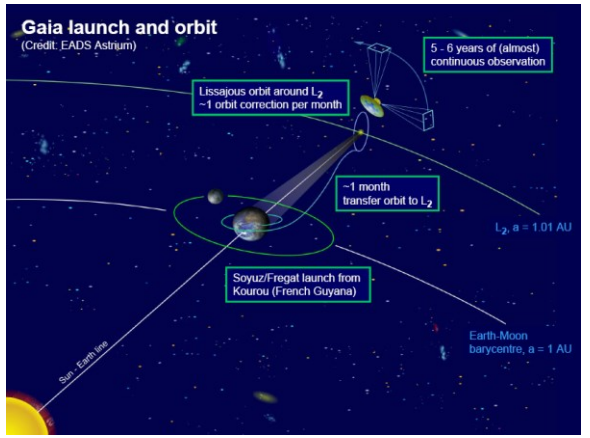
Risorse/Necessità HPC della partecipazione Italiana alla missione Gaia:

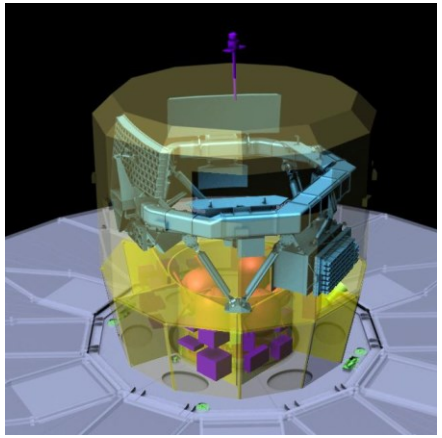
Mario G. Lattanzi (INAF-OATo)

Con li contributo essenziale di Becciani (OACt) e Vecchiato (OATo) per l'HPC, Busonero, Morbidelli, Riva (OATo) e i colleghi del DPCT@ALTEC per gli aspetti di HPC/HTC al DPCT (DB e pipeline processing)

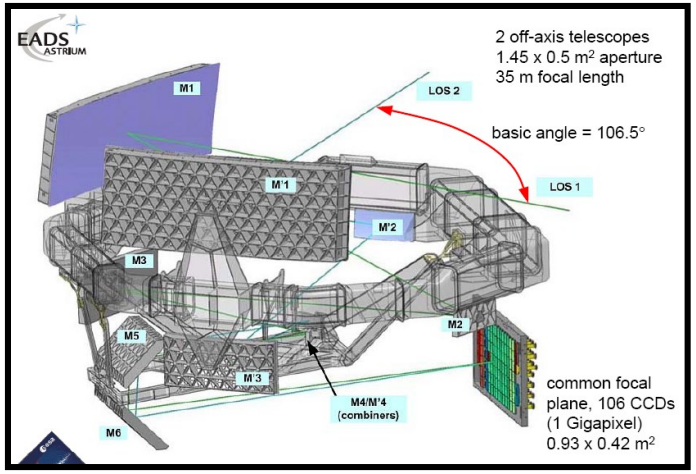
❑ The Gaia mission is unfolding!

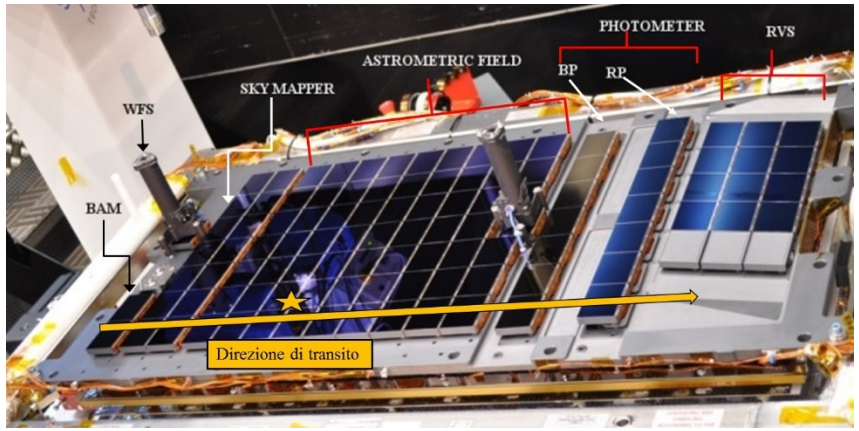
19 Dec 2013: Soyuz-Fregat launcher from CLG, French Guiana



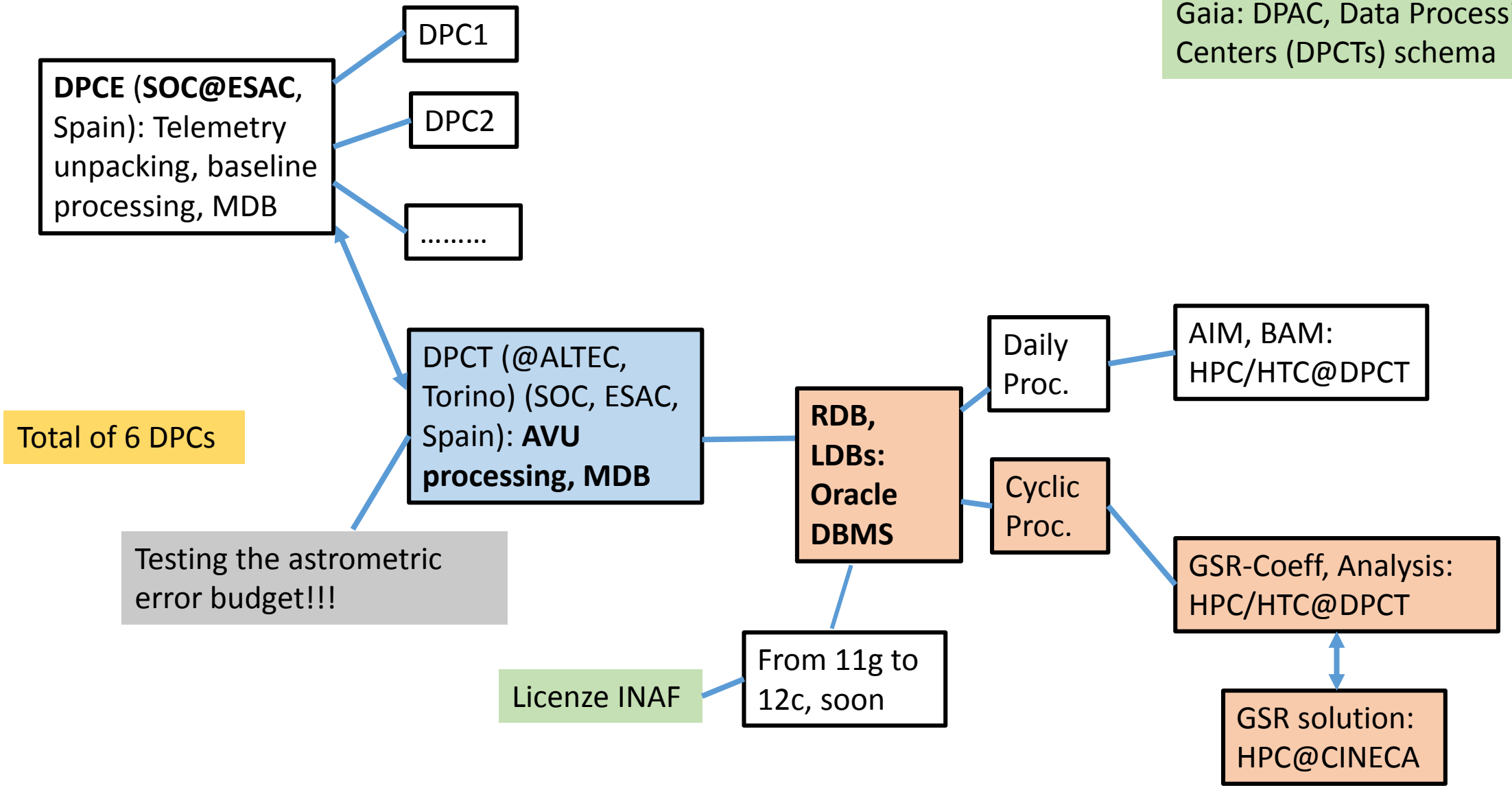


914mo giorno dal lancio!

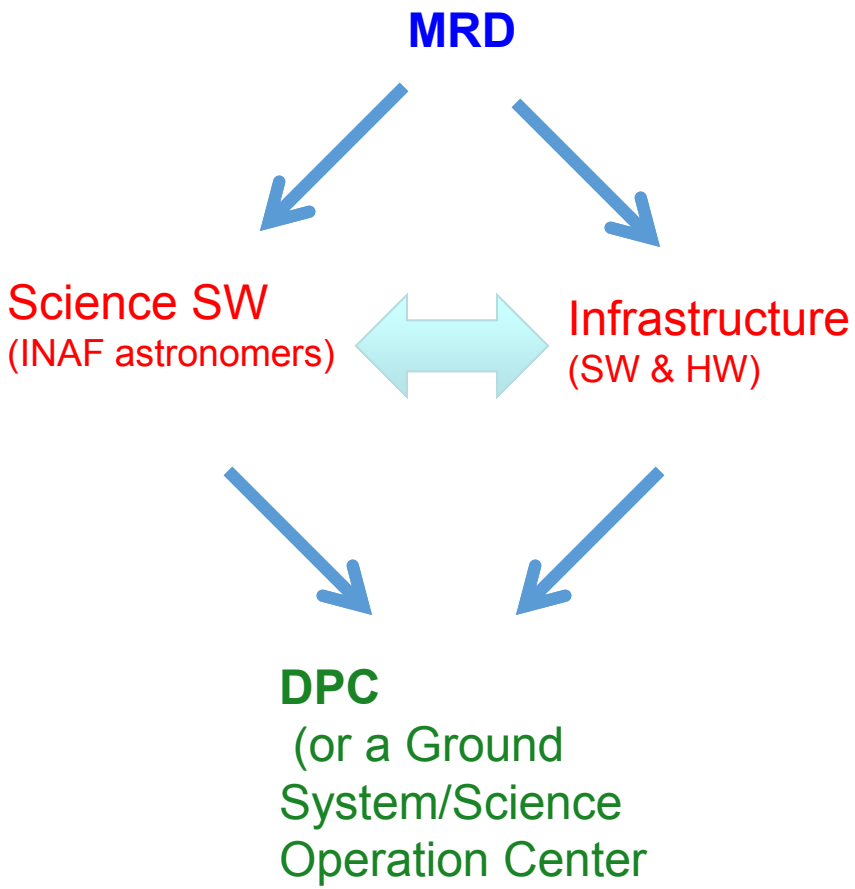




Gaia: DPAC, Data Processing Centers (DPCTs) schema



Gaia mission DP&A paradigma:



It is at this level that the INAF scientists of the night sky are required to have multidisciplinary abilities:

>> first within ourselves (INAF), then

>> must talk the language of ICT colleagues and that of SW and system engineers.

BUILD SYNERGYES

TECHNICAL REPORT

TITLE: **IMPLEMENTATION PLAN FOR THE GAIA ITALIAN DPC: 2010-2019**

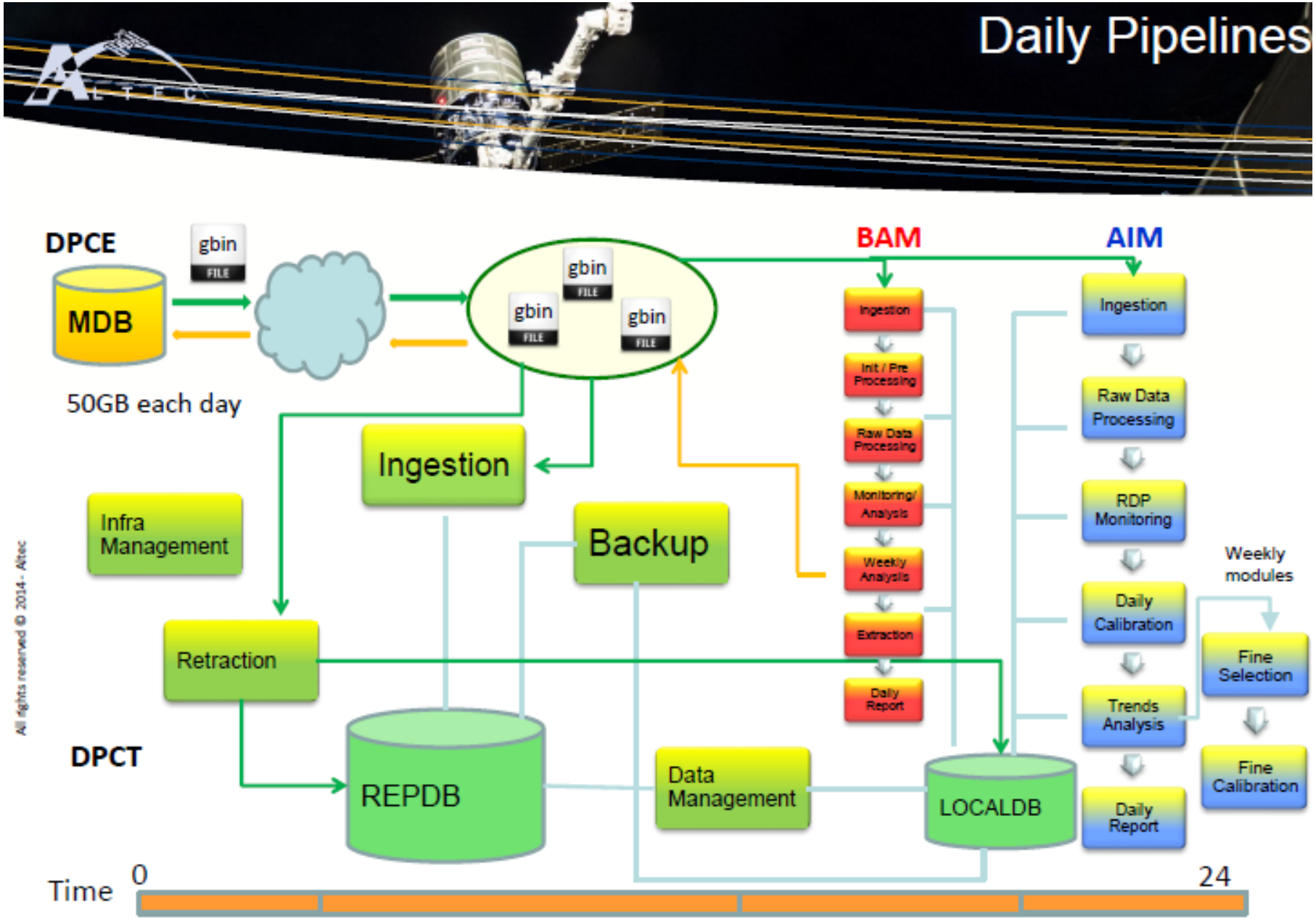
Contract: ASI I/037/08/0

SIGNATURE AND APPROVALS ON ORIGINAL

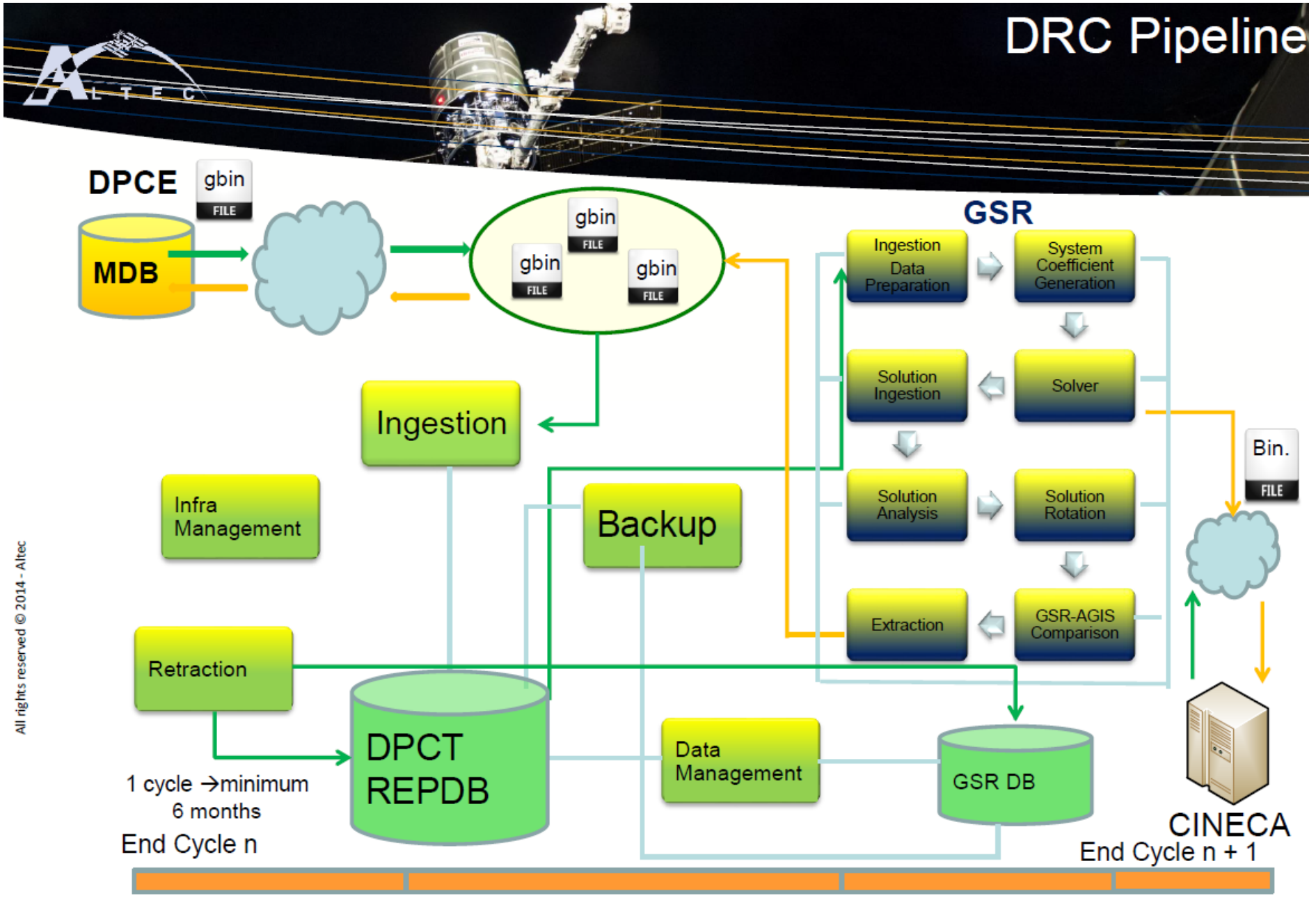
AUTHORS: MARIO G. LATTANZI (OATO-INAF), M.MARTINO (ALTEC)

APPROVALS: MARIO G. LATTANZI (PI)

Dal 2006 (4 anni di preparazione e sperimentazioni)


Daily Pipelines
DPCE
gbin
FILE
MDB
50GB each day
gbin
FILE
gbin
FILE
gbin
FILE
Ingestion
Infra Management
Backup
Retraction
DPCT
REPDB
Data Management
LOCALDB
BAM
Ingestion
Init / Pre Processing
Raw Data Processing
Monitoring/ Analysis
Weekly Analysis
Extraction
Daily Report
AIM
Ingestion
Raw Data Processing
RDP Monitoring
Daily Calibration
Trends Analysis
Daily Report
Weekly modules
Fine Selection
Fine Calibration
Time
0
24
All rights reserved © 2014 - Altec

DRC Pipeline

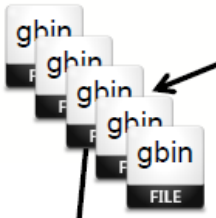


All rights reserved © 2014 - Altec

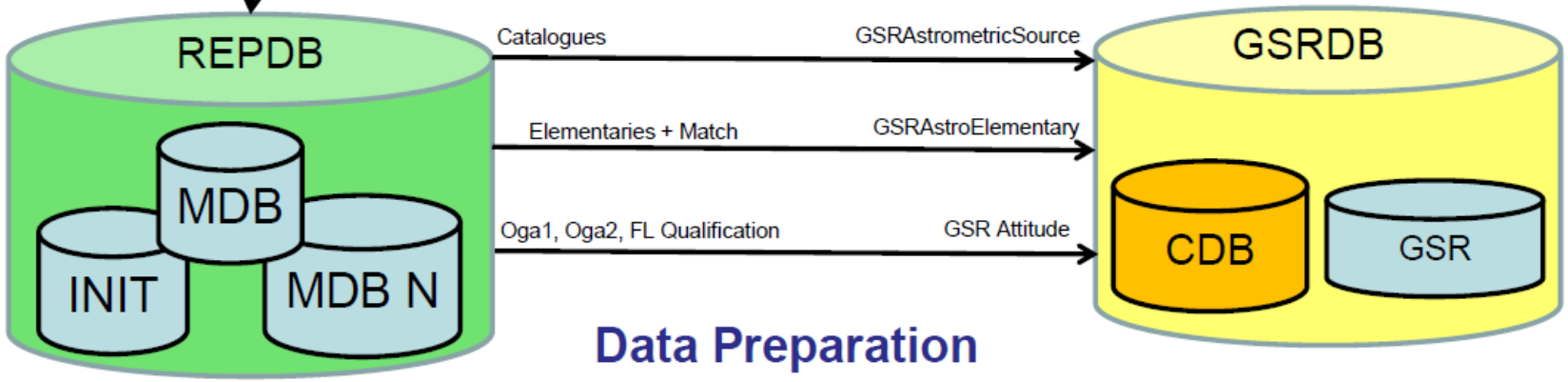
GSR Pipeline - Ingestion

From DPCE

- Oga1, Oga2, FLQualificationInfo, FLQualificationComment - MDB
- AstroElementary (IDT), Match (IDU), PhotoElementary - MDB
- AgisSource, CompleteSource - MDB N
- IgslSource - MDB INIT
- CDB

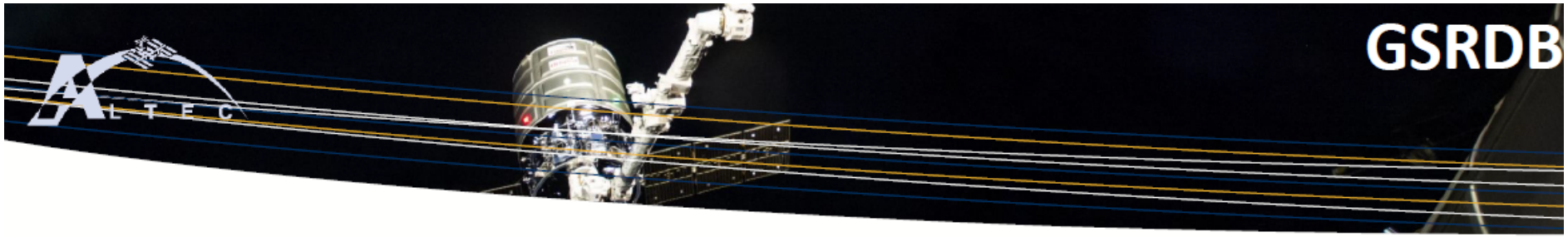


Input Ingestion

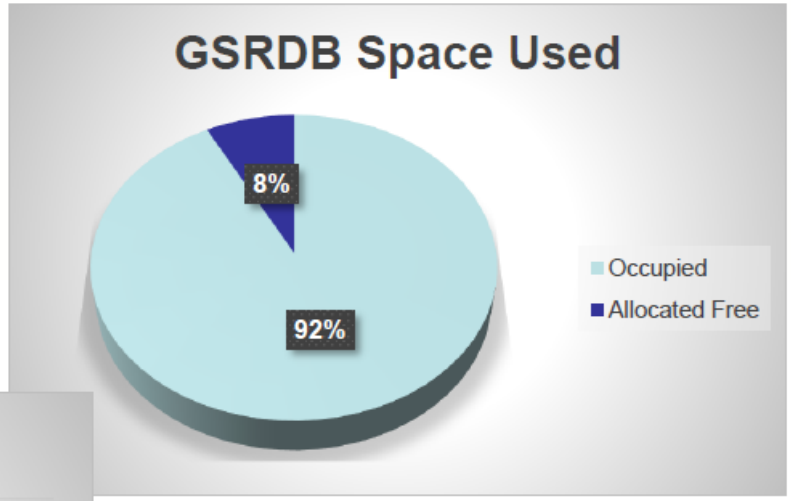


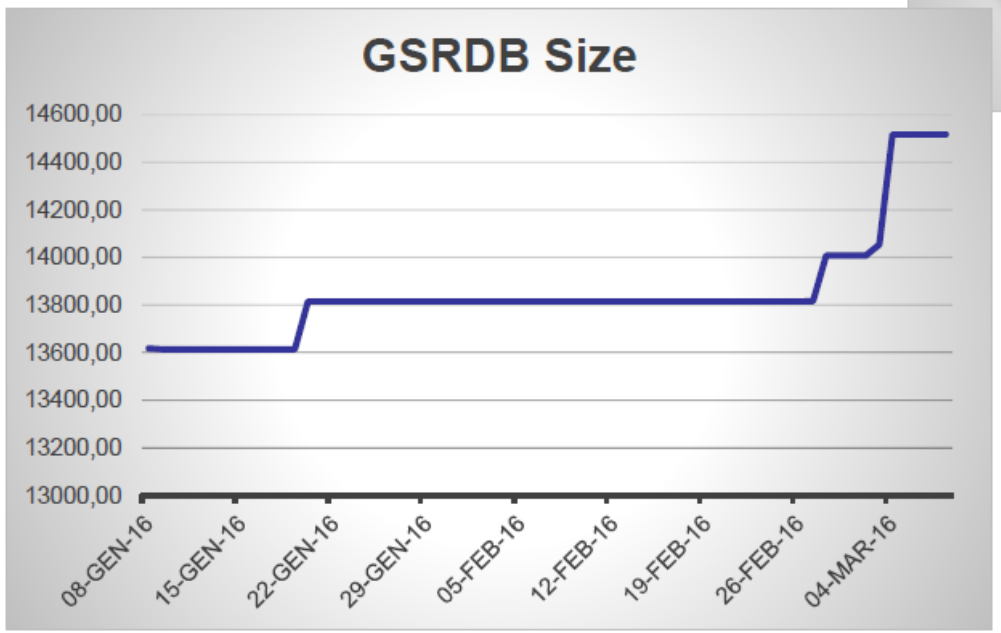
Data Preparation

All rights reserved © 2014 - Altec



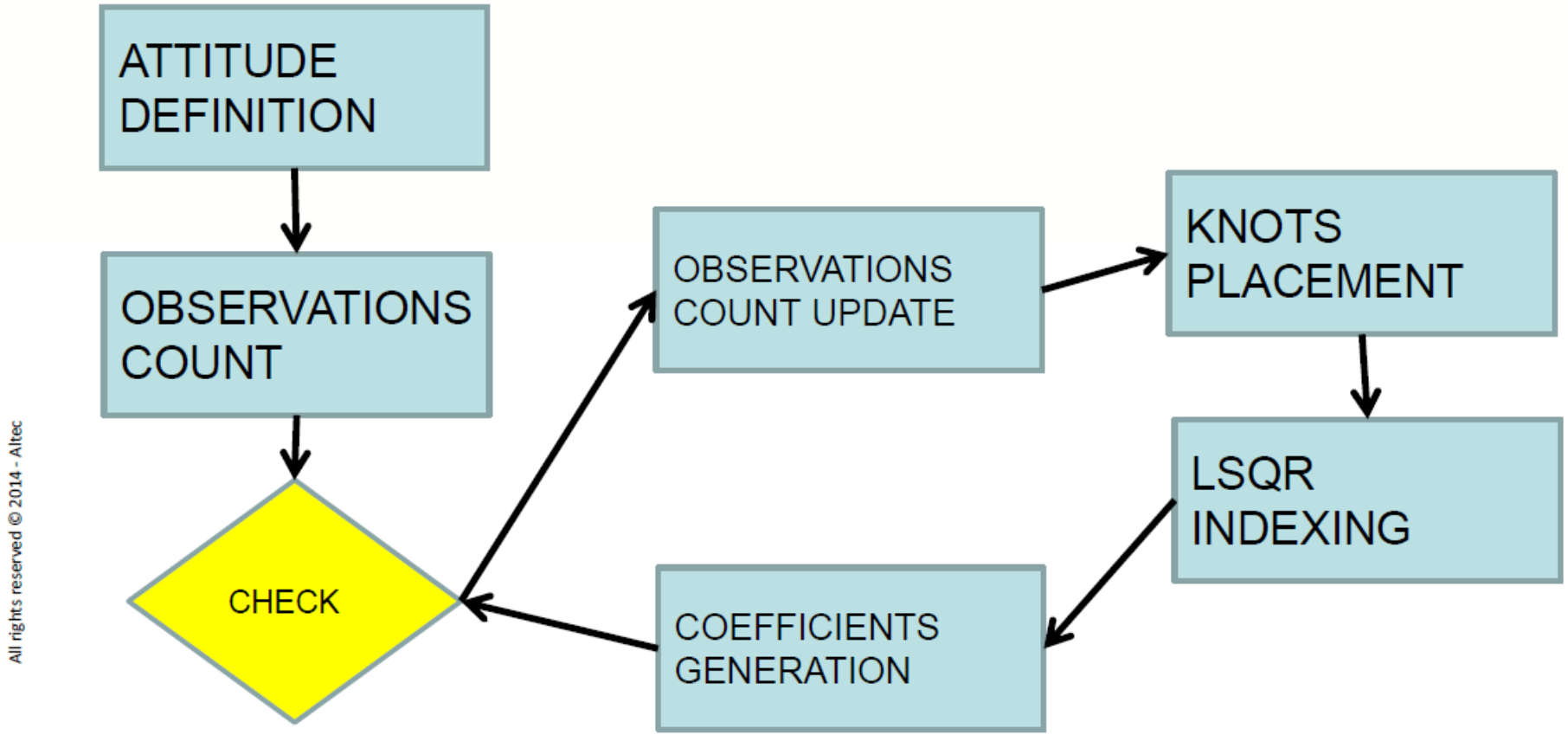
- 08/03/2016
  - Used space: *13394,52* GB
  - Allocated Free space: *1122,30* GB
  - Total: *14516,83* GB





All rights reserved © 2014 - Altec

GSR Pipeline - Coefficient Generation



All rights reserved © 2014 - Altec

gsr_iter_subworkflow, PID 40, iteration 1
gsr_iter_subworkflow, PID 45, iteration 2
gsr_iter_subworkflow, PID 46, iteration 3
gsr_iter_subworkflow, PID 55, iteration 4 -> END

HW Infrastructure

- INTERNET LINK: 1Gbps (300 Mbps guaranteed) via GARR network.
- CURRENT STORAGE CAPACITY: 700 TB overall useable disk space distributed between two HP P7400 storage units.
- FINAL STORAGE CAPACITY: more than 1 PetaByte at the end of Gaia mission on multiple storage.
- PROCESSING CAPACITY: 13 servers HP DL580 G7 with a total of 500 CPU cores and 4TB RAM.
- DB SERVERS: 4 HP DL580 G7 dedicated to the database cluster based on Oracle RAC technology.
- NETWORK CONNECTION: LAN network up to 10 Gbps.SAN network redundant at 8 Gbps.
- INFRASTRUCTURE MONITORING AND MANAGEMENT: services based on VMWare virtual environment configured with two HP DL 580 G7 servers clustered and managed by vCenter Server.
- BACKUP SYSTEM: first level on storage disks and second level on tape library (HP ESL G3/LTO-5).

All rights reserved © 2014 - Altec

- HPC INTERCONNECTION: access to HPC super computer at CINECA for dedicated processing.
- Dedicated Validation platform smaller than ops platform but similar.
- **HW replacement is needed due to obsolescence.**

HPC per Gaia

- **Classse del problema**:
  - **Risoluzione di problemi a minimi quadrati linearizzati con matrici larghe e sparse**. Questioni: **determinazione e caratterizzazione statistica della soluzione fino al calcolo completo della matrice di Covarianza**; conplessita' computazionale (solo soluzione, soluzione + varianza, soluzione + covarianza).
  - **Relatività numerica** per la modellizzazione dello spazio delle fasi della Via Lattea. (Later, possibly early 2020); complessita' TBD

- **Orizzonte temporale. Gaia da sola (operazioni, senza includere possibile estensioni, + sfruttamento): non meno di 2021 + 20 yr** (….ma vedi esperienza Hipparcos con un nuovo catalogo prodotto dopo ritrattamento dati quasi completo 10 anni dopo quello rilasciato da ESA, 1997).

- **Strumenti HPC**: HW, ovvero livello/classe dello strumento parallelo data la classe del problema (Tier?); SW: paradigma/i di parallelizzazione, numerical methods; porting in funzione di 'continuo' HW upgrade.

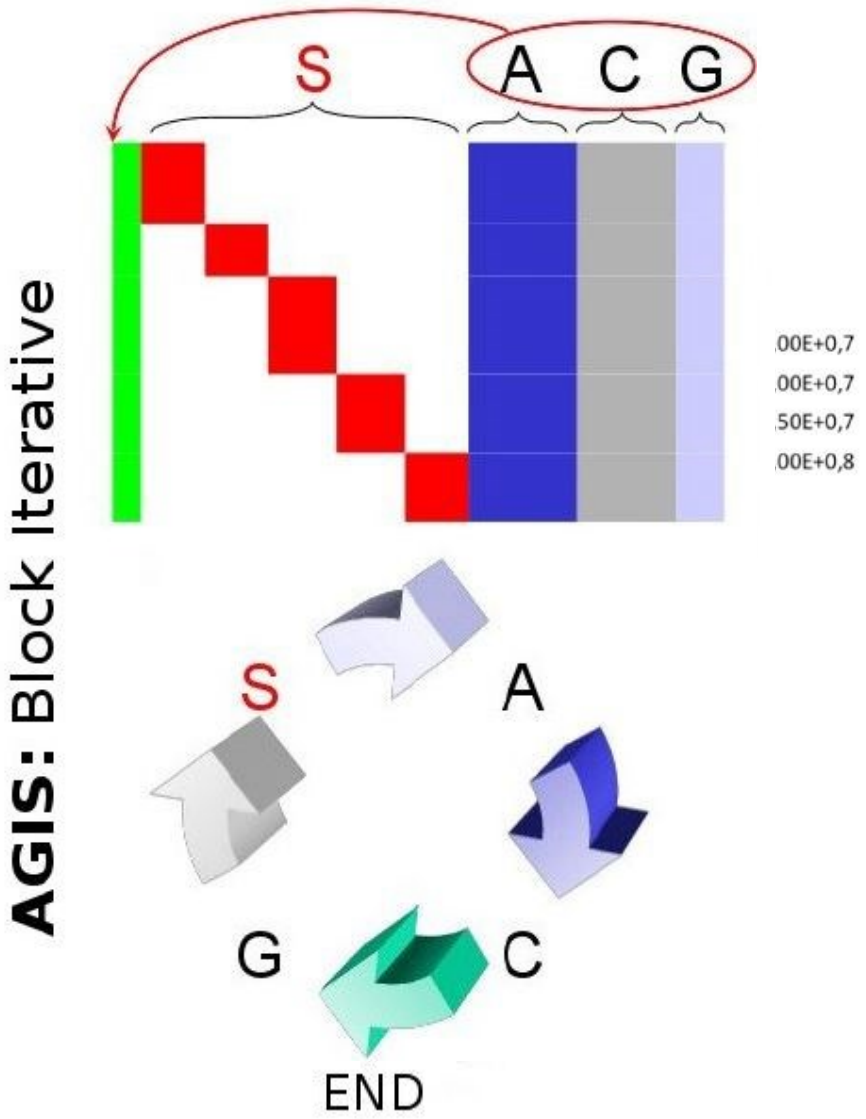
- Particolarita' legate all'esperienza Gaia: necessita' non standard, ovvero compatibilità con vincoli temporali dovute alle timelines di progetto spaziale (parte di un Science Ground System).

- Esperienza al CINECA: la facility HPC nazionale fondamentale per Gaia (giusto???); alternative?

- Risorse umane: tipologia competenze strategiche nell'Ente

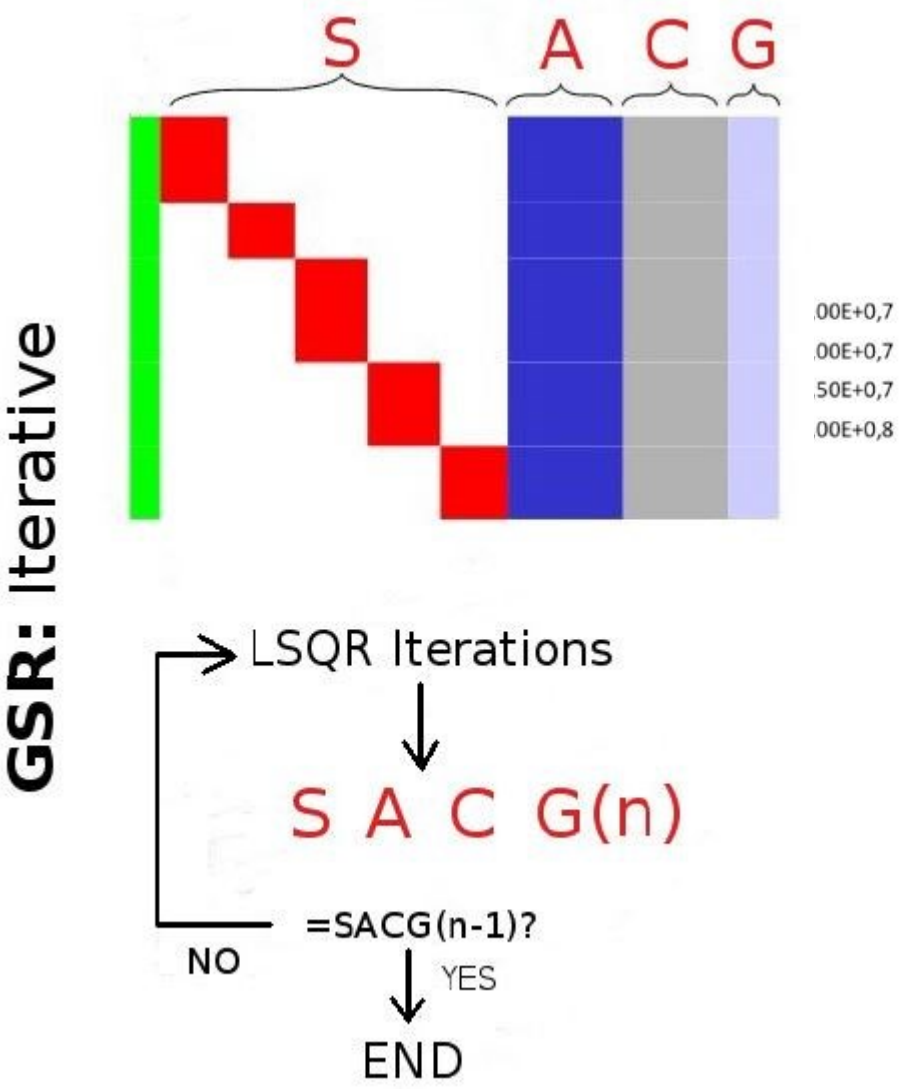
Classe del problema

- **Risoluzione di problemi a minimi quadrati linearizzati con matrici larghe e sparse**
- **Dimensioni tipiche**: 5×10^8 incognite (N), 5×10^{11} osservazioni (M), fino a 200TB per il dataset completo a fine missione, circa 20TB max. per un dataset singolo run
- **Aspetti del problema**:
  - *Costruzione del sistema e analisi della soluzione*, problema imbarazzantemente parallelo, ma HTC (interazione con DB)
  - Dal punto di vista HPC problema affrontabile con cluster, difficoltà maggiori nella gestione del dato (selezioni, ordinamenti, ecc…)
  - *Soluzione del sistema*, problema intrinsecamente parallelo (HPC ma debolmente HTC)
  - Problematiche scientifiche mai affrontate prima (assetto globale, self-calibration, fully general relativistic model)
  - Complessità computazionale:
    - metodi diretti impraticabili ($\sim 10^{11}$ PFLOPs),
    - approccio block-iterative quasi imbarazzantemente parallelo, ma limitazioni per il problema della covarianza
    - Approccio fully-iterative: 3M+7N mult. ($\sim 10^{12}$ FLOPs) per it., con 10000 iterazioni$\sim 60\,\mathrm{sec}/iteration$ (soluzione+varianza), circa 10 giorni di CPU time per full-blown solution
    - **Possibilità di estendere a soluzione+covarianza problema scientifico di frontiera sotto ogni aspetto!!!**



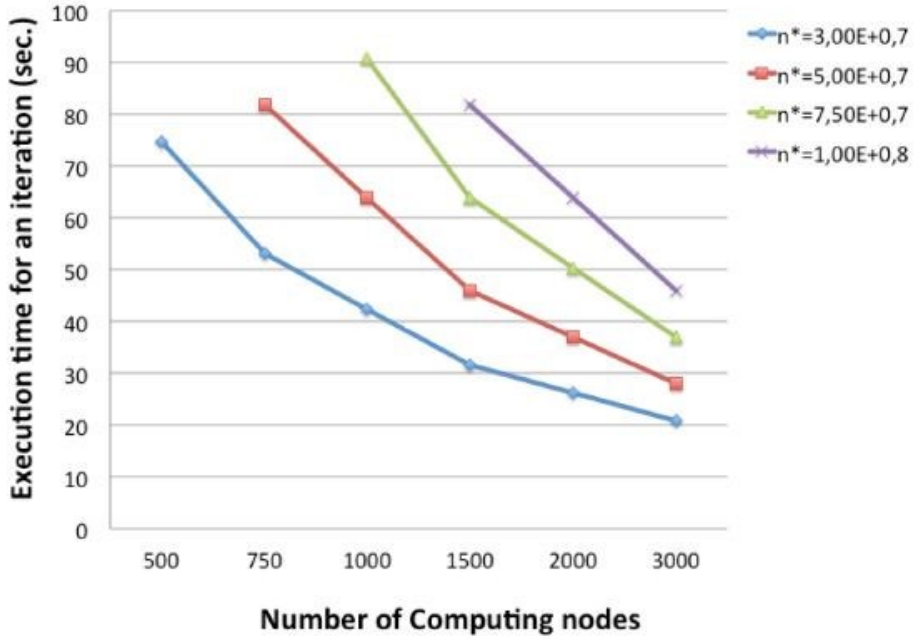
Classe del problema

- **Risoluzione di problemi a minimi quadrati linearizzati con matrici larghe e sparse**
- **Dimensioni tipiche**: 5×10^8 incognite (N), 5×10^{11} osservazioni (M), fino a 200TB per il dataset completo a fine missione, circa 20TB max. per un dataset singolo run
- **Aspetti del problema**:
  - *Costruzione del sistema e analisi della soluzione*, problema imbarazzantemente parallelo, ma HTC (interazione con DB)
  - Dal punto di vista HPC problema affrontabile con cluster, difficoltà maggiori nella gestione del dato (selezioni, ordinamenti, ecc…)
  - *Soluzione del sistema*, problema intrinsecamente parallelo (HPC ma debolmente HTC)
  - Problematiche scientifiche mai affrontate prima (assetto globale, self-calibration, fully general relativistic model)
  - Complessità computazionale:
    - metodi diretti impraticabili ($\sim 10^{11}$ PFLOPs),
    - approccio block-iterative quasi imbarazzantemente parallelo, ma limitazioni per il problema della covarianza
    - Approccio fully-iterative: 3M+7N mult. ($\sim 10^{12}$ FLOPs) per it., con 10000 iterazioni$\sim 60\,\mathrm{sec}/iteration$ (soluzione+varianza), circa 10 giorni di CPU time per full-blown solution
    - **Possibilità di estendere a soluzione+covarianza problema scientifico di frontiera sotto ogni aspetto!!!**



Classe del problema

- **Risoluzione di problemi a minimi quadrati linearizzati con matrici larghe e sparse**
- **Dimensioni tipiche**: 5×10^8 incognite (N), 5×10^{11} osservazioni (M), fino a 200TB per il dataset completo a fine missione, circa 20TB max. per un dataset singolo run
- **Aspetti del problema**:
  - *Costruzione del sistema e analisi della soluzione*, problema imbarazzantemente parallelo, ma HTC (interazione con DB)
  - Dal punto di vista HPC problema affrontabile con cluster, difficoltà maggiori nella gestione del dato (selezioni, ordinamenti, ecc…)
  - *Soluzione del sistema*, problema intrinsecamente parallelo (HPC ma debolmente HTC)
  - Problematiche scientifiche mai affrontate prima (assetto globale, self-calibration, fully general relativistic model)
  - Complessità computazionale:
    - metodi diretti impraticabili ($\sim 10^{11}$ PFLOPs),
    - approccio block-iterative quasi imbarazzantemente parallelo, ma limitazioni per il problema della covarianza
    - Approccio fully-iterative: 3M+7N mult. ($\sim 10^{12}$ FLOPs) per it., con 10000 iterazioni$\sim 60\,\mathrm{sec}/iteration$ (soluzione+varianza), circa 10 giorni di CPU time per full-blown solution
    - **Possibilità di estendere a soluzione+covarianza problema scientifico di frontiera sotto ogni aspetto!!!**



Strumenti HPC

- ALTEC (realtà industriale) come partner nel progetto (vedi sopra)
  - ALTEC: HW 8 HP DL570G7 : 4 Intel Xeon 8-cores CPUs (tot. 256 cores), 256GB RAM (tot. 2TB RAM) per Operations + piattaforma di test
  - Strutture ALTEC usate anche per testing (compreso piccolo server SGI 16 cores 256GB RAM acquistato da INAF) e soluzione di piccole sfere (fino a 20 milioni di incognite) in operations

- CINECA, tramite MoU (da rinnovare!!!!!)
  - Fermi fino a 2048 CN su 10240 (32768 cores), 16GB RAM per CN (tot. 32TB)

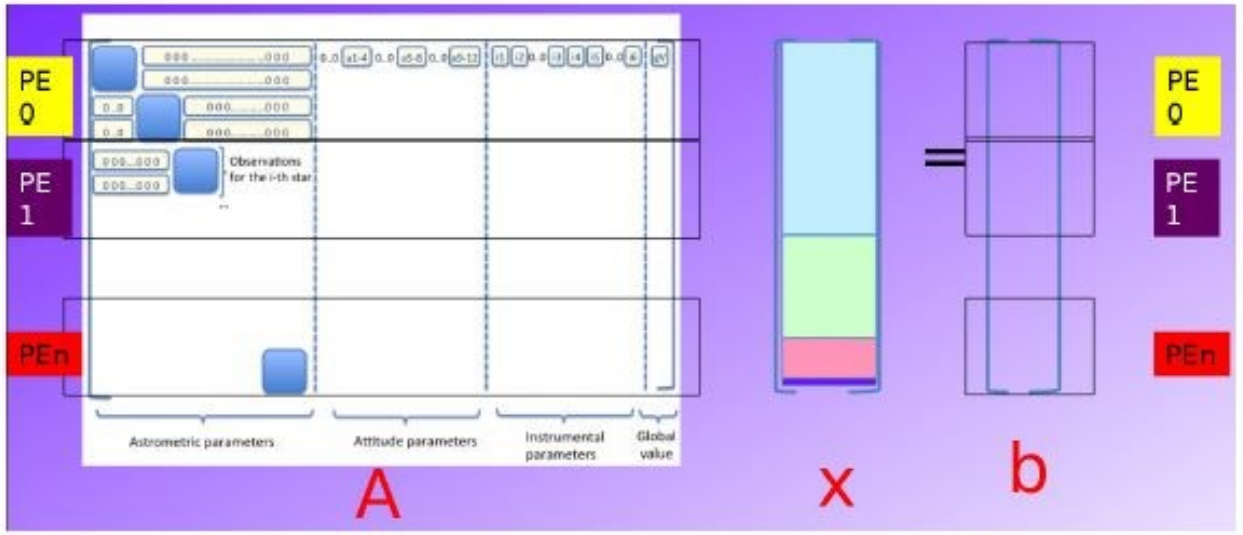
- Classe degli strumenti: Tier 0 per HPC, cluster per il resto

~40 Mhrs; for something like 50K Euro in 4 yrs

Problematiche SW

- Paradigma di parallellizzazione: C+MPI+OpenMP
- Attenzione, questa è la situazione attuale (Fermi)
- Necessario porting per prossimo switch-over a Marconi!

- Particolarità legate all'esperienza Gaia, necessità non standard **(Science Ground Center):**
  - **vincoli temporali dovuti a timelines di progetto spaziale**
  - vincoli progettuali dovuti alle necessità di interazione con un DPC complesso (6 sedi, non-disclusure agreement, ecc…)
  - problema mai affrontato prima, difficile dare inizialmente una stima affidabile delle risorse HW necessarie



  - approccio adottato ha reso accessibili diverse classi di HW senza richiedere grosse esposizioni fino alla definizione finale del problema
  - ha inoltre reso possibile cambiare «in corso d'opera» l'obiettivo del progetto, estendendone la complessità di un'ordine di grandezza rispetto a quella inizialmente prevista
  - Ciò non sarebbe stato possibile con un progetto dai confini definiti fin dall'inizio

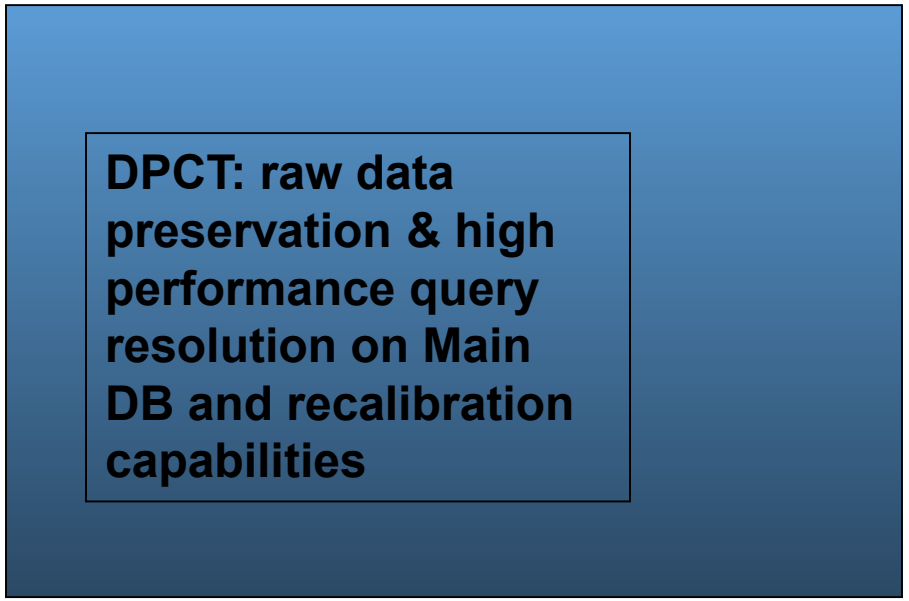
Strategia adottata e orizzonte temporale

- **Orizzonte temporale.** Gaia da sola (operazioni, senza includere possibile estensioni, + sfruttamento): **almeno 2021 + 20 yr** (…..ma vedi esperienza Hipparcos con un nuovo catalogo prodotto dopo ritrattamento dati quasi completo 10 anni dopo quello rilasciato da ESA, 1997).

- **Esperienza al CINECA: la facility HPC nazionale fondamentale per Gaia ma:**
  - Strategia «elastica» richiede un continuo monitoraggio delle risorse (Gaia è «uno» dei progetti del CINECA) e la capacità di ridiscutere la collaborazione
  - Necessaria la disponibilità di risorse di calcolo alternative, quantomeno di supporto (Tier-2?) sia per sviluppo che per sfruttamento scientifico

- **Fondamentali le risorse umane: necessario individuare la tipologia e le competenze strategiche**
  - Per l'ente necessarie figure con competenze scientifiche, numeriche e HPC finalizzate però alla copertura di interessi INAF che non si esauriscono col progetto
  - I partner (ALTEC, CINECA) devono essere selezionati non solo per la disponibilità hardware, ma per il know-how in ambiti specialistici necessario a supportare attività specifiche che per INAF sarebbero a tempo determinato (sviluppo SW, competenze DB, ecc…)

Gaia Data Legacy: tens of years (50+?)



→ IGSL

→ EUCLID (FGS, CCD's and cold electronics, SGS model)

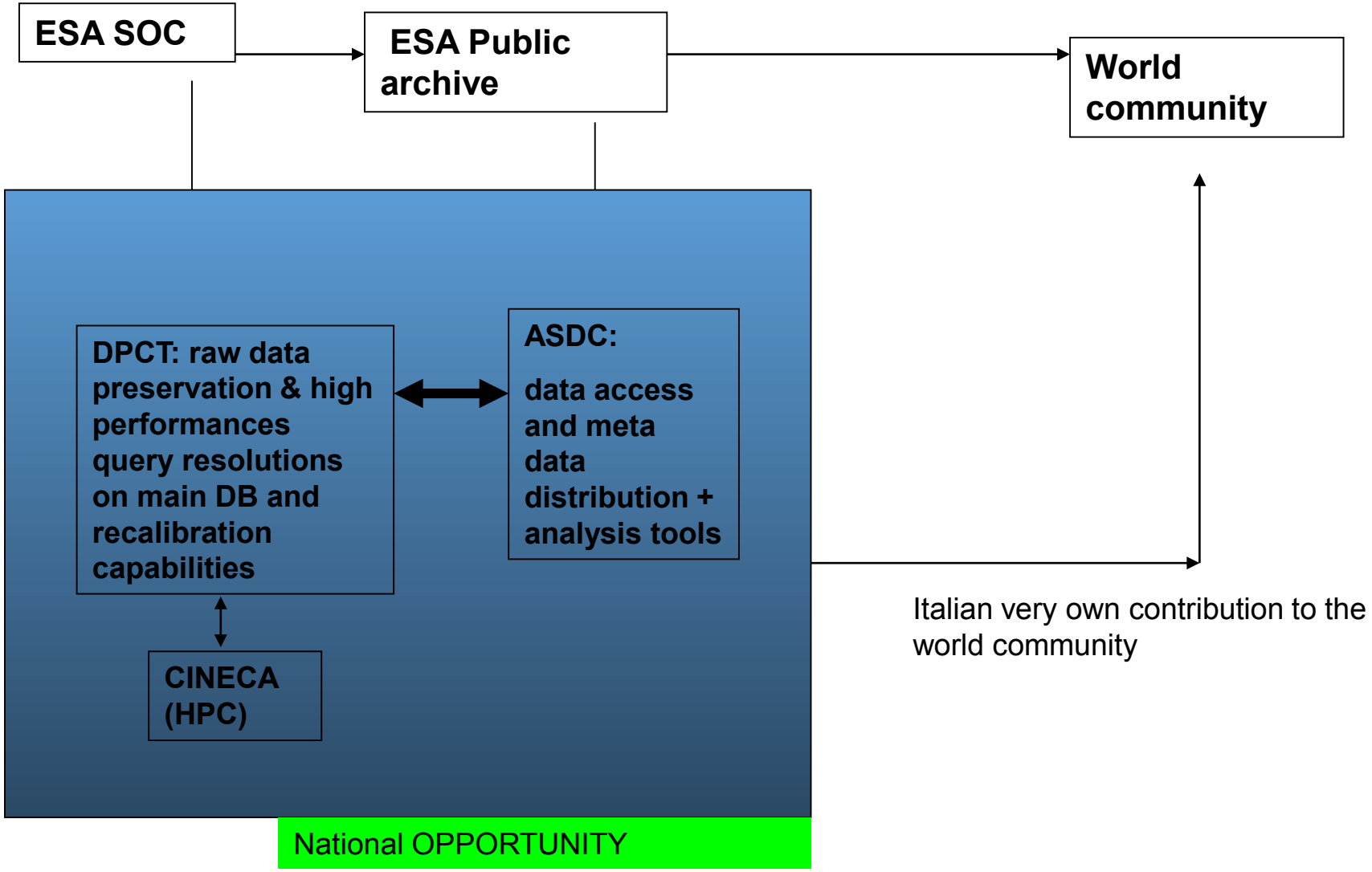
→ SPHERE

→ METIS (SGS model)

↓ PLATO

↓ Other (e.g. SKA)

- Digital Space Lab for space science and space instrumentation
- Digital (deep) sky: Sopporto progettazione nuove missioni
- Modello per futuri SGS


ESA SOC
ESA Public archive
World community
DPCT: raw data preservation & high performances query resolutions on main DB and recalibration capabilities
ASDC:
data access and meta data distribution + analysis tools
CINECA (HPC)
Italian very own contribution to the world community
National OPPORTUNITY

Conclusioni
(food for tomorrow's thoughts)

La parola d'ordine e' SINERGIA!

>>> Verso un Laboratorio (fisico+virtuale) Nazionale INAF ICCT INAF cosi' da essere uno dei principali (principale?) stakeholder in un sistema Nazionale sinergico/integrato in cui sono anche coinvolti soggetti industriali edegauati (→ Infrastruttura spesso importante almeno quanto l'HCP/HCT)

>>> Introdurre in INAF il principio dell' «entropia controllata» ☺☺

>>> Concentrare le risorse umane su specifiche competenze

>>> Mettere a frutto anche le risorse finanziarie delle missioni spaziali (coordinamento/sinergia/tavolo nazionale, long term programs and exploitations)