



Lundi 18 mai 2009

# Le CC-IN2P3

## Fonctionnement – Évolution - Extension

Réunion LCG-France - Annecy

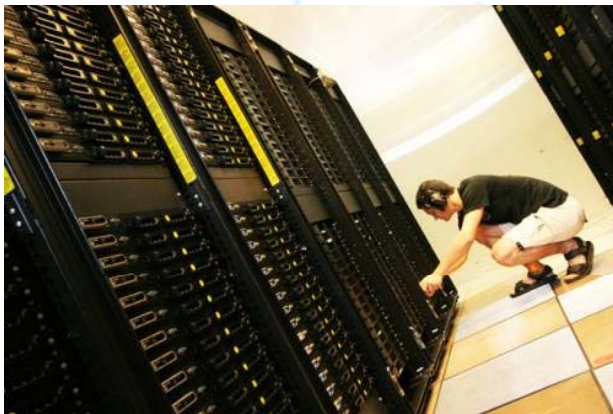


Dominique Boutigny

## Composition actuelle de la ferme de batch



- ❑ 479 DELL dual CPU / quad core INTEL 5345 @ 2.33 GHz – 6.3 MSpecInt2k
- ❑ 401 DELL dual CPU / quad core INTEL 5450 @ 3.0 GHz – 6.5 MSpecInt2k
- ❑ 412 IBM dual CPU / quad core INTEL 5430 @ 2.66 GHz – 6.0 MSpecInt2k



81 kW mesuré incluant  
l'échangeur thermique



85 kW



Refroidissement  
par porte à eau

Lundi 18 mai 2009



# Stockage disque



146 Sun / Thumper X4500 – 17 To utiles / machine → 2.5 Po  
160 Sun / Thor X4540 – 33 To utiles / machine → 5.3 Po  
**Total 7.8 Po de stockage DAS essentiellement pour dcache / xrootd**

Nous avons l'ensemble du stockage disque sur lequel nous nous sommes engagés pour 2009



Environ 15 000 disques !



IBM/DDN DCS9550  
640 To pour GPFS – 480 To pour le cache HPSS



## Stockage sur cassettes



3 robots Sun / STK SL8500 d'une capacité de 3×10 000 cassettes

La capacité théorique est de 30 Po mais:

- 1 robot plein de 9840 (20 Go)
- 1 robot plein de T10k-A (500 Go)
- 1 robot servant de trop-plein pour les T10k-A et normalement dédié au T10k-B (1 To)

Les T10k-B ne seront disponibles qu'après la mise à jour HPSS

Un 4<sup>ème</sup> robot va être installé à la fin de l'année



© CC-IN2P3 / CNRS / f. de Stefanis



# Migration HPSS



La migration de HPSS 5.1 vers HPSS 6.2 est une opération majeure

Elle va se dérouler début juin après une très grosse phase de préparation

## Avantages attendus:

- Version supportée par IBM (la 5.1 n'est plus supportée)
- Support des T10k-B
- Utilisation d'une plateforme matérielle plus performante

L'impact sur l'amélioration des performances devrait être déterminant



# Évolution du réseau



Un effort important a été fait en 2008 pour renforcer l'infrastructure du réseau local → *plus de 800 k€ investis en 2008*



Tous les workers sont connectés en Gbit

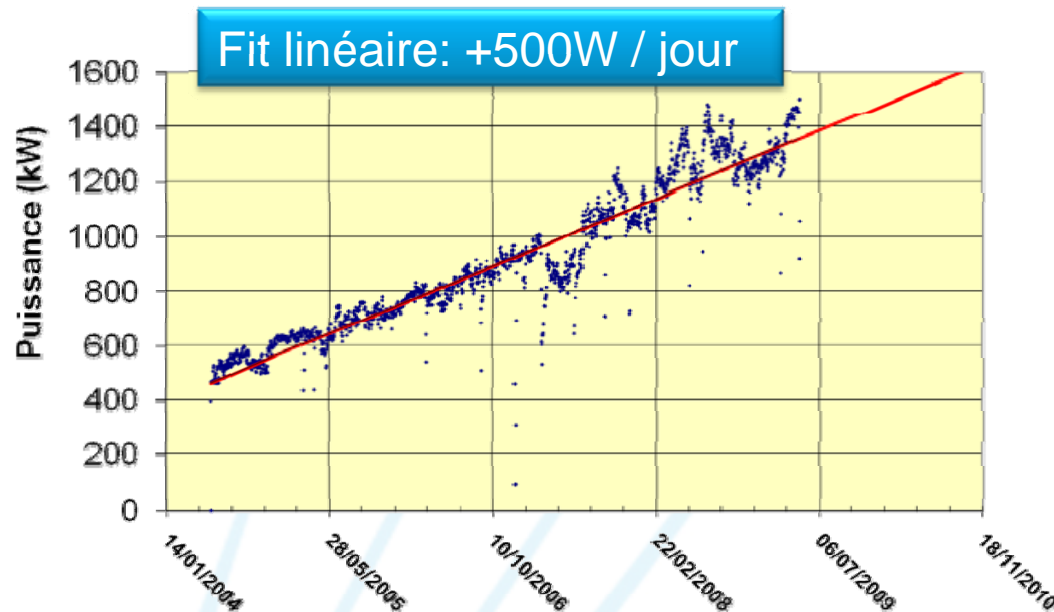
Les serveurs de données les plus sollicités sont connectés en 10 Gbits

Les autres sont en N×1Gbit



Les 3 gros switches de la salle machine sont interconnectés en 4×10 Gbits

# Consommation électrique



En principe la limite de disjonction du CC est à 1550 kW

Nous tenons grâce à un transformateur de chantier qui nous donne une marge de 400 kW (800 dans quelques semaines)

En septembre nous remplaçons les transfos par 3 nouveaux de  $3 \times 1000$  kVA ceci entrainera un arrêt important du CC – Nous essayons de trouver la solution qui impactera le moins possible les utilisateurs

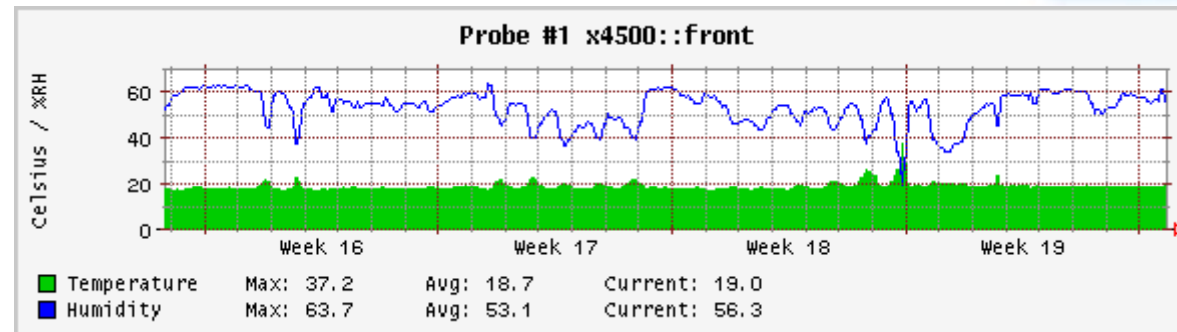
Chaque degré de température extérieure supplémentaire nous fait consommer 20 kW de plus



# Refroidissement



Nous avons actuellement 2 groupes froids capable *théoriquement* de produire chacun 600 kW de froid



Malheureusement l'efficacité des aéroréfrigérants (les gros ventilateurs sur la toiture) baisse lorsque la température extérieure augmente.

Le week-end du 3 mai les groupes froids se sont coupés en cascade en raison de la trop forte pression du fréon dans le circuit de réfrigération (22 bars !)

Un 3<sup>ème</sup> groupe froid va être installé à la fin du mois il devrait nous permettre de passer l'été

L'installation devrait se faire sans impact majeur sur les services grâce à la location d'un groupe froid et d'un groupe électrogène

En attendant la puissance CPU disponible va rester autour de 60% afin de ne pas nous retrouver de nouveau en situation critique





15 mai 2009

Investissement Transfos  
+ Groupes Froids: 550 k€



Lundi 18 mai 2009



## Refroidissement (2)



Nous étudions aussi la possibilité d'installer des refroidisseurs adiabatiques (brumisateurs) sur les aéroréfrigérants



Amélioration sensible des performances par temps chaud et sec

Nous atteignons maintenant les limites ultimes des possibilités de la salle informatique actuelle

→ L'été 2010 va être très difficile...

Nous travaillons en parallèle sur l'instrumentation du système de climatisation afin de repérer les dérives de fonctionnement avant la panne !

Nous sommes malheureusement dans un contexte industriel où la notion de "slow control" est quasiment absente !



## La stratégie pour 2009



Toute la stratégie d'acquisition de matériel au CC doit maintenant être dictée par le souci permanent d'optimiser la consommation électrique et le refroidissement

- Installation de 5 racks au CINES dès que les performances en I/O avec la nouvelle liaison à 10 Gb/s seront validées (incertitude sur GPFS)
- Remplacement de la ferme de 479 DELL PowerEdge 1950 (< 3 ans) par un système refroidi à eau  
Ce remplacement inclura l'acquisition d'un cluster parallèle InfiniBand (Pour Pistoo)
- Acquisition de racks refroidis à eau pour intégrer le restant des DELL les plus récentes
- Capotage et optimisation du refroidissement des Thors
- Migration des Thumpers vers les Thors et élimination des Thumpers
- Acquisition du matériel pour la ferme d'analyse en fonction du résultat de l'étude en cours (en utilisant le marché national DELL si possible)



# La stratégie pour 2009



Au niveau des services la stratégie est de renforcer les performances et la robustesse partout où c'est possible

Migration HPSS

Optimisation des débits Cassettes → Disques

Renforcement d'AFS

Mise en place de redondances sur les services critiques



# La nouvelle salle informatique



Nous bénéficions de 900 k€ dans le cadre du plan de relance qui doivent impérativement être dépensés cette année

Le projet avance bien et les blocages administratifs sont levés (toutefois, le terrain n'est toujours pas attribué officiellement au CNRS)

Le programme est quasiment terminé et nous allons lancer une consultation en vue d'une procédure de conception-réalisation

On vise une mise en service progressive de la nouvelle salle à partir du début 2011.  
Ce planning est tenable si toutes les opérations s'enchainent sans la moindre faille



# Future salle informatique



Surface d'environ 900 m<sup>2</sup>  
sur 2 niveaux

Le budget est très serré → le bâtiment sera extrêmement simple (type charpente métallique + bardage)

Puissance électrique limitée au début (~500 kW pour l'informatique) mais infrastructure évolutive jusqu'à >3 MW pour l'informatique (5 MW total) en plus de la salle actuelle (total ~6.5 MW)

Récupération des calories produites pour le chauffage de bâtiments universitaires



# Refroidissement de la nouvelle salle



Travail d'investigation pour identifier les meilleures solutions de refroidissement

Refroidissement à eau pour toute la nouvelle salle - A priori, confinement des allées chaudes



25 à 30 kW par rack



A priori pas de faux plancher – Électricité, eau glacée et courants faibles arrivent par le haut