Le calcul pour les Astroparticules

- Status
- Evolution des expériences actuelles
- Expériences futures
- Conclusion

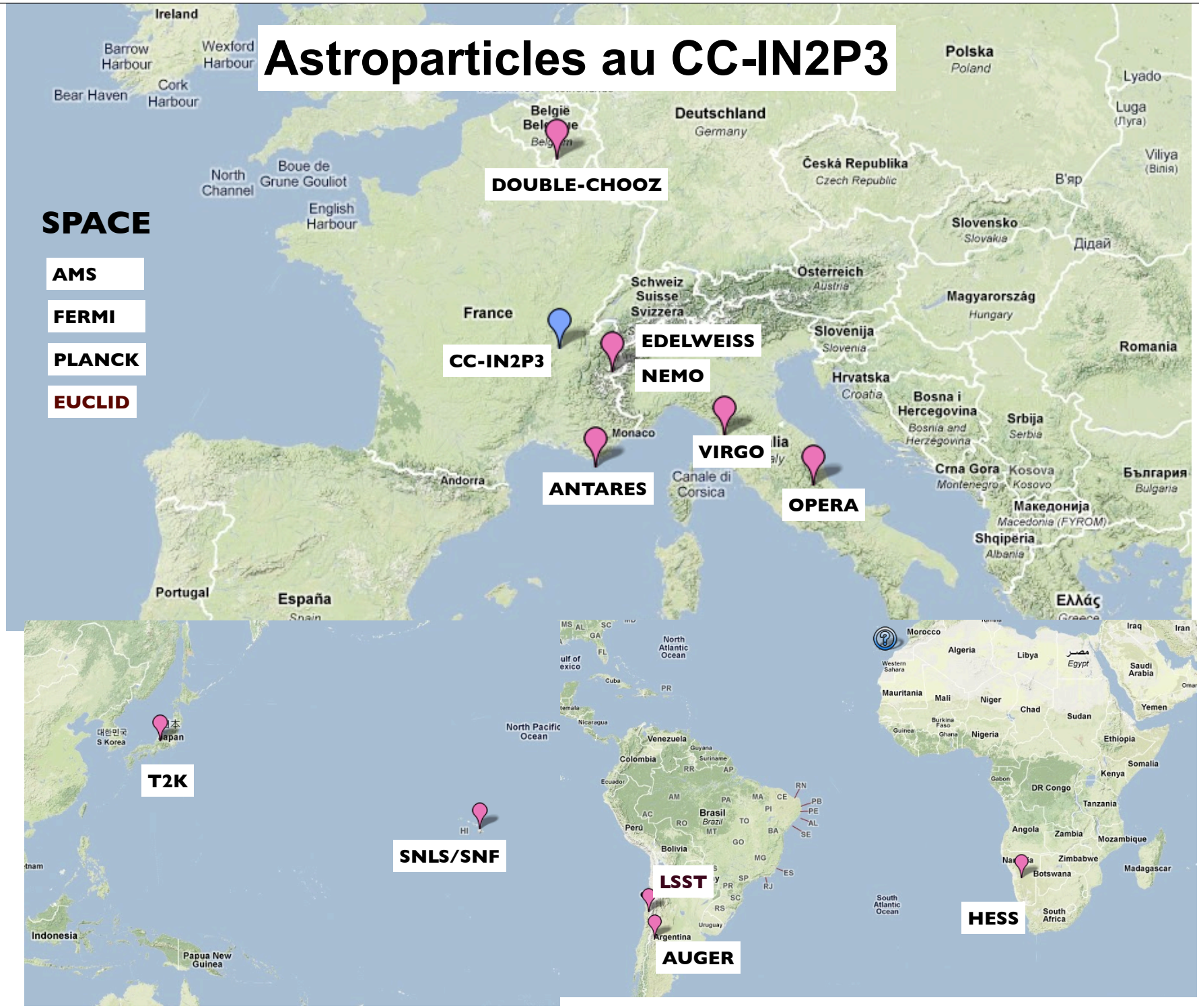
Rachid Lemrani

Journées prospectives du CC-IN2P3

Limas

Irfu
cea
saclay

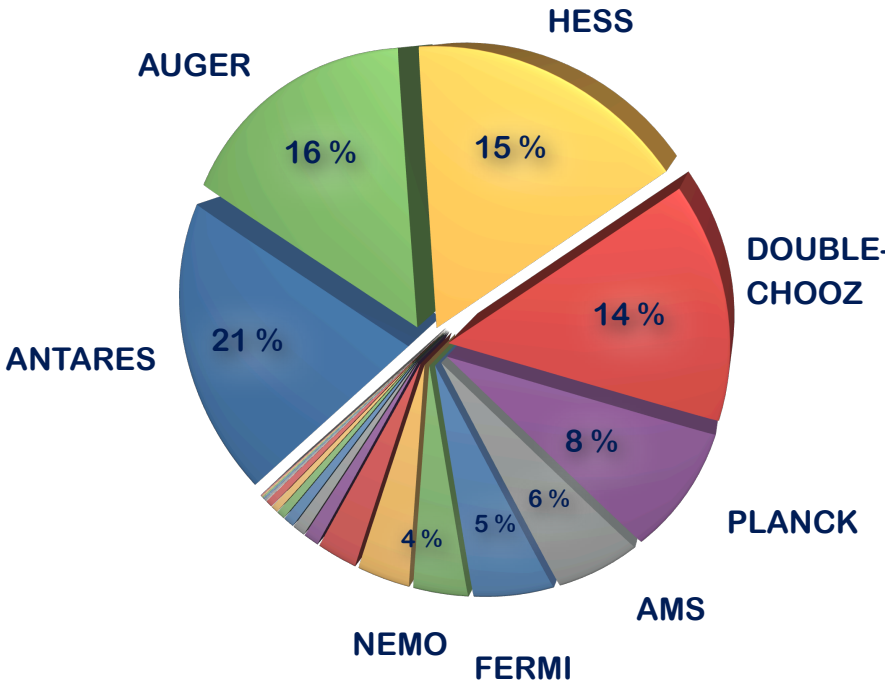
cnrs


Astroparticles au CC-IN2P3
SPACE
AMS
FERMI
PLANCK
EUCLID
DOUBLE-CHOOZ
CC-IN2P3
EDELWEISS
NEMO
VIRGO
ANTARES
OPERA
T2K
SNLS/SNF
LSST
AUGER
HESS
Ireland
Barrow Harbour
Wexford Harbour
Bear Haven
Cork Harbour
North Channel
Boue de Grune Gouliot
English Harbour
Belgie
Deutschland
Germany
Česká Republika
Czech Republic
Polska
Poland
Slovensko
Slovakia
Österreich
Austria
Schweiz
Suisse
Svizzera
France
Magyarország
Hungary
Slovenija
Slovenia
Romania
Hrvatska
Croatia
Bosna i Hercegovina
Srbija
Serbia
Monaco
Andorra
Canale di Corsica
Crna Gora
Kosova
България
Bulgaria
Македонија
Shqipëria
Albania
Portugal
España
Ελλάς
North Atlantic Ocean
North Pacific Ocean
Venezuela
Colombia
Brasil
Perú
Bolivia
Argentina
Uruguay
Japan
S Korea
Indonesia
Papua New Guinea
Morocco
Algeria
Libya
Egypt
Mauritania
Mali
Niger
Chad
Sudan
Nigeria
Ethiopia
Somalia
Kenya
DR Congo
Tanzania
Angola
Zambia
Mozambique
Zimbabwe
Botswana
Madagascar
South Africa
South Atlantic Ocean
Iraq
Iran
Saudi Arabia
Yemen

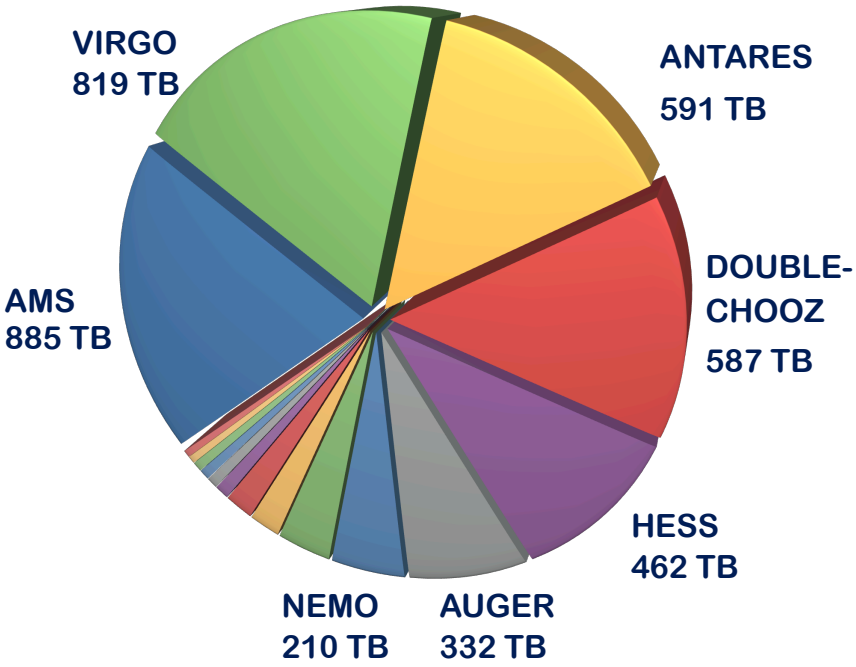
Status

2012

Calcul : 10% total



Stockage



- Disques hautes performance (GPFS) : PLANCK : ~ 190TB
  SNLS : ~ 110 TB
  SNF : ~ 60TB

Type de données

Evénements : détections de particules gamma, neutrinos, rayons cosmiques

Signaux 1D : **virgo**, **planck**

Images 2D : Telescopes CFHT, UH

Fichiers de conditions de run : calibration, monitoring

Type de Stockage

Stockage de masse sur HPSS : gros fichiers

SPS pour les analyses de données : SuperNovae, Planck ...

Disponibilités des données sur disques : Signaux pour les ondes gravitationnelles

Fichiers partagés sur AFS (throng) : programmes

Base de données : conditions de run, paramètres

Données d'**OPERA** sur 5 ans 140 TB sur Oracle

Gestions et accès aux données

Gestion avec IRODS (remplaçant SRB) : Transfert des données, gère des stockages hétérogènes : cartouches, disques, bases de données : l'utilisateur ne voit qu'une arborescence simple et utilise des commandes type unix pour copier, lister, rechercher des fichiers

Accès avec XROOTD : Accès pour lecture bien adapté aux fichiers root (autres formats de fichiers aussi possible)
Interfacé au CC-IN2P3 avec HPSS (et DCACHE)

Type de Calcul

Calcul parallèle avec la ferme de calcul // : CMB

Traitement du signal : Ondes gravitationnelles, CMB (1D → cartes 2D)

Traitement d'images : Supernovae

Traitement classique des événements : reconstruction, calibration, analyse

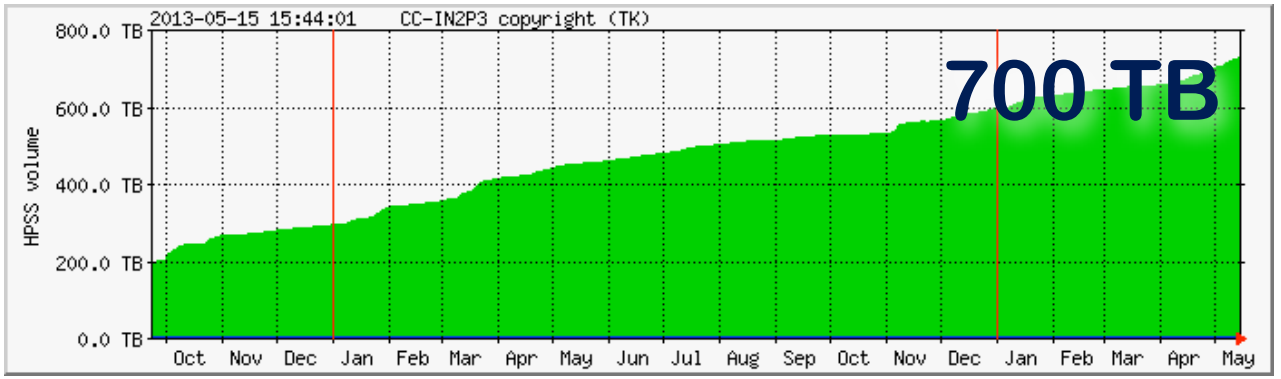
Simulations monte-carlo :

- Gerbes atmosphériques (**auger, cta**)
- Bruits de fond pour la détection d'événements rares (**edelweiss, supernemo**)
- ...

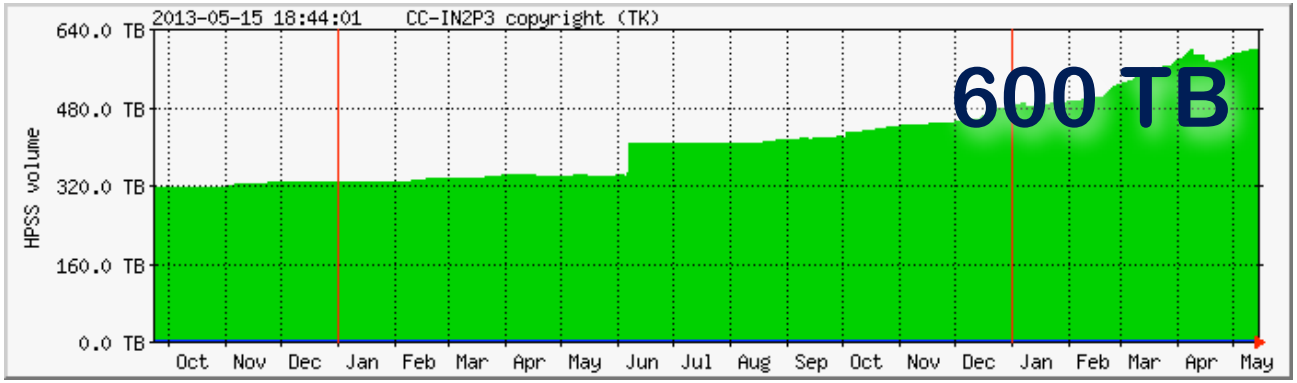
Evolution courante

Double-Chooz

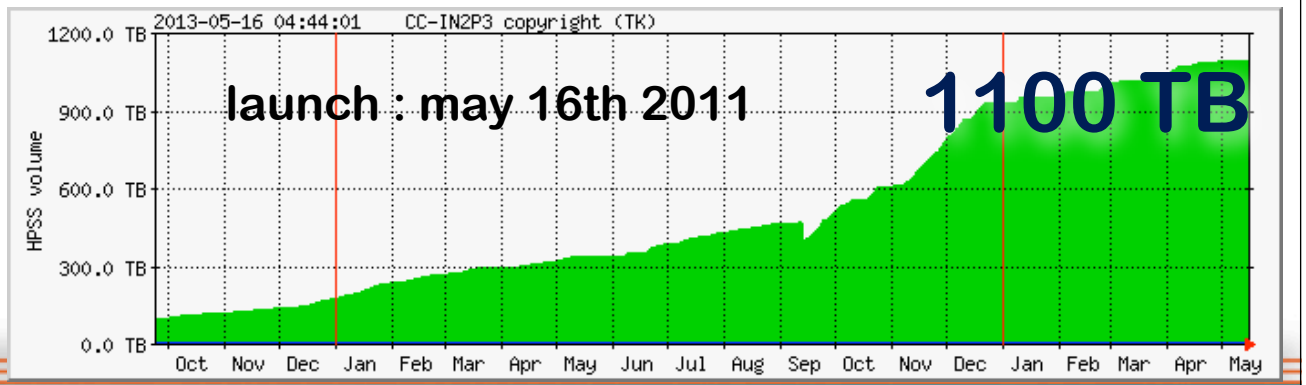
2eme détecteur
printemps 2014



Hess → Hess2



AMS02



Futur : Expériences Astroparticules

		ACTIVES	récentes	BIENTÔT	PLUS TARD
Astroparticules Energies extrêmes Origine Composition Anti-matière Matière Noire Gravité	Gamma	HESS FERMI	HESS-II		CTA
	Neutrino	ANTARES			KM3NET
	Rayons cosmiques	AUGER	AMS02		JEM-EUSO
	Ondes gravitationnelles	VIRGO		Adv VIRGO	ET
Cosmologie Matière Noire	Matière Noire	EDELWEISS			EURECA
Energie noire	SuperNovae	SNLS / SNF			LSST /EUCLID
Paramètres	CMB	PLANCK			
Neutrinos : Majorana/Masse	Double désintégration β sans ν	NEMO			SuperNEMO
Mixing	Oscillation Neutrino	OPERA	D-CHOOZ T2K		
Désintégration du proton					LAGUNA

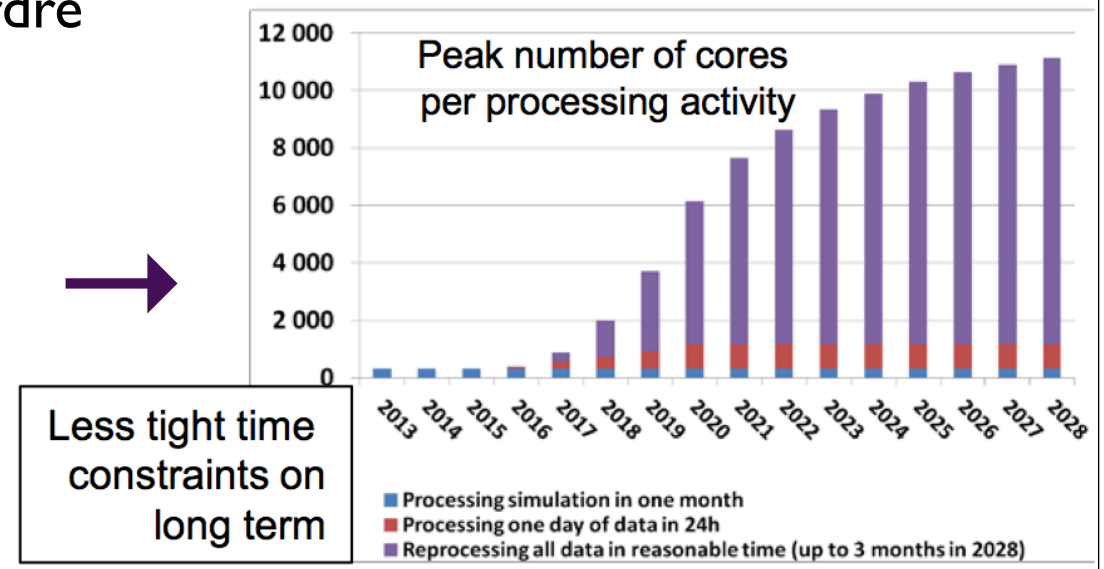
CTA

- CC-IN2P3 : via la grille
  software : SVN, RedMine, Jenkins, ...

- 2016 to 2020 : déploiement du réseau
  -> augmentation progressive des données

- Stockage : à partir de 2020 : de l'ordre de 1 à 10 Po/an

- Calcul : dépend du scenario
Exemple : reprocessing de toutes les données passées en moins de 3 mois

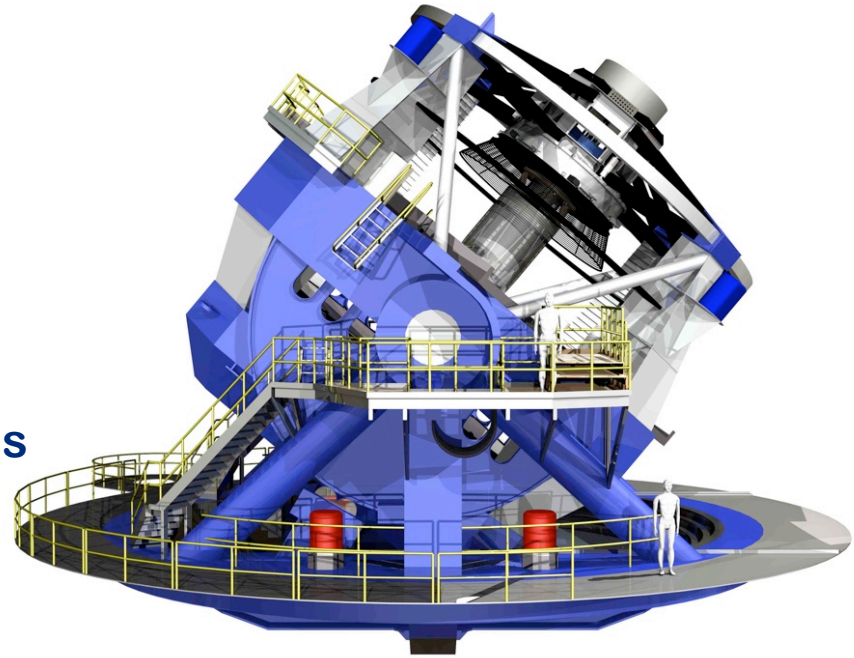


LSST

Voir Présentation de Dominique Boutigny

2018 : Commissioning
2020 : Données de physique

CC-IN2P3 : - Moitié de la production
(autre moitié par NCSA)
- Hébergement de toutes les données

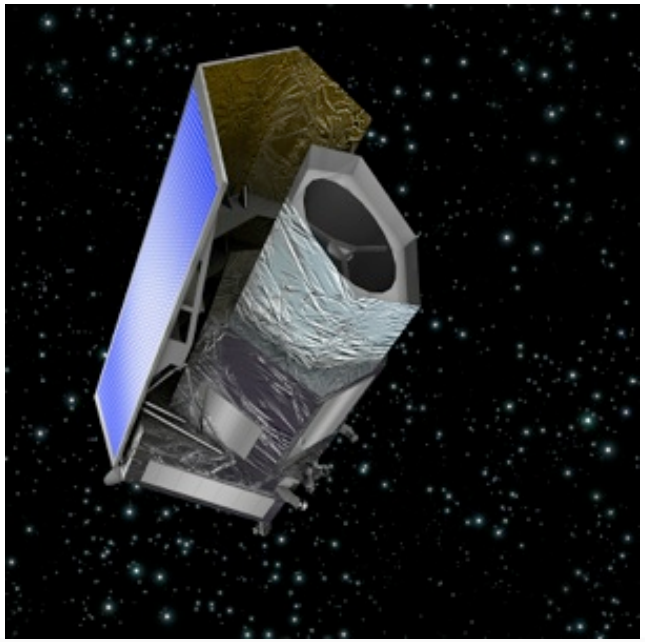


Défis : - volume de données 100s PB ?
- type d'analyse : images
- saut d'échelle par rapport à SNLS / SN Factory
- Enorme base de données: 10s PB d'objets, sources, …

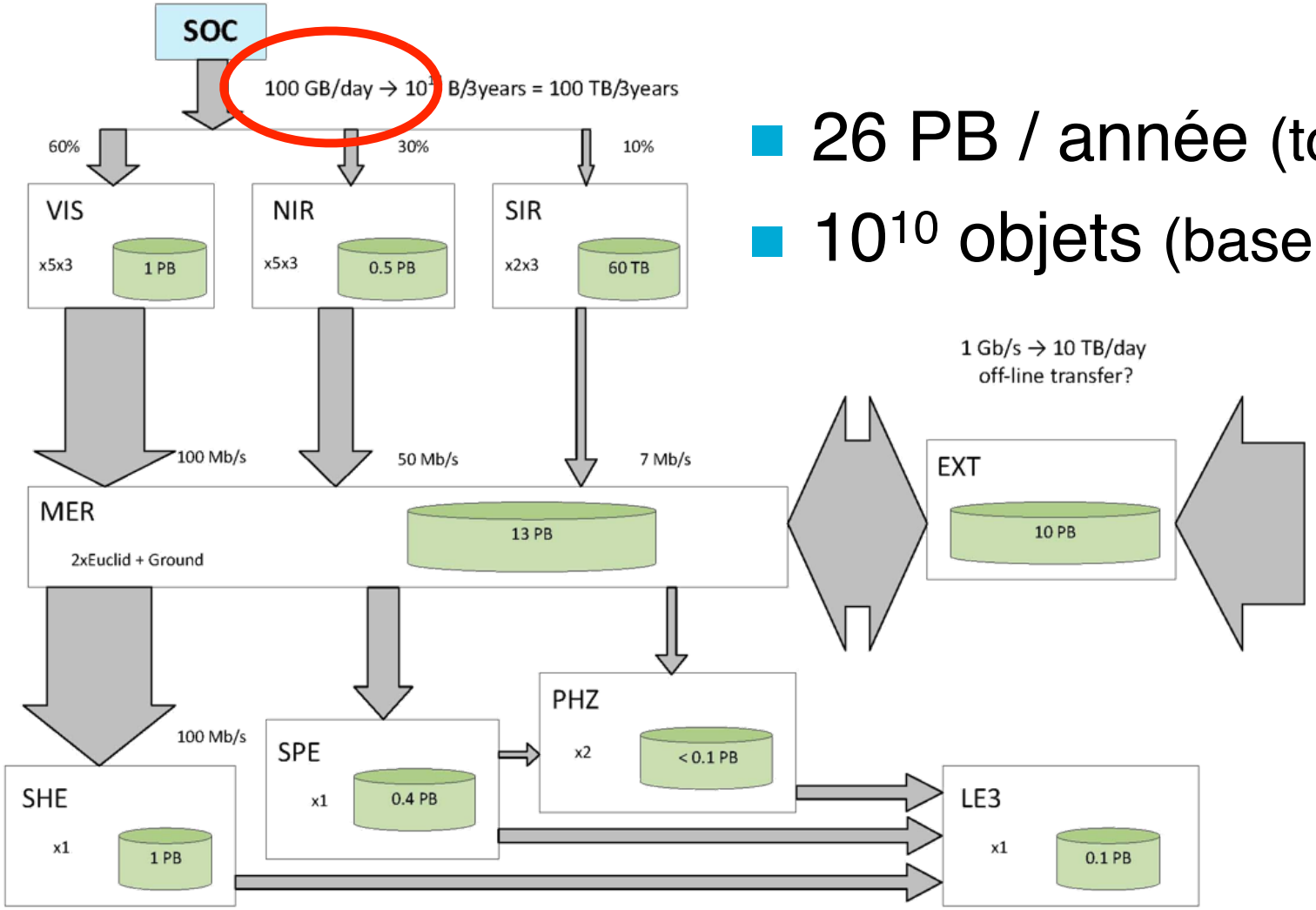
2013 : - Participation au data challenge cet été
~ 100 TB de données

Euclid

- **mission spatiale de 6 ans (~ 900 M€)**
  - étude de l'énergie / matière noire, gravité et structures à grande échelle - lancement Q2 2020
  - télescope de 1.2m avec 2 instruments
    - imagerie grand-champ dans le visible
    - imagerie / spectroscopie infra-rouge
  - Euclid Consortium (UE + NASA)
    - 109 laboratoires / > 1100 membres
      - instruments
      - traitement de données au sol

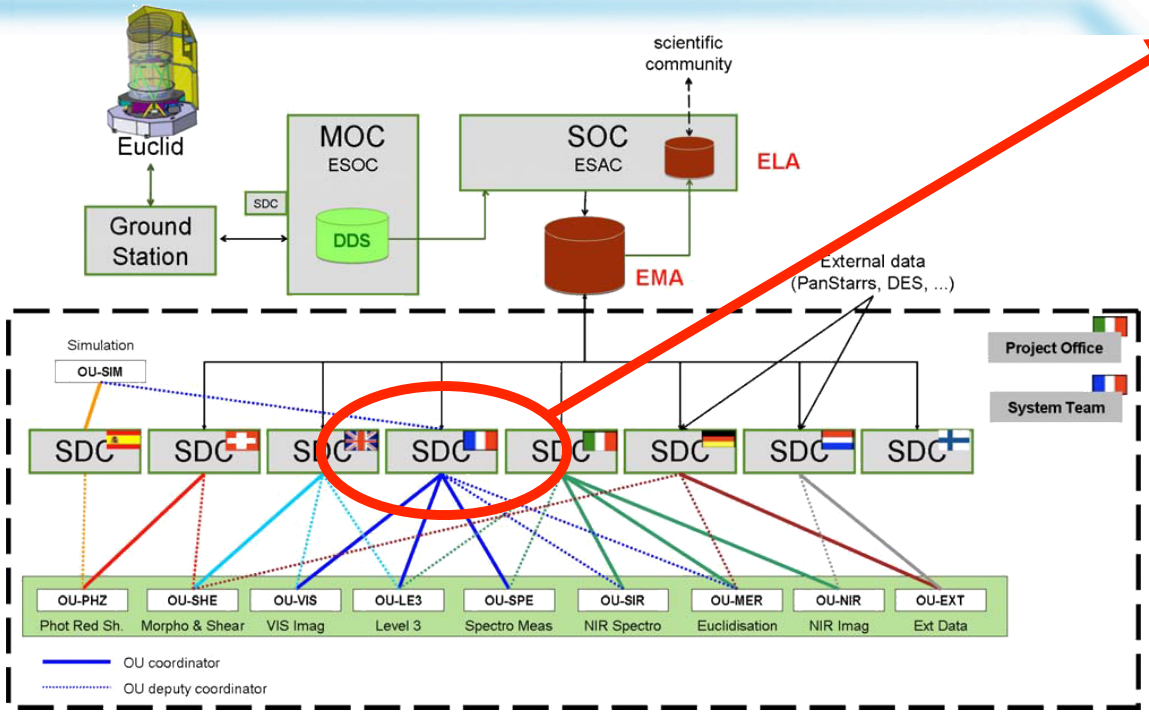


Science Ground Segment



- 26 PB / année (total)
- 10^{10} objets (base données)

Science Data Centers



CCIN2P3

SDC-PROD

- mise en oeuvre des algorithmes de traitement de données
- ressources h/w & s/w
  - DEV : développement (APC)
  - PROD : intégration & production locale (CC-IN2P3)

Euclid @ CC-IN2P3

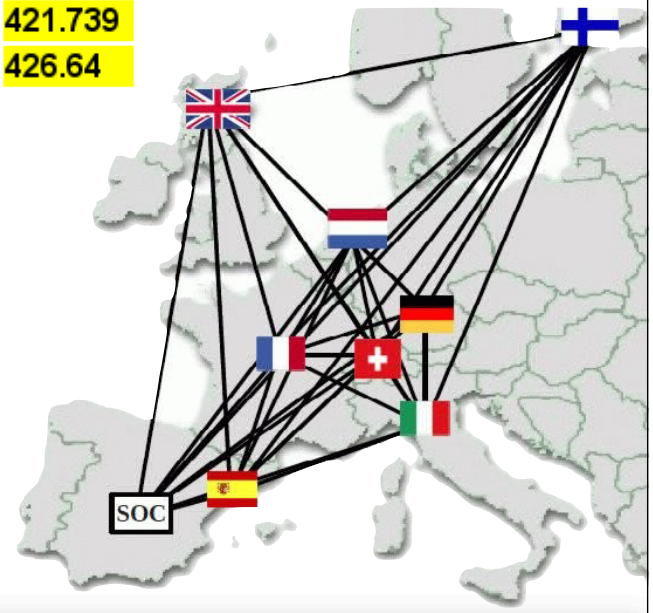
Summary, avg(Mbps), 5 clients, total avg(Mbps)=426.64±150.909

C \ S	SDC-CH	SDC-DE	SDC-ES	SDC-FI	SDC-FR	SDC-IT	SDC-NL	SDC-UK	SOC	
SDC-CH		350.791	364.815	395.924	562.077	428.592		326.867	284.608	**387.668**
SDC-DE	150.841		358.749	433.816	270.259	290.792		188.174	201.898	**270.647**
SDC-ES	268.445	668.899		500.112	786.014	417.187		430.19	607.809	**525.522**
SDC-FI										
SDC-FR	238.409	777.564	676.377	574.307		560.061		555.183	589.074	**567.282**
SDC-IT	219.054	500.527	413.674	471.274	522.48			352.885	439.034	**416.99**
SDC-NL										
SDC-UK	243.875	415.074	437.704	335.879	540.508	410.271			393.13	**396.634**
SOC	162.301	422.204	661.294	436.391	492.708	481.364		295.913		**421.739**
	213.821	**522.51**	**485.435**	**449.672**	**529.008**	**431.378**		**358.202**	**419.259**	**426.64**

client

serveur

- **Bandwidth benchmark**
  - **SDC-FR classé 1er : 528 Mb/s**



Euclid @ CC-IN2P3

- CODEEN (COllaborative DEvelopment ENvironement)
  - plateforme d'intégration continue (Jenkins + Sonar) @ APC
  - svn → source → package → installation → tests unitaires
  - test réussi d'utilisation de cceuclid comme slave Jenkins
- Catalog benchmarking
  - création et remplissage de bases de données (Oracle, FITS, SQlite, HBase, ...) avec un catalogue simulé
  - optimisation de requêtes, création d'indexes, ...
- Cloud OpenStack R&T
  - benchmarking avec Hadoop

Technologies

Cloud : oui déjà testé par LSST et Euclid

Calcul parallèle : utilisé principalement par PLANCK

GPU :

utilisé à hannovre massivement pour les ondes gravitationelles: traitement du signal des interféromètres
pourrait servir à :
la détection directe de matière noire ? traitement du signal des bolomètres
l'énergie noire ? certaines analyses d'images

Conclusion

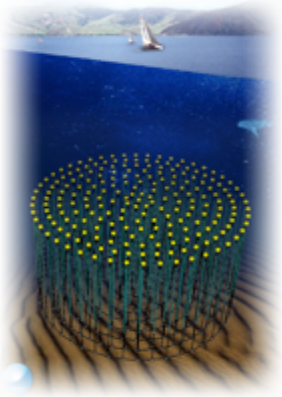
- Le calcul pour les Astroparticule augmente continuellement avec les upgrades

- Continuera à augmenter avec les nouvelles générations d'expérience : KM3NET, CTA, AUGER-extension, 1-ton DARK-MATTER, SUPERNEMO, avdVIRGO/ET, LAGUNA, JEM-EUSO, LISA, ...

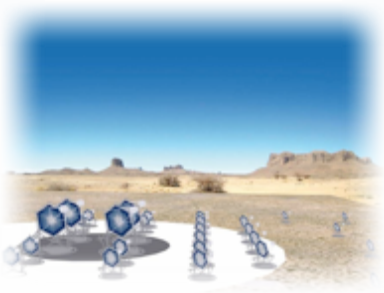
- Les expériences les plus exigeantes sont LSST and Euclid
  - Grand saut d'échelle par rapport à SNLS, SNFactory
  - Traitement d'images

Futures expériences

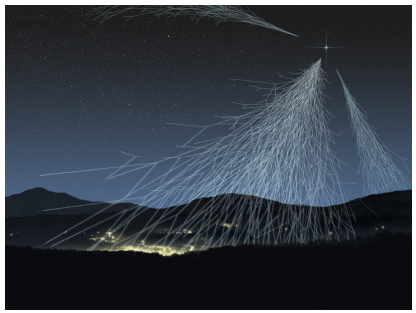
KM3NET
high energy neutrinos



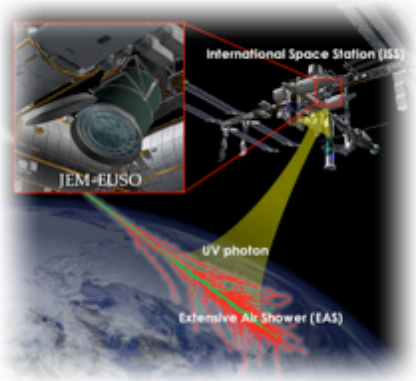
CTA : gamma rays



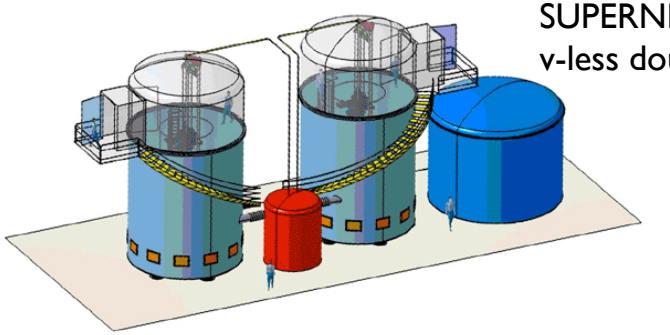
AUGER upgrade/extension



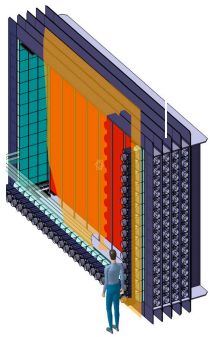
JEM-EUSO : extreme energies



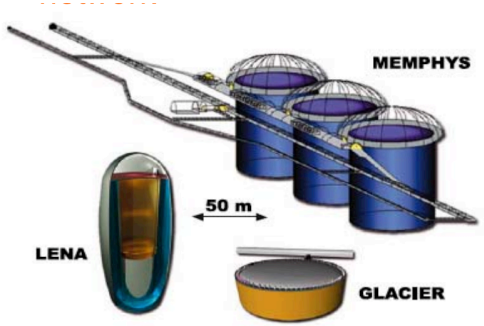
EURECA : Dark Matter



SUPERNEMO
v-less double beta decay

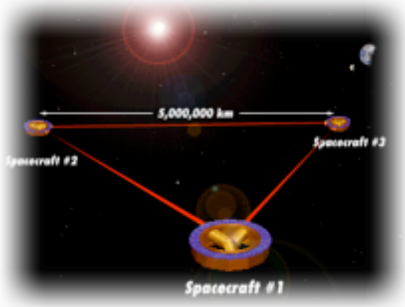


LAGUNA proton decay, neutrino

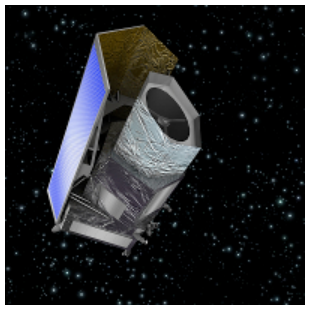
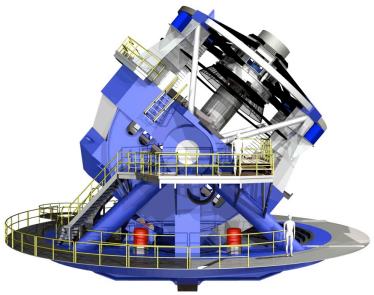


advVIRGO /ET Gravitational waves

LISA gravitational waves



Euclid,
dark energy



LSST, Chile
dark energy