THE SARDINIA RADIO TELESCOPE DATA MANAGEMENT FROM OBSERVING TO ARCHIVING

Picture by Sergio Poppi

A.Fara - Technologist INAF Osservatorio Astronomico di Cagliari - Responsible for SRT ICT

WS "Archives and data management systems in the big data era" - Bologna 26-28/02/25

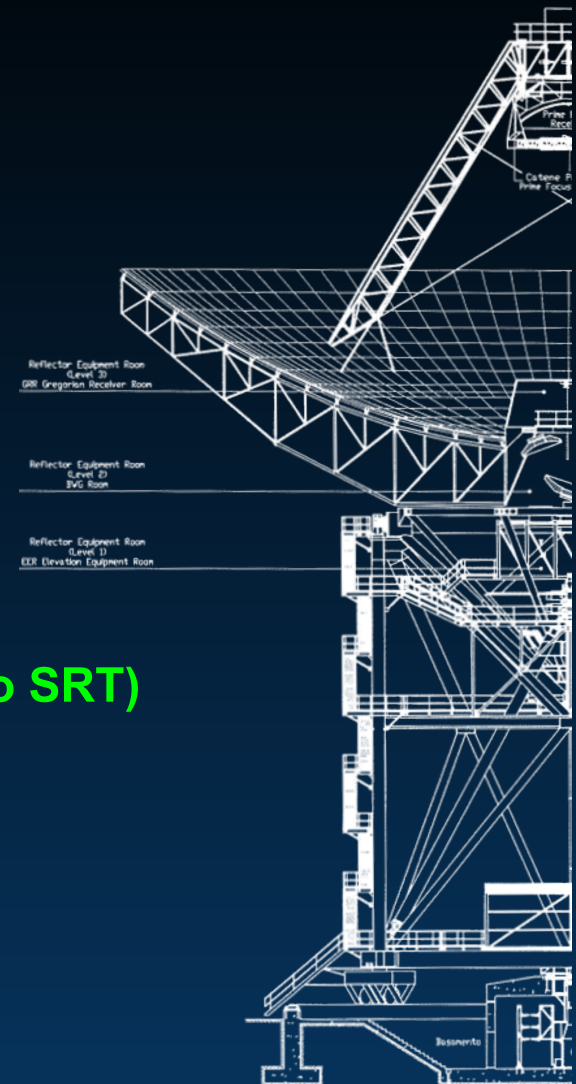
SARDINIA RADIO TELESCOPE

SINCE 2012

- 64 meters - active surface
- single and multi feed receivers / multiple backends
  - total power
  - xarcos (until 2022)
  - pdfb3 (until 2024)
  - leap (EPTA project)
    - roach1 + hpc cluster
  - sardara
    - roach2 : hpc nodes (1:1)

2019-2023 PON PIR01_00010 (Potenziamento tecnologico SRT)

- 0.3 GHz - 116 GHz
  - new receivers (1 - 2 - 7 - 16 - 19 feeds)
  - new backend recorders
    - skarab : hpc nodes (x:y)
  - new hpc and storage



HOW MANY DATA ? A BIT OF HISTORY

SCIENTIFIC COMMISSIONING - EARLY SCIENCE (2012-2016)

- total power - xarcos = 122 GB
- pdfb3 = 2 TB
- sardara + roach2 = 7.5 TB

TOT 11 TB

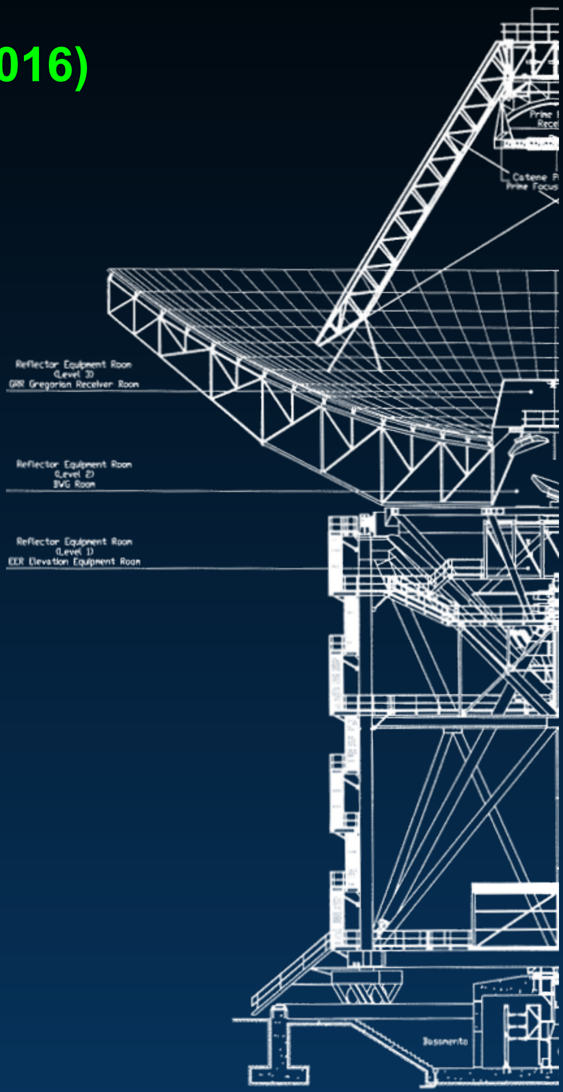
SSA MAINTENANCE (2017) - NO DATA

RECOMMISSIONING AND CALL (2017 - 2020)

- total power - xarcos = 122 GB
- pdfb3 = 9.4 TB
- sardara + roach2 = 19 TB

TOT 28 TB

PON UPGRADE (2020-2023) - NO DATA



RESTART AFTER PON UPGRADE

PRE-PON RECEIVERS WITH NEW BACKENDS

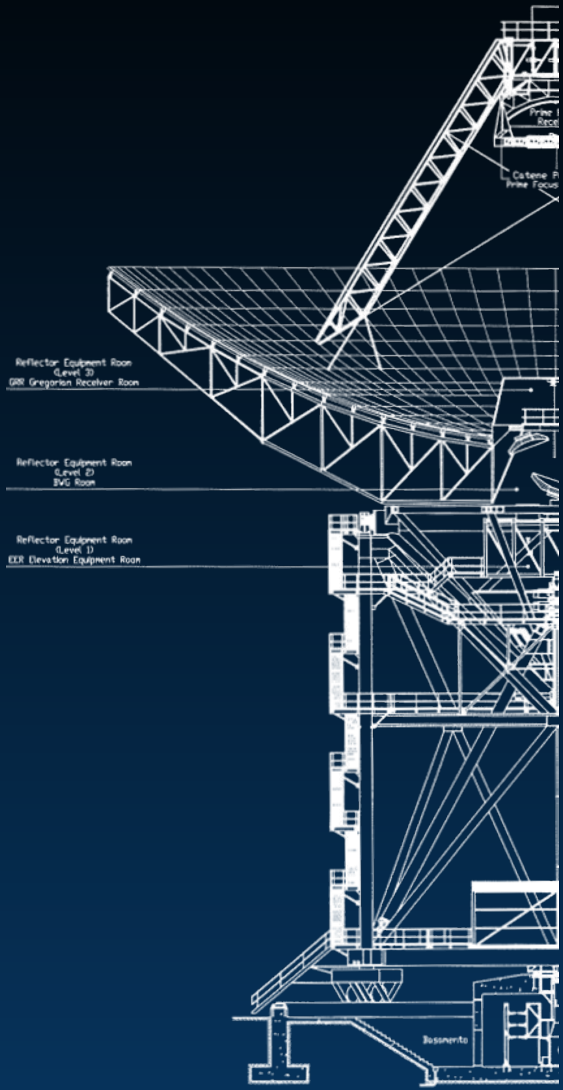
2023 (since oct) = 11 GB (only maintenance)

2024 (12 months)

- **total power - (xarcos) = 210 GB**
- **pdfb3 = 2.2 TB**
- **sardara roach2 = 3.2 TB**
- **bcalc skarab = 6.3 TB**

2025 (2 months C K receivers)

- **total power - xarcos = 0**
- **pdfb3 = 355 MB (under decommisioning)**
- **sardara roach2 = 460 GB -> 6 TB/year**
- **bcalc skarab = 557 GB -> 6 TB/year**



HOW MANY DATA IN THE FUTURE ?

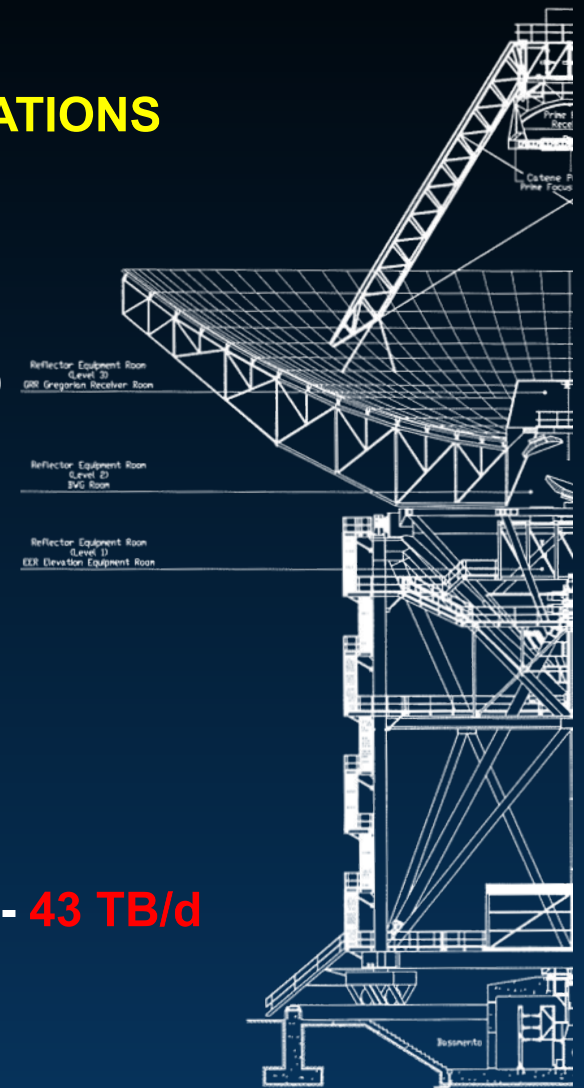
12 TB/y = UNDERESTIMATED IN FULL TIME OPERATIONS AND WITH ALL RECEIVERS

Expected maximum datarates

- **C BAND (1 feed) K BAND (7 feed) LP BAND (2 feed)**
  - **sardara roach2**
  - **125 GB/h - 2.4 TB/d**

NEW MULTIFEED RECEIVERS

- **W BAND (16 feed) - Q BAND (19 feed)**
  - **bcalc skarab**
  - **100 - 500 MB/s = 360 GB/h - 1.8 TB/h = 8.6 TB/d - 43 TB/d**
  - **MISTRAL and TRI BAND datarate = not critical**



FROM RECEIVERS TO DISCOS

DIGITALIZATION AND RECORD RAW DATA

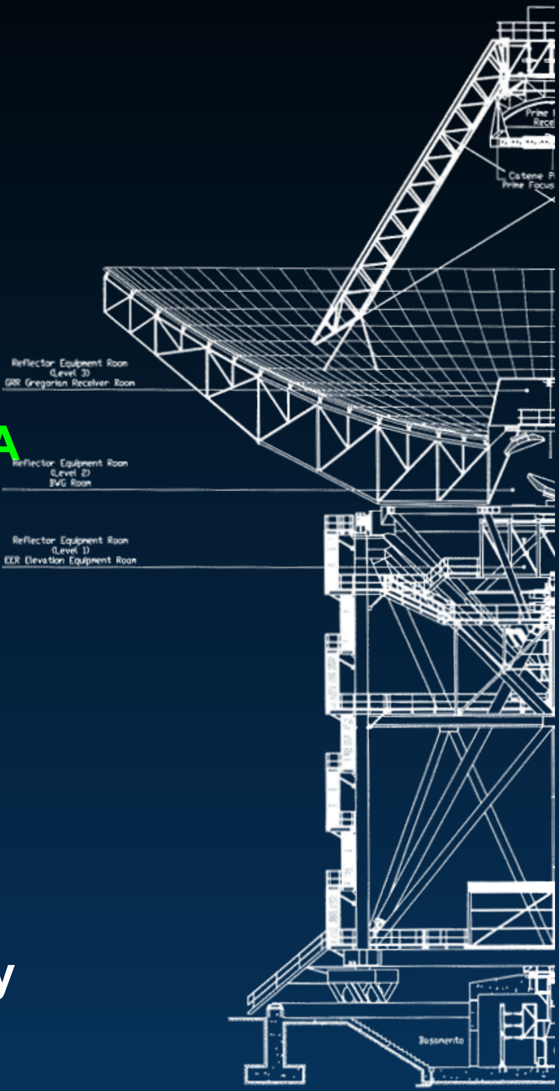
- from different receivers (single and multifeed)
- with multiple backends
  - on local ssd backend’s disks

ADD A FITS COMPLIANT HEADER TO BUILD FINAL DATA

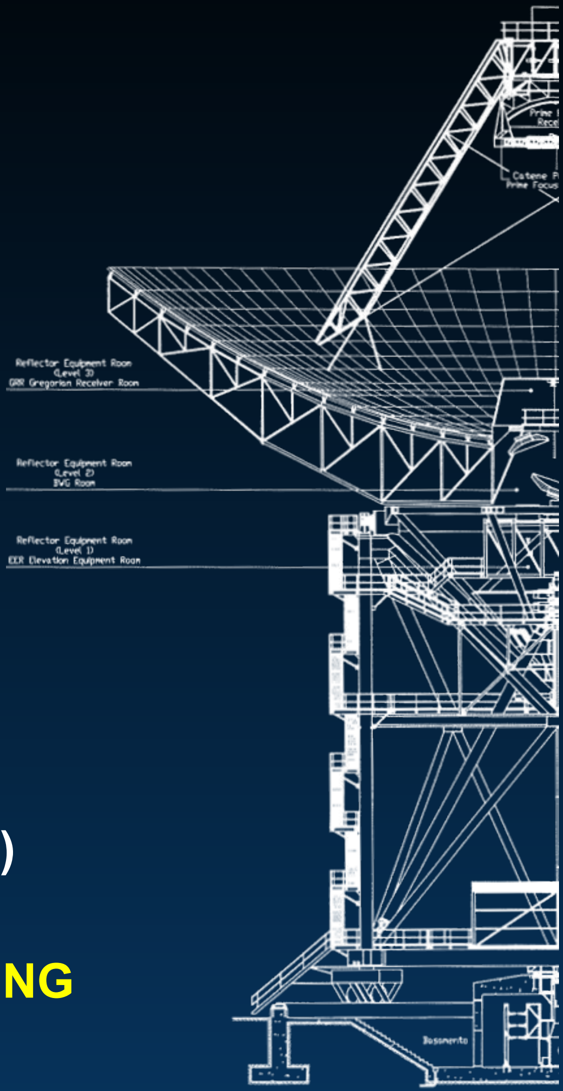
- merge by backend’s custom software
  - managed on backend’s cluster nodes
  - temporary processed on local cluster’s disks
  - final data on discos-ssd storage buffer

MAXIMUM WRITE PERFORMANCE IS REQUIRED: SSD

- pros = performance
- cons = small and expensive / death without recovery



LOCAL REAL TIME ACCESS



BUFFER DATA DISK

- 36 TB ssd (Rocky 8)
- zfs with minimum redundancy
- ready for lustre if needed
  - compliant OS kernel multiple Lustre version
  - CentOS 7 discos - Rocky8 - 9 HPC
- shared over a private 10/100 Gbps ethernet

ACCESS BY LOCAL PREPROCESSING TOOLS

- quicklook (browser)
- discos2class converter (read only nfs)
- copy to local scratch HDD (rsync - python watchdog)

WE NEED TO BALANCE BETWEEN DATA MERGING AND CONCURRENT USER'S ACCESS

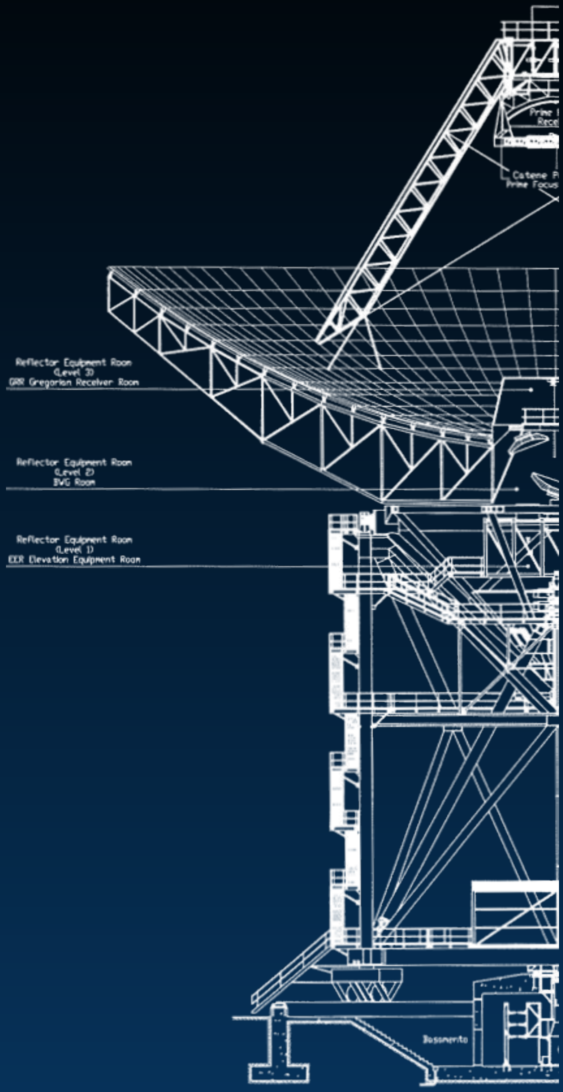
PROCESSING AND SAVING DATA

COPY (MIRROR) TO STORAGE “SCRATCH LIKE”

- 2 x 360 TB HDD (zfs + Lustre)
- interactive operations required
- ready for hpc local processing
- private project id ownership access
    - NIS authentication and mount

COPY TO SRT STORAGE

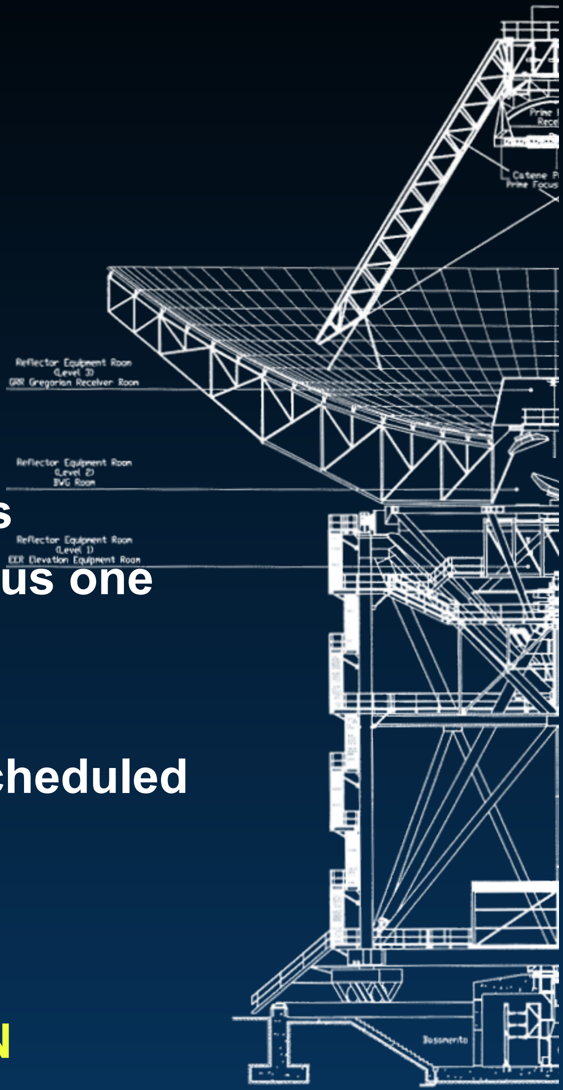
- only if the buffer-ssd is not working
- async copy from buffer-ssd to storage
    - 70 TB hdd zfs partition
- async copy from partition to partition on sftp storage
    - remote access by project id
        - populating homes
        - change permissions and ownership



CRITICAL POINTS

LOCAL MIRROR FROM SSD TO HDD

- **watchdog + pynotify + rsync (quasi real time)**
- **different write speed between ssd and hdd storage**
  - **increasing latency in mirror process**
- **rsync “freezes” a filelist**
  - **copy is not aligned with new incoming files**
- **multiple rsync process overload the server resources**
  - **a mirroring can start before the end of the previous one**



SAVE TO STORAGE

- **not enough time to extract data in case of full time scheduled telescope (*Dynamic scheduling* is in progress)**
- **manual crontab schedule is required**

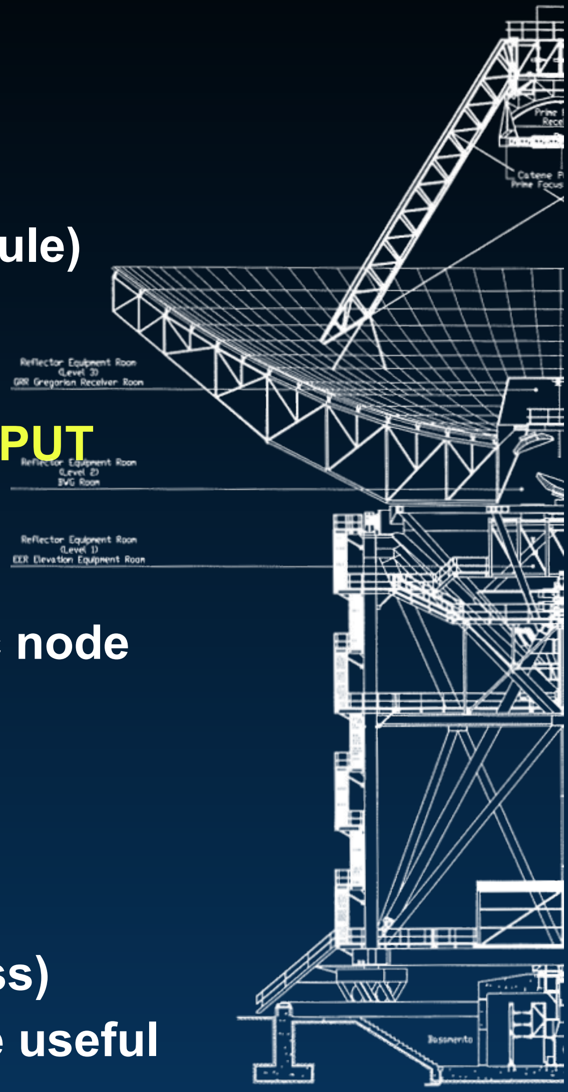
MULTIPLE SSD-BUFFERS CAN BE A SOLUTION

LONG TERM ARCHIVE

ARCHIVE TO IA2 COMPLIANCE

- SRT backends produce fits-like “custom” format
- human error can be possible (i.e wrong PID on schedule)
- header check is required

TEAM WORKING GROUP TO MADE BACKEND OUTPUT ARCHIVE COMPLIANT



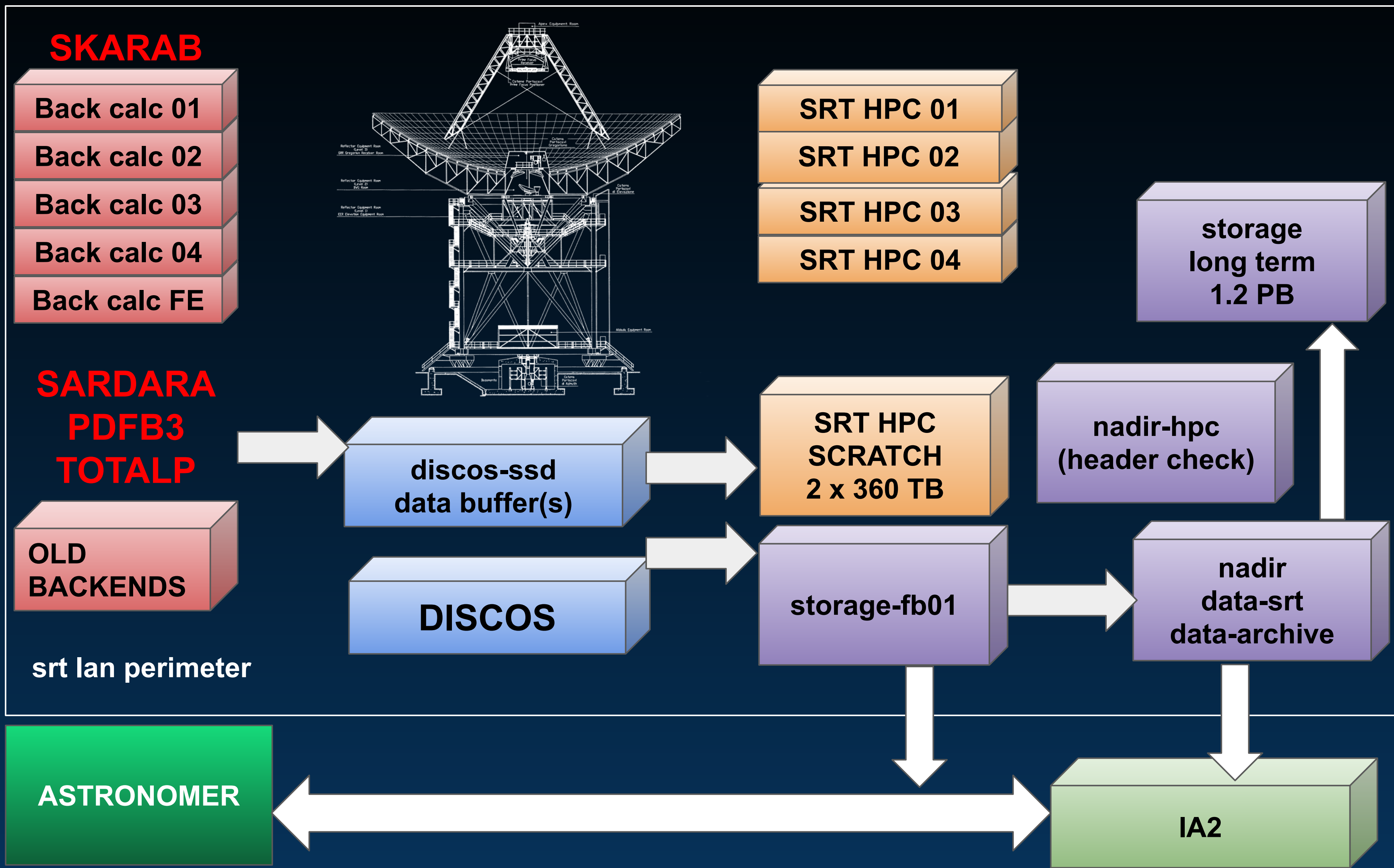
NADIR IA2 SERVER

- from storage to “raw header” folder shared with hpc node
- running checks and correct (also manual)
- ingested to IA2 and finally copied to

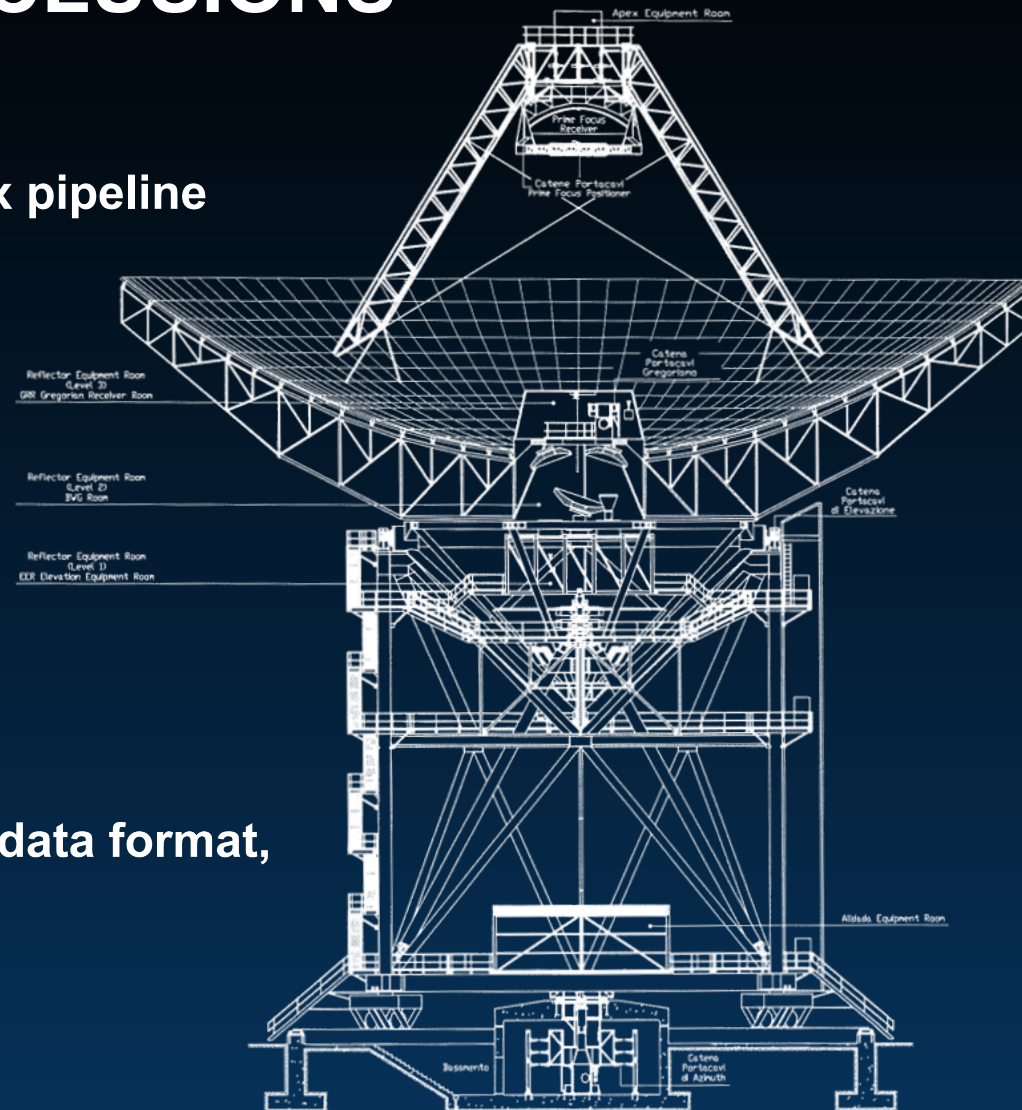
SRT STORAGE LONGTERM

- 1.2 PB hdd Lustre (not available for direct user access)
- tape library for offline very long term storage coud be useful

FINAL RECAP DETAILS



CONCLUSIONS



SRT data management is a complex pipeline that requires to balance between

- **performances**
- **real time interactive access**
- **data retrieve**
- **safe storage**
- **long term archive**

The automation of all the pipeline requires that backend and archive developers work in synergy on the data format, to reduce the manual checks and the implemetation of a reliable ICT storage infrastructure.