

News from Computing

Documentation : <https://doc.lsst.eu>

Main Slack channel: [#in2p3](#)

Data Previews - Schedule and Data Products

Rubin Baseline Data Release Scenario	Jun 2021	Jun 2022	Jul-Sep 2023	Apr-Jul 2024	Feb-Jul 2025
Data Product	DP0.1	DP0.2	DP1	DP2	DR1
	DC2 Simulated Sky Survey	Reprocessed DC2 Survey	ComCam On-Sky Data	LSSTCam On-Sky Data	LSST First 6 Months Data
Raw Images	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP Processed Visit Images and Visit Catalogs	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP Coadded Images	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP Object and ForcedSource Catalogs	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP Difference Images and DIASources	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP ForcedSource Catalogs including DIA outputs	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP Processed Visit Images	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP Difference Images	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP Catalogs (DIASources, DIAObjects, DIAForcedSources)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP Alerts (Canned)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP Alerts (Live, Brokered)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
PP SSP Catalogs	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
DRP SSP Catalogs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Finding its way through the various data facilities

- **NCSA** : Historic data facility where all the datasets are stored and where most of the software developments occur
 - NCSA is scheduled to stop operations for Rubin on Oct. 1st 2022
- **Interim Data Facility (@Google)** hosting the Rubin Science Platform and where the DP0 processing is occurring
 - Contract with Google is running up to Dec. 2023
- **US Data Facility (@SLAC)** potentially hybrid with Google
 - Will have to host the DP1 (ComCam) data for security reasons and because NCSA will be shut down
 - Cost of security : 24 M\$
- The NCSA closure + DP1 schedule makes me think that the hybrid solution (SLAC + Google) is unavoidable
- + **FrDF** (CC-IN2P3)
- + **UKDF**

Data Preview 0

DP0 is based on the DESC DC2 dataset

- DP0.1 (2021)
 - DC2 catalogs ingested in Qserv database
 - Images and catalogs also available through the butler (gen3@IDF and gen2@CC-IN2P3)
 - IDF : access to 300+ delegates through the Rubin Science Platform (RSP) only
 - CC-IN2P3 : access through the notebook platform : notebook.cc.in2p3.fr
- DP0.2 (2022): full reprocessing from raw images
 - reprocessing will be done independently at CC-IN2P3
 - the automatic job distribution system between sites will not be available
 - but various tests will be performed (PanDA...)

People involved

- At CC: C. Eloto, Ph. Cheynet, A. Georget, F. Hernandez, Q. Le Boulch, G. Mainetti, L. Schwarz, L. Tortay, F. Wernli
- + D. Boutigny (LAPP), J. Bregeon (LPSC), S. Elles (LAPP), F. Jammes (LPC), M. Paturel (LAPP)

At CC-IN2P3: B. Gounon left, 1 new permanent position opened in 2022

⇒ **Very good support from CC-IN2P3 - Much appreciated !**

Data access services associated to DP0

Qserv

- Qserv Kubernetes platform with 15 workers maintained by CC-IN2P3
 - production instance
 - test instance
- Data ingestion tool - Qserv packaging - Qserv expertise (Fabrice)
 - Qserv@CC kept as similar as possible to Qserv@IDF
 - But extra datasets at CC-IN2P3 (cosmoDC2 and skySim5000)

Rubin Science Platform (RSP)

- 2 instances deployed at LAPP in the framework of ESCAPE WP5 (ESFRI Science Analysis Platform)
 - Demo of notebooks - butler access - Firefly data display - IVOA Table Access Protocol
- 1 instance at CC-IN2P3
 - Not yet in production
 - Some work to adapt it to CC environment
 - identity management
 - access to local Qserv instance
 - ...

On hold since Bastien departure

But not yet critical for data access since there are alternatives

Work in progress

Image processing

- butler gen3 by Quentin
 - Optimization of the registry database handling (postgresql)
 - Local registry on worker (a.k.a. execution butler)
 - Update of the central registry afterward
- workflow executed on Slurm
- LSST science pipelines executed in containers for isolation

- tests of the DIRAC middleware by Johan B.



Data replication between sites

- assessment of [Rucio](#) by Lionel
 - Data Management system originally developed by ATLAS
 - tests from CERN → CC-IN2P3 (ESCAPE project) to check whether it is suitable for Rubin

⇒ Will need physicist eyes to validate the first gen3 produced catalogs at CC-IN2P3

Building the FrDF@CC-IN2P3

Regular coordination meetings between USDF, FrDF and UKDF are occurring

Work in progress on the FrDF sizing and cost model by Fabio

- First model in 2018 based on initial Rubin estimation
- New Rubin model following the decision to locate the US Data Facility at SLAC
 - Very significant increase of the storage (total dataset size 0.5 eB → 2 eB)
 - Decrease of the CPU needs but not compensating the increase in storage

BUT...

- The model is not directly transposable to the FrDF
 - large uncertainty on the storage dedicated to intermediate data
 - not clear what will be finally the IN2P3 commitment w.r.t. data replication

Once again, we need to clarify this point in the future MOA

- **Not including Rubin-LSST France scientists in the discussion is a very bad idea**

imSim image simulation at IDRIS

Proof of concept in the framework of the [FITS](#) Equipex involving IDRIS and CC-IN2P3

- 50 000 CPU hours on Jean Zay
- A similar exercise in 2020 has shown that Jean Zay was suitable for CPU intensive imSim runs
 - this time the idea is to deploy a fully automated workflow including data transfers between both sites
 - 100 Gb/s dedicated link
- complex due to very strict security constraints on the IDRIS side

A second, more challenging project is to execute to notebooks with access to GPU on Jean Zay via CC-IN2P3 notebook platform



A few announcements from CC-IN2P3

/sps/lst storage area

- The migration from GPFS to CephFS is now essentially complete
 - complex and painful but it is done, hopefully without too much disruption !
 - opportunity to re-organize the “sps” area
 - No more /sps/lst/data “junk”
 - groups area :



```
cca007[1] cd /sps/lst/groups
cca007[0] ls
FilterExch FocalPlane Monocam accounting auxtel cadence desc hsc photoz qserv transients
```

- users area
 - your own user area is available in /sps/lst/users when your account is created

pro tips

- /pbs/throng/lst/users : backed up area for your code development, limited in size
- /sps/lst/users : larger, not backed up area for data and large files to be kept from one session to the other
- /scratch : temporary area automatically cleaned up on a regular basis
- please make sure you have read the [documentation on this specific topic](#)

A few announcements from CC-IN2P3

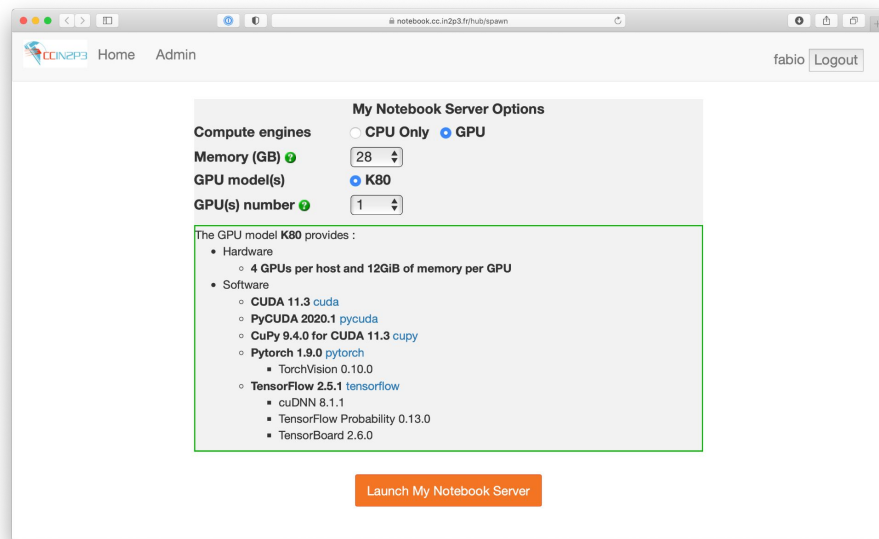
The batch system is going to change

- Grid Engine → Slurm
- See: <https://doc.cc.in2p3.fr/fr/Computing/computing-introduction.html> to migrate your scripts
- One can already submit batch jobs to Slurm
 - More workers will be progressively added until the migration is complete
- Interactive Slurm sessions are also possible (will replace `qlogin`)
- Training session for end users to be held early 2022



Notebook platform

- Regularly updated and improved
- You can run your notebook server on CPU- or GPU-equipped machines
 - 4 x GPU Nvidia K80, 12 GB each
 - expected increase in January (x2 or more)
 - needs authorization, just ask if you need
- Remaining issues
 - low I/O performance due to virtualization
 - interactive visualisation (à la DS9)
- Exploratory work ongoing to integrate Dask (for execution in Slurm cluster)



<https://notebook.cc.in2p3.fr>

DESC Computing

- Creation of the DESC Data Facility Committee
 - DESC management + NERSC (D. Bard) + France (D. Boutigny) + UK (George Beckett)
- Up to now we have not distinguished Rubin and DESC resources at CC-IN2P3
 - IN2P3 asks us to make a separate accounting of the resources
 - Probably difficult - Need to think on how to do that

CO working group (Johann C.T. was co-convenor up to recently)

- Effort to structure and organize the COmputing environment and related tools
 - Software Management Plan
 - DESC-help
 - Request for Comments (RFC) as in Rubin DM ⇒ <https://github.com/LSSTDESC/RequestForComments>
 - Start designing a system for managing intermediate & output analysis data (RFC #13)

DESC software at CC-IN2P3

The software environment is very similar to the one at NERSC (+ some specificities)

- desc-python package (H. Kelly)
 - `source /pbs/throng/lst/software/desc/common/miniconda/setup_current_python.sh` will initialize a standard DESC conda environment
 - Kept up-to-date - New versions announced on #in2p3 slack channel
- If you need specific packages on top of desc-python the simplest thing to do is to clone the DESC conda environment
 - `conda create --clone desc -p /sps/lst/users/boutigny/conda_envs/desc_test`
then:
`source /pbs/throng/lst/software/desc/common/miniconda/setup_current_python.sh`
`conda activate /sps/lst/users/boutigny/conda_envs/desc_test`

See also [documentation](#)

Notebooks

2 possibilities:

- **notebook platform** <https://notebook.cc.in2p3.fr/>
 - Simple, well maintained, easiest way to get interactive access to GPU
 - But some (temporary) limitation on memory, I/O bandwidth, incompatible with Dask
 - Create your jupyter kernel as in: <https://doc.lsst.eu/tutorial/notebooks.html> and <https://github.com/combet/desc-jupyter-kernel>
- **stackyter.py** developed by N. Chotard and maintained by Johann et. al
 - <https://github.com/nicolaschotard/stackyter>
 - `pip install stackyter`
 - doc in: <https://doc.lsst.eu/tutorial/stackyter.html>

Accessing data

3 possibilities:

- **GCRCatalogs** - <https://github.com/LSSTDESC/gcr-catalogs>
 - all the DESC catalogs are available at CC-IN2P3
 - a notebook running at NERSC should also run at CC
 - many examples here: <https://github.com/LSSTDESC/DC2-analysis/tree/master/tutorials>
 - But slow when accessing large catalogs
- **Dask** accessing Apache parquet files
 - Dask provide advanced parallelism for data analysis
 - Fully integrated with python / numpy / pandas
 - Interfaced with batch systems
 - with 1 command you can create a cluster with 100s of nodes
 - Lazy evaluation : nothing is actually executed until you decide to launch computation
 - each instruction is feeding an execution graph which will be run in parallel

Accessing data

- **Dask** (cont'd)
 - Full example with instruction to launch it [here](#)
 - It is possible to run a realistic analysis (galaxy clusters) on the full skySim5000 catalog (29 TB) in < 5 minutes
 - 100 - 200 workers with a total of 1 TB of memory can be attached to your notebook
- **Qserv**
 - Accessing data through Qserv directly can only be done at CC-IN2P3 and is open for everyone
 - DC2 object catalogs (stack processed) + truth association
 - cosmoDC2
 - skySim5000
 - Documented examples are available [here](#)
 - Also demonstrate how to launch parallel queries

So, should I use Qserv or Dask

It depends !

- Qserv will be very efficient on a complex query spatially contained in small or medium size region
 - Should also be very efficient to extract light curves / time domain series
 - Not tested by me
- Dask is incredibly fast on full sky analyses

We need use cases to better understand data access pattern

**We will be happy to help you implement your data selection / analysis
either on Qserv or on Dask**

⇒ I am available for demos

AuxTel data

Sylvie has copied some AuxTel data in:
</sps/lstt/groups/auxtel/data> for hologram analysis

We have all the tools / software to run these analyses
at CC-IN2P3

