

Raid : principes

1 PRESENTATION DE LA TECHNOLOGIE RAID

La technologie **RAID** (acronyme de **Redundant Array of Inexpensive Disks**, ou parfois **Redundant Array of Independent Disks**, traduisez Ensemble redondant de disques indépendants) permet de constituer une unité de stockage à partir de plusieurs disques durs. L'unité ainsi créée (appelée grappe) a donc une grande tolérance aux pannes (haute disponibilité), ou bien une plus grande capacité/vitesse d'écriture. La répartition des données sur plusieurs disques durs permet donc d'en augmenter la sécurité et de fiabiliser les services associés.

Les disques assemblés selon la technologie RAID peuvent être utilisés de différentes façons, appelées Niveaux RAID. L'Université de Californie en a défini 5, auxquels ont été ajoutés les niveaux 0 et 6. Chacun d'entre eux décrit la manière de laquelle les données sont réparties sur les disques:

- Niveau 0: appelé striping
- Niveau 1: appelé mirroring, shadowing ou duplexing
- Niveau 2: appelé striping with parity (obsolète)
- Niveau 3: appelé disk array with bit-interleaved data (peu employé)
- Niveau 4: appelé disk array with block-interleaved data (peu employé)
- Niveau 5: appelé disk array with block-interleaved distributed parity
- Niveau 6: appelé disk array with block-interleaved distributed parity (peu employé)

Chacun de ces niveaux constitue un mode d'utilisation de la grappe, en fonction:

- des performances
 - du coût
 - des accès disques
-

1.1.1 Niveau 0

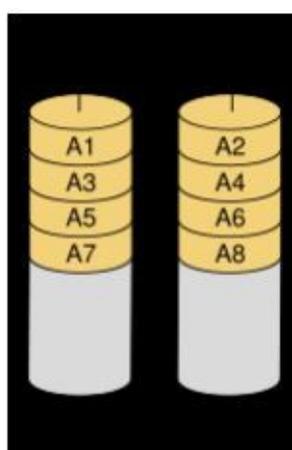
Le RAID-0, appelé aussi **striping** consiste à stocker les données en les répartissant sur l'ensemble des disques de la grappe. De cette façon, il n'y a **pas de redondance, on ne peut donc pas parler de tolérance aux pannes**. En effet en cas de défaillance de l'un des disques, l'intégralité des données réparties sur les disques sera perdue.

Cette solution est juste destinée à améliorer les temps d'accès aux disques. Étant donné que chaque disque de la grappe a son propre contrôleur, cela constitue une solution offrant une vitesse de transfert élevée.

Le RAID 0 consiste ainsi en la juxtaposition logique (agrégation) de plusieurs disques durs physiques. En mode RAID-0 les données sont écrites par "bandes" (en anglais stripes) :

Mettre en œuvre une tolérance de panne sur un stockage de données

Apport d'informations et de connaissances



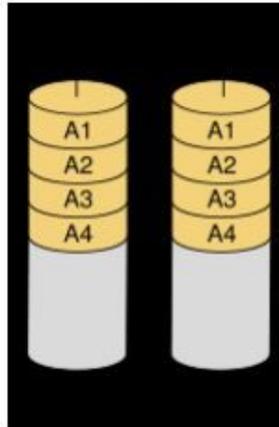
Si un des éléments de la grappe est plus grand que les autres, le système de remplissage par bande se trouvera bloqué lorsque le plus petit des disques sera rempli. La taille finale est ainsi égale au double de la capacité du plus petit des deux disques :

Deux disques de 20 Go donneront un disque logique de 40 Go.

Un disque de 10 Go utilisé conjointement avec un disque de 27 Go permettra d'obtenir un disque logique de 20 Go (17 Go du second disque seront alors inutilisés).

1.1.2 Niveau 1

Le RAID 1, appelé aussi **mirroring**, ou miroir, a pour but de dupliquer l'information à stocker sur plusieurs disques, on parle donc de mirroring, ou shadowing pour désigner ce procédé.



On obtient ainsi une plus grande sécurité des données, car si l'un des disques tombe en panne, les données sont sauvegardées sur l'autre. D'autre part, la lecture peut être beaucoup plus rapide lorsque les deux disques sont en fonctionnement. Enfin, étant donné que chaque disque possède son propre contrôleur, le serveur peut continuer à fonctionner même lorsque l'un des disques tombe en panne, au même titre qu'un camion pourra continuer à rouler si un de ses pneus crève, car il en a plusieurs sur chaque essieu...

En contrepartie la technologie RAID1 est onéreuse étant donné que seule la moitié de la capacité de stockage n'est effectivement utilisée.

Mettre en œuvre une tolérance de panne sur un stockage de données

Apport d'informations et de connaissances

En cas de panne d'un disque, il faut réaliser une manipulation qui s'appelle « casser le miroir » et le système continue de fonctionner avec un seul disque jusqu'à l'ajout d'un disque supplémentaire et la reconstruction du miroir.

Pour mettre en place un RAID 1, vous devez disposer d'un **minimum de 2 disques**, il est rare que cette technologie soit utilisée avec plus de 2 disques.

1.1.3 Niveau 2

Le niveau RAID-2 est désormais obsolète, nous la citerons pour mémoire, car il propose un contrôle d'erreur par code de Hamming (codes ECC - Error Correction Code), or ce dernier est désormais directement intégré dans les contrôleurs de disques durs.

Cette technologie consiste à stocker les données selon le même principe qu'avec le RAID-0 mais en écrivant sur une unité distincte les bits de contrôle ECC (généralement 3 disques ECC sont utilisés pour 4 disques de données).

La technologie RAID 2 offre de piètres performances mais un niveau de sécurité élevé.

1.1.4 Niveau 3

Le RAID 3 propose de stocker les données sous forme d'octets sur chaque disque et de dédier un des disques au stockage d'un bit de parité.

Disque 1	Disque 2	Disque 3	Disque 4
Octet 1	Octet 2	Octet 3	Parité 1+2+3
Octet 4	Octet 5	Octet 6	Parité 4+5+6
Octet 7	Octet 8	Octet 9	Parité 7+8+9

De cette manière, si l'un des disques venait à défaillir, il serait possible de reconstituer l'information à partir des autres disques. Après "reconstitution" le contenu du disque défaillant est de nouveau intègre. Par contre, si deux disques venaient à tomber en panne simultanément, il serait alors impossible de remédier à la perte de données.

Pour vulgariser le principe du bit de parité, prenons l'exemple suivant :

Supposons que nous ayons sur les disques 1,2 et 3, respectivement les trois nombres suivants 4 5 et 6. sur le 4^{ème} disque nous allons stocker la somme des 3 nombres soit la valeur 15.

Si le premier disque tombe en panne, pour recréer son contenu nous allons faire l'opération suivante : $15 - 5$ (valeur du disque 2) $- 6$ (valeur du disque 3) cela nous donne 4, qui est bien la valeur du disque 1 avant la panne. Voilà, vous avez compris le principe du bit de parité.

Le RAID 3 est peu employé.

1.1.5 Niveau 4

Le RAID 4 est très proche du RAID 3. La différence se trouve au niveau de la parité, qui est faite sur un secteur (appelé bloc) et non au niveau du bit, et qui est stockée sur un disque dédié.

Mettre en œuvre une tolérance de panne sur un stockage de données

Apport d'informations et de connaissances

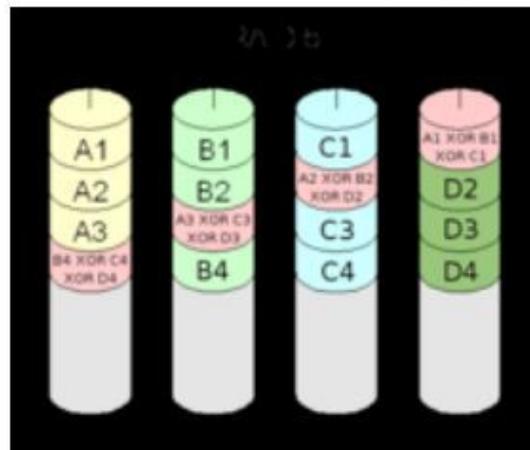
Disque 1	Disque 2	Disque 3	Disque 4
Bloc 1	Bloc 2	Bloc 3	Parité 1+2+3
Bloc 4	Bloc 5	Bloc 6	Parité 4+5+6
Bloc 7	Bloc 8	Bloc 9	Parité 7+8+9

Ainsi, pour lire un nombre de blocs réduits, le système n'a pas à accéder à de multiples lecteurs physiques, mais uniquement à ceux sur lesquels les données sont effectivement stockées. En contrepartie le disque hébergeant les données de contrôle doit avoir un temps d'accès égal à la somme des temps d'accès des autres disques pour ne pas limiter les performances de l'ensemble.

Le RAID 4 est peu employé.

1.1.6 Niveau 5

Le RAID 5 est similaire au RAID 4, c'est-à-dire que la parité est calculée au niveau d'un secteur, mais répartie sur l'ensemble des disques de la grappe.



De cette façon, le RAID 5 améliore grandement l'accès aux données (aussi bien en lecture qu'en écriture) car l'accès aux bits de parités est réparti sur les différents disques de la grappe.

Le mode RAID-5 permet d'obtenir des performances très proches de celles obtenues en RAID-0, tout en assurant une tolérance aux pannes élevées, c'est la raison pour laquelle c'est le mode RAID le plus intéressant en termes de performance et de fiabilité.

Pour mettre en place un RAID 5, vous devez disposer d'un **minimum de 3 disques**.

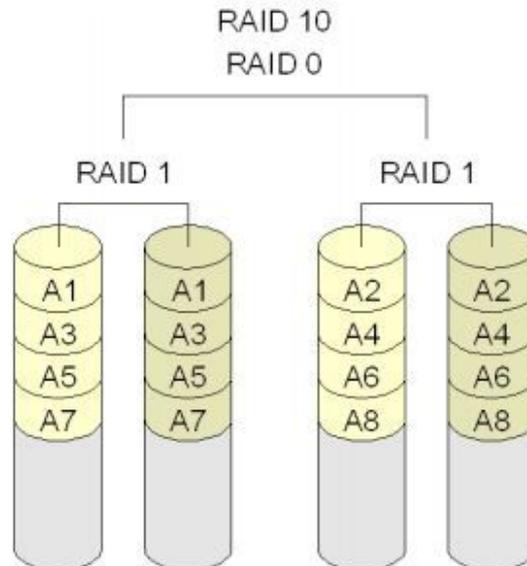
Avec le RAID 2 c'est le mode le plus employé.

1.1.7 Niveau 6

Le niveau 6 a été ajouté aux niveaux définis par Berkeley. Il définit l'utilisation de 2 fonctions de parité, et donc leur stockage sur deux disques dédiés. Ce niveau permet ainsi d'assurer la redondance en cas d'avarie simultanée de deux disques. Cela signifie qu'il faut au moins 4 disques pour mettre en œuvre un système RAID-6.

1.1.8 Les RAID combinés

Il est possible de combiner plusieurs niveaux de RAID, par exemple le RAID 1 et le RAID 0, on parle alors de RAID 10.



Cette combinaison est rarement employée sur un poste de travail et est plutôt réservée aux systèmes serveurs.

1.2 comparaison

Les solutions RAID généralement retenues sont le RAID de niveau 1 et le RAID de niveau 5.

Le choix d'une solution RAID est lié à trois critères :

- **La sécurité** : RAID 1 et 5 offrent tous les deux un niveau de sécurité élevé, toutefois la méthode de reconstruction des disques varie entre les deux solutions. En cas de panne du système, RAID 5 reconstruit le disque manquant à partir des informations stockées sur les autres disques, tandis que RAID 1 opère une copie disque à disque.
 - **Les performances** : RAID 1 offre de meilleures performances que RAID 5 en lecture, mais souffre lors d'importantes opérations d'écriture
 - **Le coût** : le coût est directement lié à la capacité de stockage devant être mise en œuvre pour avoir une certaine capacité effective. La solution RAID 5 offre un volume utile représentant 80 à 90% du volume alloué (le reste servant évidemment au contrôle d'erreur). La solution RAID 1 n'offre par contre qu'un volume disponible représentant 50 % du volume total (étant donné que les informations sont dupliquées).
-