# Calcul sur grille pour l'astronomie Gamma

Mathieu de Naurois (+ de nombreuses contrib.)
LLR - Ecole Polytechnique
denauroi@in2p3.fr
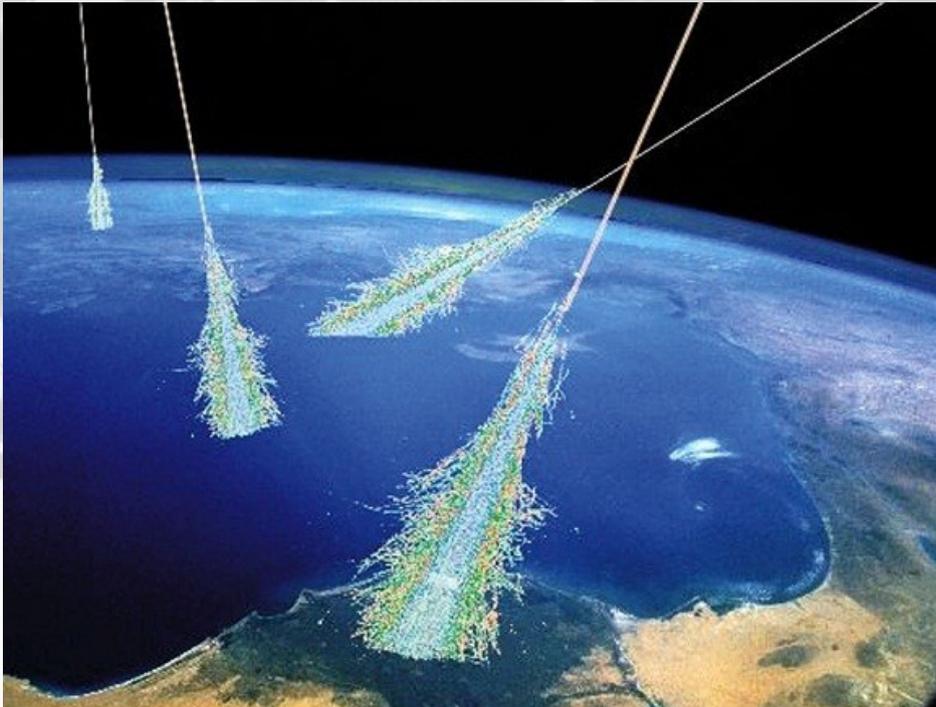
- ☞ *Contexte*
- ☞ *Outils*
- ☞ *Ressources*
- ☞ *Performances*
- ☞ *Perspectives*

Contexte

Mathieu de Naurois
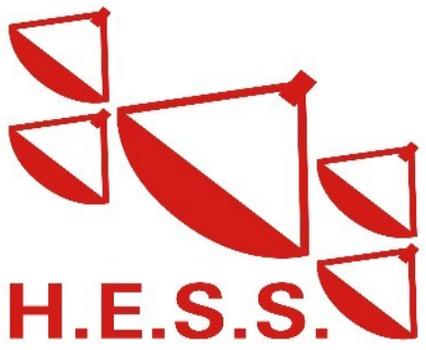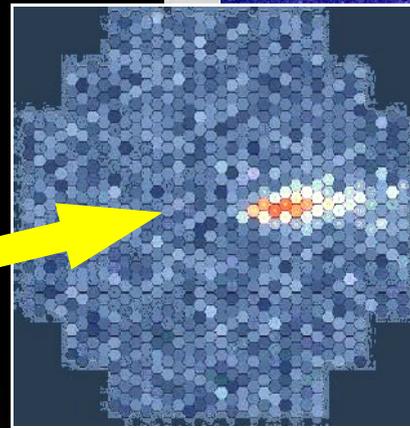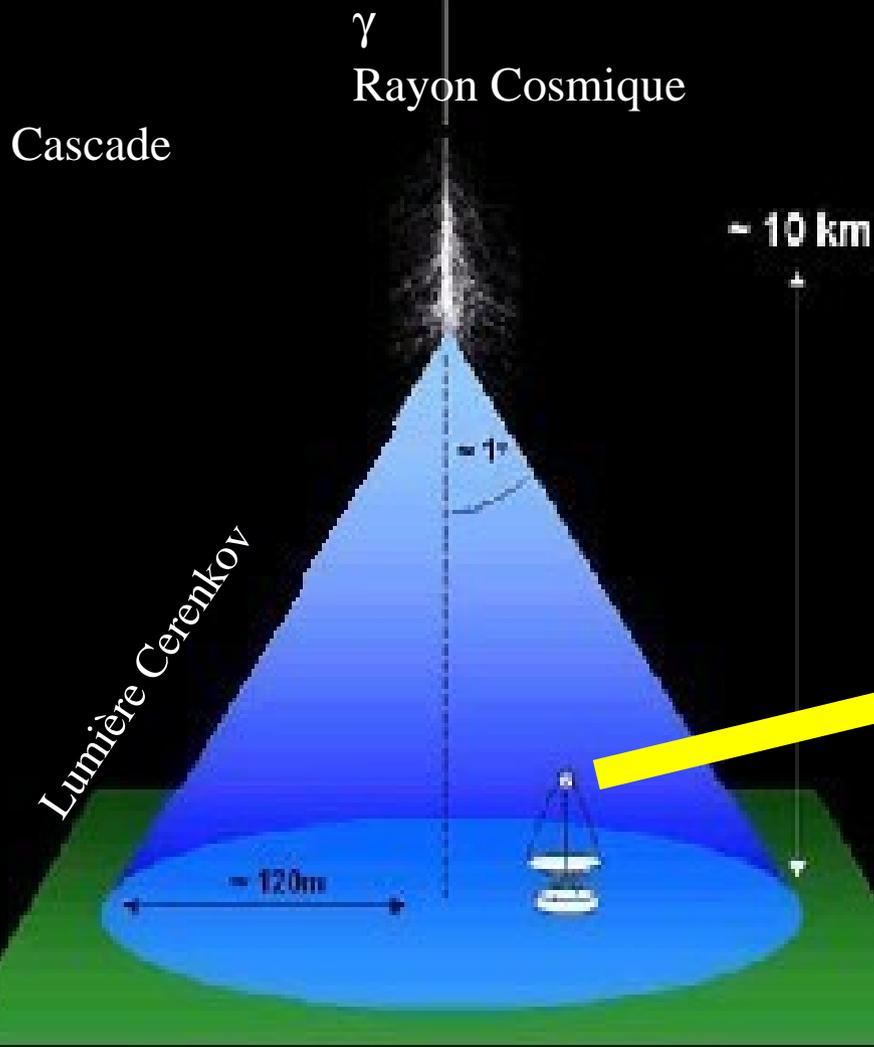
Sky & Telescope

- *Consortium international, mené par Allemagne + France*
- *4 télescopes de 14m + 1 télescope de 28m au centre du réseau*
- *Site: 23°16" S, 16°30" E, 1800 m asl, 100 km de Windhoek (Namibie)*
  - *très bonne qualité optique du ciel*
 - *hémisphère peu observé, grande partie du plan galactique observable*

Gamsberg
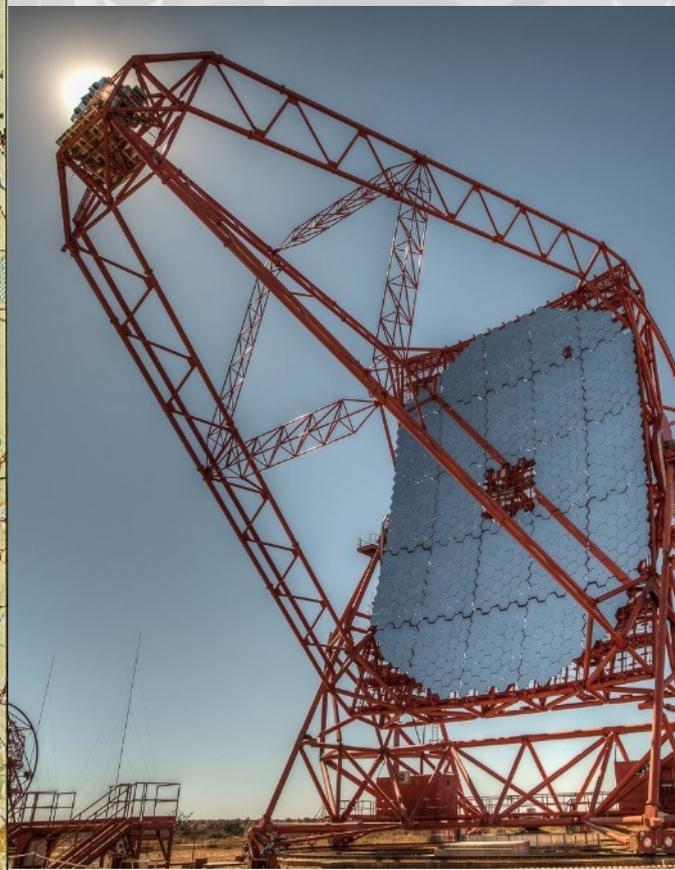
MPI Kernphysik, Heidelberg
Humboldt Univ. Berlin
Ruhr-Univ. Bochum
Univ. Hamburg
Landessternwarte Heidelberg
Ecole Polytechnique, Palaiseau
APC Paris
LPNHE Univ. Paris VI-VII
CEA Saclay
CESR Toulouse
LPTA Montpellier
LAOG Grenoble
LAPP Anecy
Observatoire de Paris
Durham Univ.
Dublin Inst. for Adv. Studies
Charles Univ., Prag
Yerewan Physics Inst.
North-West Univ., Potchefstroom
Univ. of Namibia, Windhoek

Mathieu de Naurois,

- *~ 1000 h d'observation par an*
- *10 GB → 100 GB / heure*
- *~ 400 TB d'archive à Lyon*
- *Traitement majoritairement effectué au CCIN2P3, ~ 150 coeurs en permanence*
- *Quelques % de la capacité totale de Lyon*

CTA - Cherenkov Telescope Array
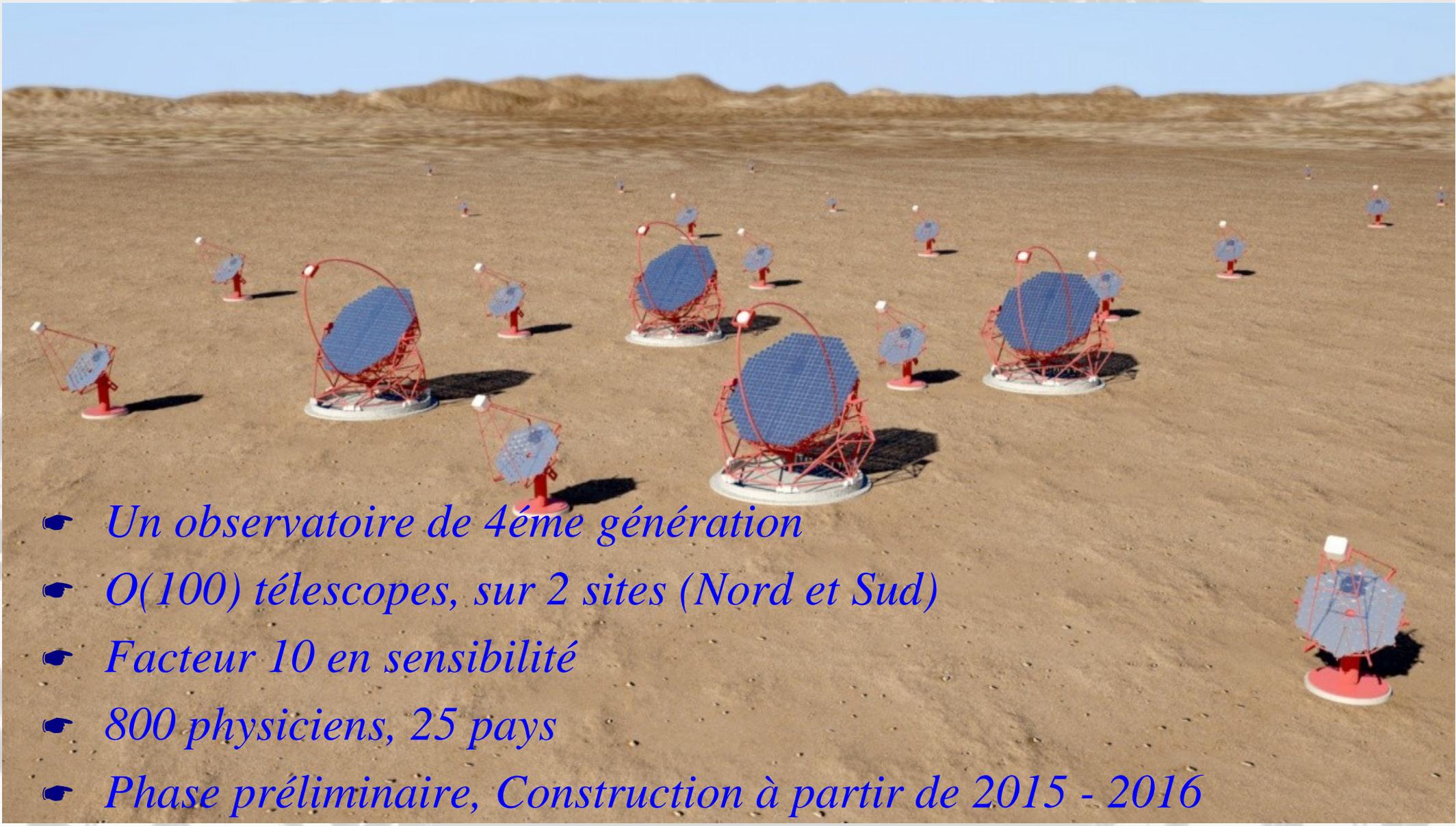
- ☛ *Un observatoire de 4éme génération*
- ☛ *O(100) télescopes, sur 2 sites (Nord et Sud)*
- ☛ *Facteur 10 en sensibilité*
- ☛ *800 physiciens, 25 pays*
- ☛ *Phase préliminaire, Construction à partir de 2015 - 2016*

- *Accès ouvert (partiellement) aux données*
- *Outils de haut niveau, standards astrophysiques, communauté mondiale*
- *Volume de données : PB/an (Taux 0.4 - 5.3 GB/s)*
- *~ 1000 cœurs à plein temps ou 10 000 à 10 % de temps (intermittence)*
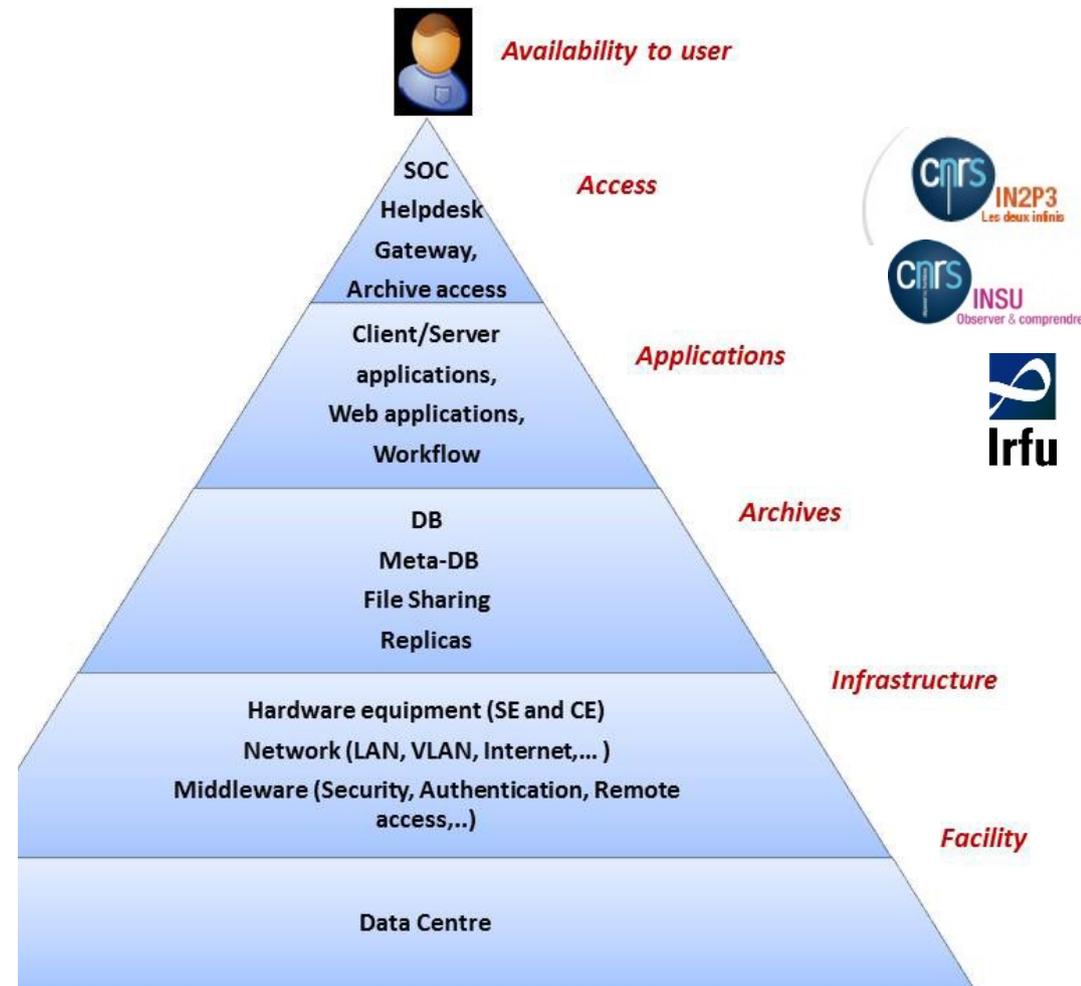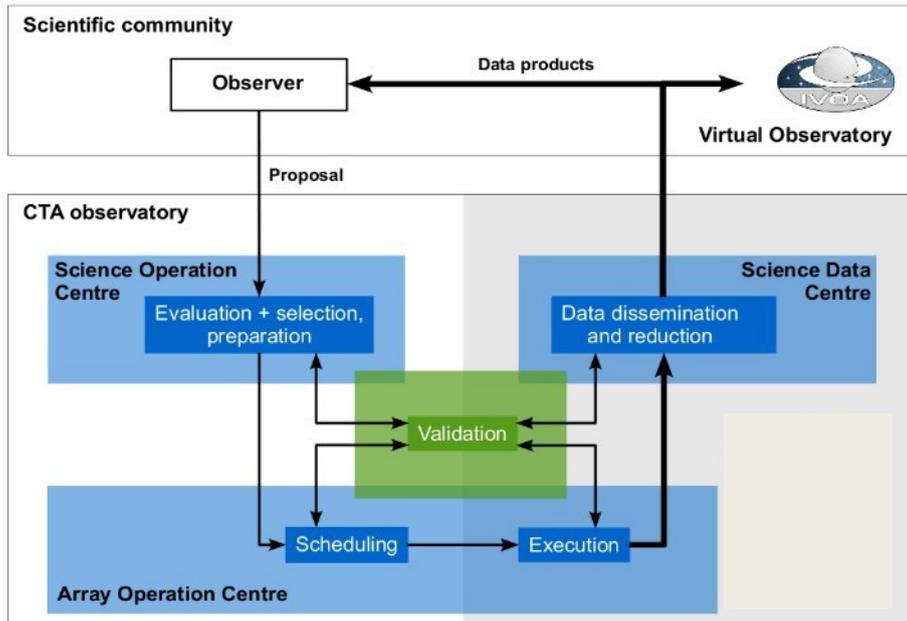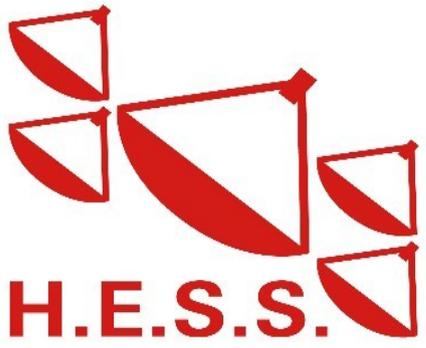
Outils

- *Développement d'un middleware propriétaire en python*
    - *Gestion des systèmes de batch : Grid Engine, Grid (glite), Sun, Unix Batch, ... permettant d'utiliser des ressources hétérogènes*
    - *Interfacé sur glite*
    - *Gestion des accès aux fichiers : HPSS, DISK, XROOTD, SSH, SRM, LFC, ...*
    - *Intégré aux logiciels HESS et à leurs interfaces graphique (pour une configuration plus aisée)*
    - *Implémentation récente des catalogues LFC*
- *Base de données mysql (bookkeeping) + interface php*

☛ *Configuration au CCIN2P3*

☛ *Configuration sur EGI*

☛ *Soumission via WMS ou directement sur CE.*

☛ *Interfacage avec les outils glite*

Interopérabilité

☛ *Utilisation de fichiers générés sur la grille depuis un job batch à Lyon:*

Bookkeeping en ligne

☛ *CTA évalue depuis 2011 le framework DIRAC, développé originellement pour LHCb*

  ☛ *Installation spécifique (dirac0x.pic.es + CCIN2P3), gère la soumission des jobs et le monitoring, (python API), redondance, load management & fault tolerance*

  ☛ *Book-keeping des jobs (Web portal), ...*

  ☛ *Catalogue de fichiers*

Portail DIRAC

Ressources

Mathieu de Naurois

Ressources H.E.S.S.

- *Computing Elements (CE):*
  - *GRIF: APC, LLR, CEA, LPNHE, LAL, IPNO*
  - *→ 9425 slots CPU, partagés entre VO.*
  - *+ LUPM récemment*
- *Storage Elements (SE):*
  - *GRIF (LAL, LLR): 100 TB*
  - *CCIN2P3: 2.1 TB de cache (dCache), produits migrés automatiquement sur HPSS.*
- *→ Pratique pour utiliser produits grilles sur EGI et en batch depuis Lyon (transparent pour l'utilisateur) !!*
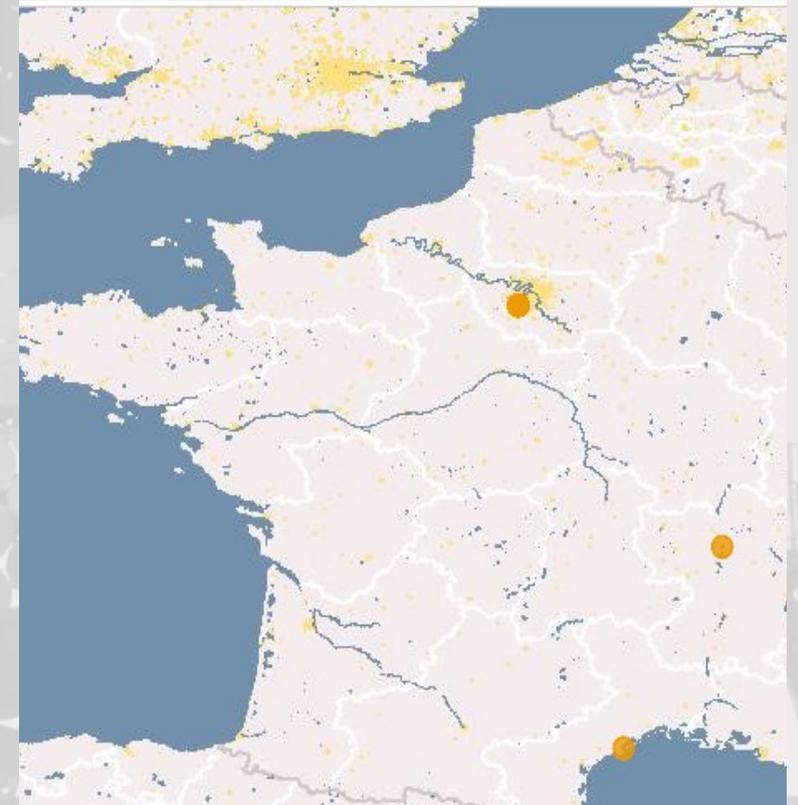
- *17 sites sur 6 pays (vo.cta.in2p3.fr)*
- *18k CPUs, partagés avec d'autres VOs*
  - *1000 – 2000 CPUs pour CTA*
  - *Stockage: plusieurs TB sur chaque site, ~500TB avec les bandes*
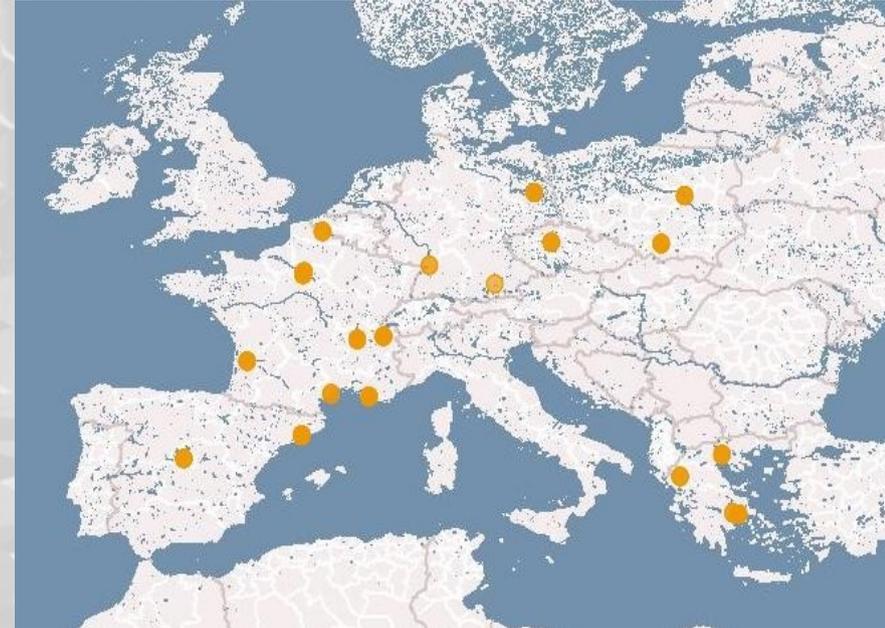
	available CPUs	for CTA	proton simulations
IN2P3-LAPP	925	200	yes
IN2P3-CC	3992	200...600	yes
GRIF-IRFU	1472		no
GRIF-LAL	1229	200	yes
GRIF-APC	100	80	yes
GRIF-LLR	988		small jobs
GRIF-LPNHE	372		no
LPTA, Montpellier	104	100	small jobs
MPPMU	856		no
DESY-ZN	700	400	yes
UNI-DORTMUND	1832		no
PIC	2304	200	yes
CIEMAT-LCG2	0		no
CYFRONET-LCG2	3324	several 100	yes

Mathieu de Naurois,   *LLR*

Performances

Mathieu de Naurois

☞ *http://gstat.egi.eu/gstat/vo/vo.hess-experiment.eu/*
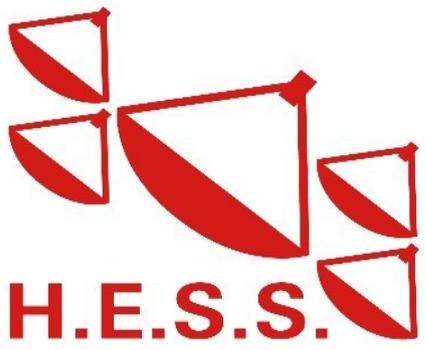
 ☞ *Jusqu'à 3k jobs en pic, ~300–500 jobs en moyenne*

Décembre 2012
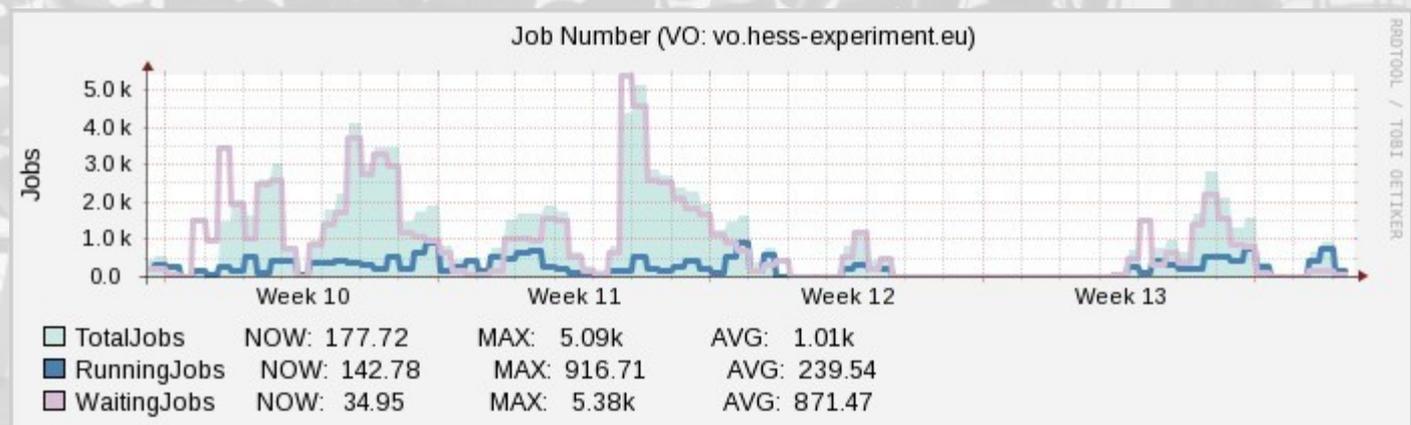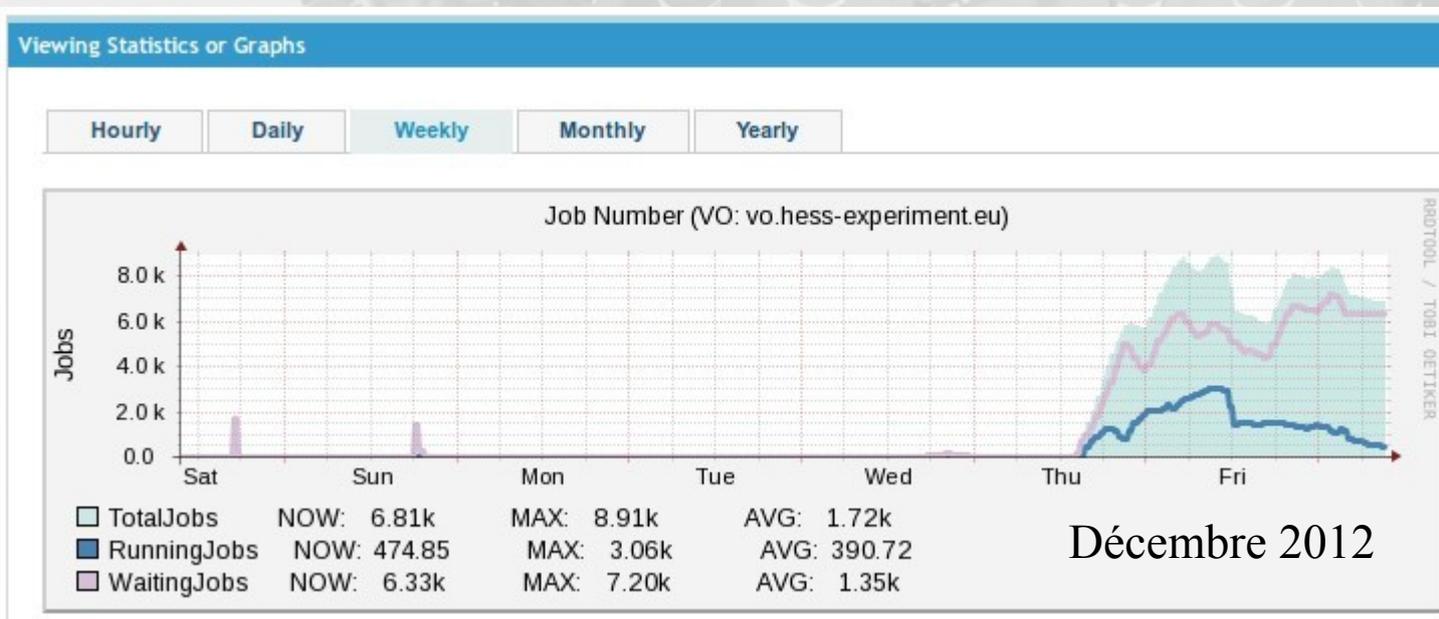
☛ *Distribution des fichiers (prod):*
En 1.5 mois (14/02/2013 → 04/04/2012):
> 150.000 jobs
~70 TB générés

File size distributions

- pSpectrum_paris_0-8-24 - Total=37.55 TB; nb of entries=50464.
- gSpectrum_paris_0-8-24 - Total=19.95 TB; nb of entries=54629.
- eDiffuseSpectrum_paris_0-8-24 - Total=14.13 TB; nb of entries=48094.
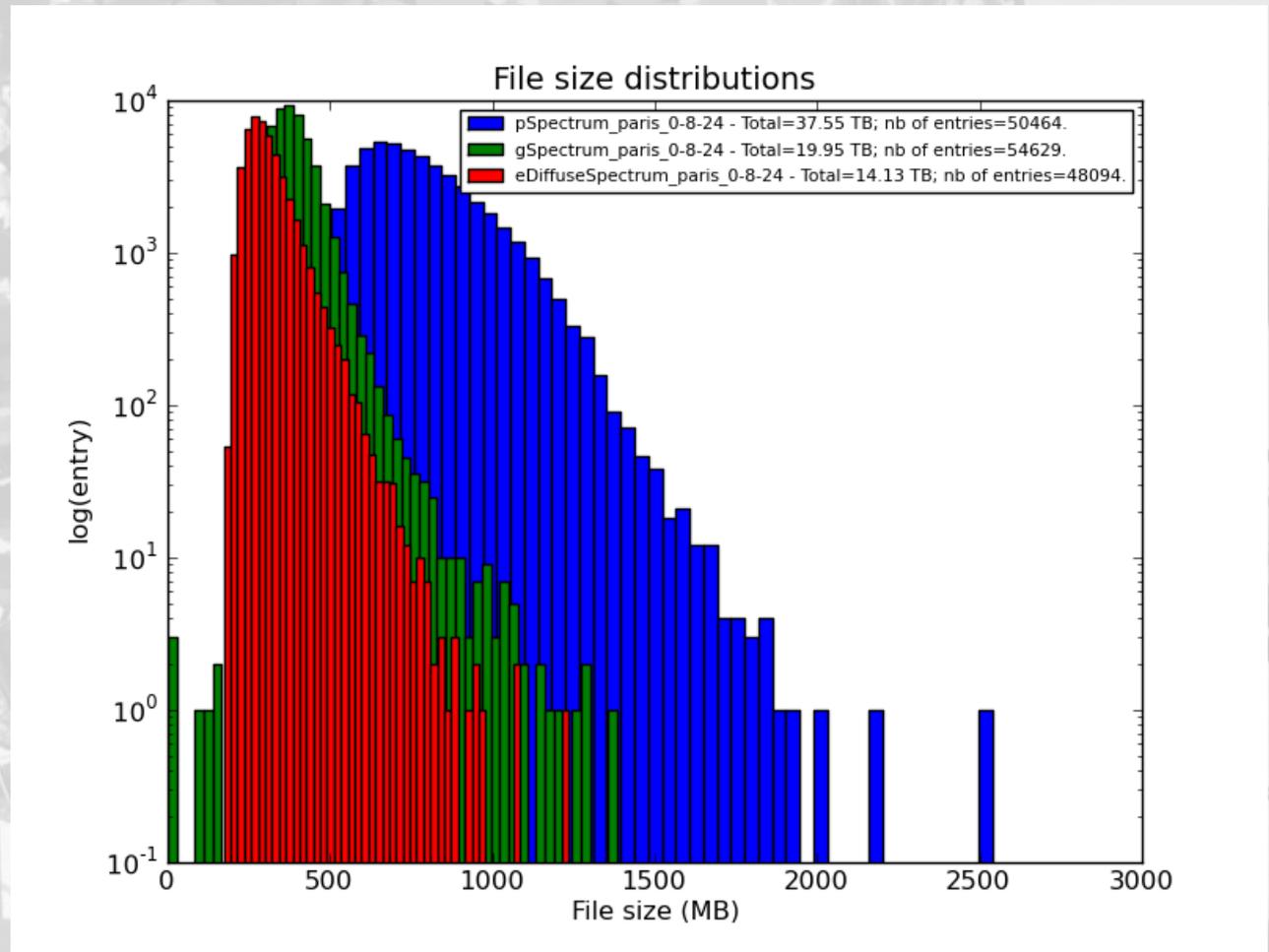
☞ *Premier retour d'expérience :*

- ☞ *Simulation, reconstruction, analyse*
- ☞ *Certaines tâches (simulation) relativement aisées*
- ☞ *D'autres nécessitent de nombreuses étapes avec des fichiers intermédiaires => complique la tache, de nombreux modules DIRAC interdépendants*

☛ *Test grandeur nature : réponse de l'instrument en fonction de la position de la source*

- ☛ *4 semaines, 180 000 jobs, 97 % de taux de succès*
- ☛ *400 TB processés*
- ☛ *12 TB de données produites*

Jobs by Site — 16 Weeks from Week 26 of 2012 to Week 42 of 2012

Max: 931, Average: 64.8

LCG.DESY-ZEUTHEN.de	77.5%	LCG.CYFRONET.pl	4.1%	LCG.MSFG.fr	0.9%
LCG.PIC.es	7.1%	LCG.IN2P3-CC.fr	2.1%	DIRAC.PIC.es	0.0%
LCG.GRIF.fr	6.5%	LCG.LAPP.fr	1.8%		

Generated on 2012-10-18 14:2

Cumulative Input data by Site — 16 Weeks from Week 26 of 2012 to Week 42 of 2012

Max: 392, Average: 178, Current: 392

LCG.DESY-ZEUTHEN.de	306.7	LCG.CYFRONET.pl	14.2	LCG.MSFG.fr	3.3
LCG.PIC.es	27.8	LCG.IN2P3-CC.fr	8.5	DIRAC.PIC.es	0.0
LCG.GRIF.fr	24.5	LCG.LAPP.fr	6.5		

Generated on 2012-10-18 14:27:07 UTC

User Web Client and/or VM for data searching
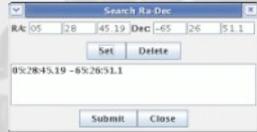
Web Server workload management

Grid jobs management

Client application for analysis job submission and results info./ file

downloading

VOMS Server (CCIN2P3)

Regional Network

Workload Management System (WMS)

LFC (CCIN2P3)

LAPP Internal Network

Backup Server

High-level data request

Storage Element (SE)

Computing Element (CE)

EGEE Grid services

Worker Nodes

Storage
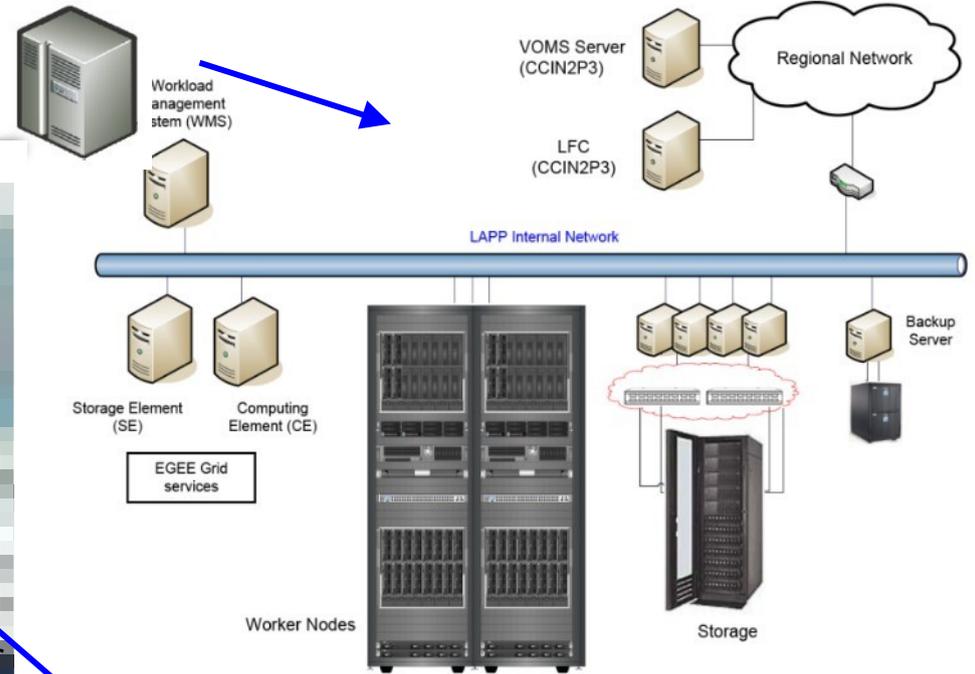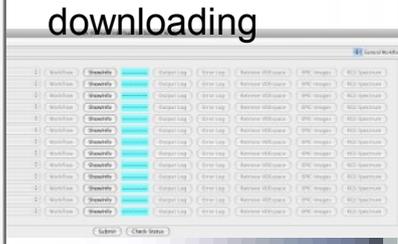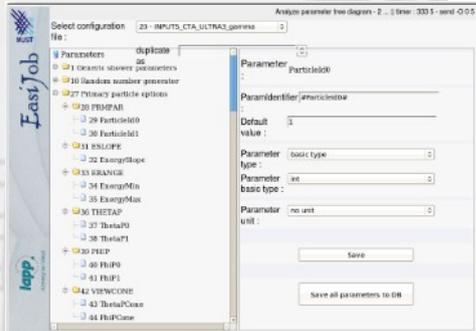
Interface layer with archive DB
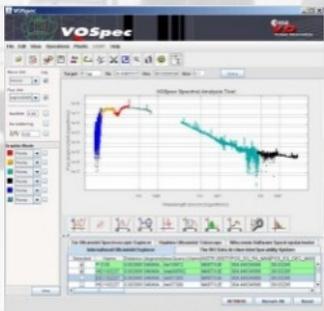
FITS and VObs data access

Science Archive, DB and MetaDB systems, notification and results uploading

FITS and VO compliant results VObs tools (VOSpec, Aladin)

Conclusions

Mathieu de Naurois

- *Deux approches assez différentes :*
  - *HESS : middle-ware propriétaire, très intégré au soft HESS.*
    - Nettement plus simple pour l'utilisateur (GUI), gestion assez transparente
    - Mais un travail de développement plus lourd
  - *CTA : utilisation d'un middle-ware extérieur (DIRAC)*
    - Plus simple au niveau développement
    - Mais la complexité de certaines tâches rend l'utilisation des outils pas toujours aisée
- *Dans les deux cas, retour d'expérience très positif*
- *Le point le plus difficile reste la disponibilité des espaces de stockage*