POOL ET ESPACES DE STOCKAGE-CLIENT

# INTRODUCTION

La gestion des disques physiques sur un système d’exploitation peut se faire de plusieurs manières, ainsi à partir d’un disque physique nous pouvons créer plusieurs disques logiques.

Il est préférable de pouvoir séparer le système et les données en les plaçant sur des partitions distinctes. Seulement si le disque dur tombe en panne, les différentes partitions seront inaccessibles.

L’idéal serait d’utiliser deux disques durs physiques distincts.

Mais même dans ce cas, le risque demeure de tout perdre.

La solution va venir de l’utilisation d’un système de stockage permettant de ne pas perdre ses données même en cas de panne d’un disque dur.

# Définition

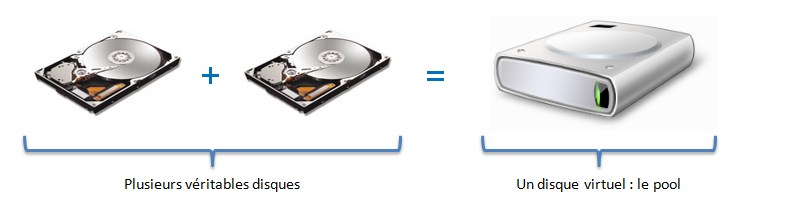
## Pool

Un pool, que l’on pourrait traduire par « groupe » ou même « grappe » (comme une grappe de raisin, par exemple), est l'agrégation de plusieurs disques durs physiques en un seul disque dur logique. Autrement dit, vous prenez plusieurs vrais disques durs et vous faites croire à Windows qu’il n’y en a qu’un seul.

Avec 3 disques de 5 Giga, la création d'un pool de stockage permettra d'obtenir un seul disque de 15 Giga.

Une des forces du système proposé par Windows est qu’il peut associer dans un pool plusieurs disques durs de marques et de capacités différentes. Ces disques peuvent être internes ou externes (branchés en USB par exemple).

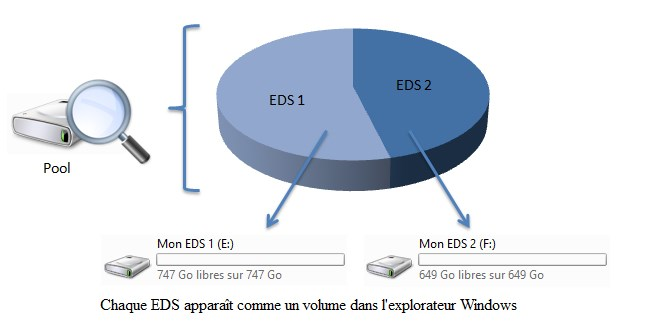
Les disques ajoutés au pool de stockage doivent être vides



## Espace de stockage

L’EDS est à un pool ce qu’une partition est à un vrai disque dur. Un EDS est donc une section du pool, qui aura sa lettre de lecteur dans l’explorateur Windows (D:, E:, F:, etc.).

Chaque EDS apparaît comme un volume dans l'explorateur Windows

Il est possible de créer plusieurs EDS par pool, comme c’est le cas sur le schéma ci-dessus, mais il est courant de n’avoir qu’un seul EDS occupant la totalité de l’espace du pool.  


# PROTECTION

En informatique, le terme **résilience** désigne la capacité d’un système, quel qu’il soit, à continuer de fonctionner en cas de panne. Dans le cas qui nous intéresse ici, le système à protéger est l’intégrité de nos données, les pools et espaces de stockage ont cet avantage de pouvoir « survivre » à la panne d'un (ou même plusieurs) disque(s) dur(s). Pour cela, il faut choisir l’une des quatre stratégies de résilience proposées par Windows :

* Miroir double ;
* Miroir triple ;
* Parité ;
* Simple (sans résilience).

On applique l’une de ces stratégies à un EDS au moment de sa création.

Pour illustrer les différences entre ces stratégies, je vais prendre le cas où un seul EDS est créé au sein d’un pool de disques durs, mais le principe est le même si plusieurs EDS sont définis.

#### RAID

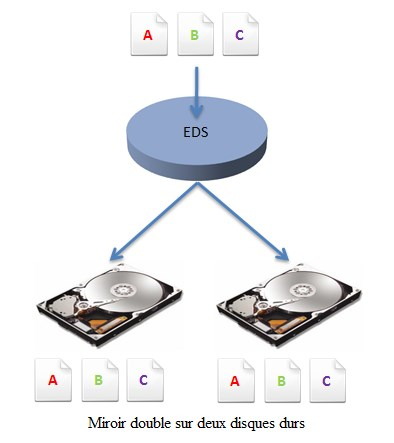
On compare habituellement les pools et espaces de stockage au « RAID », qui est lui aussi un système permettant de protéger les données ou d'y accéder plus rapidement (ou les deux). Mais attention, ce que propose Windows n'est pas à proprement parler du RAID. Redundant Array of Independent Disks, ce qui signifie « regroupement redondant de disques indépendants » Le coût important des disques physiques à hautes capacités a poussé les ingénieurs à rechercher des solutions, ainsi il a été possible en combinant plusieurs disques de petits formats d’obtenir pour un même espace de stockage des performances meilleurs et un coût inférieur à un disque de grand format. Aujourd’hui la baisse des prix de l’octet de sauvegarde a permis de se focaliser sur d’autres qualités que le coût immédiat des disques.

RAID 0, 1 et 5 sont les différents niveaux les plus courants de RAID que l’on peut rencontrer.

### « Miroir double »

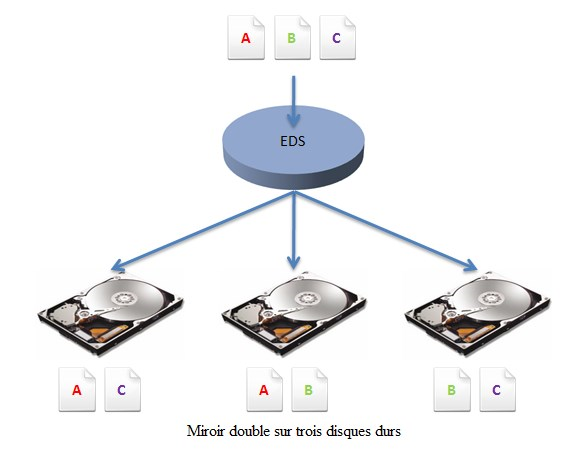
Pour s’assurer que ses données ne seront pas perdues en cas de problèmes, la solution la plus simple est la sauvegarde. Avec une résilience en miroir double, les données sont copiées sur deux disques à la fois au sein du pool.

Dans l’exemple ci-dessous, nous enregistrons trois fichiers (notés A, B et C) sur notre EDS, composé de deux disques durs :



Les fichiers sont enregistrés sur le premier disque dur et sur le second. Dans notre EDS, nous ne voyons bien sûr les fichiers qu’une seule fois, mais ceux-ci sont en réalité dupliqués automatiquement. Dans ce cas, si un disque dur tombe en panne, il est toujours possible de retrouver les deux fichiers. Ce type de résilience nécessite donc que le pool contenant l’EDS soit composé d’au moins deux disques durs.

Si le pool contient plus de disques, les fichiers sont toujours enregistrés en deux endroits, pas plus. Par exemple, avec trois disques :



Dans ces conditions, on peut toujours survivre à la panne d’un disque dur, mais pas plus. Chaque fichier n’est toujours copié qu’en deux endroits différents. Si l’un des disques durs tombe en panne, les fichiers A, B et C sont toujours accessibles. En revanche :

si les disques 1 et 2 tombent, le fichier A n’est plus présent nulle part ;

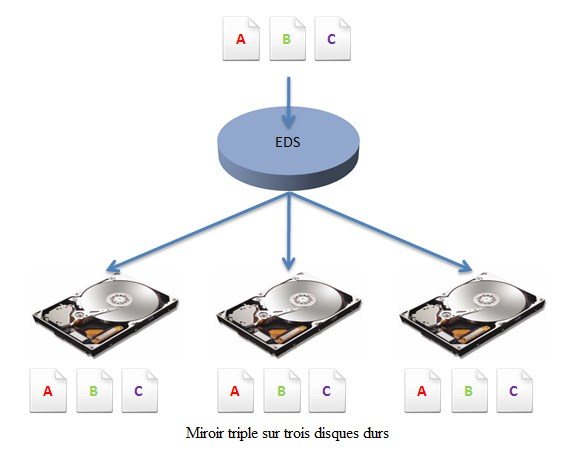
si les disques 1 et 3 tombent, le fichier C n’est plus présent nulle part ;

si les disques 2 et 3 tombent, le fichier B n’est plus présent nulle part.

Pour plus de sécurité, il faut donc copier les fichiers… trois fois ! C’est le type de résilience miroir triple.

### « Miroir triple »

Avec une résilience en miroir triple, le principe est exactement le même, si ce n’est que les données sont copiées trois fois. Ce type de résilience nécessite donc au moins trois disques durs mais est capable de protéger vos données même si deux disques tombent en panne en même temps.



### « Parité »

Le type de résilience parité est le plus intéressant de la série car il combine deux améliorations : la protection des données et leur rapidité d’écriture sur le pool de disques.  
Ecrire sur un disque dur est très long pour un ordinateur, c’est l’une des choses les plus lentes qui demeurent (ouf le SSD est arrivé). Lorsque l’on dispose d’un pool de disques (et donc d’EDS), il devient possible d’accélérer les choses en écrivant **en parallèle** sur plusieurs disques. Ainsi, si on mettait un temps T à écrire un fichier sur un disque dur, on met alors un temps T/2 à écrire ce même fichier sur deux disques dur en même temps (en théorie bien sûr). C’est ce qu’il se passe avec le mode de résilience parité.

Mais si la rapidité est à l’honneur, la sauvegarde reste présente en nécessitant au moins 3 disques.

Lorsqu’un fichier est écrit sur l’EDS, il est réparti de cette façon sur les disques du pool :

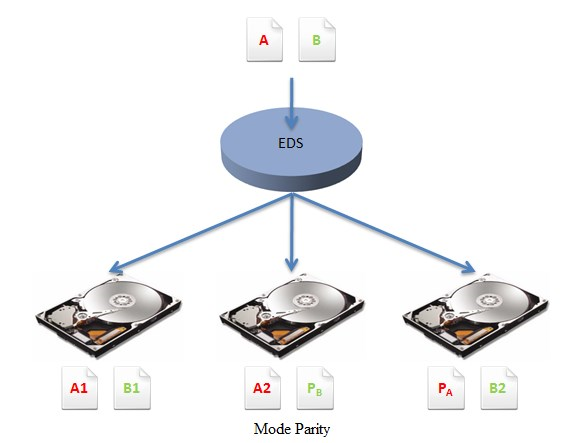
un morceau de fichier (la moitié dans ce cas) est écrit sur le premier disque ;

un autre morceau du fichier (l’autre moitié) est écrit sur le deuxième disque ;

des « informations de parité » sont écrites sur le troisième.

Ces informations de parité sont des données de récupérations, qui permettent de retrouver les morceaux de fichiers perdus, en les associant aux morceaux de fichiers qui sont toujours présents. Ainsi, les données peuvent être récupérées si un (et un seul) des disques tombe en panne.

Prenons le cas de deux fichiers, écrits sur un pool de trois disques durs :



Le fichier A est découpé en deux morceaux, A1 et A2, qui sont répartis sur les disques du pool comme le montre le schéma ci-dessus. Les informations de parité associées sont notées PA et sont placé sur le troisième disque. Idem pour le fichier B

Selon la répartition faite sur le schéma ci-dessus, voyons ce qui se passe en cas de pannes…

Si le disque 1 tombe en panne :



Rappelons qu’un fichier ne peut être complet que si ses deux parties sont disponibles.  
Concernant le fichier A, le morceau A1 est perdu. A2 est toujours disponible sur le disque 2 et les informations de récupérations PA sont toujours sur le disque 3. Grâce à la combinaison de A2 et PA, A1 peut être retrouvé. On a donc A1 et A2. On a donc A tout entier !  
La logique est exactement la même pour le fichier B, on a donc bien pu récupérer nos deux fichiers.

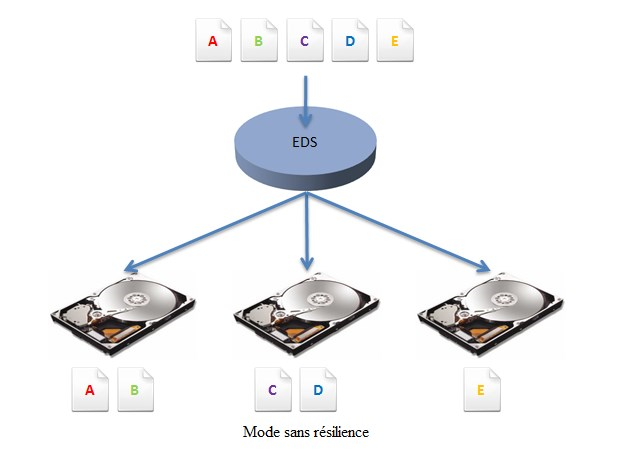
Si deux disques tombent en panne en même temps :



Prenons l’exemple où les disques 1 et 3 tombent en panne simultanément. Cette fois, A1 est inaccessible et ne peut pas être retrouvé car les informations de récupération PA sont également perdues. Concernant B, seules PB est toujours en état, ce qui ne nous avance pas beaucoup plus.  
Conclusion : le mode de résilience parité ne permet de récupérer les données que si l’un des disques tombe en panne, mais pas plus.

### « Simple (sans résilience) »

Les termes sans résilience sont clairs : aucune protection des données ne sera appliquée. Si l’un des disques durs tombe en panne, l’ensemble des données seront inaccessibles. L’EDS utilise l’ensemble des disques sans aucune redondance.



## Introduction au RAID

Les différents niveaux de RAID :

**RAID 0** utilise la répartition (striping) des données sur plusieurs disques pour accélérer le débit de données, notamment pour les fichiers volumineux, dans les environnements ne nécessitant pas de redondance des données.

**RAID 1** exploite la mise en miroir des disques afin d'écrire les données simultanément sur deux disques physiques. Il est particulièrement adapté aux bases de données de petite taille ou aux applications peu gourmandes en ressources, mais nécessitant une redondance totale des données.

**RAID 5** combine répartition sur disques et stockage de données de parité sur tous les disques physiques (parité distribuée) afin de fournir un haut débit et une redondance des données, particulièrement pour les accès aléatoires aux données peu volumineuses.

**RAID 6** est une évolution du RAID 5. Il utilise un bloc de parité supplémentaire. Il met en place une répartition au niveau du bloc, avec deux blocs de parité répartis sur tous les disques membres du groupe. RAID 6 protège des situations où deux disques tombent en panne, ainsi que des pannes qui surviennent lorsqu'un seul disque est en cours de reconstruction. Si vous utilisez une seule matrice, la mise en place d'un système RAID 6 est plus efficace que l'utilisation d'un disque de rechange.

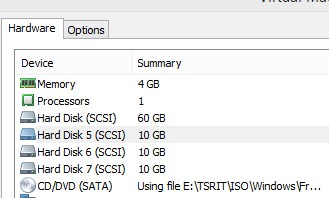
**RAID 10** est une combinaison de RAID 0 et de RAID 1 et utilise la répartition sur des disques mis en miroir. Il fournit un haut débit et une redondance totale des données. RAID 10 prend en charge jusqu'à huit matrices et jusqu'à 32 disques physiques par matrice.

**RAID 50** est une combinaison de RAID 0 et RAID 5 dans laquelle une matrice RAID 0 est répartie sur plusieurs éléments RAID 5. RAID 50 nécessite au moins six disques.

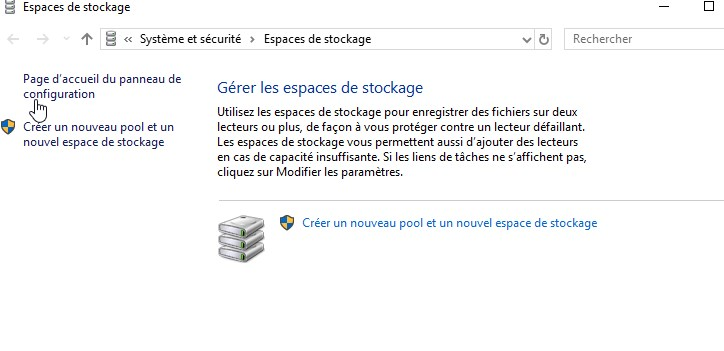
**RAID 60** est une combinaison de RAID 0 et RAID 6 dans laquelle une matrice RAID 0 est répartie sur plusieurs éléments RAID 6. Ce niveau nécessite au moins 8 disques.

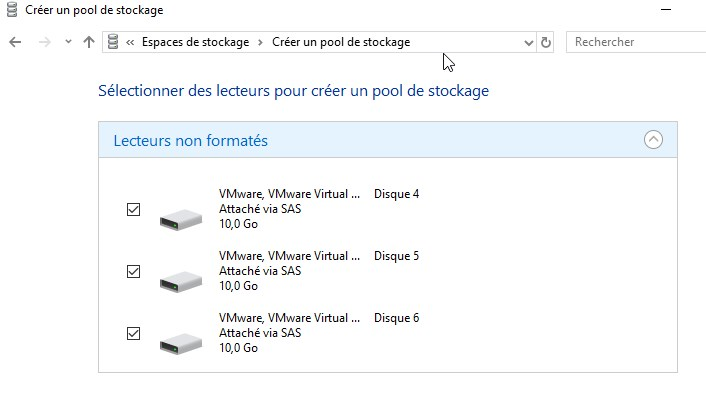
TP

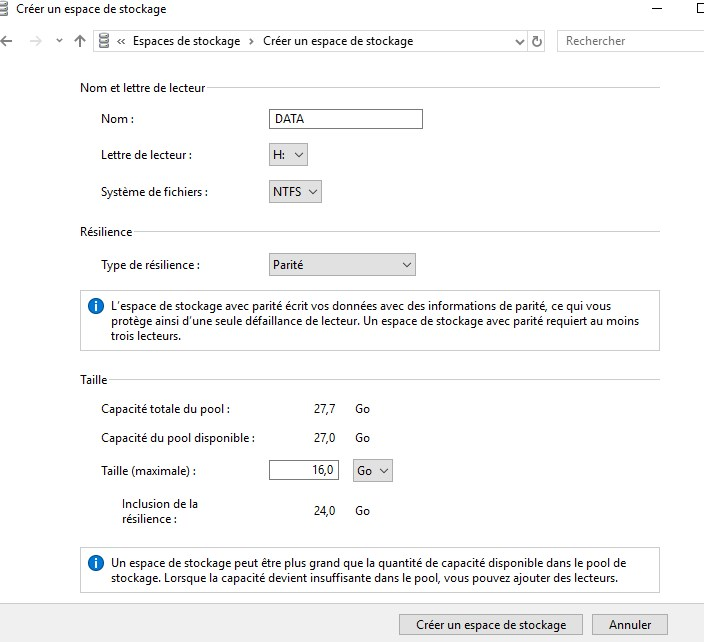
Sur vmware : créer 3 data disk de 10Gb chacun

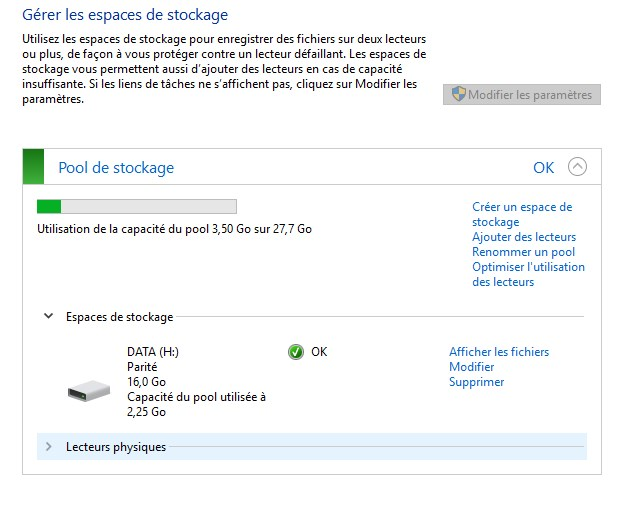


Créer un nouveau pool









(Facultatif) Pour accélérer la préparation des lecteurs, empêchez votre PC de passer en veille. Tapez **Alimentation et mise en veille** dans la zone de recherche, puis sélectionnez **Paramètres d’alimentation et de mise en veille**. Sous **En cas de branchement sur le secteur, mettre le PC en veille après**, sélectionnez **Jamais**.

Si vous avez créé un pool dans Windows 10 ou mis à niveau un pool existant, vous pouvez supprimer un lecteur dans ce dernier. Les données stockées sur ce lecteur seront déplacées vers d’autres lecteurs du pool et vous serez libre d’utiliser ce lecteur pour autre chose.