



1 LES SAUVEGARDES

Il existe plusieurs systèmes de sauvegardes, permettant de gagner en rapidité et en performances. En effet, sauvegarder plusieurs machines via un réseau, la nuit génère une contrainte de durée.

1.1 SAUVEGARDE TOTALE

La sauvegarde totale est simple, il suffit de prendre toutes les données et de les transférer vers le serveur de sauvegarde. Toutefois, il peut y avoir un problème de performance :

Par exemple, un réseau 1Gbps, comptant 10 machines ayant un disque dur de 1To à sauvegarder, prendra 22 heures

*Volume total : 1 To * 10 machines = 80 000 000 000 000 bits (calculs approximatifs)*

Durée totale : volume total / débit = 80 000 secondes, soit plus de 22 heures

Pour résoudre ce problème, il est possible d'effectuer les sauvegardes pendant le week-end... mais durant la semaine, il faut utiliser d'autres méthodes.

Vidéo : [robot de stockage StorageTek \(2000\)](#)

1.2 SAUVEGARDE INCRÉMENTALE (Δ /JOUR)

L'idée est de ne sauvegarder pendant la semaine, que les éléments modifiés. Dans une base de données, **on enregistrera que les nouvelles requêtes du jour**. Sur un espace de disque, on ne conservera que les octets modifiés (les fichiers pouvant être considérés comme des éléments binaires).

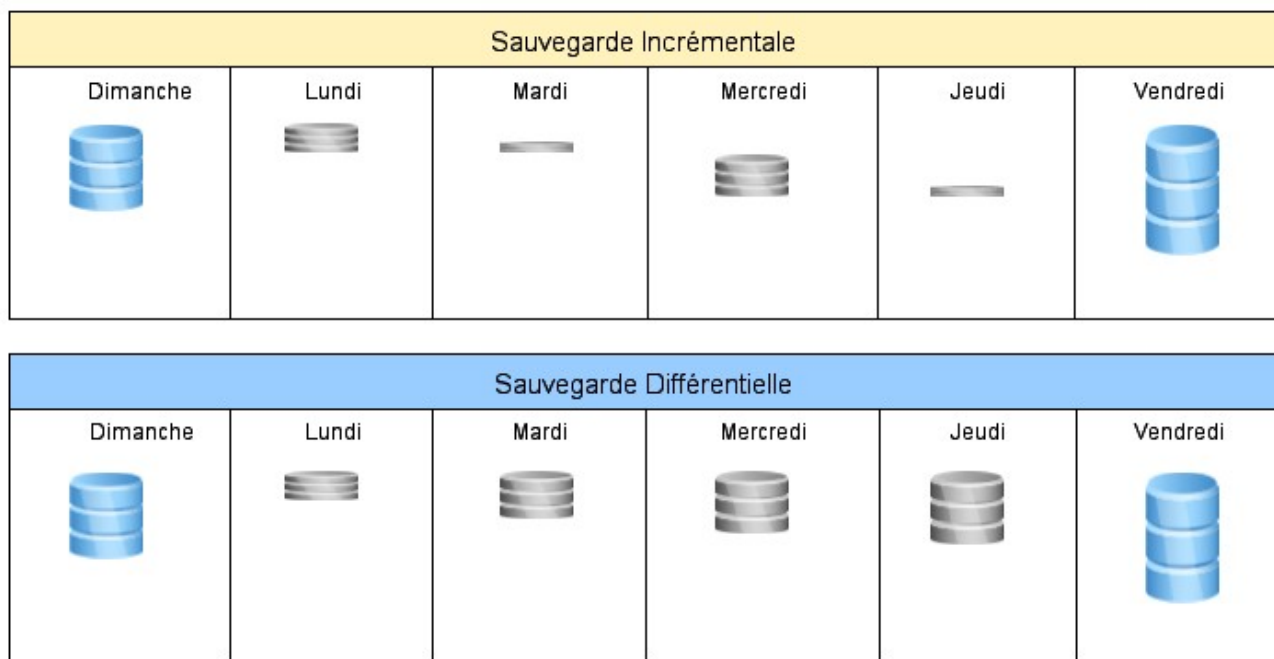
Cette sauvegarde présente cependant l'inconvénient de devoir restaurer toutes les modifications depuis la dernière sauvegarde totale, puisqu'il y a chaque jour des modifications et que celle de la veille n'est plus sauvegardé. La restauration est donc plus longue.

1.3 SAUVEGARDE DIFFÉRENTIELLE (Δ /TOTAL)

Cette méthode est un compromis entre les deux systèmes de sauvegarde : on **sauvegarde chaque jour toutes les modifications depuis la dernière sauvegarde totale**. Ainsi, le jeudi, on dispose de toutes les actions sauvegardées de la semaine.

1.4 RÉSUMÉ DES SAUVEGARDES

Le schéma ci-après, montre l'intérêt des trois types de sauvegardes.



La sauvegarde totale est coûteuse en temps, s'exécute donc plutôt les jours fermés (vendredi soir, samedi soir).

La sauvegarde différentielle ne génère pas beaucoup de flux mais implique un travail laborieux de restauration.

La sauvegarde incrémentielle occupe de plus en plus de temps chaque jour, mais permet une restauration depuis la dernière sauvegarde seulement.



1.5 LES SYNCHRONISATIONS

Certains outils permettent de synchroniser des disques ou des répertoires. Avec l'arrivée du cloud, les solutions permettent même de le faire en ligne.

1.5.1 Synchronisation automatique

Un programme fonctionne en arrière-plan sur l'ordinateur. Dès qu'un fichier est modifié (date, taille, attributs), ce programme répercute la modification sur le système de fichiers distant.

Exemple : Yadis ! Backup, FreeFileSync...

1.5.2 Synchronisation manuelle

D'autres programmes permettent de synchroniser les données sur demande : c'est notamment le cas des outils de type GIT !

L'utilisateur vérifie que les documents sur son espace local sont à jour, puis, après modifications, il doit transmettre celles-ci sur l'espace distant. On utilise le terme dépôt (en anglais, repository) et les actions sont "tirer" (pull) ou "pousser" (push) vers le cloud.

1.6 LA GESTION DE VERSIONS

C'est la force des outils GIT ou SVN : ils enregistrent les changements et permettent ainsi de connaître les différentes révisions d'un document. Chaque fois qu'un utilisateur fait un "commit", il doit remplir une description des modifications effectuées.

Ce système permet, lors de débogages importants, de retrouver quelle modification a entraîné une régression dans le code.

1.7 LA REDONDANCE

Il est utile d'avoir des systèmes fiables, et l'un des moyens sur les serveurs de sauvegarde, est l'usage des technologies RAID (Redundant Array of Independent Disks). Il s'agit principalement d'agréger des disques durs pour obtenir de meilleures performances, tout en ayant une certaine tolérance aux pannes. Voici les plus classiques (mais il existe d'autres RAID¹).

1.7.1 RAID-0 : rapidité

Cette technique consiste simplement à utiliser deux disques simultanément, pour gagner en performances. L'astuce consiste à partager l'écriture d'un flux, sur deux disques, séparément. L'inconvénient, est qu'en cas de panne d'un des disques, on ne peut récupérer qu'une partie de l'information.



Les disques sont dits "entrelacés".

1.7.2 RAID-1 : miroir

C'est la technique la plus simple à comprendre : l'information est écrite sur les deux disques en même temps. En cas de panne de l'un des disques, on peut récupérer l'information complète sur l'autre.



Les disques sont dits "en miroir".

1.7.3 RAID-5 : redondance partagée

Cette technique est la plus complexe, car elle implique l'usage de trois disques au minimum. Chaque disque contient un bloc de données ou un bloc de parité. Si un disque tombe en panne, on perd un bloc de données (calculable avec un autre disque grâce au bloc de parité) ou un bloc de parité (contrôle d'erreur, calculé à partir de deux blocs de données). Les blocs sont en rotation, ce n'est pas toujours le même disque qui contient les blocs de parité.

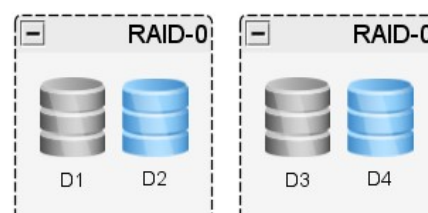


Le bloc de parité est calculé par l'opération booléenne XOR entre deux blocs de données (exemple $D1 \text{ XOR } D2 = P1$). Si le bloc D2 est perdu sur le disque en panne, $D2 = P1 \text{ XOR } D1$.

Les disques sont dits "agregés avec parité"

1.7.4 RAID-10 : rapidité et miroir

Il s'agit de la concaténation de RAID-1 et RAID-0 et non d'une nouvelle technique. Il faut donc 4 disques au minimum (les disques sont appairés en RAID-0 et chaque groupe participe à RAID-1).



1 https://fr.wikipedia.org/wiki/RAID_%28informatique%29



1.8 ARCHIVAGE ET DURÉE DE VIE

S'il paraît simple d'effectuer une sauvegarde, l'opération de restauration dépend de ce critère : en effet, restaurer un document quelques jours ou quelques mois plus tard n'est pas risqué mais que se passe-t-il 10, 20, 30 ou 50 ans plus tard ?

1.8.1 Équipements

Les technologies évoluent et la première difficulté est de trouver le matériel capable de lire le support :

- Lecteur de disquette 5" 1/4
- Lecteur de disquette 3" 1/2 double-face ? Double-densité ?
- Lecteur de CD ou DVD
- Disque-dur SCSI ? IDE ? SATA
- Clé USB 1.0 ? 2.0 ? 3.0 ? mini-USB ? USB-C ?

Il faut donc prévoir de conserver un système complet compatible avec les sauvegardes de l'époque.

1.8.2 Fiabilité des supports

Après les disquettes et les bandes magnétiques, les disques optiques semblaient être la solution... mais les technologies de gravure et les matériaux employés ont rapidement montré que la durée de vie d'un tel support ne dépasserait pas 20 ans environ.

Le document suivant indique l'espérance de vie des supports connus.



THE LIFESPAN OF STORAGE MEDIA

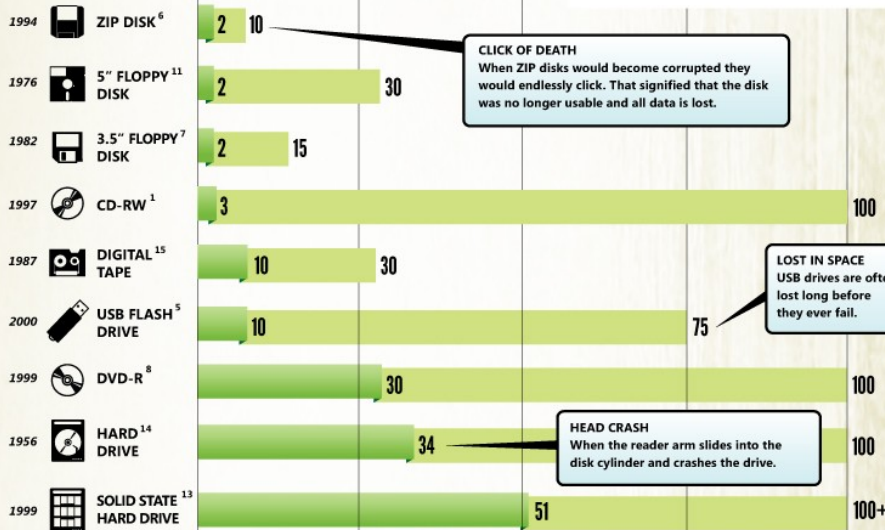
FOR AS LONG WE HAVE BEEN INVENTING WAYS TO CAPTURE AND PRESENT DATA, we have been developing media on which to store it. With each new innovation, we've seen greater convenience, increased storage capacity, and longer lifespans. But at some point, even with the utmost care, media fails... leaving us with a useless piece of plastic, tape, or metal.

The Cloud, with its extensive network of server farms, leverages redundant storage media, ensuring that in the event of hardware failure, the data lives on, stored safely in another part of the cloud. By way of comparison, we've assembled some lifespan statistics for storage media throughout the digital ages.

FAST REDUNDANT RELIABLE ACCESSIBLE ANYWHERE FOREVER



COMPUTER MEDIA

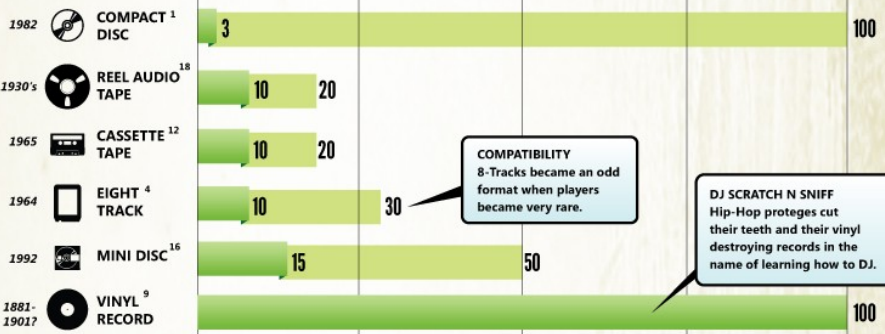


CLICK OF DEATH
When ZIP disks would become corrupted they would endlessly click. That signified that the disk was no longer usable and all data is lost.

LOST IN SPACE
USB drives are often lost long before they ever fail.

HEAD CRASH
When the reader arm slides into the disk cylinder and crashes the drive.

AUDIO MEDIA

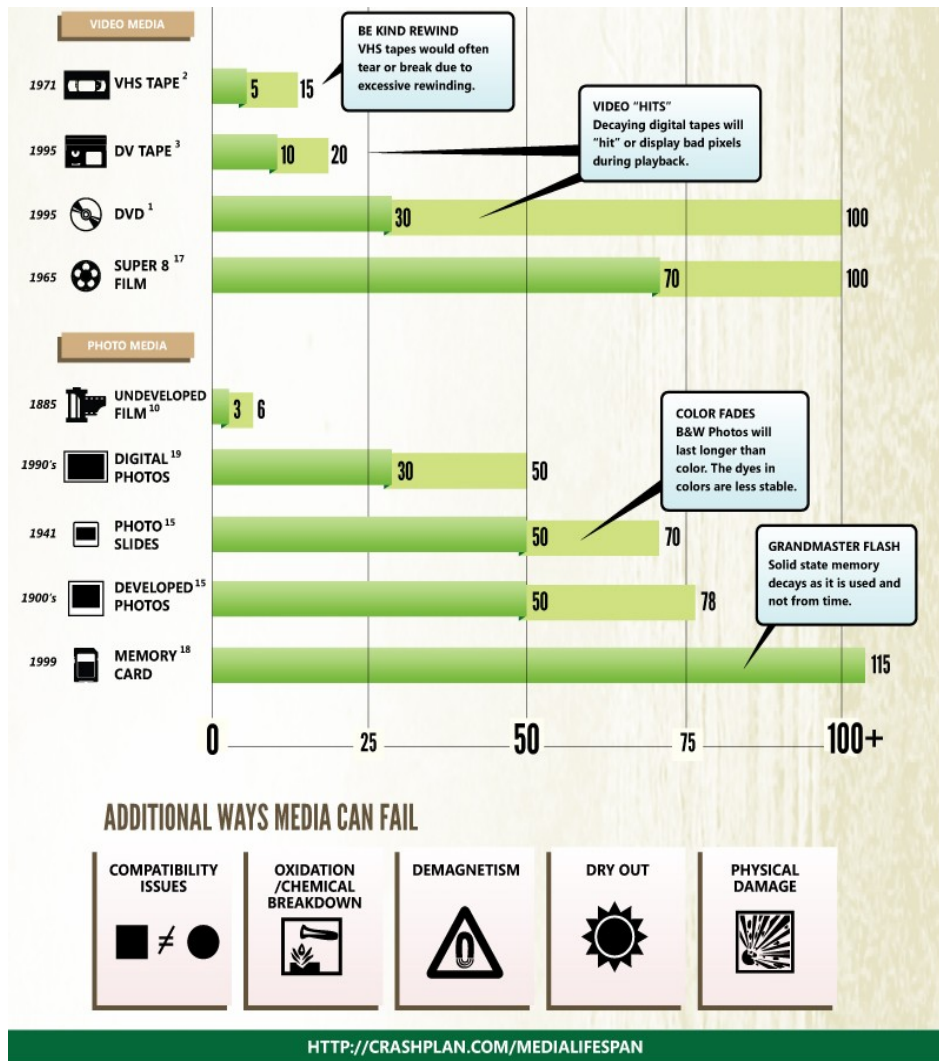


COMPATIBILITY
8-Tracks became an odd format when players became very rare.

DJ SCRATCH N SNIFF
Hip-Hop proteges cut their teeth and their vinyl destroying records in the name of learning how to DJ.



402 Les sauvegardes et archivages



SOURCES:

- 1 archive.gov, osta.org 2 videotransferboston.com/shelf-life-of-video-tape/
- www.audiokarma.org 3 http://www.cllr.org/pubs/reports/pub54/4life_expectancy.html 4 www.dresource.com 5,6,7,8 webdev.cac.edu/talkin/Storage_Quick_Reference.pdf
- <http://webdev.cac.edu/talkin/storage.htm> 9 <http://www.musicstack.com/articles/can-vinyl-last-forever-tips-on-care-vinyl-records> 10
- engineering.dartmouth.edu/microeng/processing/_/NR93000PY.pdf 11 http://www.atarimagazines.com/creative/v9n12/205_Floppy_disk_handling_and_php
- 12 www.audiokarma.org. 13 <http://www.storagesearch.com/ssdmyths-endurance.html>
- 14 The Western Digital MTBF specification is 300,000 hours. 15 <http://www.caps-project.org/cache/DigitalMediaLifeExpectancyAndCare.html>
- 16 <http://forums.sonyinsider.com/topic/23903-mini-disc-lifespan/> 17 <http://www.bonofilm.com/super8reality.asp>
- 18 <http://www.whatdigitalcamera.com/equipment/advice/517064/understanding-memory-cards.html>
- 18 <http://www.whatsbestforum.com/showthread.php?1619-Reel-to-Reel-Tape-Lifespan> 19 <http://printers.toptenreviews.com/photo-only/how-to-make-your-photos-last-longer.htm>



©2012 CODE 42 SOFTWARE, INC., ALL RIGHTS RESERVED

