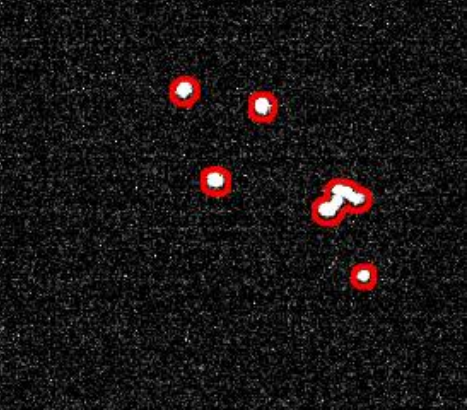
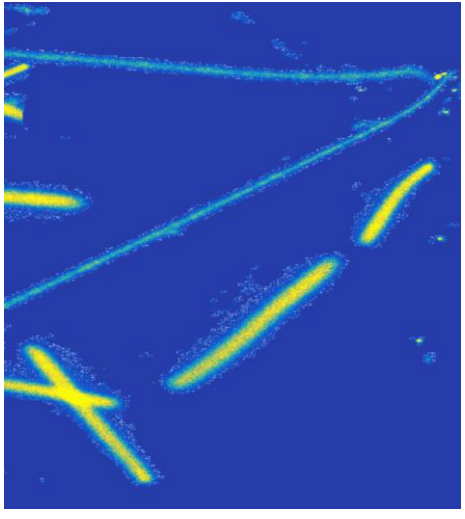


Status of Computing and Reconstruction code



Giorgio Dho

Istituto Nazionale di Fisica Nucleare (INFN-LNF), Frascati (RM), Italy

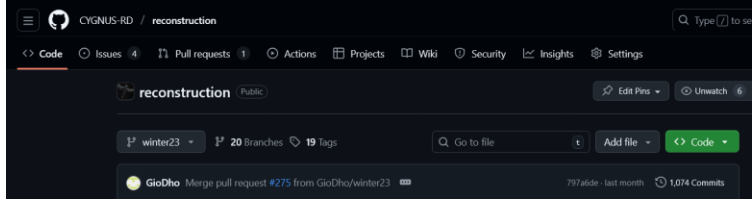
CYGNO Collaboration meeting

Reconstruction

Reconstruction Flow

- The current more updated branch is still **Winter23**.

- Newest release is available with stable operation (Saladin)

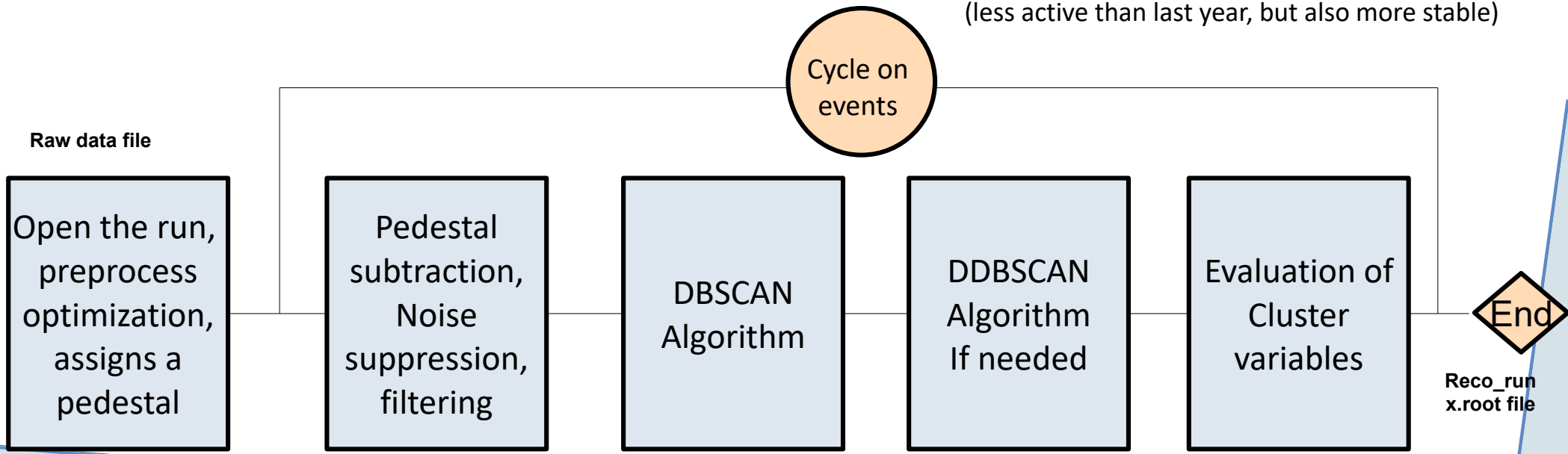


Comparing the same screenshot of 1 year ago:

7 PR merged

34 new commits

(less active than last year, but also more stable)



Stability and fixes

- Discord channel and PR method proved to be a very effective tool

Hopefully I was responsive a quick enough with who had reco problems

No indirect crashes introduced in LIME reconstruction by PRs

- Most of the fixes were about:
  - Adapt to new download path and CVMFS environment
  - Slimmer reconstruction folder and code (cleaning and vignetting map download not required anymore)
  - Addition of oxygen variables

Missing: Barrel Correction

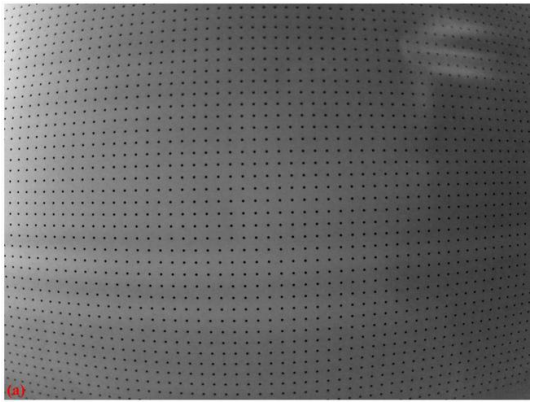
Method: https://discorpy.readthedocs.io/en/latest/tutorials/methods.html#correcting-perspective-effect

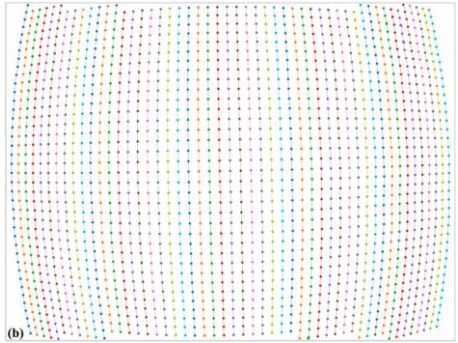
Usage: https://discorpy.readthedocs.io/en/latest/usage/demo_05.html

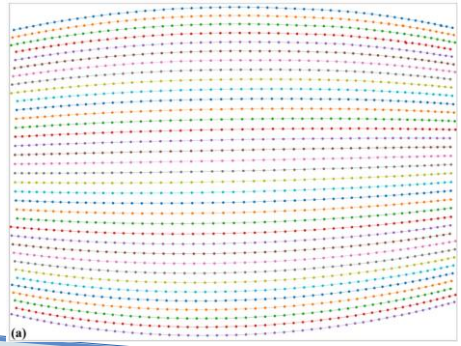
- Discorpy library (used by Migdal group and Australian CYGNUS) seems to be the way to go

- Study started but unfinished yet. **Should be addressed**

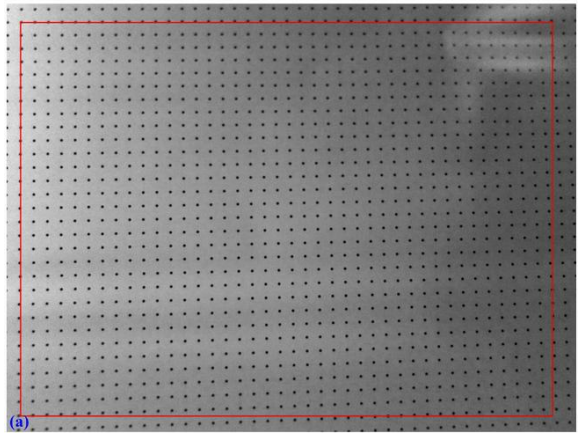
Image







Corrected



Missing: After Reco?

- With the reco_run files some cluster variables are calculated, but it is not a full analysis
- For CYGNO-04, a different type of files will be generated by DAQ and the cluster analysis will be separated
- A new repository with postprocessing code is needed
- The idea is to provide the user a template to read the files and a group of functions the user can add to their custom code

Future

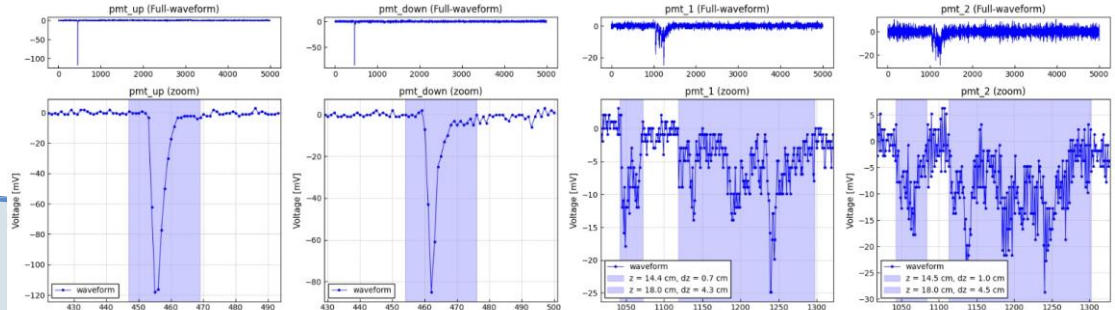
- Disuniformity map could also be added in future once understood

- Focus will shift towards coping with new DAQ data format:

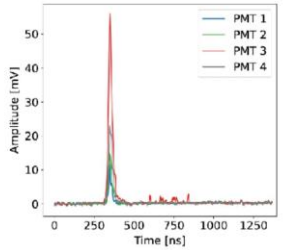
    - Virtual memory usage for trigger level analysis and event builder (started)

    - Restructuring of variables for analysis

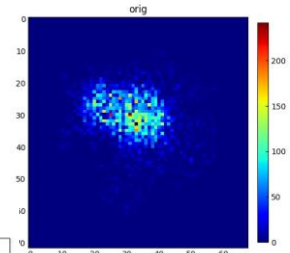
    - Improvements for PMT algorithm (started by Scamporlino)



Event builder and Analysis Pipeline

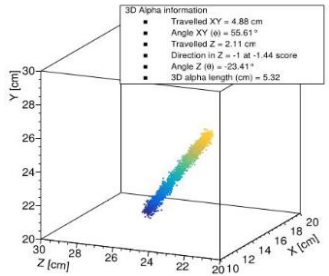


- Frontend trigger provides series of signal pixels and waveforms and timing [Igor and Stefano's presentation]



- DAQ PC (hopefully) will provide a built event merging timing information and a fast association (BAT or ML)

  An event will have a redpix block and a probable sets of waveforms and thier time. Saved in files on cloud.



- Reco-Layer0: Full association with 3D reconstruction and correct clustering

  Output stored as 3D redpix, timestamp and env

- Reco-Layer1: Higher level analysis. Stored output: directionality, Head-tail, energy, z, PTOH

Computing

logical units, “composed” services

docker Compose

test and development setup at LNF

production setup at LNGS

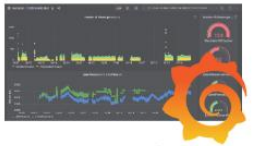
Mariadb replica for metadata
sql.cygno.cloud.infn.it

S3 storage
minio.cloud.infn.it

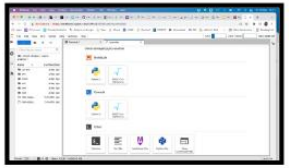
messaging
kafka.cygno.cloud.infn.it

Identity and Access Management
iam.cloud.infn.it

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



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

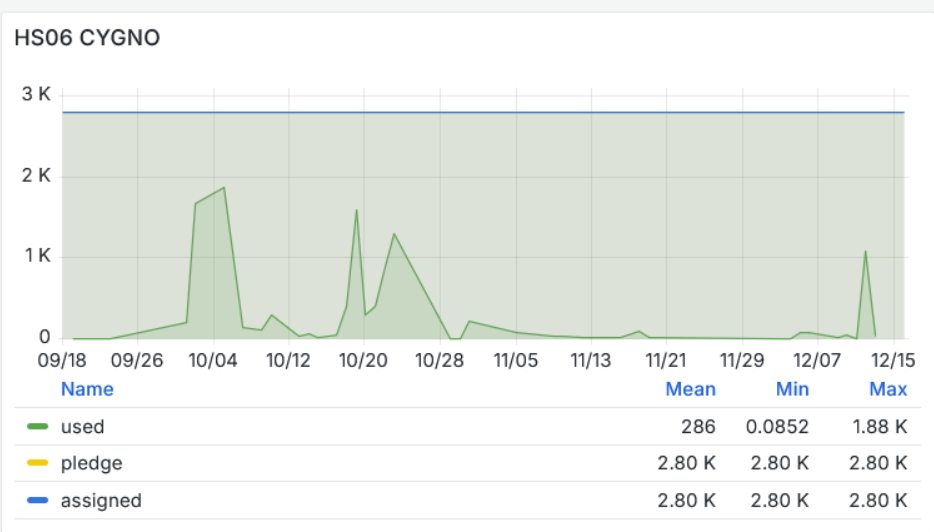
pre analysis and data quality
sentinel.cygno.cloud.infn.it

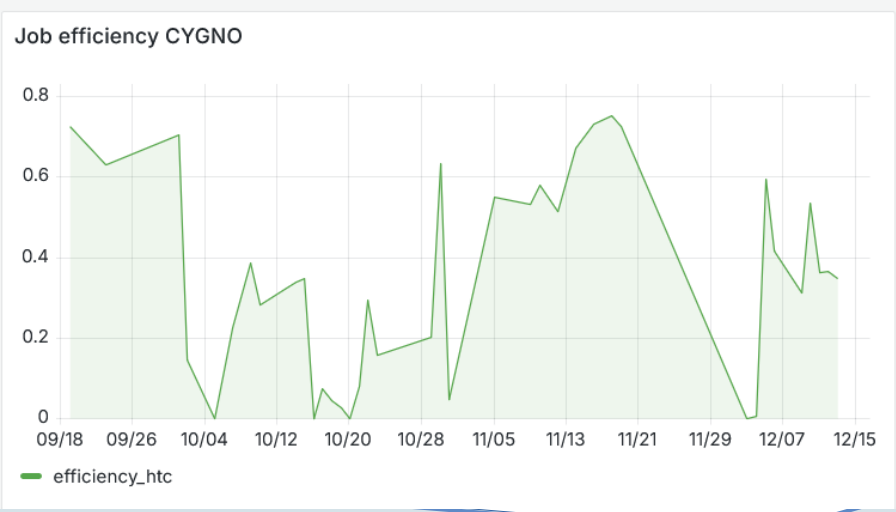
Pledge Resourses

- Status of requested computing resources 2026

HS06 (CPU)

- 5000
  - 3200 To be used for autoreco and notebooks (more res will given to notebooks)
  - 2800 Allocated to Tier1 CNAF queues
- 8000 Through the PSGE CSN2 resources (not dedicated only to CYGNO)





Pledge Resourses

- Status of requested computing resources 2026

HS06 (CPU)

- 5000
  - 3200 To be used for autoreco and notebooks (more res will given to notebooks)
  - 2800 Allocated to Tier1 CNAF queues
- 8000 Through the PSGE CSN2 resources (not dedicated only to CYGNO)

Storage

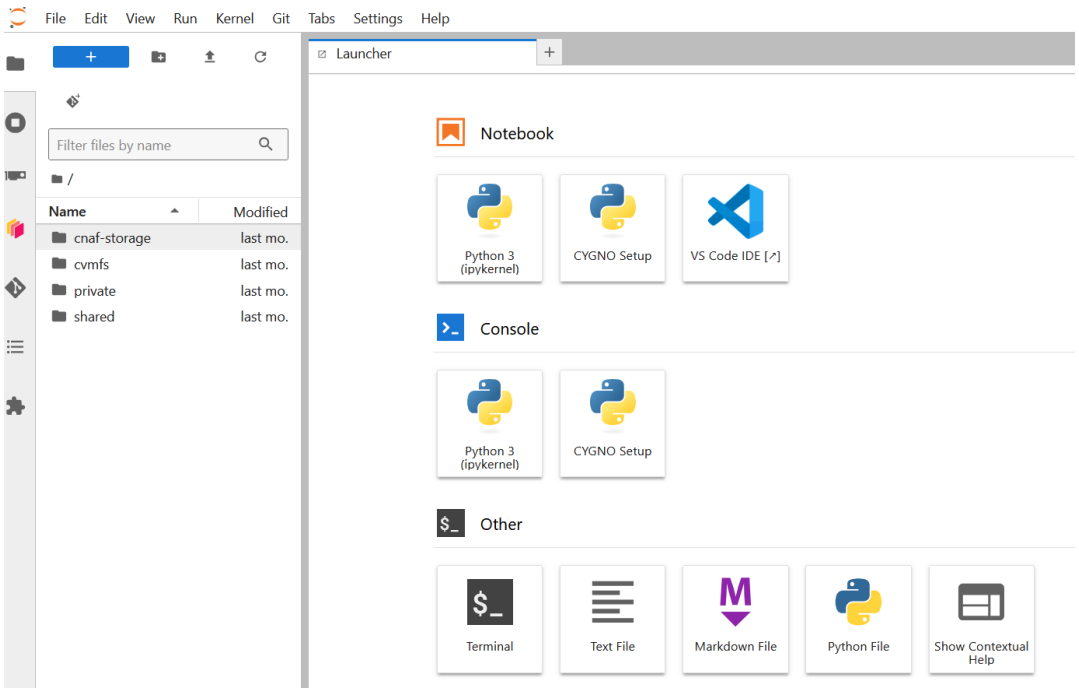
- 190 TB (30 more than last year)
  - 10 TB for analysis (2.4 occupied)
  - 20 TB for simulation (3.6 occupied)
  - 160 TB for data (108 occupied now)

Tape

- 300 TB (100 more than last year)
  - 180 occupied now

Notebook V2

- Notebook structure was upgraded



- VS code plug-in added
- Python notebook with *cygno_setup* at start
- Cnaf disks mounted
- CVMFS software integrated
- Personal bucket
- Ubuntu22 image as base
- Distributed notebook storage
- Storage webapp with rclone
- X-term will come

CNAF Storage Migration

- Migration to new s3 storage class (called CNAF storage by us) has been completed
- Now all storage is of the same type (no more MINIO or Bari)
- All codes (reconstruction, cygnolibs, digitisation) and notebooks are now correctly interfaced to new storage

CVMFS

- CERN-based repository with standard software in various frozen versions (sft.cern.it:)
- Python ROOT and Geant4 for example are now regularly used by all cygners on notebooks and queues
- Scalable, customizable, low-maintenance
- A CYGNO dedicated CVMFS is already available sft-cygno.infn.it (all required python and C++ packages, submit software, soon maybe also reconstruction and digitization!)
- Can be mounted on any PC and prevent any new user from installing new software: it is already there!

CNAF Tier1 Queues

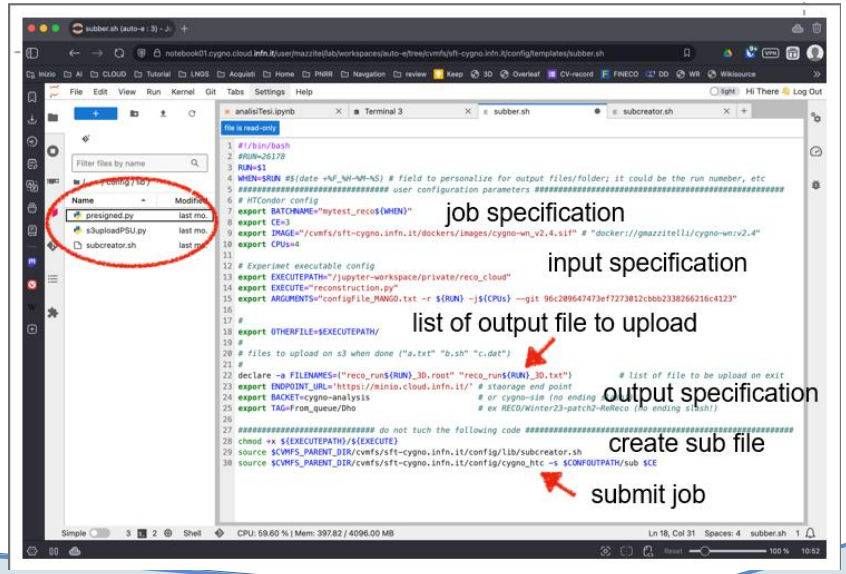
https://agenda.infn.it/event/48044/ on how to use them!

- 6 different CE (queues equivalent)

```
(base) dho@notebook:/private# cygno_htc -h
Usage:
  -t/-c --tier1/--cloud, configure/switch between htc@tier1 and htc@cloud
  -s --submit, submit a job:  -s <subfilename> <ceid> [only for tier1 ceid=1-7 default ce02]
  -f --tranfer, tranfer files: -f <jobid> <ceid> [only for tier1 ceid=1-7 default ce02]
  -r --remove, remove jobs: -r <jobid> <ceid> [only for tier1 ceid=1-7 default ce02]
  -q --monitor, monitor jobs: -q <ceid> [only for tier1 ceid=1-7 default ce02]
  -m --myjobs, monitor my jobs (dho): -m <ceid> [only for tier1 ceid=1-7 default ce02]
  -j --jobs, monitor all jobs: -j <ceid> [only for tier1 ceid=1-7 default ce02]
  -h --help, show this help
(base) dho@notebook:/private#
```

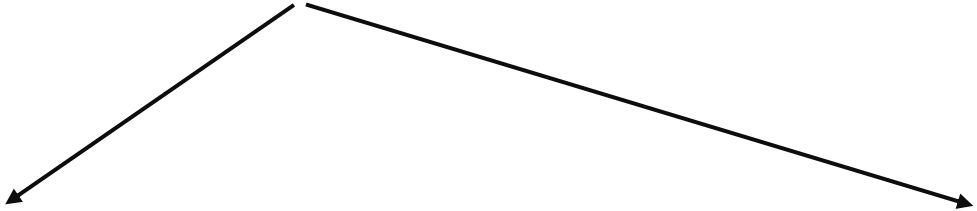
```
(base) dho@notebook:/private# cygno_htc -q 3
htc@tier1


-- Schedd: ce03-htc.cr.cnaf.infn.it : <131.154.192.51:9619?... @ 07/21/25 13:05:18
OWNER    BATCH_NAME          SUBMITTED   DONE   RUN    IDLE  TOTAL JOB_IDS
cygno003 mytest_reco13487    7/12 10:44      _      _      _      1 2677747.0
cygno003 mytest_reco13489    7/12 10:44      _      _      _      1 2677748.0
cygno003 mytest_reco13490    7/12 10:45      _      _      _      1 2677749.0
cygno003 mytest_reco13492    7/12 10:45      _      _      _      1 2677757.0
cygno003 mytest_reco13493    7/12 10:46      _      _      _      1 2677758.0
cygno003 mytest_reco13703    7/16 11:35      _      _      _      1 2698969.0
cygno003 mytest_reco13706    7/16 11:35      _      _      _      1 2698970.0
cygno003 mytest_reco13707    7/16 11:35      _      _      _      1 2698971.0
cygno003 mytest_reco13709    7/16 11:35      _      _      _      1 2698972.0
cygno003 mytest_reco13710    7/16 11:35      _      _      _      1 2698973.0
cygno003 mytest_reco13704    7/16 11:36      _      _      _      1 2698974.0
cygno003 mytest_reco13701    7/16 11:36      _      _      _      1 2698977.0

Total for query: 12 jobs; 12 completed, 0 removed, 0 idle, 0 running, 0 held, 0 suspended
Total for cygno003: 12 jobs; 12 completed, 0 removed, 0 idle, 0 running, 0 held, 0 suspended
Total for all users: 7927 jobs; 2246 completed, 0 removed, 573 idle, 5107 running, 1 held, 0 suspended
```

- Loads CVMFS exactly as notebooks
- Submit preparation sort-of userfriendly with tutorial!
- Job output sent automatically to CNAF storage
- Considering the usage we have, there is almost no limit to usage (only Melba found the limit of the space!)



Update Autoreco

- Autoreco software was updated last week to cope with new queue images and submit files
- New version is ready to run on new infrastructure for CYGNO-04
- It is about to be tested with the rereco of Run5



Rereco will have the PMT waveforms saved and oxygen inside the reco variables

It will start as soon as a bug in queues highlighted by our usage will be fixed

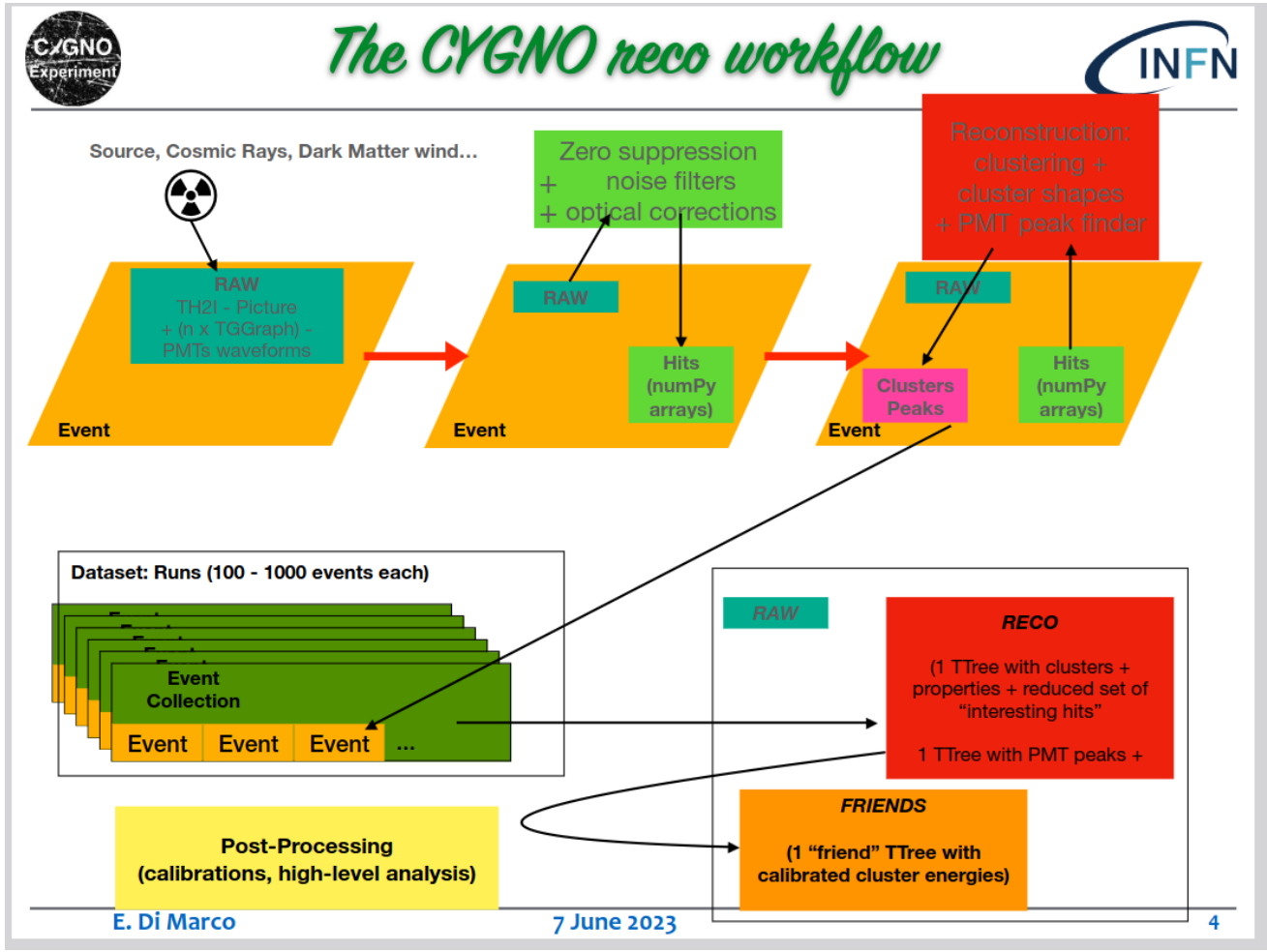
Utilities

- Rclone mount expertise has increased.
  - Notebooks mount disks via Rclone
  - Anyone can mount CNAF disks on their PCs (https://github.com/CYGNUS-RD/middleware/blob/master/help/Rclone.md)
  - Soon new script to mount disks on read-only mode as generic folder
- IAM CYGNO
  - Authentication now indepentent from CNAF, but relying on experiment directly (faster)
- RUCIO
  - Manager regulates the copy of files from local DAQ to storage and tape
  - Internal database which helps smoothing all copying and data management operations

Conclusions

- Reconstruction code has seen a relatively quiet year
- Reconstruction is still lacking some improvemets (mainly barreling)
- Adaptation for CYGNO-04 has started with design of Event builder and autoreco adaptation
- Large amount of work was performed this year in the improvement of the computing infrastructure
- Notebook, CNAF tier queues, CVMFS among the most impactful highlights

Extra 1


The CYGNO reco workflow
INFN
Source, Cosmic Rays, Dark Matter wind...
RAW
TH2I - Picture
+ (n x TGGraph) -
PMTs waveforms
Event
Zero suppression
+ noise filters
+ optical corrections
RAW
Hits
(numPy
arrays)
Event
Reconstruction:
clustering +
cluster shapes
+ PMT peak finder
RAW
Clusters
Peaks
Hits
(numPy
arrays)
Event
Dataset: Runs (100 - 1000 events each)
Event
Collection
Event
Event
Event
...
RAW
RECO
(1 TTree with clusters +
properties + reduced set of
"interesting hits"
1 TTree with PMT peaks +
FRIENDS
(1 "friend" TTree with
calibrated cluster energies)
Post-Processing
(calibrations, high-level analysis)
E. Di Marco
7 June 2023
4