

Salon Linux: DataDirect Networks, les tera-octets à la pelle...

A l'origine de DataDirect Networks, PMI qui compte aujourd'hui 150 personnes, fut Megadrive, en 1988, un fabricant français de serveurs à architecture sécurisée (RAID) avec interface à très haut débit (

Fibre channel). En 1998, la société acquiert Impact Data, un spécialiste de la virtualisation -déjà- de partage de réseaux SAN en environnement OS hétérogène. La fusion des équipes donnera naissance à DataDirect Networks. La nouvelle entité s'est très vite concentrée sur les unités de stockage de grande capacité (30 To en juin 2000 à la NASA) avec lecture/écriture très rapide, des unités capables de soutenir la lecture de fichiers vidéo en temps réel (*streaming*). En 1998, sur cahier des charges de la NASA, fut conçue une 'appliance' (S2A, *Silicon storage appliance*) pour des réseaux locaux de stockage (SAN), très performante en écriture, utilisant des architectures redondantes de disques (RAID). L'objectif visé fut tenu: des débits de 1000 à 1500 Mo/sec (sur 'théorie' de 1.600) en lecture/écriture. « *Ces dix dernières années, les capacités de stockage ont été multipliées par 300. Mais, paradoxalement, les temps de latence n'ont pas changé: 6,5 milli-secondes. Et, pire, le temps de reconstruction d'un secteur défectueux, par exemple, a presque doublé, grimant de 1,5 seconde à 2,5 secondes* », constate Laurent Thiers, dg de DataDirect Networks France. **Et l'apport de Linux dans ce contexte? Sans être la panacée, Linux a apporté beaucoup.** « *Il est reconnu que les 'clusters' sous Linux [serveurs en grappes] sont les plus rapides du monde comme BlueGene/L de Lawrence Livermore utilisant Linux et 1 péta-octets de nos disques* », souligne L. Thiers. « *Cet 'operating system' colle bien avec les traitements processeurs 'concurrents' et le parallélisme des ports de nos «S2A»* ». Parmi les témoins des bonnes performances obtenues figure Bull. L'application réalisée **au CEA avec «Tera 10»**, et présentée à la presse il y a quelques semaines [cf. notre article], l'illustre bien: 8000 disques de 146 Go, pilotés par 54 modules seulement d'entrée/sortie, des «S2A 9500» de DataDirect Networks. En arrière plan c'est un Linux renforcé par Bull qui supporte **Lustre**, le gestionnaire de fichiers distribué pour grappes de serveurs (de 'CFS', Clustered File Systems). « *Lustre fut à l'origine un projet de recherche de l'Université Carnegie Mellon (projet « Coda »), puis poussé par le DoD et le DoE américains (Department of Defense, Department of Energy respectivement) notamment Los Alamos, Sandia et Lawrence Livermore. C'est un système de fichiers 'open source'. Lustre introduit entre autres un concept révolutionnaire: une répartition des fichiers parallèle par objets et non pas par paquets, ainsi qu'un mécanisme de « file locking » direct (verrouillage de fichiers) entre le client et le stockage, ce qui accélère les lectures/écritures sur les disques et offre une évolutivité sans précédent* », explique L. Thiers. Nous sommes ici au coeur des développements Open source. Des « locomotives » ont contribué financièrement, comme les laboratoires américains, notamment Lawrence Livermore National Lab. Le code est gratuit mais une partie du système ouvert 'open source' est repris par des 'distributions' qui maintiennent et font évoluer les développements. Ces solutions viennent directement **concurrencer GPFS d'IBM** ou encore, dans une certaine mesure, CXFS de SGI voire GFS de Redhat. L'objectif pour les entreprises reste le même: monter en capacité, sans perdre en performances. « *Nous pouvons, avec nos appliances supporter jusqu'à 1.120 disques* », pour un coût situé entre 25 et 500 K-euros et à une fraction de l'investissement initial du S2A », argumente Laurent Thiers. Cette tendance s'est affirmée sur le Salon Linux: « *On voit la percée des systèmes avec interfaces 'Fibre channel' vers SATA', offrant de larges*

capacités à partir de 10 tera-octets, des systèmes pilotés par un seul contrôleur (S2A 8500) et intégrant Linux pour la partie serveur». En clair, les architectures sont simplifiées: moins de contrôleurs, moins de serveurs. Alors, pas de doute, les grands clients, gros consommateurs de puissance de calculs ou de simulations à grande échelle, en comprennent vite l'intérêt. Il suffit de voir les applications de « pompes » vidéo (la vidéo sur demande), ou les baies de stockage à très gros volumes (comme au CEA, avec Bull-Tera 10, où l'on atteint 5 peta-octets pour la partie archive sur S2A 9500 « double parité » et disques SATA !). Ces applications s'appuient sur ces solutions où le prix des disques commence à rivaliser avec celui des bandes – donc en concurrence directe avec les solutions disques d'EMC ou bandes de StorageTek. Bref, rien ne va plus... refaites vos jeux, les tera-jetons se redistribuent!