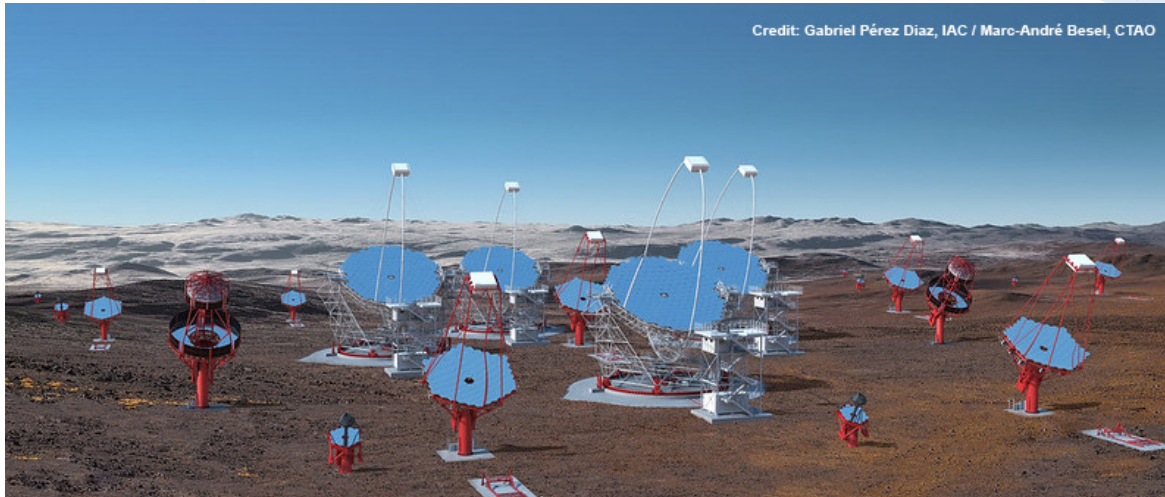# Cherenkov Telescope Array

Nadine Neyroud, Agustin Bruzzese, Frederic Gillardo, Jordi Delgado,
Luisa Arrabito, Berkay Turk & Gonzalo Merino
(Gareth Hughes)

Experimental Setup

Credit: Gabriel Pérez Díaz, IAC / Marc-André Besel, CTAO

[Science of CTA](#)

- Two sites taking data with the data centres in Europe
- O(10TB) of raw reduced data required to be transferred 'daily'
- Data must be duplicated at data centres before removal at origin

- Raw data is processed to science data
- Raw data is stored long term on tape
- Raw data must be periodically reprocessed

CTA DAC21 Use Cases

- ## Use Case 1: Long-haul Transfer
  - Raw data transferred from La Palma to CTA-Rucio instance at PIC
  - Six Tests defined each building on the previous

- ## Use Case 2: Data Reprocessing
  - Using DIRAC: data taken from tape and processed to level science data
  - Two tests defined

- ## Use Case 3: Data Analysis
  - Science data discovered on data lake and analysed

CTA DAC21 Use Case 1 Results

- Setup:
  - CTA Rucio instance at PIC
  - Two types of protocol used (gridFTP & xrootd)
  - Simulated raw CTA data
  - ~2 GB filesize

- Results of tests: ([notebook](#))
  - 1: 300GB from LP to PIC (gridFTP) Disk
  - 2: 300GB from LP to PIC (xrootd) Disk
  - 3: 10TB from LP to PIC (gridFTP & xrootd) Disk
  - 4: 40TB from LP to PIC (gridFTP) Disk
  - 5: 10TB from LP to PIC (gridFTP) Tape
  - 6: 10TB form LP to PIC (gridFTP) and replication to CNAF Disk

- **Test were largely successful**

- Minor issues understood or to be retested
  - 4: Large dataset could not finish in time (ambitious)
  - 6: Issue with syncing of new certificates

- Next steps
  - File deletion at source
  - Second replication site
  - Longer range transfer tests
  - File size range
  - Priority data products

CTA DAC21 Use Case 2 Results

- Using the Belle2 Rucio DIRAC plugin
- Test initially delayed
  - PIC Rucio instance upgraded
    - Second instance being created to be able to run tests simultaneously
- Successes, able to:
  - ingest data using DIRAC
  - launch test jobs on worker nodes
  - Launch CTA production scripts (today!)
- Ongoing work:
  - Ingest using Rucio directly does not work if you want to use it with DIRAC
- Goal: to identify what functionality is needed to integrate (CTA)DIRAC with Rucio
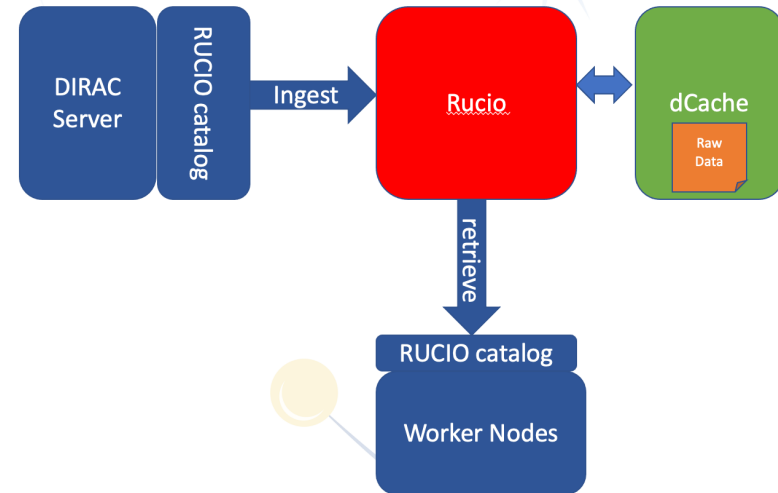
CTA DAC21 Use Case 3 Results

- DAC21 [notebook](#)

- Software not yet fully onboarded so was registered to Science Platform by hand

- Publicly available example notebook and data chosen to run test

- Data uploaded to ESCAPE-Rucio instance using containerized client

- From Science Platform:
  - Software and notebook could be launched
  - Data location could be found and added to shopping basket
    - However data could not be downloaded

- From DLaaS:
  - Data could be downloaded/replicated
  - software needed to installed by hand

Discussion Points
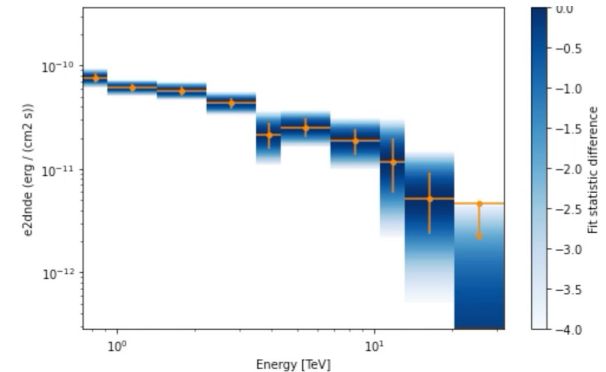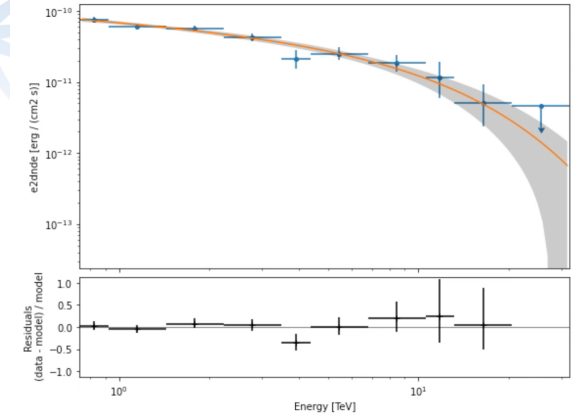
- Things that worked and have been identified to help ESFRI/RIs on Data Management and Data Access:
  - support of OIDC
  - support of metadata complex (as json format)
    - CTA still needs to validate the performance of such features
  - Tape storage
  - Belle2 DIRAC interface

Discussion Points

- Identified barriers to adopt the DIOS model, services or tools:
  - The setup of the entire DIOS solution is still very complex
    - Rucio + RSEs installation + FTS + grafana …

- Suggestions for general improvements on DIOS: model, services, tools:
  - provide a self-contained DIOS as a docker container to ease the setup
    - for proof of concept only, not production
  - A discussion on the OAIS (Open Archival Information System) model for archives in relationship to the data lake
    - does it fit? does it need updating? what is missing from the data lake?
  - Ability to expose data from REST API
  - Documentation & guides
    - some very useful information exists as google docs, perhaps could be formalized

Discussion Points

- Your ESFRI/RI plans regarding the technologies exposed in DIOS, will you consider continue exploring or adopting?
  - Yes, exploring: CTAO is actively correlating our experiences in ESCAPE to our computing requirements (Bulk data and Science data archives)
  - Many requirements have been met, some have been discovered others need further investigation
    - Final decision will be made by CTAO at a later time
  - DAC21 carried out on PIC/CTA-Rucio instance, second instance being built now
    - Follow-up testing will continue throughout 2022
    - including DIRAC integration

- Did you identified security issues? Any specific security related worries to name (present or future) ?
  - No issues identified
  - security always a concern, we would consider adding some tests for this year
    - embargoed data

Discussion Points

- Is your ESFRI/RI Interested in a longer term existence of an ESCAPE or an ESCAPE-like infrastructure?
  - Yes, CTAO will be responsible for the management of its bulk and science archives
  - However having a longer-term existence of the ESCAPE Rucio instance will:
    - good for EOSC-future
    - excellent way to share data across ESFRI/RIs
    - incentives further collaboration between ESFRI/RIs
  - CTA will have the Rucio instance @PIC, but having a longer-term existence of the ESCAPE instance is very useful
    In order to validate the performance, using the following feature/option:
    - testing several RSEs with a large amount of storage available
    - using KB8 to deploy daemons on several PODs
    - using the Grafana monitoring tools

Discussion Points

- Is your ESFRI/RI interested establishing standing collaborations, channels, joint efforts? On which specific topics?
    - Yes! Benefits of the ESCAPE project are clear.
        - Large ESFRIs adopting similar technologies also bring advantages
    - Channels should be selected according to what infrastructure is available and timeframe
    - This would help re-enforce current joint efforts on other topics, with which we are involved

    - CTAO is open to all topics
        - an example would be furthering the DIRAC/Rucio interface, could be a joint effort with other ESFRI/Ris
        - embargoed data

Conclusion

- CTA has learnt a lot from the DIOS experience

- We will continue to work & perform tests through 2022

- We have interest in joint common projects with other ESFRIs (e.g. Rucio-DIRAC)

- Ongoing effort within CTA to transfer knowledge to the observatory

- Thank you to:
  - ESCAPE-CTA, CTA and external colleagues for their hard work
  - DIOS for their active support and engagement