

# A proto-data processing centre

Cécile Cavet, Antoine Petiteau, Maude Le Jeune, Eric Plagnol, Etienne Marin-Martholaz,

Jean-Baptiste Bayle

François Arago Centre, APC, Univ. Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs. de Paris, Sorbonne Paris Cité,

13 rue Watt, 75013, Paris, France



## Abstract

The LISA project preparation requires to study and define a new data analysis framework, capable of dealing with highly heterogeneous CPU needs and exploiting the emergent information technologies. In this context, we started to develop and provide a prototype of the mission's **Data Processing Centre (DPC)**. The DPC is designed to efficiently manage computing constraints and is able to offer a common infrastructure where the collaboration-wide development work could start. Several tools such as continuous integration have already been delivered to the collaboration and are used now for simulations and performance studies. This poster presents the progress made regarding this collaborative environment and also discusses the possible next steps towards an on-demand computing infrastructure. This activity is supported by the CNES as part of the French contribution to LISA.

## Context

► **DPC**: a set of tools provided to ease the challenging data analysis (DA) tasks of LISA.

- First observations of this kind: DA techniques to be invented.
- Fluctuation of CPU charge to be handle by new technologies (virtualization, DevOps...).

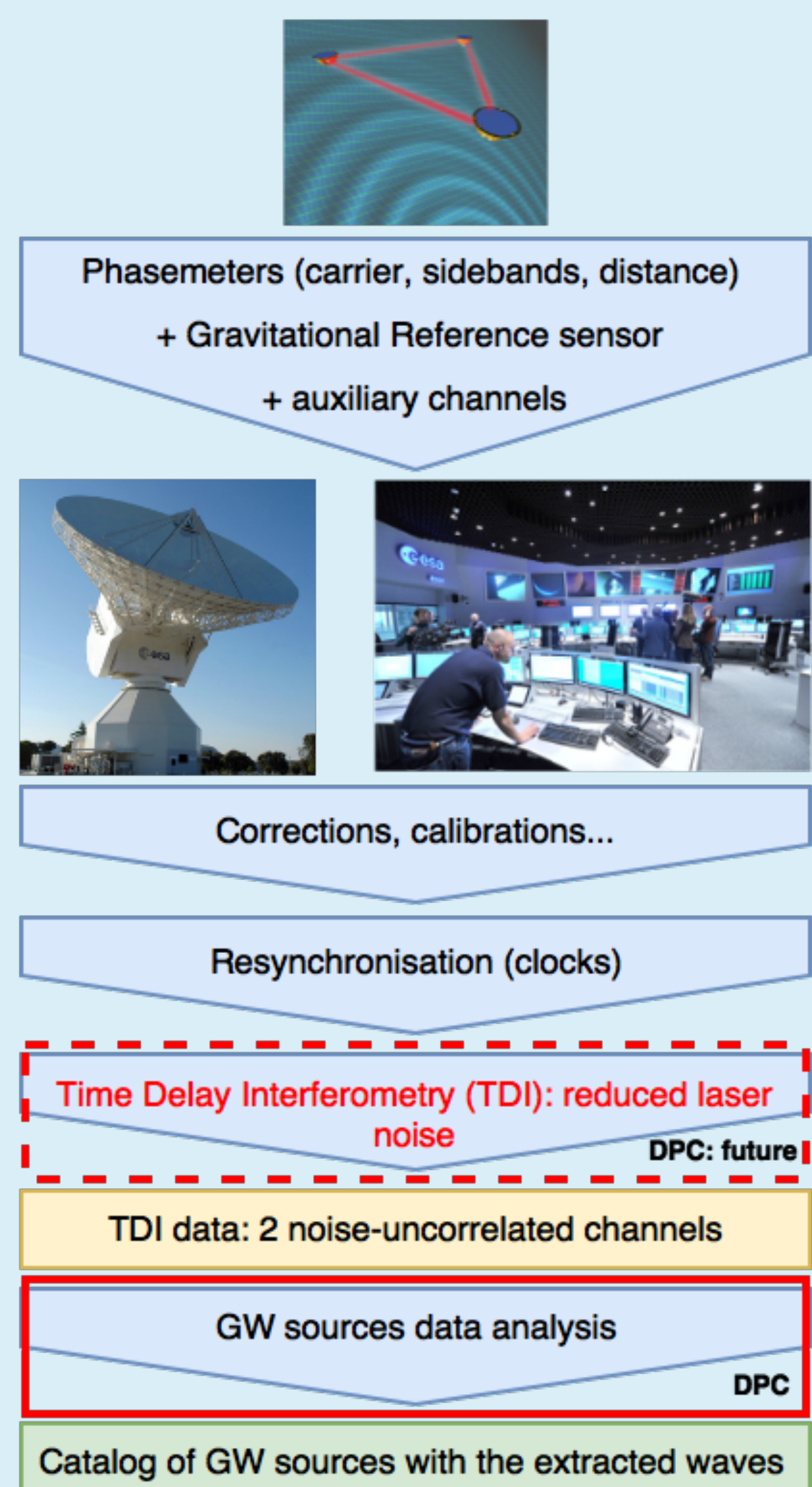
► The **DPC** has to be flexible and easily upgradeable until the end of the mission.

► **Why now?**

- Need for a place to start a collaborative work on DA and simulation.
- DA challenges: restart MLDC with more realistic simulations (noise from LisaPathfinder, more GW sources) to encourage the development of new DA techniques.
- End-to-end simulations: assess the industrial proposals by 2017.

► The **proto-DPC** is the framework for LISA simulations and for next DA collaborative activities as MLDCs.

## Data Analysis (DA)



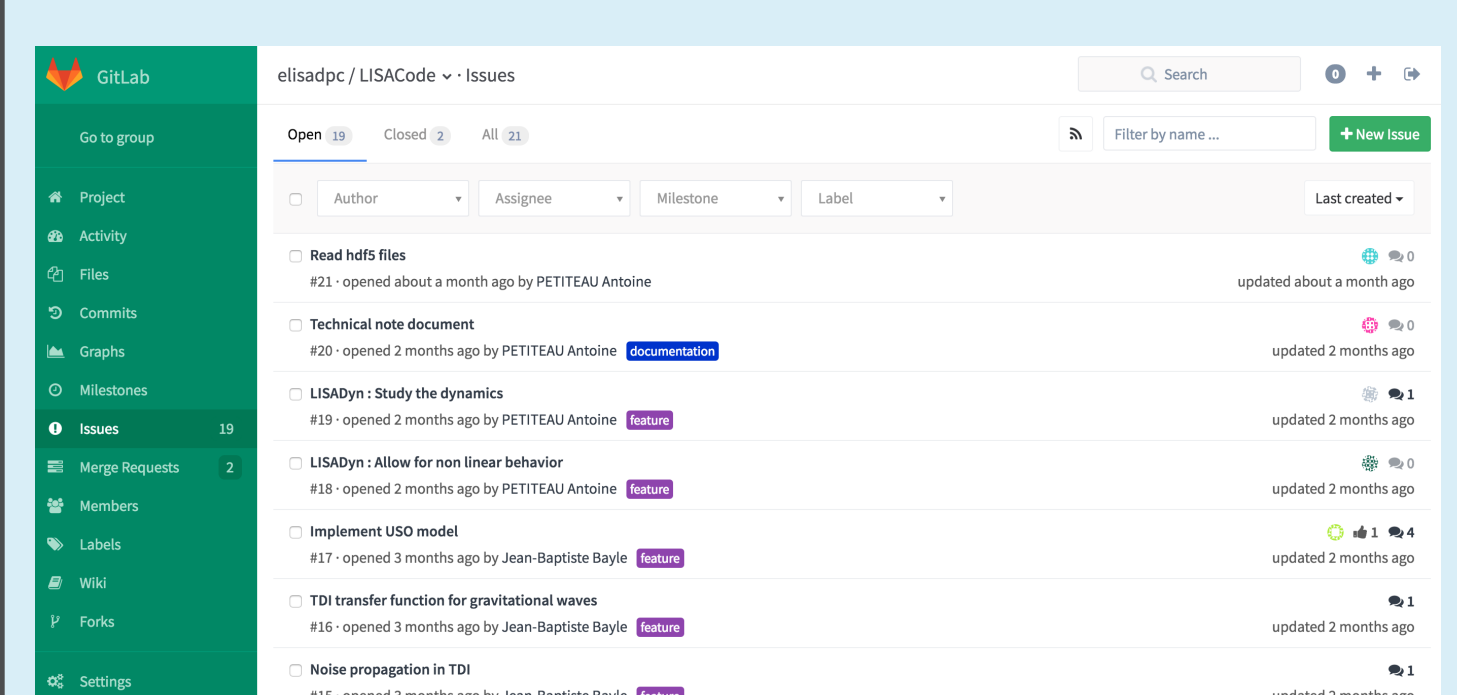
Workflow of DA in LISA.

## LISA simulation

► **Goals:**

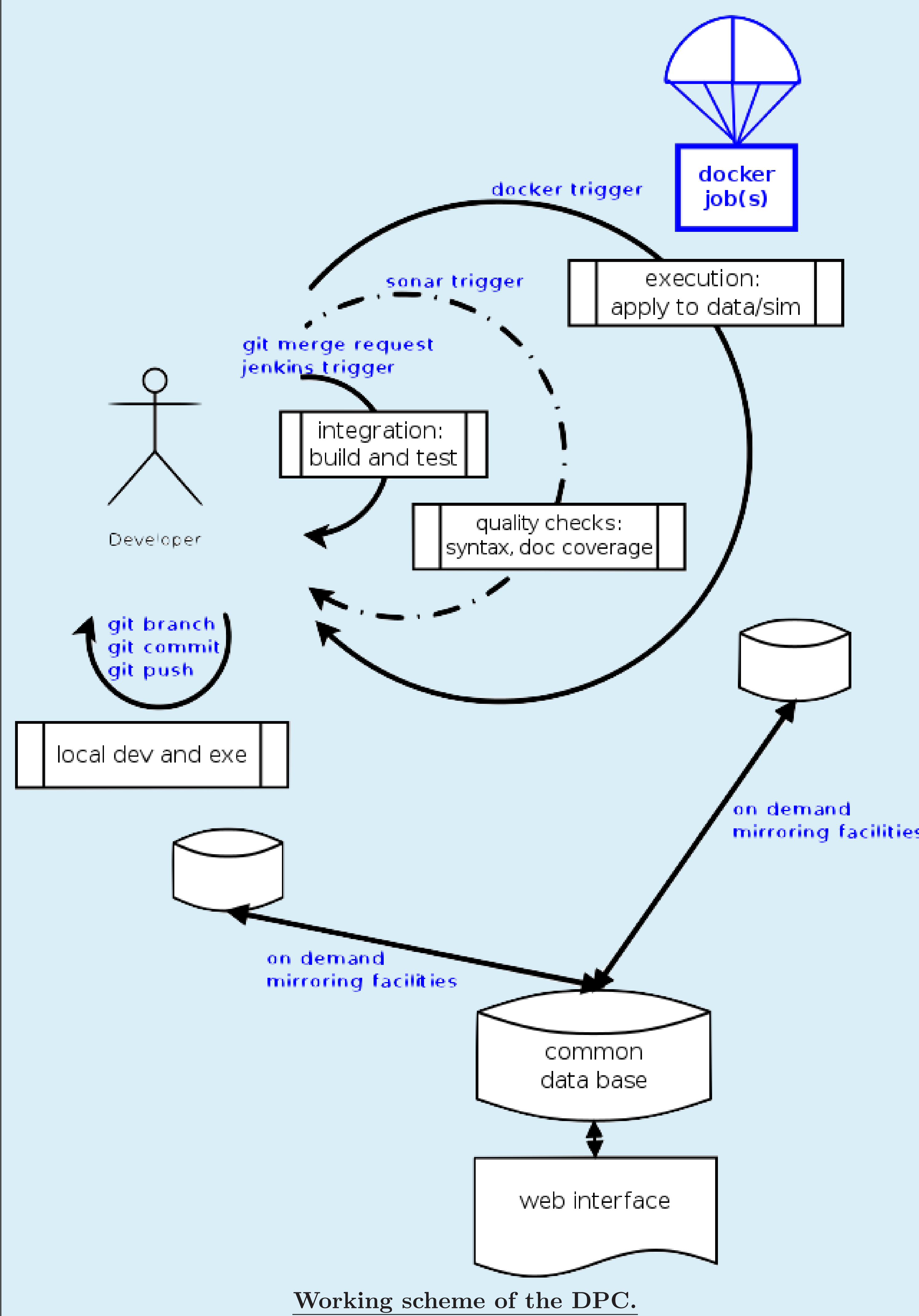
- The mission simulator (end-to-end simulations).
- "Quick" performance study: final design, need for the mission's phase A.
- Tool for the performance control.

► **LISACode** [2], the first basis of the end-to-end simulator is already managed by the **CI** platform of the **DPC**!



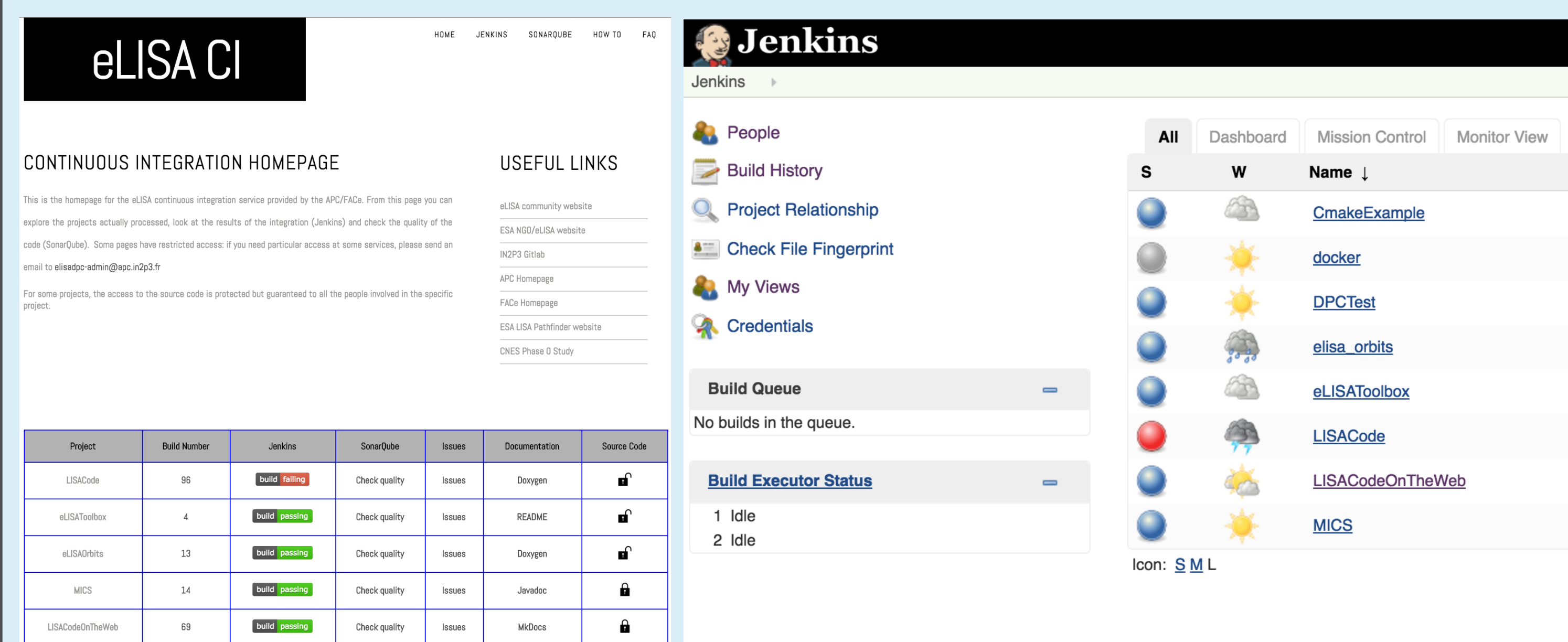
GitLab repository of LISACode.

## Status



Working scheme of the DPC.

## Continuous integration (CI)



CI platform (left) and Jenkins interface (right).

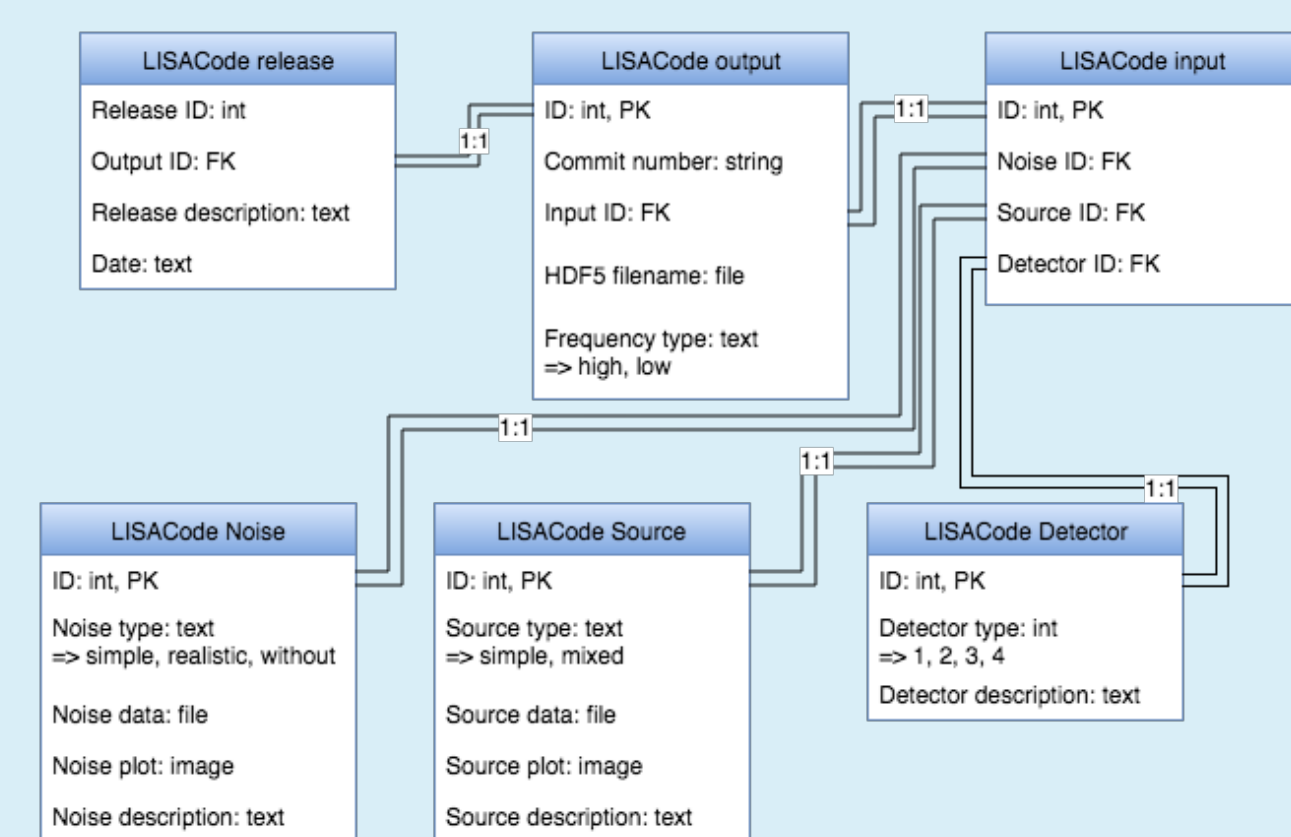
► **How the CI works:**

- Code repository on **GitLab@CC-IN2P3** (Computing Center of IN2P3 institute) [2].
- CI platform@FACe (François Arago Center, spatial mission data processing center of APC laboratory) [1] with **Jenkins**: building, execution, automatic documentation and unit tests; and **Sonar** for code quality: coding standard, documentation coverage.
- Code development and execution with **Docker** [3]: automatic build of Docker image on the DockerHub, Docker container as reference environment for LISACode simulator.
- Use of the academic **IaaS cloud**: virtual machines and Docker containers.

## Mock Lisa Data Challenge

► **How the MLDC works:**

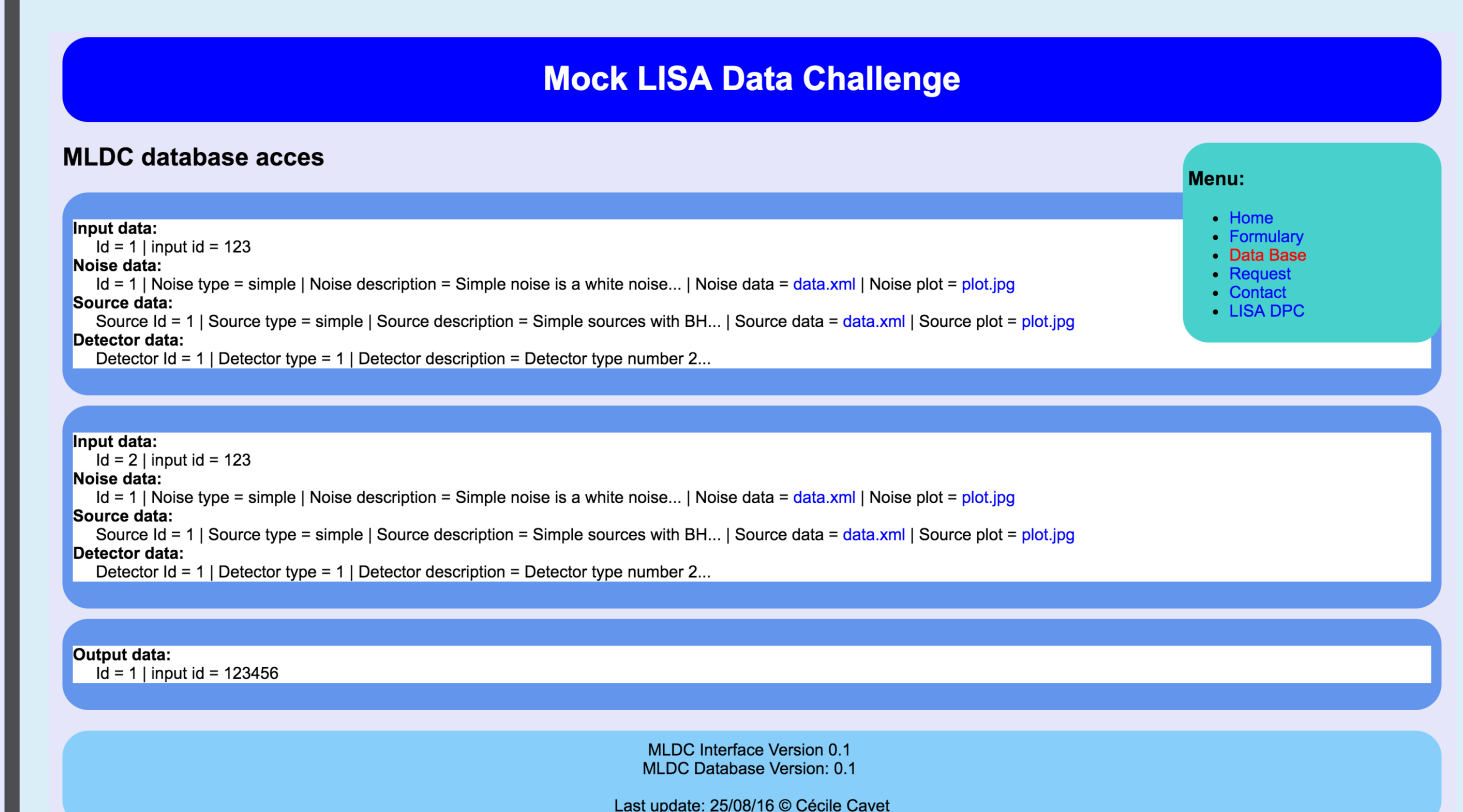
- simulated data = several sources + instrumental noise.
- teams have to analyse the data to retrieve the source parameters.
- results are compared to the input between the teams.
- increasing complexity of the simulated data.



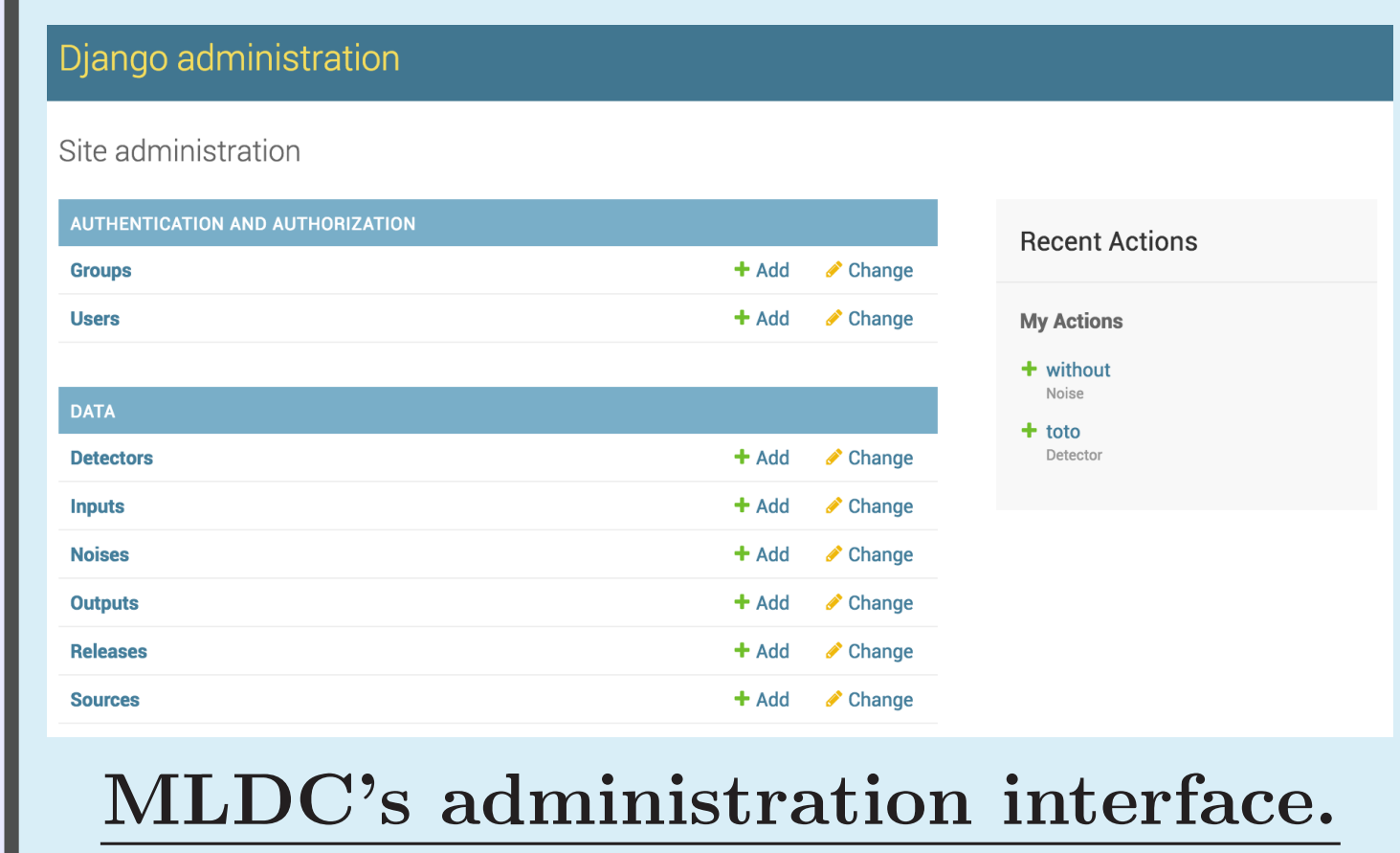
Scheme of the MLDC's data base.

► **Running the next MLDC within the DPC:**

- need a common data base to store the simulation and the team results.
- handling of the simulator versions and the parameters by the configuration management.
- comparison software hosted by the DPC and accessible to everyone.
- hosting the DA code as much as possible to improve *in fine* data quality and CPU performance.



MLDC's Proto-Web application.



MLDC's administration interface.

## How to contribute?

► **Open to suggestions / contributions/collaborations:**

- DPC platform: <http://elisadpc.in2p3.fr/>
- Create an account on GitLab@CC-IN2P3: <https://gitlab.in2p3.fr>
- Send an email to petiteau at apc.univ-paris7.fr (with your name /email).
- Other projects can be integrated.

## References

- [1] DPC platform: <http://elisadpc.in2p3.fr>
- [2] LISACode GitLab repository: <https://gitlab.in2p3.fr/elisadpc/LISACode>
- [3] DockerHub: <https://hub.docker.com/u/elisadpc>