

Ajout de disques dans les serveurs CPU - un gain ?

P.Hennion

LCG-FRANCE – juin 2018



Constat au LLR:

On LLR on achète :

- Des serveurs de stockage contenant 36 DD
- Des serveurs CPU contenant 1 DD et 2 ou 3 emplacements vides pour des DD supplémentaires.

Question :

Pourquoi ne pas utiliser ces emplacements vide pour notre stockage ?

Remarques:

Plus cher d'utiliser les serveurs ayant seulement 2 emplacements vides -> pas étudié .

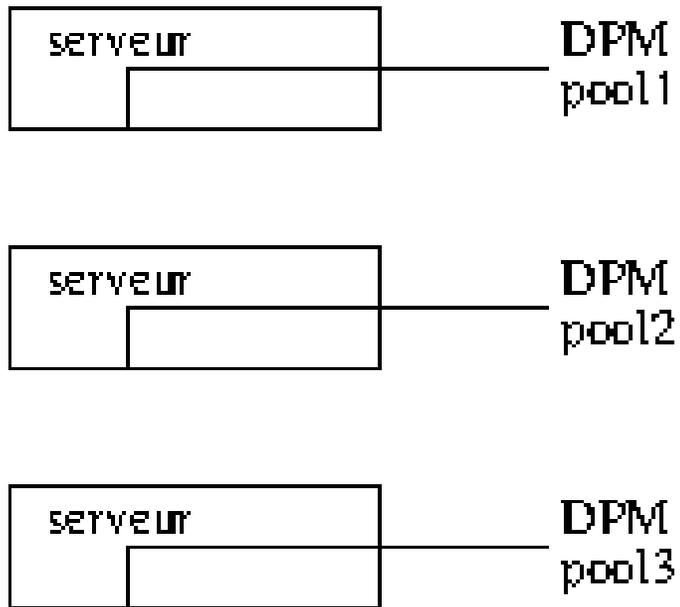
Perte de densité pour les serveurs de calcul ayant 3 emplacements vides

Au LLR, les serveurs de stockage sont ondulés, le calcul ne l'est pas

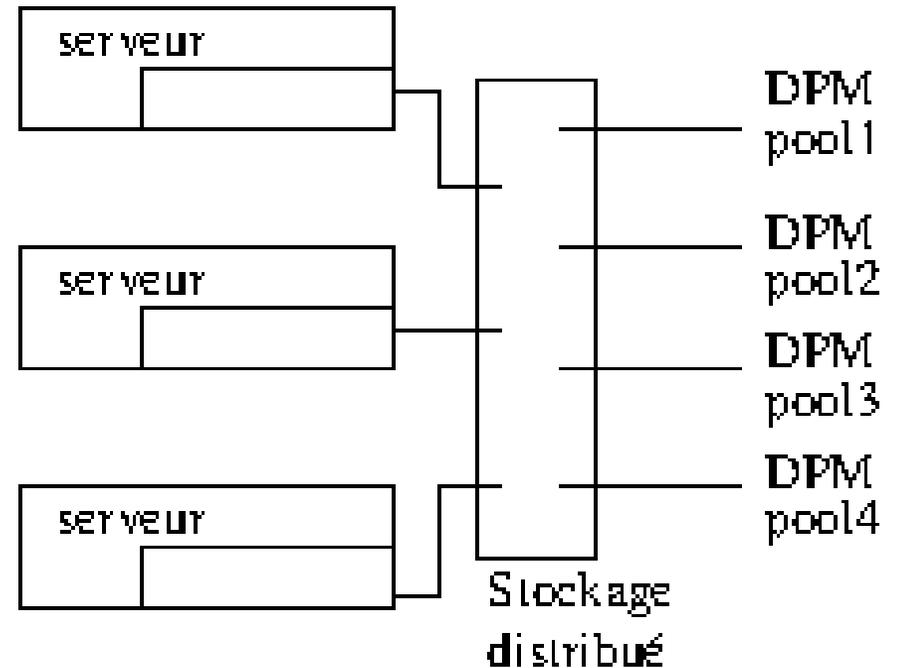
Etude faite pour le LLR. A revoir au cas par cas.

Que peut-on faire de ces disques ?

RAID 5 sur chaque serveur



Stockage distribué



Si on ne compte que le prix des disques:

- Raid5 sur les 3 disques durs. Chaque serveur de calcul devient aussi un nœud de stockage
 - 11% d'économie
- Stockage réparti entre les serveurs de calcul
 - Si on utilise 20% des disques pour répliquions/erasure-coding (à vérifier) :
26% d'économie

Mais (Raid 5)

A ce prix il faut ajouter

- Onduleur pour tout le calcul
- La partie du CPU qui sera utilisé par le stockage
- La perte de densité

Et on perd en sécurité :

- Raid 5 au lieu de Raid6
- Multiplie par 15 le nombre de serveurs -> multiplie d'autant le risque de perte de données
- Installation de DPM sur tous les serveurs

Mais (stockage réparti)

A ce prix il faut ajouter

- Onduleur pour tout le calcul (sauf si stockage réparti sur plusieurs sites)
- La perte de densité
- La partie du CPU qui sera utilisé par le stockage (ajouter ~5% au prix du serveur)
- La RAM utilisée par le stockage (ajouter ~5% au prix du serveur)
- L'augmentation de la bande passante réseau (réseau privé dédié ?)
- L'ajustement des paramètres de répliquions/erasure-coding qui permettent d'obtenir la même sécurité qu'actuellement (calcul fait avec 20%)

Et on gagne en complexité

Conclusion

Une alternative serait d'acheter des serveurs avec les mêmes mémoires et CPU que nos serveurs CPU actuels mais dans lequel on peut mettre jusqu'à 16DD

Actuellement, pour le LLR, le surcoût serait de 10%

D'autres scénarios doivent être envisagés

Etude à faire par chacun des sites

CEPH recommandations

- 1 CPU/DD
- 1Go mémoire /1To de DD
- Ceph monitor
- Disques SSD
- Réseau privé 40Go