

SNT

TP17 : stockage des données

Le volume de données créées en 2018 est estimé à 33 zettaoctets¹.
33 mille milliards de milliards d'octets ($33 \cdot 10^{21}$ octets).
En 2035, on envisage 65 fois plus : 2 142 zettaoctets².
Comment stocker une telle quantité d'informations ?

Historique des supports de stockage : <https://youtu.be/4hx6FqR6oPs> ³
Présentation de quelques uns d'entre eux.

A/ Stockage local sur disque dur

Ouverture d'une tour ou d'un pc portable (*service tice*) et localisation du disque dur, identification de composants élémentaires.

Ouverture d'un disque dur, tête de lecture et plateaux.

Mise sous tension d'un disque dur ouvert.

Nécessité d'effacement voire de destruction avant élimination.

Le disque dur de mon ordinateur personnel a une capacité de 1 To (T = tera = 10^{12} octets).
Son épaisseur est de 2,3 cm.

Combien de disques durs de ce type faudrait-il pour stocker la quantité de données produites en 2018 ?

Si on empilait ces disques durs les uns sur les autres, quelle serait la hauteur de la pile ?

A titre de comparaison, le diamètre de la Terre est de 12 760 km et la Lune est à environ 380 000 km de la Terre.

Réponses : 33 milliards de disques durs. $7,6 \cdot 10^8$ m = $7,6 \cdot 10^5$ km soit 2 fois la distance Terre-Lune.

B/ Stockage distant : *cloud computing* et *data center*

Visite virtuelle d'un data center *Google* : <https://youtu.be/avP5d16wEp0>

1/ Exemple de l'ENT

Les données de l'ENT ne sont pas conservées localement, elles ne sont pas sur votre ordinateur.

Se connecter à l'ENT.

Cliquer sur l'aide en haut à droite (?).

Quel est le nom de l'ENT ?

Quel est le nom de la société qui le fournit ?

Utiliser l'adresse <http://verif.com> pour obtenir la localisation de la société et le groupe qui la détient.

Ce groupe a racheté un hébergeur de données (possédant des data center).

Utiliser internet pour obtenir le nom de cet hébergeur et sa localisation.

Pour obtenir une localisation plus précise du data center, utiliser l'adresse suivante :

<https://www.datacentermap.com/search.html>

Quel(s) problème(s) voyez-vous à la divulgation de la localisation exacte d'un data center ?

Ce data center est classé TIER III par l'organisme Uptime Institute.

Est-ce un bon classement ?

1 Source : Statista Digital Economy Compass 2019 -

<https://static2.statista.com/download/pdf/DigitalEconomyCompass2019.pdf>

2 <https://fr.statista.com/infographie/17800/big-data-evolution-donnees-numeriques-creees-dans-le-monde/>

3 « Le stockage informatique, de la carte perforée au cloud »— par AFP

Donner la durée moyenne d'interruption par an pour ce classement.
Qu'appelle-t-on « redondance » des composants ? A quoi sert-elle ?

2/ Problèmes structurels dans l'exploitation des data center

① L'alimentation électrique

Rechercher la consommation électrique annuelle de la ville d'Epernay
Comparer à celle des data center français en 2015 : 3 TWh (terawattheure).
Conclure.

② L'échauffement

Manifestation de l'effet *Joule* : caméra thermique, bouilloire électrique
Des températures élevées entraînent une détérioration des composants électroniques, une diminution des performances (vitesse de transfert, temps d'accès...) des systèmes.
Solution: le refroidissement
Free cooling (air), liquid cooling (eau, immersion dans un liquide non conducteur) :
<https://youtu.be/mm7Ojgc6G5M> ⁴

3/ Utilisation du *Big Data*

L'existence de si grands volumes de données pousse à en extraire des informations.
Décrire une application concrète du *Big Data*.

Dérives du *Big Data*⁵ : <https://youtu.be/5otaBKsz7k4>

⁴ « How to Prepare a Server for Submer's Immersion Cooling »

⁵ Episode #15 de la chaîne *Data Gueule*. « Big data : données, données, donnez-moi ! »