Uses Rucio JupyterLab Extension in Replica mode (i.e. TPC to local storage)
 - Connected to ESCAPE Data Lake(escape-rucio.cern.ch)
 - Automatically preconfigured to use token (OIDC)
 - FUSE mount ability to the local storage services
 - Making files available means creating a replication rule to move files to EULAKE-1
 - Direct download mode is still possible

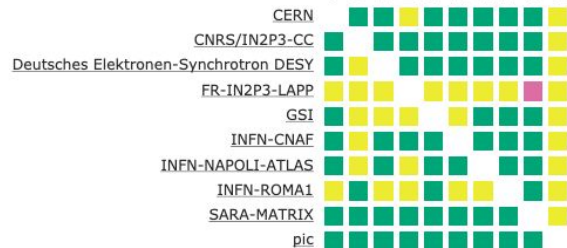
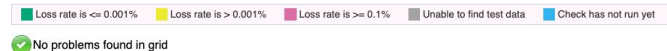


Achievements (4/6) - Enabling Knowledge Transfer

- Set-up a transversal deployment and operations team, composed by participating **partners and experiments members**
 - Looking after Data Lake operations. Crucial for project sustainability
- Data transfer **commissioning tools** and **network health dashboard** in place

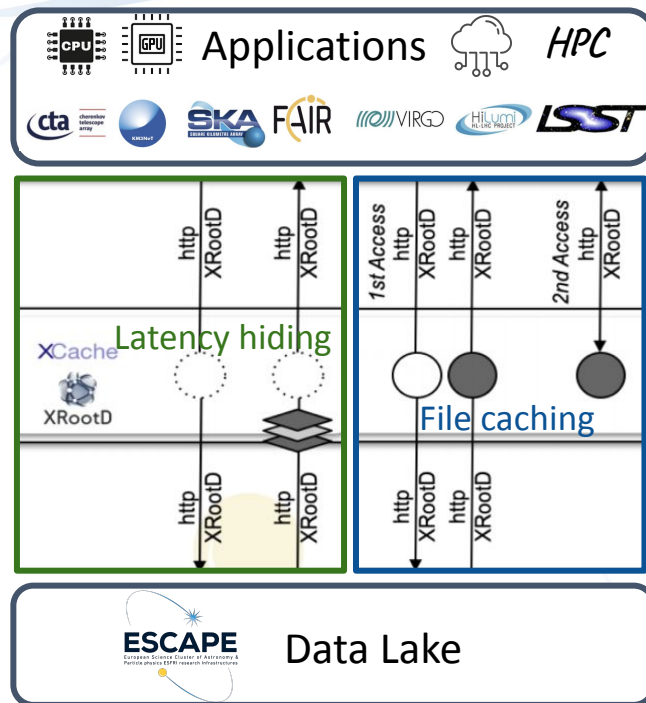


ESCAPE Mesh Config - ESCAPE Latency - Loss



Achievements (5/6) - Opening Data Access

- Content delivery and latency hiding prototypes
 - Streaming caches demonstrated potential on **latency hiding and file re-usability** in HEP
 - Caches enabled via notebook with the DLaaS
 - Potentially instrumental** technology to enable external compute resources and HPC initiatives **FENIX** project and PRACE/GEANT/SKA/CERN initiative
- Effort made to reduce complexity for **easy deployment**
 - containers and open source repositories (possibility for WP3)



Achievements (6/6) - **A common identity**

- **AAI infrastructure deployed and consolidated:**
 - ESCAPE IAM instance **deployed and integrated** with EduGAIN
 - **Strong user enrollment.** GSI/VOMS authN/Z in place and working, token-based access deployment progressing
- **Namespace** organization proposal and deployment strategy defined.
 - Incremental steps towards fine-grained authz (e.g. enabling implementation of embargo data)
- Developed a **test suite** assessing AAI infrastructure works as expected
 - Data access and namespace ACLs, storage endpoints endpoints auth, obtaining credentials, etc.



Welcome to **escape**

Sign in with your escape credentials

	xespinal	
		

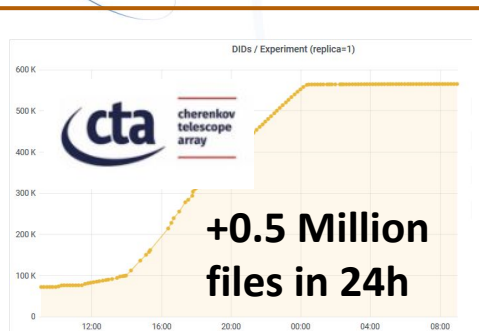
Sign in

[Forgot your password?](#)

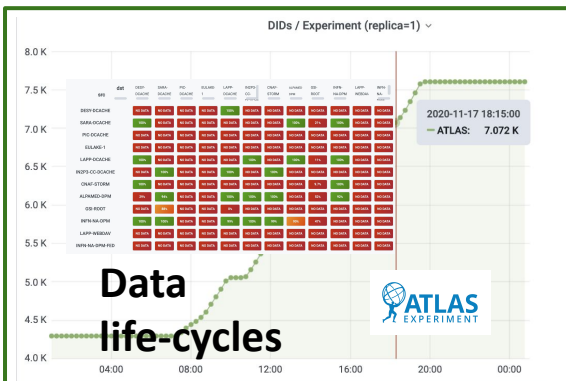
Or sign in with



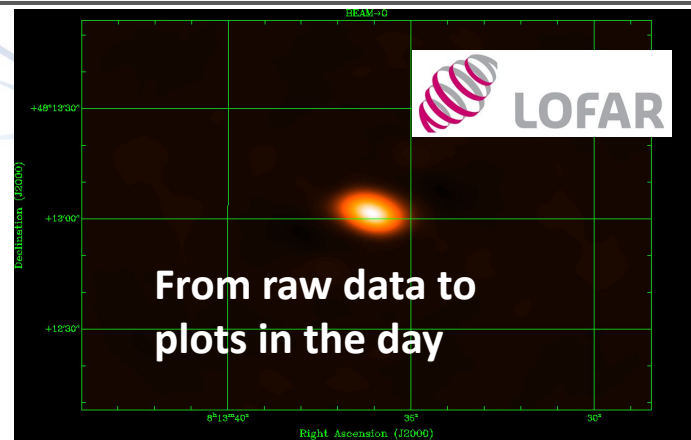
Data Lake 24-hour Dress Rehearsal 17 Nov 2020



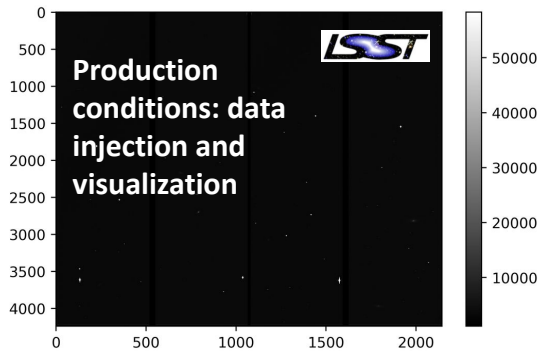
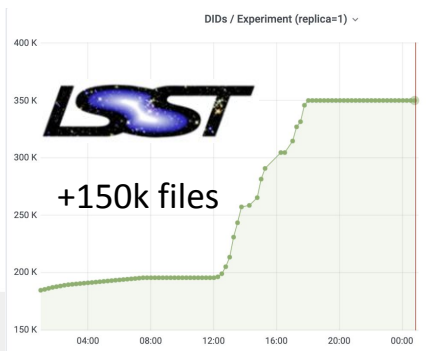
CTA: Simulate a night data captured from telescope in Canary Island for 6 h: ingest 500 Dataset of 10 files.



ATLAS: Storage QoS functionality tests: upload files from LAPP cluster to ALPAMED-DPM (FRANCE) and INFN-NA-DPM (ITALY), then request transfer to 1 RSE **QoS=SAFE** and 2 RSEs **QoS=CHEAP-ANALYSIS**



LOFAR: astronomical radio source 3C196 made using LOFAR data. The raw visibility data was downloaded via rucio from the EULAKE-1 and processed on Open Nebula at surfsara using the container based LOFAR software

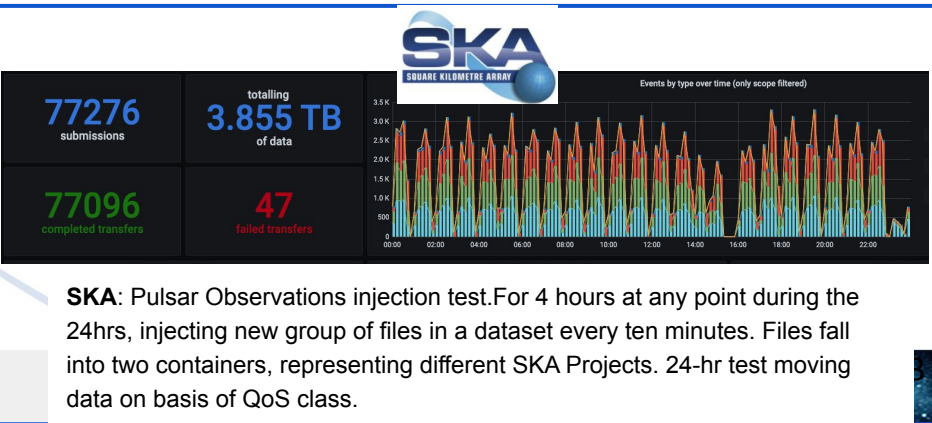
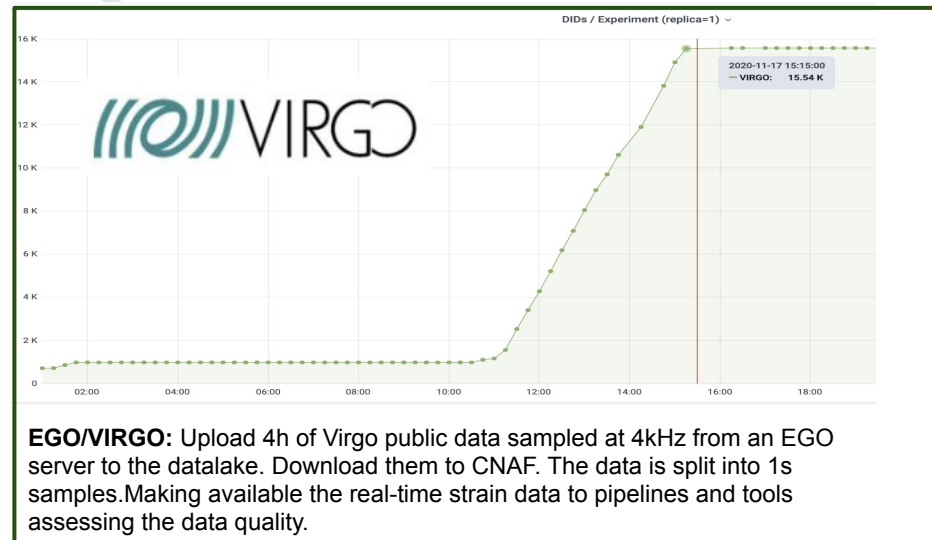
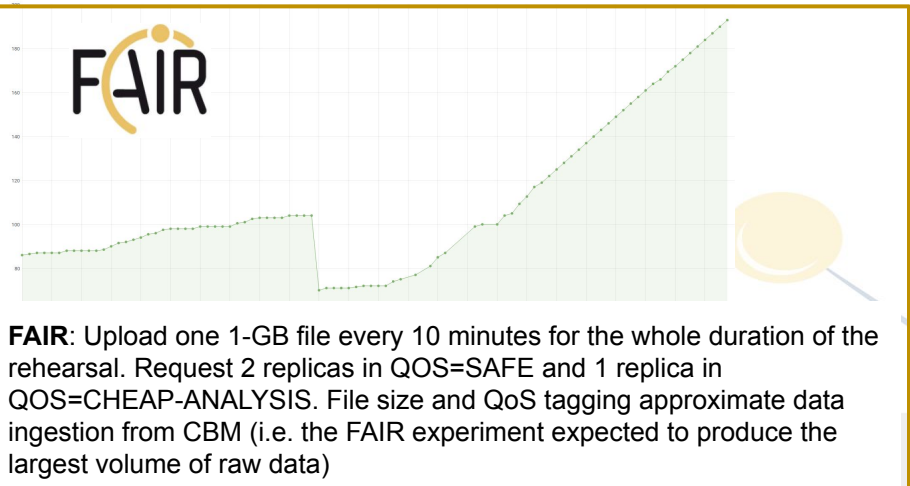
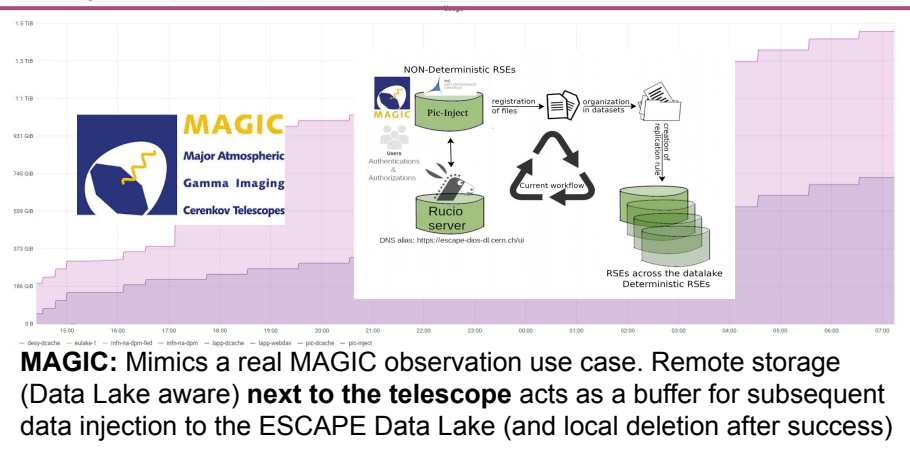


LSST: Simulate production conditions: ingest the HSC RC2 dataset from CC-IN2P3 local storage to the Data Lake, **at a realistic LSST data rate (20TB/24h)**. Then **confirm integrity and accessibility of the data via a notebook**.

→ The image is a reconstruction drawn within a Jupyter Notebook accessing the data used in the Full Dress Rehearsal.



Data Lake 24-hour Dress Rehearsal 17 Nov 2020

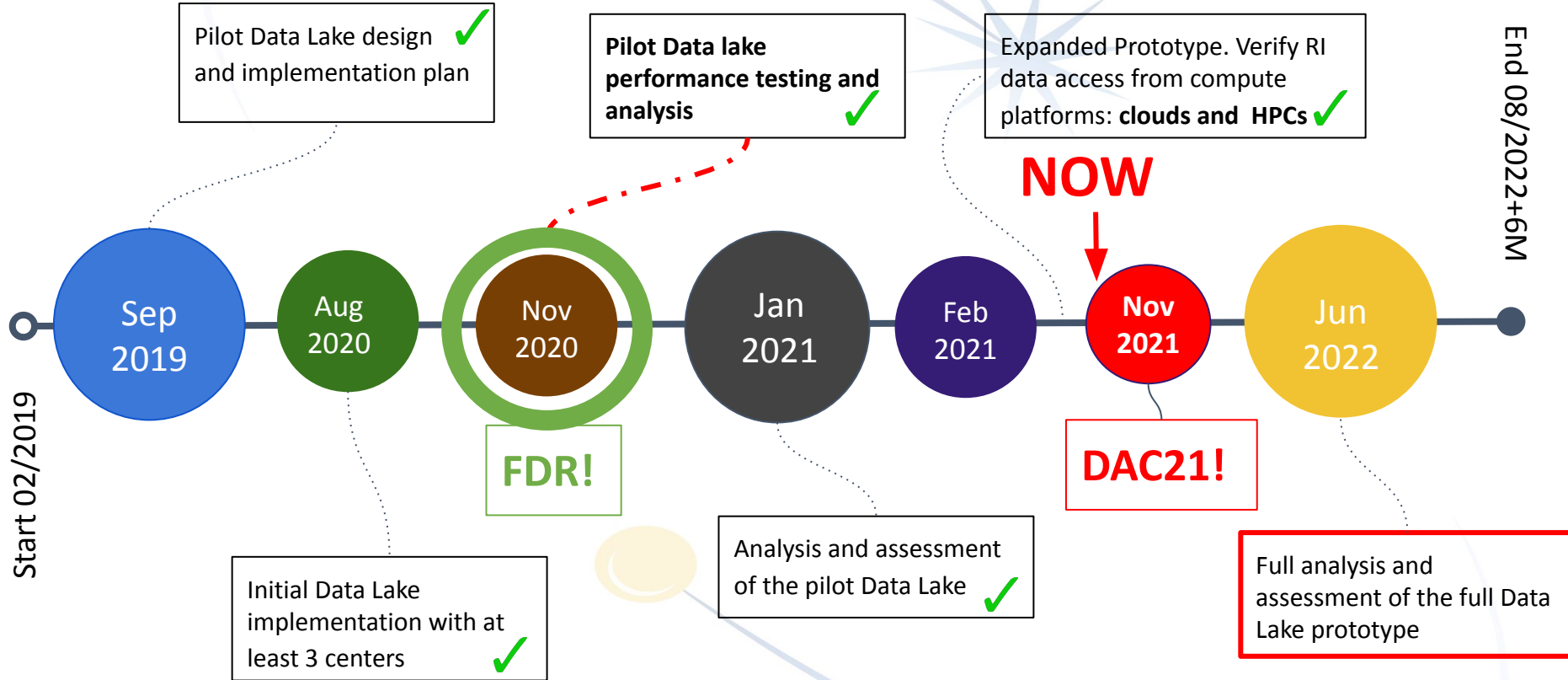


Why ESCAPE DIOS is instrumental now and tomorrow?

- **ESCAPE is providing the framework** for the sciences and participating sites to explore and influence on the development of next generation data management tools
 - Managing large datasets and large data volumes
 - RUCIO instances being deployed for by the experiments themselves: SKA, CTA, LSST, MAGIC
 - Specific metadata extension needs for astro-particle and radio-astronomy RUCIO
 - HL-LHC, CTA and SKA active on RUCIO workshops and community (**and IVOA as seen yesterday!**)
 - Data ingestion from remote sources to large shared storage infrastructures Data Lake
 - Vera Rubin Observatory and MAGIC, but also SKA south-africa and Australia
 - Understanding future of storage services at the sites, different storage backends and **technologies fulfilling the needs of experiment's data lifecycles**
 - Several experiment lifecycles tested in the FDR and a pillar activity for the DAC21
- In a challenging future ESCAPE provided an invaluable starting point. An all-together spirit for next generation scientific data needs.
- ESCAPE is setting the scene for further projects and collaborations with the right spirit: **sites, experiments and service providers together.**



Summary WP2 Roadmap



Next steps (1/2)

- DAC21 (=extended FDR, in extension, scope and goals)
 - **Running production** workloads of the participating experiments, including data management and analysis , in the prototype Data Lake (w/WP3,WP4,WP5)
 - Data acquisition, injection, replication, life-cycles and data policies
 - **Data Processing:** Data Preparation and User Analysis (=>w/WP3)
 - **Demonstrate Data Lake sustainability** after ESCAPE (=> towards EOSC)
 - ie. some experiment's deploying their own RUCIO integrated with the ESCAPE data lake.
 - End-to-End AAI proven for all sciences and workflows exercised at the DAC21.
 - Assessment ranging from the experiment experts to end users. Token integration is progressing adequately.
 - Connection with networking groups, develop long-haul data transfer plans: inclusion of Australian storage in ESCAPE rucio instances for SKA.



Next steps (2/2)

- **Data Lake as a Service: a step further on User vs. Data Lake experience.**
 - Explore Further integration with ESAP, catalogues and data/metadata sources
 - Expand deployments and usage, key abilities are implemented and are functional:
 - OIDC (token based AAI) integration
 - Data browsing, file moving, data download and upload
 - Storage cache layer interplay
 - FUSE mount (posix filesystem mount of the local storage), terminal capabilities (RUCIO-cli)
 - Provide the end-user with a Notebook ready-to-be-used and fully integrated with the Data Lake
 - DLaaS demo shown [yesterday](#) [available [here](#)]
- **Synergies with external projects and scientific communities**
 - **FENIX** (HPCs) infrastructure (<https://fenix-ri.eu/>) - interest in collaborating with Science Clusters. Collaboration ongoing to enable ESCAPE Data Lake from/to FENIX infrastructure. A data transfer demonstrator with CSCS HPC ongoing. Accounts and storage space created. Token delegation/orchestration ongoing.
 - **CS3MESH4EOSC** - EC project. Federated sync&share file service providing a collaboration platform for hundreds of thousands of users. Commonalities for open science data access, specially on DLaaS and notebook/analysis approaches.
 - **PANOSC/EXPANDS**- EC project. Common PaN ESCAPE Data Management Workshop (12/20) [[link](#)] Potential synergies enlarging usage scope originally planned for the ESCAPE Data Lake prototype
 - **PUNCH4NFDI**: German National Research Data Infrastructure consortium of particle, astro-, astroparticle, hadron and nuclear physics. ESCAPE invited to give a webinar.





ESCAPE
European Science Cluster of Astronomy &
Particle physics ESFRI research Infrastructures

