

Commentaires sur les chiffres et la présentation de Les Robertson au RRB
Avril 2004 / DL

Pièces jointes, outre cette note :

- un tableau dynamique des coûts estimés 2002-2008 de LCG au CC
- un tableau des ressources qui ne fait que reprendre les données de Les Robertson au RRB

Le rapport Hoffmann a fait des prévisions pour 2007 qui étaient d'environ :

~30 PB bande, 10 PB disque, 75 M SI2K

Chaque tier1 assurant ~10% de chacun de ces chiffres, soit :

~3 PB bande, 1 PB disque, 7 M SI2K

Les prévisions de Les pour 2008 :

18 PB bande, 20 PB disque, 45 M SI2K

restent dans les mêmes ordres de grandeur, elles font cependant la part belle au disque.

Si l'on table sur 20% au lieu de 10% pour un Tier1 (conséquence du petit nombre de Tier1 compensé par un grand nombre de Tier"1,5"), cela donne

3,5 PB bande, 4 PB disque, 9 M SI2K

cela reste compatible avec le rapport HH, sauf pour le disque.

selon le tableau de coûts ci-joint 2002-2008, j'arrive à un total de : 8 Millions € de hardware étalés sur les 3 ans si le CC doit disposer des ressources prévues par Les R. Cette somme représentant 50% des investissements CPU/disque/bande du CC. En ajoutant un petit % pour le reste des équipements, on approche de 20M. Disons 5-6 M par an sur cette période.

Cela fait une grosse augmentation de budget pour le CC.

En 2004 déjà, le budget du CC ne permet pas de répondre à la demande et les expériences vont sérieusement le ressentir.

Pour y arriver, en supposant que LCG sature à 50% du CC après une phase de montée en puissance, il faut :

DISQUE : je dois faire croître le disque plus vite que prévu, un facteur 2 (=+100%) au lieu de 70% sur toute la période 2002-2008.

Remarques : 1) ceci n'est pas compatible avec la réalité 2004.

2) On peut aussi penser à accélérer en fin de période, en gardant 70% jusque 2006 et plus que 100% au-delà. Le tableau permet de jouer avec les paramètres

3) Qu'est-ce-que veut dire un modèle où disque=bande, autant faire du tout sur disque dans ce cas-là, ce serait plus simple que l'usine à gaz d'un stockage de masse légitimé par le gain financier.

On a beaucoup critiqué Babar et son modèle Disque/Bande = 30% (contre 5-10% max. pour les autres)

Le rapport Disque/Bande n'est pas très différent au SLAC (Tier0+1) et au CC-IN2P3 (Tier1). Or au Cern, Disque/Bande=25%

Il faudra qu'on m'explique ces 80 M€ (120 MCHF) de disque.

4) Egalement, si disque/bande=1, je ne pense pas qu'il soit nécessaire de prévoir autant de serveurs de disques car pas mal de données ne seront que peu accédées. Le coût disque/serveurs peut sans doute être réduit. Les coûts du tableau correspondent à de grosses bandes passantes en I/O, nécessitant pas mal de serveurs et de quincaillerie.

CPU : le facteur 2 annuel est ici OK, on peut même réduire un peu, à 1,9 par an.

Remarque : la demande CPU au CC ne faiblit pas, elle reste à +100%/an depuis dix ans.

BANDES : vu la réduction prévisionnelle de Les, le facteur 2 annuel peut être réduit de 100% à 75%.

Remarque : la demande bande au CC, grâce à un bon HSM (HPSS), accélère plutôt que de ralentir, nous sommes au-dessus de +100%/an.

Dans le tableau excel joint, je te laisse "jouer" avec les cases violettes.

Comme signalé, les "annual increases" CPU-Tape-Disk du tableau de l'an dernier étaient de 2,0 - 2,0 - 1,7, Les R. inverse le taux de croissance disque/bande. C'est pour marquer la différence Tier0 versus Tier1, mais cela se vérifiera-t-il dans les Tier1, qui sont tous dotés d'un gros Mass storage, j'en doute fort au vu des coûts. Les bandes à 1TB devraient succéder en 2005-2006 aux bandes actuelles à 200GB (75€).

Dans le tableau des ressources, les coûts estimés par le CERN tablent en 2008 sur :

3,25 CHF / GB disque (2,2 €)

0,6 CHF / GB tape (0,4 €), soit un rapport 5 assez comparable à aujourd'hui

1150 CHF / K SI2K (750 €), soit 1500 € pour un bipro (3 GHz) de la puissance actuelle en 2008, ça me paraît beaucoup.

Dans mon tableau de prix, les coûts en 2008 sont estimés à

1,8 € disque (LR : 2,2€) OK

0,25 € bande (LR : 0,4 €) un an de décalage (Moore)

172 € / K SI2K (LR : 750 €) 3 ans de décalage

Si les chiffres de Les R. sont les bons (et que je n'ai pas fait d'erreur ☺), il faut revoir encore à la hausse les budgets du CC-IN2P3.

Je n'ai pas tenu compte de l'incidence éventuelle d'un Tier2 à Clermont.

DL