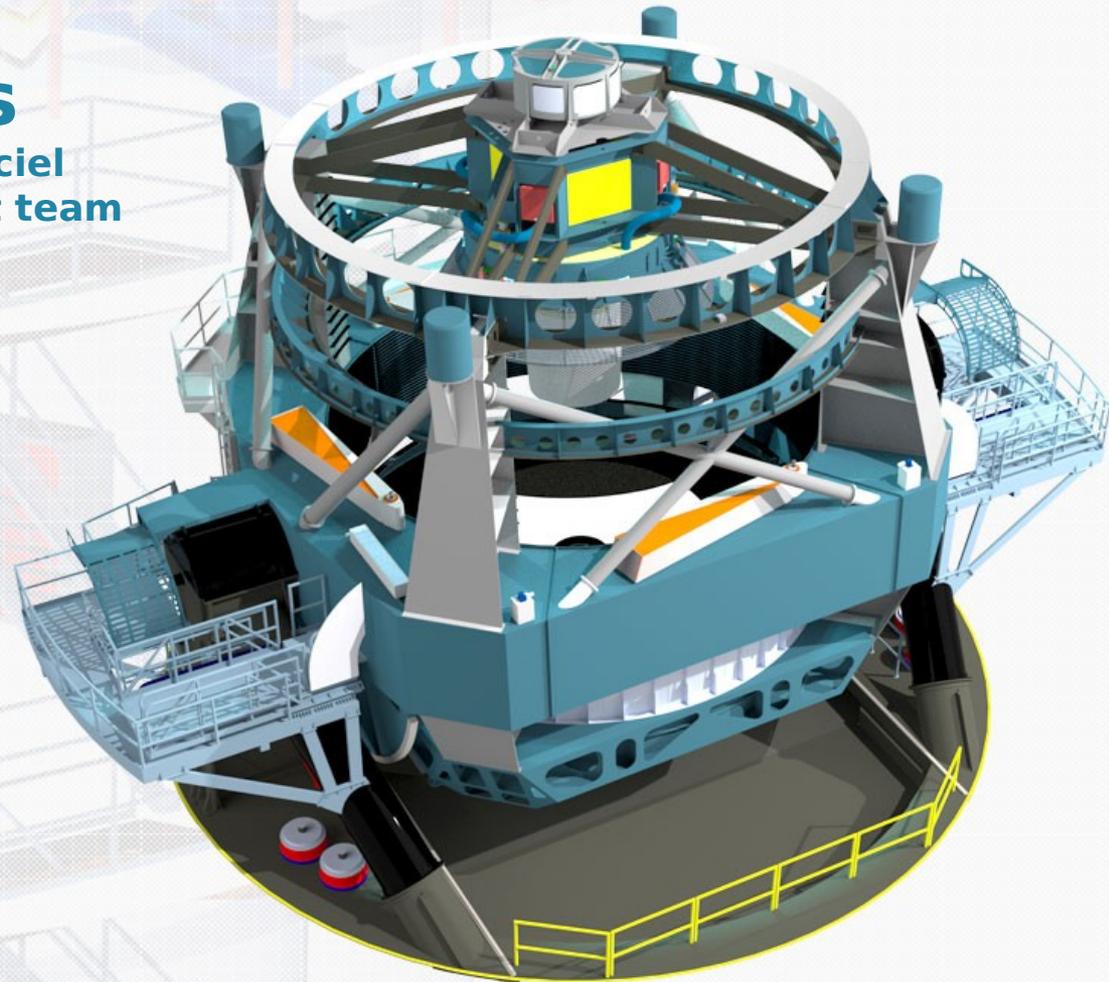


# Analyse de données à l'échelle du PetaOctet avec Qserv

**Fabrice Jammes**

Expert en développement logiciel  
IN2P3/LSST Data-management team

**Avril 2015**



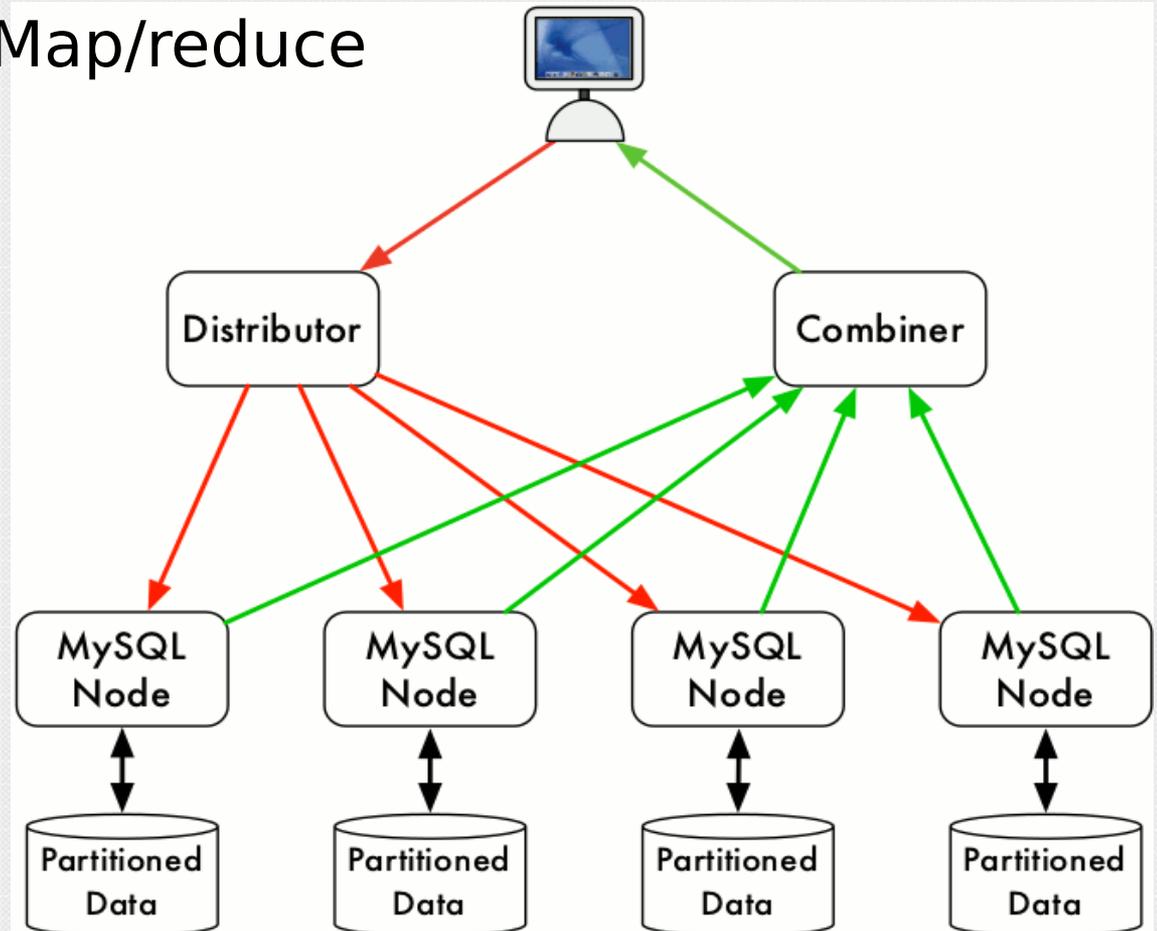
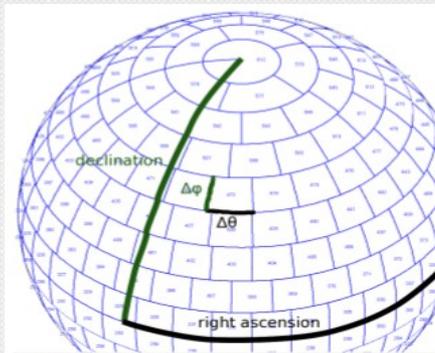
## images

- ~38 PB (persistantes)
- ~ $\frac{1}{2}$  EB (virtuelles)

- ~3 milliards de visites
- ~37 milliards d'objets

- Plus grande table: ~5 PB
- Plus haute table: ~50 billions de lignes
- Total (toute data releases, compressé): ~83 PB

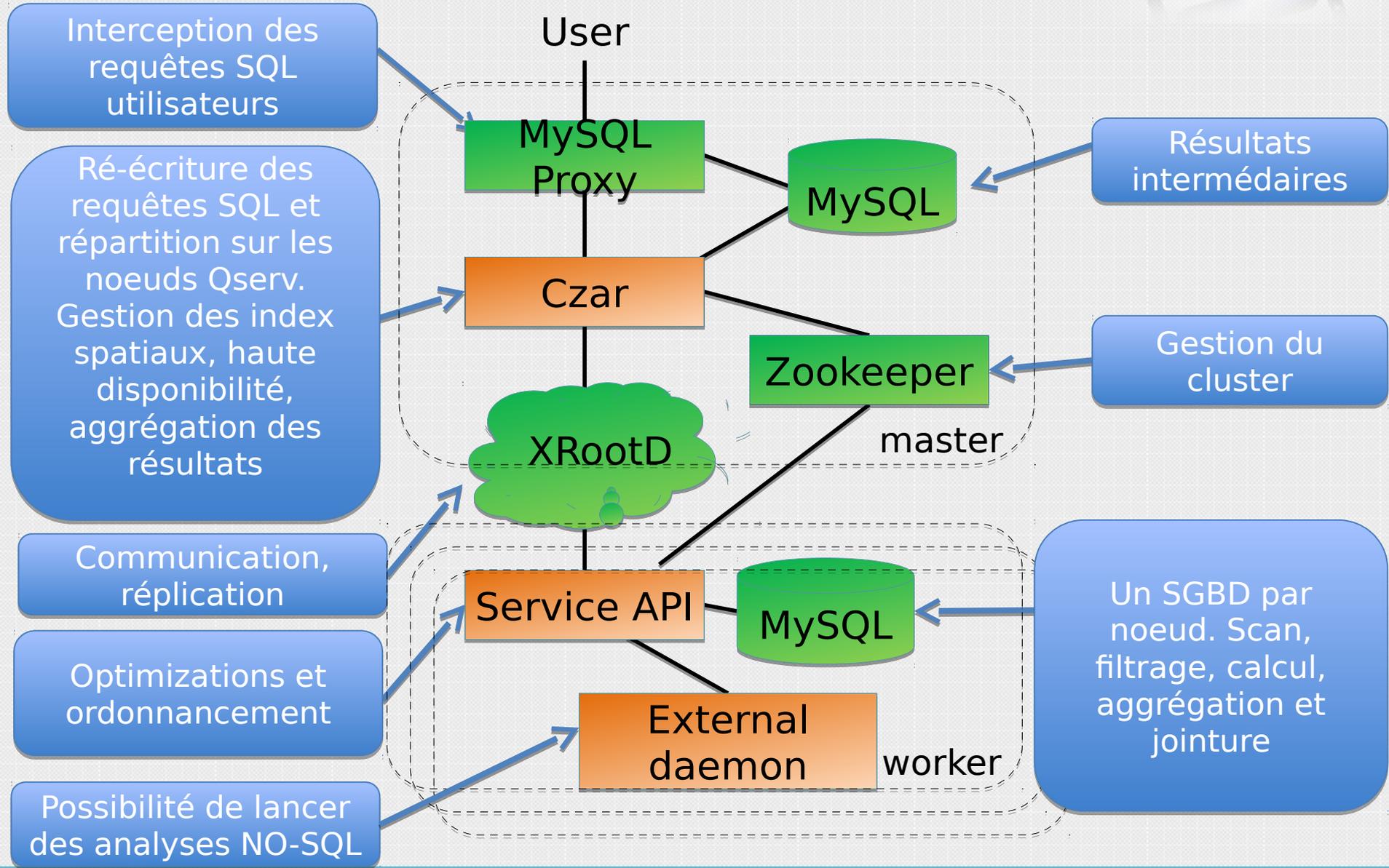
- Base de données relationnelle, partitionnée spatialement avec des recouvrements
- Modèle de calcul Map/reduce





- Pas de produit sur étagère disponible
- SLAC développe Qserv, avec l'aide de l'IN2P3
  - Basé sur des composants existants intégrés de manière spécifique
- 100% open source

# Architecture technique



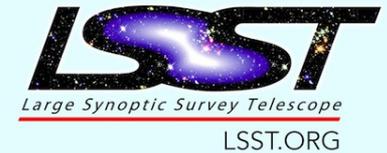


- Géométrie sphérique performante
  - Gestion de la distortion au pôles, et de la recherche par polygones convexes, calcul de distance précis, test d'appartenance à une région sphérique (cercle, ellipse, boîte, polygone convexe)
  - UDFs spécifiques <https://github.com/wangd/scisql>
- Jointures spatiales optimisées pour les recherches par voisinage
  - Partitionnement sphérique avec recouvrement
  - 2 niveau, second niveau matérialisé à la volée
- Scan partagé des données
  - Lecture continue et séquentielle à travers les données
  - (Non-interactif) requêtes attachées à un scan spécifique du disque
- La complexité interne est transparente pour les utilisateurs



- Prototype fonctionnel, utilisable, mais pas encore à l'épreuve des balles
- Sous-ensemble de fonctionnalités non implémentées (sous-requêtes)
- Chargeur de grandes masses de données en cours de validation
- Version optimisée de xrootd intégrée au produit
- Cluster de développement de 50 noeuds au CC-IN2P3 (offert par DELL )
  - => Chargement de 400 TB de données en cours
- Contributeurs extérieurs bienvenus:  
<http://www.slac.stanford.edu/exp/lstt/qserv/>

# LSST IS HIRING



WE'RE SEEKING TOP TALENT TO WORK IN A TEAM ENVIRONMENT THAT INSPIRES EXCELLENCE.



## JOIN US IN:

LSST HEADQUARTERS  
TUCSON, AZ

SLAC/STANFORD  
MENLO PARK, CA

PRINCETON UNIVERSITY  
PRINCETON, NJ

NCSA / UIUC  
URBANA-CHAMPAIGN, IL

UNIVERSITY OF WASHINGTON  
SEATTLE, WA

LSST OBSERVATORY SITE  
CERRO PACHÓN, CHILE

## ABOUT US

LSST IS A PUBLIC-PRIVATE PARTNERSHIP AND THE TOP-RANKED LARGE-SCALE GROUND-BASED PROJECT FOR THE NEXT DECADE AS RECOMMENDED BY THE NRC'S ASTRO2010 DECADAL SURVEY. LSST WILL SCAN THE SKY FOR 10 YEARS, PRODUCING A PETABYTE-SCALE, NON-PROPRIETARY DATABASE DESIGNED TO ADDRESS THE MOST PRESSING QUESTIONS IN ASTRONOMY AND PHYSICS, WHILE DRIVING ADVANCES IN BIG-DATA SCIENCE AND COMPUTING.

## CAREERS

LSST IS A NEW PARADIGM FOR LARGE SCIENTIFIC FACILITIES: OPEN SOURCE, OPEN DATA, AND AN OPEN, FLEXIBLE WORK ENVIRONMENT. ALL LSST WORK SITES OFFER EXCEPTIONAL BENEFITS PACKAGES AND ROOM FOR PROFESSIONAL GROWTH.

## OUR TEAM

- PROJECT OFFICE
- DATA MANAGEMENT
- TELESCOPE & SITE
- CAMERA
- EDUCATION & PUBLIC OUTREACH
- SYSTEM ENGINEERING

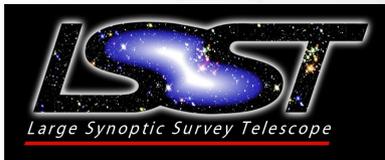


[LSST.ORG/HIRING](https://lsst.org/hiring)



# PetaSky

Programme de la MI du CNRS : mégadonnées (Mastodons → GDR Madics)  
dans le contexte astrophysique

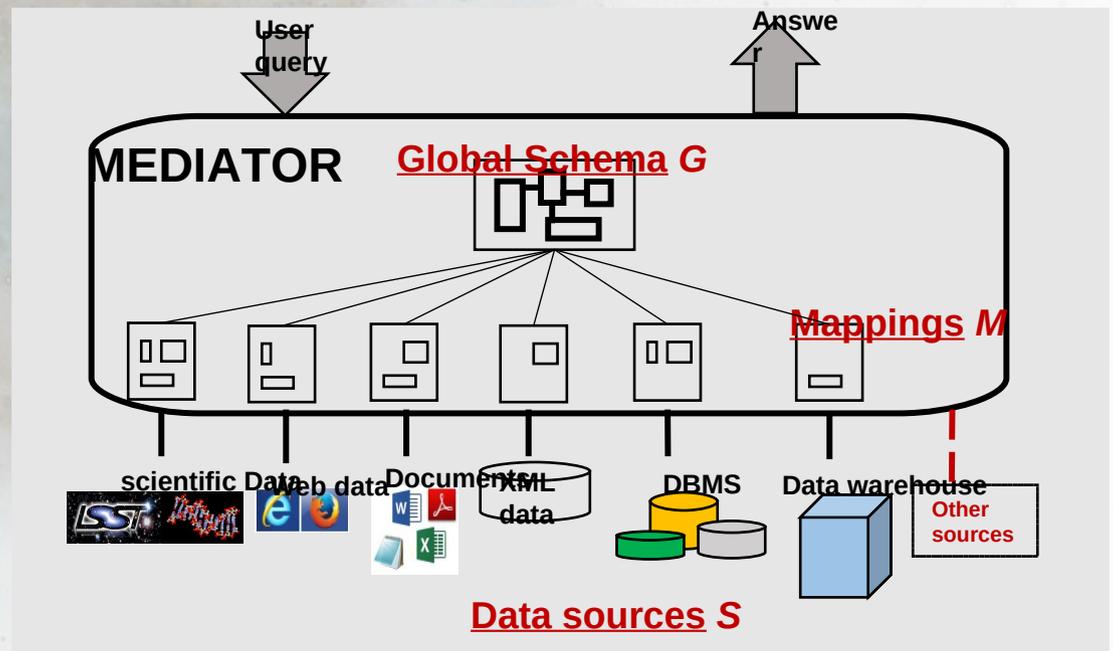
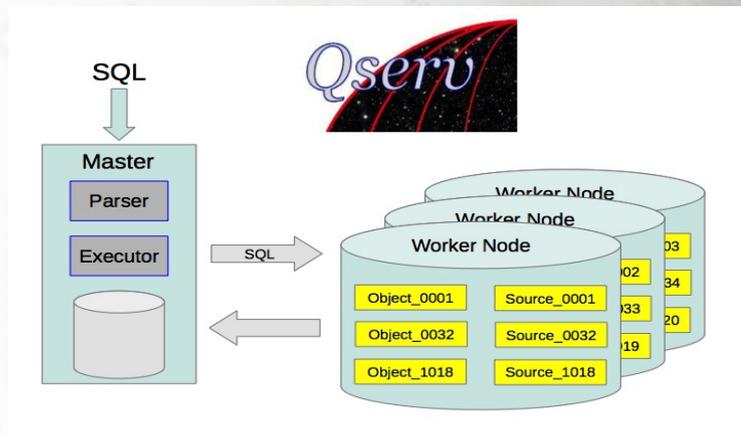


- IN2P3
  - LPC, CCIN2P3 → Qserv, Bases de données
  - APC → photo-z
  - LAL → visualisation d'images
- INSU
  - LAM → Exoplanètes, Euclid
  - OBSPM, U. Besançon → GAIA (Voie Lactée, Astéroïdes, CU9)
- INS2I
  - LIMOS, LIRIS, PriSM → Base de données, fouille
  - LIF, LIRMM → fouille visuelle
  - LaBRI → algorithmique
- INSMI
  - Ceremade, LPP → statistique

~ 45 participants

# R&D Bases de données

- **Qserv** : baseline de référence
  - benchmarking, analyse des limitations
- Contexte théorique des  **systèmes d'intégration**



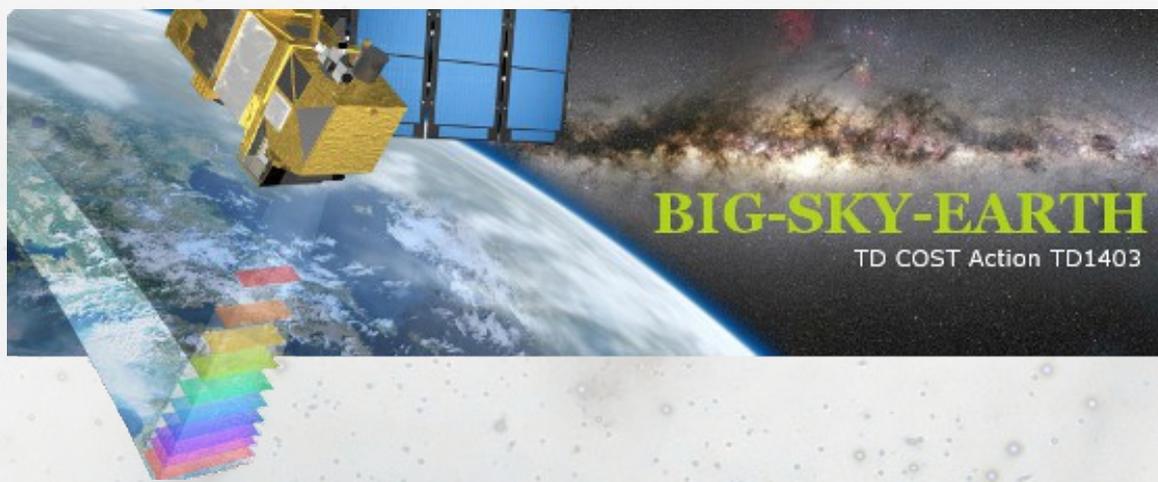
→ prototypage (InTheMidst), étude des optimisations

# R&D Bases de données

- *Approches exploratoires*
  - Modèle relationnel pour LSST ?
    - difficulté d'insertion, valeurs peu accédées
  - **Map/Reduce** (HadoopDB, Hive, ...)
    - Map : pousser les traitements vers les données
    - Reduce : pousser les données vers les traitements
    - benchmarking : mise en évidence de limitations
      - Optimisation complexe
      - Coût réseau
      - Division en 2 traitements
  - **BSP** (Bulk Synchronous Parallel)
    - Méthode basée sur les graphes
      - Algorithmique en cours de développement
- Mise en place d'une *plateforme de test*

# Fouille de données

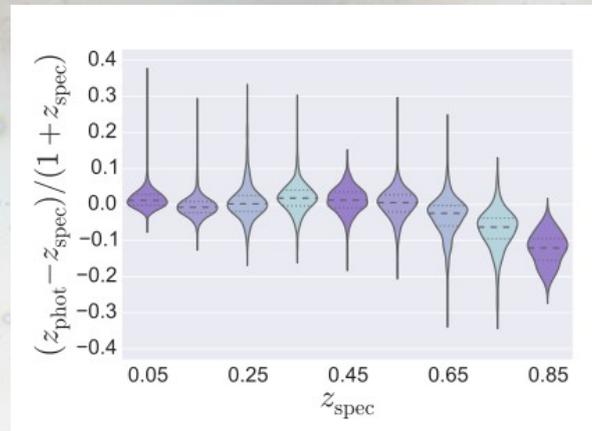
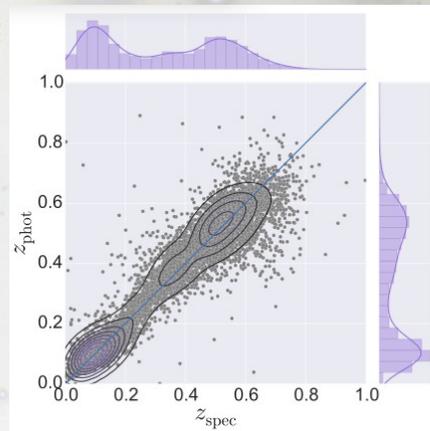
- **Photo-z** *comme cas d'école*
  - Quelques travaux préliminaires côté info
    - Convergence culturelle lente...
  - Neural network → **Deep Learning**
- *Algorithmiques passant à l'échelle*
  - Ou comment ne pas regarder tout, mais seulement ce qui est intéressant, au prix d'approximations
    - **Fouille guidée par l'expert** (généralisation d'hypothèses ou de relations)
    - **Abstraction des données** (représentations résumées )
  - encore préliminaire ...



- *Consortium européen Big Data Astro et Géo*
  - Kick-off 2015...
  - 24 pays, 3 communautés...
    - un peu auberge espagnole, mais...
  - Contacts avec *Astrostatistique*, communauté des *VO*
  - Présence de *LSST au niveau européen*
    - Anglais 120 PI (Victor Debattista)
    - Serbie 4 PI (Darko Jeremovic) → transients de LSST (AlertSim)
    - Espagne (Johan Knappen)
    - Allemagne (qq dizaines de PI, non présents dans le COST)
    - + autres ... (Hongrie, Coratie, Pologne, R. Tchèque, Slovénie, Autriche)
  - *Workshops et Training Schools*



- Initiative hébergée par l' IAA (International Astrostatistics Association)
- ~30 participants, surtout post-docs 
- *Statistique* (Approximated Bayesian Computation, Generalized Linear Models, Gaussian Mixtures...)
- *Appliquée à la cosmo* (photo-z, estimation des paramètres cosmo, supernovae, galaxy clusters, ...)
- Ex : « The Overlooked Potential of Generalized Linear Models in Astronomy-II: Gamma regression and photometric redshifts » arXiv 1409.7699



- Code rapide (quelques secondes sur laptop), ~2% catastrophiques (SDSS)

# En guise de conclusions

- **Collaborations pluridisciplinaires**
  - *Ouverture sur d'autres communautés*
    - Temps d'acculturation important
    - Enrichissement scientifique (méthodes, outils)
    - Intérêt pour les données LSST
  - *Retour d'expérience*
    - Modèle de travail
      - Faire nos analyses, tirer parti des idées des autres.
      - Guider la conception des outils (via la nature des données)
    - Moyens associés (missions, RH sur appels d'offre...)