

Sauvegarde et Archivage

Par Lauréline Reynaud

*D'après le cours de Yann Philippe
donné à Gobelins l'école de l'image
Et avec l'aide de quelques informaticiens spécialisés.*

Dossier publié pour l'article «Minimalisme numérique et productivité:
une méthode simple pour sauvegarder ses données ? »
du blog [Par les yeux de la coccinelle](#)
Merci de ne pas redistribuer.

Introduction

Différence entre sauvegarde et archivage

- Le stockage correspond aux données vivantes, actuellement utilisées ou nécessaires. Il est dépendant de l'infrastructure (applications, configurations, OS, etc). C'est ici que l'on travaille.
- L'archivage correspond au fait de mettre de côté et classer des informations (n'ayant plus un intérêt immédiat).
- La sauvegarde est le moyen de protéger des informations contre toute atteinte qui leur serait portée.

Dans la «vraie vie» l'archive est un carton, rangé au fond du grenier et contenant tes bulletins de notes de CP et les jolis colliers de nouilles que tu as offerts à ta mère. La sauvegarde serait plutôt le fait de garder le même carton à trois endroits à la fois: chez ta mère, chez ton père, mais aussi à l'école (celle qui t'a fourni le bulletin).

En informatique, l'archivage de tes données correspond en général au stockage sur ton disque dur principal, de travail, ou à l'impression de tes épreuves. Les sauvegardes elles, sont une copie de tous tes fichiers importants, de manière régulière. Elles sont de préférence au nombre de 3, dont deux supports différents et un délocalisé. La sauvegarde permettra de récupérer n'importe quel travail important, à n'importe quelle heure, n'importe quel jour, et de préférence depuis n'importe où.

Archivage et Stockage des fichiers

1. Les différents supports de stockage

Il existe différents supports de stockage, classés en deux catégories: mémoire mécanique ou mémoire flash (+ le cloud en bonus, mais ces clouds sont stockés sur des disques physiques).

<i>Mémoire mécanique</i>		<i>Mémoire Flash</i>	
Disques durs classiques (HDD de 2.5 ou 3.5 pouces) <u>Avantages:</u> Economique pour une grande capacité de stockage	<u>Inconvénients:</u> Commence à disparaître des configurations les plus simples au profit des SSD. Particulièrement fragiles à cause de leur tête de lecture qui vient inscrire et lire les données sur des plateaux tournant à plusieurs milliers de tours/min.	SSD: solid state drive <u>Avantages:</u> 150x plus rapide qu'un HDD, fiabilité, économie d'énergie. Plus solide car non mécanique. Plus léger et plus fin.	<u>Inconvénients:</u> Même si les tarifs baissent, la capacité de stockage reste très inférieure à celle d'un HDD. Durée de vie plus courte.
CD/DVD/Blu-ray <u>Avantages:</u> Léger, facile à ranger, rapide à graver, peu cher.	<u>Inconvénients:</u> Bientôt dépassé par la mémoire flash, les cloud et le streaming. Durée de vie loin d'être infinie (ne dépasse que très rarement les 5 ans pour les disques gravés) Impossible d'affirmer que les supports de lectures seront toujours disponibles dans les années qui viennent. (pensons, avec nostalgie, à la disquette....).	Cartes mémoires et clés USB <u>Avantages:</u> Léger, petit, peu cher.	<u>Inconvénients:</u> Durée de vie limitée, facile à perdre. Ce sont des supports qui doivent servir à l'échange de données et non au stockage.

Les mémoires mécaniques sont adaptées à la sauvegarde. Les SSD sont fragilisés par les opérations d'écritures, préférez donc y installer OS et logiciels (que la vitesse soit vraiment utile quoi).

/! Confidentialité des données et fiabilité des supports, principalement sur les clouds. Les données transférées dans le cloud ne sont pas forcément présentes sur le territoire national, et sont sujettes aux lois qui s'y appliquent. L'utilisation faite des données (je pense aux drives de Google, Microsoft ou Dropbox) et leur confidentialité dépendent donc des lois de ces pays. De plus, confier ses données à une entreprise proposant un service, c'est risquer qu'elle le ferme, sans raison, sans préavis, etc. Enfin, les données sont stockées chiffrées (normalement) dans des data centers très gourmands en énergie et en eau.

Archivage et Stockage des fichiers

2. Formatage et systèmes de fichiers

Quand on achète un support de stockage externe, il est préformaté dans une norme. C'est pour cela que l'on trouve des disques durs «pour Mac»... On se retrouve donc régulièrement à ne pas pouvoir écrire ou lire une clé venant d'un autre OS.

- **FAT32**: les fichiers sont limités à 4Go, ils sont très rapidement fragmentés. Lisible et inscriptible par les OS d'Apple et de Microsoft.
- **exFAT**: amélioration du FAT32, les limitations de tailles sont éliminées. Lisible et inscriptible par tous les OS. En revanche il pose parfois des problèmes avec certains équipements (TV, box, etc.). C'est donc le système de fichier le plus recommandé pour nous.
- **NTFS**: mis au point par Windows, fonctionne à l'image d'une base de données. Les fichiers ne peuvent pas faire plus de 16To. Les Mac ne peuvent pas y écrire, mais peuvent le lire.
- **HFS+**: format des MAC, équivalent au NTFS ou aux formats Linux. N'est pas visible par les Windows (ou reconnu comme élément non formaté).
- Sur les distributions Linux, les formats de fichiers sont légèrement différents, mais très souples. Ils gèrent sans soucis les autres standards.

Liens utiles pour plus d'informations:

- [Stockage des données informatiques \(très complet\)](#)
- [C'est quoi un disque dur ?](#)
- [Vos données et Google](#)
- [Le vol des données sur le cloud](#)
- [Cloud et données personnelles](#) (Live Youtube avec [Korben](#))
- [Articles et explications sur le RGPD](#), le nouveau règlement général de protection des données européen

Archivage et Stockage des fichiers

3. Entretenir son matériel

Ejection

Le fait d'éjecter son disque dur («retirer en toute sécurité») permet l'arrêt de toutes les écritures en cours sur le disque, ferme les fichiers ouverts, alerte les programmes, etc. Si un programme est toujours en train de travailler sur une image (généralisé Photoshop), une fenêtre s'ouvre, du type «impossible d'éjecter le disque car il est en cours d'utilisation». NE PAS FORCER L'ÉJECTION ! Chercher d'abord à quel niveau le disque est utilisé, si le fichier est fermé, mais que Photoshop est toujours ouvert, fermer Photoshop. Si le problème persiste, éteindre l'ordinateur. Sinon vous risquez au mieux de perdre la table d'index, au pire de vous retrouver avec quelques fichiers corrompus ou de griller complètement le disque.

Stockage et Rangement

Préférer un rangement au sec, à la verticale, sans poussières. Éviter les écarts de température, ou le soleil. Faire tourner le disque régulièrement (1 à 2 fois par an). Regarder aussi du côté des boîtiers ignifugés.



*Armoire de rangement pour 60 disques dur,
vendu chez Cdiscount à 500€*



*Boite de rangement en PVC pour disque dur
3.5", vendu chez Cdiscount à 10€*

Fragmentation et défragmentation

Sur certains systèmes de fichiers/OS, les fichiers sont répartis un peu partout, éparpillés, façon puzzle, en plein de petits morceaux, là où il y a de la place. Quand un programme veut accéder à ses fichiers, il est donc nécessaire de retrouver chaque petit morceau (cf table d'index), ce qui prend du temps et ralentit l'ordinateur. Réaliser une défragmentation va permettre de «réorganiser» ses morceaux de fichiers et de les regrouper.

Lors d'une défragmentation, il est préférable d'avoir suffisamment d'espace disque disponible afin de pouvoir «déplacer» temporairement de gros fichiers (les cases du bas sur le GIF). Pour réaliser une défragmentation, je vous conseille Defragler sur Windows (plus performant que l'outil natif), et IDefrag est conseillé pour Mac.

/!\ La défragmentation est fortement déconseillée (interdite) sur des disques de type SSD. Elle entraîne une usure prématurée des blocs d'écritures, alors que la fragmentation des données sur un SSD n'a pas d'incidence sur les performances.

1	2	3	4	5	6	7	1	2	3
4	8	9	10	1	2	3	1	2	3
4	5	6	5	4	5	1	2	3	4
5	6	7							

Disque fragmenté

1	2	3	4	5	6	7			
	8	9	10	1	2	3	1	2	3
4	5	6		4	5	1	2	3	4
5	6	7							

Un élément est supprimé

1	2	3	4	5	6	7	1	2	6
7	8	9	10	1	2	3	1	2	3
4	5	6	8	4	5	1	2	3	4
5	6	7							

D'autres éléments sont dis-
patchés dans les «trous»

1	2	3	4	5	6	7	8	9	10
7				1	2	3	1	2	3
4	5	6	8	4	5	1	2	3	4
5	6	7							
						6	2	1	

Défragmentation en cours

1	2	3	4	5	6	7	8	9	10
1	2	1	2	3	4	5	6	7	8
1	2	3	4	5	6	1	2	3	4
5	6	7							

Défragmentation terminée

Archivage et Stockage des fichiers

4. Organisation et nomenclature

Si nous étions bien organisés, nous n'aurions pas besoin de fouiller dans nos dossiers. Il suffirait simplement de lancer une recherche dans notre «finder» favoris avec certains mots-clés pour trouver notre précieux.

Pour cela, la meilleure méthode selon l'expérience de Yann, consiste à organiser son arborescence par date.

```
2016
2017
  01.JANVIER
    PHOTOGRAPHE_AGENCE_TYPE_TITREBOULOT_CLIENT(_part01)
      RAW
        SELECT
        TRASH
      PSD
      TIFF (definitifs)
        RVB
        CMJN
      JPEG (HD, sans compression)
      MAILS ENVOYES (jpeg basse def de validations)
        20170123
      RESSOURCES (brush, textures...)
      BRIEF (infos, mails, pdf)
      ADMIN (devis, factures, cessions de droits, embauches...)
  02.FEVRIER
  03.MARS

2018
```

Le dossier JPEG permet une recherche sur le terme JPEG donnant un aperçu global de tout le travail effectué. Ce dossier est LE dossier à synchroniser sur un Cloud (de préférence automatiquement).

Astuce:

Créer à la racine une arborescence vierge (tous les dossiers, sans fichiers) qu'on viendra coller au bon endroit et renommer selon le projet, et ainsi gagner du temps !

Sauvegarde

1. Comment faire une sauvegarde ?

Les différentes méthodes de sauvegarde

Il existe trois méthodes de sauvegarde, proposées par la plupart des logiciels d'automatisation.

- **La sauvegarde complète:** copie de toutes les données indépendamment des modifications (ou non) depuis la sauvegarde précédente.
- **La sauvegarde incrémentielle:** copie des documents nouveaux ou modifiés depuis la dernière sauvegarde.
- **La sauvegarde différentielle:** seulement les documents nouveaux ou modifiés depuis la dernière sauvegarde COMPLÈTE sont copiés.

En plus de ces méthodes, on trouvera la synchronisation (ce qui est fait d'un côté se répercute automatiquement de l'autre, genre un fichier supprimé par inadvertance n'est pas récupérable¹) et la copie pure et dure à la main (qu'on va tenter d'éviter, hein!).

¹ Sauf certains cloud/drive qui gardent temporairement une «corbeille» où il est possible d'aller rechercher des fichiers...

Pour rappel:

Une bonne sauvegarde suit la règle 3-2-1. 3 copies des données, stockées sur 2 supports différents et 1 copie délocalisée. Mais ça je vous l'ai déjà dit. Les deux supports différents permettent de contrer les défauts de matériel, l'obsolescence, l'arrêt de production des périphériques de lecture, etc. La copie hors site assure une sauvegarde des données même en cas d'inondation/feu/vol, etc.. <http://yeux-coccinelle.fr/blog>

Le RAID

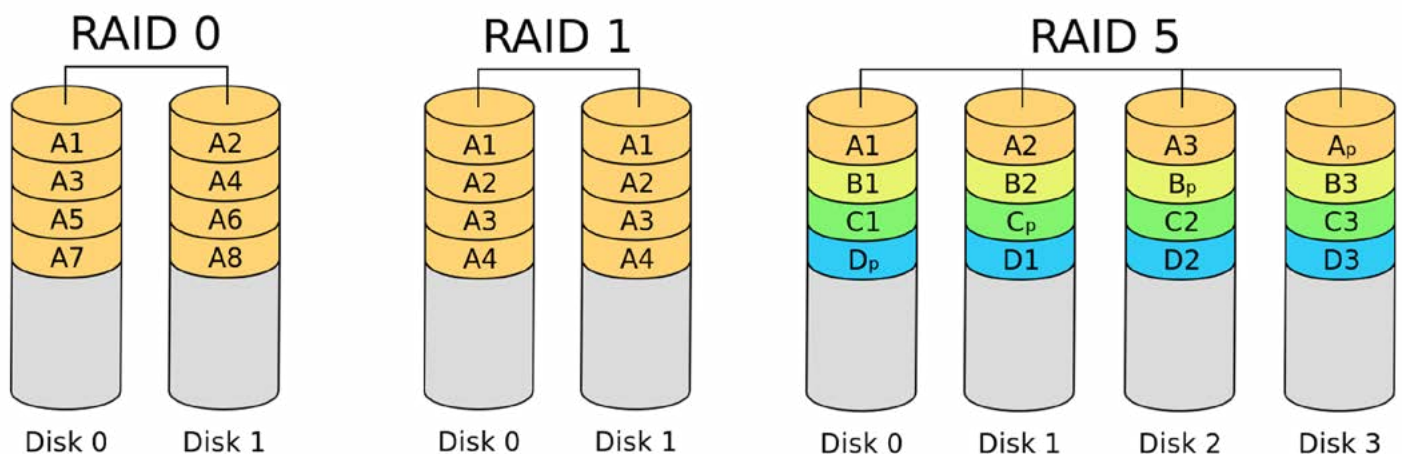
De manière plus matérielle, le RAID va nous permettre de mettre en place notre stratégie de sauvegarde en terme physique. Le terme RAID (Redundant Array of Inexpensive Disks) définit une technologie de regroupement de disques durs physiques indépendants pour ne former qu'une seule unité. Ils permettent d'améliorer les performances, la sécurité et la tolérance aux pannes.

On peut combiner des disques au sein d'une unité centrale (dans un ordi) ou dans une baie (une boîte faite pour). Cependant il existe plusieurs types de RAID, plus ou moins «sécuré» à utiliser pour nos sauvegardes:

- **RAID 0**: On combine tous les disques pour qu'ils ne forment qu'un seul gros. L'espace disponible correspond au nombre de disques multiplié par la PLUS PETITE capacité des disques. Il est donc conseillé de prendre des disques de même tailles. Ce type de RAID n'offre aucune sécurité face aux pannes: si un seul disque tombe en panne, on perd toutes les données.

- **RAID 1** ou mirroring: comme son nom l'indique il consiste à copier les données en «miroir». Il peut se faire à partir de deux disques: l'un est la copie conforme de l'autre. Si l'un des disques tombe en panne, pas de soucis, il y a le second qui possède toutes les données. En revanche ce système coûte assez cher puisqu'il faut exactement la même capacité sur les deux disques. On achète 2x 2To (4To en tout), mais seulement 2To sont disponibles.

- **RAID 5**: trois disques minimum sont requis ici. Les données sont partagées entre les disques, avec une information «générale» pour chaque donnée sur un autre disque, ce qui permettra en cas de panne de recréer le disque manquant à partir des autres. C'est la méthode la plus sécurisée des trois, mais elle nécessite l'achat de 3 disques.



Les différents logiciels

Il existe de très nombreux logiciels de sauvegarde. Leur principal avantage: automatiser les tâches, et vous libérez de cette corvée!

- Backup en ligne:

- *Crashplan*: service de sauvegarde en ligne disponible sur Windows, Mac et Linux. Stockage illimité, versionning (garde des «versions» des fichiers, en fonction des modifications), «Corbeille» infinie. Backup sur disque externe possible. /!\ Arrêt de l'outil pour les particuliers en octobre 2018. CrashPlan se concentre maintenant sur les entreprises. 10€/mois.
 - *Pixelapse*: créer pour les designers s'est fait intégrer par Dropbox et est maintenant fermé.
 - Encore ouverts, *Carbonite* et *BlackBaze* sauvegardent tout type de fichiers, proposent un chiffrement par clé privée et gèrent les versions. *Mozy* ne propose pas de stockage illimité et est trop limité pour nos besoins de photographes. *SpiderOak One* propose des offres allant de 5€/mois pour 100Go à 25€/mois pour 5To, en sauvegarde et/ou synchronisation, au choix. Clé privée et versionning sont de la partie.
 - Pour nous, étudiants à Gobelins, qui n'avons pas peur de céder nos informations personnelles à Google: *Google Drive*. Nous avons la possibilité d'utiliser le Drive gratuitement et de manière illimitée. Outil de Sauvegarde et/ou de Synchronisation (qui change de nom toutes les semaines depuis 4 mois).
 - Les services de Microsoft, Dropbox, Hubic (OVH)....
- Les vrais **logiciels de sauvegarde**, qui gèrent les tâches, le versionning (normalement), le choix des dossiers, etc.
- *Cobian Backup*: il est moche, mais ultra puissant, et gratuit. Mais disponible seulement sur Windows
 - *Acronis true image*: disponible sur Windows, Mac et Linux, mais payant.
 - *Time Machine*: le soft préinstallé Made in Apple
 - *Carbon Copy Cloner*: disponible uniquement sur Mac, deux espaces, un complet et un simple (très utile en shooting par exemple).
 - Les softs propriétaires que vous avez eus avec vos disques durs ou NAS (chez Synology ou Western Digital par exemple)

Des outils physiques adaptés: les NAS

Un NAS (Network attached storage) est un espace de stockage relié au réseau. Il permet de sauvegarder ou synchroniser des données depuis plusieurs ordinateurs, mais aussi de diffuser photos, vidéos et musiques via les box ou Airplay, par exemple. De plus, l'accès est possible et sécurisé (normalement) depuis l'extérieur à partir d'un simple navigateur web. Un drive «chez soi»; pas sur les serveurs de Google quoi !

En NAS grand public on trouve les Western Digital (Mycloud), les Synology (avec en bonus le DiskStation Manager, système d'exploitation Web complet,) ou les Qnap (plus orienté expert).



NAS Western Digital My Cloud EX2 Ultra
2x2To à 300€ chez Materiel.net



NAS Synology DiskStation DS216se 2 baies
140€ chez Materiel.net sans disque dur

Le magazine «[Workflow](#)» n°14 en plus de proposer un guide sur les disques durs, parle des NAS et de comment les installer chez soi. Je vous le conseille! Même s'il est un peu vieux (mars 2016).

Lauréline Reynaud

Sauvegarde

2. Intégrité d'une sauvegarde

Une sauvegarde ne sert à quelque chose que si elle est 1. à jour et 2. en bonne forme. Vous allez vouloir récupérer votre sauvegarde en cas de problème matériel, mais aussi en cas d'infections. Or, si ces infections ont aussi touché votre sauvegarde, tout est perdu.

Il existe donc des solutions pour vérifier si les fichiers sauvegardés sont identiques aux fichiers de base (déjà), s'ils sont bien lisibles, et pour éviter qu'ils chopent une MST.

Sortir Couvert

On n'ouvre pas les pièces jointes envoyées par des inconnus (on vérifie les adresses email), on ne clique pas sur les pop-up de pub, on évite les sites pornos, on télécharge nos logiciels directement sur le site de l'éditeur, surtout pas chez 01net/clubic/softonic/logiciel.net/winportal, etc. Quand on installe quelque chose, on fait bien gaffe à ce qui est coché (on oublie les toolbar, les logiciels autres que celui qu'on voulait, etc.). On ne lance pas les .exe à l'aveuglette... Une image n'a pas comme extension un .exe . Un film non plus!

Idéalement, un poste informatique exposé ne doit pas pouvoir accéder aux sauvegardes de lui même...

Et bien sûr, on utilise des antivirus, antimalware (même sur Mac hein, si les virus agissent moins sur l'OS d'Apple que sur celui de Microsoft, il peut quand même en être vecteur, et le refiler aux clés USB par exemple), etc.. Et on fait ses mises à jour.

Vérifications régulières

- Régulièrement, tous les 2/3 mois min on fait un test manuellement: on télécharge quelques fichiers parmi plusieurs dossiers (anciens et récent) et on voit s'ils sont toujours exploitables.
- [MD5 Checksum](#): permet de s'assurer que le fichier copié est exactement le même que l'original en vérifiant une chaîne d'octet. Je ne connais pas du tout le principe, je vous fais découvrir en même temps que moi. Je sais que c'est utilisé dans les boîtes de sécurité informatique. Je ne vais pas m'étendre au risque de dire des (grosses) conneries. Juste un détail pas top: si le fichier original est corrompu après la sauvegarde, le bon résultat est celui de la sauvegarde... Mais ça, les softs ne le précisent pas...

Sauvegarde

3. Que faire si je n'ai pas fait de sauvegarde ou si elle est corrompue ?

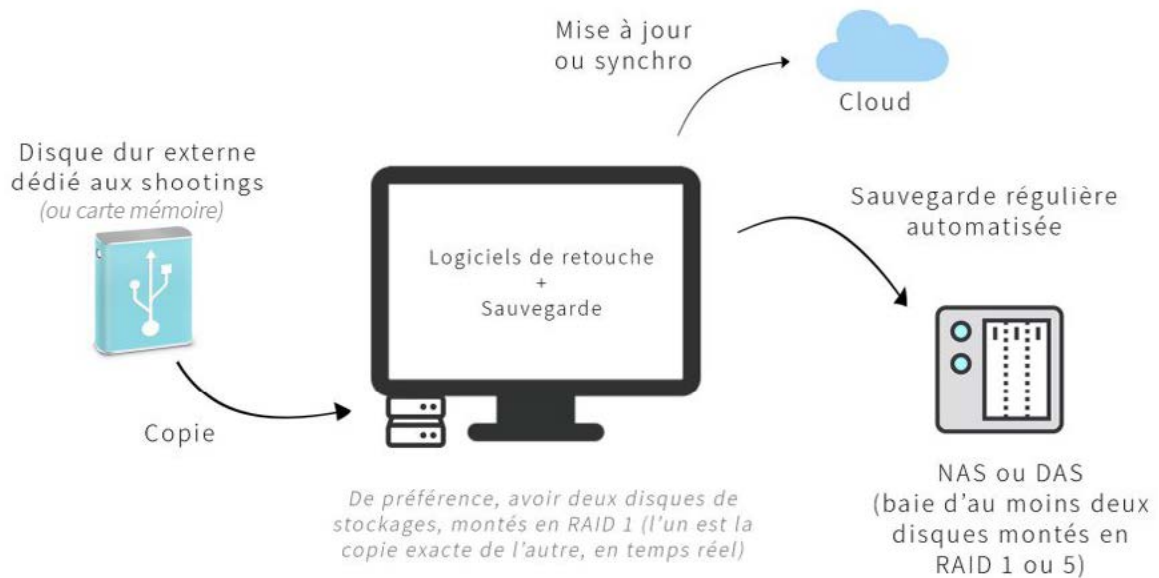
On pleure.

Des logiciels de récupérations de données existent. Par millions. La semaine dernière encore j'ai utilisé Recuva pour retrouver les images d'une carte mémoire: la carte était lisible par l'ordi, les noms de mes fichiers apparaissaient, mais impossible de les ouvrir ou de les copier. J'ai pu en récupérer une grosse partie (tous les JPEG) parfaitement. Les RAW sont passés à la trappe par contre... Des logiciels de reconstruction de table d'index existent aussi... Attention par contre, éviter le formatage ou pire, réécrire par dessus avant de tenter une récupération.

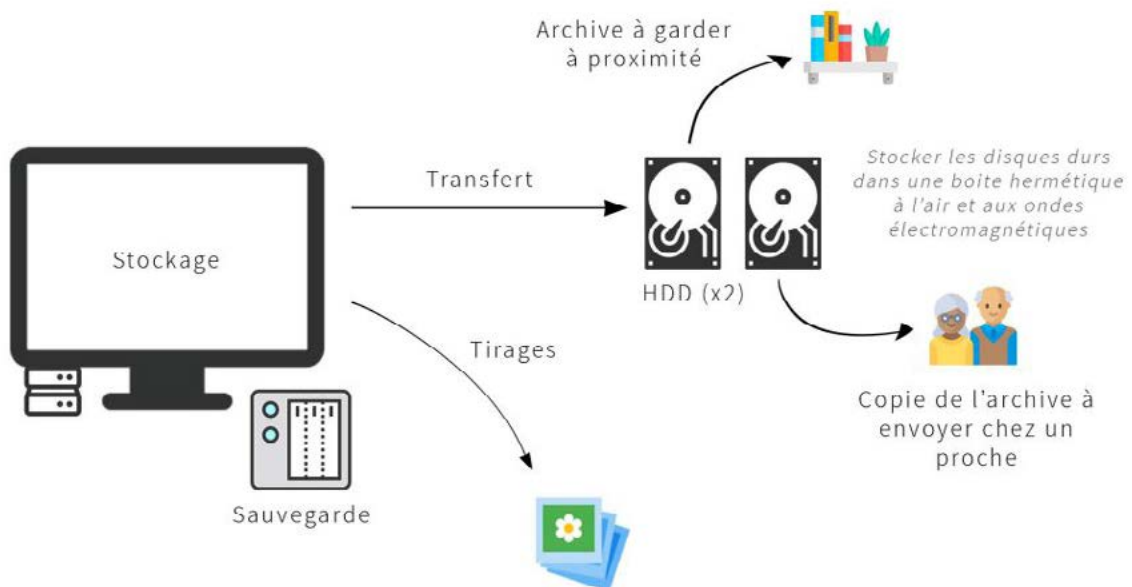
Et si jamais rien de fonctionne, les salles blanches sont là pour ça, sur devis (en général à partir de 300€ pour des petites capacités).

Sauvegarde

[4. Ma méthode de sauvegarde \(lien cliquable vers le blog\)](#)



Sauvegarde depuis le poste de travail



Archivage 1x/an