

Dark energy in the GRID

The PAU survey project

Presenter: **Nadia Tonello** (PIC)

Co-authors: C. Neissner (PIC)

S. Serrano (IEEC)

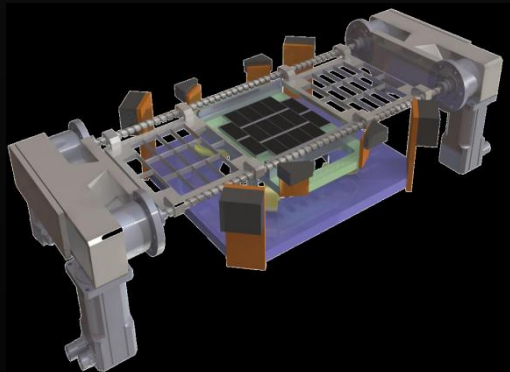
Outline

- Description of the PAU survey project
 - Features
 - Requirements
- Activity at PIC related to the project
- Data flow
 - Pipeline
 - Jobs management
- Future projects: database



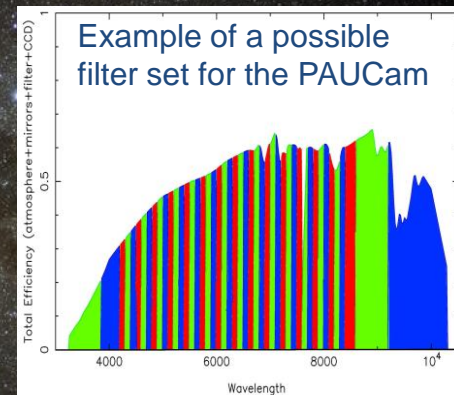
Physics of the Accelerating Universe

Aim put constraints on cosmological parameters for the study of the acceleration of the universe: DARK ENERGY



Method Large field of view telescope camera with low-resolution spectroscope $\sim 200 \text{ deg}^2$ of astronomical images in fits format, taken with ~ 40 narrow band filters $\sim 100 \text{ \AA}$ wide

Data Reduction process Catalog of a large sample of galaxies in a huge volume (spectral energy distribution)



Physics of the Accelerating Universe

Requirements:

Storage: simulated data – results in tapes

Security: two independent Roles into the same vo:

PAUS : DES images in fits format, used as a testing tool for the data pipeline

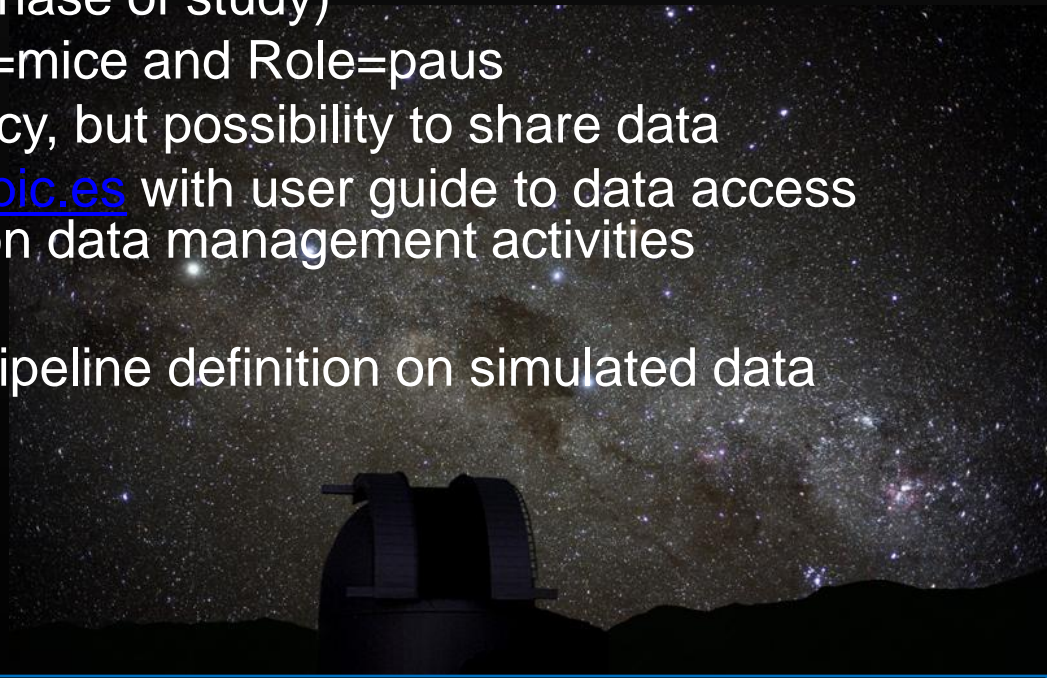
MICE : dark matter particles evolution in huge volume (Marenostrum supercomputing center, Barcelona)

Scalability of pipeline: data reduction of simulated data (FITS files) must be scalable to analyze future real data

In design phase

Activity at PIC

- **Data Management**
 - D0T1 permanent raw data storage (simulation files)
 - D1TX rapid access to results of the reduction process
 - \nfs buffer disk for temporal access during data process (direct reading from dCache in phase of study)
 - VO: vo.paus.pic.es Role=mice and Role=paus
Different data access policy, but possibility to share data
 - Web page <https://cosmo.pic.es> with user guide to data access and general information on data management activities
- **Software for data reduction** pipeline definition on simulated data
- **Database** (future project)



Reduction process

- Main script in python
- Configuration files with all the parameters selected to run the entire analysis.

Part 1 : **INITIALIZATION and PRE-PROCESS** (wms-job)

Create list of files stored in tapes to work with

Pre-stage (dcap)

Copy files to /nfs/

Master-BIAS (filter independent)

Part 2 : **CALIBRATION** (DAG jobs: one for filter)

Instrumental calibration and masking

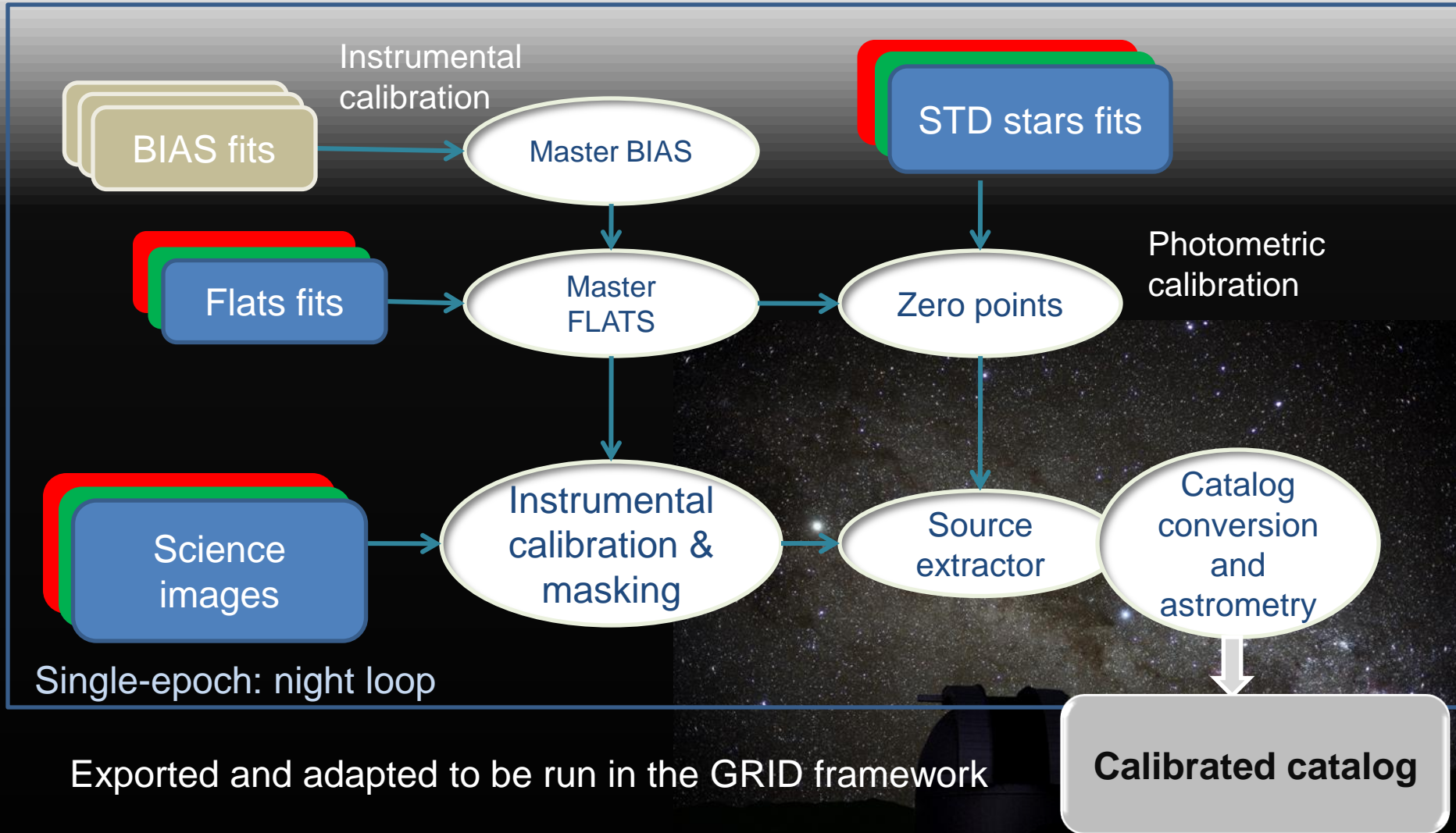
Photometric calibration

REDUCTION (parallel jobs)

Reduction software applied to scientific images

The Pipeline

Coadd: Loop over nights of observation



Exported and adapted to be run in the GRID framework

Calibrated catalog

Part 0: Pre-processing



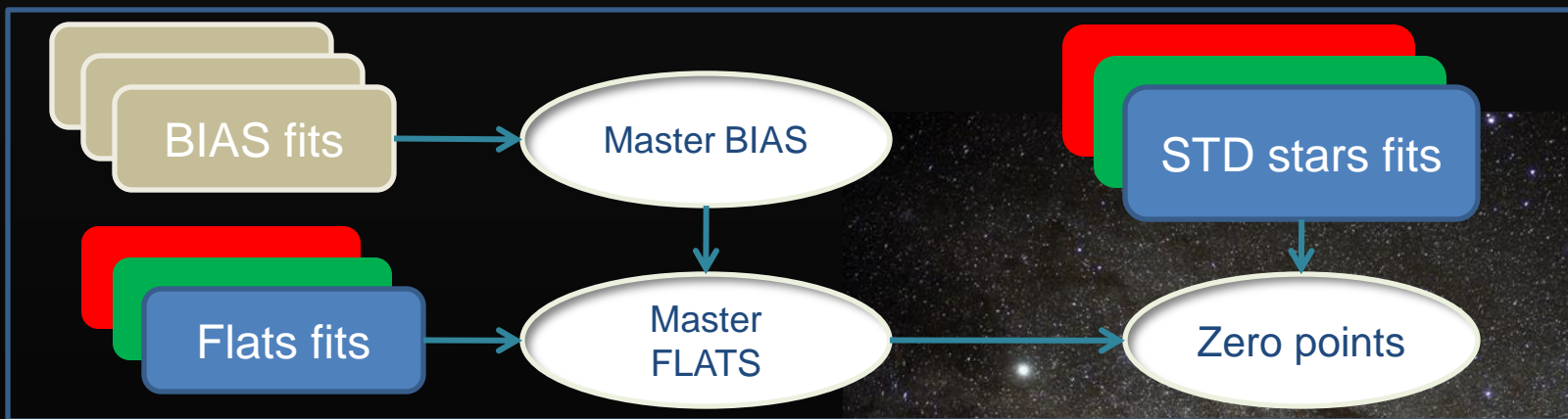
- From configuration file: parameters for the analysis
 - List of files to copy from tape to /nfs/
 - Prestage of files D0T1
 - srmCP to /nfs
 - Decompress fits.gz files



PART 1: Calibration

Instrumental calibration

- Bias runs
- flats runs
- masking cosmetics, cosmic rays traces, satellites traces



Master MASK: cosmetics,
cosmic-rays, saturated pix.

Photometric calibration

Standard stars images (SDSS Stripe 82)
Calibration: source extraction and matching
Magnitude calibration factor
Headers of simulated files

PART 2: Scientific images reduction

Reduction pipeline for each image

- Instrumental calibration
- Source extraction (SExtractor)
- Catalog conversion and Astrometry (SCAMP)
- Production of a calibrated catalog



JOBS MANAGEMENT

main.py

- Initialization of configuration parameters

JOB: Pre-stage and Master BIAS

- Check job status function
- Configuration of DAG-job jdl (one per filter)



Parallel jobs

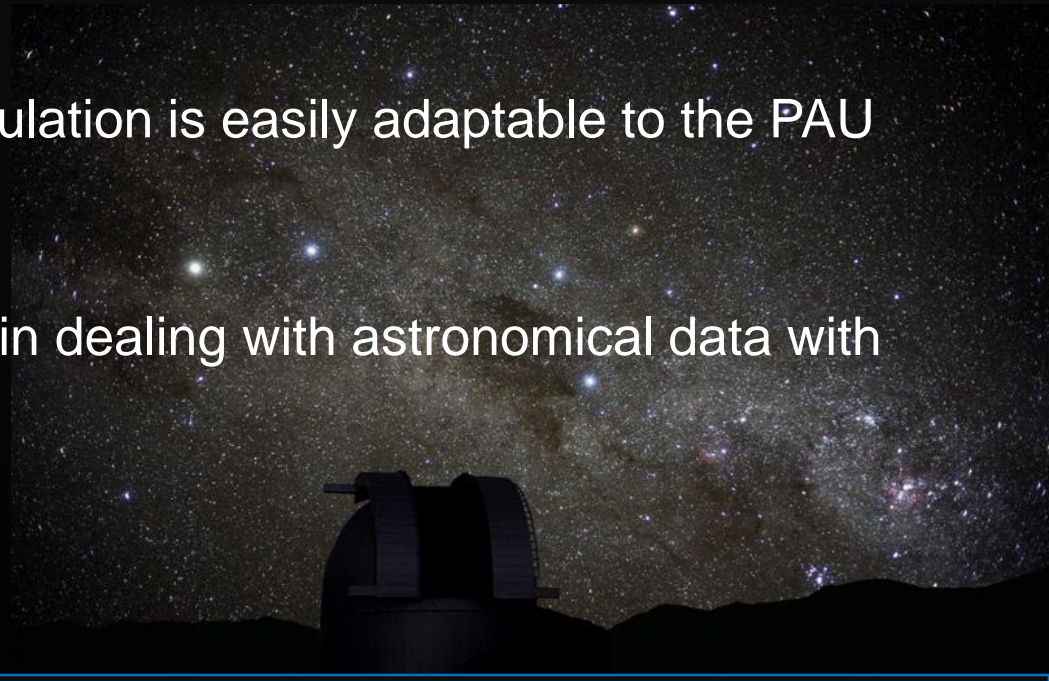
Database

- In design phase
- Information about:
 - Scientific objects (galaxies: position, magnitude, shape, etc..)
 - DM metadata (file location, access permissions, etc...)
- Based on DES database, with some differences:
 - Access limited to 10-20 users
 - Low query rate
 - High number of objects and parameters (10^9 entries, 200 params per object)
 - Estimate size: 1TB



IMPACT

- Availability of data to all members of the user community
- Data chain parallel to the 'official' DES reduction code, to analyze simulations by the DES collaboration
- Pipeline used for the simulation is easily adaptable to the PAU survey data structure.
- Exchange of experience in dealing with astronomical data with MAGIC collaboration



Summary

- The PAU survey will take data in some years, producing information of a huge number of objects. Results will be organized in catalogs.
- Simulation files are managed at PIC and used to make preliminary studies and test the data reduction pipeline.
- Data pipeline we are designing and testing in the GRID environment will be flexible enough to fit the requirements of the real data.
- JOBS manager is written in python and allow us to run and check the status of many jobs running in parallel and in an easy way for the user.
- A database for the resulted catalogs is in project.