

PetaSky – Experience, Feedback

PREDONx 2015 : Atelier sur la Préservation des Données Scientifiques
Mercredi 9 décembre 2015,
Observatoire Astronomique de Strasbourg

C. Surace Laboratoire d'Astrophysique de Marseille

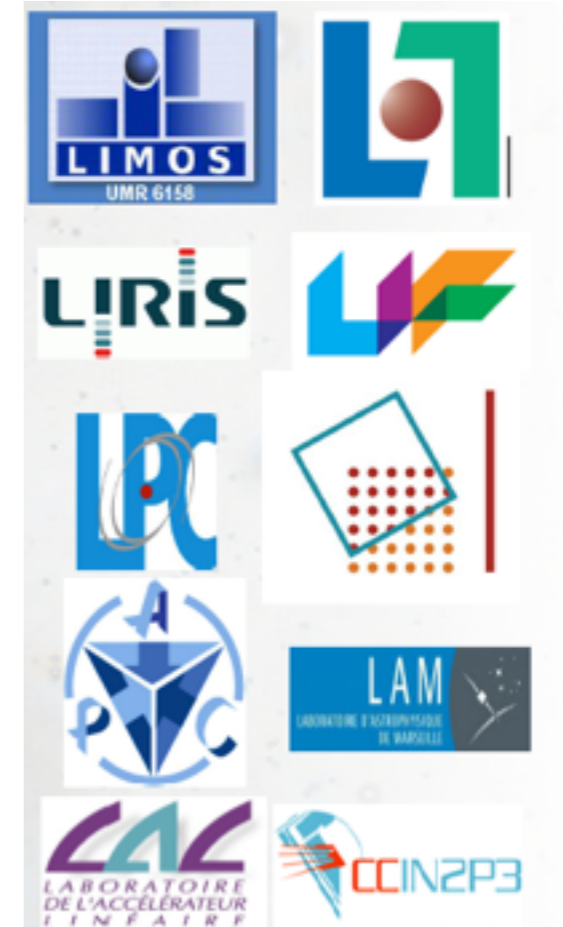
Résumé extrait de :

présentation WC meeting - Cost BigSky Earth, 2015 (K. Zeitouni)
présentation MASTODONS 2014-2015 (F. Toumani, E. Gangler)
pour le projet PETA SKY

Petasky

Le consortium

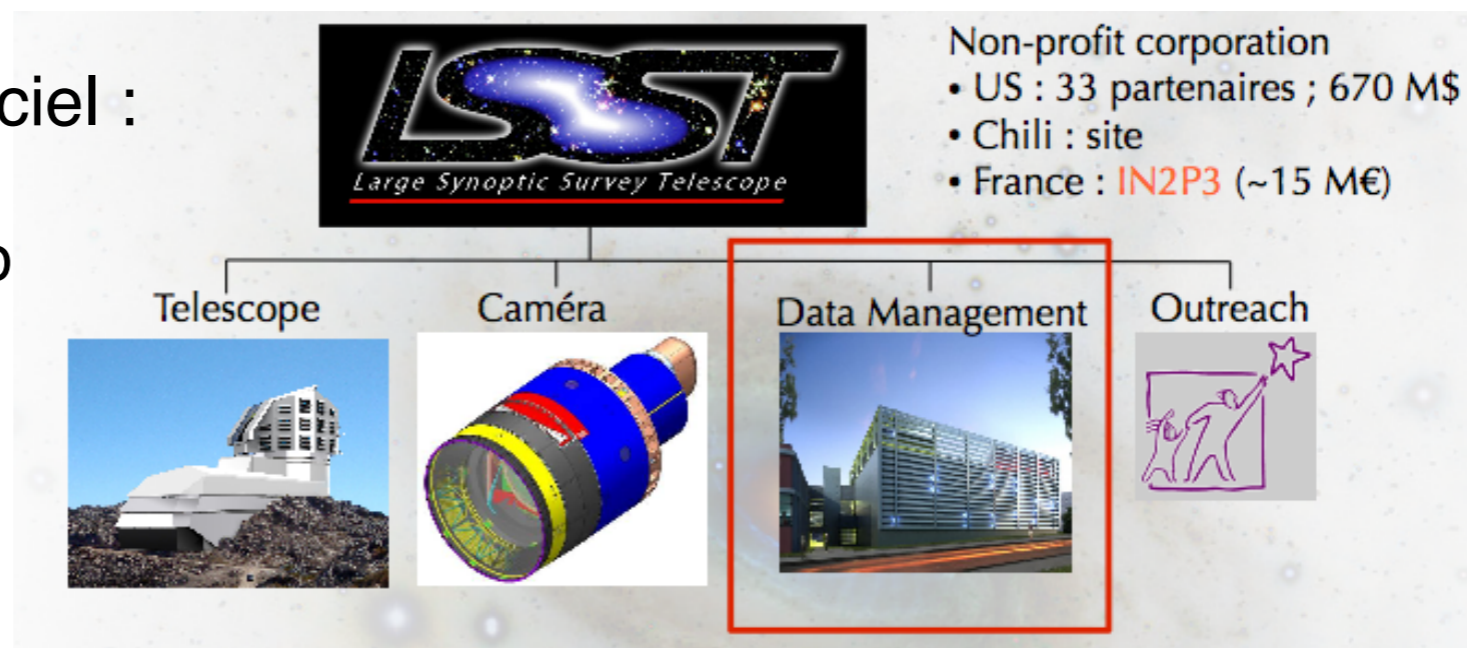
- **INS2I**
 - LIMOS (UMR 6158, Clermont-Ferrand)
 - LIRIS (UMR 5205, Lyon)
 - LaBRI (UMR 5800, Bordeaux)
 - LIF (UMR 7879, Marseille)
 - LIRMM (UMR 5506, Montpellier)
 - **IN2P3**
 - LPC (UMR CNRS 6533, Clermont-Ferrand)
 - APC (UMR CNRS 7164, Paris)
 - LAL (UMR CNRS 8607, Paris)
 - Centre de Calcul de l'IN2P3/CNRS (CC-IN2P3).
 - **INSU**
 - LAM (UMR 7326, Marseille)
- 19 chercheurs ;
 - 8 ITA ;
 - 5 doctorants



LSST

Une nouvelle fenêtre sur le ciel :

- Télescope de 8,4 m
- Astronomie très grand champ
- Cerro Pachon (Chili)
- Démarrage 2020
- CNRS partenaire du projet

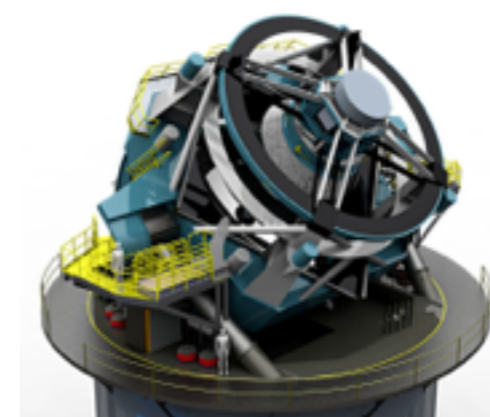


Tout le ciel visible en 6 bandes (ugrizy) (20000 sqdeg)

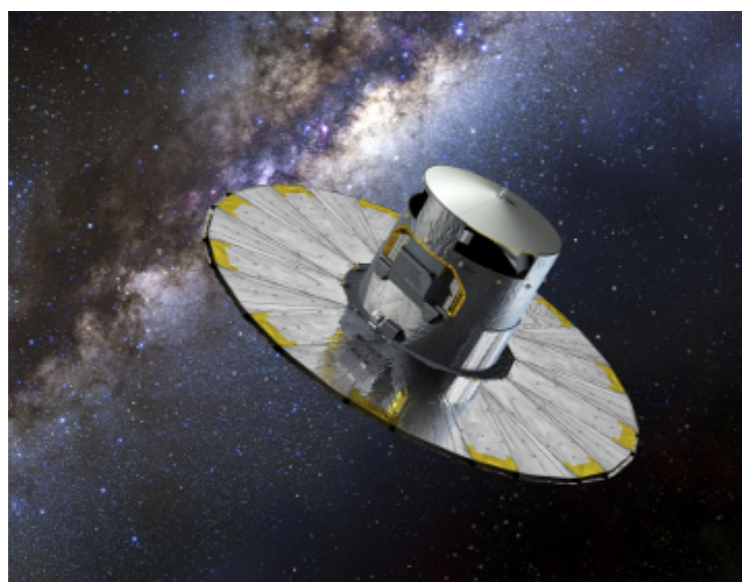
- Poses de 15 s, 1 visite / 3 jours
- **10 ans, 60 Pbytes de données brutes**

Paradigme pour l'analyse : Astrominformatique
(characterize first, analyze later)

- Comment accéder efficacement à ces données ?
- Quels outils pour l'analyse et la visualisation



Gaia



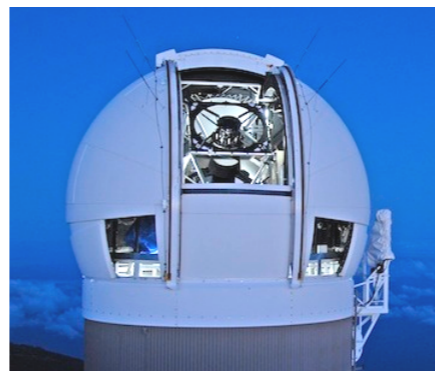
GAIA space telescope

Launch date: 19. December 2013

1PB of data after **5 years** (40 million observations a day!) (free data catalogue access policy)

Applications: 3D catalogue of ~1 billion astronomical objects

The era of Big Data has arrived!



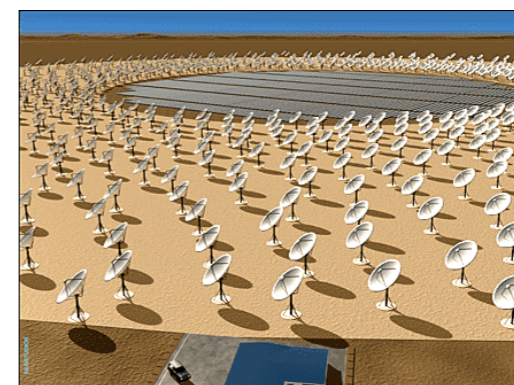
Pan-STARRS (NEO defence)
>100 TB of data



LOFAR (radio telescope):
1 petabyte per year



Large Synoptic Survey Telescope:
30TB of imaging data each night



Square Kilometre Array:
1.5 exabytes per year



The Virtual Observatory (VO): provides standards describing all astronomical resources worldwide and supports standardized discovery and access to these collections

extrait de la presentation COST

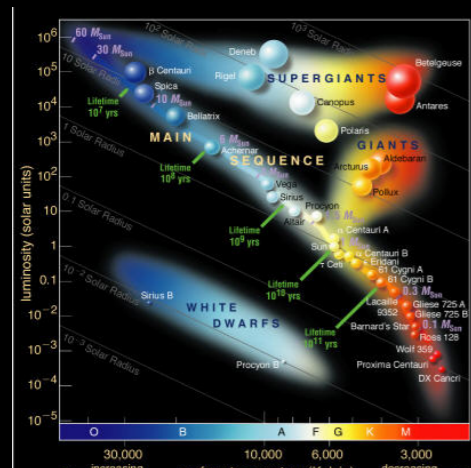
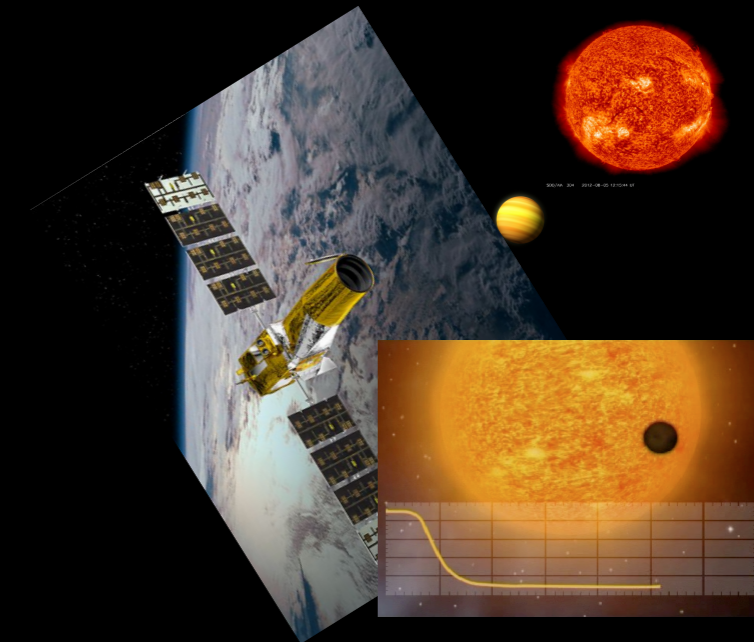
EXODAT

Datamining dans le cas d'exoplanètes observées par CoRoT

COROT

Satellite lancé en 2006, pour 6 ans

But de la mission (exoplanètes) :
Détection d'exoplanètes par méthode de transit.

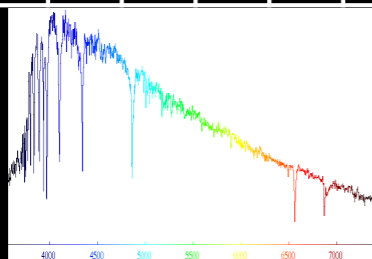


Données EXODAT (<http://cesam.lam.fr/exodat/>)

150 000 étoiles dont :

- 95000 avec une photométrie précise
- 104 000 avec une classification d'activité
- 11500 avec une classification spectral
- 531 avec une signature de transit
- 30 planètes confirmées

U	B	V	R	I	J	H	K
13,5	13,0	12,4	11,6	12,6	13,5	14,6	15,0



Data challenges

Shared challenges: data tsunami

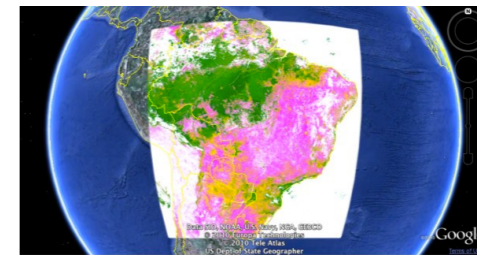


Digital curation and data access

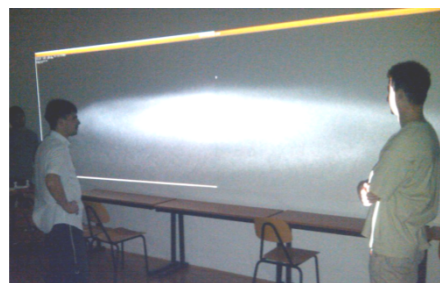
- store, maintain & preserve huge amounts of data
- large **multidimensional & highly interrelated** datasets = paradigm change: **push the computing to the data**

Visualization

- visualizing large quantities of data with: low signal-to-noise ratio, high dynamic range, multidimensional parameter space, multi-layered time-dependent, ...



Google Earth Engine



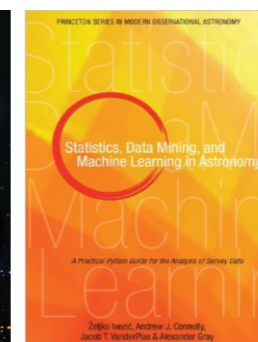
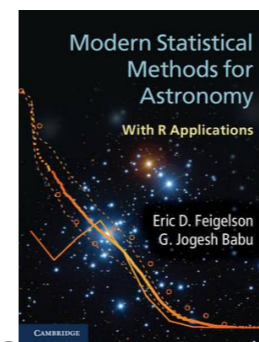
GPUs as numerical co-processors

Adaptation to new high performance computing (HPC) technologies

- heterogeneous supercomputing environments
- new programming techniques for GPUs and cloud computing

Training of a new generation of scientists

- astrophysics, geoinformatics, bioinformatics
- natural sciences + IT/CS = exploration with statistics



New books and online courses

extrait de la presentation COST

Quelques résultats

Rapidement

Analyse de données rares par contre exemple
Data Mining sur les données expérimentales

En ce qui concerne PREDON

Accès aux données
Analyse de données

Analyse des données

- **Le point de vue de l'astronome : analyse**
 - **Catégorisation** des objets (pertinence, erreurs)
 - Détection de **rareté** (découverte, anomalie)
 - Réduction de dimensionnalité
 - **Mesure** de paramètres (régression, statistique)
 - Prise en compte des incertitudes
- **Le point de vue de l'informaticien : fouille**
 - Données sous forme tabulaire (**interface avec SGBD**)
 - **Algorithmique efficace** en dimension élevée (> 1000 , $> 10^{10}$ entrées)
 - Recherche de relations (implications, **dépendances fonctionnelles**)
 - Visualisation des données

Etudes en cours

Exploration interactive de masse de données basée sur les contre-exemples

processus par Itération

- Sélection automatique d'un sous-ensemble des données sur lesquelles l'hypothèse pourrait être vérifiée (apprentissage)
- Reformulation de l'hypothèse
- Prototype iSQL en prod pour les Données ExoDAT (Corot)

Reconstruction des redshifts photométriques

- Passage à l'échelle des méthodes de template-fitting
- Optimisation du nombre de noeuds de la couche cachée d'un RNA à l'aide de méthodes de clustering

• Clustering et énumération

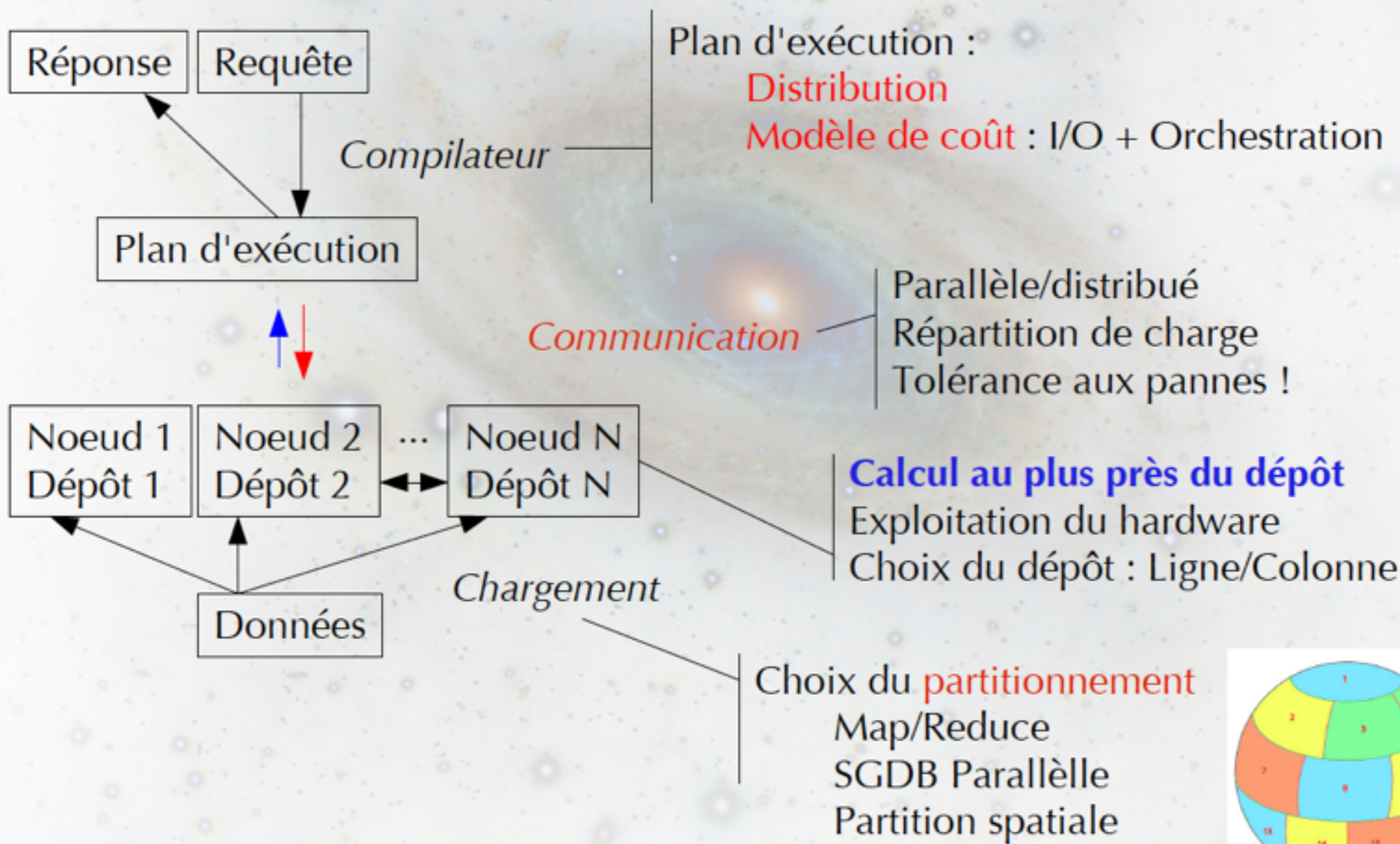
- Algorithmes à une passe pour la recherche de clusters
- Conception d'une nouvelle technique d'énumération incomplète

• Dépendances fonctionnelles

- Algorithme parallèle générique pour l'extraction des DF
- Optimisation des requêtes skyline multidimensionnelles
- Plateforme de visualisation des DF

Preservation et accès des données

- SGBD distribué



Accès aux données

Proposer une architecture distribuée pour stocker quelques dizaines de **PO** de données

- Open Source
- Pouvoir évaluer aussi bien des requêtes simples (quelques **secondes** de calcul) que des requêtes complexes (des **heures** de calcul)
- Possibilité d'accéder à des objets en utilisant des indexes ou en procédant à un parcours (scan) complet des grosses tables (**>> 1 PO**)

Test

Infrastructure

Mise en place d'un environnement d'expérimentation avec 100 machines virtuelles.

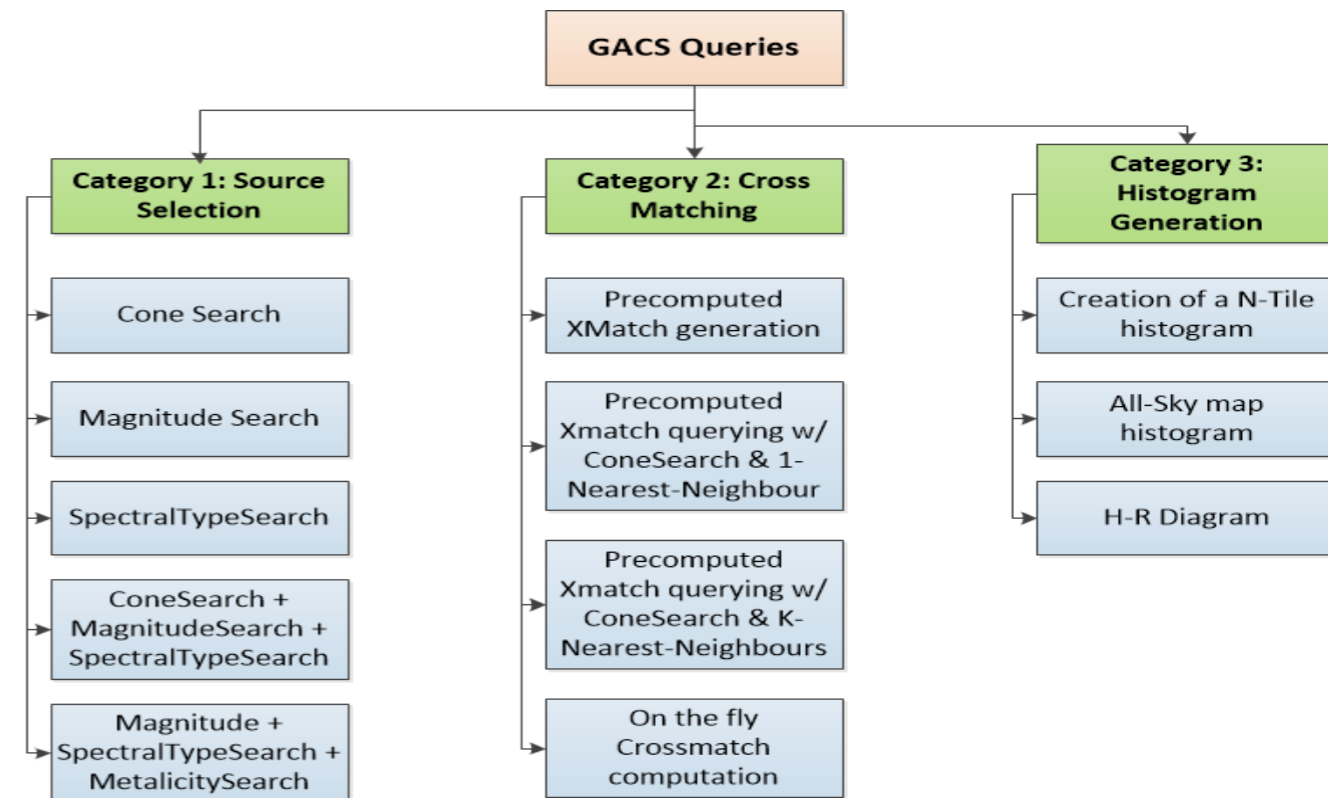
- 350 Go d'espace disque
- 8 Go de RAM
- Systèmes de gestion de données
- **Centralisés**: Mysql, PostgreSQL et DBMS-X
- **Distribués** (à la Map/Reduce): Hive et HadoopDB
- Caractérisation des requêtes selon la difficulté des traitements

Requêtes types

id	Syntaxe SQL
Q1	select * from source where sourceid=29785473054213321;
Q2	select sourceid, ra, decl from source where objectid=402386896042823;
Q3	select sourceid, objectid from source where ra > 359.959 and ra < 359.96 and decl < 2.05 and decl > 2;
Q4	select sourceid, ra, decl from source where scienceccdexposureid=454490250461;
Q5	select objectid, count(sourceid) from source where ra > 359.959 and ra < 359.96 and decl < 2.05 and decl > 2 group by objectid;
Q6	select objectid, count(sourceid) from source group by objectid;
Q7	select * from source join object on (source.objectid=object.objectid) where ra > 359.959 and ra < 359.96 and decl < 2.05 and decl > 2;
Q8	select * from source join object on (source.objectid=object.objectid) where ra > 359.959 and ra < 359.96;
Q9	SELECT s.psfFlux, s.psfFluxSigma, sce.exposureType FROM Source s JOIN RefSrcMatch rsm ON (s.sourceId = rsm.sourceId) JOIN Science_Ccd_Exposure_Metadata sce ON (s.scienceCcdExposureId = sce.scienceCcdExposureId) WHERE s.ra > 359.959 and s.ra < 359.96 and s.decl < 2.05 and s.decl > 2 and s.filterId = 2 and rsm.refObjectId is not NULL;

Comparaison méthodes Map Reduce

	Hadoop [Dean et al, 04]	Hadoop++ [Dittrich et al, 10]	HAIL [Dittrich et al, 12]	HadoopDB [Abouzeid et al, 09]	Hive [Thusoo et al, 10]
Query language	procedural	Procedural	Procedural	Declarative	Declarative
Simple indexes	Not supported	supported	Supported	supported	Supported
Multi indexes	Not supported	Not supported	supported	Supported	Supported
Complex indexes	Not supported	Not supported	Not supported	supported	Supported
Storage System	HDFS	HDFS	HAIL	Classical DBMS	HDFS
Index choice	-	Manual	Manual	Automatic	Manual

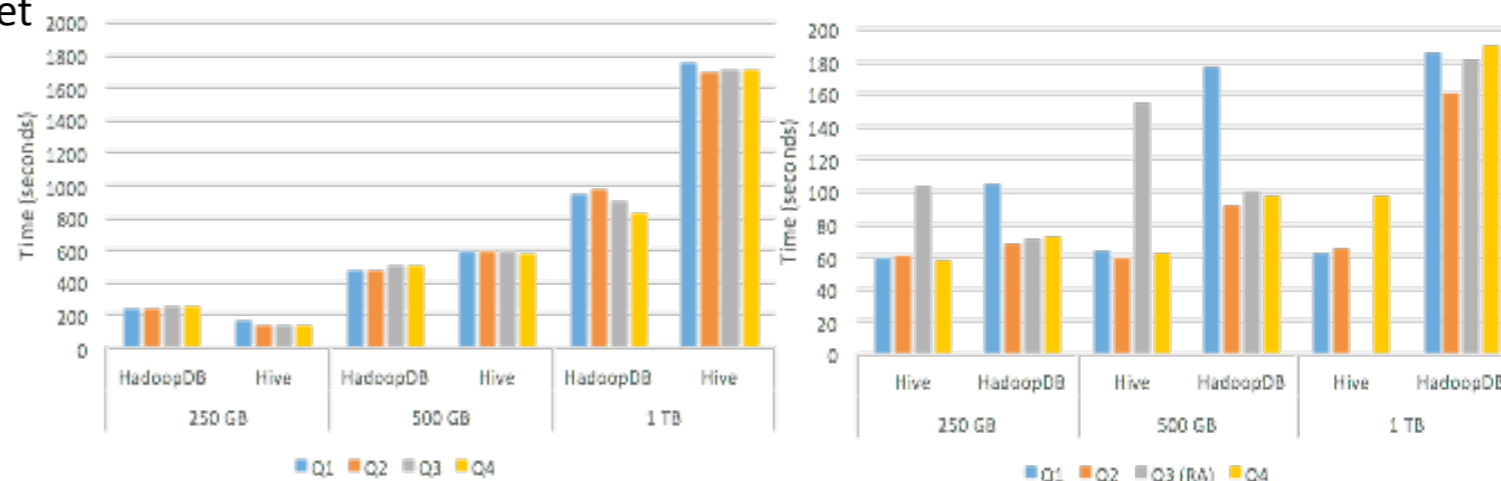


Résultats

Différence (Hive >> HadoopDB): 3X pour 25 machines et 2X pour 50 machines

- Partitionnement personnalisé pénalise HadoopDB
- Hive:
 - 2X de temps pour 2X de volumes de données
 - 25 >> 50 machines: même temps
- HadoopDB:
 - 2X de données: 90 %-120 % supplémentaires
 - 25 >> 50 machines: 25 % gain

Query processing Hive vs. HadoopDB (2/6)



Predicate	D250GB	D500GB	D1TB	D2TB
SourceID =id	1	1	1	1
ObjectID =id	43	43	43	43
$2 < DECL < 2.05$	1.6×10^5	3.3×10^6	6.6×10^6	13.2×10^6
$359.959 < ra < 359.96$	14×10^3	28×10^3	57×10^3	127×10^3
$359.959 < ra < 359.96$ & $2 < decl < 2.05$	21	43	86	172
ScienceCcdExposureId = id	3.6×10^3	7.3×10^3	14.6×10^3	29.2×10^3

Remarque 1: Hive ne supporte pas l'utilisation d'index pour évaluer les requêtes de jointure

- Remarque 2: HadoopDB ne supporte pas plus d'une jointure
- Hive vs HadoopDB
- Hive est nettement meilleur que HadoopDB même lorsque HadoopDB utilise l'index
- Hive parallélise le traitement des jointures
- HadoopDB utilise une seule machines pour construire les résultats de la jointure

Conclusions

HADOOP n'est pas encore la solution finale

Chargement des données assez long
Dépendant du partitionnement des données
Nécessite un « tuning important » et spécialisé

Besoin en Répartition spatiale des bases distribuées

- des demandes qui sont basées sur des cones search
- des cross match
- des densités

Approches à l'étude

- Hadoop GIS
- Spatial hadoop
- Pigeon
- GeoSpark

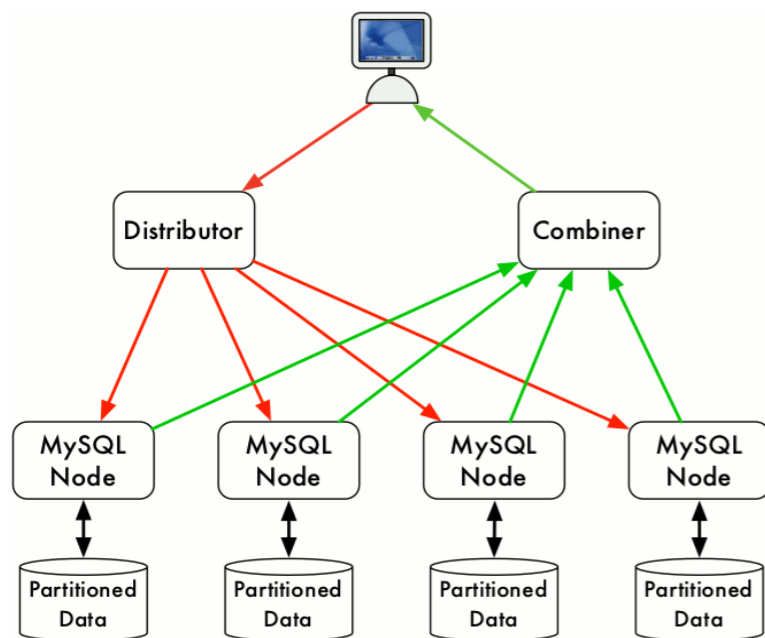
Quelques réponses

- Pas de concordance avec des systèmes orienté GIS
 - Pas de géométrie sphérique implémentée
 - Spatial Hadoop offre un rapport 1/100 par rapport à Hadoop mais est limité en requêtes
 - GeoSpark est prometteur

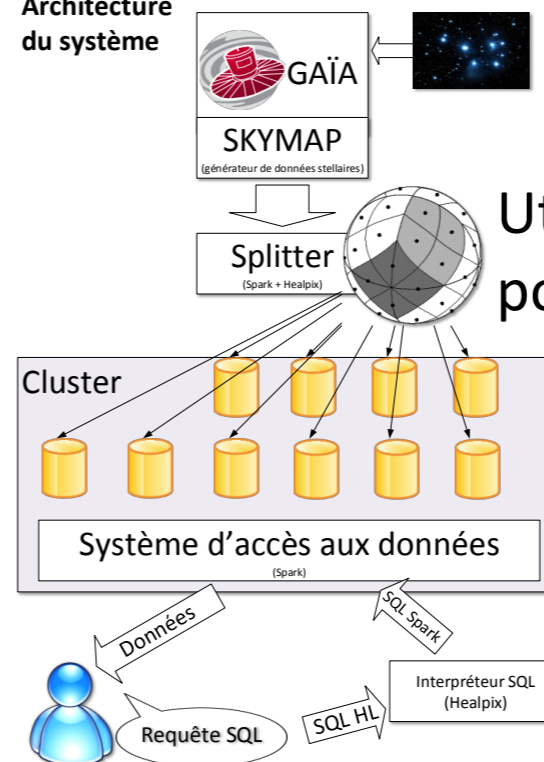
- Extension de SPRK dans un framework QServ

<http://ls.st/lpj> (Database Design doc)

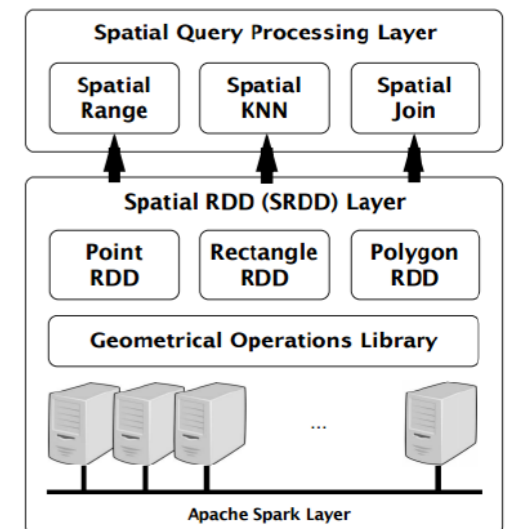
<http://ls.st/6ym> (User Manual)



Architecture du système



Utilisation de tessellation Healpix pour partitionnement des requêtes



Finalemment

Les algorithmes basés sur Map reduce peuvent permettre une implémentation optimisée des requêtes

Les systèmes hybrides de stockage par colonne/tuples sont intéressant à étudier pour la répartition des requêtes

Nécessité de travailler sur :

L'abstraction des données astro

l'optimisation des techniques (partitionnement, tuning, cache)

Prendre en compte les requêtes

Prendre en compte la mise à jour des information pour optimisation

[1] Mesmoudi, Amin, Mohand-Saïd Hacid, and Farouk Toumani. **Benchmarking SQLon MapReduce systems using large astronomy databases**. Distributed and Parallel Databases (2015): 1-32.

Production scientifique :

16 publications, 7 séminaires invités, 5 co-encadrement doctoraux ...

PetaSky a servi de levier pour

- Projet européen **COST BIG-SKY-EARTH** : Big Data Era in Sky and Earth Observation
- 2 membres de PetaSky sont au Management Committee de l'action
- Proposer une action structurante au sein du **GDR MaDICS**
- 2 projets-projets ANR
- L'intégration de LSST et des Big Data dans le projet AUDACE du **CPER 2015-2020 de la région Auvergne**