

R

ENDEZ-VOUS

P.80 Logique & calcul

P.86 Art & science

P.88 Idées de physique

P.92 Chroniques de l'évolution

P.96 Science & gastronomie

P.98 À picorer

LE MONDE NUMÉRIQUE PASSE AU ZETTA

Les capacités de calcul et de stockage de l'information continuent de croître. Globalement, où en est-on aujourd'hui? Les estimations font apparaître l'ordre de grandeur de 10^{21} octets, c'est-à-dire le zettaoctet.

L'AUTEUR



JEAN-PAUL DELAHAYE
professeur émérite
à l'université de Lille
et chercheur au
laboratoire Cristal
(Centre de recherche
en informatique, signal
et automatique de Lille)



Jean-Paul Delahaye a notamment publié : **Les Mathématiciens se plient au jeu**, une sélection de ses chroniques parues dans *Pour la Science* (Belin, 2017).

Il y a une vingtaine d'années, plusieurs articles de cette rubrique traitaient de la quantité d'informations numériques sur Terre et de la capacité cumulée de calcul des dispositifs électroniques. Le nombre 10^{20} (100 milliards de milliards) était présenté comme un repère approximatif commode. La quantité d'informations numériques sur Terre – disques durs, mémoires des ordinateurs et téléphones, CD, DVD, etc. – était vers l'an 2000 d'environ 10^{20} octets (1 octet est égal à 8 bits d'information) et la capacité de calcul des dispositifs électroniques atteignait en ordre de grandeur 10^{20} instructions par seconde (IPS).

La loi de Moore, sous sa forme la plus générale, est l'affirmation que «les capacités des dispositifs informatiques pour un prix donné doublent tous les deux ans». Cette forme générale ne correspond pas vraiment à la loi énoncée par Gordon Moore, laquelle ne portait que sur le nombre de transistors par centimètre carré des puces électroniques dont ce cofondateur d'Intel pensait qu'il doublait environ tous les ans (énoncé de 1965), puis tous les deux ans (énoncé de 1975). Nous prendrons cependant comme repère la loi générale concernant tous

les dispositifs informatiques, en la critiquant quand c'est nécessaire. Notre question principale sera : qu'est devenu le 10^{20} proposé il y a deux décennies?

L'ÈRE DU ZETTA, VOIRE DU YOTTA

Un doublement tous les deux ans pendant vingt ans donne une multiplication par 1000 environ, car $2^{10} = 1024 \approx 10^3$, soit un ordre de grandeur gagné tous les sept ans environ. En sommes-nous alors à 10^{23} pour la mémoire numérique cumulée et pour la puissance de calcul mesurée en nombre d'instructions par seconde?

Il est difficile de répondre avec certitude. Les évaluations disponibles, nos calculs et les divers raisonnements formulés pour produire des réponses, nécessairement approximatives, ne permettent pas des affirmations certaines. Cependant, les conclusions que nous allons proposer sont justes à un ordre de grandeur près, parfois deux pour les plus délicates. Même s'il est impossible d'avoir des évaluations fines, nous pensons qu'il est utile de disposer des chiffres que nous indiquerons, plutôt que croire à tort que, par exemple, 10^{10} , 10^{15} ou 10^{30} sont des évaluations globales pour l'information ou le calcul.

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