

Nous allons voir que pour la mémoire, 10^{23} est sans doute un peu trop optimiste, mais que pour le calcul, ce nombre est dépassé. L'état du monde numérique aujourd'hui a donc franchi le repère 10^{21} , d'où le nom proposé d'«ère du zettaoctet». En effet, après les préfixes kilo (10^3), méga (10^6), giga (10^9), téra (10^{12}), péta (10^{15}) et exa (10^{18}), vient zetta (10^{21}). Si l'on continue à gagner plus d'un ordre de grandeur tous les dix ans, comme cela s'est produit en gros pendant cinquante ans, nous serons dans l'ère suivante, l'ère numérique du yotta (10^{24}), avant 2040.

Cependant, la loi de Moore semble se fatiguer et donc l'ère du yotta se fera peut-être attendre un peu. La divergence entre les chiffres de la mémoire et ceux de la puissance de calcul, dont la croissance est plus rapide, risque d'ailleurs de se confirmer et de produire de nouveaux effets. Après le yotta, on aura, dans un futur encore lointain, le bronto (10^{27}), le geop (10^{30}), puis le sagan (10^{33}) (ces préfixes ne sont pas encore officiellement définis ni admis par tous).

Les calculs et évaluations pour le futur proposés ici ne prennent en compte ni les gains qu'on pourrait tirer de l'utilisation de la mécanique quantique, ni ceux qui proviendraient de la maîtrise du stockage massif d'informations grâce à l'ADN. Ces deux technologies progressent et il n'est pas exclu qu'elles finissent par changer le cours de l'histoire. Aujourd'hui toutefois, aucune de ces deux sources potentielles de progrès n'est sur le point de rattraper l'extraordinaire avance de l'électronique qui, sur sa lancée et grâce à plus d'un demi-siècle de progrès faramineux, nous précipite vers l'ère du yotta.

UTILITÉ DES ORDRES DE GRANDEUR

Connaître les ordres de grandeur actuels de la mémoire et du calcul permet par exemple de savoir ce qui est envisageable comme calcul massif pour mener des attaques de codes cryptographiques par la méthode de la force brute. Grâce à l'évaluation de la mémoire de masse globale, on évalue ce qu'un organisme étatique ou un *data broker* (firme qui réunit des données pour les revendre) peut stocker comme informations sur chacun d'entre nous. On peut aussi fixer, grâce à la connaissance des ordres de grandeur, ce qu'il est raisonnable de tenter comme tables mathématiques, géographiques, météorologiques, etc.

Si quelqu'un vous dit par exemple qu'il veut faire la table de tous les nombres premiers s'écrivant avec 50 chiffres ou moins, vous pouvez lui rire au nez. En effet, une telle table exigerait environ $50 \times 10^{50} / \log 10^{50} = 10^{50} / \log 10$ octets, soit plus de 10^{49} octets, sachant qu'il faut 50 octets en moyenne pour mémoriser un nombre premier de 50 chiffres et que la densité des nombres premiers autour de n est $1/\log(n)$.



UNE AMPOULE ALLUMÉE PENDANT UNE HEURE?

On lit ou entend parfois qu'une requête sur le moteur de recherche de Google équivaut, en dépense d'électricité, à la consommation d'une ampoule pendant une heure. C'est faux, même si le calcul est assez délicat à mener.

Il est possible de majorer le coût d'une requête par le raisonnement suivant. Google consomme 12 TWh/an pour faire fonctionner les serveurs de tous ses centres de données et répondre à 7 milliards de requêtes chaque jour. Si l'on suppose que toute cette électricité est consacrée aux réponses aux requêtes (ce qui bien sûr donnera une valeur largement supérieure à la réalité), cela donne une dépense moyenne de $12 \times 10^{12} / (7 \times 10^9 \times 365) = 4,7 \text{ Wh} = 0,0047 \text{ kWh}$ par requête, c'est-à-dire de quoi allumer une ampoule de 100 W pendant moins de 3 minutes, électricité qui vous sera facturée moins de un millième d'euro.

Cette majoration est assez loin du chiffre de une heure. En 2009, date de l'apparition de la légende urbaine, Google a répondu en proposant sa propre évaluation. Google trouve

que l'ampoule ne restera allumée que durant 10 secondes. Le calcul date d'il y a plus de dix ans : c'est probablement 5 ou 10 fois moins aujourd'hui.

Le numérique est coûteux en énergie, mais n'oublions pas qu'il progresse encore, d'où une baisse de la dépense d'électricité par opération, et surtout qu'il fait économiser de l'énergie en limitant l'impression et la circulation de documents matériels, ainsi que les déplacements physiques pour se rendre dans les centres de documentation et les réunions.

Le coût du numérique dépend par ailleurs des quantités de données que l'on fait circuler sur les réseaux, quantités aujourd'hui majoritairement liées à l'usage de la vidéo et non aux courriels ou aux requêtes sur des moteurs de recherche.

La pollution engendrée par le numérique est aussi due à la fabrication des machines (ordinateurs, téléphones, etc.) et leur destruction. Pour limiter cette pollution, le plus efficace est de ne pas changer trop souvent ses téléphones, tablettes et ordinateurs.

Si, grâce à un peu de codage, on réussit à n'utiliser qu'un bit par nombre premier de la table, alors la table envisagée nécessitera $10^{50} / (8 \times \log 10^{50})$ octets, soit plus de 10^{46} octets. Dans aucun des deux cas, une table aussi volumineuse n'est envisageable dans un avenir prévisible.

Si la NSA (la *National Security Agency* des États-Unis) gardait un gigaoctet d'informations sur chaque individu de la planète, cela ferait environ $10^9 \times 10^{10} = 10^{19}$ octets. Ce n'est pas totalement impossible, mais difficile : le Cern, qui collecte une quantité colossale d'informations pour l'analyse des résultats de ses expériences, n'a atteint 10^{18} octets que récemment. En 2021, Facebook en est à 300 pétaoctets, soit 3×10^{18} octets (<https://kinsta.com/blog/facebook-statistics/>). Toujours pour 2021, le site *statista.com* a évalué à 2×10^{21} octets le total de la mémoire de tous les centres de données.

La population humaine totale est évaluée à 7,7 milliards, mais il est impossible d'avoir plus de précision ; dans le cas de la mémoire numérique totale sur Terre, les difficultés sont bien

2

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.?

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...

