



ARCHIVER Project

# Technical Summary

---

## PIC MIXED FILE REMOTE STORAGE

### Problem Definition:

We want to substitute the current in-house tape library storage for both initial and derived datasets. Each instance of the service to be purchased is the 5-year safe-keeping of a yearly dataset and its derived datasets according to the specifications below.

### Lifecycle - Workflow Characteristics:

- A daily dataset is 1 TB/day on average, with x2 fluctuations (from none to 2 TB/day)
- The source makes a daily dataset available all at once. Remote storage of a daily dataset must be completed within 8 hours of notice of its availability (250 GB/hour = 69 MB/s ≈ 0.5 Gbps)
- An initial version of derived datasets will be included in the daily datasets, becoming part of the yearly dataset. (Typical total yearly dataset size 300 TB)
- In addition, anytime during the 4 years following the creation of the data, additional versions of derived datasets may need to be uploaded from [PIC](#) to the Remote storage, becoming part of the yearly dataset.
- Remote storage of an additional version of the derived data would be done from a workflow running at PIC which re-processes a yearly dataset in one month (50TB in 375k files in 30 days, 1.7 TB/day and 12k files/day)
- Source of data would initially be PIC, possibly moving later to a mixture of PIC and [ORM](#), Canary Islands
- Sources are connected to the Spanish academic network (RedIRIS) and through it to Géant at a nominal speed of 10 Gbps
- Recall of data:
  - Entire yearly dataset (including “raw” and “derived”):
    - A remotely stored yearly dataset is recalled as a whole (except for compliance verification, see below). Recall rate is low and usually takes place less than once per year.
    - A prior notice could be given to the supplier prior to the start of a recall (to be negotiated)

- Recall must bring an entire yearly dataset back to disks at PIC in 30 days (10-15 TB/day=417-625 GB/hour=116-174 MB/s≈1.5 Gbps)
  - Derived dataset component of a yearly dataset
    - A given version of the “derived” part of a remotely stored yearly dataset is recalled as a whole (except for compliance verification, see below). Recall rate is low and usually takes place less than once per year.
    - A prior notice could be given to the supplier prior to the start of a recall (to be negotiated)
    - Recall must bring an entire yearly dataset back to disks at PIC in 5 days (10 TB/day=417 GB/hour=116 MB/s≈1 Gbps)

### **Authentication and Management Functions:**

The producers and expert consumers of the data will be either human experts from PIC or automated processes running on PIC’s servers. A very limited number of identities is needed but they should be quite secure. Any reasonable industry-standard mechanism for authentication which is compatible with the Interface Characteristics can be considered.

Management functions: Web and Command Line interfaces should allow a manager to specify an instance and get basic information about the remote storage: Number of files stored, total volume stored, list of files stored, access to compliance evidence. Alarm and monitoring information can be sent by email. A desirable but not mandatory feature is to provide endpoints to PIC’s Icinga alarm and monitoring system.

### **Data and Metadata Characteristics:**

- Yearly dataset volume per instance: 300-400 TB
- Data are immutable once produced (Read Only)
- Data are not compressible (they are already compressed on origin)
- Data are organized as:
  - 250 TB of “raw data” in 2 GB binary files (one year≈125K files, one day≈625 files)
  - 50 TB of “derived data” in smaller binary files (one year≈375k files). Additional versions generate and equal amount of data.
  - Their content is in proprietary format and the supplier doesn’t need to look inside
- Derived data are generated by private workflows whose input is either raw data or a given type of derived data, and whose output is a given type of derived data. There are about 5 types of derived data. Each successive type or derived data is roughly a factor of 10 smaller than the previous one.
- Metadata: Filename: 256 character string, Checksum: 8 character Adler32 checksum
- Metadata are immutable once produced (Read Only)
- Neither data nor metadata contain personal or sensitive information
- Data and metadata must be kept private, only accessible to authorized PIC servers

### **Interface Characteristics:**

- put/get interface with full error recovery/correction using filename as id

- Programmatic interface to initiate a put/get operation and to check the status. Python bindings and REST-like API are desirable but not mandatory.
- Fully documented Command Line Interface to allow manual or bash script put/get
- Full documentation in order to allow us as customers to write an interface to emulate a Tape Library connected to PIC's Enstore or dCache software.
- Emulation as a Virtual Tape Library is desirable but not mandatory.

**Reliability Requirements:**

- A single-bit error anywhere within a file renders it completely useless
- Maximum tolerable loss over the 5-year storage period is one file of "raw" data and three files of "derived data" out of the production of each year (≈150k "raw" files, ≈375k-1.1M "derived" files )

**Compliance and Verification:**

- Scrubbing: Supplier must provide evidence that every stored file has been read and its checksum freshly calculated at least once per year, without any intervention by the buyer. An alarm should be raised if the freshly calculated checksum does not match the original.
- Random 1% buyer's monthly check: Each month, the buyer will give a one week notification simulating a recall of a yearly dataset. The buyer will then ask to recall a randomly chosen 1% subset of the dataset, which must be accomplished within 24 hours.

**Cost Requirements:**

The total cost of an instance of the service must be less than 40k € total for the nominal data volume of 300 TB stored for 5 years, with only one version of derived data stored, plus 7.5k € flat fee per additional version of derived data (up to 50 TB) stored. A desired cost structure would consist of a fixed fee plus a variable fee proportional to the volume of data actually stored.

**Initial Data Management Plan:**

DMP Topic	What needs to be addressed
Data description and collection or re-use of existing data	Raw, calibration and derived datasets will be handled in the ARCHIVER project. These data will be from scientific projects in which IFAE participates and which have explicitly given their consent. Initially, data will be from the MAGIC Telescopes <a href="https://magic.mpp.mpg.de/">https://magic.mpp.mpg.de/</a> . Data will be re-used in the context of multi-instrument gamma-ray astrophysics.

<p>Documentation and data quality</p>	<p>Data documentation will be done by adapting the internal Data Management and Preservation Plans (DMPP) of the scientific projects who own the data. MAGIC has already given permission to use their DMPP plan. Data quality is measured continuously by the scientific instruments and data quality tags will be available as metadata.</p>
<p>Storage and backup during the research process</p>	<p>Today, PIC is using its dCache/Enstore based service to archive 'cold data' in a HSM like (disk + automated tape) service. That layer is going to be replaced/extended with ARCHIVER services using an additional copy of the data. (cleanup at the end of the project requires no further data movement).</p>
<p>Legal and ethical requirements, codes of conduct</p>	<p>Citation should be given when data is used by permission.</p>
<p>Data sharing and long-term preservation</p>	<p>Easy to use data sharing, both amongst access-controlled groups as well as Open Data, is a major objective of placing the data in the ARCHIVER-developed platforms. Long-term preservation will be tested by simulating in an accelerated manner the Data Management and Preservation cycle.</p>
<p>Data management responsibilities and resources</p>	<p>High-level data management will remain the responsibility of IFAE, while low-level data management and resources will be provided by the vendors.</p>