LISA Computing

HTC / HPC

Antoine Petiteau for the LISA group

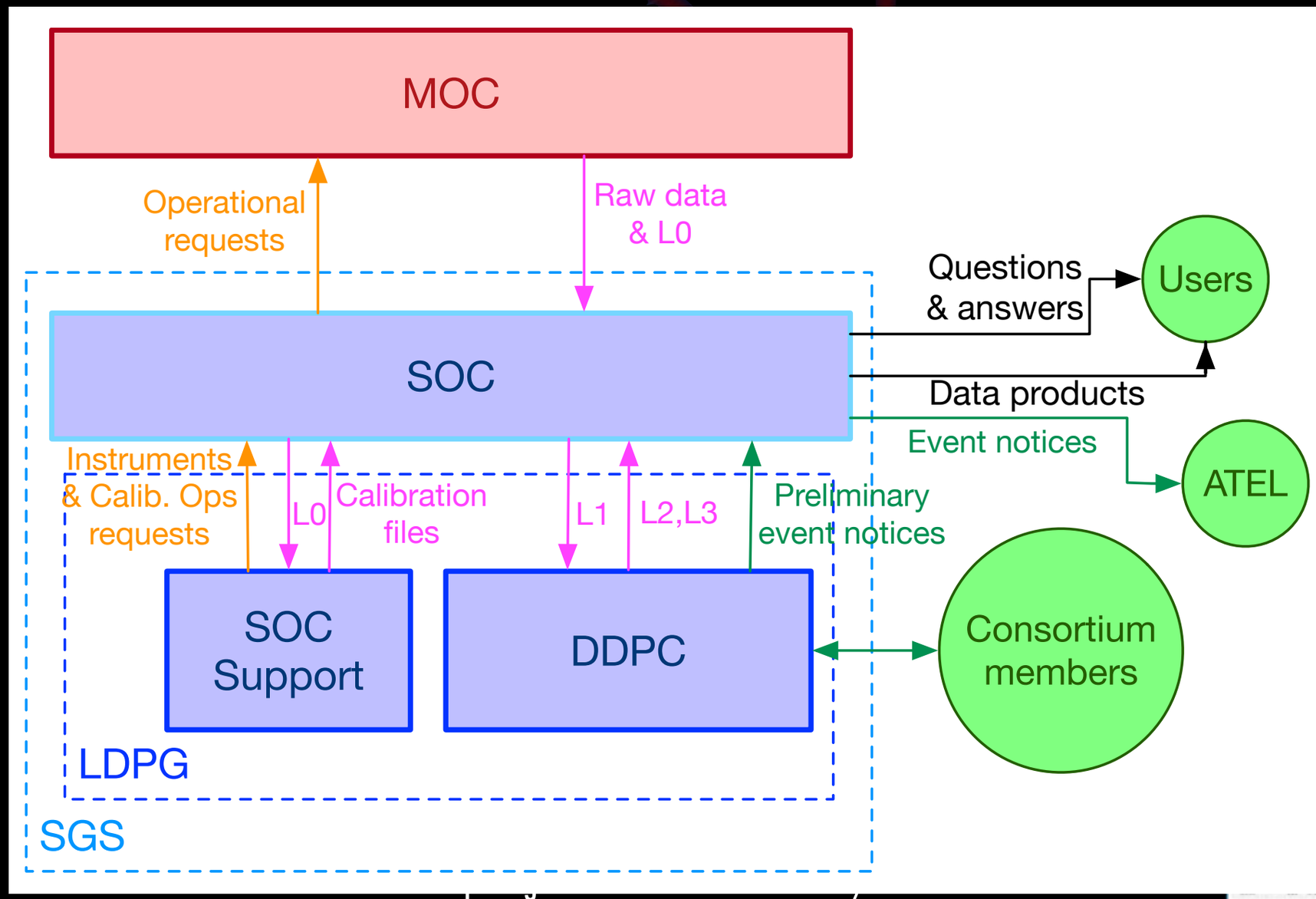
(APC - Université Paris-Diderot)

DANTE APC/IN2P3

20th december 2018

LISA Ground Segment

- The major part of the Ground Segment from the LISA Consortium is the Distributed Data Processing Center: French responsibility with APC lead

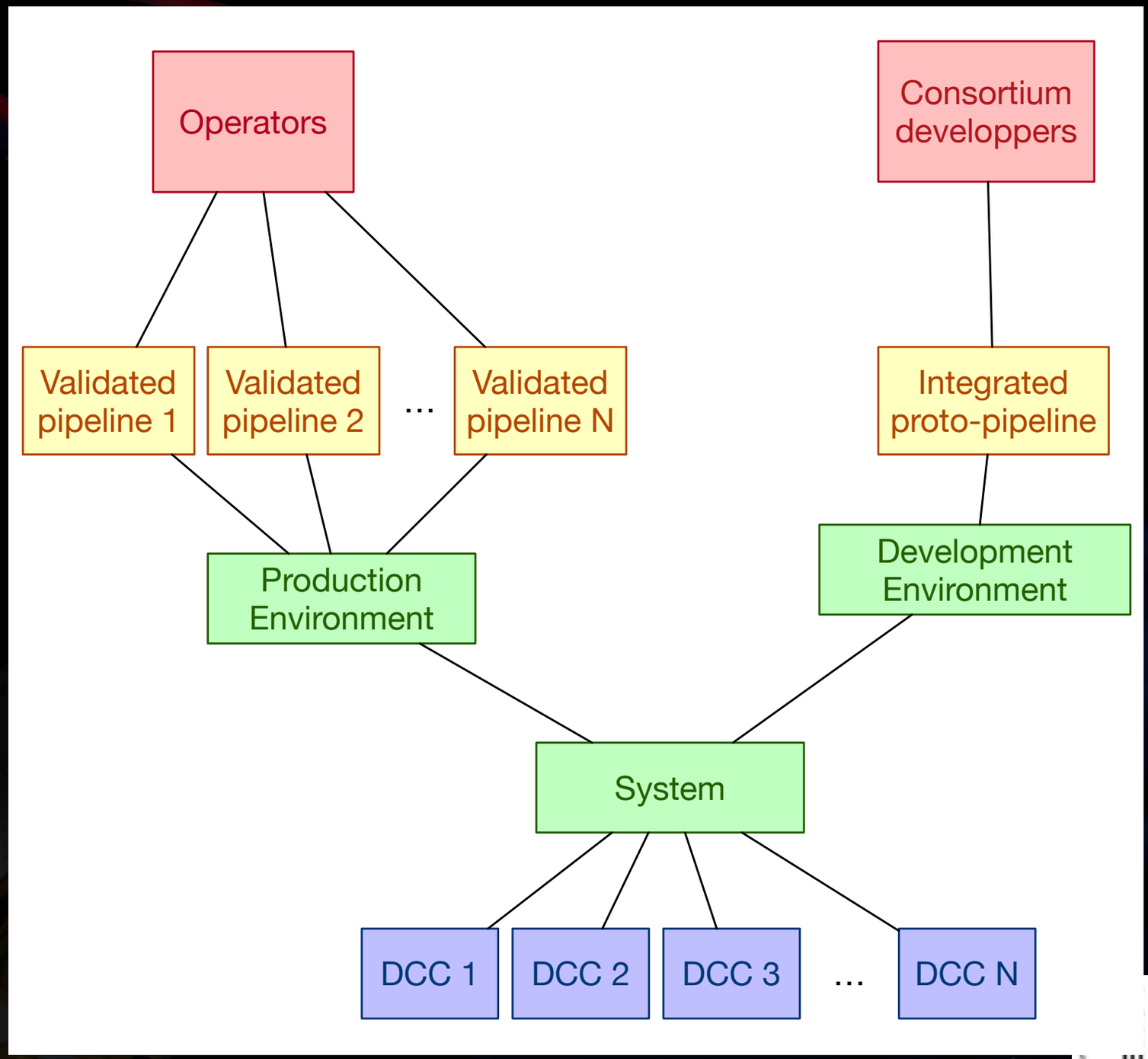


LISA Phase A/B1

- Part of the tasks for phase A (2018-2020) and B1 (2020-2022):
  - Design architecture, organisation, etc of GS with CNES and other partners $\Rightarrow$ consolidated design and costs
  - Prototyping ... using support activities ...
  - Support of the Consortium and Agencies activities:
    - Simulation $\Rightarrow$ Instrument design including on-ground pre-processing
    - Data analysis pipelines: design, first implementation, tests
    - Data management: data basis
  - Services

Common system: dev./prod.

- First ideas based on a common system:
  - short cycle between dev. & prod.
  - distributed hardware on DCCs (Data Computing Centres)
  - cloud compatibility



Proto-DPC: basics

- Development environment: in production
  - Collaborative work, reproducibility of a rapidly evolving & composite DA pipeline; Keep control of performance, precision, readability, etc
  - Use existing standard tools (version control, Continuous Integration, Docker)
- Data basis & data model: in R&D
  - Data sharing, a lot of information (search engine, DB request, tree view);
  - Context: Not very big data volume for data itself but large number of sub-products, simulations, ... => LDC, simulations, LPF data
- Execution environment: in R&D (singularity, ...)

Data analysis & simulations

- Simulations:
  - Simulations at different scales: micro-sec to years in reasonable time
  - Coherently simulate control loops, integrate discretization/ interpolation, precisions, …
- Data pre-processing: clock, ranging, TDI
- Data processing: extracting science
  - For the matched filtering: optimisation of likelihood computation, variety of samplers, possibly large number of parameters, evolving number of parameters, …
  - Orchestration of multiple pipelines in parallel
  - Keep track of all produced data
  - Incremental data: new data to integrate every day
  - Fast pipeline for alerts, …

LISA computing

- Data analysis: extraction of GW sources:
  - Most of the technics based on matched filtering: large number of template - likelihood computation in parallel:
    - parallel (mostly MPI) with limited exchange between processes (direct scaling => more core always better)
    - template can be heavy: few GBytes of RAM per process
    - I/O limited: load data at start
  - Growing interest for machine learning:
    - could be interesting to have some CPU/GPU (TBC)
    - ...
- Simulations: sequential, Monte Carlo, (future: parallel)

LISA data

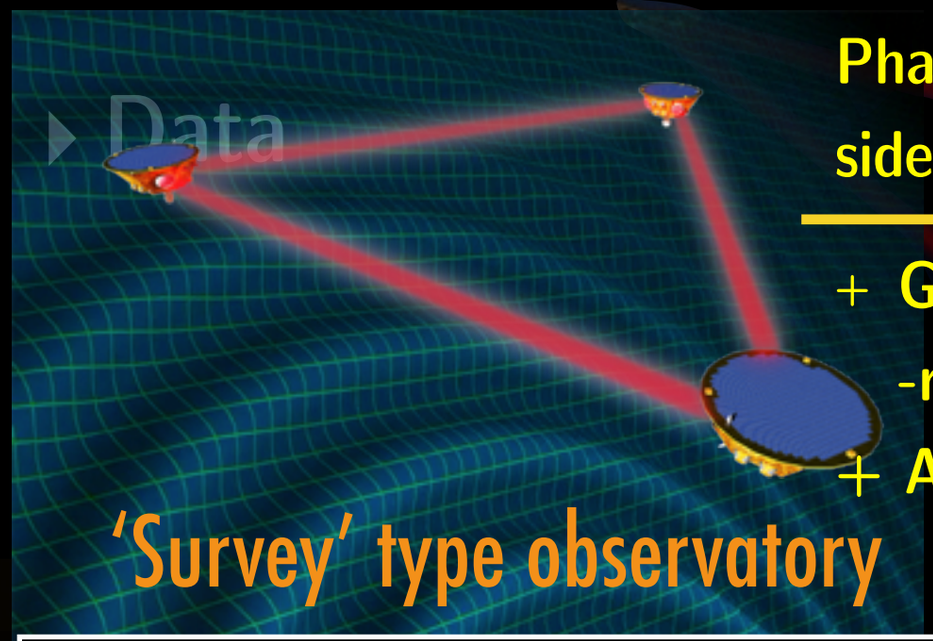
- Not big volume for the mission data: 1Go/day for the telemetry
- But large number of simulations
- Requires database to be manage over long period > 20 years
- Data sharing

Conclusion

- **LISA Data Processing**:
  - French responsibility
  - APC lead with support of CCIN2P3
- DANTE could be the **prototype** of the core center for LISA
- **Computing ressources requested** now:
  - science case: detection, parameter estimation,
  - pipeline development and test (LISA Data Challenge),
  - simulation
- **Future**: one of the main (or the core) center

Merci

LISA data



'Survey' type observatory

Phasemeters (carrier, sidebands, distance)

+ Gravitational Refe--rence Sensor

+ Auxiliary channels

L0 Calibrations corrections

Resynchronisation (clock)

Time-Delay Interferometry
reduction of laser noise

L1 3 TDI channels with 2 "~independents"

L2 Data Analysis of GWs

L3 Catalogs of GWs sources with their waveform

Gravitational wave sources emitting between 0.02mHz and 1 Hz

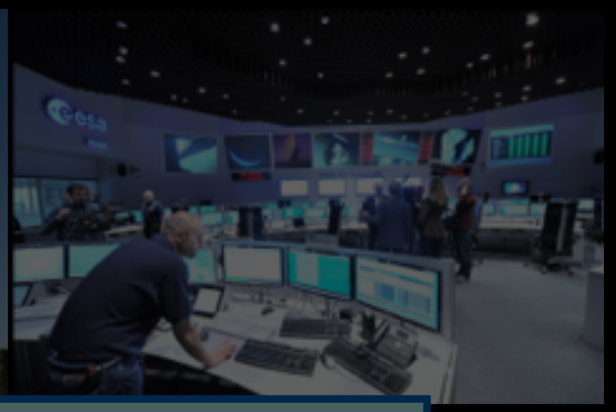
LISA data flow

Mission Operation Centre

Phasemeters (carrier, stance)

+ Gravitational Reference Sensor
+ Auxiliary channels

'Survey' type observatory



L0

Calibrations corrections

Resynchronisation (clock)

Time-Delay Interferometry
reduction of laser noise

L1

3 TDI channels with 2 "~independents"

Science Operation Centre

- 6 x10^7 galactic binaries
- 10-100/year SMBHBs
- 10-1000/year EMRIs
- larae number of Stellar Oriain

Distributed Data Processing Centre

- Unknown sources

L2

Data Analysis of GWs

L3

Catalogs of GWs sources
with their waveform