

Le stockage de données de TERA

Le Centre de Calcul TERA produit quotidiennement un volume de données équivalent à celui de la Bibliothèque Nationale de France à Paris, soit plusieurs Teraoctets¹. Cette production soutenue a engendré, au fil des mois et des années, un volume colossal de données de 6 Petaoctets qui doit rester en ligne, accessible à tout instant par les utilisateurs. Aussi, une architecture de gestion de données à plusieurs niveaux, incluant des composants matériels et logiciels dédiés, a été associée au calculateur Tera-10.



Baie de disques (sous-ensemble)

Le premier composant est une batterie de 10 000 disques destinés à héberger les fichiers les plus récemment créés ou utilisés. Ces disques constituent le premier niveau de la hiérarchie de stockage dans laquelle les données évoluent.

Les fichiers ne peuvent résider à demeure sur des disques - il en faudrait plusieurs millions -, c'est pourquoi ils sont enregistrés sur des bandes magnétiques très capacitives, de l'ordre de 800 Go par bande, qui sont conservées dans un silo de stockage. Ce second composant constitue le deuxième niveau de stockage.

Chaque silo dispose de plusieurs milliers d'alvéoles contenant les médias. Huit mains robotisées, véritables "bibliothécaires automatisés", viennent chercher les bandes voulues pour l'insérer dans les lecteurs qui leur ont été assigné par le logiciel de gestion des données.



Silos Sun/StorageTek



Intérieur d'un silo Sun/StorageTek : alvéoles de stockage des cartouches et main (au centre)

Le troisième composant est logiciel, il s'agit de HPSS - High Performance Storage System -, un gestionnaire de stockage hiérarchique. Il pilote l'ensemble des disques, bandes et robots de stockage utilisés. En fonction de la fréquence d'utilisation des fichiers, de leur ancienneté, de leurs tailles, les fichiers sont "sédimentés" sur des médias de plus en plus capacitifs mais de moins en moins rapides d'accès. Cet ensemble constitue une hiérarchie de stockage de données de très grande capacité.

Ainsi, un fichier va vivre sa vie dans ce système, naissant sur disque pour ensuite vieillir sur les bandes, libérant ainsi l'espace disque pour de nouveaux fichiers. Ce logiciel HPSS s'exécute sur un petit « cluster » de serveurs qui lui est dédié.

Le réseau de stockage interconnecte les calculateurs et les systèmes de stockage ; il est l'ultime composant de cet ensemble. A base de technologie InfiniBand, il offre un débit très important ainsi qu'une latence faible, et constitue l'artère à grande vitesse indispensable à la circulation des flux de données engendrés.

¹ 1 Gigaoctet= 10⁹ octets ; 1 Teraoctet= 10¹² octets ; 1 Petaoctet= 10¹⁵ octets