



“ARCHIVE” is not a simple “repository”!

Low-level Big-Data Archival facility needed!

CTA-ASTRI and Miniarray experience!

by Stefano Gallozzi & ASTRI/Miniarray team

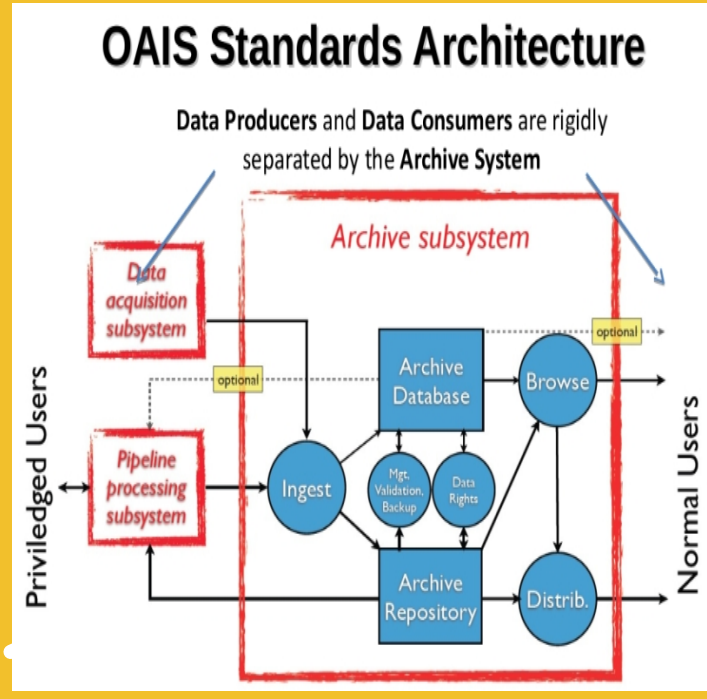
“Archive” is ALWAYS central!



VHE Observatory Subsystems works **around the Archive** facility.

To store, manage and give an easy access to a big amount of data a dedicated (distributed) infrastructure is almost-always needed because horizontal scalability is a fundamental requisite. It also guarantees high availability and no loss of performance nor any single point of failure.

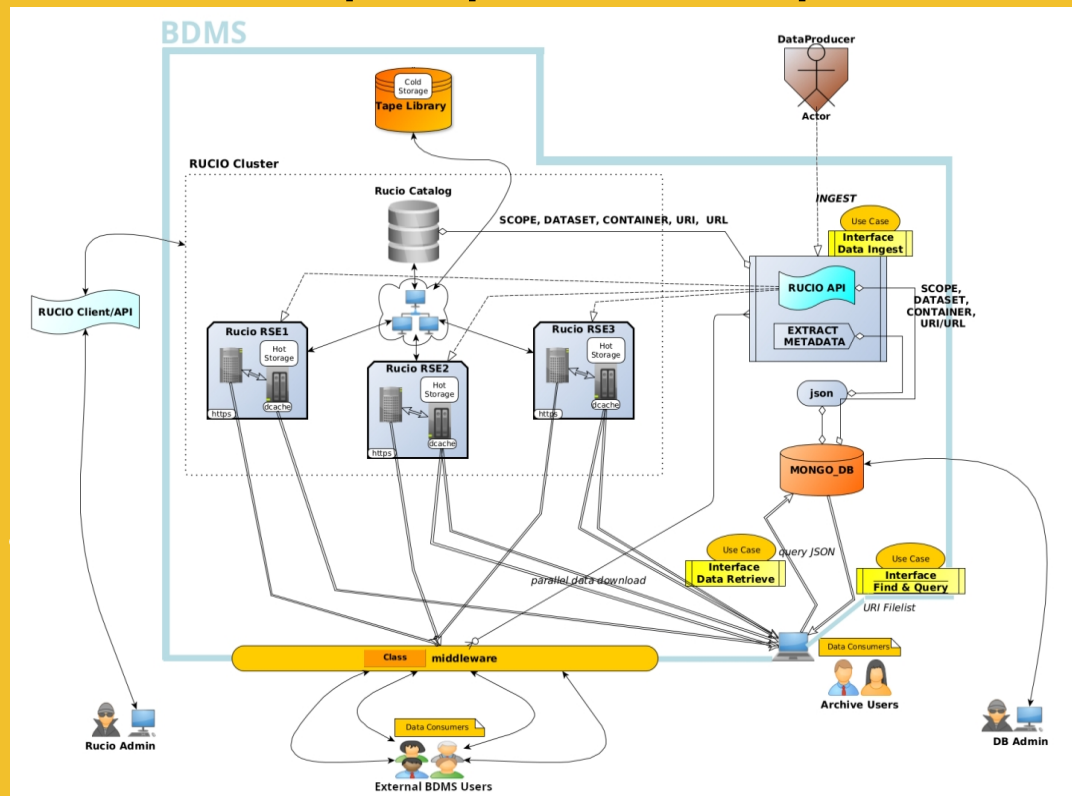
Astronomical observatories usually follows and reflects the standard given by **OAIS-Open Archival Information System**





-Horn & Mini-Array

“our experience”:
a multi-purpose & “open”-archival infrastructure!



**4 European “distributed”
Datacenters** (in Italy, Barcelona, Berlin & Lugano) seen as single Archive Facility to manage, process and replicate low level data. (→ **tens of PB /yr**)

1 centralized Scientific-Archive (in Desy-Zeuthen) to access at high-level Scientific-Data. (→ **few TB/yr**)



Technological "Choices"!

01

Distributed-Storage

Distributed Storage guaranteed by the **RUCIO Scientific Data Management** (by CERN)

02

DB Management System

Distributed high performance and high-availability Database guaranteed by a document oriented DB family **RethinkDB Cluster** with a DB node for each datacenter.

03

Scientific Data Access

Scientific data access to P.I. and generic high-level data consumers is guaranteed by a AMAZON-like cloud-storage service **Minio Storage System** with A&A access by **project LDAP facility**.

04

Services and Gateway

The common link to put together scientific services and high level search-query&Retrieve interfaces are given by a user-access-Portal based on the **DJANGO Content Management System**.

Powered by:

python
django
RethinkDB
MINIO





THANK YOU

ASTRI-horn, Mini-Array and CTA italian team

