

plus grandes pour une série de raisons. Explicitons-en quelques-unes.

D'abord, il faut préciser ce qu'on veut mesurer; pour nous, il s'agira de toutes les informations codées sous un format numérique (disques durs, disques SSD, clés USB, CD, DVD, puces de mémoire, etc.), à la condition qu'elles restent lisibles par des dispositifs appropriés encore disponibles. Pour une évaluation précise directe, il faudrait donc connaître les chiffres de tous les supports numériques produits par toutes les firmes concernées, qui ne publient pas nécessairement les données dont on aurait besoin.

DES DIFFICULTÉS POUR ESTIMER LA MÉMOIRE NUMÉRIQUE

Plus difficile encore, il faudrait pouvoir évaluer parmi tous les dispositifs numériques de stockage qui ont été produits quelle proportion est restée accessible et quelle proportion est maintenant perdue: ordinateurs mis au rebut, disques tombés en panne, etc. Heureusement certains types de dispositifs sont négligeables devant d'autres. Par exemple, les CD sont aujourd'hui passés de mode et comme ils ont des capacités assez modestes, comparés aux disques durs ou aux disques SSD, on peut se dispenser d'évaluer leur contribution avec précision.

En raisonnant par des méthodes indirectes, en prenant des conclusions de rapports déjà publiés il y a quelques années et en

extrapolant leurs résultats, on obtient une série d'évaluations dont on tire une conclusion plausible. Nous avons mené une dizaine de raisonnements de ce type et en avons tiré la conclusion que 10^{22} octets est un ordre de grandeur raisonnable à retenir pour toute la mémoire numérique sur Terre en 2021 (voir l'encadré 3).

Sachant qu'une heure de vidéo de qualité moyenne exige environ 1 gigaoctet, soit 10^9 octets, combien de vies entières de 80 ans filmées pourrait-on conserver avec les 10^{22} octets de la mémoire numérique terrestre? Le calcul, facile, aboutit à un résultat surprenant: $10^{22}/(10^9 \times 24 \times 365 \times 80)$, soit environ 14 millions de vies entières.

Si l'on partageait équitablement la capacité d'enregistrement numérique entre les 7,7 milliards d'humains, chacun disposerait de $(10^{22})/(10^9 \times 7,7 \times 10^9) \approx 1300$ heures de vidéo pour raconter sa vie et en conserver le film des moments les plus importants. Encore un petit effort et nous serons en mesure d'enregistrer le film de la vie de tous les humains!

Si la mesure de la mémoire numérique totale du monde n'est pas facile, celle de la capacité de calcul de tous les dispositifs électroniques est encore plus délicate. Il y a d'abord le problème de l'unité de puissance: opération logique par seconde, instruction par seconde pour un processeur moyen, opération entre nombres entiers codés sur 8 octets par seconde, opération en virgule flottante par seconde, hertz, etc.?

2

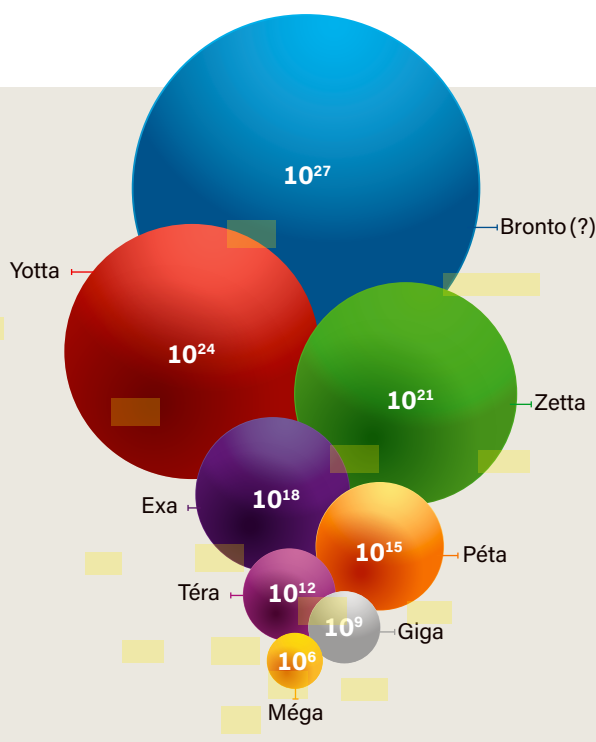
LES OCTETS ET LEURS MULTIPLES

Un octet de mémoire (byte en anglais) correspond à 8 bits d'information, c'est-à-dire à 8 choix entre 0 ou 1. Un octet peut coder un nombre entier compris entre 0 et 255, ou un caractère pris dans un alphabet de 256 caractères.

Le kilooctet vaut 2^{10} octets, c'est-à-dire 1 024 octets, mais on l'assimile souvent à 1 000 octets. En multipliant par 1 024 à chaque fois, on trouve ensuite :

- le mégaoctet : environ 1 million d'octets, c'est-à-dire le texte d'un livre de 400 pages environ, ou une image de qualité moyenne ;
- le gigaoctet : environ 1 milliard d'octets, qui est l'espace mémoire nécessaire pour une heure de vidéo de qualité moyenne ;
- le téraoctet : environ 10^{12} octets, de quoi stocker un million de livres ou 1 000 heures de vidéo ;
- le pétaoctet : environ 10^{15} octets, soit un milliard de livres ou un million d'heures de vidéo ;
- l'exaoctet : environ 10^{18} octets ; c'est la capacité de stockage du Cern, près de Genève, et le tiers de la capacité de Facebook ;
- le zettaoctet : environ 10^{21} octets, soit en gros la capacité de mémorisation de tous les disques durs produits en 2020.

Il y a ensuite le yottaoctet, puis, avec des préfixes qui ne font pas encore l'unanimité, le brontoctet, le geopctet, le saganoctet, l'alphaoctet...



Procédant comme pour la mémoire, nous avons mené différents raisonnements et utilisé différentes sources d'informations pour produire des estimations variées. Les résultats ne concordent qu'à un ou deux ordres de grandeur près, mais conduisent à affirmer qu'en 2021, on a dépassé 10^{23} opérations par seconde en quantité globale de calcul effectivement mise en œuvre et 10^{24} en capacité globale (non nécessairement active à chaque instant). Un ordre de grandeur au moins sépare donc maintenant mémoire et calcul.

Pourquoi la capacité de calcul s'accroît-elle plus vite que la capacité de stockage ? Plusieurs explications sont possibles.

Du côté des besoins, nous sommes de plus en plus des utilisateurs de calculs dont il n'est pas nécessaire de garder des traces. Ainsi, quand nous parcourons des pages internet, la plupart des informations que l'ordinateur reçoit, manipule (par exemple en préparant les images) et affiche sont rapidement effacées et oubliées, car nous sautons de page en page rapidement jusqu'à trouver ce que nous cherchons. Dans le cas du *streaming* audio ou vidéo, la situation est portée à son comble, car d'un instant au suivant tout est effacé. Déplacer les données, les décompresser et les afficher, tout cela demande du calcul, mais très rapidement la mémoire nécessaire à ces opérations est récupérée et réutilisée. Il en résulte que la pression pour avoir plus de calcul dépasse la pression pour avoir plus de mémoire.

Dans le domaine des algorithmes, de nombreux programmes opèrent massivement des calculs pour ne retenir que le résultat final qui prend peu de place comparé à la taille des calculs menés. Dans le cas de la circulation de l'information sur les réseaux, chaque serveur-relais oublie aussitôt ou assez rapidement la majeure partie de ce qui passe par lui. C'est clair, le développement de l'informatique des réseaux exige bien plus une accélération du calcul qu'une augmentation des capacités de stockage.

LE BITCOIN, UN CAS EXTRÊME

Le cas de la cryptomonnaie Bitcoin est caricatural : le « minage » qui permet d'attribuer les nouvelles unités créées à certains comptes consiste à essayer de produire, quasiment au hasard, une certaine configuration gagnante, comme quand on cherche à obtenir un triple « 6 » avec trois dés en les lançant jusqu'à réussir. L'objectif des « mineurs » de bitcoins est beaucoup plus difficile qu'un triple « 6 » et, chaque seconde, le réseau Bitcoin effectue le lancer des dés numériques environ 10^{20} fois en calculant 10^{20} fois la fonction notée SHA256, dont il efface instantanément le résultat si la configuration attendue n'est pas obtenue, comme quand on ramasse les dés parce que le

3

COMBIEN DE MÉMOIRE AU TOTAL ?

Voici ce qui semble le meilleur raisonnement pour évaluer la mémoire totale des dispositifs électroniques en 2021. On commence par repérer les sources les plus importantes de mémoire de masse créées et vendues en 2021.

- Les diverses catégories sont, par ordre d'importance décroissante :
- (a) Les « disques durs » (DD), qui sont des dispositifs de stockage magnétique sur disques rotatifs **A**.
 - (b) Les « disques » SSD (*solide state drive*), c'est-à-dire la mémoire de masse sur puces électroniques **B**. Ces « disques » remplacent progressivement les DD, et ils équipent la plupart des nouveaux ordinateurs portables.
 - (c) Les mémoires NVM (*non volatile memory*) utilisées pour des accès plus rapides au cœur des ordinateurs.
 - (d) Les mémoires optiques (CD, DVD, etc.) **C**, de moins en moins utilisées.
 - (e) Les bandes magnétiques, réservées à certains usages professionnels **D**.

D'après le site Datanami (<https://www.datanami.com/2020/02/19/storage-in-the-exabyte-era/>), la partie (a) de la mémoire numérique est la plus importante et représente 60 % de toute la mémoire nouvelle créée en 2021.

Connaître (a) donne donc accès à une évaluation de toute la mémoire créée en 2021. Le site [statista.com](https://www.statista.com/statistics/398951/global-shipment-figures-for-hard-disk-drives/) indique que plus d'un zettaoctet de DD a été fabriqué et vendu en 2020 (<https://www.statista.com/statistics/398951/global-shipment-figures-for-hard-disk-drives/>).

En arrondissant, on arrive à 2,5 zettaoctets pour le total de la mémoire fabriquée et vendue en 2020, et, à quelques pourcents près, à la même valeur pour 2021.

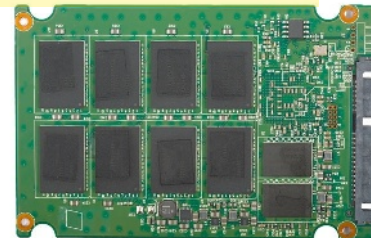
Reste à évaluer la quantité de mémoire produite les années précédentes et encore utilisée en 2021. C'est la partie la plus délicate. Le coefficient à appliquer est certainement situé entre 4 et 10, ce qui conduit donc à une valeur comprise entre 10^{22} et $2,5 \times 10^{22}$ octets pour la mémoire numérique totale sur la planète en 2021.

Le site [import.io](https://www.import.io/post/idc-predicts-the-datasphere-to-reach-160-zettabytes-by-2025/) propose la valeur 7×10^{22} octets et le site [statista.com](https://www.statista.com/statistics/871513/worldwide-data-created/) indique 6×10^{22} pour 2020, mais sans donner de détails sur les raisonnements mis en œuvre pour parvenir à ces chiffres (voir <https://www.import.io/post/idc-predicts-the-datasphere-to-reach-160-zettabytes-by-2025/> et <https://www.statista.com/statistics/871513/worldwide-data-created/>).

A



B



C



D

