



Vendredi 12 octobre 2007

# Présentation du CC-IN2P3

Réunion IN2P3 - DSM / DAPNIA

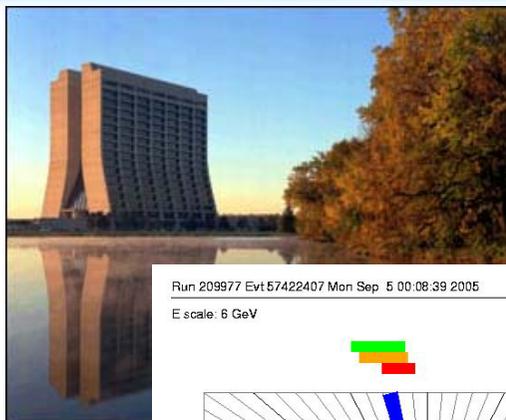


dapnia  
cea  
saclay

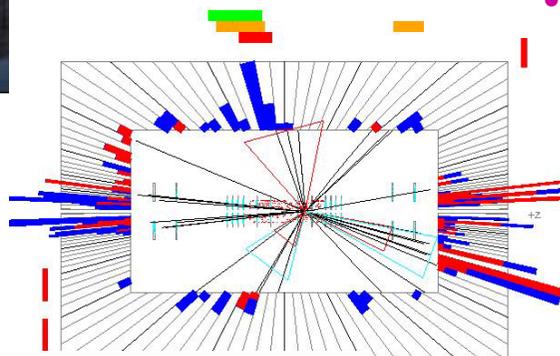
CNRS  
CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE



# Mission du CC-IN2P3 (2)



Run 209977 Evt 57422407 Mon Sep 5 00:08:39 2005  
E scale: 6 GeV



## Expériences de physique

- Physique Nucléaire
- Physique des particules
- Astro-particules



## Recherche fondamentale



## Analyse des résultats



## Masse de données

```

101000 100111 0001001010001
00011101 100010001111 00010
101000 100111 0001001010001
00011101 100010001111 00010
101000 100111 110001 111010
0001001010001101000 100111
0001001010001 00011101 1110
100010001111 00010 101000 00
11 0001100111 0001001010001
00011101 100010001111 00010
101000 100111 0001001010001

```

## Traitement des données



## Publications

FERMLAB-CONF-  
CDF/PUB/CDF/PUBI  
November

Electroweak, Top and Bottom Physics at the Tevatron

FUMIHIKO UKEGAWA (CDF Collaboration)  
Institute of Physics, University of Tsukuba  
Tennoudai 1-1-1, Tsukuba-shi, Ibaraki-ken 305-8571, Japan  
E-mail: ukogawa@hep.pr.tsubu.ac.jp

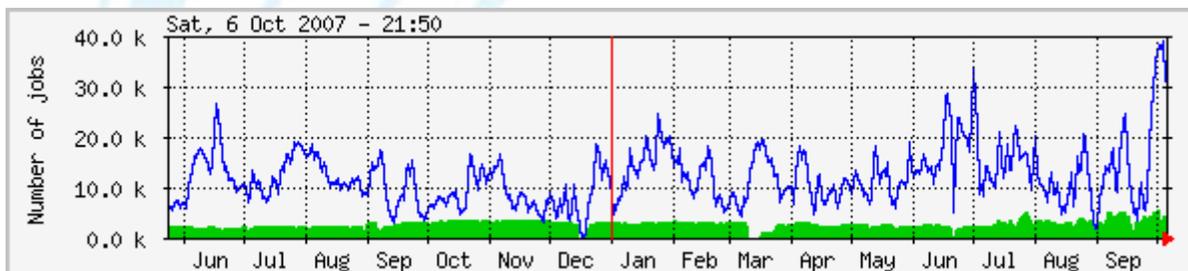
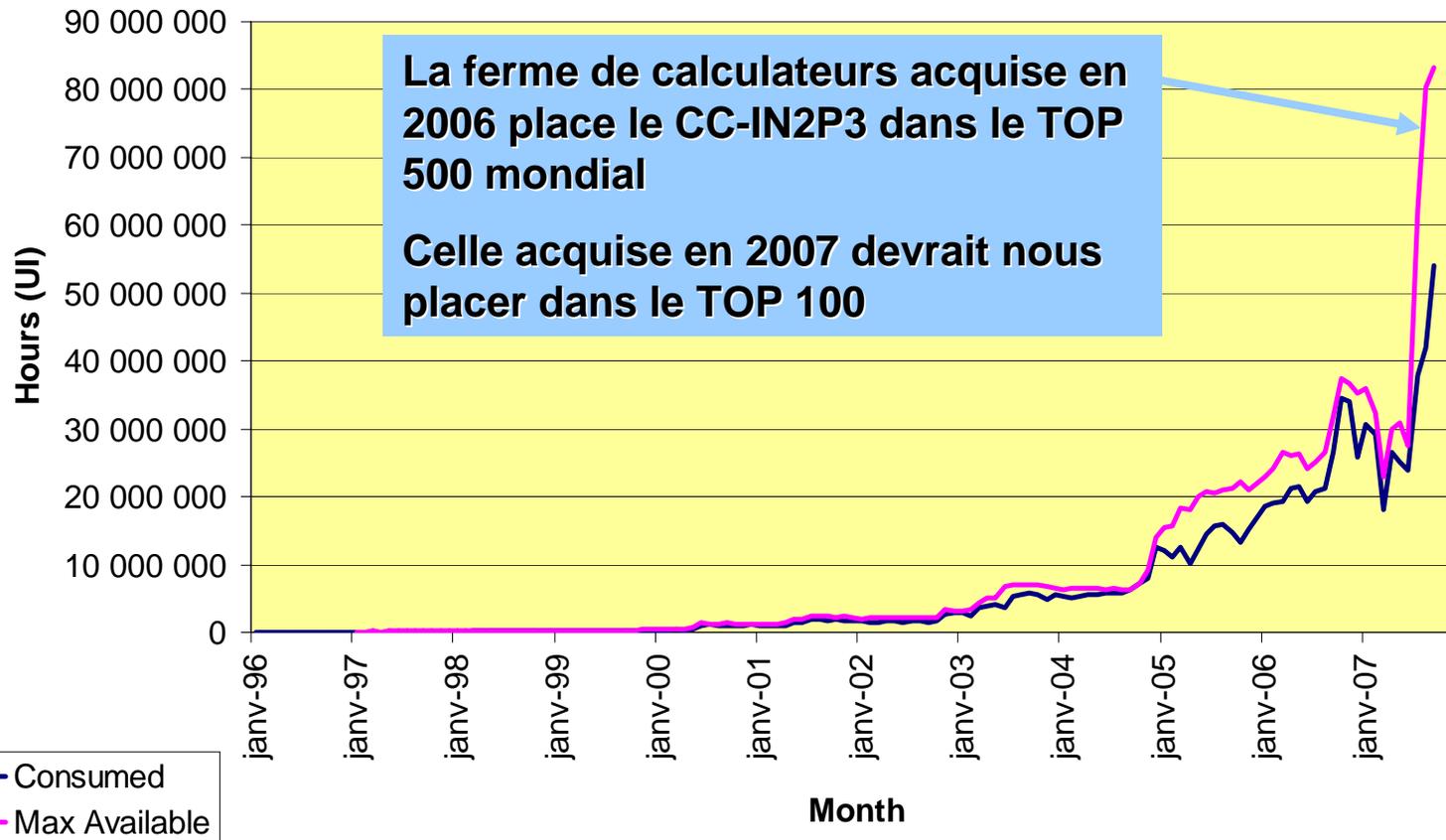
representing the CDF and D0 collaborations

### ABSTRACT

The Tevatron Run-II program has been in progress since 2001, and the CDF and D0 experiments have been operational with upgraded detectors. Coupled with recent improvements in the Tevatron accelerator performance, the experiments have started producing important physics results and measurements. We report these measurements as well as prospects in the near future.

0411012 v2 12 Nov 2004

# Évolution de la puissance de calcul disponible au CC-IN2P3

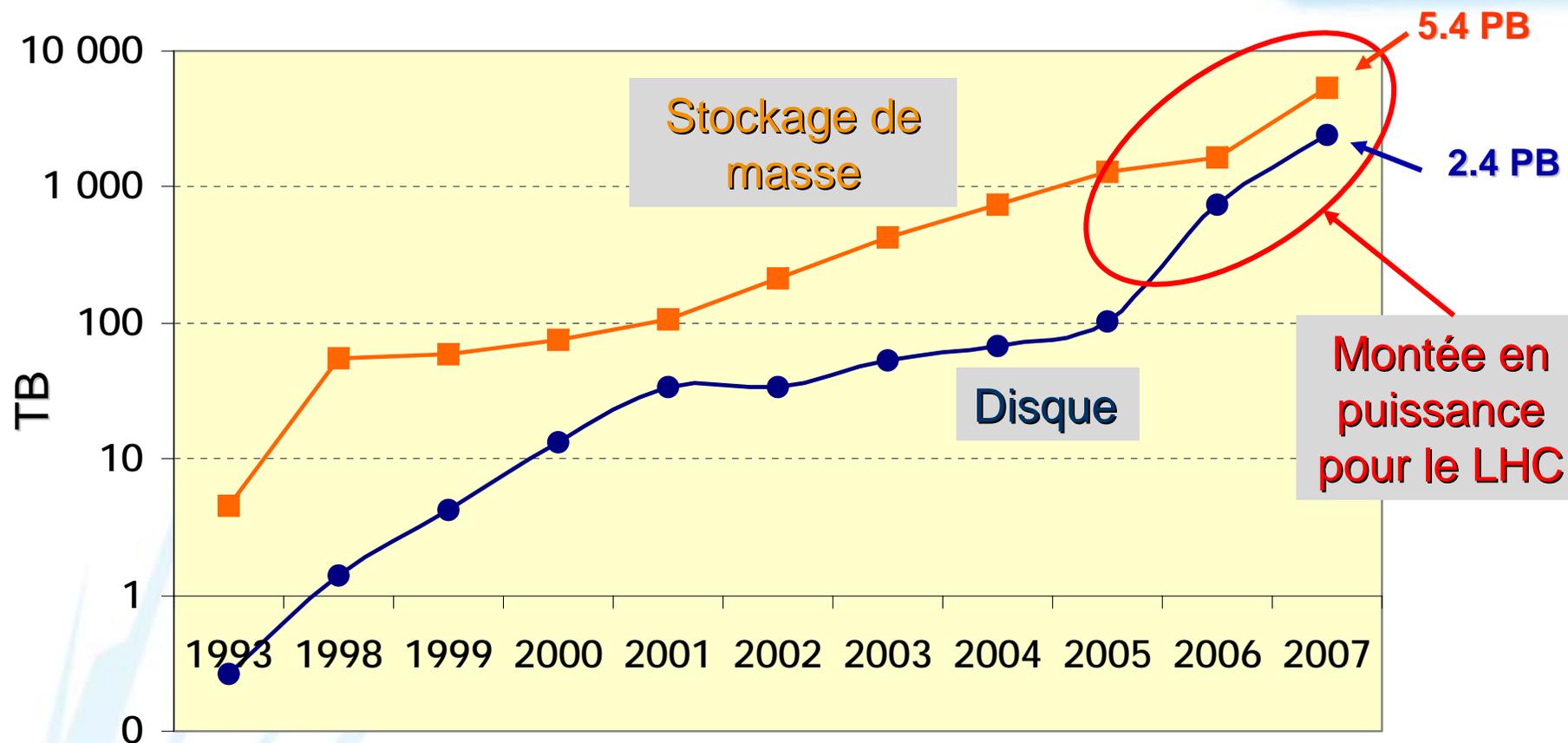


**5500 tâches en //**

**20 000 – 30 000 tâches en attente**

**20 000 tâches exécutées par jour**

# Évolution de la capacité de stockage au CC-IN2P3



# Pourquoi de tels besoins informatiques en physique des hautes énergies ?



- Sophistication des expériences
  - Très grand nombre de canaux d'électronique
    - Masse de données
- Collaborations internationales
  - Distributions des données
  - Analyse distribuée géographiquement
    - Développement du réseau longue distance
    - Importance des relations avec RENATER
- Nécessité de simuler les expériences
  - Sophistication de la modélisation de la réponse des détecteurs
    - Temps de calcul et masse de données

# Une dimension internationale



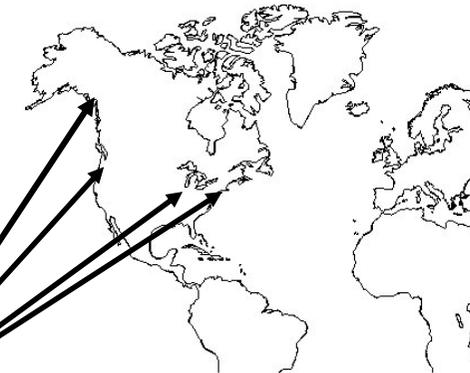
Le CC-IN2P3 fait partie d'un réseau mondial de grands centres de calcul pour la physique des hautes énergies



- CERN
- RAL (Grande-Bretagne)
- CNAF (Italie)
- FZK et DESY (Allemagne)
- PIC (Espagne)
- NIKHEF (Pays Bas)



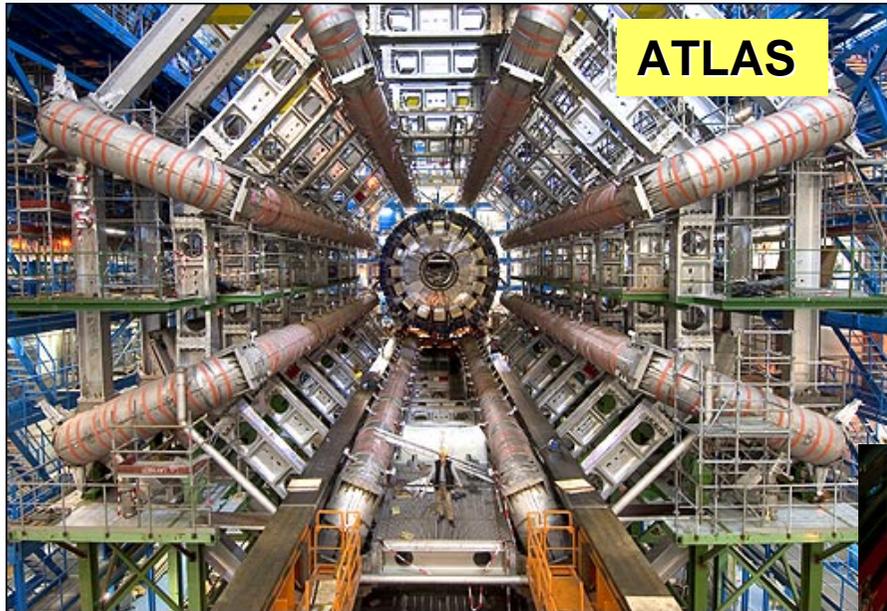
→ Collaboration active dans de nombreux domaines



- NDGF (Pays Nordiques)
- TRIUMF (Canada)
- SLAC (Stanford)
- Fermilab (Chicago)
- Brookhaven (Long Island)
- KEK (Japon)
- ASCC (Taïwan)

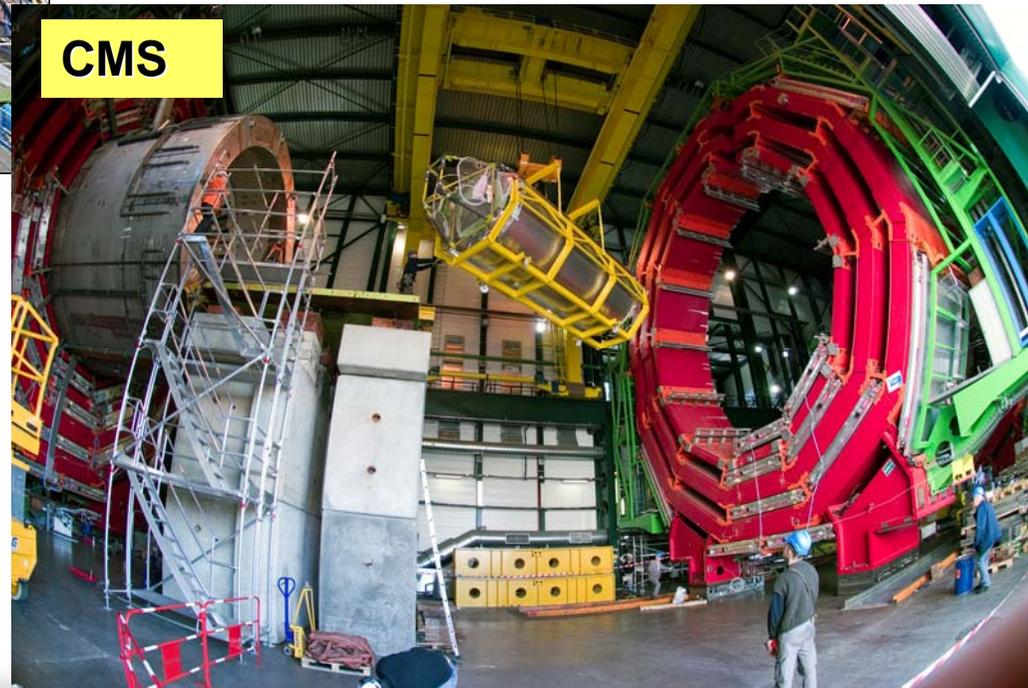


# Le LHC – Un nouveau défi



**ATLAS**

**4 expériences d'une complexité sans précédent**



**CMS**

**10 milliards de collisions enregistrées / an**

**15 Po de données / an**



# 15 PetaOctets ?

1 disque dur classique: 500 Go

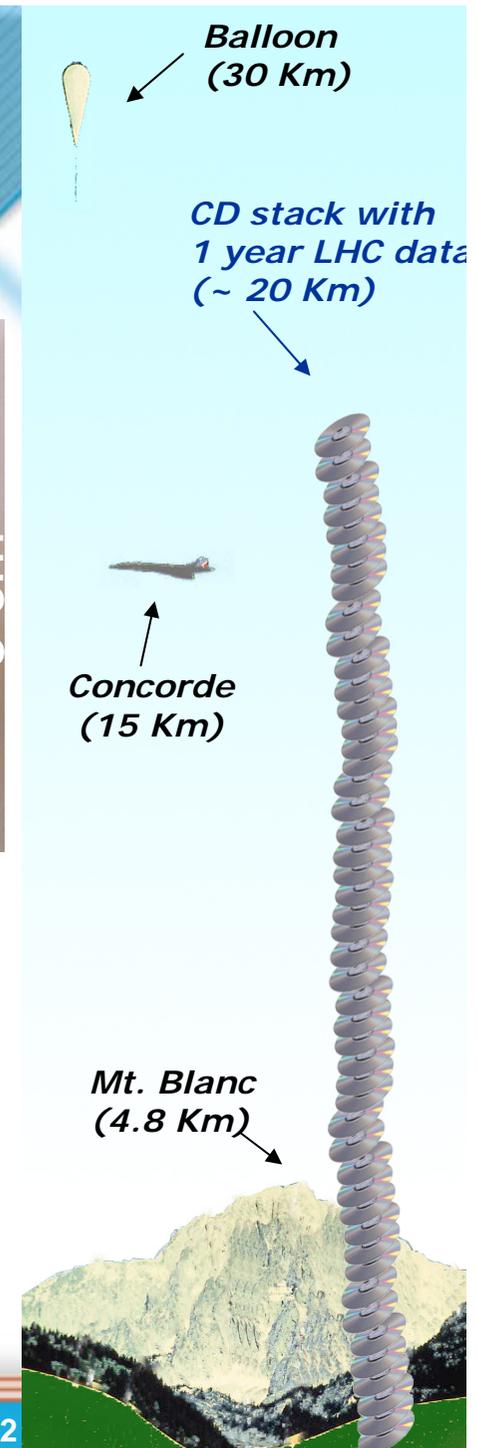
1 an de LHC → 30 000 disques durs

100 millions de SpecInt2000 sont nécessaires pour traiter ces données

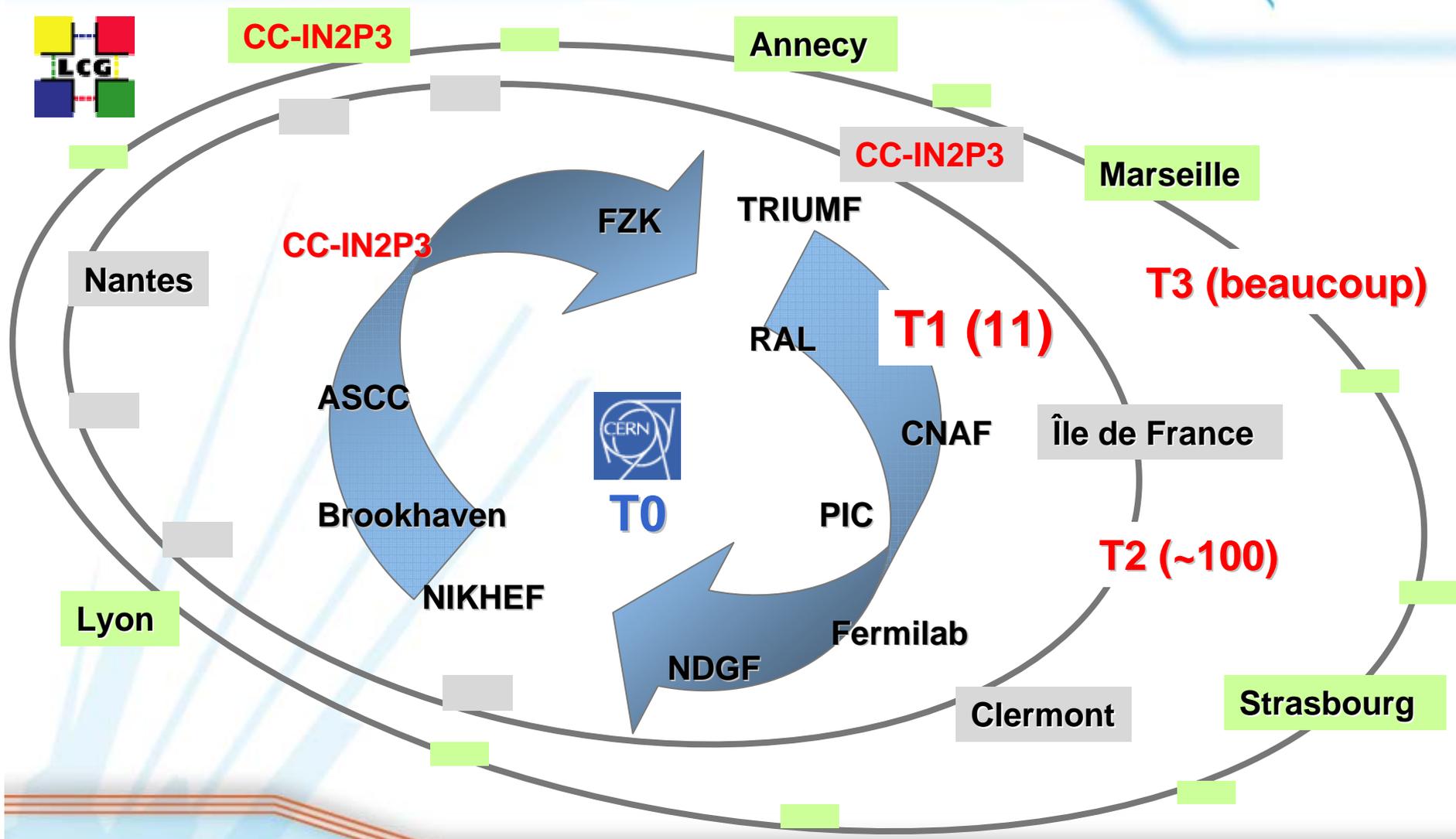
Soit environ 70 000 CPU d'aujourd'hui

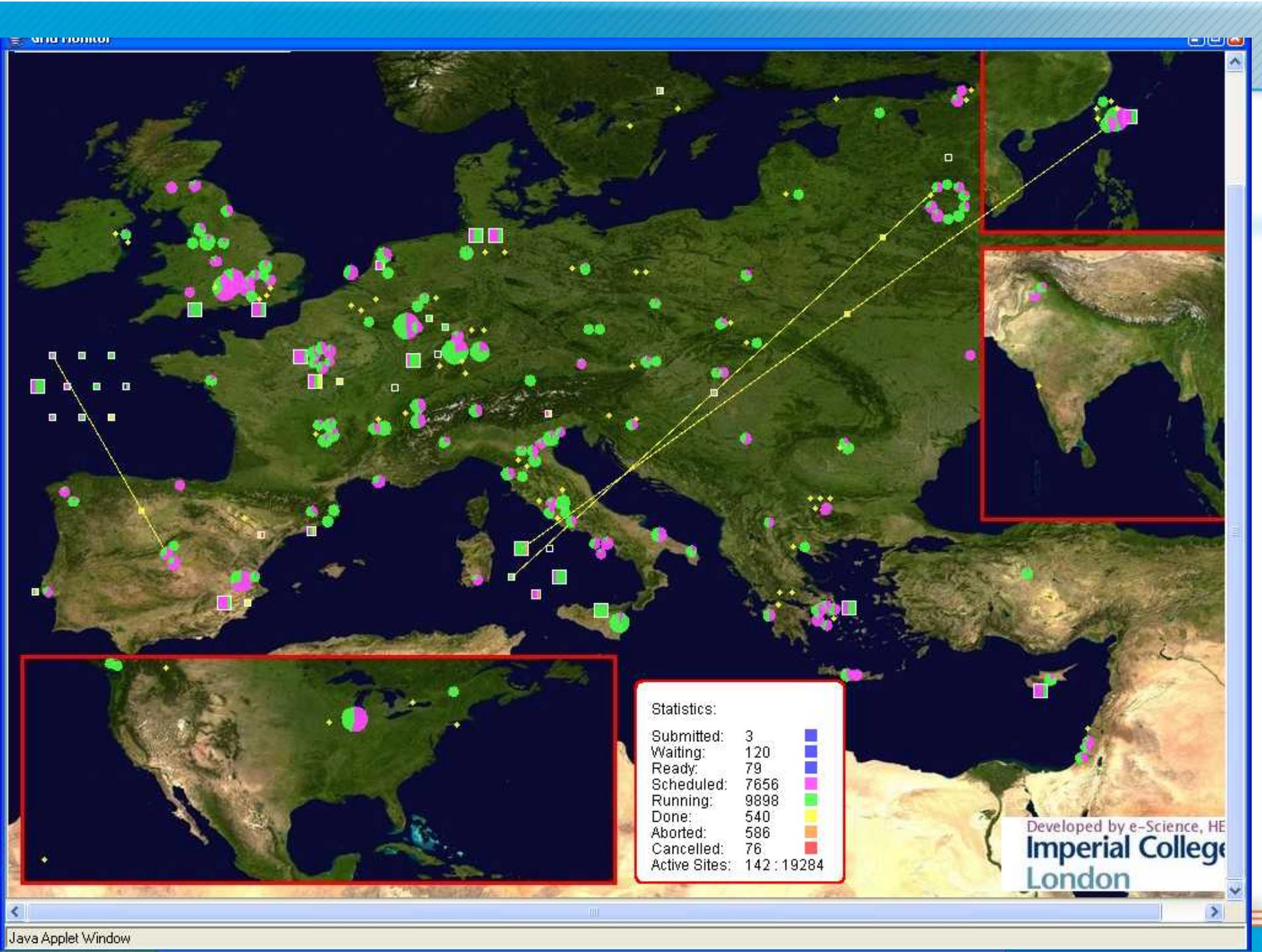
Un traitement centralisé des données n'est pas envisageable

→ Il est nécessaire de coupler la puissance de calcul disponibles dans de nombreux centre de calcul



# Une architecture de Grille de calcul dans le cadre d'un projet mondial





# Le CC-IN2P3 et les grilles de calcul



Grille pluridisciplinaire EGEE, pilotée par le CERN

Rôle important du CC-IN2P3

11-12 personnes financées par EGEE

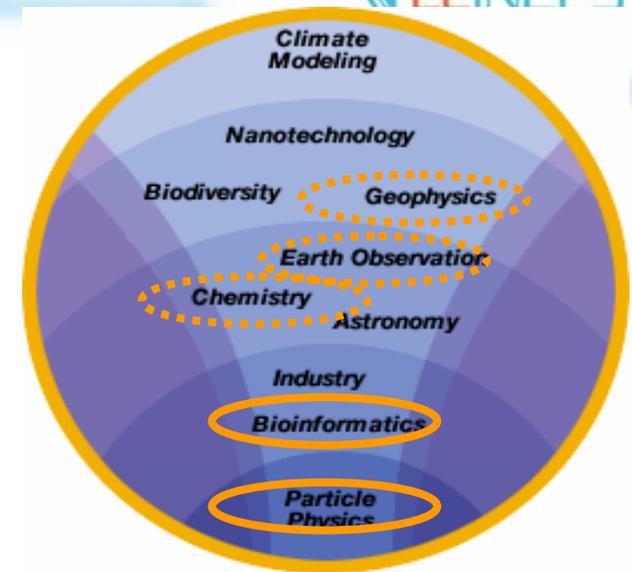
Critique pour le succès du calcul LHC en France

→ Expertise dans l'exploitation de la grille

Rôle pilote dans plusieurs projets de grilles

- OpenPlast : Grille industrielle liée au pôle de plasturgie
- RUGBI : Grille Bioinformatique

Projet ANR en cours sur l'interopérabilité des grilles et les transferts massifs de données



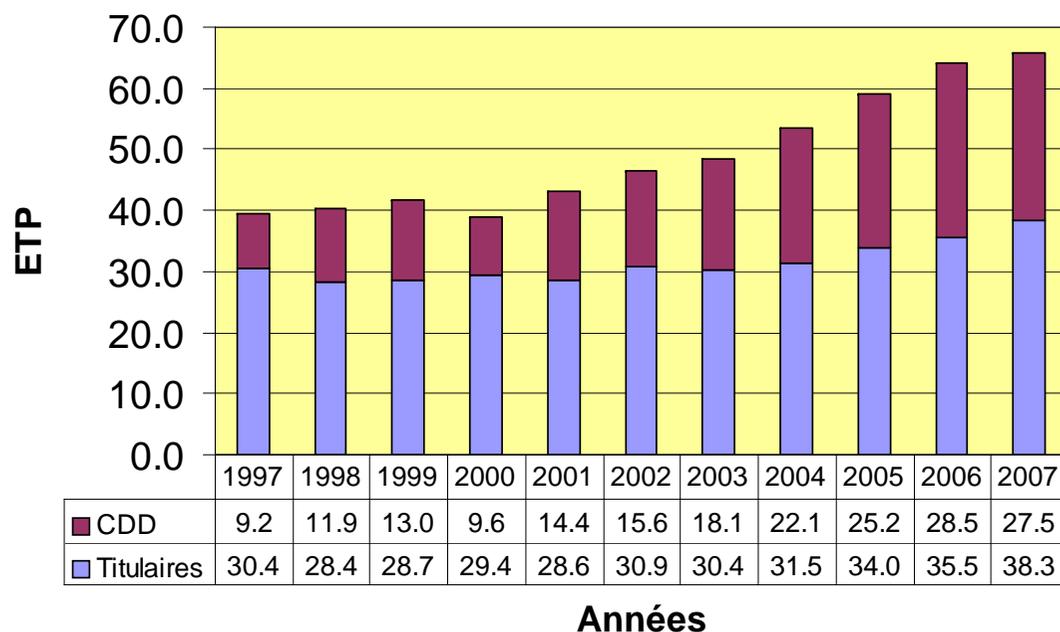
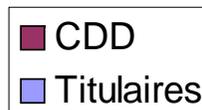
→ Souhait de faire du CC-IN2P3 la tête de pont d'une grille régionale



# Évolution du personnel



### Personnel ETP du CC-IN2P3



Recrutement important ces dernières années +4 / an essentiellement lié au projet LCG

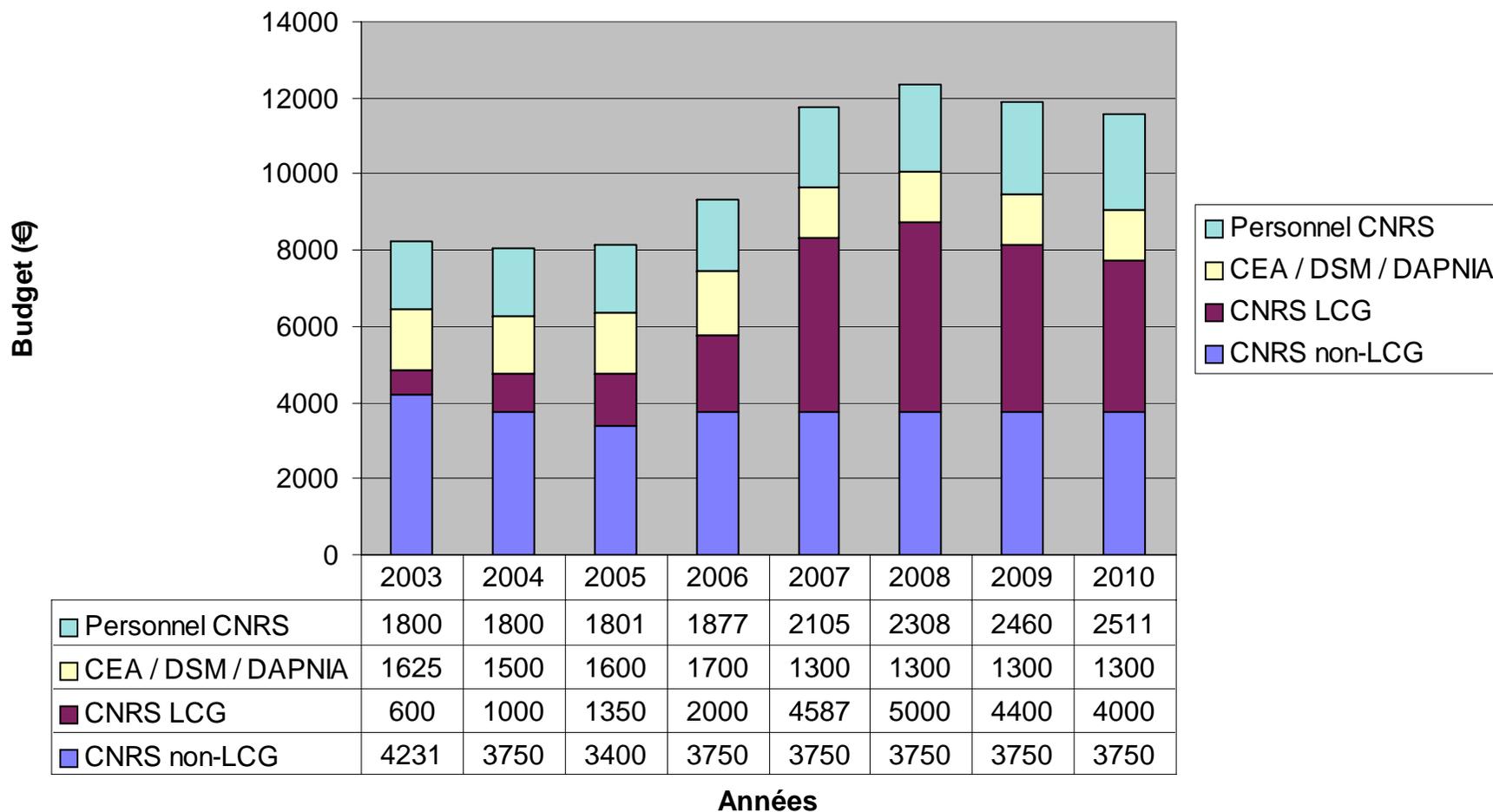
Rapport CDD / Permanents

- 2006: 45%
- 2007: 42%

Les locaux sont maintenant saturés

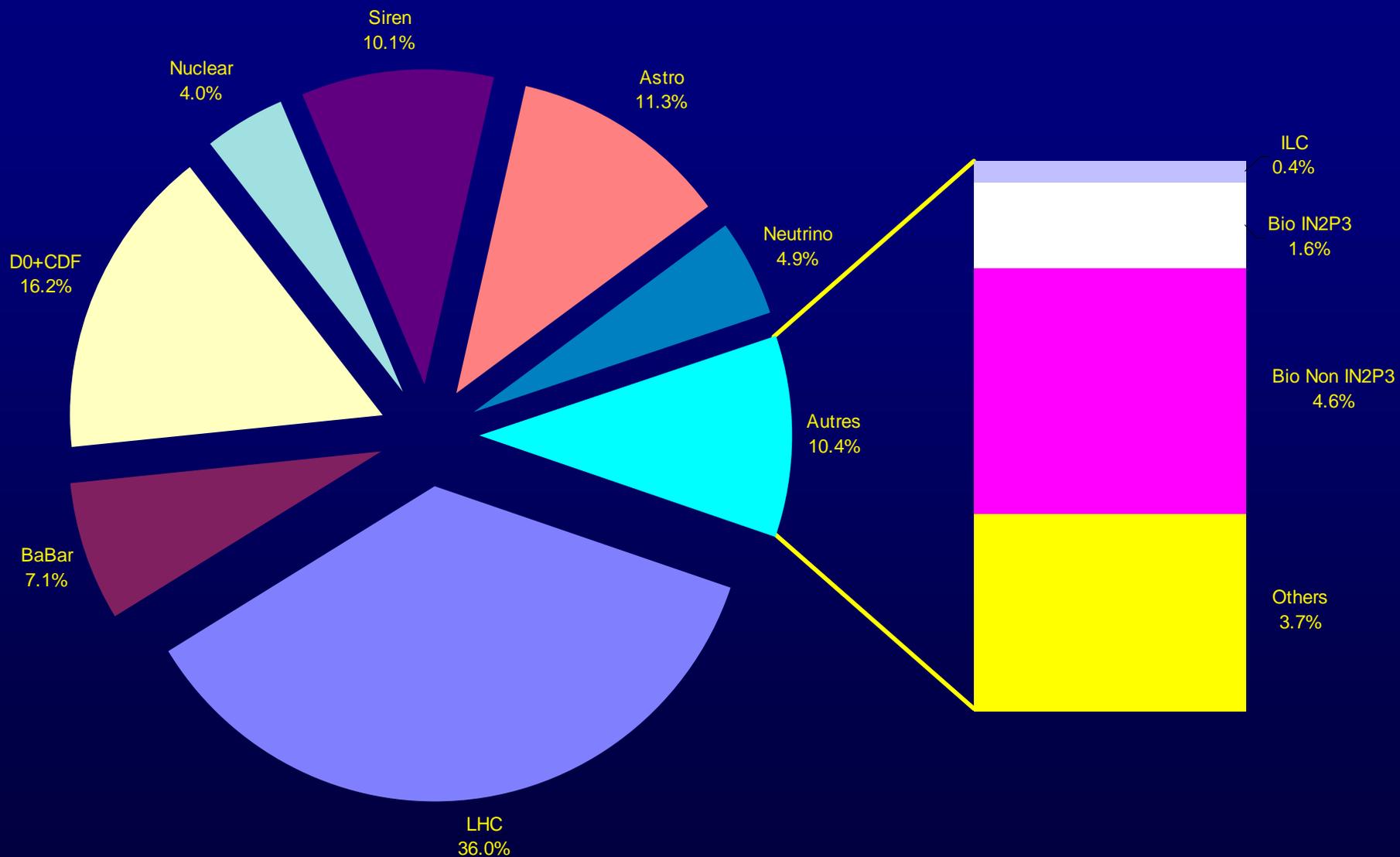


# Budget du CC-IN2P3

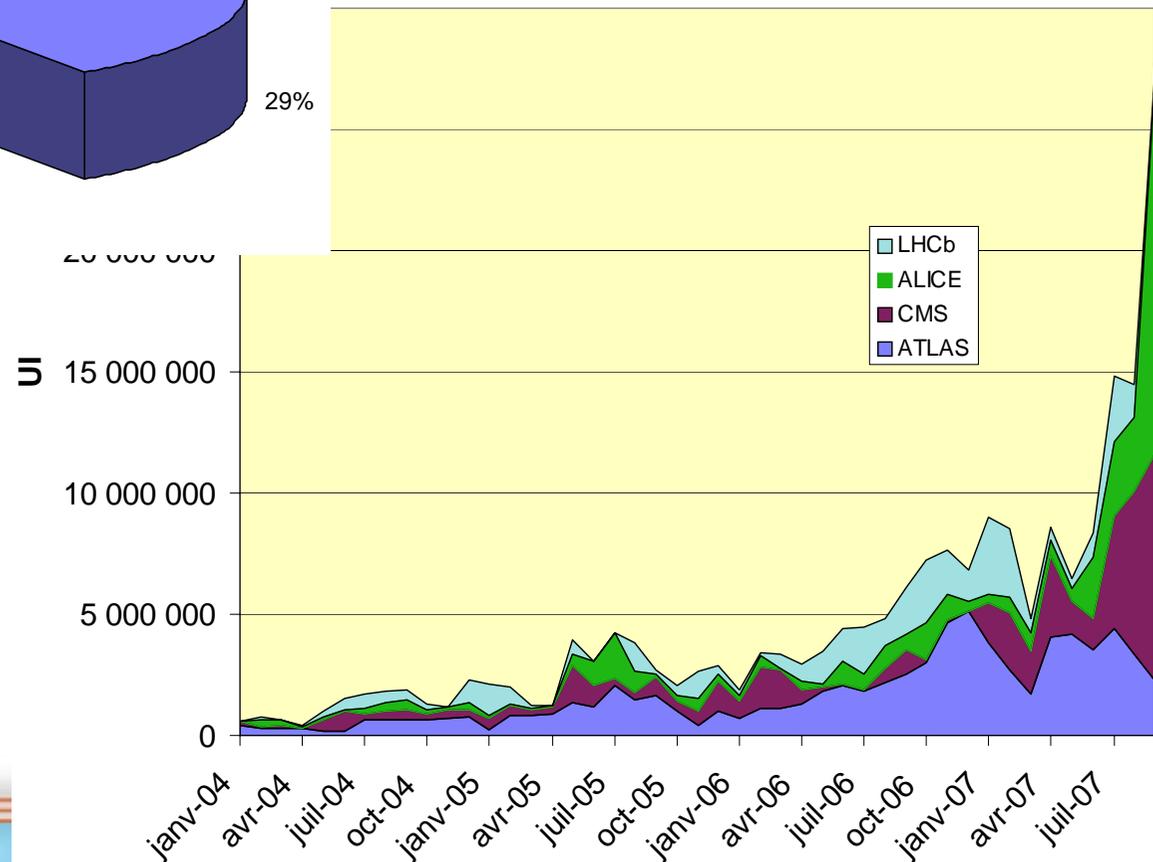
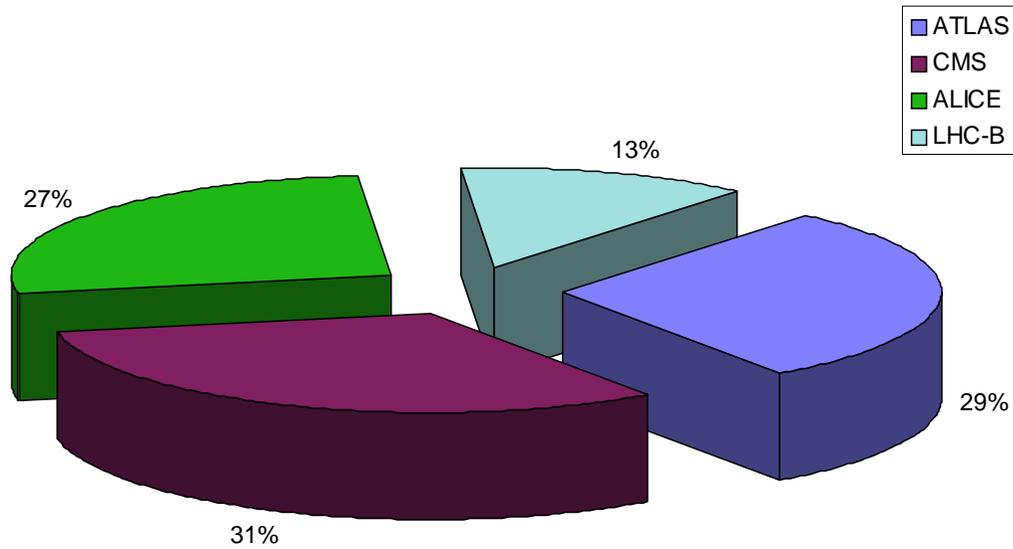


+ budget EGEE ~500 k€/an

# Utilisation du CPU en 2007



# Répartition du CPU entre les expériences LHC

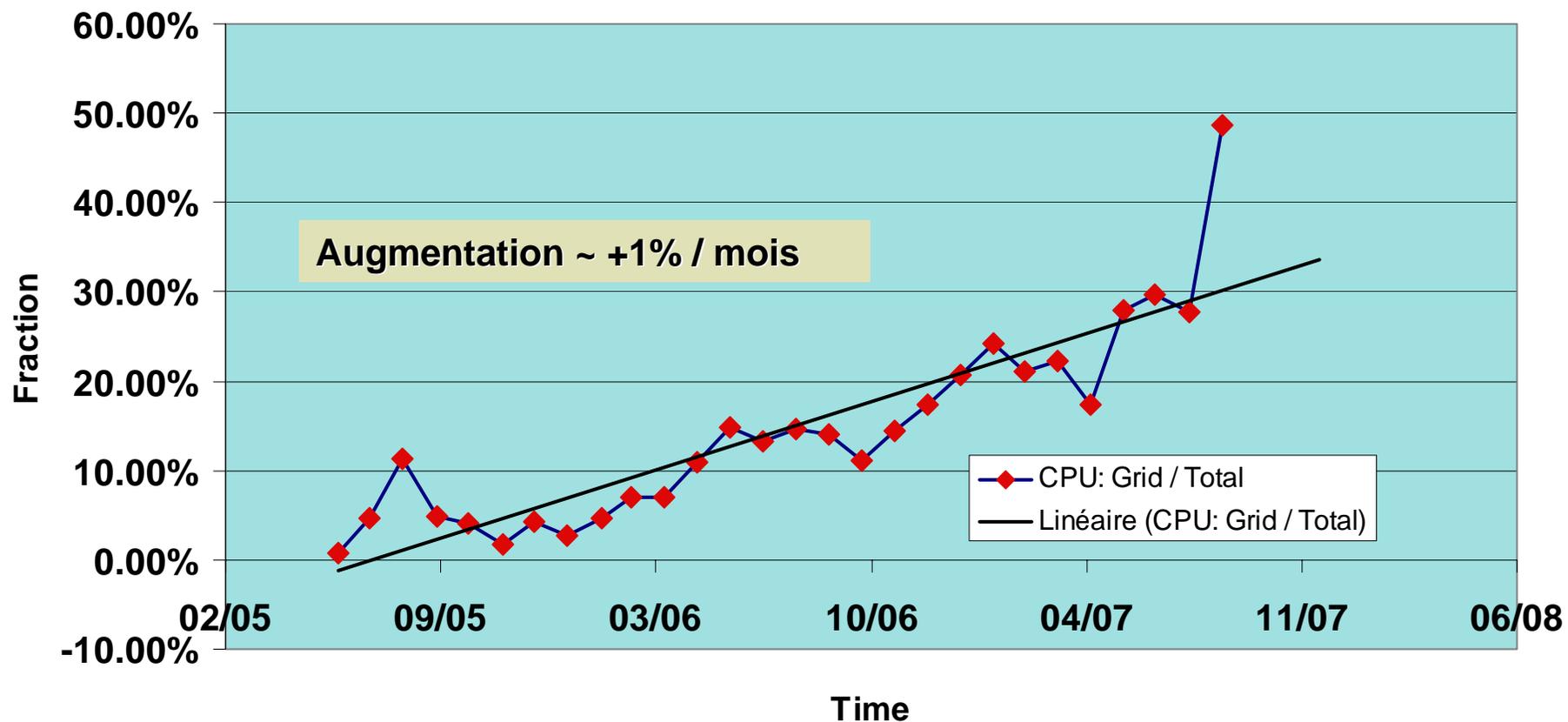




# Utilisation de la grille



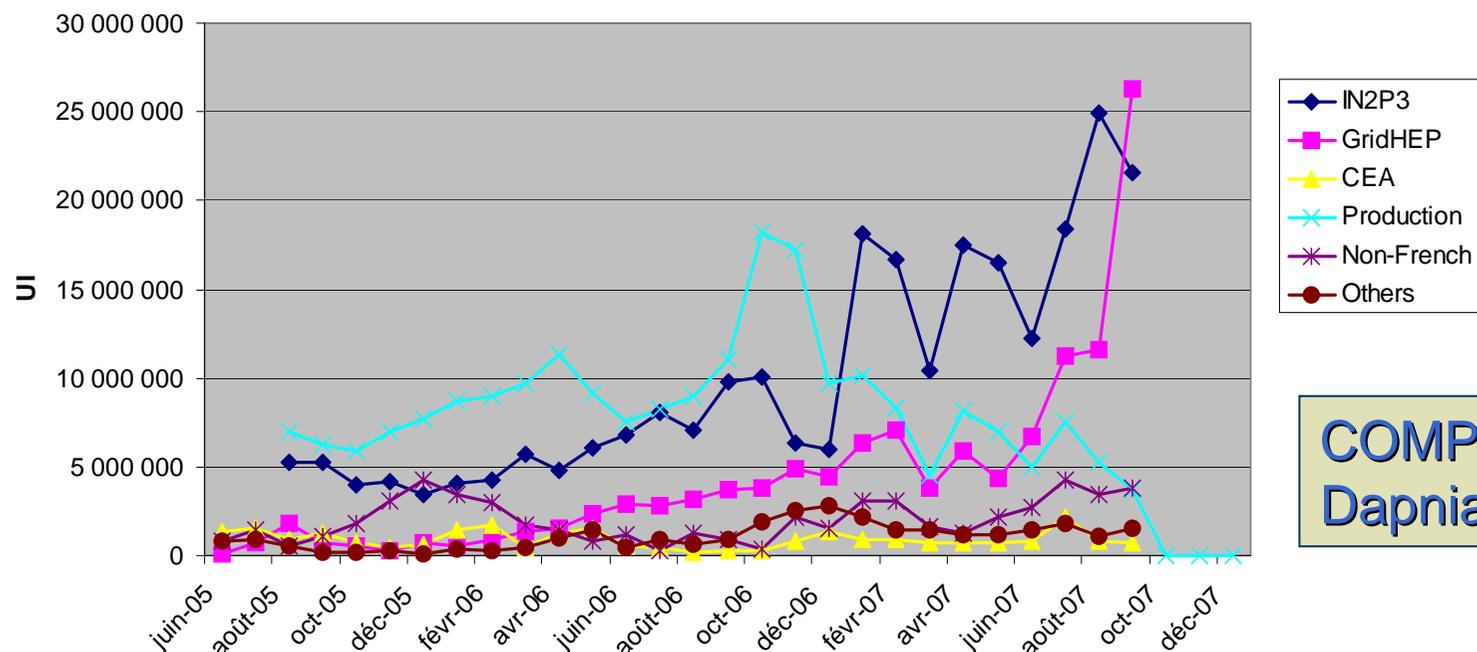
## CPU: HEP Grid / Total





- Plusieurs systemes de stockage sont disponibles au CC-IN2P3
  - D pend du type d'utilisation
  - dCache – xrootd – GPFS – HPSS
- Capacit  totale sur disque: ~2 Po
  - L'essentiel du stockage sur disque au CC-IN2P3 est bas  sur des serveurs SUN X4500
- Capacit  totale sur cassettes: ~10 Po (~2.6 Po utilis e)
  - Augmentation r cente de la capacit  du stockage de masse par l'adjonction d'une nouvelle roboth que

# Utilisation du CC par le Dapnia



COMPASS: 100%  
Dapnia

Moyenne d'utilisation:  $\text{CEA} / (\text{Total} - \text{Production} - \text{Grille})$ : **8.3%**

**Sur les 12 derniers mois:**

- 844 comptes actifs
- 537 comptes actifs hors grille et production
- 41 comptes actifs Dapnia



# Utilisation du CC par le Dapnia



Le CPU n'est pas une bonne mesure de l'utilisation du Centre de Calcul

Principaux utilisateurs du Dapnia:

- Procurs (CLAS): 3 800 kUI
- Elalaoui (COMPASS): 1 900 kUI

Principaux utilisateurs ATLAS:

- Bachacou: 160 kUI
- Lancone: 12 kUI

**L'essentiel du CPU pour ATLAS passe par la Grille**

Mode de fonctionnement des "grosses" expériences:

- Productions réalisées par un petit nombre de comptes

Très grandes masses de données → Fichiers destinés à l'analyse (n-tuples)

**Le CC-IN2P3 est essentiellement un centre de production et de stockage de données**



## Principaux achats récents



### ■ 2006:

- 265 bi-pro / bi-cœur IBM – 685 k€
- Robothèque STK / SL8500 – Capacité de 10 000 cartouches – 840 k€
- 400 To serveurs de stockage SUN X4500 – 640 k€

### ■ 2007:

- 479 bi-pro / quadri-cœur DELL – 1.2 M€
- 1.6 Po serveurs de stockage SUN X4500 – 1.6 M€
- ~500 To stockage GPFS – 1.1 M€
- Achat complémentaire CPU – 800 k€
- Cartouches magnétiques – 800 k€

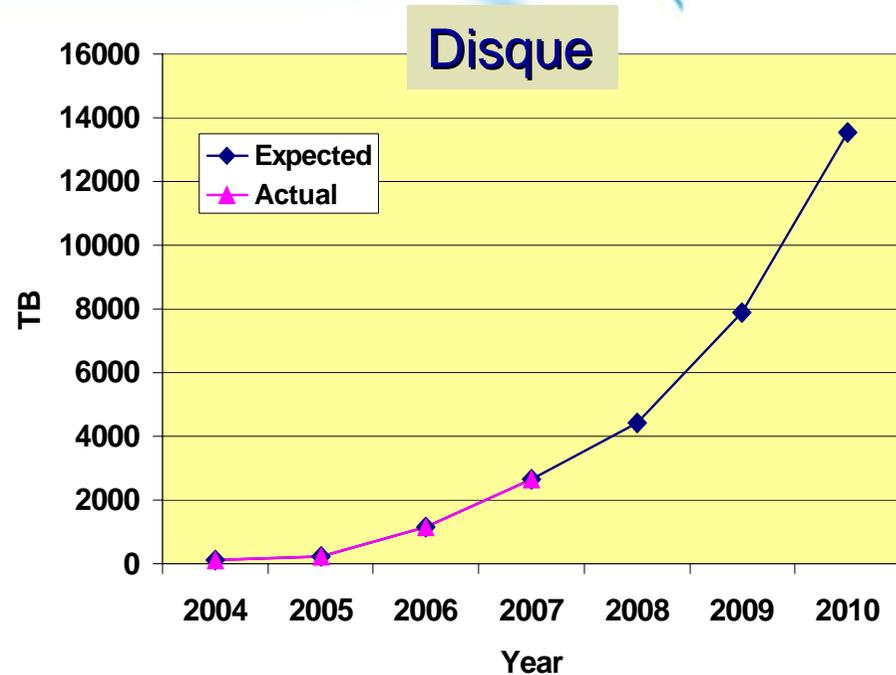
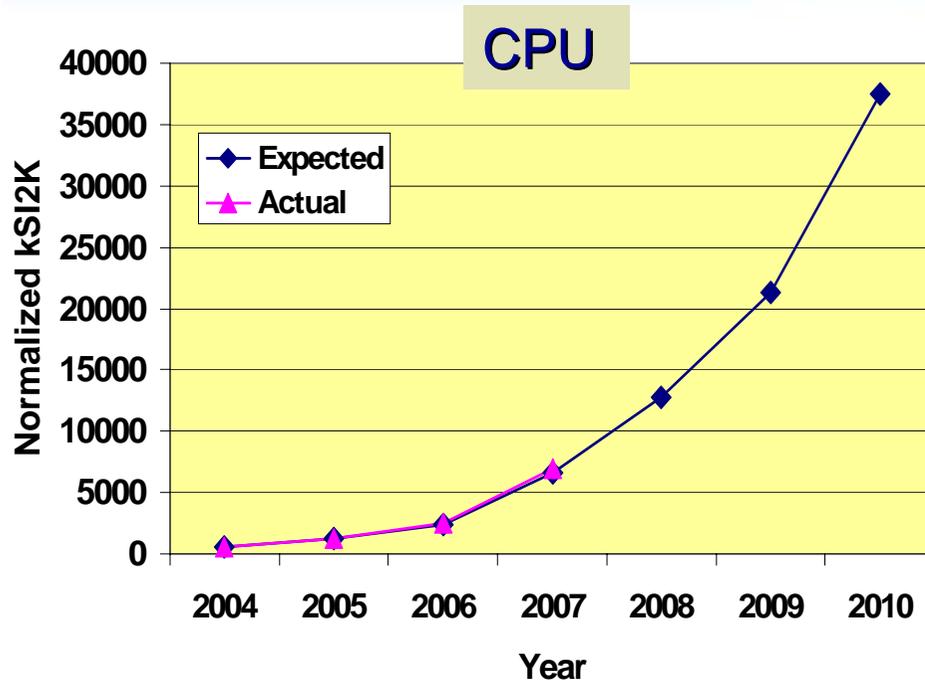


# Gestion des achats



- 1 Ingénieur de haut niveau est chargé de coordonner la procédure d'achat
  - Tests de matériel
  - Négociations avec les fournisseurs
  - Rédaction du cahier des charges
  - Soumission au service marché de la DR7
  - Installation et réception du matériel
  - Mise en service
- Chaque fois que c'est possible, nous faisons en sorte que ces négociations profitent aux labos in2p3 et au Dapnia
  - Achat serveurs de calcul en 2007
  - Négociation de l'achat des licences GPFS en 2006

# Évolution de la capacité informatique du CC-IN2P3



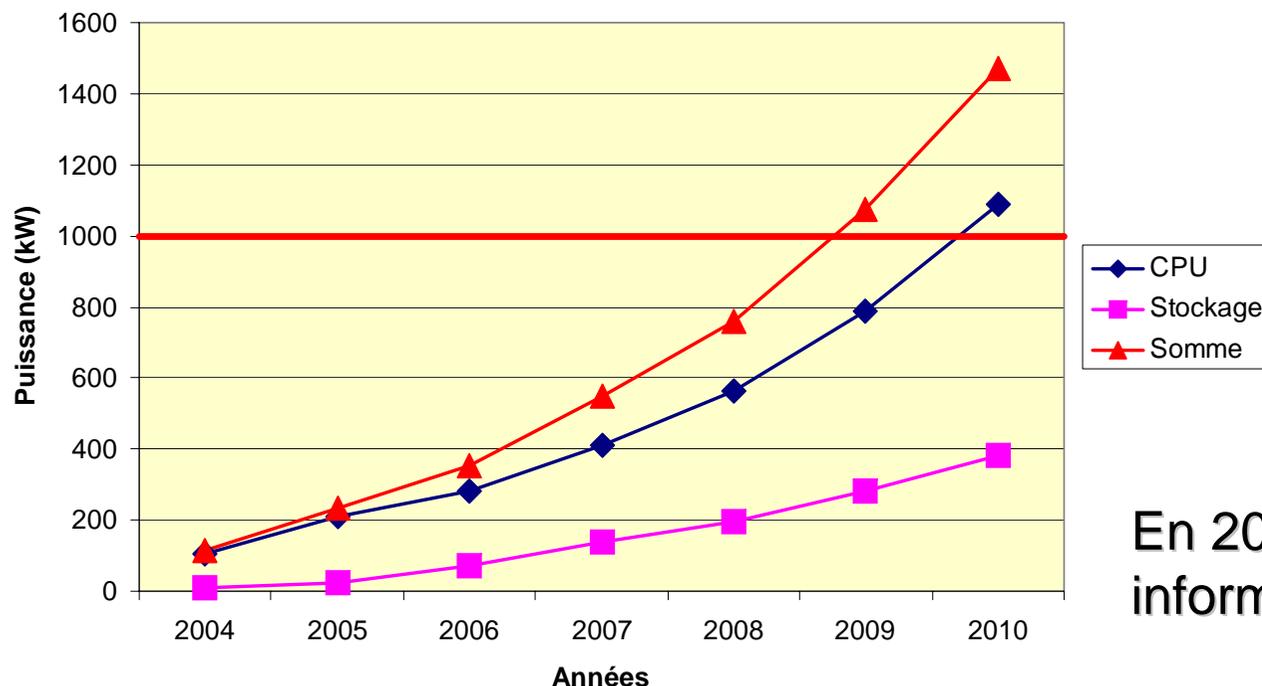
La capacité CPU et disque augmente plus vite que la loi de Moore

→ Implique un changement d'échelle au niveau de l'infrastructure du CC-IN2P3

# Évolution de la puissance électrique



Puissance électrique



En 2009 la salle informatique sera saturée

1 MW correspond à la capacité maximum de la salle machine actuelle (après l'investissement de 1.5 M€ en 2007 pour sa mise à niveau)



# Le projet d'extension



- Doublement en m<sup>2</sup> de la salle informatique
  - 850 m<sup>2</sup>
- Capacité
  - Initiale: 1 MW supplémentaire
  - Extensible jusqu'à 2.5 MW (total 3.5 MW)
  - Nécessite une superficie de locaux technique équivalente à celle de la salle informatique
- Espace bureaux + amphi : ~890 m<sup>2</sup>
- Études de scénarii d'implantation réalisées avec une architecte
  - Les superficies sont réalistes



## Coût



- Une fraction importante du coût est liée aux équipements électriques et climatiques
    - Groupes froids et climatisation
    - Onduleurs
    - Groupes électrogènes
    - Installation électrique spécifique
- ~3700 k€**
- 
- Coût total: **9900 k€ HT (11.6 M€ TTC)**

## Difficultés concernant le projet d'extension



- Toujours en attente de l'arbitrage du CPER
- Le budget risque d'être en deçà des besoins
- Difficulté au niveau de la cession du terrain par l'Université (Loi sur l'autonomie des Universités)
- La non disponibilité de l'extension pour 2010 est identifiée comme un risque majeur pouvant mettre en danger le calcul LHC en France et remettre en cause les engagements du CNRS
  - COS du CC-IN2P3
  - CS de l'IN2P3



# Approche écologique



- Approche responsable et citoyenne pour un équipement qui va consommer une quantité énorme d'énergie et engendrer une forte dissipation thermique
- En ligne avec le projet d'éco-campus de l'Université Lyon I
- Plusieurs pistes:
  - Valorisation de la chaleur produite
    - Circuit d'eau chaude du campus
  - Configuration de la salle machine
    - Isolation (terrasse végétalisée par exemple)
  - Technique de refroidissement
    - Utilisation du froid extérieur en hiver
  - Sources d'énergie alternatives
    - Photovoltaïque ???



## Ouverture du CC-IN2P3



- Plusieurs groupes de biologie / biomédecine travaille activement au CC-IN2P3
  - Génomique
  - Imagerie cérébrale fonctionnelle
  - Embryogenèse
  - Recherches sur la polyarthrite rhumatoïde
  - Etc...
- Le CC-IN2P3 apporte son savoir faire dans le domaine du stockage et des fermes de calcul
  - Rôle fédérateur et structurant
- Changement complet des possibilités des groupes de recherche
  - 1 an de travail sur un PC → quelques jours ou semaines
- Stratégie d'ouverture essentiellement tournée vers la Région et l'Université
  - Projet de création d'une grille régionale dont le CC-IN2P3 serait la tête de pont
- Ouvertures également vers des entreprises de la région



- Hébergement de nombreux équipements
  - Bureau virtuel de l'étudiant
  - Nœud Régional RENATER
  - Machines de la DSI (Labintel...)
  - Équipement réseaux AMPLIVIA – LYRES – LYONIX
- Hébergement de plus de 200 sites Web
- Service de Visio-conférence
- Webcast