

## 1 Introduction

RAID ou «Redundant Array of Inexpensive / Independent Disks » a été inventé en 1987 par trois chercheurs de l’université californienne de Berkeley, dans le but de pouvoir utiliser des disques de faibles capacités, donc peu coûteux (Inexpensive) de telles façons qu’ils soient vus comme un disque unique. De cette manière, le RAID peut être décrit comme étant l’opposé du partitionnement : dans un cas, on crée plusieurs unités logiques à partir d’un seul disque, dans l’autre, on crée une seule unité logique à partir de plusieurs disques physiques.

Le RAID est un ensemble de techniques de virtualisation du stockage permettant de combiner plusieurs disques durs (HDD ou SSD) afin d’améliorer les performances, la sécurité ou encore la tolérance aux pannes dans un système informatique.

Cette technique offre différents avantages. On peut tout d’abord noter que son prix est largement inférieur comparé au prix d’un seul et même périphérique de stockage pour des capacités de stockage, d’écriture et de lecture équivalente. En effet, le RAID permet aussi d’améliorer les vitesses de lecture et d’écriture des données sur les disques. Certains systèmes de RAID permettent quand à eux de limiter très fortement la perte de données. Quand on dispose d’un seul périphérique de stockage sur un système, nous sommes beaucoup plus sujet à la casse du disque dur que si l’on possédait des données réparties de façon “miroir” comme nous le verrons.

Il existe différents niveaux de RAID, nommés par leur numéro, possédant chacun leurs avantages et inconvénients.

Cependant, on peut les classer suivant leur objectif

- Palier la panne en utilisant un système de redondance ( RAID 1 )
- Améliorer les performances ( RAID 0 )
- Les deux à la fois, mais avec une moins bonne efficacité ( RAID 5 )

## 2 Parité et redondance, les concepts du RAID

La redondance consiste à dupliquer des données afin d’augmenter la fiabilité d’un système informatique. Une technique principalement utilisée pour l’architecture RAID et la régénération des données manquantes à partir de données restantes si un ou plusieurs disques tombent en panne.

Le système de redondance le plus simple et le plus utilisé est le calcul de parité. Ce système repose sur l’opération logique XOR. Supposons que l’on possède une bande de données, ces données sont réparties sur 2 blocs de disques et le troisième disque possède quand à lui le bloc de parité résultant de l’opération XOR entre les bits des données des deux autres disques. Sur  $n$  bits, si le nombre de 1 est pair, alors le bit de parité vaut 0. Si le nombre de 1 est impair alors le bit de parité est 1. On possède alors  $n + 1$  bits. Si un des  $n + 1$  bits se retrouve perdu, il est possible de récupérer la valeur de celui-ci grâce au

bit de parité. Pour généraliser, si on possède  $N$  blocs de données  $A_1, A_2, \dots, A_N$  de même taille alors, en notant  $\oplus$  l'opération XOR, on calcul la parité par  $A_1 \oplus A_2 \oplus \dots \oplus A_N = X$ . Si un bloc de données  $A_i$  se retrouve défaillant, on peut le retrouver par  $A_i = X \oplus A_1 \oplus A_2 \oplus \dots \oplus A_{i-1} \oplus A_{i+1} \oplus \dots \oplus A_N$ .

### 3 Les différents niveaux de RAID

#### 3.1 RAID 0, volume agrégé par bande

Le RAID 0, également nommé "entrelacement de disque" ("striping" en anglais), est une configuration RAID permettant d'augmenter les vitesses de lecture et d'écriture des données en faisant travailler au moins 2 disques durs en parallèle. C'est l'architecture RAID la plus primitive car les données sont distribuées par bloc sur l'ensemble des disques. Notons que le RAID 0 ne possède aucun mécanisme de redondance, il ne peut donc pas être considéré comme un RAID, d'où l'appellation de RAID 0.

Raid 0 de 3 disques

Disque 1	Disque 2	Disque 3
A	B	C
D	E	F
G	H	I
J	K	L

**Capacité** Afin de constituer un RAID 0, il faut au minimum posséder 2 disques. La capacité totale du RAID 0 ne peut pas dépasser la capacité du disque possédant le moins de stockages. En effet, pour cette méthode de RAID, les données sont réparties par bande d'une taille fixe appelé granulation sur chaque disque dur. En d'autres termes, tous les disques possèdent les mêmes données. Cela permet de paralléliser l'écriture et la lecture des données en sollicitant tous les disques durs.

**Fiabilité** Le principal avantage du RAID 0 en fait aussi son plus gros inconvénient. Le système d'agrégation par bande ne fonctionne plus si un seul des disques durs tombe en panne.

**Coût** Le coût pour un RAID 0 dépend du nombre de disque parallélisé. Étant donné que la vitesse d'écriture et de lecture est un facteur important sur un unique disque dur, le RAID 0 n'est pas une configuration négligée pour ceux qui souhaitent avoir de grandes performances.

**fonctionnement** On considère ici une parallélisation sur 2 disques durs avec un RAID 0 ayant une taille de granulation de 64 kb. Si l'on veut écrire un fichier de 200 kb sur l'architecture de stockage, le fichier va d'abord être découpé en fonction de la taille de granulation

$$\frac{200}{64} = 3 \times 64 + 8$$

Ainsi 4 bandes seront utilisées pour écrire le fichier. Les bandes seront ensuite réparties de la façon suivante :

Disque 1 : 1,2

Disque 2 : 3,4

L'écriture de fichier pourra alors se faire simultanément sur chacun des disques en un temps équivalent à l'écriture de 128 kb, ce qui revient presque à doubler le débit d'écriture. Il faut comprendre que ce n'est pas l'accumulation des vitesses d'écriture de chaque disque dur qui augmente les performances d'écritures et de lectures mais plutôt la répartition des tâches sur chaque disque dur.

Ainsi sur un RAID de  $n$  disques, chaque disque ne doit lire que  $\frac{1}{n}$  des données.

**Utilisateur** Ce type de RAID est parfait pour des applications demandant un traitement rapide des données avec des disques durs. Il est notamment utilisé pour accélérer les applications de type montage vidéo, traitement d'image ou encore modélisation graphique. Il peut être aussi utilisé afin d'augmenter la bande passante d'un serveur nécessitant des support de stockage capables de supporter de très grande quantité d'écritures.

**Avantages :** Performances en lectures et en écritures très importante.

**Inconvénients :** Aucune tolérance à la panne, il est donc impossible de récupérer les données en cas de panne d'un disque.

### 3.2 RAID 1, disques en miroir

Le RAID 1 est la forme du RAID la plus simple, elle consiste à l'utilisation de  $n$  disques redondant, chaque disque contenant à tout moment exactement les mêmes données d'où l'appellation miroir.

### Raid 1 de 3 disques

Disque 1	Disque 2	Disque 3
A	A	A
B	B	B
C	C	C
D	D	D

**Capacité** Comme le RAID 0, la capacité totale ne peut pas dépasser la capacité du disque possédant le moins de stockage. Il est donc conseillé d'utiliser des éléments de capacité identique pour améliorer la rentabilité de l'architecture.

**Fiabilité** Cette architecture est basée sur la sécurité. Elle offre un excellent niveau de protection des données et fonctionne jusqu'à ce que tous les disques tombent en panne simultanément, ce qui est très rare si l'on parallélise plusieurs disques durs.

**Coût** Le coût de stockage est directement proportionnel au nombre de miroir utilisé, bien que la capacité reste inchangée.

**Performances** Le débit de lecture des données n'est pas significativement changé. En effet, l'accès aux données se fait sur le disque le plus facile d'accès à un moment donné. En revanche, la vitesse d'écriture se retrouve diminuée puisque chaque disque doit se mettre à jour et la vitesse d'écriture du disque le plus lent limite les performances.

**Avantages :** Grande fiabilité, la perte de données est minimisée grâce à l'architecture miroir.

**Inconvénients :** Performances de lecture et d'écriture réduites et perte de capacité de stockage.

### 3.3 RAID 2,3 ET 4, des architectures peu utilisées

**RAID 2** est un RAID distribué et avec parité. La distribution des données est effectuée par "1 bit". Donc la taille du bloc est de 1 bit. La parité est calculée par le code Hamming (Code de correction d'erreurs). Les données de la parité sont stockées sur un disque indépendant. Ce type de RAID est considéré comme historique car il n'est pas utilisé ni commercialisé.

**RAID 3** est un RAID distribué par bytes. Les disques membres du Raid 3 sont synchronisés. La parité est calculée par la fonction XOR et stockée sur un disque séparé au niveau de la même position. Un nombre minimal de 3 disque est nécessaire pour construire un RAID 3.

**RAID 4** Comme le RAID 3, le RAID 4 est un RAID distribué, mais par blocs, avec une parité dédiée sur un disque séparé. Les disques membres du RAID 4 sont synchronisés. La parité est calculée par la fonction XOR et stockée sur un disque séparé au niveau de la même position. Un nombre minimal de 3 disque est nécessaire pour construire un RAID 4.

### 3.4 RAID 5, volume agrégé par bande à parité répartie,

Le RAID 5 combine la méthode du volume agrégé par bandes à une parité répartie. La parité se retrouve répartie circulairement sur les différents disques. Si on possède  $n + 1$  disques, chaque bande est constitué de  $n$  blocs de données et d'un seul bloc de parité. Le principale avantage de cette technique est que le bloc de parité permet de récupérer (calculer) les données potentiellement perdues sur un disque. Le RAID 5 ne supporte donc la perte que d'un seul disque. Si 2 disques tombent en pannes simultanément, il ne sera plus possible de calculer les données manquantes sur tous les blocs des disques.

Raid 5

Gauche, Synchrone

Disque 1	Disque 2	Disque 3
D1	D2	P
D4	P	D3
P	D5	D6

Raid 5

Gauche, Asynchrone

Disque 1	Disque 2	Disque 3
D1	D2	P
D3	P	D4
P	D5	D6

**Fonctionnement** Le Raid 5 utilise les techniques de distribution des blocs de données ainsi que la parité. Le flux de données est divisé par bloc dont la taille est le fameux paramètre taille du bloc (granulation). Sur chaque ligne, il y a un seul bloc de parité. Ce bloc est calculé par la fonction "XOR" des données des deux autres blocs. En cas de panne d'un disque, les données manquantes

peuvent être déduites par la fonction "XOR" à partir des blocs de données disponibles.

**Exemple** Si on possède 4 disques de tailles identiques A,B,C et D, le système va répartir les premiers blocs de la première bande sur les disques A,B et C comme pour le RAID 0. Sur le disque D, un bloc de parité va être mis en place calculer à partir de l'opération *OU* exclusif entre les autres disques ( $A \text{ xor } B \text{ xor } C$ ). Ensuite, pour la deuxième bande, les blocs de données vont être répartis sur les disques D,A et B et la parité ( $A \text{ xor } B \text{ xor } D$ ) sera placé sur le disque C. Et on continue le principe en permutant circulairement les disques pour chaque bande.

**Avantage** Le RAID 5 étant une dérivé du RAID 0, il possède les performances de lectures élevées. Mais il possède l'avantage de tolérer la perte d'un disque.

**Inconvénient** Le RAID 5 demande de calculer la parité pour écrire des données; la vitesse d'écritures est donc ralenti. De plus les blocs de parité prennent la place de données qui pourraient être directement stocké, au total nous perdons l'équivalent de 1 disque pour n disques de même capacité de stockage utilisés. Notons aussi que si l'on perd un disque, le temps de reconstruction peut s'avérer très long si la capacité de stockage des disques est importante, les disques étant beaucoup sollicité sur cette opération, il se peut qu'une deuxième panne arrive et que toutes les données soient perdues.

### 3.5 Raid 6

Le Raid 6 est un Raid à double parité P et Q. La parité P est calculée par la fonction "XOR", tout comme le RAID 5. La parité Q est calculée généralement par le code Reed-Solomon ou bien une de ses variantes. Le code Reed-Solomon dépend de l'ordre d'arrivé de données à coder. Ce qui génère quatre variétés selon la rotation de la parité (gauche ou droite) et selon l'ordre de calcul des parités P et Q. L'objectif de la double parité est d'aider le RAID 6 à résister à une double panne de deux disques durs simultanément. Par contre, cette double parité affecte les performances en écritures. Une étude sur le RAID 6 a constaté une chute de performance de l'ordre de 30% par rapport au RAID 5.

Raid 6, gauche, P en tête

Disque 1	Disque 2	Disque 3	Disque 4
Q	Data	Data	P
Data	Data	P	Q
Data	P	Q	Data
P	Q	Data	Data
Q	Data	Data	P
Data	Data	P	Q
Data	P	Q	Data
P	Q	Data	Data

Raid 6, gauche, P en queue (Q en tête)

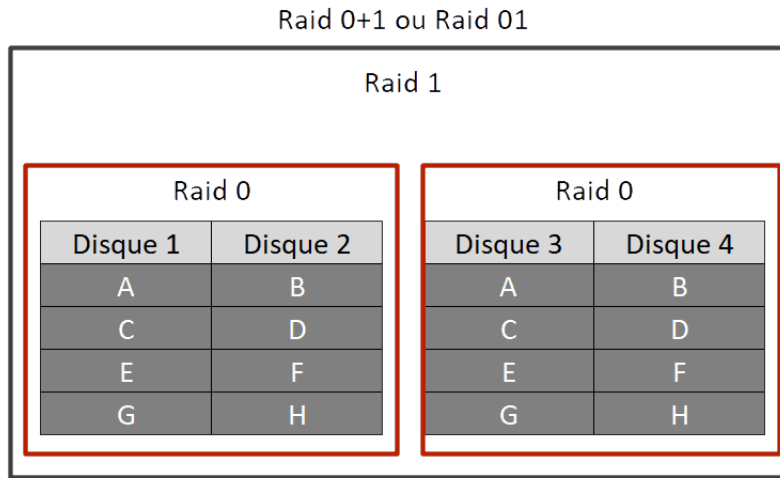
Disque 1	Disque 2	Disque 3	Disque 4
P	Data	Data	Q
Data	Data	Q	P
Data	Q	P	Data
Q	P	Data	Data
P	Data	Data	Q
Data	Data	Q	P
Data	Q	P	Data
Q	P	Data	Data

## 4 Combinaison de RAID

Chaque système RAID présente des avantages mais aussi des défauts. Il est possible de palier à ces défauts en associant deux systèmes RAID différents à plusieurs disques ou au contraire de profiter de leurs avantages. Le système le plus utilisé dans les combinaisons est le RAID 0. S'il est associé avec un système RAID de niveau 1, 3 ou 5, le système obtenu possède alors un niveau de performance très élevé ainsi qu'une tolérance à la panne. On appelle les regroupements de plusieurs disques dans un même RAID une grappe.

### 4.1 RAID 01

Le RAID 01 est une combinaison de deux types de RAID : RAID 0 et RAID 1. Ce RAID 01 est basé sur les techniques du miroir et de distribution des blocs de données. Quatre disques, au moins sont nécessaires. D'abord, deux disques seront utilisés pour créer le RAID 0. Ensuite, celui-ci sera dupliqué (ou miroité) sur une autre grappe de deux disques.



**Capacité** La capacité de la grappe du RAID 01 est déterminée par la capacité du plus petit disque multiplié par le nombre de disques de la grappe de base en RAID 0.

**Performances** La vitesse de lecture du RAID 01 de N disques est N fois plus rapide qu'un seul disque. Ce RAID 01 regroupe les avantages des performances du RAID 0 et la sécurité du RAID 1. Ce type de RAID est considéré comme le plus coûteux car la moitié de la capacité des disques est utilisée pour la redondance. Dans le cas d'une panne, le temps de reconstitution des données est élevé. En effet, il faut rétablir les données de chaque disque. C'est pourquoi le RAID 01 n'est que peu utilisé.

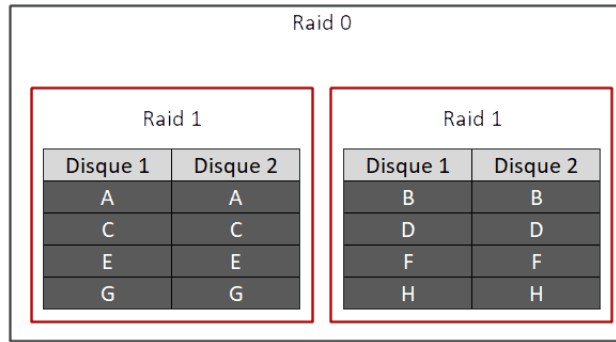
## 4.2 RAID 10

Le RAID 10 associe une architecture de RAID 1 avec celle d'un RAID 0. On regroupe d'abord les disques deux par deux pour créer  $\frac{n}{2}$  système RAID 1, puis on configure les paires de disque en RAID 0. En cas de panne d'un disque, il n'y a pas de problème. En effet, il y a juste une paire de disque qui fonctionne en mode dégradé. Si un deuxième disque tombe en panne, mis à part le cas où il s'agit du second disque de la même paire, il n'y a toujours pas de problème. La probabilité d'avoir les deux disques de la même paire défectueux est très faible. Dans ce cas toutes les données seraient perdues comme pour un simple RAID 0. Dans le cas d'une panne, il n'y a alors qu'un seul disque à reconstituer. Ce système offre les bonnes performances du RAID 0 tout en offrant la tolérance aux pannes du RAID 1. Ce mode de RAID a toujours le même inconvénient que le RAID 1, c'est à dire qu'il est coûteux en capacité de stockage à cause du mirroring. Il nécessite au minimum 4 disques durs : le RAID 0 nécessite au moins 2 disques qui dans ce cas sont les 2 paires de disques du RAID 1. Etant



donné la baisse actuelle des prix des disques, cette solution est de plus en plus envisagée.

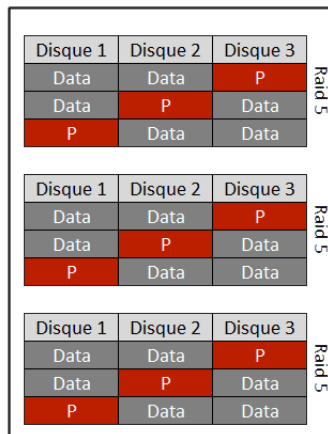
Raid 1+0 ou Raid 10



### 4.3 RAID 50

Le RAID 50 est une combinaison des RAID 5 et 0. Les données sont distribuées par blocs de RAID 0 sur plusieurs grappes de RAID 5. Le RAID 50 offre un bon compromis entre la performance, la capacité disponible et la sécurité. Il est réputé d'être parmi les RAID les plus solides. Néanmoins, il a également ses inconvénients.

Raid 0



**Capacité** Le RAID 50 a besoin d'au minimum 6 disques. Si on note  $G$  le nombre de grappe (contenant au minimum 3 disques) et  $N$  le nombre totale de disque, la capacité disponible vaut alors  $C_T = (N - G) \times C$ , si les disques possèdent tous la même capacité de stockage  $C$ . À comparer avec la capacité d'un RAID 10 qui perd 50% de sa capacité de stockage.

**Fiabilité** avec un RAID 5, la probabilité de panne de deux disques simultanément augmente avec le nombre des disques. Avec un RAID 50, la sécurité augmente avec le nombre de disques, à la condition que deux disques de la même grappe ne soient en panne simultanément.

**Performance** Généralement le RAID 50 fonctionne moins bien que le RAID 10, mais beaucoup mieux qu'un Raid 5. La vitesse maximale en lecture et écriture est donnée par  $V_M = N \times (G - 1) \times V$  où  $V$  est la vitesse d'écriture\lecture du disque le moins performant dans l'architecture.