

Les méthodes de sauvegarde

Il existe plusieurs méthodes de sauvegarde. Toutes les méthodes définies ci-après ont un fonctionnement basé sur des attributs de fichiers spécifiques. Certains systèmes d'exploitation s'appuient sur l'attribut d'archivage (Microsoft) alors que d'autres travaillent sur les dates de fichiers (Unix)

1. Le bit d'archivage

Dans les systèmes d'exploitation se basant sur cet attribut, chaque fichier dispose d'un bit particulier dit « d'archivage ». Lorsque le fichier est créé ou modifié puis refermé, le système d'exploitation positionne ce bit à 1 pour indiquer qu'il a été créé ou modifié depuis la dernière sauvegarde. Le bit d'archivage permet aussi aux logiciels de sauvegarde d'identifier les fichiers à enregistrer, puis de repositionner le bit à 0 une fois la sauvegarde effectuée.

Pour les systèmes d'exploitation qui s'appuient sur la date des fichiers, le principe est beaucoup plus simple. Tous les fichiers possèdent une date de création et une date de dernière modification, ce qui permet au système de sauvegarde de les comparer à celle de la dernière sauvegarde. Les différents attributs de fichiers, dont les dates, sont stockés dans les inodes (index et nodes). Le système de sauvegarde va donc comparer cette date à celle de la dernière sauvegarde. Si elle est antérieure à la date de modification du fichier alors il enregistrera le fichier, sinon il sera ignoré lors de la sauvegarde.

2. La sauvegarde complète

La sauvegarde complète est très simple, elle ne se base pas sur l'attribut d'archivage pour choisir quels sont les fichiers à sauvegarder, elle les sauvegarde tous, qu'ils soient anciens, modifiés ou non. L'objectif est de réaliser une copie conforme des données à sauvegarder, sur un support séparé. Si les données ne changent jamais, chaque nouvelle sauvegarde complète est identique à la précédente.

Toutefois, en fin d'exécution, l'attribut d'archive du fichier est réinitialisé pour mémoriser le fait qu'il a été enregistré. Si la date est utilisée, celle de la dernière sauvegarde est mémorisée, de sorte à pouvoir différencier les fichiers qui ont été sauvegardés des autres, en fonction de leur date de dernière modification.

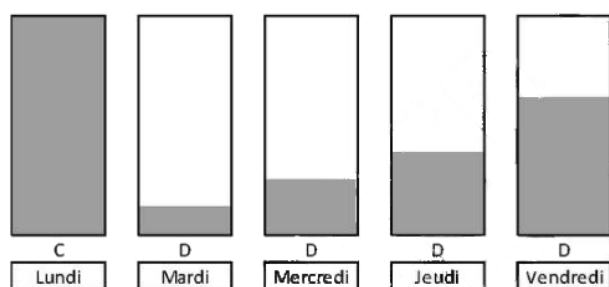
En ce qui concerne la restauration, il est nécessaire de restaurer la totalité de la sauvegarde complète en cas de pertes des données.

3. La sauvegarde différentielle

Contrairement à la sauvegarde complète, la sauvegarde différentielle (ou cumulative) se focalise uniquement sur les fichiers modifiés depuis la dernière sauvegarde complète appelée « base différentielle »

Lors de son exécution, elle archive tous les fichiers ayant l'attribut d'archivage positionné à 1 ou dont la date de modification/création est postérieure à celle de la dernière sauvegarde, mais ces attributs ne sont pas réinitialisés après la copie. Par conséquent, à chaque nouvelle sauvegarde différentielle, les modifications précédentes plus celles du jour sont prises en compte. C'est pourquoi cette méthode est dite cumulative.

Le plus souvent, la stratégie choisie dans ce contexte consiste à combiner une sauvegarde complète effectuée périodiquement de concert avec des sauvegardes différentielles fréquentes.



C : Sauvegarde complète

D : Sauvegarde différentielle

Une sauvegarde différentielle effectuée peu de temps après sa base est beaucoup plus petite que sa base différentielle, ce qui permet un gain de temps de sauvegarde et d'espace de stockage. En revanche, au gré des modifications apportées aux données, la différence entre cette dernière et la base différentielle augmente. Plus le temps écoulé entre une sauvegarde différentielle et sa base augmente, plus la taille de la sauvegarde peut être importante. A terme, cela signifie que la taille des sauvegardes différentielles peut atteindre celle de la base différentielle.

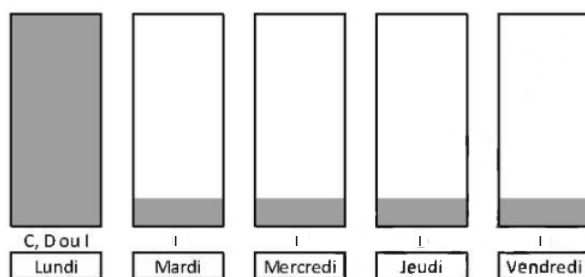
Pour restaurer l'ensemble des données, il faut se munir de la dernière sauvegarde complète (la base différentielle) puis de la dernière sauvegarde différentielle.

4. La sauvegarde incrémentale

Le fonctionnement de la sauvegarde incrémentale est semblable à celui de la sauvegarde différentielle puisqu'il consiste à copier tous les éléments créés ou modifiés depuis la dernière sauvegarde quel que soit son type (complète, incrémentale ou différentielle).

La différence vient du fait qu'en fin de traitement, les attributs d'archive du fichier sont réinitialisés pour signifier qu'ils ont été sauvegardés.

Par exemple, dans le cadre de sauvegardes quotidiennes, si une sauvegarde complète est réalisée un jour J, la sauvegarde incrémentale du jour J+1 sera faite en référence au jour J, celle du jour J+2 en référence au jour J+1, etc...



C : Sauvegarde complète

D : Sauvegarde différentielle

I : Sauvegarde incrémentale

Pour restaurer l'intégralité des données, il faut donc se munir de la dernière sauvegarde complète ainsi que toutes les sauvegardes incrémentales effectuées depuis.

Certains logiciels de sauvegarde utilisent notamment une combinaison de sauvegardes complètes et incrémentales pour optimiser le temps de sauvegarde des données.

Sources : La sauvegarde de données en entreprise, Sylvain Gasnier, 2013