

Update on middle ware and DAQ stuff

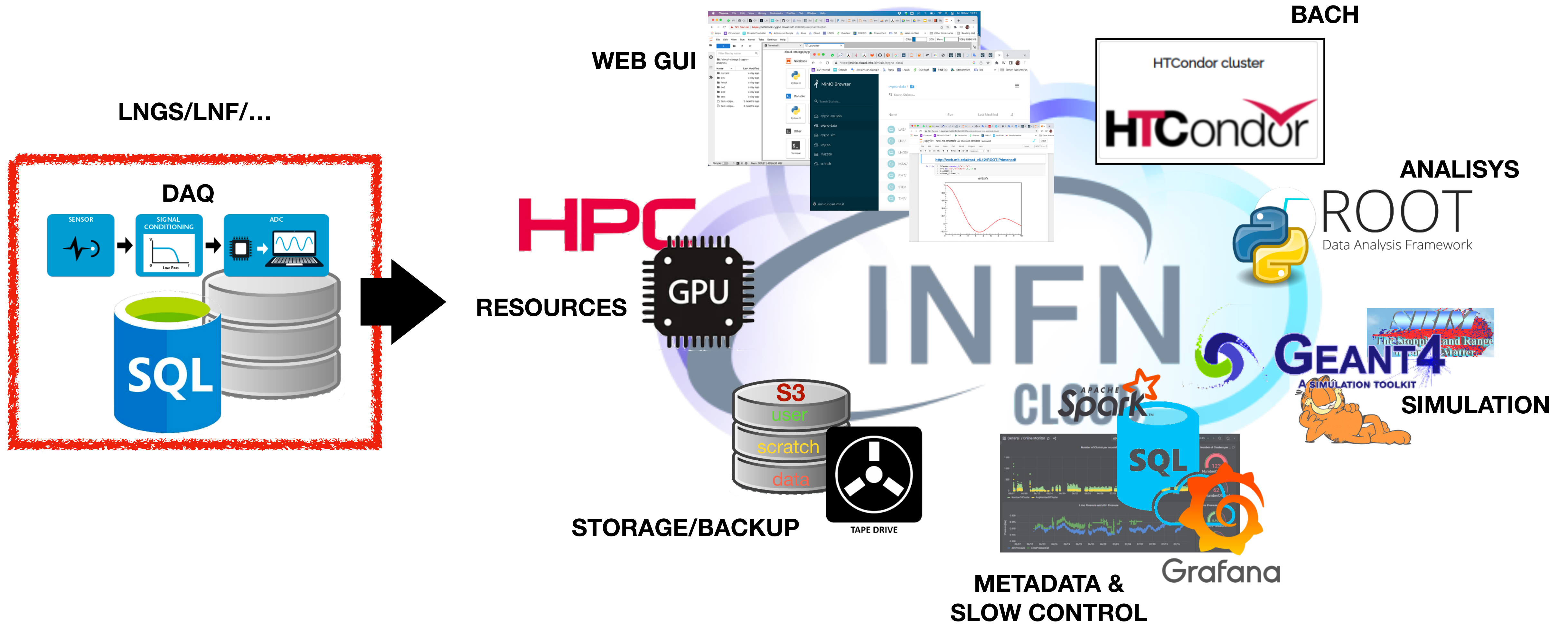
G. Mazzitelli

the middle ware CYGNO project

data management and online data validation and qualification

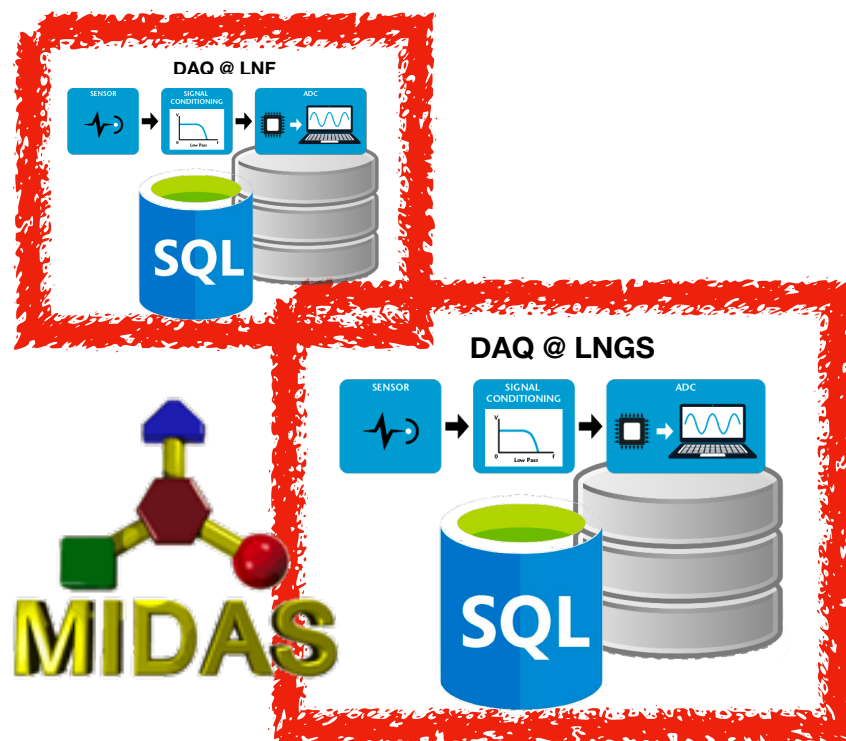
- full experiment data management;
- full experiment raw and high level metadata production and management;
- slow/fast remote experiment monitor without access to LAN DAQ (shift workers from all over the world);
- data online reconstruction and pre-analysis;
- data validation and qualification;
- information, alarms and warning dispatcher via discord experiment channel

CYGNO... computing model



logical units, “composed” services

test and development setup at LNF

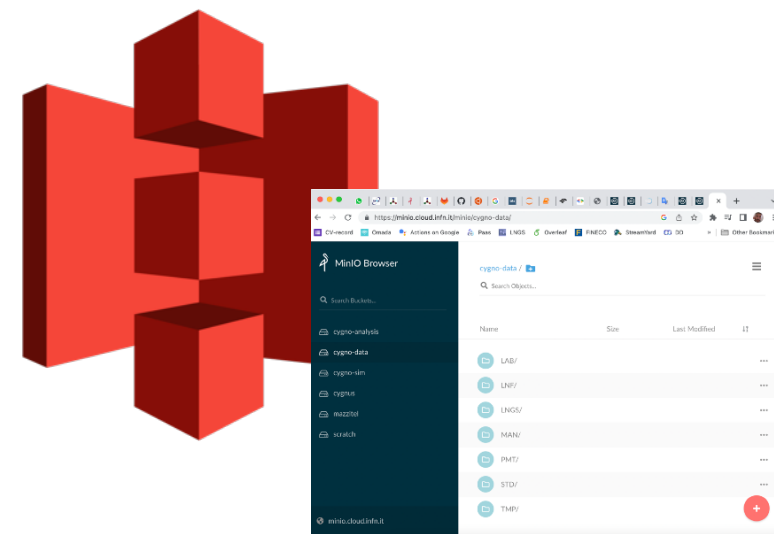


production setup at LNGS

Mariadb replica for metadata
sql.cygno.cloud.infn.it



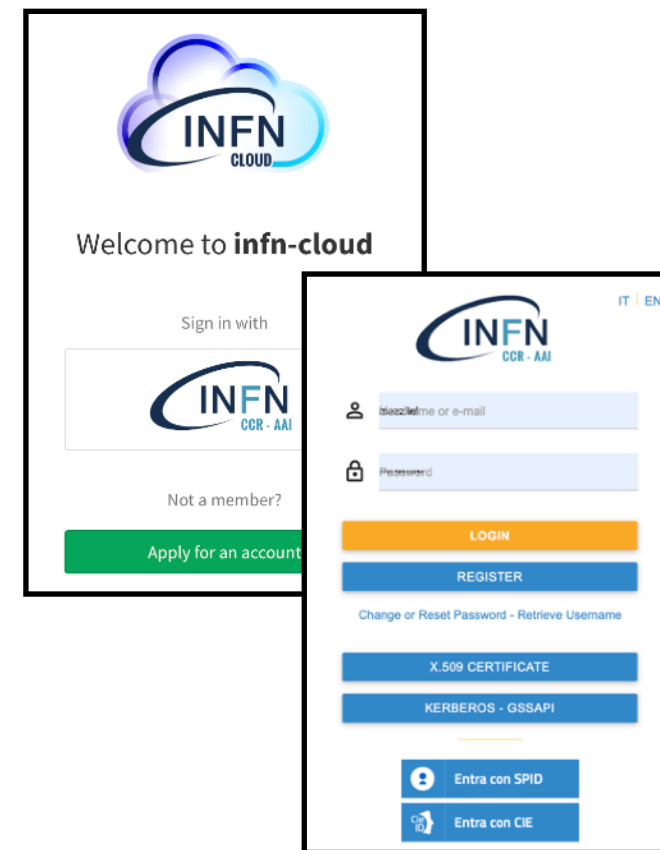
S3 storage
minio.cloud.infn.it



messaging
kafka.cygno.cloud.infn.it



Identity and Access Management
iam.cloud.infn.it

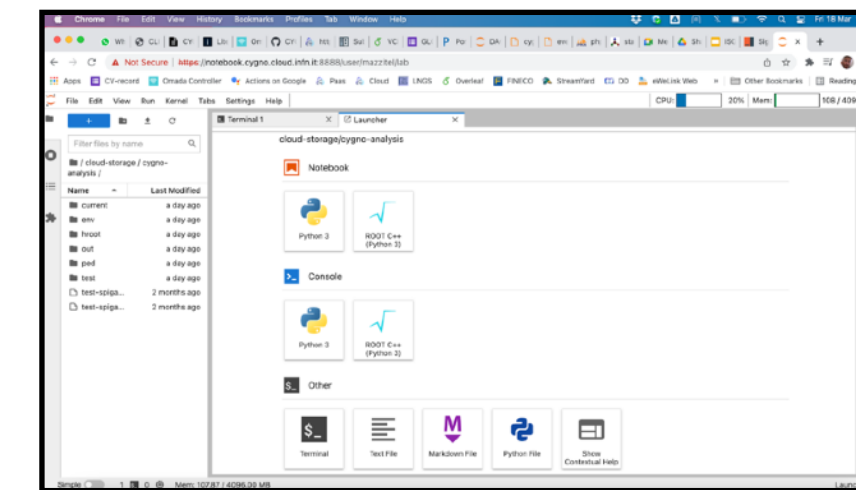


data and metadata monitor
grafana.cygno.cloud.infn.it

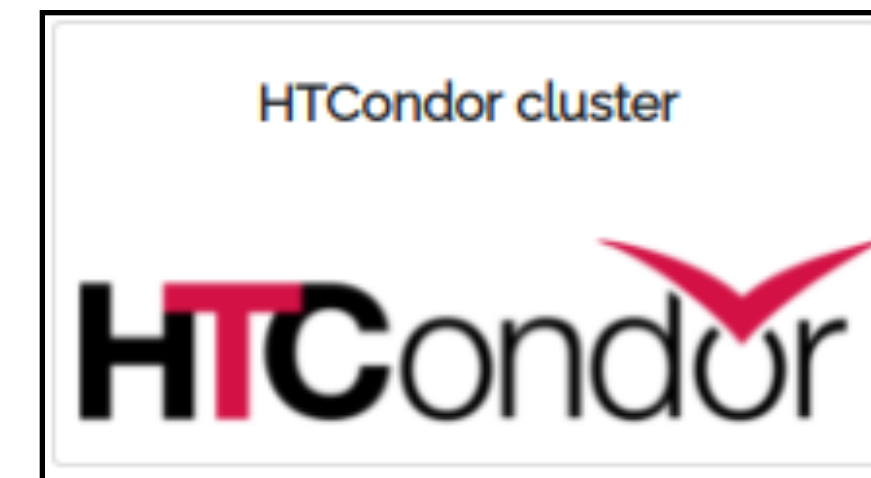


Grafana

analysis and simulation web interfaces
notebook01.cygno.cloud.infn.it
notebook02.cygno.cloud.infn.it



batch queues
condor01.cygno.cloud.infn.it
condor02.cygno.cloud.infn.it



backup
tape.cygno.cloud.infn.it



TAPE DRIVE

pre analysis and data quality
sentinel.cygno.cloud.infn.it



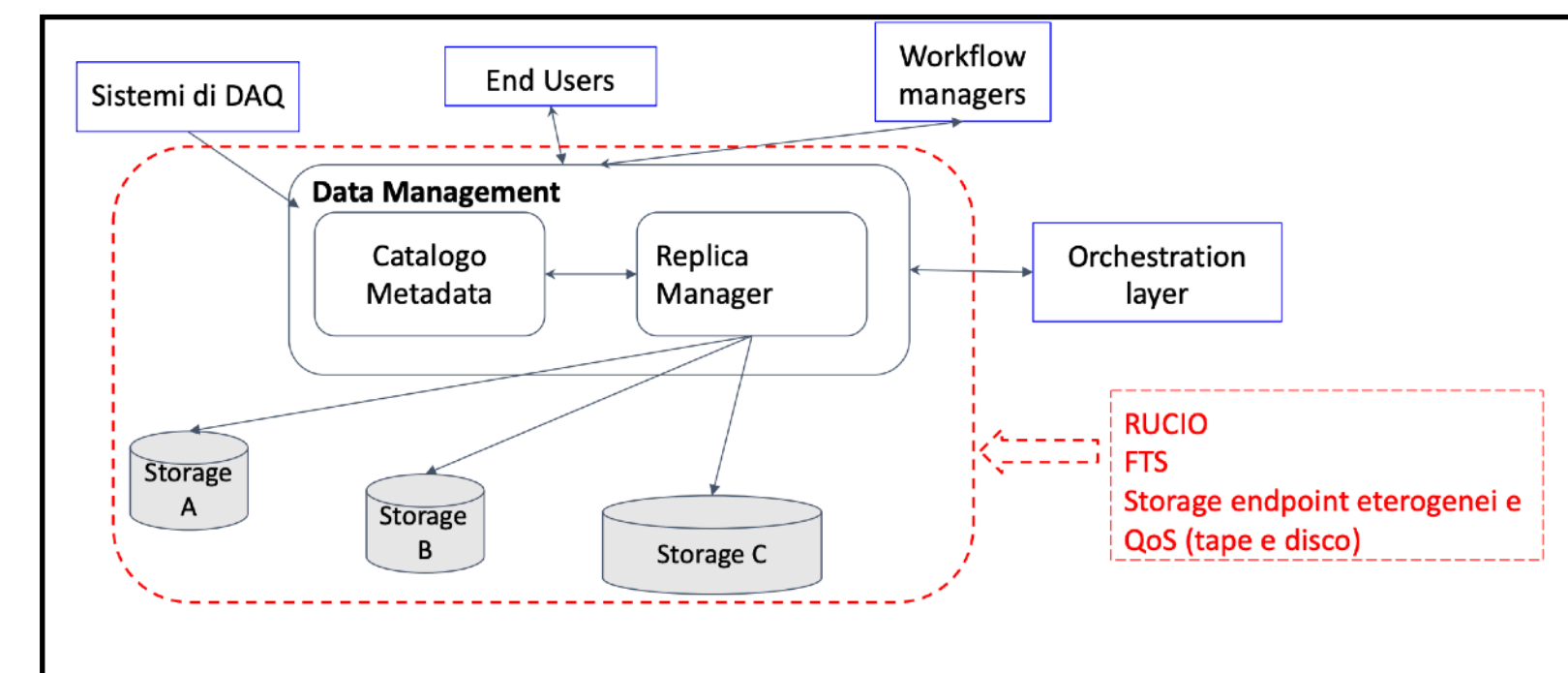
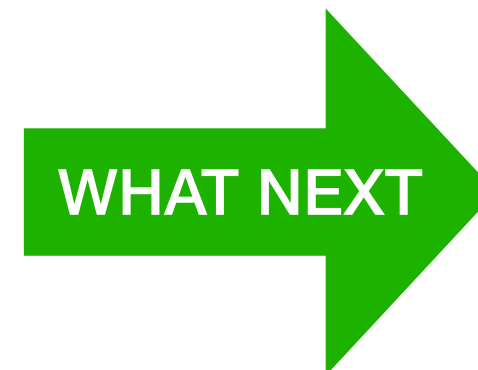
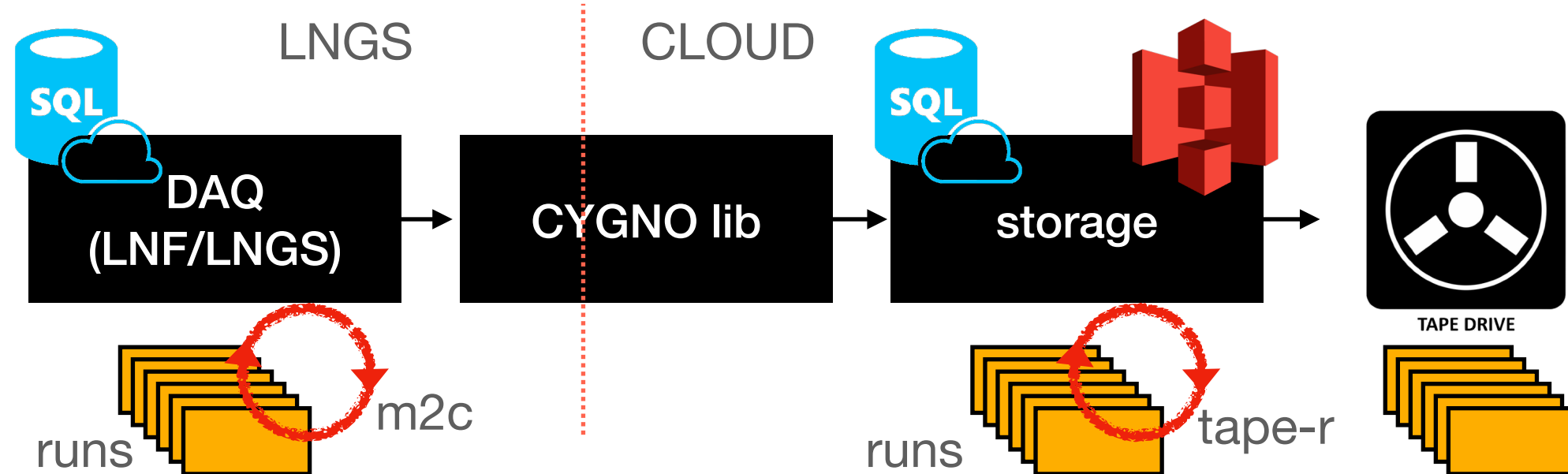
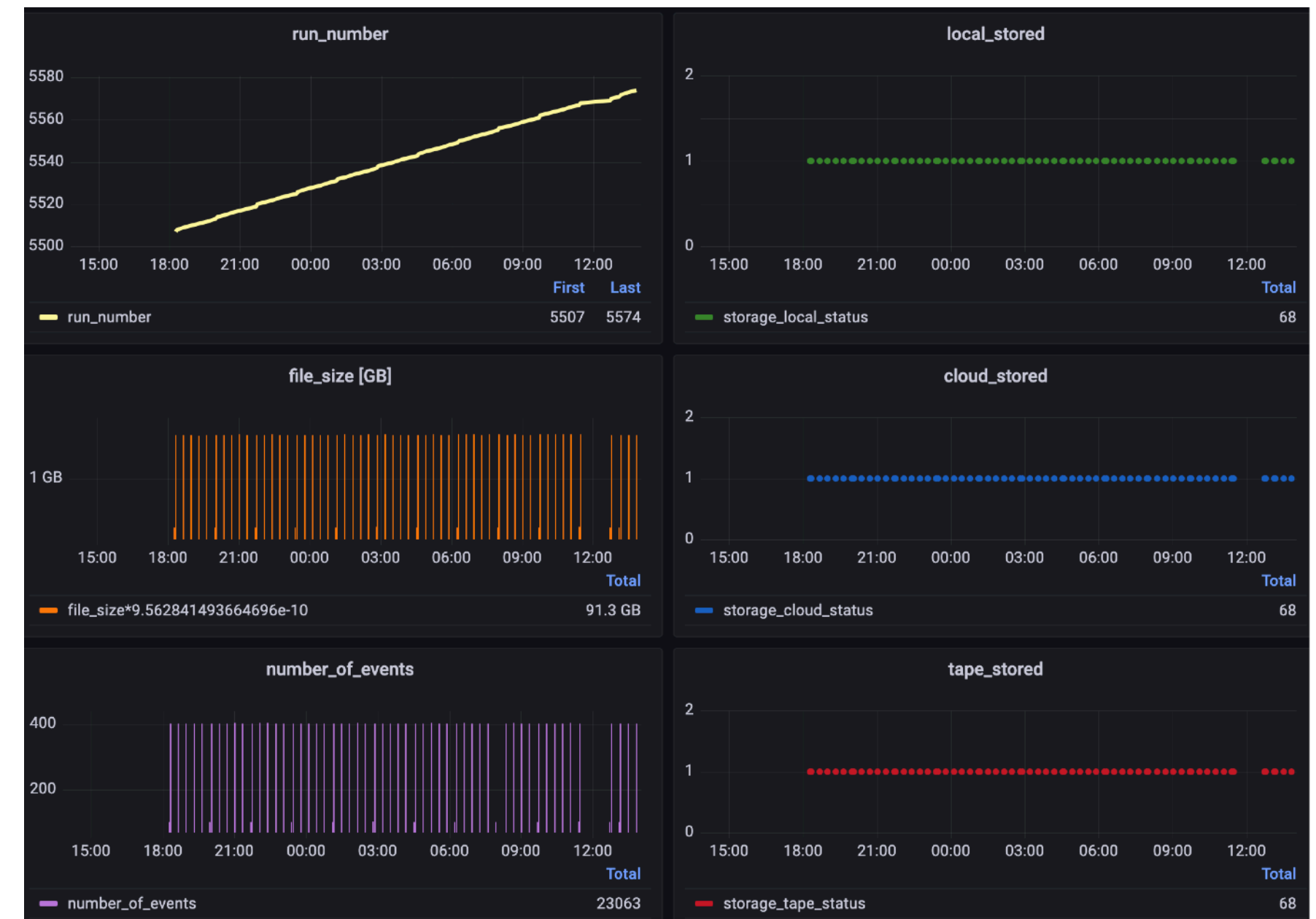
data management

the “tape-r”

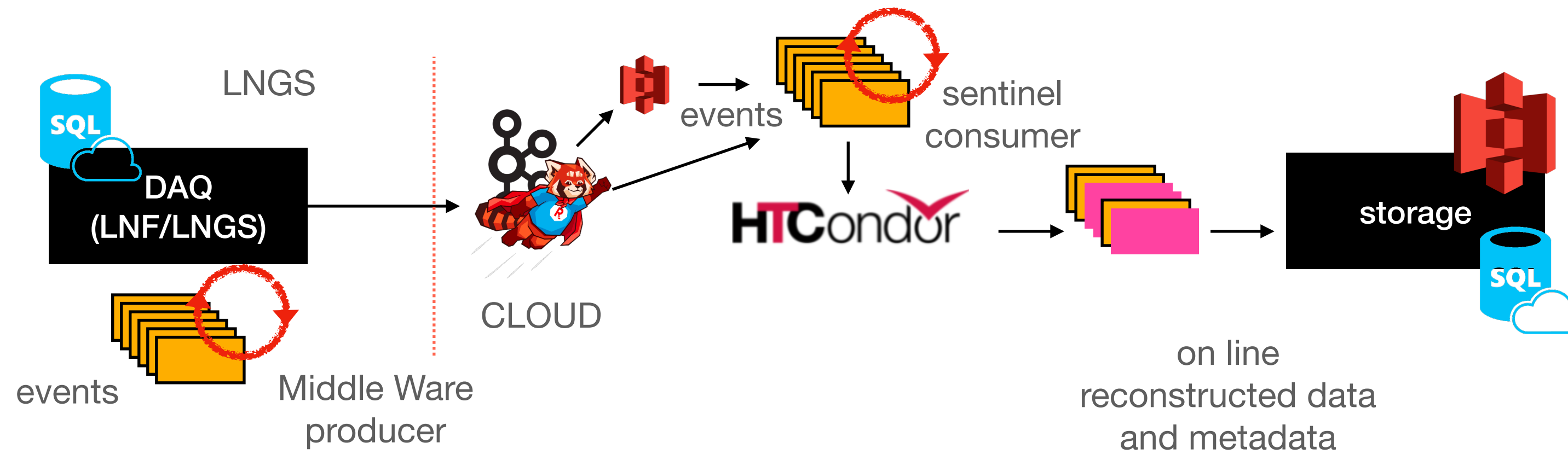
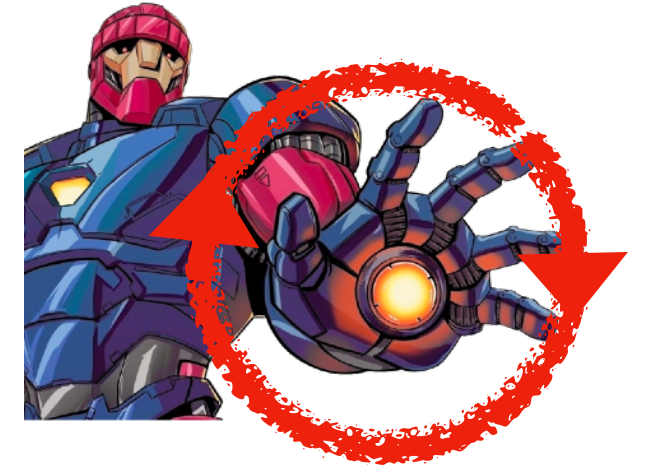


- data by means **m2c process**, bunched in runs, are copied on **S3 object storage**, as well as **metadata**, locally stored and replicated on cloud MariaDB;
- a few second after the run is close is available for **full reconstruction** on the cloud HTcondor queue and can be **download** with various tools (web, rest api, POSIX, ecc);
- the “**tape-r**” process replicate **data** on tape and update **metadata** of the run status;
- tape up to now is accessible via **GFAL, RUCIO and FTS** is going to be implemented as cloud services for more complete and generalised data management system

data replica dashboard



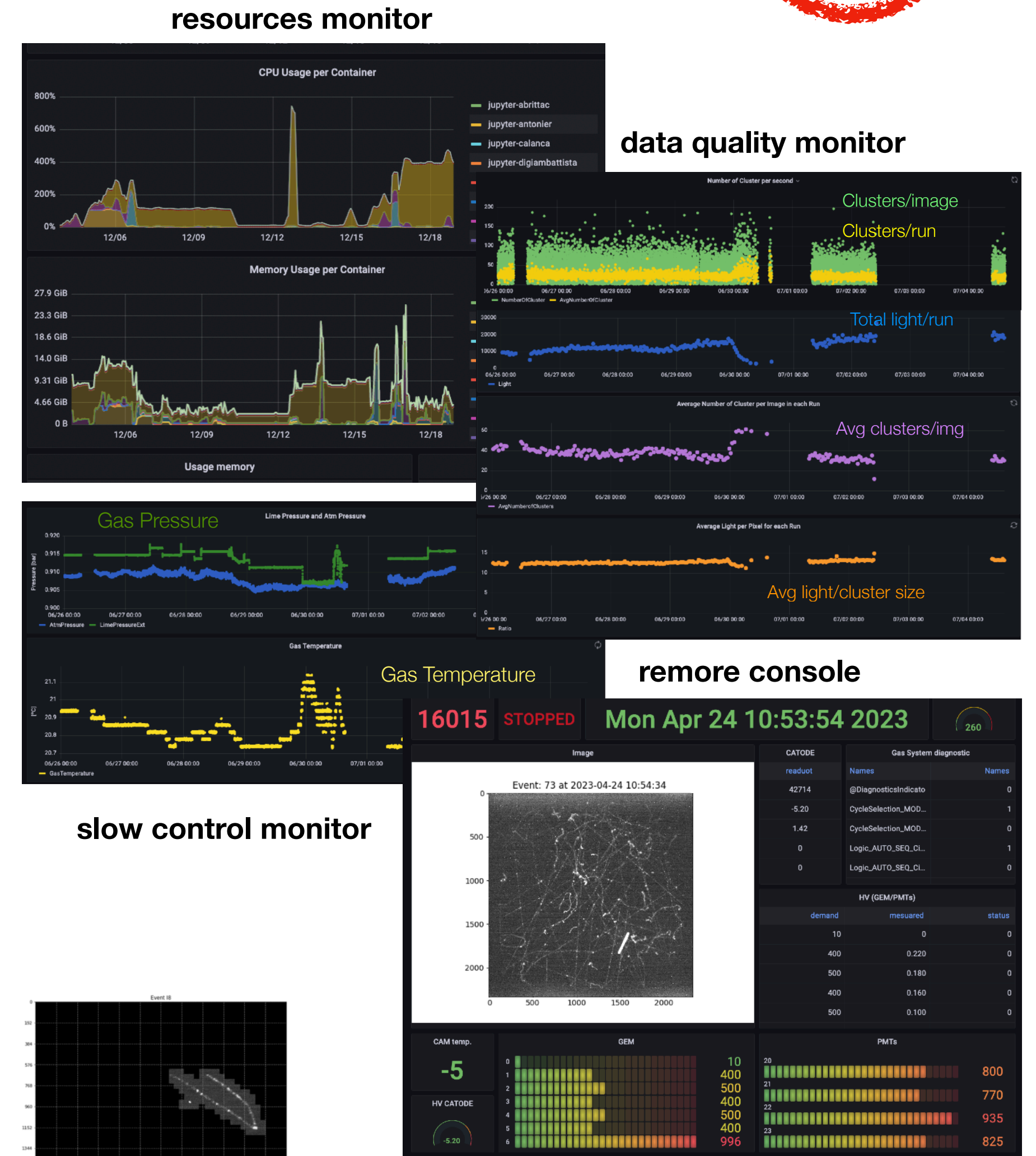
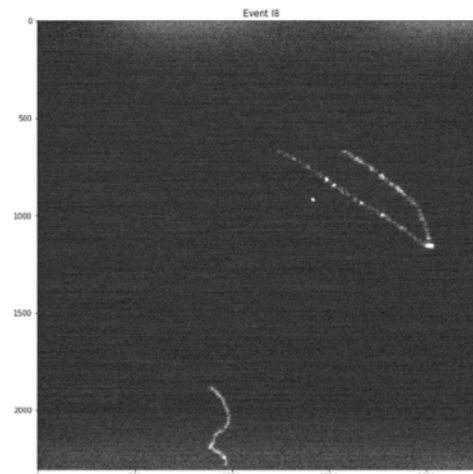
online data reconstruction middle ware and the “sentinel”



- parallel to run data management, single events are sent to cloud by means of **kafka producer**
- the **sentinel** process consumes data **parallelising the events reconstruction** on the HTCondor queues
- data and metadata are stored and **presented** for on line monitoring

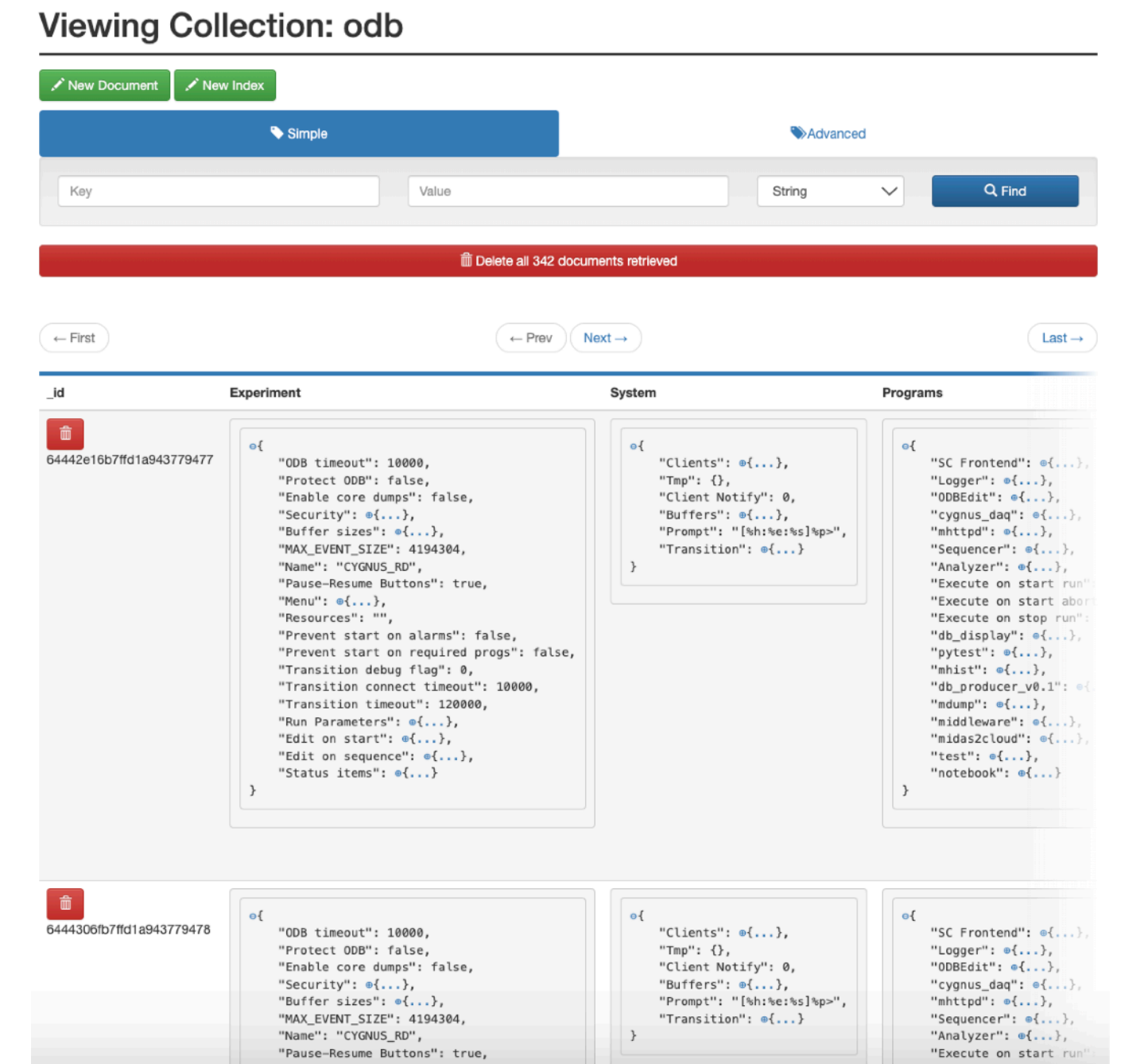


optimise/scale architecture to completely be able to provide online reconstruction
implement **data compression** (triggerless ML/GPU algorithms are under study)



on going Thu

- start Mariadb replica
- update delete local file
- update m2c (remove replica, faster list)
- test delete cloud file
- setup **remote console** (what we need?)
- **middle ware** slow channels history (what we need? uptime? MongoDB?)
- slow channel (*.idf *.idx *.hst 300MB/day), slow channel toll?



ongoing before restart

- S3 EVENT buffer cleaner (service task)
- start fast record sentinel
- implement full reco queuing (it's the same of point 2, do we really need?)
- LNGS network issues