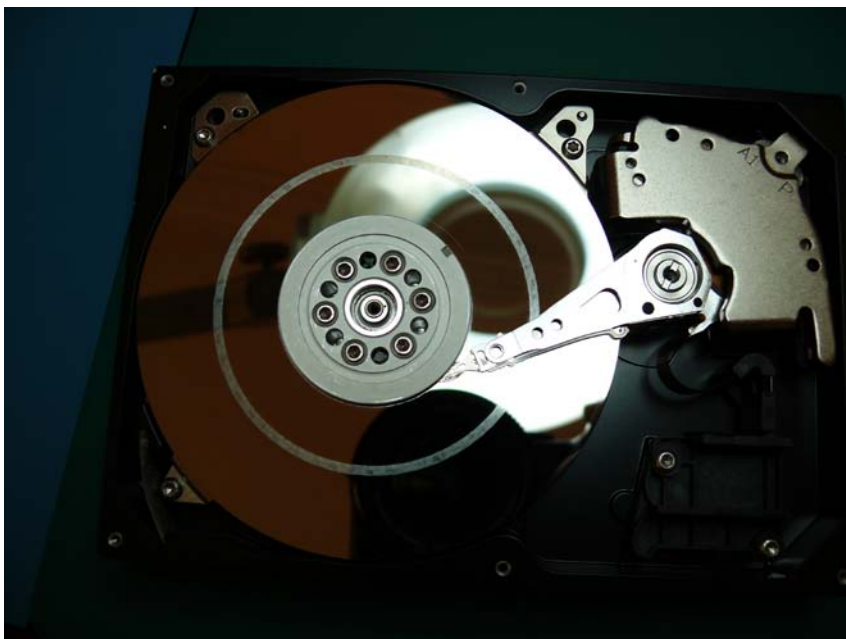




1. CE DISQUE DUR A ÉTÉ ENDOMMAGÉ par la tête de lecture qui a touché le matériau ferromagnétique à la surface du disque et tracé le sillon circulaire. Les données enregistrées sur le disque et situées dans le sillon sont définitivement perdues.

Qui possède encore un lecteur de disquette souple ou un lecteur de disque Zip ou Jaz ? De nombreux supports sont devenus obsolètes parce qu'ils ont été remplacés par des modèles plus performants ou parce qu'ils ont été des échecs commerciaux.

Ces deux aspects de la pérennité de l'information peuvent être réglés avec un peu d'organisation et de bon sens. Transposer au numérique le dépôt de métadonnées qui était déjà essentiel pour l'analogique n'est pas hors de portée ; de plus, on peut parfaitement choisir des systèmes d'exploitation, des logiciels et des formats de fichiers ouverts – dont le code est accessible et que chacun peut utiliser – ou très répandus, qui offrent des perspectives raisonnables de pérennité (voir l'article de Roberto Di Cosmo dans ce numéro).

Ces deux premiers problèmes sont importants, mais ils ne doivent pas masquer le troisième, qui est bien plus grave. On peut prendre toutes les précautions que l'on voudra en créant des métadonnées parfaites et en utilisant des formats robustes : si les données elles-mêmes disparaissent, cela ne sert évidemment à rien ! Ce problème a longtemps été sous-estimé et négligé. Hélas, nous allons voir que cet aspect de la pérennité des données pose des difficultés techniques et économiques qui sont loin d'être résolues.

La variété des dispositifs de stockage disponibles dans le commerce et leur évolution rapide peuvent donner l'impression qu'il existe de nombreuses méthodes différentes pour stocker et sauvegarder de l'information numérique. Certains supports sont-ils adéquats pour de l'archivage de longue durée ? En réalité, ces dispositifs fonctionnent sur un petit nombre de principes physiques et se rangent dans quelques catégories simples qui subissent toutes des détériorations en quelques années.

Tous les supports s'altèrent

La première catégorie fait appel à l'enregistrement magnétique : le système d'enregistrement est conçu pour orienter de façon contrôlée des petits domaines dans une couche magnétique déposée sur une surface plane. La lecture s'effectue ensuite par une tête électromagnétique sensible à cette orientation. C'est ainsi que fonctionnent les disques durs, les bandes magnétiques (ou LTO, pour *Linear Tape Open*) et même les disquettes de formats divers, dont la plupart ont maintenant disparu à cause du succès des mémoires flash. Les disques durs présentent de grandes capacités de stockage pour un coût faible, et avec un accès rapide aux données. Ils sont cependant fragiles et

soumis à des pannes soudaines et imprévisibles (ou *crash*) pouvant entraîner une perte totale des données (par exemple, si la tête de lecture touche la couche magnétique, voir la figure 1). Pour ne pas perdre ses données, une possibilité est la sauvegarde sur plusieurs disques en parallèle (tels les systèmes redondants RAID, pour *Redundant Array of Independent Disks*). La longévité des disques durs est difficile à évaluer, car elle dépend de nombreux facteurs indépendants. On s'accorde à penser néanmoins que, hors *crash*, une durée de vie de cinq ans est un bon ordre de grandeur. Pour optimiser cette longévité, il faut laisser tourner les disques durs en permanence, ce qui minimise les incidents. Les grands centres de données appliquent cette méthode coûteuse en électricité. On estime que les centres de données contribuent ainsi à plus de deux pour cent de la production de dioxyde de carbone d'origine humaine.

Il y a quelques années, les bandes magnétiques étaient considérées comme obsolètes, car moins performantes que les disques durs. Mais, précisément du fait de la nécessité de l'archivage des données, les bandes magnétiques ont effectué un retour spectaculaire sur la scène du numérique : elles conservent les données de façon passive, sans consommation électrique. Cependant, leur durée de vie est aussi limitée, de l'ordre de cinq ans, temps au bout duquel il devient indispensable d'effectuer une « migration » – c'est-à-dire de recopier les données sur une bande neuve, si on ne veut pas les perdre. Il existe d'ailleurs des robots dont la fonction est de gérer un grand nombre de bandes, d'y suivre le nombre d'erreurs et d'effectuer une migration vers des bandes neuves dès que c'est nécessaire.

La seconde catégorie de supports numériques fait appel à un enregistrement électrique et non magnétique : dans les mémoires flash, de très nombreux petits transistors stockent une charge électrique qui définit la valeur du bit enregistré. Les mémoires flash sont devenues d'usage universel, dans les appareils photo, les caméscopes, les ordinateurs, les téléphones, etc. Elles autorisent des accès aux données encore plus rapides que les disques durs, mais le prix par mégaoctet de données est plus élevé. Quant à la longévité des données stockées sur des mémoires flash, on dispose de peu d'informations à ce sujet. On sait cependant que les rayons cosmiques