

Moshe Yanai, IBM XiV, ou la ré-invention du stockage grâce au 'grid'...

Moshe Yanai : tous les pros du 'data storage' connaissent cette grande figure à qui l'Histoire reconnaît au moins trois hauts-faits : le principe des architectures de disques avec blocs de data fixes, sur les 'mainframes', c'est lui. La conception du Symmetrix (le haut de gamme d'EMC), c'est encore lui. Et depuis 2003, il s'est ré-investi dans une invention originale: un système de stockage « grid ». Il a parachevé son développement et son industrialisation au sein d'une start-up, démarrée à Tel-Aviv -la société 'XiV' (*) qu'IBM est venu lui racheter début 2008.

Chez Big Blue, son innovation a été prise très au sérieux puisqu'elle participe du tournant technologique que le groupe désigne depuis quelques mois comme « L'architecture du 21^è siècle ».



Moshe Yanai nous explique tout d'abord que face à l'explosion des données à stocker (de 161 exabytes en 2006 à 988 exabytes en 2010, source IDC), on doit s'inquiéter de 5 facteurs déficients ou manquants: la fiabilité des solutions, les performances, les fonctionnalités, l'« administrabilité » de ces solutions et l'évolution des coûts de façon globale.

Critique appuyée des disques 'SSD'

Notre expert continue de croire dur comme fer... aux disques durs et à l'évolution encore possible de leurs technologies. Il conteste vivement les affirmations qui se répandent et soutiennent que les 'disques flash' ou SSD (solid state disks) vont remplacer les disques durs.

« Il sera toujours plus rapide d'accéder aux données, en grand volume, sur un disque, disposées électromagnétiquement de façon séquentielle et linéaire, que sur une cartouche 'flash', car ici, à l'inverse, chaque

lecture des trames d'information nécessite des déplacements sur des « chemins » de circuits électroniques, donc des temps d'accès physiquement incompressibles. »

Il ajoute pour être bien compris :

« Même si les mémoires 'flash » continueront de se développer sur les appareils de grande consommation (appareils photos, typiquement) avec des capacités décuplées (de 10 Go on atteindra 3 ou 5 fois plus), elles ne remplaceront pas les disques. Je suis prêt à prendre le pari sur les 20 années qui viennent!« .

Alors, là, vous n'insistez plus... Votre interlocuteur sait fort bien qu'EMC, parmi d'autres leaders, promeut des disques SSD. Mais chez IBM, le parti semble pris: leur usage sera limité ou réservé à certaines fonctions de complément mais sûrement pas à du stockage de masse!

Un 'storage grid' : pourquoi révolutionnaire?

Prêt à relever les défis du 21^e siècle, Moshe Yanai se passionne à expliquer quel tournant radical représente son invention, imaginée au début des années 2000 et concrétisée au sein de sa société XiV, aujourd'hui rachetée par IBM.

« La plupart des constructeurs, explique-t-il, conçoivent les solutions de stockage de données dans une logique de cycle ILM (information life management), louable en soi notamment pour hiérarchiser les données, mais la démarche a ses limites. On classifie, certes mais il faudrait 10 catégories, peut-être et on ne résoudrait pas pour autant cette inflation quasi-incontrôlable du volume et donc de la capacité et du 'thin provisioning' à mettre en face. Car on risque toujours de n'utiliser réellement que 50% des ressources que l'on a achetées!« .

Face à la virtualisation...

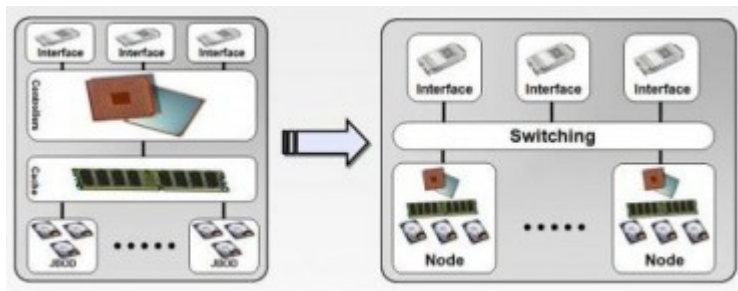
Certes, lui rétorque-t-on, mais les solutions de virtualisation sont censées résoudre une bonne partie du problème, non?

« Oui est non, car beaucoup installent cette virtualisation au mauvais endroit; donc c'est souvent inefficent! »

Bref le salut passerait pas une solution « grid », la seule capable de prendre le relais, efficacement, des systèmes en 'cluster'.

Qu'en est-il au juste?

« Dans les clusters de systèmes de stockage [ou grappes de serveurs dédiés], le process CPU, la mémoire et l'intelligence sont centralisées et pilotent l'ensemble des ressources systèmes (qui peuvent être réparties, redondantes, virtualisées..) mais tout est rapporté sur un système central unique. Notre architecture 'grid' [matricielle, répartie, avec des process massivement parallèles] change radicalement la donne, car elle consiste à placer en commutation permanente et en temps réel autant de sous-ensembles intelligents, autonomes et redondés, chaque sous-ensemble étant constitué d'un serveur avec sa propre mémoire, son ou ses disques avec ses contrôleurs et son interface vers le reste de la structure 'grid' » en inter-commutation au niveau de 'enregistrement des blocs de données« .



A gauche, une architecture de cluster et, à droite, l'architecture 'grid' de XiV (IBM)

Quels avantages, en pratique?

On comprend donc que cette architecture fédère des unités redondantes et redondées, grâce à du processing massivement parallèle, grâce à des algorithmes de distribution et de répartition des données sur ce 'grid', avec multi-écritures simultanées, c'est à dire une réplication des données localement ou à distance.

« Il n'y a plus cette fragilité d'un système en cluster où la mémoire cache est centrale et unique » .

Les avantages d'un tel dispositif seraient multiples :

« On peut réellement faire du 'thin provisioning' [ou allocation fine des ressources disques] ; entre autre raison « parce que les volumes de données dissociés ou 'orphelins' n'existent plus » ; on peut faire de la réplication à distance. L'ensemble est plus simple à administrer. Et les économies d'énergies sont au rendez-vous. Un rack XiV, avec des disques SATA de 146 GO ou plus ne consomme que 42 watts, contre 380 w sur des configurations en grappe équivalentes » .

Cette solution offre également des fonctions d'auto-réparation, d'auto-réglage et d'évolutivité [scalability] dynamique.

Bref, on comprend que cette solution vise le stockage de très gros volumes, a priori en configurations de réseaux SAN (Storage area networks) en évitant le coûteux Fibre Channel. L'offre originelle de XiV s'appelait Nextra ; Ce système de stockage en architecture massivement parallèle autorise des capacités allant jusqu'à plusieurs peta-octets (milliers de tera-octets!).

(*) XiV a rejoint la division System Storage d'IBM fin 2008.