

De l'importance des IOPS dans le choix des systèmes de stockage

Posted on [30 janvier 2012](#) by [Nicolas Bouyssy](#)



En matière de virtualisation, le stockage tient une part prépondérante dans les performances générales des VMs.

En effet les problèmes de performances liées à la surcharge des hyperviseurs sont faciles à détecter, et les problèmes liés au dimensionnement du réseau sont rares même en ethernet 1gbps, dans le cas où les serveurs sont équipés de cartes en nombre suffisant et que l'hyperviseur est bien configuré pour les gérer intelligemment.

Le goulot d'étranglement principal est potentiellement le nombre d'IOPS (Input/output per seconds) que le datastore va pouvoir gérer en même temps.

Voilà pourquoi on assiste parfois à des « boot storms », des « antivirus storms » ou des « cron storm », qui se produisent quand des milliers de VM tentent d'accéder en même temps au stockage, par exemple pour un check antivirus.

Ceci a pour effet direct de saturer la baie de stockage en I/O et donc de générer des latences sur les disques des VMs. Les performances de toutes les VMs situées sur le datastore saturé subissent une énorme dégradation de performance.

Il est bien sûr important d'étaler dans une journée toute action automatisée sur les VMs qui aurait pour effet de faire beaucoup d'accès disques, plutôt que de tout lancer à la même heure. Mais il faut aussi s'intéresser de très près à la capacité de ses baies de stockage à délivrer un grand nombre d'IOPS simultanément.

Outre les capacités mémoire ou de CPU des diverses baies de disque, ou le type de transport et protocoles utilisés (iSCSI/Fibre/NFS+ethernet), le premier point à étudier est probablement la capacité des disques à délivrer un grand nombre d'IOPS, ainsi que le niveau de RAID utilisé.

Chaque disque peut délivrer un nombre donné I/O par seconde, en fonction des paramètres suivants:

- **La vitesse de rotation du disque**

- **la latence moyenne en ms:** soit le temps que la tête va prendre pour être positionnée sur le secteur à lire ou à écrire.

- **le seek time moyen en ms:** soit le temps dont le disque a besoin pour effectivement lire ou écrire sur le secteur. (différents pour l'écriture et la lecture)

Une formule communément utilisée pour calculer les IOPS d'un disque est la suivante:

$$\text{IOPS Estimated} = 1 / ((\text{seek time} / 1000) + (\text{latency} / 1000))$$

exemple pour un disque seagate ST3600057SS Cheetah 15K.7 SAS 6-Gb/s 600-GB Hard Drive vendu pour

Average latency 2.0ms

Random read seek time 3.4ms

Random write seek time 3.9ms
 soit moyenne seek time 3.65ms
 $IOPS = 1 / ((3.65/1000) + (2.0/1000)) = 176,99$ IOPS

Autre élément à prendre en compte, le niveau de RAID utilisé. En effet par exemple le RAID 1 (mirroring simple) va faire 1 I/O pour chaque lecture, mais 2 pour chaque écriture ! Et c'est bien pire en RAID 5 ou 6:

I/O Impact		
RAID level	Read	Write
RAID 0	1	1
RAID 1 (and 10)	1	2
RAID 5	1	4
RAID 6	1	6

Et afin de compliquer encore un peu les choses, la pénalité due au raid ne s'applique que sur les écritures (voir tableau ci dessus) et le rapport entre nombre d'accès read/write dépend également de l'usage de chacun.

La formule suivante permet de calculer le nombre d'IOPS réel d'une VM ou d'un serveur ESX coté baie de stockage:

$(Total\ Workload\ IOPS * \% \text{moyen de Read IO}) + (Total\ Workload\ IOPS * \% \text{moyen de Write IO} * \text{Pénalité due au RAID})$

disons qu'une VM a une moyenne de 50IOPS , 40% en read 60% en write et que l'on a une baie en RAID 5 (pénalité 4 en write)

$(50 * 0.4) + (50 * 0.6 * 4) = 140$ IOPS

Donc cette VM générant 50IOPS coté hyperviseur va en réalité en générer 140 coté baie de stockage!

On comprend donc tout de suite qu'à-fin de maximiser le nombre d'IOPS il convient plutôt de trouver le juste équilibre entre performance IO individuelle des disques/taille des disques/niveau de raid , plutôt que de se dire simplement qu'il vaut mieux prendre les plus gros disques possibles, et ceux qui tournent le plus vite.

Ci dessous une table des IOPS généralement constatées par type de disque:

RPM	IOPS
SSD	6000
15K	175
10K	125
7200	75
5400	50

Il est donc primordial de connaître son infrastructure virtuelle et surtout de bien comprendre ce qui s'y passe au niveau IO, principalement quand on utilise des disques non SSD. C'est une clé essentielle pour obtenir de bonnes performances globales.

Au vu des chiffres ci-dessus, quelque chose me dit que le SSD ne va pas rester confidentiel très longtemps...