

# Traitement des données Gaia

Aperçu, organisation

Participation française

----

F. Mignard

OCA/Cassiopée



- **Le traitement de données**
- **Formation & organisation du DPAC**
- **Participation Française**



I- Les Données et le traitement

## ■ Trois instrument fournissant trois flots de données

- ◆ Données astrométriques des CCD du champ principal
- ◆ Données photométriques des deux photomètres (BP, RP)
- ◆ Données spectroscopiques

## ■ Toutes les données arrivent sous forme de paquets de télémétrie

## ■ Volume de données

- ♦ télémétrie compressée 250 Tb
- ♦ données brutes au sol 100 TB
- ♦ données traitées et archives ~1 PB

## ■ Dimension du traitement

- ♦  $1.5 \times 10^{21}$  FLOPs

## ■ Puissance de calcul au sein du DPAC

- ♦  $> 20$  TFLOP/s  $\rightarrow$  2 yr CPU pour  $10^{21}$  FLOPs

## ■ Transfert des données

- ♦ Flux quotidien vers le sol 50 GB/day
- ♦ Échange entre les DPCs : toujours un défi

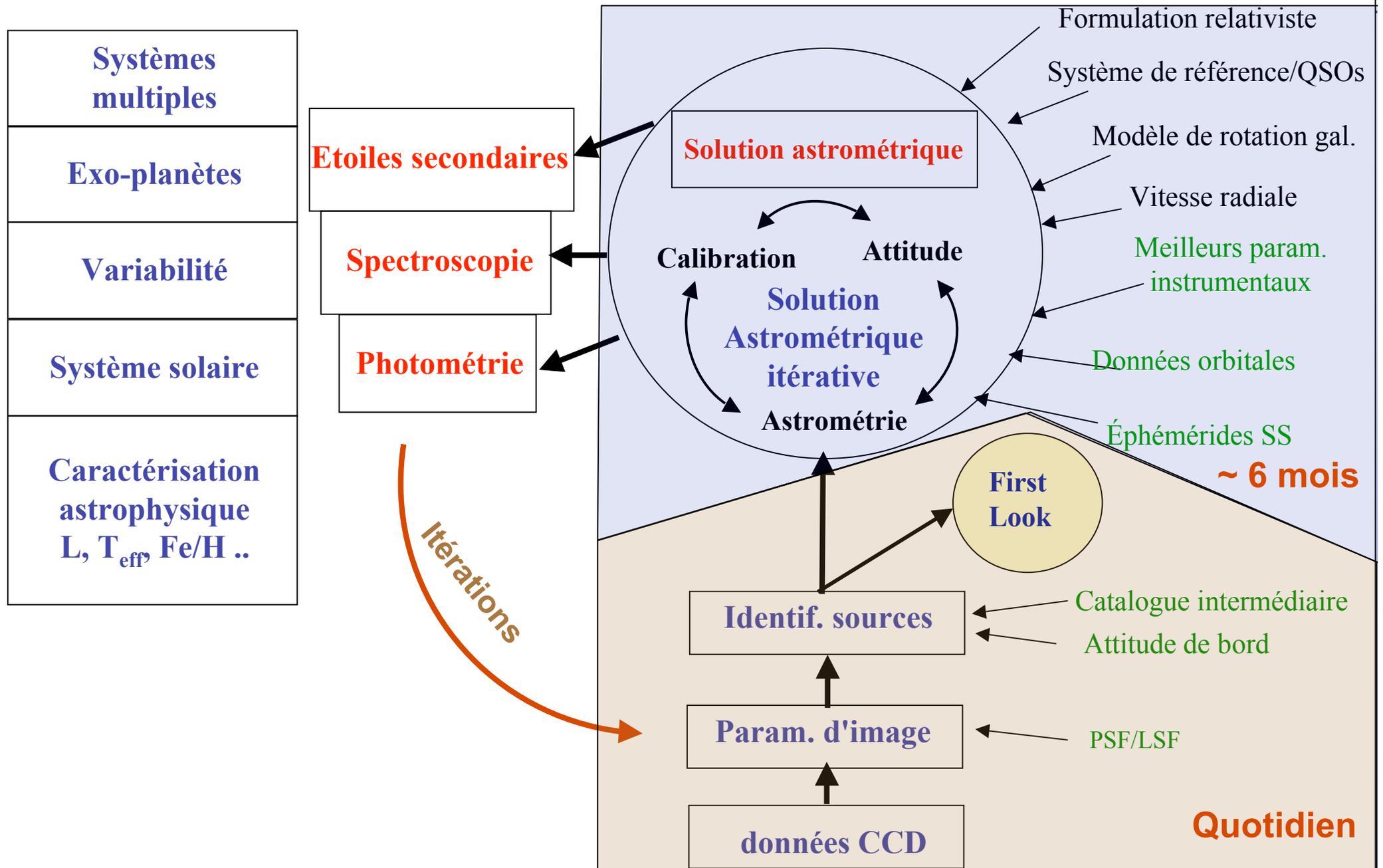
## ■ Traitement global :

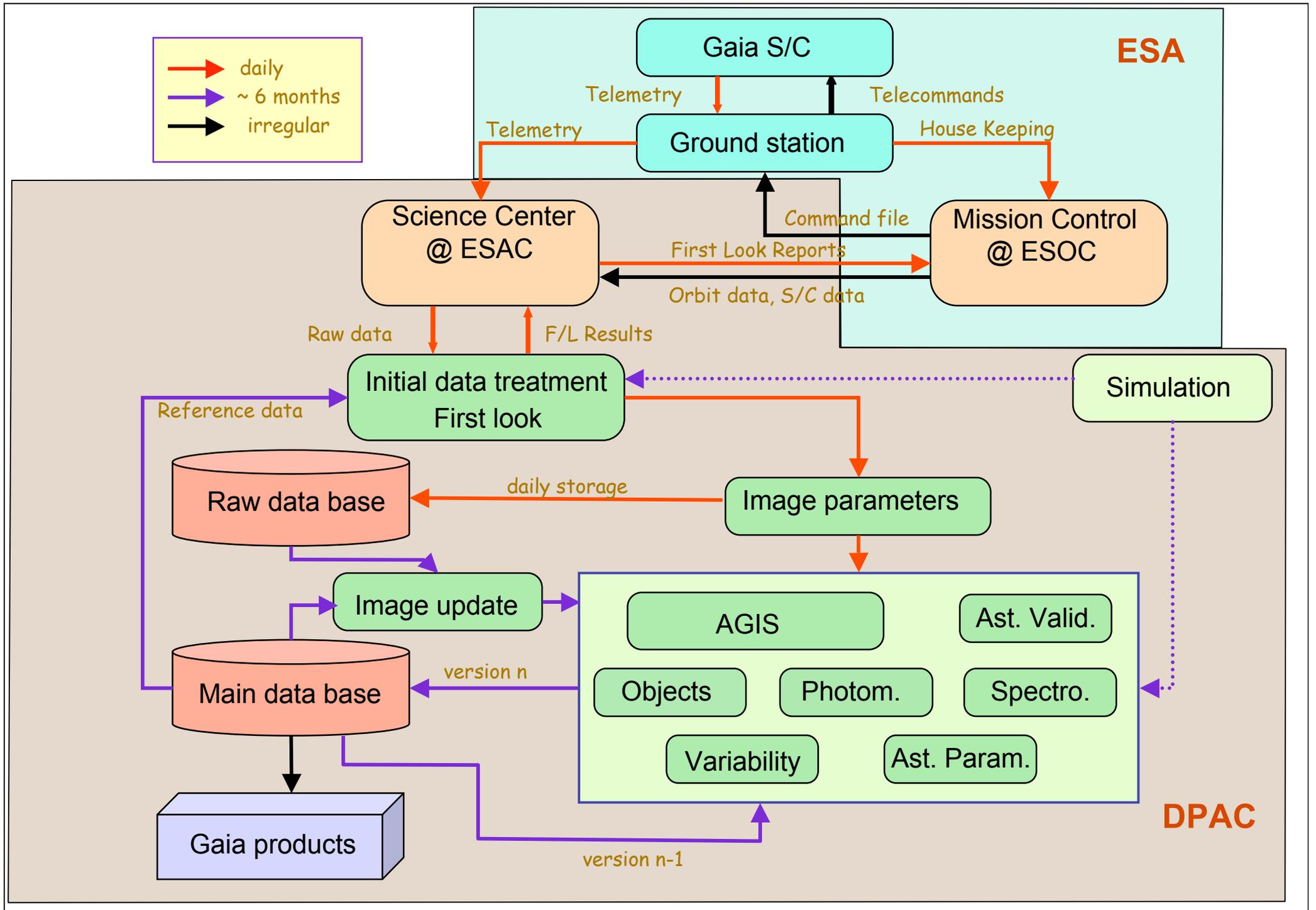
- ◆ Réception des données, calibration, attitude, système (AGIS)
- ◆ Mise à jour de la base de données, stockage

## ■ Données scientifiques :

- ◆ Traitement général par objet
  - solution astrométrique des étoiles secondaires
  - photométrie, analyse des variables
  - analyse des données RVS
  
- ◆ Objets particuliers
  - étoiles doubles et multiples
  - galaxies non résolus
  - quasars
  - objets du système solaire
  
- ◆ Analyse scientifique globale

# Structure d'ensemble du traitement de données





## II- Formation & organisation du DPAC

## ■ Pas de doublement du traitement

- ♦ principes acquis avec Hipparcos
- ♦ ressource non disponibles pour une duplication
- ♦ mais certains éléments spécifiques pourraient l'être
  - c'est un élément de la validation

## ■ Pas de possibilité d'avoir un institut spécialisé

- ♦ Le développement du S/W doit être effectué en de nombreux endroits
- ♦ cela implique une grosse coordination

## ■ H/W dans un nombres très limité de centre

- ♦ instituts où sera fait le traitement effectif

## ■ Le traitement de données est très largement à la charge de la communauté

### ♦ le projet finance :

- le satellite, le lancement, les opérations,
- réception et archivages des données
- le traitement initial et l'astrométrie 'core'

### ♦ ESA ne finance pas l'analyse scientifique

### ♦ Financement à trouver auprès des agences nationales → Consortium Européen

## ■ Un prototype de traitement de données a été développé en 2003-2006

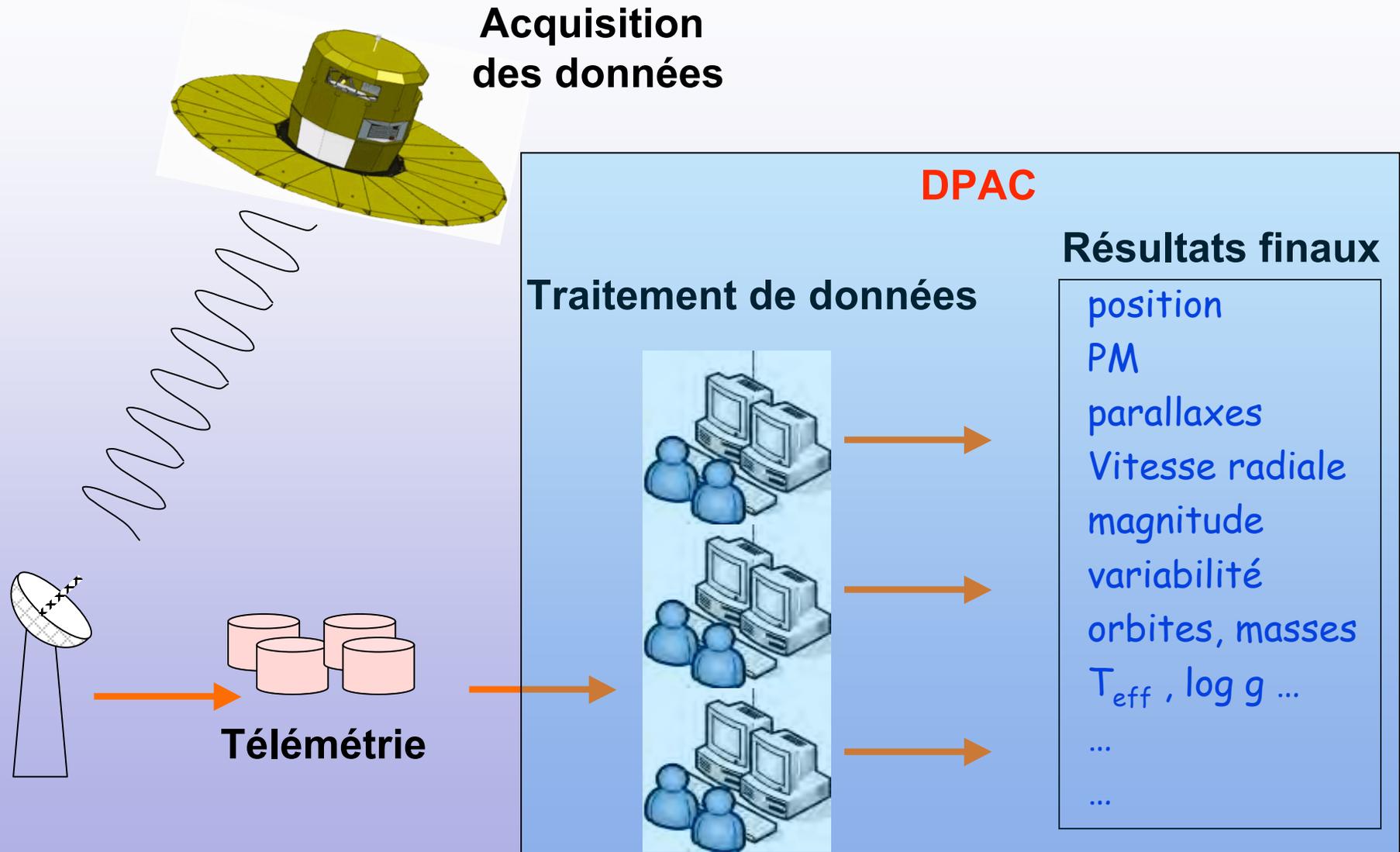
- ♦ objectif : évaluer la faisabilité des approches actuelles
- ♦ aborde : le système, le matériel, la base de données, les accès
- ♦ définition des algorithmes, intégration et tests

## ■ L' ESA a émis un appel d'offre vers la communauté

- ♦ publié en novembre 2006
- ♦ il ne concerne que le traitement de données

## ■ Un consortium s'est formé dans le but de répondre

- ♦ DPAC = Data Processing & Analysis Consortium
- ♦ Il doit transformer les données brutes en résultats astronomiques
  - production d'un catalogue astrométrique, photométrique et spectroscopique
  - sur les étoiles, les objets du SS, les quasars
- ♦ Officiellement sélectionné par le Comité des Programmes Scientifiques de l'ESA en mai 2007



## ■ Préparation de l'analyse des données

- ♦ choix de la structure générale
- ♦ définition et programmation des algorithmes

## ■ Production de données simulées pour assister :

- ♦ la conception de l'ensemble
- ♦ le développement des méthodes
- ♦ les tests à toutes les échelles

## ■ Conception, développement et mise en place de la chaîne de traitement

- ♦ gestion des flots de données entre les groupes
- ♦ test et intégration des logiciels
- ♦ acquisition, maintenance et mise en place du matériel (H/W)
- ♦ production des données intermédiaires et finales
- ♦ Contrôle qualité, documentation, interfaces

- En juillet 2007 : 330 membres → ~ 180 FTE
- 16 pays représentés
- Essentiellement postes stables dans le monde universitaire

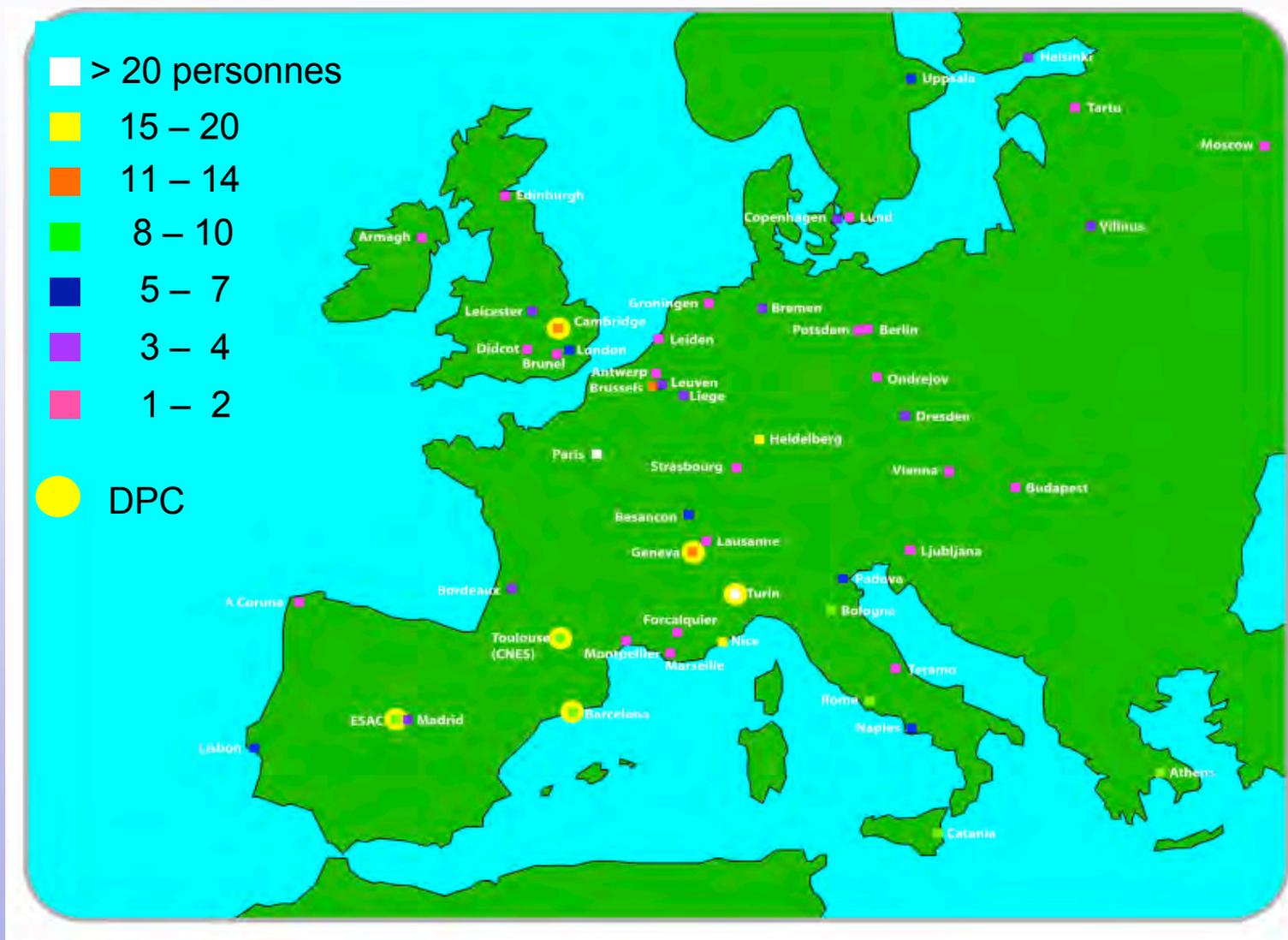
## Distribution sur les CUs

CU1	CU2	CU3	CU4	CU5	CU6	CU7	CU8	total
34	67	69	58	55	50	46	64	443

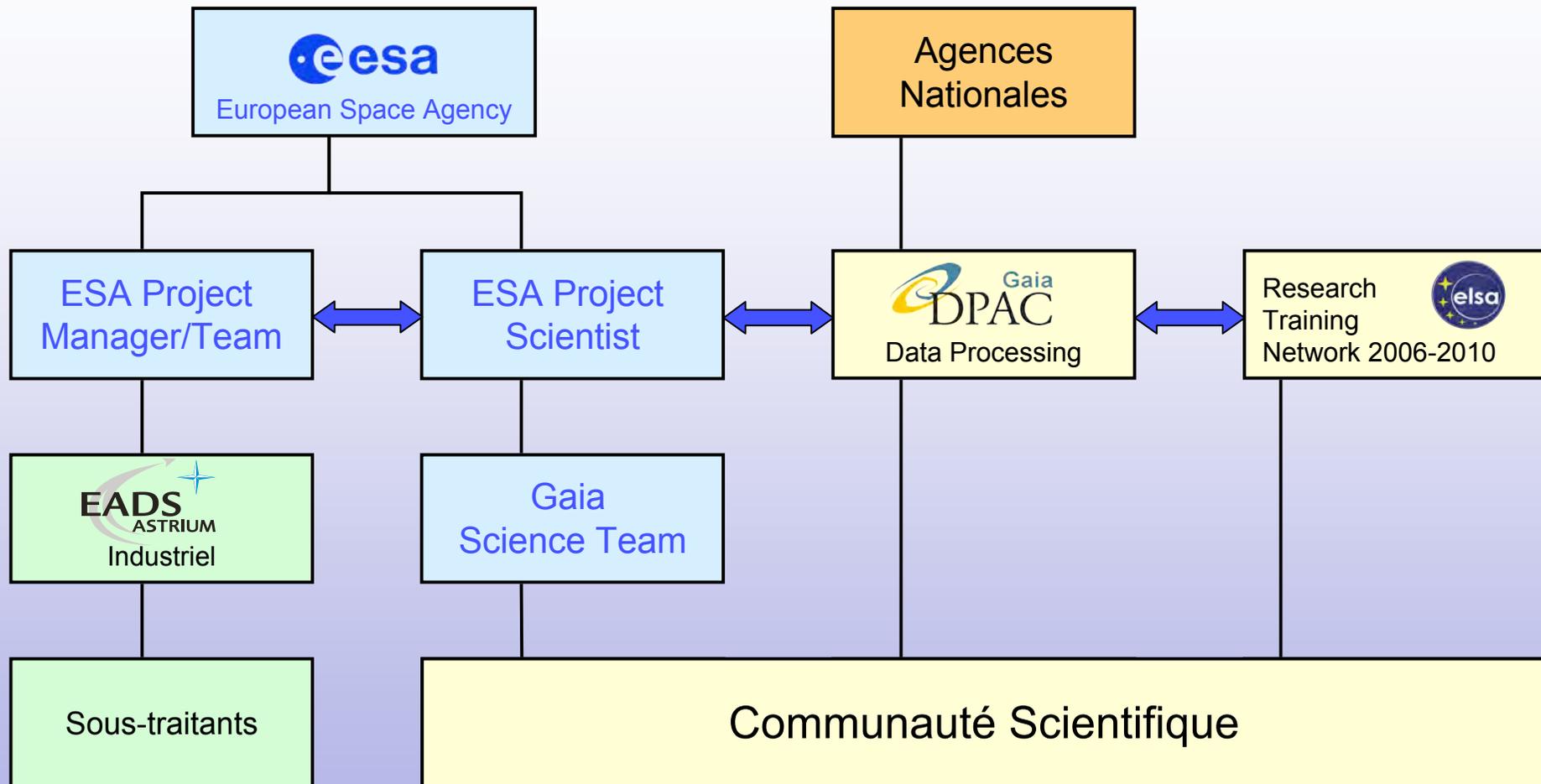
## Les contributeurs les plus importants

fr	it	de	be	uk	es	ch	gr	esac
82	62	28	23	26	21	13	9	14

# Les Instituts Participants



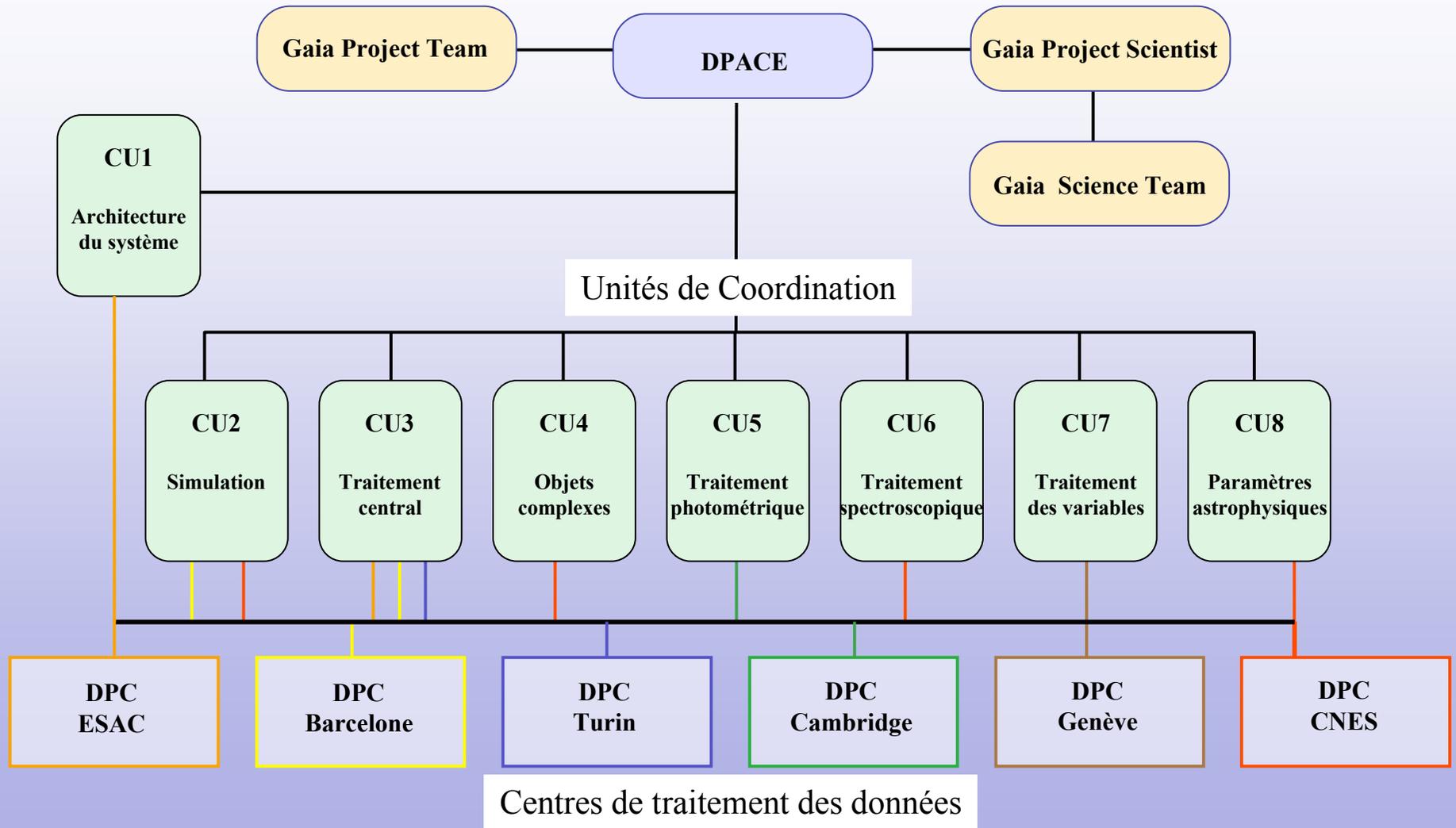
# Le DPAC dans la structure mission



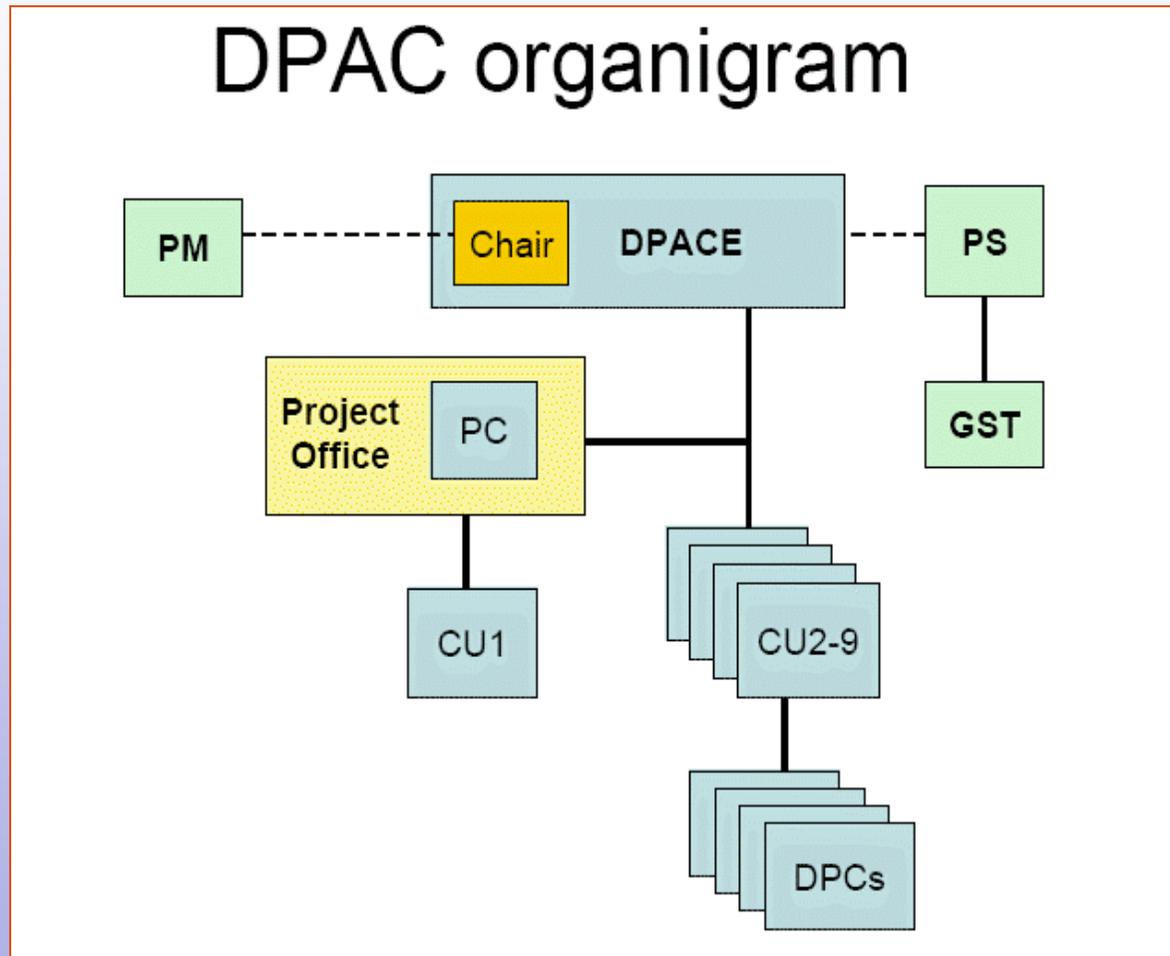
"Gaia People Finder" recense environ 500 personnes (excluant l'industrie)

# Organigramme général du DPAC

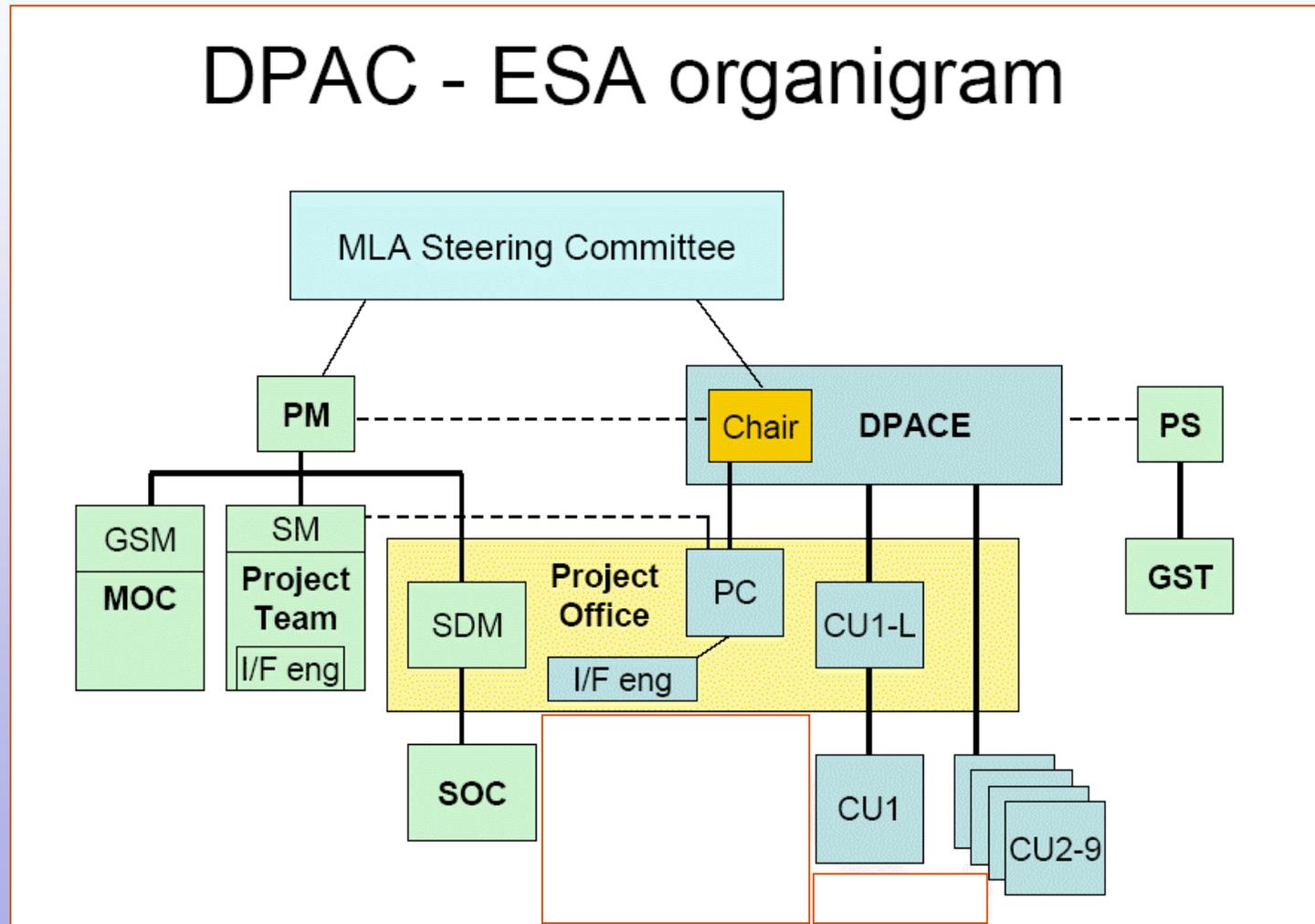
■ Pas de relation simple : une CU  $\leftrightarrow$  un DPC

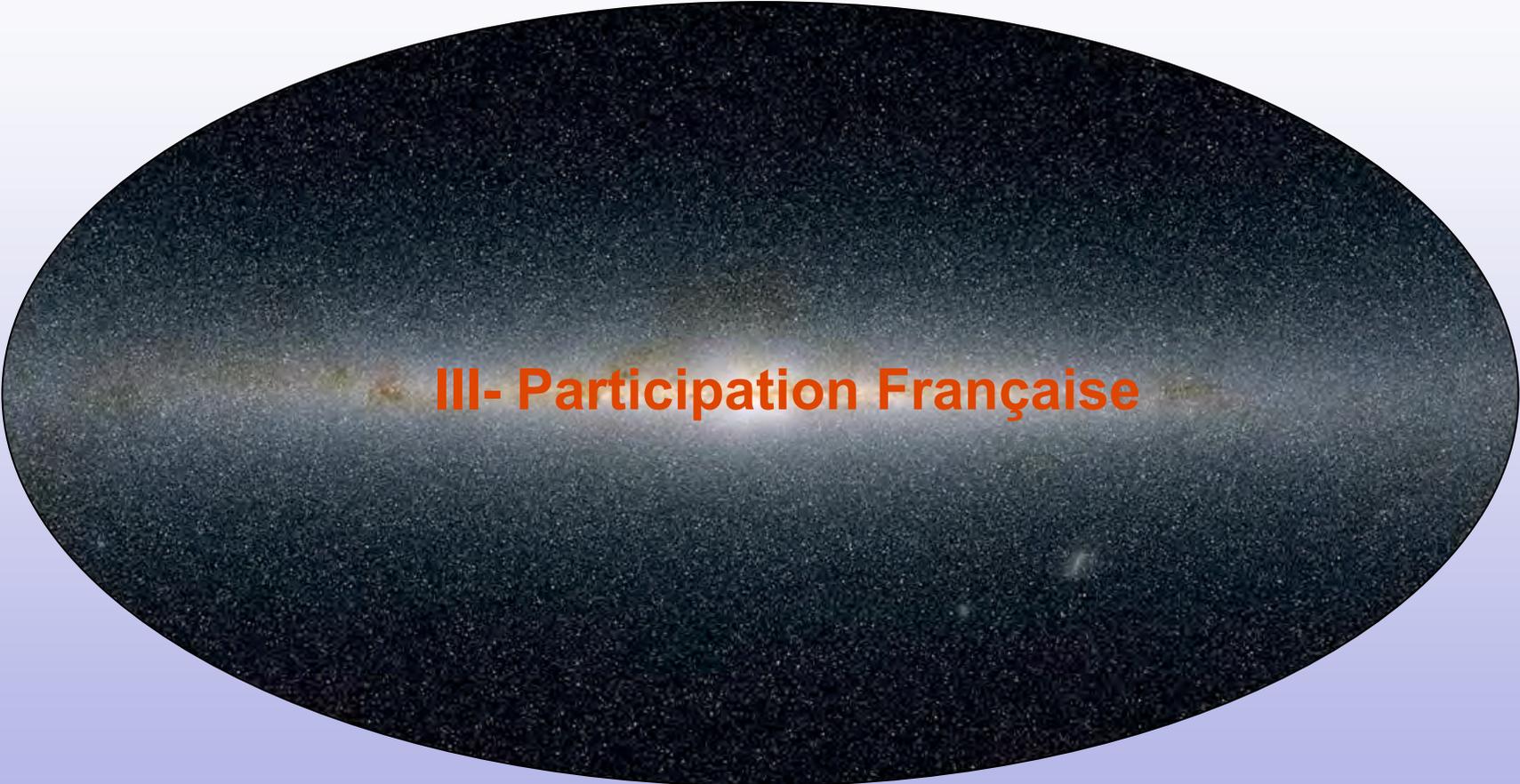


■ Pas de relation simple : une CU  $\leftrightarrow$  un DPC



■ Pas de relation simple : une CU  $\leftrightarrow$  un DPC





### III- Participation Française

## ■ Equipes scientifiques :

### ♦ 3 pôles :

- Observatoire de Paris (GEPI, IMCCE, SYRTE, LESIA)
- Observatoire de la Côte d'Azur (Cassiopée)
- CNES/CST

### ♦ Autres centres

- Observatoires de Besançon, Bordeaux, Marseille, Montpellier, Strasbourg
- IAP

### ♦ 80 membres du DPAC, 25% de l'effectif

## ■ Effectif

- ♦ OBSPM 32 (dont IMCEE 8)
- ♦ OCA 17
- ♦ CNES 10 (+ 4 externalisés)
- ♦ Obs. Besançon 5
- ♦ Obs. Strasbourg 4
- ♦ IAP 4
- ♦ Obs Bordeaux 2
- ♦ OAMP 2
- ♦ UNSA 2
- ♦ GRAL 1

Dans les CUs:	1	2	3	4	5	6	7	8
	8	28	7	29	1	29	4	17

## ■ Responsabilités

- ♦ Management général (2 DPACE, présidence)
- ♦ Responsabilité du DPC-C (CNES)
- ♦ 5 coordinateurs ou co-coordinateurs de CUs (CU1, 2, 4, 6 et 8)
- ♦ 6 membres des Comités de Pilotage des CUs (CU4 et 6)
- ♦ Responsabilité du GBOG (Coordination des observations au sol)

## ■ Soutien aux équipes

- ♦ CNRS, Ministère
- ♦ CNES
- ♦ Laboratoires d'accueil
- ♦ Commission Européenne avec ELSA
- ♦ Collectivités territoriales (en province)

## ■ Thématiques

- ◆ Etudes instrument RVS
- ◆ Logiciel de bord pour la détection
- ◆ Simulation:
  - analyse générale, modèle d'Univers
  - GIBIS, spectro, système solaire, QSOs, validation, spectres synthétiques
- ◆ Préparation analyse de données
  - spectro (définition du système, calibration lambda, analyse des spectres)
  - systèmes multiples, binaires spectro, système solaire
  - classification automatique, paramètres atmosphériques
- ◆ Système de référence, tests physique fondamentale
- ◆ Données auxiliaires (stellaires, physique, éphémérides)



**Merci de votre attention**

## ■ Le traitement nécessite des expertises spécifiques et un fort engagement :

- ♦ Analyse globale d'un projet complexe
- ♦ besoin de développeurs S/W professionnels
- ♦ contrôle qualité indispensable
- ♦ engagement des équipes sur ~ 10 à 15 ans
- ♦ capacité à travailler dans un environnement très contraint

## ■ L'organisation type n'existait pas

- ♦ Ce n'était plus le même type d'activité que dans les WG
- ♦ le niveau de coordination est sans commune mesure
- ♦ l'analyse scientifique n'est pas la première tâche du DP mais un but final
- ♦ les responsabilités doivent être claires et hiérarchisées

## ■ Solution:

- ♦ Mise en commun des efforts et des ressources dans un consortium

# Qu'est ce que le DPAC

- **Group transnational structuré pour préparer et exécuter le DP**
- **Ce groupe a répondu à l'appel d'offre de l'ESA**
  - ◆ **il doit être en mesure de fonctionner pendant ~ 13 années**
- **Ce groupe doit avoir une visibilité :**
  - ◆ **pour le projet en général (ESA, science program, AWG/SPC)**
  - ◆ **nationalement pour les financements et les postes**
  - ◆ **au niveau européen pour les programmes post-doctoraux**
- **C'est un élément de visibilité du projet au niveau international**
- **le DPAC doit avoir sa propre organisation et ses ressources pour assurer le traitement dans son intégralité**
- **Les spécifications ont été données dans l'AO et le SMP**
- **Le DPAC a fait une proposition à l'ESA**

# Gaia – Etat du projet et calendrier

