



pracelab **2**

IDAC-NCBJ STATUS

Agnieszka Pollo

team: Krzysztof Nawrocki, Bartosz Zięba, Tomasz Fruboes, Adam Zadrożny, Orest Dorosh

Installation Site of Light IDAC



- Light IDAC will be installed at NCBJ cluster at PSNC, Poznań, Poland
- Light IDAC will be part of larger system
 KMD/pracelab2 25PB, 5-8k CPU cores
 system + GPU
 - Resources to be increased the following years
- KMD is developed by consortium of institutions including NCBJ





Specification of Light-IDAC (POL-NCB S1)

Specification of light IDAC (according to the proposal):

- 500 CPU cores (20-30 cores for system/db, 470-480 cores for users)
- 0.5 TB of storage space for Object Light Catalogue
- 4.5 TB for catalogues
- 2 x 0.25 FTE support stuff
- proposed center could serve 500-1000 users
- and 250 users using center concurrently

Under consideration:

- 10 nVidia GPU providing ~50 TFLOPS

As specified in guidelines: <u>https://rtn-003.lsst.io/</u>







Roadmap

- Q3.2023 open equipment tenders
- Q3-Q4.2023 fully working demonstrator with RSP and dedicated HPC solutions
- Q4.2023 offer selection and hardware purchasing (1st phase)
- Q4.2023 hardware installation
- Q4.2023-Q2.2024 Data Center v1.0 fully operational with configuration from demonstrator implemented
- Q2-Q3.2024 start of operations
- (all dates <u>subject to change</u>, if memorandum of agreement not signed before <u>Aug 31</u>)

Science Cases

- Science Pipeline Development in the LSST
 - Galaxies Science Collaboration
 - Dark Energy Science Collaboration
 - AGN Collaboration



- Other science cases submitted by local teams inside the LSST-PL colaboration
 - Solar System
 - Time Domain
 - Stellar physics

Science Pipeline Development in the LSST Galaxies Science Collaboration

 Directable software development effort (the CIGALE SED fitting tool)

- Contribution Lead Katarzyna Małek
- User-generated datasets: catalogs of galaxy properties (and, possibly, photometric redshifts)
- Catalogs will be made accessible to Galaxies SC members through appropriate data centers, among them the local IDAC

Science Pipeline Development in the LSST Dark Energy Science Collaboration

- Developing software for wide-angle photometric redshift galaxy catalogs
- Selection and photometric redshift (PZ) estimation of galaxies from the LSST data products
- Catalogs will be made accessible to DESC members through appropriate data centers, among them the local IDAC





Contribution Lead Maciej Bilicki Science Pipeline Development in the LSST AGN Collaboration

- Software development effort in the general area of LSST AGN variability science analysis
- The code will be assimilated into the LSST pipeline
- Spectroscopic monitoring of selected quasars with 10-m class Southern African Large Telescope
- Photometric monitoring with OGLE, SMARTS and BMT telescopes
- Multi-method time delay determination In data analysis for this monitoring

POL-NCB-S6



Contribution Lead Bożena Czerny

Demonstrator

- Demonstrator system being developed and tested in Supercomputing Center Świerk (CIŚ)
- Two modes of using resources:
 - Rubin Science Platform
 - Multipurpose science platform (for high performance/dedicated tasks) developed by us
 - Kubernetes / SLURM
 - Jupyter / SSH / Desktop access



Note: NCBJ experience in serving high availability and reliability computing infrastructures and services, eg. for CERN experiments: Tier-2 site for LHC-B (going to become Tier-1 in 2023) and Tier-2 site for CMS (going to become Tier-1 in 2024)



Hardware configuration used for demonstrator

C jupyterhub Home

- blade server using virtual machines
 - 10 CPU cores
 - 40 GB RAM
- fully working Rubin Science Platform
- developed Multipurpose Science
 Platform for HPC
- currently in process of testing science cases for full IDAC

Location: Supercomputing Center Świerk



Demonstrator Science Case (POL-NCB-S3)

- Galaxy model fitting based on DP0.2 catalogue
- Test for a set of galaxies (1k, 10k, 100k) with redshift
 - redshift, flux in filters u, g, r, y, i, z
- Galaxy model fitting using CIGALE was done on RSP version of demonstrator
- Tested on restricted (100) and typical number of spectra (28k) in the grid
- Generated data from fitting would be served via database as part of Object Data Catalogue



Analysis of all galaxies in LSST DR1

• total time of analysis: 10,2-17,5 day

Assumptions used:

- assumed number of galaxies: 4B
- assumed number of cores used: 400 (70 left for users and 30 for system/db)
- number of models to fit: 28644
- 50 separate processing streams each with 8 cores assigned
- each stream will take series of bulks of 100k galaxies each
- 10% overhead time on measured operations

(conservative estimation - the final setup is expected to be faster than that)

Summary

- we are on time with IDAC roadmap
- currently our demonstrator is using
 - Rubin Science Platform
 - Multipurpose Science Platform HPC
- we in process of testing science cases for full IDAC
- we expected have fully operational center v1.0 in Q2 2024
- start of operations can be planned planned starting Q3 2024 (assuming signed MoA in Q3 2023)

