The Italian Gaia Data Processing Center and beyond: From Idea to product

INAF experience and the challenges of private-academic collaboration

Deborah Busonero
for the Italian Gaia DPC collaboration
INAF – Osservatorio Astrofisico di Torino

gaia DPCT
Data Processing Center Torino

ALTEC

Osservatorio Astrofisico di Torino – INAF Istituto Nazionale di Astrofisica

- The DPCT: one of the 6 data processing centers (DPC) of the Gaia SGS, hosted in ALTEC in Turin. Under dedicated ASI industrial and scientific contracts, its construction and operation is the result of the work of an integrated INAF-OATo / ALTEC team.
- DPCT provides the infrastructure support (in terms of HW, DB and software framework) to run the CU3 sw systems that are part of the AVU (Astrometric Verification Unit): AIM, BAM and GSR from raw data to catalogue data.
- Concept and design of the DPCT and of the AVU data reduction pipelines started 15 years ago.

OVERVIEW

- The INAF personnel involved in the activities mainly refer to INAF-OATo leader for the design and implementation of advanced calibration and data reduction techniques for wide-field or survey-oriented astrometry at the microarcsec level and beyond and global sphere recostruction, and INAF-OACT for HPC codes and exa-scale porting

LONG-TERM GOAL: The project aims to look beyond the operation of the Gaia mission, with the aim of providing to INAF and non-INAF community a Data Center with data analysis and reprocessing capabilities for astrophysics and space science and technologies .

- OPS4-TLS project towards the creation of one of the most advanced archiving, distribution, processing, analysis and exploitation systems for Big Data dedicated to the investigation of the near and distant Universe

Expression of interest for the National Center

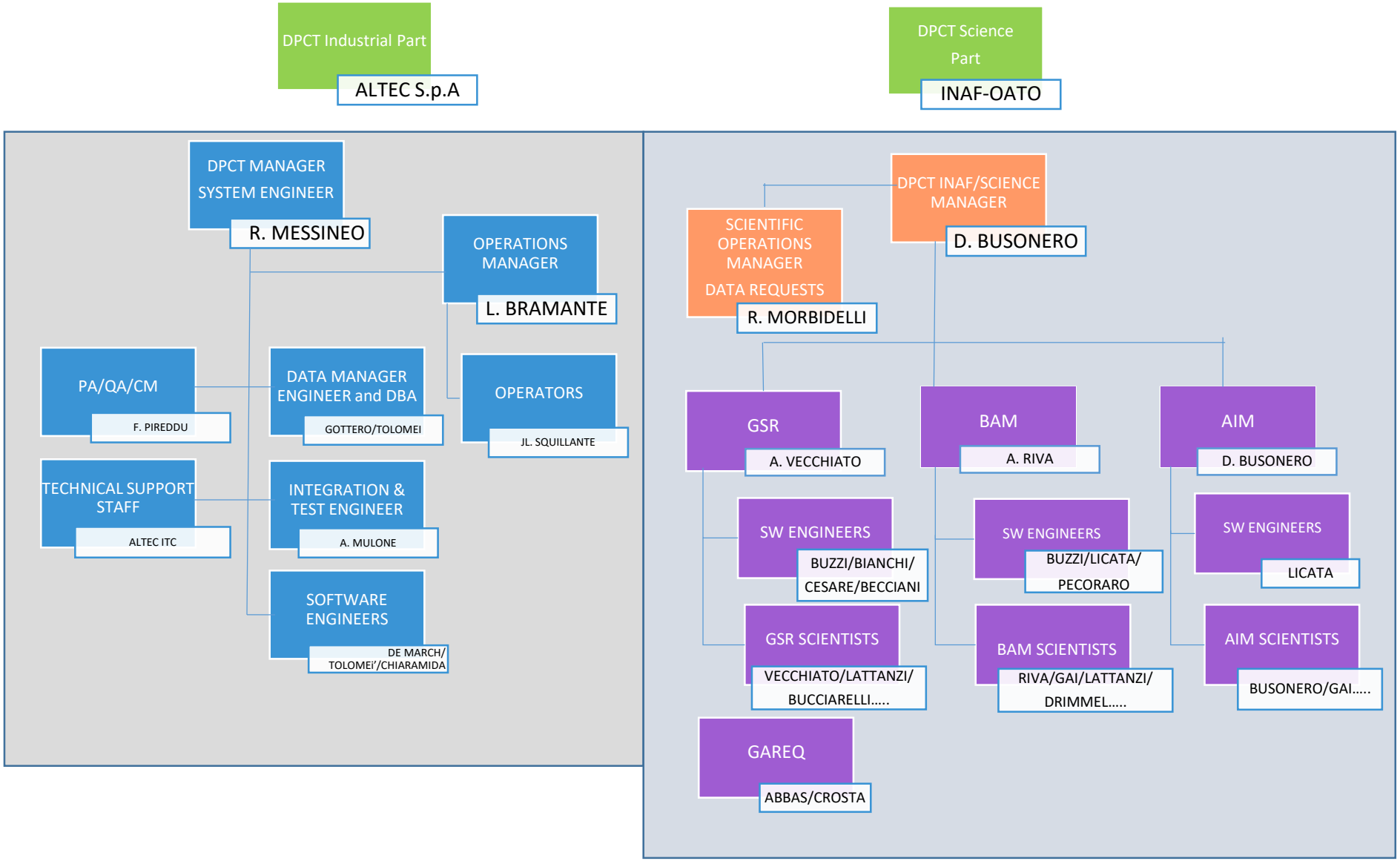
ORGANIZZAZIONE

INAF-OATo	definition, design, development, testing and validation of the 3 AVU systems and coordination of all scientific activities necessary for the development and operation of the DPCT
ALTEC	definition, design, implementation, testing, validation and operations of the HW and SW infrastructure of the DPCT, integration into the operating environment and operations of all the software provided. Coordination of the technical activities necessary for the operation of the Center.

Team skills

- Software system engineering,
- DataBase management
- creation, implementation and operability of data processing systems typical of a ground segment of a space mission,Quality e product assurance,
- data science and data analytics,
- porting su pre-Exascale e HPC code engineering.

INAF-OATo / ALTEC joint effort: THE TEAM



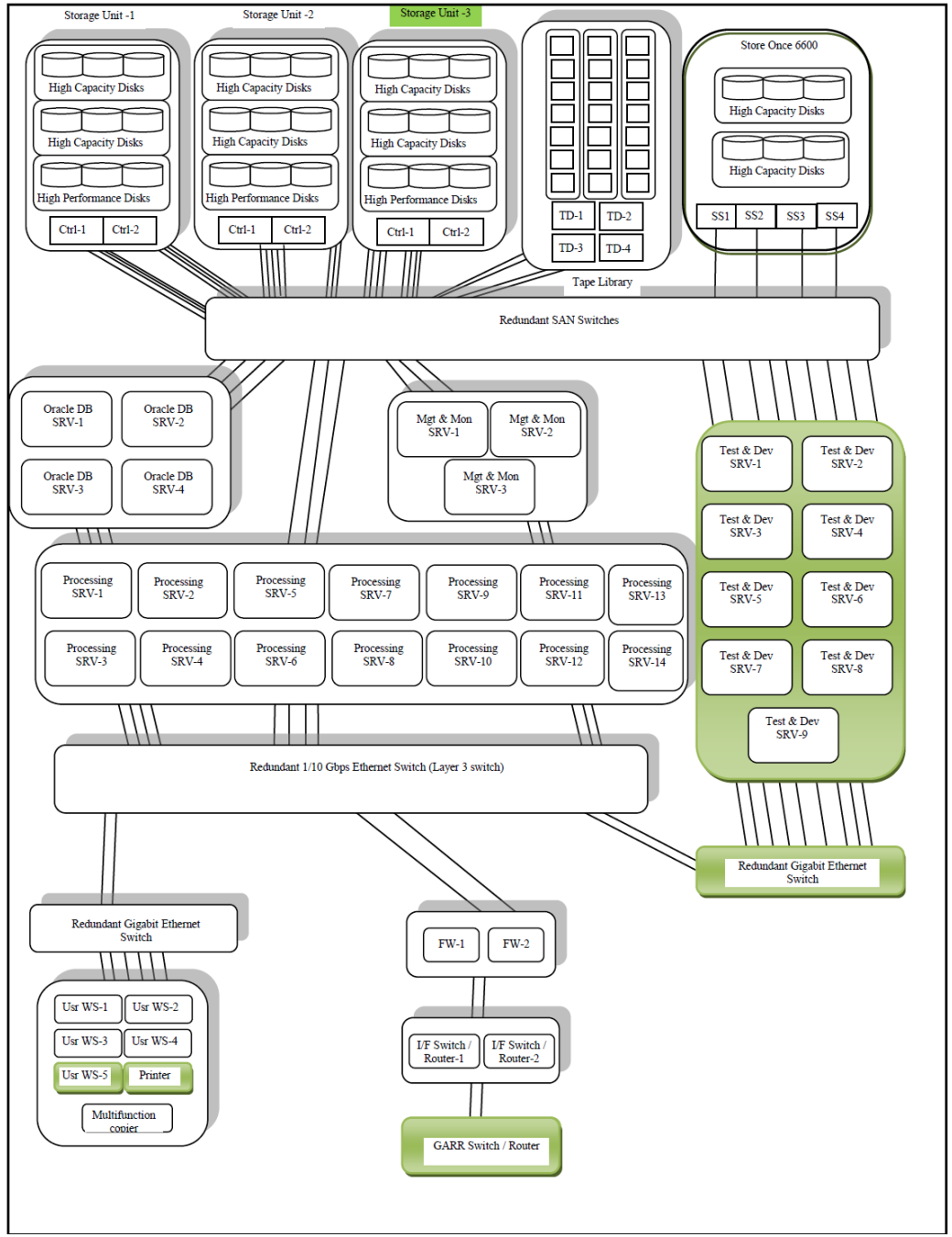
DPCT HW Infrastructure

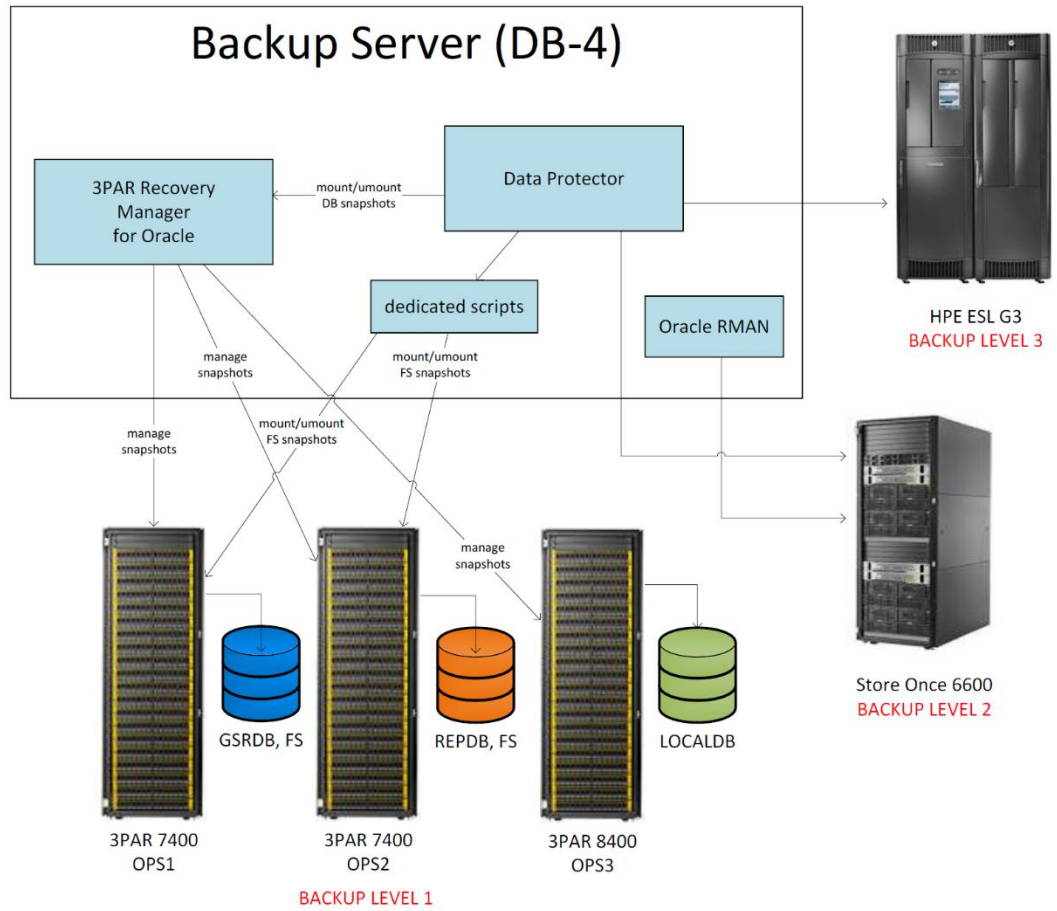
- Operational and test & development platform
- Procurement performed incrementally according to mission needs

INTERNET LINK : 1Gbps (300 Mbps guaranteed) via GARR
STORAGE CAPACITY: **2.5 PB overall raw disk space** distributed between two HP P7400 storage units and **one P8400**.
COMPUTING : **14 servers** HP DL580 G7/G9 with a total of about **600 CPU cores and 4.5TB RAM**.
DEV & TEST: 7 servers HP
DB SERVERS: **3 servers** HP DL580 G7 (**32 cores**, 256MB RAM each) based on Oracle RAC technology (**DBMS** Oracle).
NETWORK CONNECTION: LAN network up to 10 Gbps. SAN network redundant at 8 Gbps.
SECURITY SERVICE: redundant firewall based on pfSense, enabling secure remote access via VPN.
INFRA MONITORING AND MANAGEMENT: services based on VMWare virtual environment configured with two HP DL 580 G7 servers clustered and managed by vCenter Server.
BACKUP SERVERS: HP DL580 G7 dedicated to DB and filesystem backups from data volume snaphots.
3 LEVELS BACKUP : L1 on primary storage array, L2 on disks (StoreOnce 6600) and L3 on tape libraries (HP ESL G3).
HPC INTERCONNECTION: access to HPC super computer at CINECA for dedicated processing.

Direct link to CINECA

DPCT Overall HW infrastructure





Credits: ALTEC

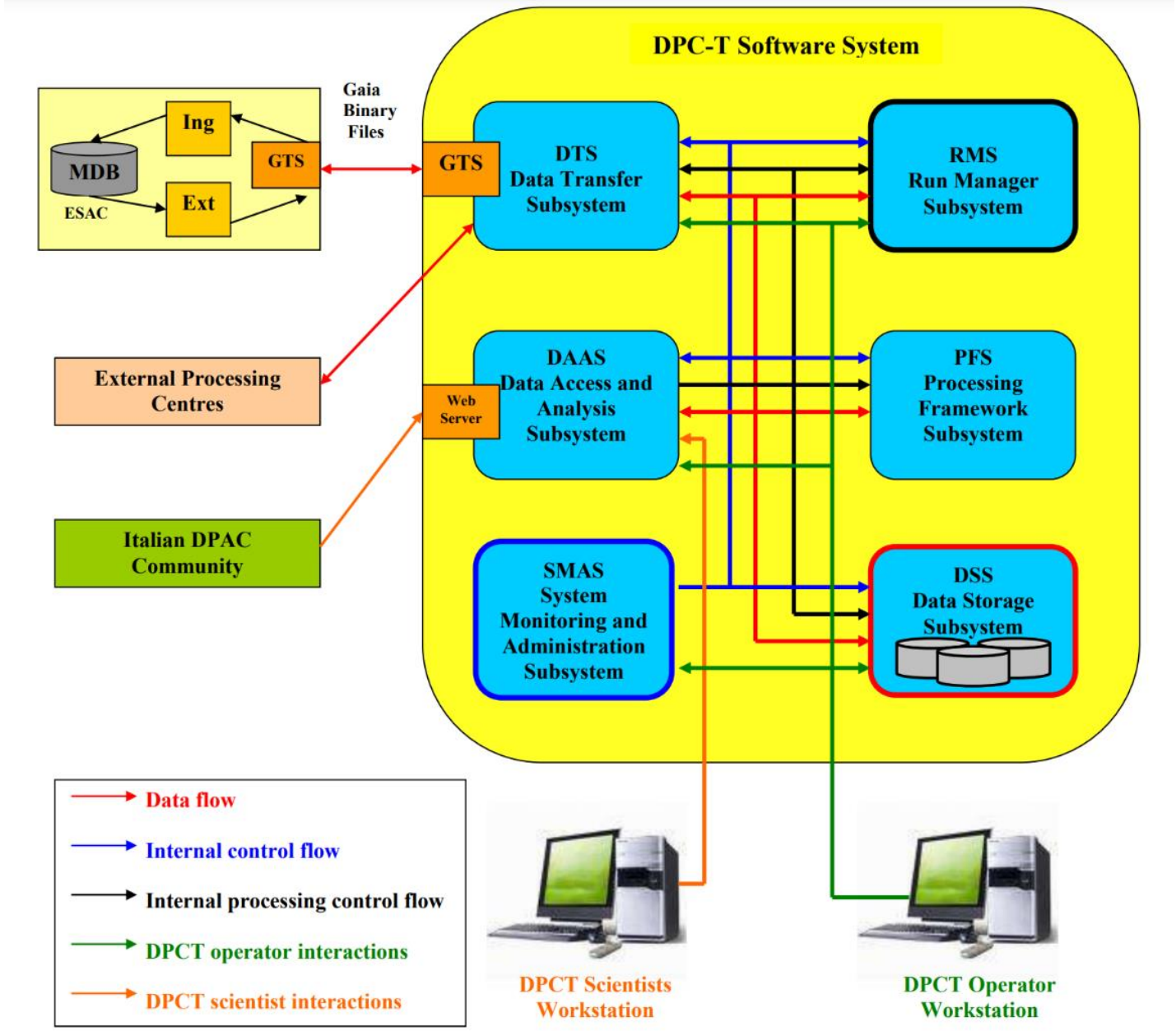
The DPCT is provided with an operations area of 55 m2, with access control and air conditions, in order to host the console operators and the DPCT scientists for their daily activity. All equipment in the server room and operations room shall be under Uninterruptable Power Supply system (UPS).

The DPCT hardware infrastructure is based on distributed environment, including a computational grid, a database grid and a storage area network.

The DBMS choice is Oracle that provides advanced availability and scalability features. Oracle allows multiple computers to run the Oracle DBMS software simultaneously while accessing a single database, thus providing a clustered database. The database grid will use the following Oracle products:

- Oracle Server
- Oracle RAC
- Oracle Partitioning
- Oracle ASM to manage storage used by database

DPCT SW Infrastructure



Heterogeneous and complex software system, composed by many subsystems structured in an open architecture in which each subsystem is a peer responsible for a different class of services.

DPCT SW Infrastructure

DTS: The Data Transfer Subsystem manages the regular data transfer with ESAC DPC. This subsystem interfaces the GTS for all required data exchange.

DSS: it provides the capabilities to access and use the DPCT storage in an efficient manner

EMS is in charge of managing messages exchanged among subsystems and notify users with alert on subsystems status (including itself)

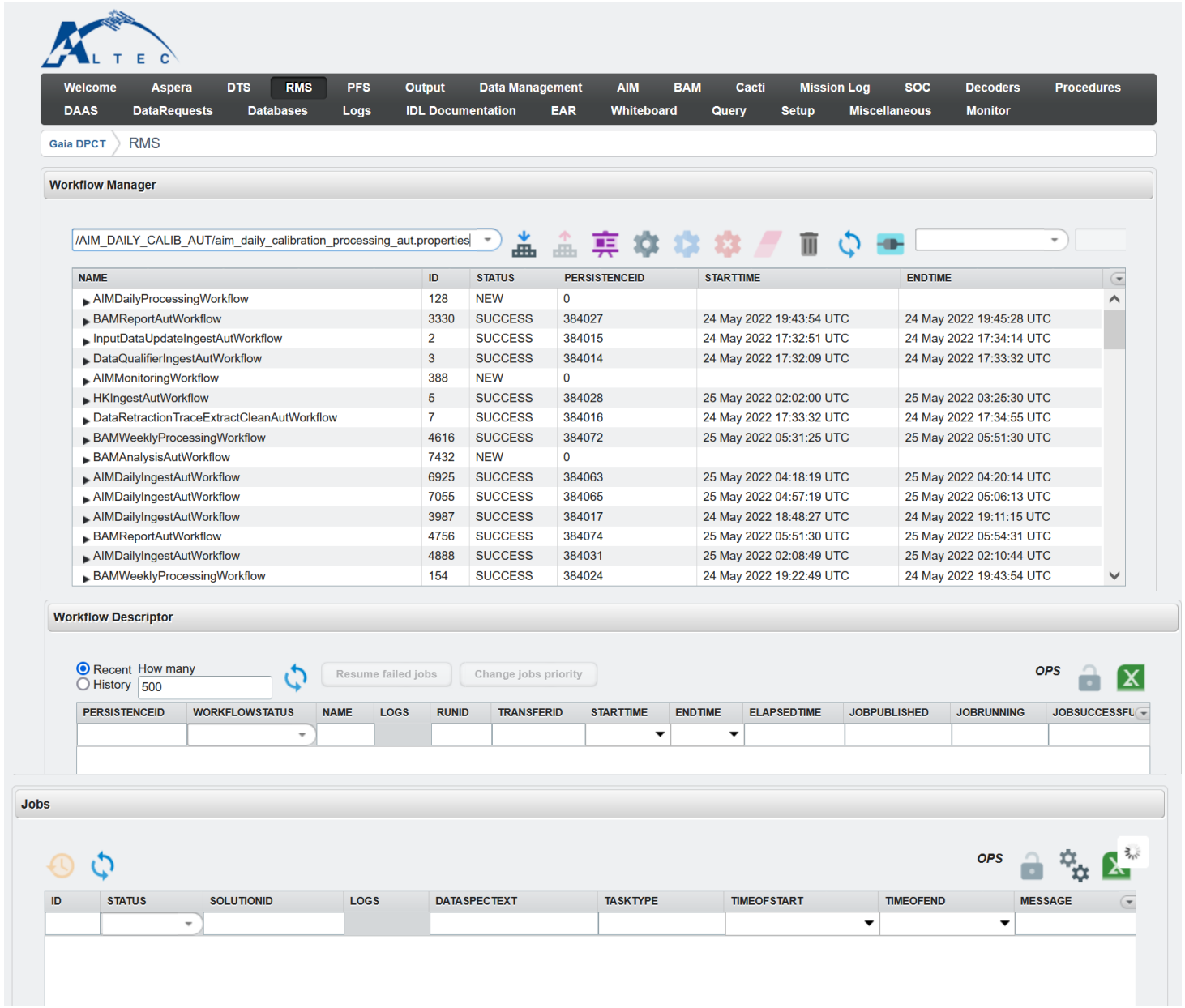
DAAS: It provides to scientists the required capabilities to make use of data and analyse them also outside the data reduction pipeline.

SMAS: it is responsible to administer and monitor the whole system, including hardware and COTS components. Also the authentication and authorization mechanism shall be provided by SMAS.

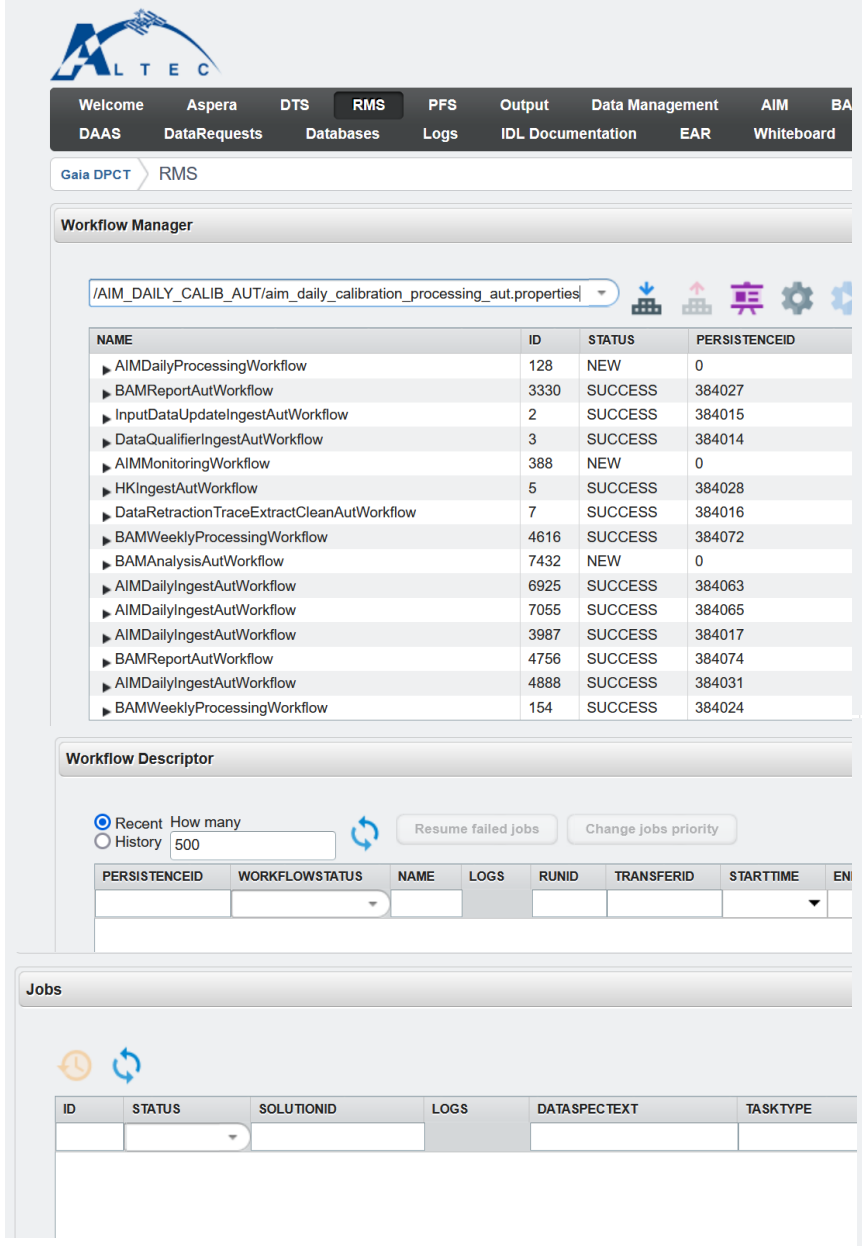
RMS: The brain of the sw infrastructure. It provides the control logic to coordinate and manage the software execution infrastructure, represented by the PFS subsystem. Manager, workflow, step, task.

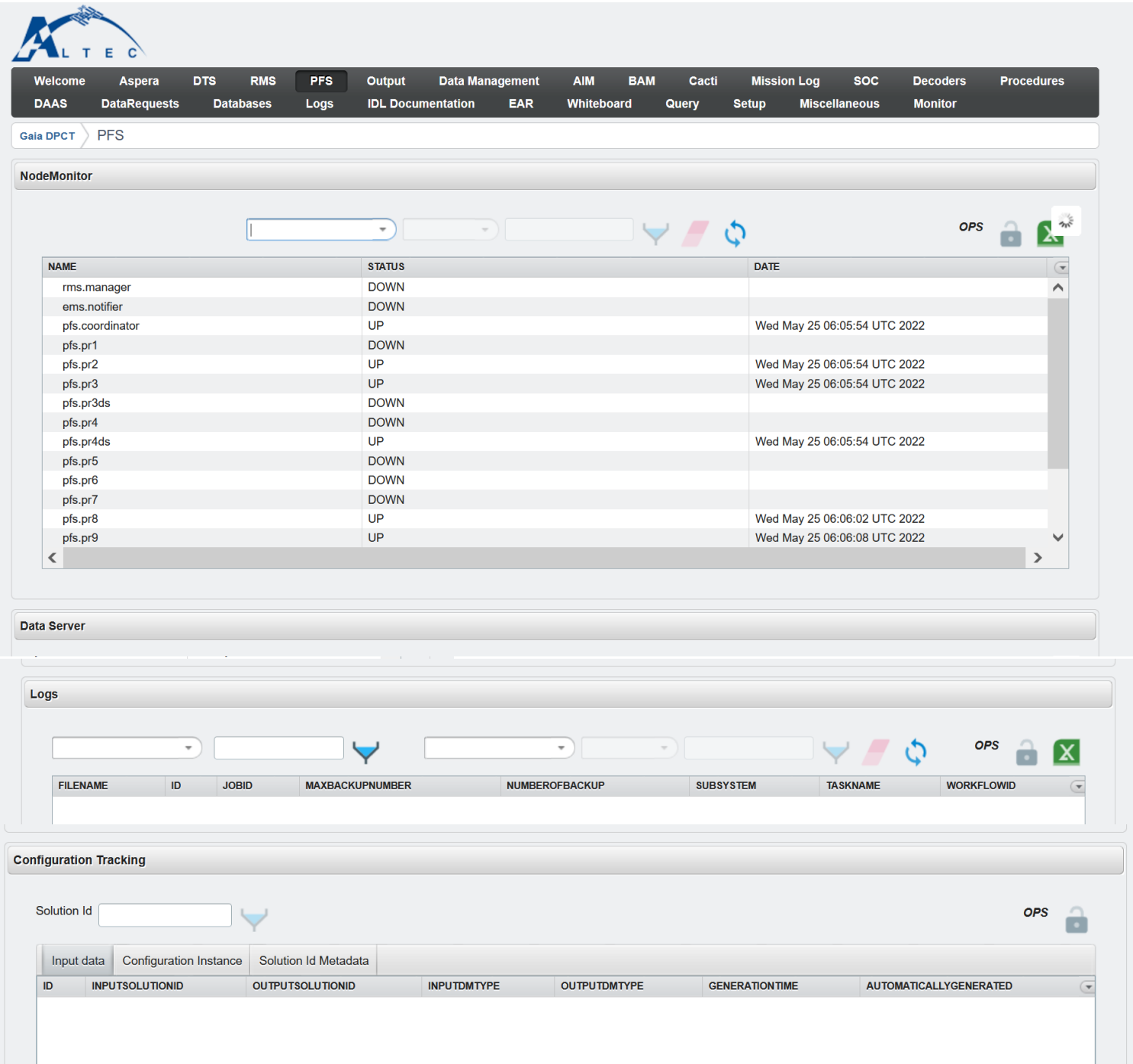
PFS: The processing core of the DPCT. It is in charge of running and controlling the execution of all defined jobs based on the input provided by the RMS. The PFS provides also the interfaces to run and control execution of scientific jobs developed compliant with Gaia Infrastructure Framework interfaces.

DPCT WEB INTERFANCE

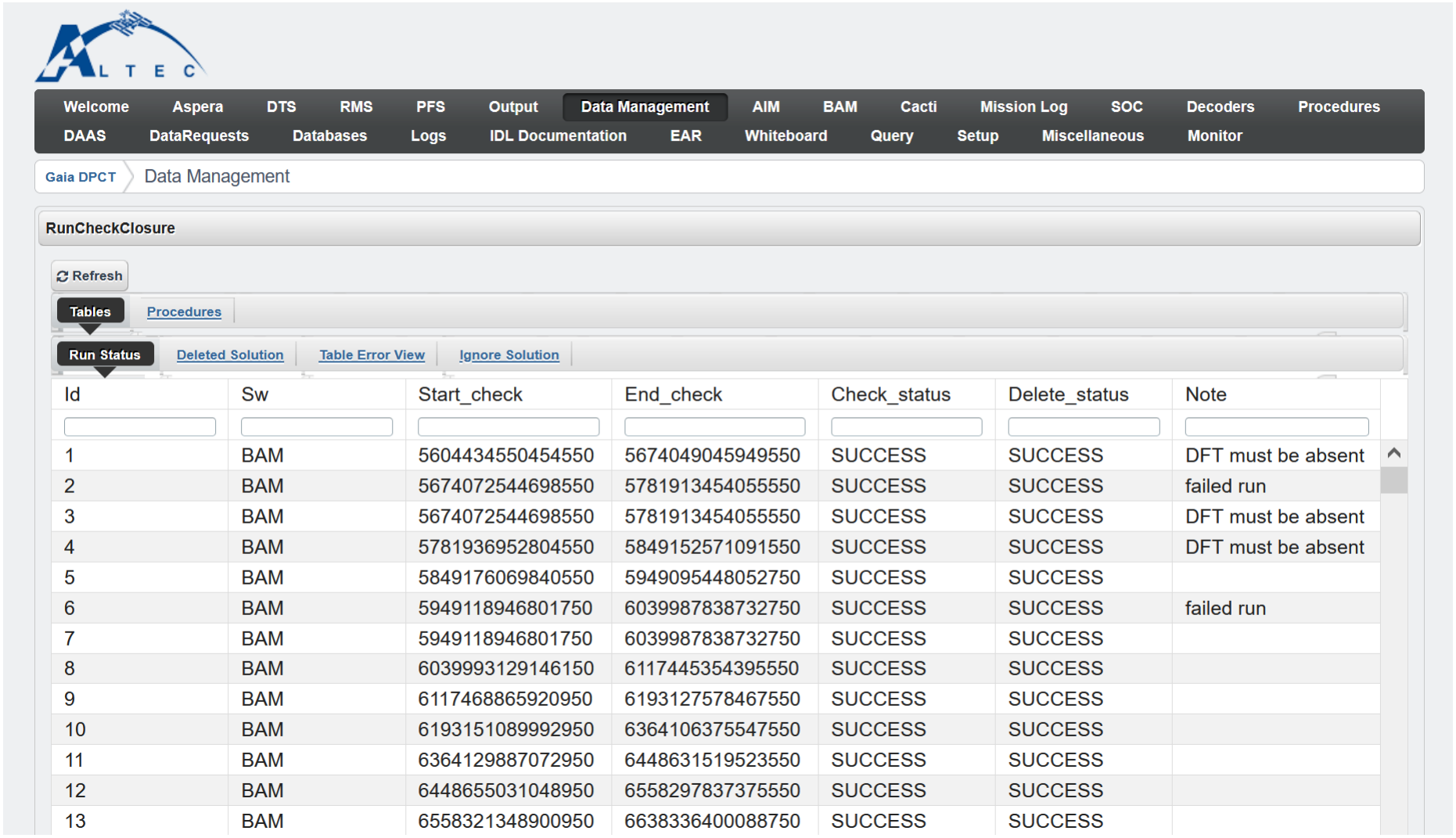


DPCT WEB INTERFANCE



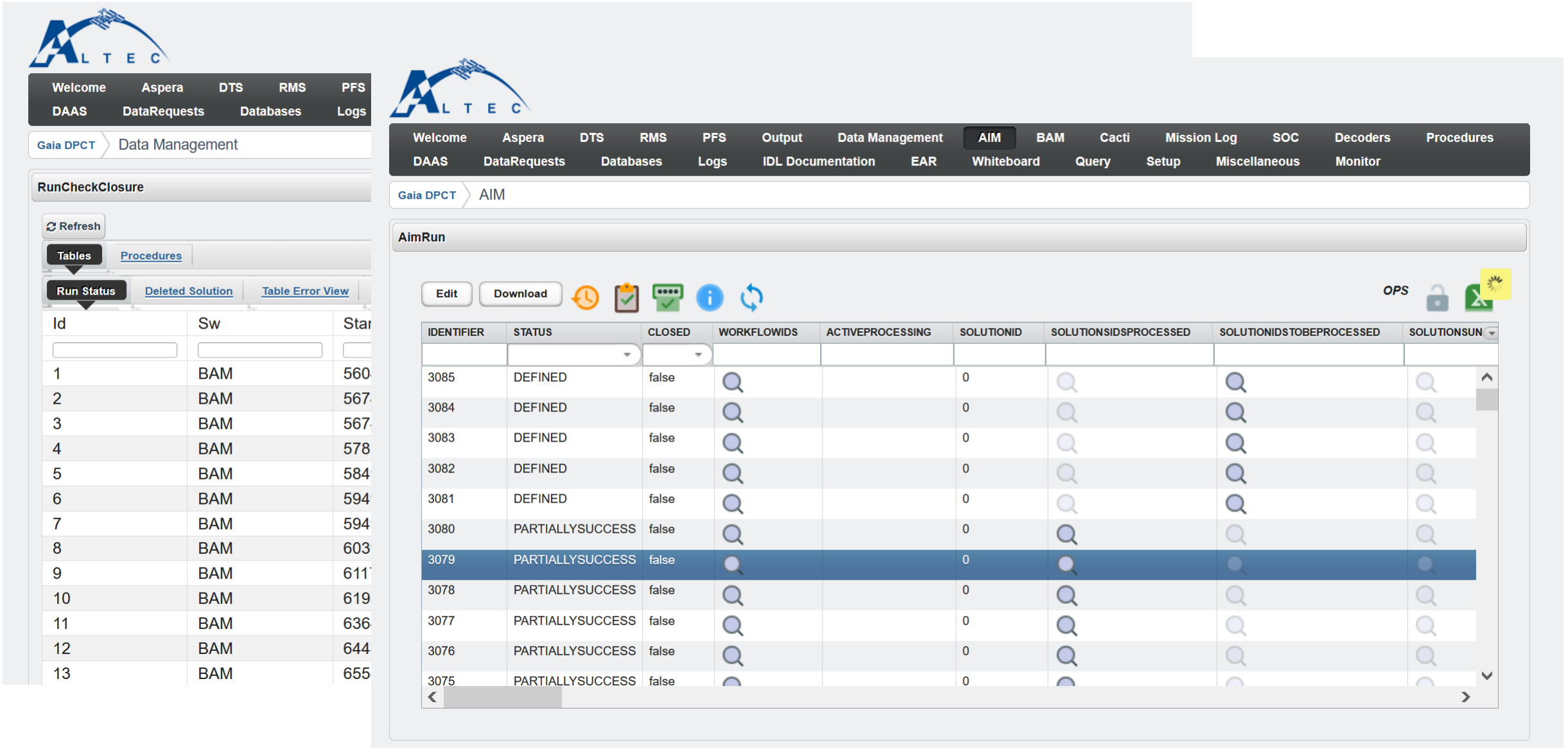


DPCT WEB INTERFANCE

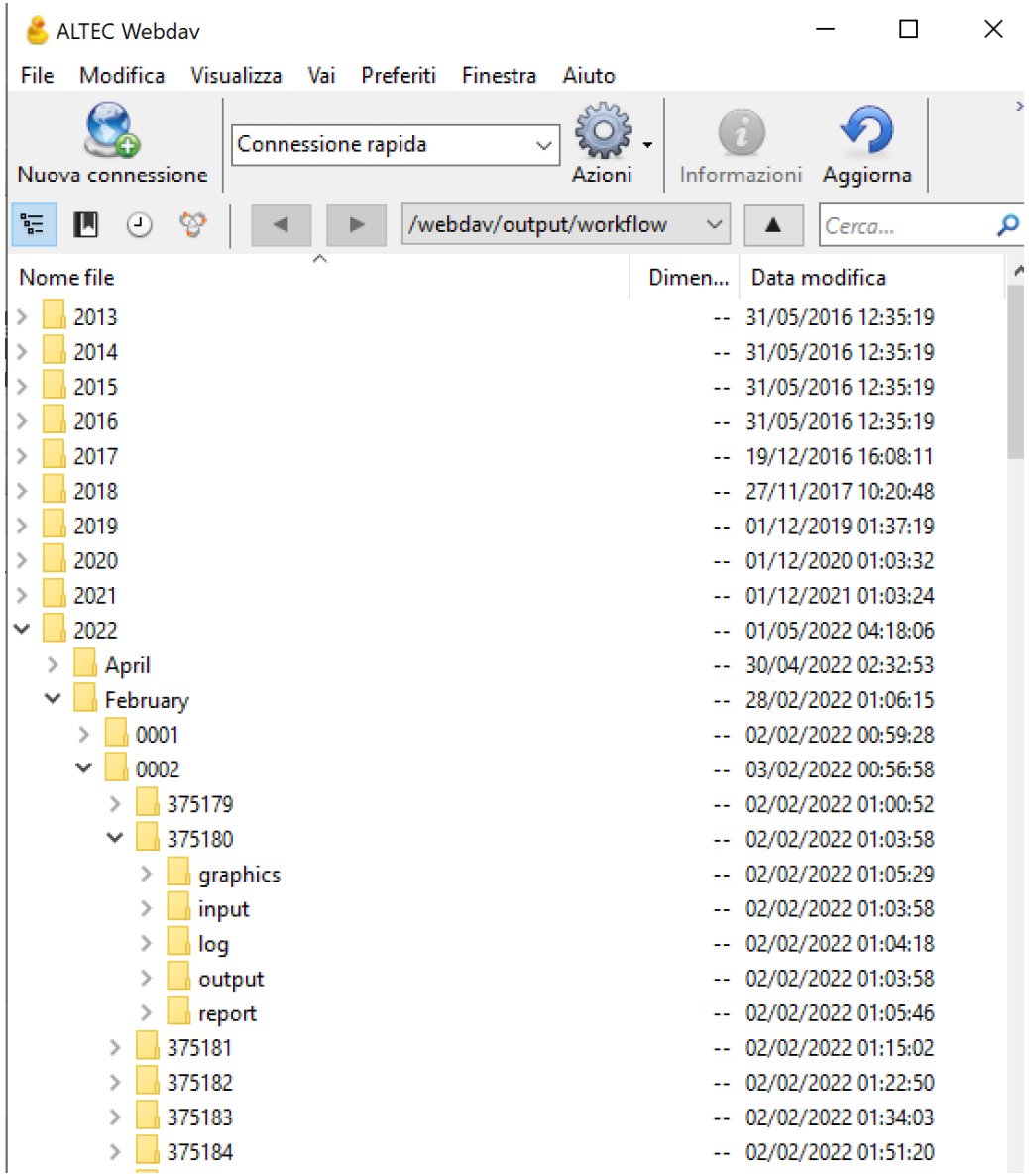


Id	Sw	Start_check	End_check	Check_status	Delete_status	Note
1	BAM	5604434550454550	5674049045949550	SUCCESS	SUCCESS	DFT must be absent
2	BAM	5674072544698550	5781913454055550	SUCCESS	SUCCESS	failed run
3	BAM	5674072544698550	5781913454055550	SUCCESS	SUCCESS	DFT must be absent
4	BAM	5781936952804550	5849152571091550	SUCCESS	SUCCESS	DFT must be absent
5	BAM	5849176069840550	5949095448052750	SUCCESS	SUCCESS	
6	BAM	5949118946801750	6039987838732750	SUCCESS	SUCCESS	failed run
7	BAM	5949118946801750	6039987838732750	SUCCESS	SUCCESS	
8	BAM	6039993129146150	6117445354395550	SUCCESS	SUCCESS	
9	BAM	6117468865920950	6193127578467550	SUCCESS	SUCCESS	
10	BAM	6193151089992950	6364106375547550	SUCCESS	SUCCESS	
11	BAM	6364129887072950	6448631519523550	SUCCESS	SUCCESS	
12	BAM	6448655031048950	6558297837375550	SUCCESS	SUCCESS	
13	BAM	6558321348900950	6638336400088750	SUCCESS	SUCCESS	

DPCT WEB INTERFANCE

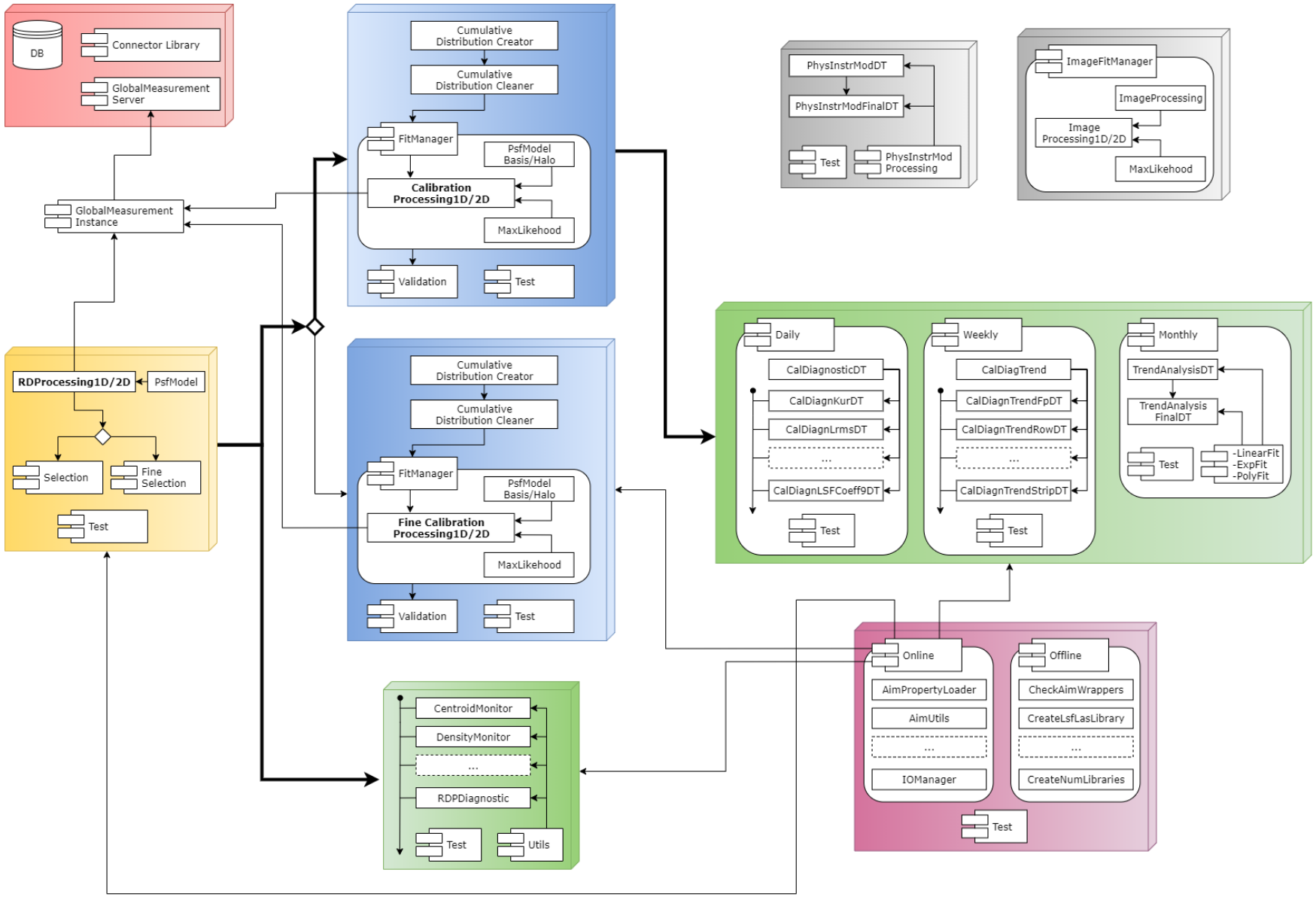


FILES MANAGER



Accesso da remoto tramite VPN sulle WS; utilizzo di X2go

HIGH LEVEL ACTIVITY DIAGRAM of AIM PIPELINE:



AIM PIPELINE

AIM daily pipeline:

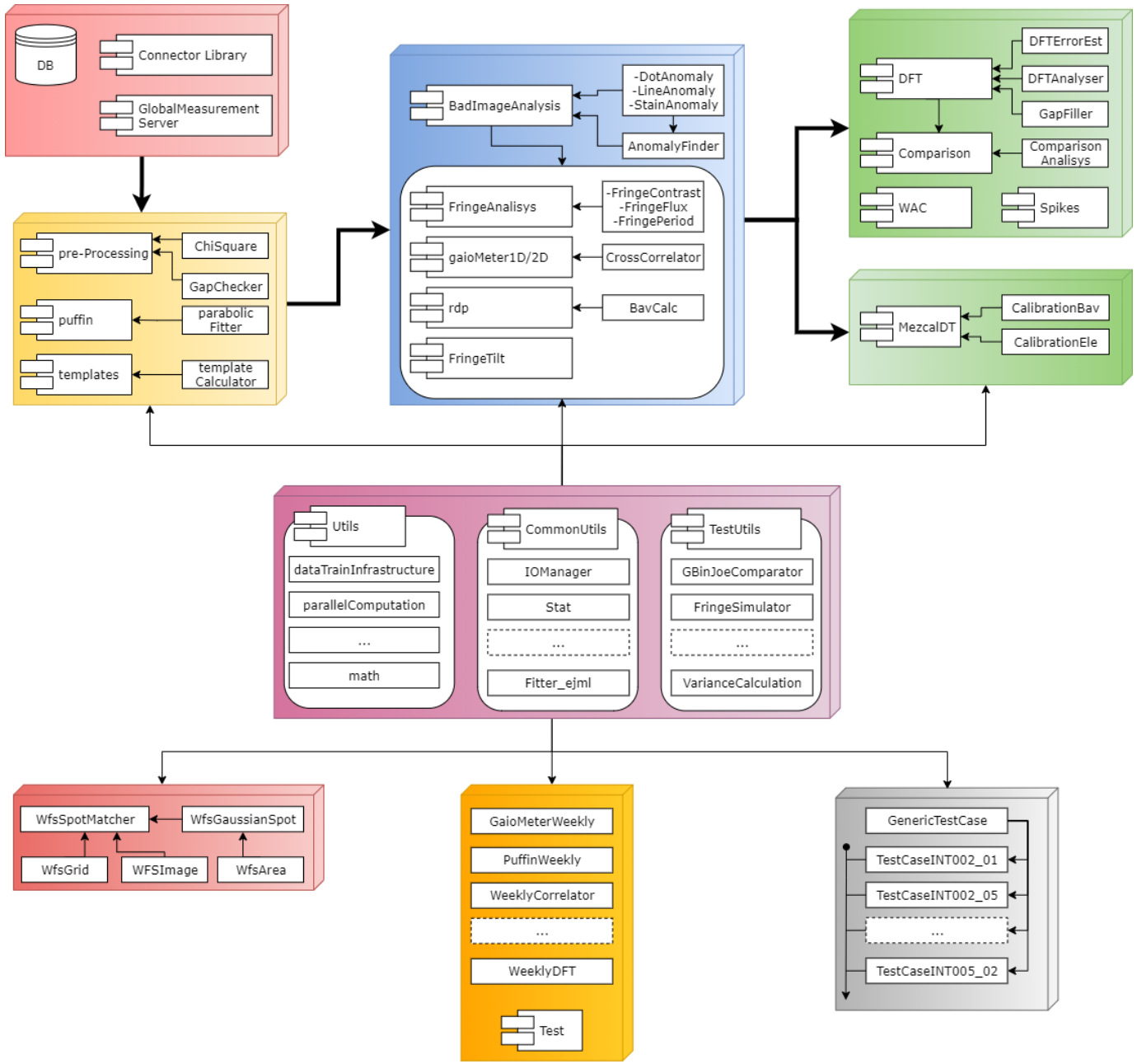
raw data processing, image parameter determination, LSF/PSF modelization and calibration, astrometric instrument monitoring and diagnostics throughout the mission lifetime

- Daily sw version 100.000 lines of code
- Cyclic sw version about 100.000 lines of code
- 24 hours of raw data each run: from $2x10^6$ to $15x10^6$ raw images
- Complex structure of the pipeline:

10 sw modules managed by a coordinator in an automatic way, the output of one run become the input of the next one

- 6 hours of time execution on the DPCT Operation platform for each run
- AIM cyclic version of the software aims to image reprocessing for calibration improvement and updated calculation

HIGH LEVEL ACTIVITY DIAGRAM of BAM PIPELINE:



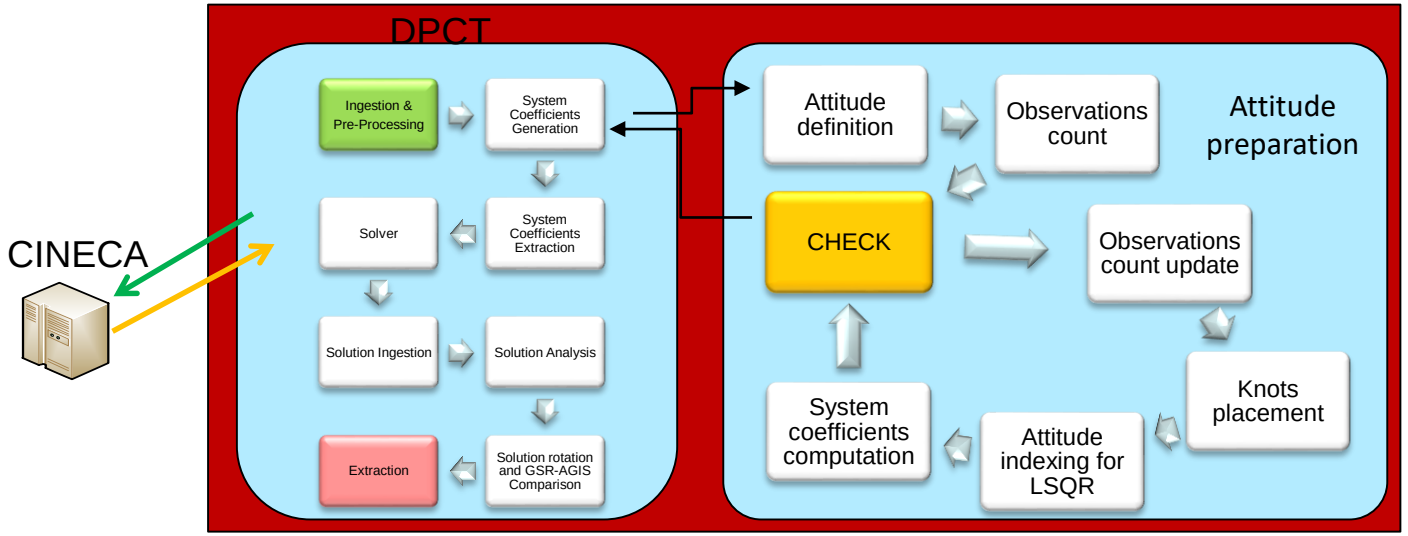
BAM PIPELINE

AVU/BAM daily pipeline:

raw data coming from the Basic Angle Monitoring (BAM) instrument, i.e. fringes, for monitoring and analyzing the instrument behaviour thoughout the mission and performing the BAV calibration.

- Daily sw version 100.000 lines of code
- Cyclic sw version circa 50.000 lines of code
- 24 hours of raw data each run: almost 8x104 images
- 1-2 hours for each run
- The pipeline output is sent to DPCE and ingested into the MDB
- AVU/BAM cyclic version of the software aims to fringes reprocessing for calibration improvement

HIGH LEVEL DIAGRAM OF GSR PIPELINE



The Global Sphere Reconstruction (GSR) solves a linearized system of equations who's result gives the global astrometric reference system (position, parallax, proper motions).

Starting from 10^7 to 10^8 objects for each run

- Very complex pipeline structure
- **130.000 lines of code in Java + 30.000 in C, C++ for the Solver running at CINECA**
- Final GSR output sent to DPCE in the MDB
- The Solver module run at CINECA which is managed as one processing node of the DPCT
- The whole process could be iterated for Non-Linearity
- **One run takes from 3 to 6 days** on 10^7 objects.

JUST A FLAVOUR OF THE OPERATIONAL NUMBERS.....

AVU DPCT activities since the Commissioning phase:

~ 350.000 workflows
Most than 40.000.000 processing jobs
~50 GB Gaia telemetry data each day (SOC),
~16 GB to DPCT each day,
~ 114.5 x 10^9 di Osservazioni ricevute,
3000 AVU/BAM runs,
3000 AIM runs,
24.000 Mission Log entries

GSR DRC started at the end of cycle 2 (2017) with 74 processing runs

The overall DB tablespace allocation is about **614** TB:

~400 TB nel database Repository (RDB),
~110 TB nel LOCAL database (LDB),
~104 TB nel GSR database (GSRDB).

With a total data amount of **1.5 PB**

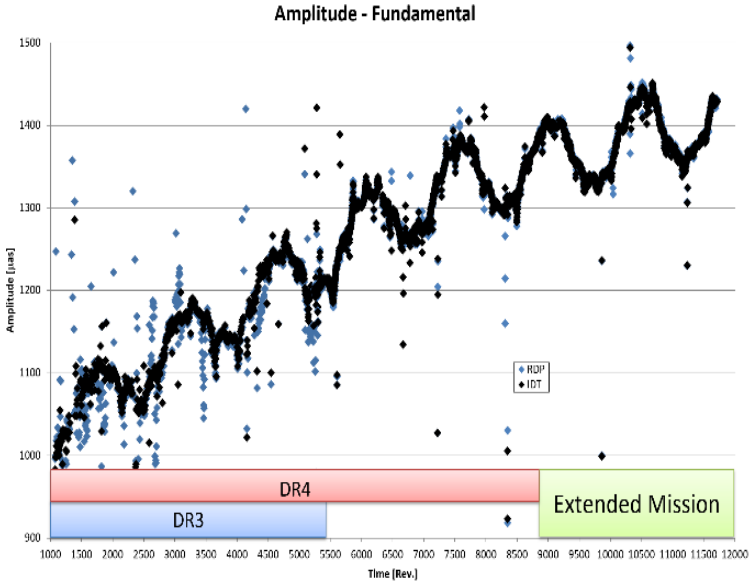
Science products:

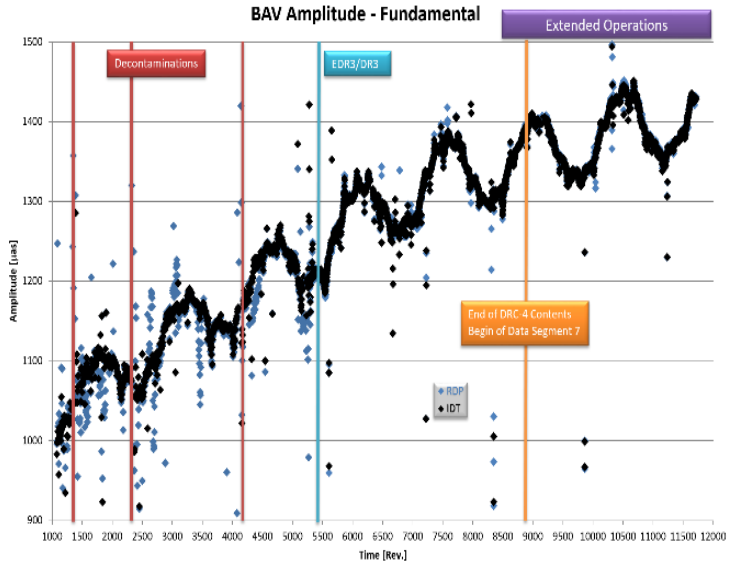
AIM e BAM attivi fin dalla prima fase di Commissioning;

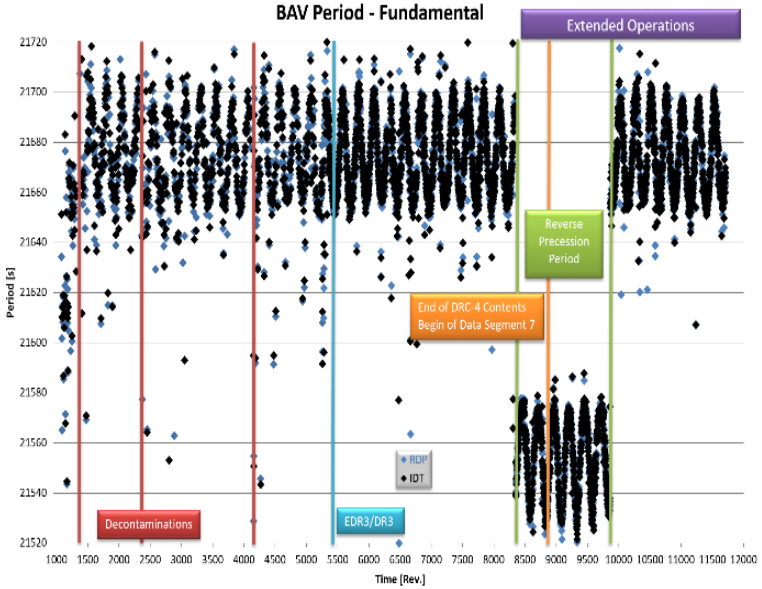
Output utilizzati dai Payload Expert italiani per l'individuazione e risoluzione delle problematiche riscontrate sul satellite.

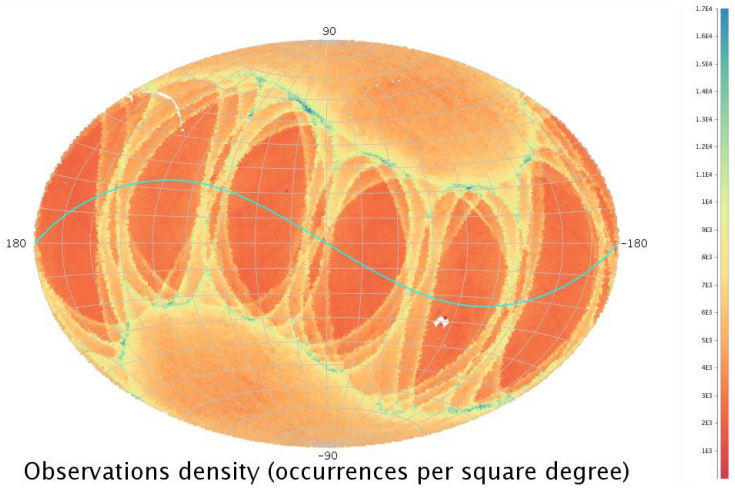
GSR attivo dalla fine del ciclo 2

- 5500 rapporti tecnico-scientifici giornalieri,
- 100 rapporti tecnico-scientifici settimanali,
- 100 documenti ECSS quali SRN, STR, SRS, STS, SDD, SUM, ICD, tra cui il DPCT development plan, DPCT operations plan, DPCT Procedure Handbook DPCT design document
- una ventina di note tecniche DPAC descriventi le soluzioni architetturali e algoritmiche una ventina di pubblicazioni referate e non (SPIE) specifiche reperibili su ADS, oltre alla documentazione accompagnante le 3 data release avvenute fino ad oggi (GDR1, GDR2, GEDR3).
- A queste si aggiungono gli articoli della Gaia Collaboration pubblicati in occasione del rilascio di ogni Data Release.
- Per il progetto TLS sono state prodotte 3 note tecniche INAF, mentre per l'espansione OPS4 è stata prodotta a inizio 2021 una nota tecnica dedicata

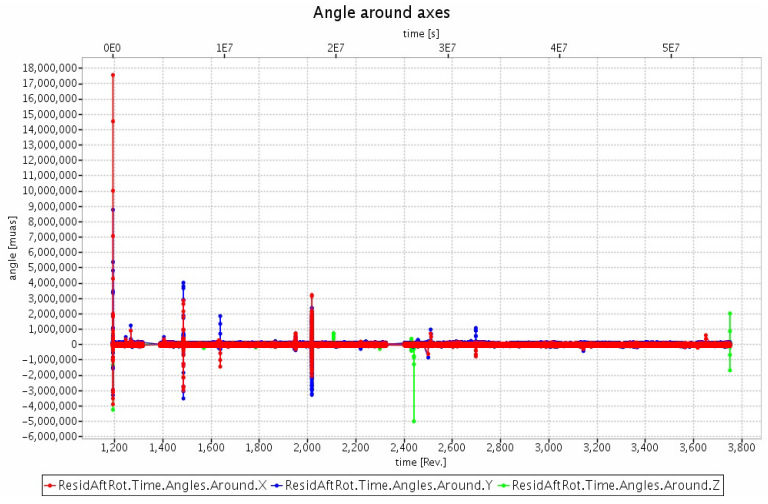


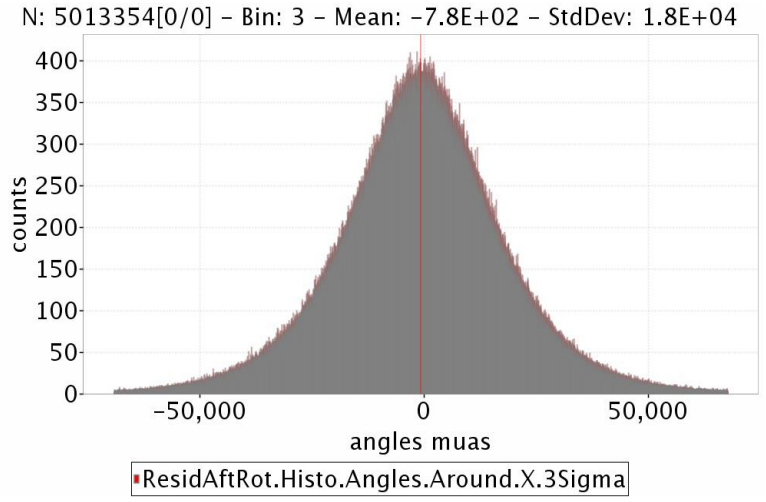




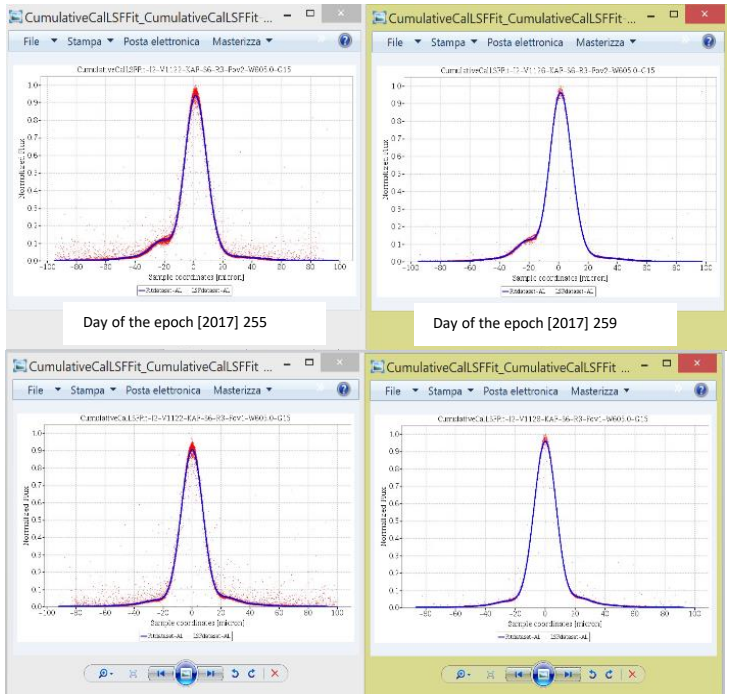


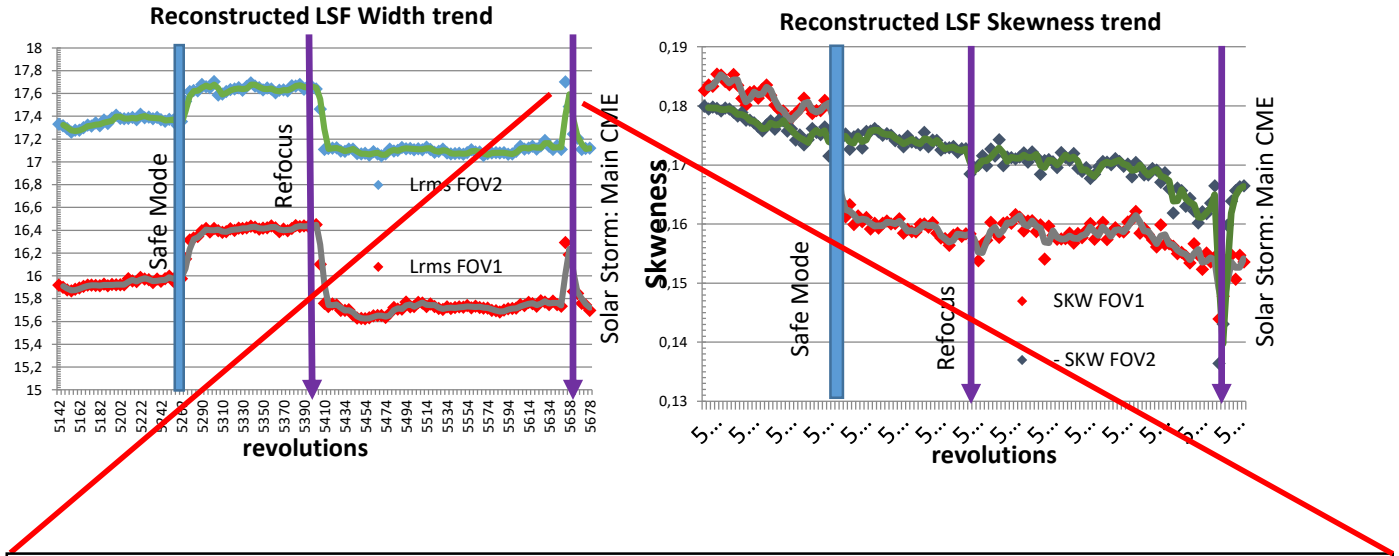
Observations density (occurrences per square degree)

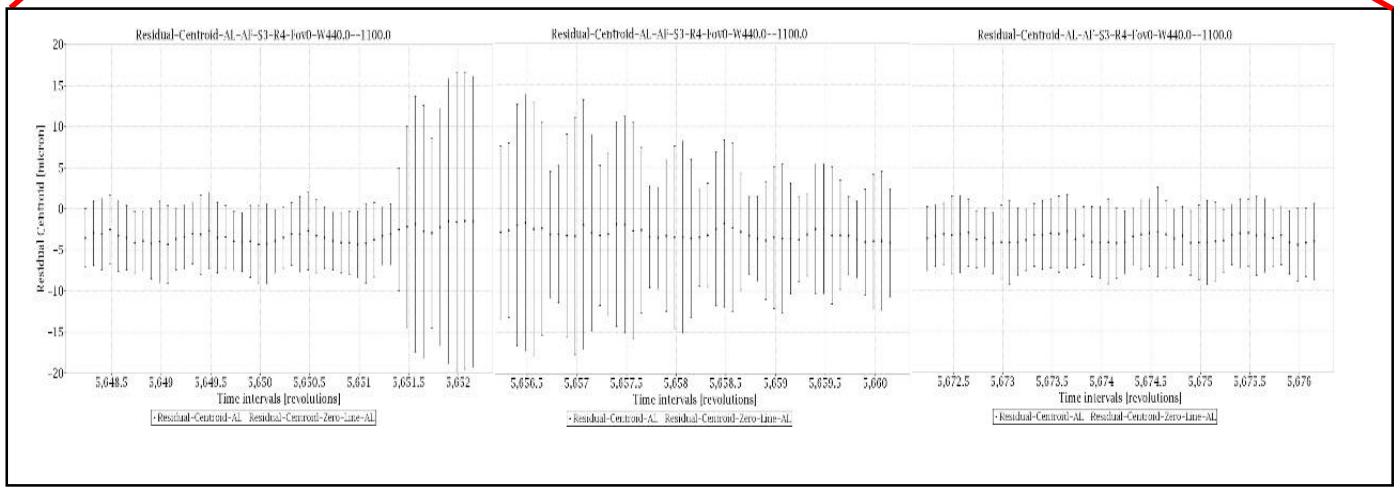




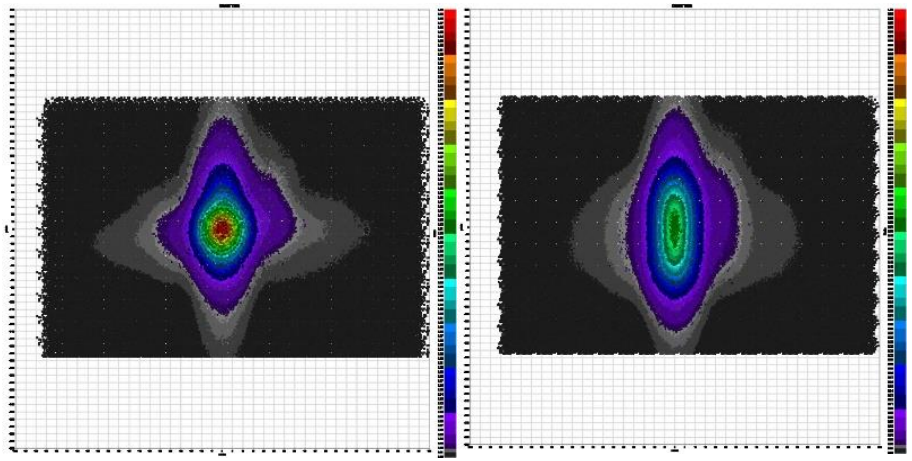
Solar flares effects on the gaia instrumentall response







Gaia PSF reconstruction



BUT…..

Paradigm change need for data management and exploitation.
Identification of an adequate data model for the management of the various use cases of interest both to the astrophysical community and not.

3 POC (Proof of Concept) derives from the intention to conceive, design and implement investigation methods that can be traced back to the formulation of a scientific requirement and to the measures related to it, aimed at validating it in a "usable" and “smart” way.
A system that, at its core, is not dependent on the development of ad hoc s/w, but rather based on the proximity of the methods of "navigation" of the data extracted from the mission database, organized in explicit types (no blobs are allowed , raw or other non-explicit encodings)

DPCT-OPS4 towards the Legacy?

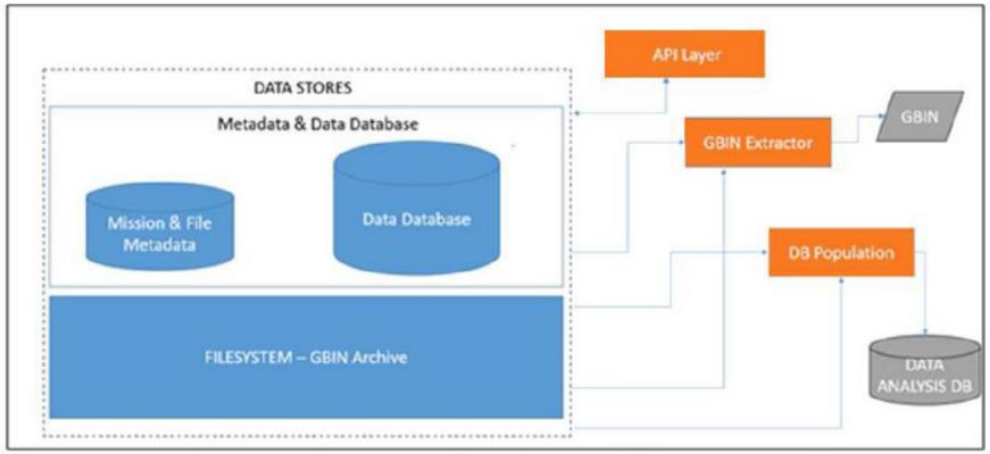


Figure 2. High-level (logical) design of the new OPS4@DPCT (see subsec. 4.1. for viable implementation options).

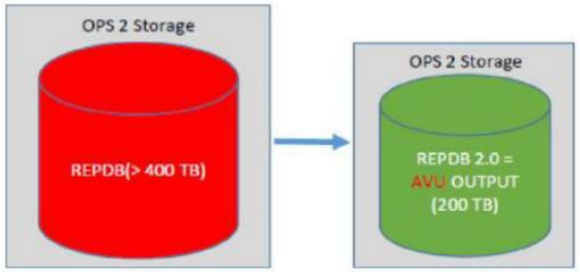
Figure 3. Sizing of OPS4 Mission DBs and file system.

INPUT: Mission & File MetaData and Data DBs	
FS DAILY	180.0
FS DRC	261.4
FS + MISSION METADATA	5.0
DATA-SELECTION DAILY	52.9
DATA-SELECTION DRC	21.1
Tot TB	520.4

Table 1. Details of the data that led to the size estimations for the mission repositories in Fig. 4.

DATA ANALYSIS (*)	
DATA REQUEST	74.08
Tot TB	74.08

Figure 4. OPS4 Data Analysis DB details. (*) As estimated from verification requests/queries received by AVU scientists.



Initial REPDB on OPS2 Final REPDB on OPS2 (REPDB 2.0)

Figure 5. OPS2 hosting REPDB2.0 after the one-time migration to OPS4.

OUTPUT: REPDB2.0	
DB DAILY OUTPUT	180.0
DB DRC OUTPUT	20.0
Tot TB	200.0

Table 2. Details of the REPDB2.0 size after current REPDB data migration to OPS4.

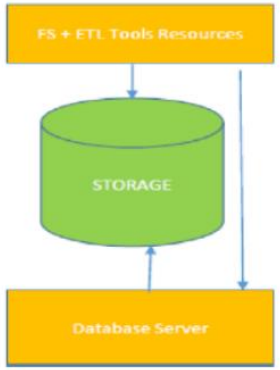


Figure 7. Summary schematics of the OPS4 system: storage, DB servers and Data Analysis servers.

ESPERIENZA E CONCLUSIONI

- Esperienza della collaborazione con l'industria di base positive soprattutto se si mantiene il giusto equilibrio nel rapport con l'Azienda definendo I campi di operatività corretti.

- **Adeguati profili interni INAF** per garantire la gestione e lo sfruttamento scientifico e tecnologico di enormi moli di dati quali quelle di un Data Base tipo Gaia che è unico attualmente nel suo genere per complessità.

- I Centri di processamento dati da considerarsi infrastrutture INAF, visto l'interesse scientifico e tecnologico della comunità tutta nello sfruttamento dei dati, ben oltre la conclusione della fase operativa di una missione. Per poter garantire ciò e non aver speso sforzi invano dobbiamo **coordinare in modo più organico i gruppi in INAF con tale expertise e pensare a un piano a lungo termine sia per lo sviluppo di tali centri di calcolo di science ground segment per missioni spaziali, che per il loro mantenimento e sfruttamento.**