

Le stockage en quête de technologies de rupture

La Terre va-t-elle se couvrir de datacenters pour stocker l'ensemble de nos données ?

IDC estime à 175 zettaoctets (un zettaoctet = un milliard de téraoctets) le volume de données qui sera stocké en 2025, cinq fois plus qu'aujourd'hui.

Pour un exaoctet de données (1024 petaoctets), il faut construire un datacenter d'un million de mètres carrés, pour un coût de plusieurs milliards d'euros et une consommation électrique annuelle de l'ordre de 10 TWh. « En 2018, il fallait 33 000 centres de ce type pour héberger l'ensemble des données stockées sur la Terre.

En 2040, il en faudra entre cinq à trente millions de fois plus ! » explique François Képès, chercheur et membre du groupe de travail « ADN : lire, écrire, stocker l'information » de l'Académie des technologies ([Document PDF à télécharger](#)). Le chercheur souligne qu'à moins de changer de paradigme, en 2040, 1/1000 e des terres émergées de la planète seront couvertes par ces datacenters...

Des technologies alternatives pour archiver les données froides

Marc Antonini, chercheur au laboratoire d'informatique, signaux et systèmes du CNRS de Sophia Antipolis souligne qu'une partie de ces données n'est que rarement sollicitée, part qui s'accroît de 60 % par an. « La densité de stockage des technologies actuelles n'augmente que de 20 % par an, ce qui oblige à accroître la taille et la quantité des datacenters pour stocker ces données froides.

La durée de vie des dispositifs de stockage actuels est limitée. Il faut les remplacer régulièrement pour assurer la pérennité des données : tous les cinq ans pour les disques durs et 20 ans au maximum pour les bandes magnétiques, »

Il devient urgent de trouver des solutions de rupture pour remplacer les bandes magnétiques et disques optiques.

Marc Antonini évoque notamment les travaux de Microsoft pour graver des données sur des disques de quartz, un média quasi indestructible qui résout le problème de la durée de vie du support, mais pas celui de la densité : « Dans le cadre de [Project Silica](#), Microsoft a dévoilé en 2019 un carré de 7,5 cm sur 7,5 cm et 2 mm d'épaisseur pour porter 75,6 Go de données, de quoi stocker le film Superman de 1978. Cette solution est capable de conserver des données pendant plusieurs centaines d'années. »

La densité d'écriture reste plutôt faible si on la compare aux [cartes microSD](#) actuellement commercialisées. L'équipe de Microsoft Research Cambridge et celles de Microsoft Azure travaillent sur une autre approche.

Plutôt que de graver des voxels sur une surface de quartz, les chercheurs utilisent des

hologrammes pour stocker des données dans un volume 3D. La lecture des données enchâssées dans le bloc de cristal est assurée par des lasers et des capteurs photo comparables à ceux de nos smartphones actuels.

ADN : un stockage éprouvé depuis... des millions d'années !

Pour le chercheur, une autre solution sérieuse pourrait bien se trouver dans notre propre ADN : «on estime qu'il est possible de stocker 215 pétaoctets sur un gramme d'ADN ! L'ADN est ultracompact, et peut être conservé des centaines, voire des milliers d'années s'il est conservé dans de bonnes conditions. »

Théoriquement, la totalité des données numériques de l'humanité de 2018 pourrait tenir dans 66 grammes d'ADN seulement. En 2040, 10 kg d'ADN suffiront. « La densité informationnelle est un point fort du stockage moléculaire , en pratique, on ne pourrait pas faire tenir l'ensemble de la sphère globale des données en 66 g » , précise toutefois François Képès. « En pratique, la sphère globale des données de 2018 ne tiendra pas dans 66 g, mais dans un espace e la taille d'une fourgonnette et, en 2040, dans un camion. »

Une densité qui reste extraordinaire si on la compare aux centaines de milliers de datacenters qu'il nous faudrait construire pour héberger des disques durs !

Accélération d'un facteur deux tous les six mois

Actuellement, on estime que le coût de lecture devra être divisé par 1 000 et le coût de l'écriture divisé par 100 millions pour que [le stockage d'ADN](#) devienne viable d'un point de vue économique. Un objectif qui devrait être rapidement atteint, car les chercheurs ont amélioré les performances de lecture-écriture d'un facteur un million ces dix dernières années, et l'accélération est actuellement d'un facteur deux tous les six mois, bien plus de la loi de Moore des composants électroniques.

D'ici à 2024, une machine pourra ainsi lire et écrire 1 To de données par jour. Les chercheurs estiment même que d'ici cinq à dix ans, la technologie pourra être utilisée sur des marchés de niche, puis de manière plus globale d'ici dix à vingt ans.

par **Alain Clapaud**