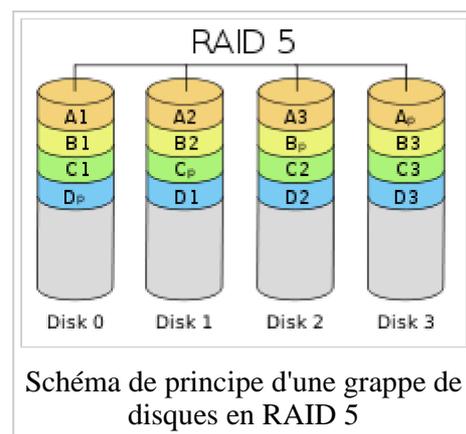


# RAID (informatique)

En informatique, le mot **RAID** désigne les techniques permettant de répartir des données sur plusieurs disques durs afin d'améliorer soit la tolérance aux pannes, soit la sécurité, soit les performances de l'ensemble, ou une répartition de tout cela.

L'acronyme **RAID** a été défini en 1987 par l'Université de Berkeley, dans un article nommé *A Case for Redundant Arrays of Inexpensive Disks (RAID)*<sup>1</sup>, soit « chaîne redondante de disques peu onéreux ». Aujourd'hui, le mot est devenu l'acronyme de **Redundant Array of Independent (or inexpensive) Disks**, ce qui signifie « chaîne redondante de disques indépendants ». Le coût au mégaoctet des disques durs ayant diminué d'un facteur 1 300 000 en 29 ans, aujourd'hui le RAID est choisi pour d'autres raisons que le coût de l'espace de stockage<sup>2</sup>.



## Sommaire

- 1 Historique
- 2 Description et concepts
  - 2.1 Comparaison RAID/SLED
  - 2.2 Parité et redondance
- 3 Les différents types de systèmes RAID
  - 3.1 Le RAID logiciel
    - 3.1.1 Avantages
    - 3.1.2 Inconvénients
    - 3.1.3 Diverses implémentations
  - 3.2 Le RAID pseudo-matériel
    - 3.2.1 Avantages
    - 3.2.2 Inconvénients
  - 3.3 Le RAID matériel
    - 3.3.1 Avantages
    - 3.3.2 Inconvénients
- 4 Les différents niveaux de RAID
  - 4.1 Les niveaux standard
    - 4.1.1 RAID 0 : volume agrégé par bandes
    - 4.1.2 RAID 1 : Disques en miroir
    - 4.1.3 RAID 5 : volume agrégé par bandes à parité répartie
  - 4.2 Les niveaux de RAID peu courants
    - 4.2.1 NRAID (ou JBOD - Just a Bunch Of Disks) : concaténation de disques
    - 4.2.2 RAID 2 : volume agrégé par bandes à parité
    - 4.2.3 RAID3 et RAID4
    - 4.2.4 RAID 6
    - 4.2.5 RAIDn
    - 4.2.6 RAID DP
  - 4.3 Les niveaux de RAID combinés
    - 4.3.1 le RAID 01 (ou RAID 0+1)
    - 4.3.2 RAID 10 (ou RAID 1+0)
    - 4.3.3 RAID 05
    - 4.3.4 RAID 15
    - 4.3.5 RAID 50
    - 4.3.6 RAID 51
  - 4.4 Les niveaux de RAID spéciaux
    - 4.4.1 RAID 5E

- 4.4.2 RAID 5EE
  - 4.4.3 RAID 5DP
  - 4.4.4 RAID TP
  - 4.4.5 Double parité
  - 4.4.6 RAID 1.5
  - 4.4.7 RAID 7, RAID ADG
  - 4.4.8 RAID S
  - 4.4.9 Matrix RAID
  - 4.4.10 IBM ServeRAID 1E
  - 4.4.11 Sun RAID-Z
  - 4.4.12 X-RAID (NetGear)
- 5 Fiabilisation et réglages d'une installation RAID
    - 5.1 Choix du contrôleur
    - 5.2 Redondance de contrôleurs
    - 5.3 Enfichage à chaud (hotplug/hotswap)
    - 5.4 Disques de rechange (spare/hotspare)
    - 5.5 Cache
    - 5.6 Granularité
  - 6 Les possibilités et les limites du RAID
    - 6.1 Ce que peut faire le RAID
    - 6.2 Ce que ne peut pas faire le RAID
    - 6.3 Comment le RAID peut vous trahir
  - 7 Références
  - 8 Voir aussi
    - 8.1 Liens externes
  - 9 Développeur de Raid Hardware

## Historique

En 1978, un employé d'IBM, Norman Ken Ouchi, déposa un brevet<sup>3</sup> concernant un « Système de récupération de données stockées dans une unité de stockage défectueuse », et dont la description était ce que deviendrait plus tard le RAID 5. Ce brevet fait également mention du miroitage de disque (qui sera appelé plus tard RAID 1), ainsi que de la protection avec une parité dédiée (qui sera appelé plus tard RAID 3 et 4).

La technologie RAID a été élaborée par un groupe de chercheurs de l'Université de Berkeley (Californie) en 1987. Ces derniers étudièrent la possibilité de faire reconnaître deux disques durs ou plus comme une seule unité par le système. Ils obtinrent pour résultat un système de stockage aux performances bien meilleures que celles des systèmes à disque dur unique, mais doté d'une très mauvaise fiabilité. Les chercheurs s'orientèrent alors vers des architectures redondantes, afin d'améliorer la tolérance aux pannes du système de stockage.

En 1988, les différents RAID, de type 1 à 5, étaient formellement définis par David Patterson, Garth Gibson et Randy Katz dans la publication intitulée « *A Case for Redundant Arrays of Inexpensive Disks (RAID)* »<sup>4</sup>. Cet article introduisait le terme « RAID », dont l'industrie du disque s'est immédiatement emparée, dont elle proposait cinq niveaux différents, en les comparant au « SLED », chacun d'eux ayant ses avantages et ses inconvénients.

## Description et concepts

### Comparaison RAID/SLED

---

Depuis sa création, la particularité principale de l'architecture RAID est sa capacité à combiner de nombreux périphériques de stockage bon marché et d'une technologie courante dans une matrice unique, de sorte que ce groupement offre une capacité, une fiabilité et/ou des performances accrues, ce pour un coût largement inférieur à un périphérique de stockage unique équivalent exploitant des technologies de pointe. L'architecture RAID s'oppose donc à l'architecture SLED (*Single Large Expensive Disk*), qui est fondée sur l'utilisation d'un seul et même disque dur de

grande capacité, donc d'un prix élevé, car celui-ci doit non seulement pouvoir stocker beaucoup d'informations, mais il doit de plus être d'excellente qualité pour garantir au mieux la pérennité et l'accessibilité de son contenu.

En effet, dans une architecture de type SLED, la bonne conservation des données est dépendante de la moindre défaillance du disque dur. Lorsqu'une panne survient, non seulement le système est inexploitable le temps du remplacement du matériel défectueux, mais la seule manière de récupérer les données est de procéder à une restauration de la dernière sauvegarde, ce qui peut prendre plusieurs heures durant lesquelles le système est toujours inutilisable.

Si un tel temps d'inactivité est acceptable pour l'ordinateur d'un particulier, il est en revanche rédhibitoire pour le système informatique d'une entreprise, pour qui une telle panne peut avoir des conséquences non négligeables sur sa santé financière. L'utilisation d'une architecture RAID, du moins dans la plupart de ses niveaux fonctionnels, permet justement d'apporter une réponse à ces besoins car non seulement la défaillance d'un des disques de la grappe ne gêne pas le fonctionnement des autres disques, ce qui permet au système de continuer de fonctionner, mais de surcroît, une fois le disque en panne échangé, son contenu est reconstruit à partir des autres disques pendant le fonctionnement normal du système. Ainsi, l'activité de l'entreprise continue de façon ininterrompue et transparente pendant toute la durée de l'incident.

Le RAID, suivant ses niveaux fonctionnels, s'il donne des temps de réponse identiques à ceux des disques s'ils étaient utilisés individuellement, offre des débits particulièrement soutenus, même en utilisant des disques durs bons marchés et de performances moyennes, tout en garantissant une bien meilleure fiabilité (sauf pour le RAID 0 qui lui la **réduit** d'autant que le nombre de disques). Dans de telles situations, les architectures RAID se révèlent donc idéales, tant du point de vue de leurs performances que de leur fiabilité. Dans tous les cas, le RAID reste complètement transparent à l'utilisateur qui, quel que soit le nombre de disques physiques utilisés pour construire le RAID, ne verra jamais qu'un seul grand volume logique, auquel il accédera de façon tout à fait habituelle.

## Parité et redondance

---

Le miroitage s'avère être une solution onéreuse, puisqu'il est nécessaire d'acquérir les périphériques de stockage en plusieurs exemplaires. Aussi, partant du principe que plusieurs unités de stockage ont une faible probabilité de tomber en panne simultanément, d'autres systèmes ont été imaginés, dont ceux permettant de régénérer les données manquantes à partir des données restant accessibles et d'une ou plusieurs données supplémentaires, dites de redondance.

Le système de redondance le plus simple et le plus largement utilisé est le calcul de parité. Ce système repose sur l'opération logique XOR (OU exclusif) et consiste à déterminer si sur  $n$  bits de données considérés, le nombre de bits à l'état 1 est pair ou impair. 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 vaut 1. Lorsque l'un des  $n + 1$  bits de données ainsi formés devient indisponible, il est alors possible de régénérer le bit manquant en appliquant à nouveau la même méthode sur les  $n$  éléments restants. Cette technique est utilisée dans les systèmes RAID 5.

Il existe des systèmes de redondance plus complexes et capables de générer plusieurs éléments de redondance afin de supporter l'absence de plusieurs éléments. Le RAID 6 utilise par exemple une technique de calcul de parité fondée sur des polynômes.

## Les différents types de systèmes RAID

Le système RAID est :

- soit un système de redondance qui donne au stockage des données une certaine tolérance aux pannes matérielles (ex : RAID1)□;
- soit un système de répartition qui améliore ses performances (ex : RAID0)□;
- soit les deux à la fois (ex : RAID5).

Le système RAID est donc capable de gérer d'une manière ou d'une autre la répartition et la cohérence de ces données. Ce système de contrôle peut être purement logiciel ou utiliser un matériel dédié.

## Le RAID logiciel

---

En RAID logiciel, le contrôle du RAID est intégralement assuré par une couche logicielle du système d'exploitation. Cette couche s'intercale entre la couche d'abstraction matérielle (pilote) et la couche du système de fichiers.

### Avantages

---

- C'est la méthode la moins onéreuse puisqu'elle ne demande aucun matériel supplémentaire.
- Cette méthode possède une grande souplesse d'administration (logicielle).
- Cette méthode présente l'avantage de la compatibilité entre toutes les machines équipées du même logiciel de RAID (c'est-à-dire du même système d'exploitation)

### Inconvénients

---

- L'inconvénient majeur réside dans le fait que cette méthode repose sur la couche d'abstraction matérielle des périphériques qui composent le volume RAID. Pour diverses raisons, cette couche peut être imparfaite et manquer de certaines fonctions importantes comme, par exemple, la détection et le diagnostic des défauts matériels et/ou la prise en charge du remplacement à chaud des unités de stockage.
- La gestion du RAID monopolise des ressources systèmes (légèrement le processeur et surtout le bus système) qui pourraient être employées à d'autres fins. La baisse de performances due à la gestion logicielle du raid est particulièrement sensible dans des configurations où le système doit transférer plusieurs fois les mêmes données comme, par exemple, en RAID1, et, assez faible, dans des configurations sans redondance : exemple, le RAID 0.
- L'utilisation du RAID sur le disque système n'est pas toujours possible.

### Diverses implémentations

---

La plupart des systèmes d'exploitation grand public permettent déjà de mettre en œuvre le RAID logiciel, qu'il s'agisse de Microsoft Windows, d'une distribution Linux quelconque ou de Mac OS X.

- Microsoft Windows XP gère logiciellement le RAID 0, 1 et peut gérer le RAID 5 moyennant une petite adaptation (Explications sur le site d'Ere numérique [1])
- Microsoft Windows 2003 Server gère logiciellement le RAID 0, 1, et 5.
- Mac OS X gère logiciellement le RAID 0, 1 et la concaténation.
- Le noyau Linux (>=2.6) gère logiciellement le RAID 0, 1, 4, 5, 6, et 10 ainsi que les combinaisons de ces modes.

Les RAIDs logiciels de Microsoft Windows et de Linux sont incompatibles entre eux.

## Le RAID pseudo-matériel

---

L'extrême majorité des contrôleurs RAID bon marché intégrés à de nombreuses cartes mères récentes en 2004/2005 gèrent souvent le RAID 0 et 1 sur des disques IDE ou SATA. Malgré le discours marketing qui tend systématiquement à induire en erreur sur ce point, il ne s'agit pas de RAID matériel à proprement parler mais plutôt d'un contrôleur de disque doté de quelques fonctions avancées.

D'un point de vue strictement matériel, cette solution hybride n'est pas différente d'un RAID logiciel. Elle diffère cependant sur l'emplacement des routines logicielles de gestion du RAID.

### Avantages

---

L'intérêt principal de ce type de RAID est d'apporter une solution au troisième problème du RAID logiciel, à savoir qu'il ne peut pas toujours servir à héberger les fichiers du système d'exploitation puisque c'est justement ce dernier qui permet d'y accéder.

Dans ce type de RAID, la présence d'un BIOS intégrant les routines logicielles basiques de gestion du RAID permet de charger en mémoire les fichiers essentiels du système d'exploitation (le noyau et les pilotes essentiels).

Puis, le pilote du contrôleur intègre les mêmes routines logicielles de gestion du RAID et fournit alors aux couches supérieures de l'OS non pas un accès aux périphériques mais un accès au volume RAID qu'il émule.

## Inconvénients

---

En dehors de cet avantage important, ce type de RAID cumule les défauts des deux autres approches :

- Les limitations de performances sont les mêmes que pour le raid logiciel car il s'agit effectivement d'un RAID logiciel camouflé.
- Un problème important posé par ces contrôleurs hybrides est leur piètre gestion des défauts matériels et leurs fonctionnalités BIOS généralement limitées.
- L'interopérabilité est très mauvaise surtout si l'on considère qu'il s'agit généralement de matériel intégré aux cartes mères des ordinateurs. Pire, le changement de carte-mère (voire simplement de version de bios), si la nouvelle utilise des jeux de puces différents, peut imposer de reconstruire le RAID entièrement. De manière générale, une reconstruction est possible si l'on reste dans des contrôleurs RAID de même marque mais de modèles différents, mais il n'existe pas de règle définie de compatibilité.
- La fiabilité annoncée de ces dispositifs est assez controversée. [citation nécessaire]

## Le RAID matériel

---

Dans le cas du RAID matériel, une carte ou un composant est dédié à la gestion des opérations. Le contrôleur RAID peut être interne à l'unité centrale (carte d'extension) ou déporté dans une baie de stockage.

Un contrôleur raid est en général doté d'un processeur spécifique, de mémoire, éventuellement d'une batterie de secours, et est capable de gérer tous les aspects du système de stockage RAID grâce au microcode embarqué (*firmware*).

Du point de vue du système d'exploitation, le contrôleur RAID matériel offre une virtualisation complète du système de stockage. Le système d'exploitation considère chaque volume RAID comme un disque et n'a pas connaissance de ses constituants physiques.

## Avantages

---

- Les contrôleurs RAID matériels permettent la détection des défauts, le remplacement à chaud des unités défectueuses et offrent la possibilité de reconstruire de manière transparente les disques défaillants. Mais les systèmes d'exploitation évolués permettent également cela si le matériel le permet.
- La charge système (principalement l'occupation du bus) est allégée. (surtout dans des configurations avec beaucoup de disques et une forte redondance)
- Les vérifications de cohérence, les diagnostics et les maintenances sont effectués en arrière plan par le contrôleur sans solliciter de ressources système.

## Inconvénients

---

- Les contrôleurs RAID matériels utilisent chacun leur propre système pour gérer les unités de stockage. En conséquence, au contraire d'un RAID logiciel, des disques transférés d'un système à un autre ne pourront pas être récupérés si le contrôleur RAID n'est pas exactement le même (*firmware* compris). Il est donc conseillé de posséder une deuxième carte en cas de panne de la première.

- Les cartes d'entrée de gamme possèdent des processeurs de puissance bien inférieure à celle des ordinateurs actuels. On peut donc avoir de bien moins bonnes performances qu'un RAID logiciel.
- Le contrôleur RAID est lui-même un composant matériel complexe, qui peut tomber en panne, et son logiciel (*firmware*) peut contenir des bugs, ce qui constitue un risque de panne supplémentaire (un nouveau *single-point-of-failure*).
- Les différents fabricants de contrôleurs RAID fournissent des outils de gestion logicielle très différents les uns des autres (et de qualité parfois inégale). À l'opposé, les outils de gestion du RAID logiciel fournis avec un système d'exploitation sont généralement bien intégrés dans ce système.
- La durée du support d'un contrôleur RAID par son constructeur (correction de bugs dans le *firmware*, par exemple), parfois liée à l'arrivée de nouveaux produits rendant les anciens obsolètes, peut être moins longue ou plus volatile que le support du RAID logiciel par le fournisseur du système d'exploitation. Le constructeur peut même disparaître (ce qui est assez rare parmi les fabricants de systèmes d'exploitation).
- Une moindre souplesse par rapport au RAID logiciel, qui dispose d'une couche d'abstraction permettant de gérer du RAID au-dessus de tout types de périphériques blocs supportés par le système d'exploitation, locaux ou distants (ATA, SCSI, ATA over Ethernet, iSCSI... et toutes les combinaisons possibles entre eux). Les contrôleurs RAID sont spécialisés pour un seul type de périphérique bloc.

## Les différents niveaux de RAID

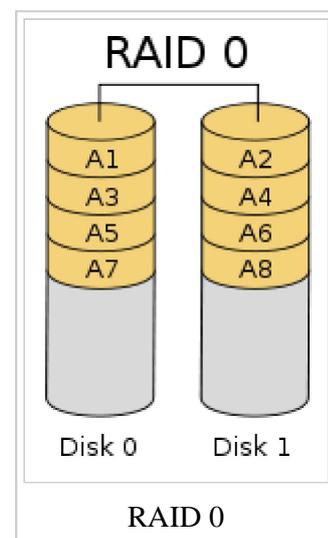
### Les niveaux standard

Les différents types d'architecture RAID sont numérotés à partir de 0 et peuvent se combiner entre eux (on parlera alors de RAID 0+1, 1+0, etc.).

### RAID 0 : volume agrégé par bandes

Le RAID 0, également connu sous le nom d'« entrelacement de disques » ou de « volume agrégé par bandes » (*striping* en anglais) est une configuration RAID permettant d'augmenter significativement les performances de la grappe en faisant travailler  $n$  disques durs en parallèle (avec  $n \geq 2$ ).

- Capacité :  
La capacité totale est égale à celle du plus petit élément de la grappe multiplié par le nombre d'éléments présent dans la grappe. L'espace excédentaire des autres éléments de la grappe restera inutilisé. Il est donc conseillé d'utiliser des éléments identiques. [réf. nécessaire]
- Fiabilité :  
Le défaut de cette solution est que la perte d'un seul disque entraîne la perte de toutes ses données.
- Coût :  
Le RAID 0 n'apportant pas de redondance, tout l'espace disque disponible est utile.



Dans cette configuration, les données sont réparties par bandes (*stripes* en anglais) d'une taille fixe. Cette taille est appelée granularité (voir plus loin la section granularité).

Exemple : avec un RAID 0 ayant une bande de 64 kio et composé de deux disques (disque *Disk 0* et disque *Disk 1*), si l'on veut écrire un fichier A de 500 kio, le fichier sera découpé en 8 bandes (car  $7 < \frac{500}{64} \leq 8$ ), appelons-les 1, 2, 3, 4, 5, 6, 7 et 8, elles seront réparties sur l'ensemble des disques de la façon suivante :

Disk 0 : 1, 3, 5, 7  
Disk 1 : 2, 4, 6, 8

Ainsi l'écriture du fichier pourra être effectuée simultanément sur chacun des disques en un temps équivalent à

l'écriture de 256 kio.

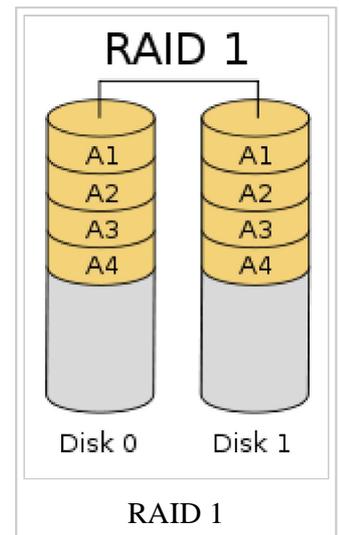
Ainsi, sur un RAID 0 de  $n$  disques (avec  $n \geq 2$ ), chaque disque ne doit lire et écrire que  $\frac{1}{n}$  des données, ce qui a pour effet de diminuer les temps d'accès (lecture et écriture) aux données; les disques se partageant le travail, les traitements se trouvent accélérés.

Ce type de RAID est parfait pour des applications requérant un traitement rapide d'une grande quantité de données. Mais cette architecture n'assure en rien la sécurité des données ; en effet, si l'un des disques tombe en panne, la totalité des données du RAID est perdue.

### RAID 1 : Disques en miroir

Le RAID 1 consiste en l'utilisation de  $n$  disques redondants (avec  $n \geq 2$ ), chaque disque de la grappe contenant à tout moment exactement les mêmes données, d'où l'utilisation du mot « miroir » (*mirroring* en anglais).

- Capacité :  
La capacité totale est égale à celle du plus petit élément de la grappe. L'espace excédentaire des autres éléments de la grappe restera inutilisé. Il est donc conseillé d'utiliser des éléments identiques.
- Fiabilité :  
Cette solution offre un excellent niveau de protection des données. Elle accepte une défaillance de  $n - 1$  éléments.
- Coût :  
Les coûts de stockage sont élevés et directement proportionnels au nombre de miroirs utilisés alors que la capacité totale reste inchangée. Plus le nombre de miroirs est élevé, et plus la sécurité augmente, mais plus son coût devient prohibitif.

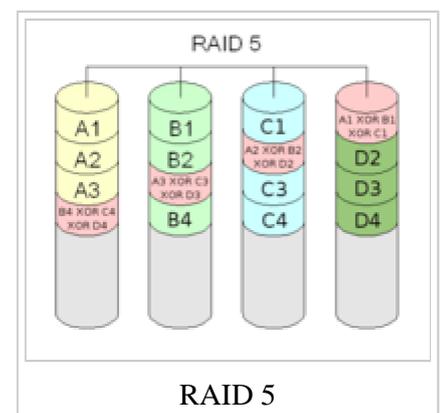


Les accès en lecture du système d'exploitation se font sur le disque le plus facilement accessible à ce moment-là. Les écritures sur la grappe se font de manière simultanée sur tous les disques, de façon à ce que n'importe quel disque soit interchangeable à tout moment.

Lors de la défaillance de l'un des disques, le contrôleur RAID désactive, de manière transparente pour l'accès aux données, le disque incriminé. Une fois le disque défectueux remplacé, le contrôleur RAID reconstitue, soit automatiquement, soit sur intervention manuelle, le miroir. Une fois la synchronisation effectuée, le RAID retrouve son niveau initial de redondance.

### RAID 5 : volume agrégé par bandes à parité répartie

Le RAID 5 combine la méthode du volume agrégé par bandes (*striping*) à une parité répartie. Il s'agit là d'un ensemble à redondance  $N + 1$ . La parité, qui est incluse avec chaque écriture se retrouve répartie circulairement sur les différents disques. Chaque bande est donc constituée de  $N$  blocs de données et d'un bloc de parité. Ainsi, en cas de défaillance de l'un des disques de la grappe, pour chaque bande il manquera soit un bloc de données soit le bloc de parité. Si c'est le bloc de parité, ce n'est pas grave, car aucune donnée ne manque. Si c'est un bloc de données, on peut calculer son contenu à partir des  $N - 1$  autres blocs de données et du bloc de parité. L'intégrité des données de chaque bande est préservée. Donc non seulement la grappe est toujours en état de fonctionner, mais il est de plus possible de reconstruire le disque une fois échangé à partir des données et des informations de parité contenues sur les autres disques.



On voit donc que le RAID 5 ne supporte la perte que d'un seul disque à la fois. Ce qui devient un problème depuis

que les disques qui composent une grappe sont de plus en plus gros (1 To et plus). Le temps de reconstruction de la parité en cas de disque défaillant est allongé. Il est généralement de 2 h pour des disques de 300 Go contre une dizaine d'heure pour 1 To. Pour limiter le risque il est courant de dédier un disque dit de *spare*. En régime normal il est inutilisé. En cas de panne d'un disque il prendra automatiquement la place du disque défaillant. Cela nécessite une phase communément appelée "recalcul de parité". Elle consiste pour chaque bande à recréer sur le nouveau disque le bloc manquant (données ou parité).

Bien sûr pendant tout le temps du recalcul de la parité le disque est disponible normalement pour l'ordinateur qui se trouve juste un peu ralenti.

Exemple pratique : Considérons quatre disques durs A, B, C et D, de tailles identiques. Le système va enregistrer les premiers blocs en les répartissant sur les disques A, B et C comme en mode RAID 0 (*striping*) et, sur le disque D, le résultat de l'opération OU exclusif entre les autres disques (ici A xor B xor C). Ensuite le système va enregistrer les blocs suivants en les répartissant sur les disques D, A et B, puis la parité (soit D xor A xor B) sur le disque C, et ainsi de suite en faisant permuter circulairement les disques, à chaque bloc. La parité se trouve alors répartie sur tous les disques.

En cas de défaillance d'un disque, les données qui s'y trouvaient pourront être reconstituées par l'opération xor. En effet, l'opération XOR ( $\oplus$ ) a la propriété suivante : si on considère  $N$  blocs de taille identique  $A_1, A_2 \dots A_N$  et si  $A_1 \oplus A_2 \oplus \dots \oplus A_N = X$  alors  $X \oplus A_2 \oplus \dots \oplus A_N = A_1$ , et de façon générale,  $A_1 \oplus \dots \oplus A_{k-1} \oplus X \oplus A_{k+1} \oplus \dots \oplus A_N = A_k$ . C'est-à-dire que n'importe quel bloc de données  $A_k$  perdu à cause d'un disque défaillant sur un RAID 5 de  $N + 1$  disques peut-être récupéré grâce au bloc  $X$  de données de contrôle.

On voit donc que si on veut écrire dans un bloc, il faut lire le bloc à modifier. Lire le bloc de parité de la bande. Écrire le bloc de données et le bloc de parité. L'opération xor permet heureusement de calculer la nouvelle parité sans avoir besoin de lire les  $N$  blocs de données de la bande. Augmenter le nombre de disque d'une grappe RAID 5 n'allonge donc pas le temps de lecture ou d'écriture. Cependant si plusieurs processus veulent écrire simultanément dans un ou plusieurs blocs de données d'une même bande la mise à jour du bloc de parité devient un point de blocage. Les processus concurrents sont suspendus à la libération du bloc de parité et de fait cela limite le débit d'écriture. Plus le nombre de disque d'une grappe RAID 5 augmente plus le temps de reconstruction d'un disque défaillant augmente. Puisque pour reconstituer le bloc manquant d'une bande il faut lire tous les autres blocs de la bande et donc tous les autres disques.

Ce système nécessite impérativement un minimum de trois disques durs. Ceux-ci doivent généralement être de même taille, mais un grand nombre de cartes RAID modernes autorisent des disques de tailles différentes.

La capacité de stockage utile réelle, pour un système de  $X$  disques de capacité  $c$  identiques est de  $(X - 1) \times c$ . En cas d'utilisation de disques de capacités différentes, le système utilisera dans la formule précédente la capacité minimale.

Ainsi par exemple, trois disques de 100 Go en RAID 5 offrent 200 Go utiles ; dix disques, 900 Go utiles.

Ce système allie sécurité (grâce à la parité) et bonne disponibilité (grâce à la répartition de la parité), même en cas de défaillance d'un des périphériques de stockage.

Il existe une variante : le « RAID 5 orthogonal » où chaque disque a son propre contrôleur. Toutes les autres fonctionnalités sont identiques.

On a souvent tendance à croire qu'un système RAID 5 est totalement fiable. Il est en effet généralement admis que la probabilité de défaillance simultanée de plusieurs disques est extrêmement faible — on parle évidemment d'une défaillance entraînant la perte de données définitive sur plusieurs disques et non d'une simple indisponibilité de plusieurs disques. Cela est vrai pour une défaillance générale d'une unité de disque. Cependant, cela est faux si l'on considère comme "défaillance" un seul secteur devenu illisible.

En effet, dans la pratique, il est très rare que toutes les données d'un volume soient lues régulièrement. Et quand bien même ce serait le cas, la cohérence de la parité n'est que très rarement vérifiée pour des raisons de performances. Il est donc probable que des défauts tels que des secteurs de parité illisibles ne soient pas détectés pendant une très

longue période. Lorsque l'un des disques devient réellement défectueux, la reconstruction nécessite de parcourir l'intégralité des disques restants. On peut alors découvrir des défauts qui étaient restés invisibles jusque-là.

Tout ceci pourrait ne pas être bien grave et occasionner la perte d'une quantité de données minime (un secteur de disque), cependant, l'extrême majorité des contrôleurs RAID est incapable de gérer les défaillances partielles : ils considèrent généralement qu'un disque contenant un secteur illisible est totalement défaillant. À ce moment-là, 2 disques sont considérés défaillants simultanément et le volume RAID 5 devient inutilisable. Il devient extrêmement difficile de récupérer les données.

Un système RAID 5 doit donc être vérifié et sauvegardé très périodiquement pour s'assurer que l'on ne risque pas de tomber sur ce genre de cas.

- Avantages :

- performances en lecture aussi élevées qu'en RAID 0 et sécurité accrue
  - surcoût minimal (capacité totale de  $n - 1$  disques sur un total de  $n$  disques)

- Inconvénients :

- pénalité en écriture du fait du calcul de la parité
  - minimum de 3 disques

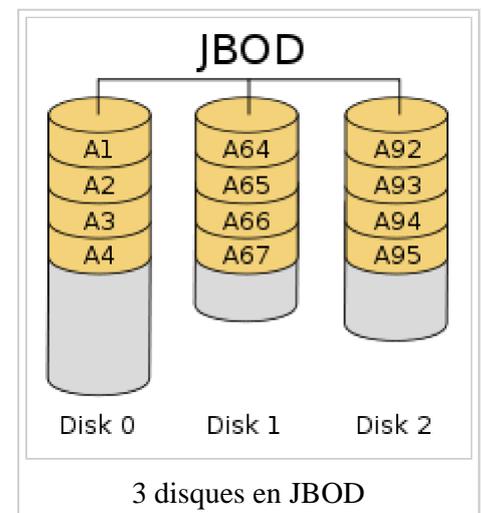
## Les niveaux de RAID peu courants

### NRAID (ou JBOD - Just a Bunch Of Disks) : concaténation de disques

*NRAID : Near/Non Redundant Array of Inexpensive/Independent Disk*

La concaténation de disques consiste à additionner les capacités de plusieurs disques durs en un volume logique d'une taille équivalente à la somme des tailles de disques durs. Cette méthode utilise une méthode d'écriture séquentielle : les données ne sont écrites sur le disque dur suivant que lorsqu'il ne reste plus de place sur le précédent.

Le NRAID n'est pas à proprement parler un RAID, et il ne permet d'ailleurs aucune redondance de données, mais il offre cependant une tolérance aux pannes supérieure au RAID 0. On le rencontre souvent sous le nom de JBOD (*Just a Bunch Of Disks*).



### RAID 2 : volume agrégé par bandes à parité

Le RAID 2 est aujourd'hui obsolète. Il combine la méthode du volume agrégé par bande (*striping* en anglais) à l'écriture d'un code de contrôle d'erreur par code de Hamming (code ECC) sur un disque dur distinct. Cette technologie offre un bon niveau de sécurité, mais de mauvaises performances.

### RAID3 et RAID4

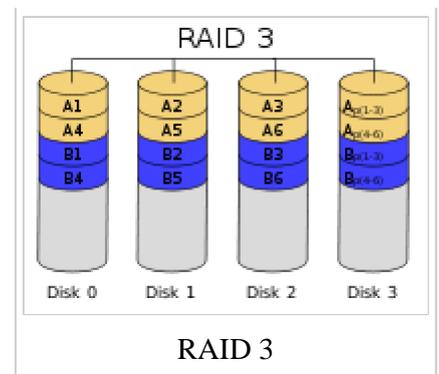
Le RAID3 et le RAID4 sont sensiblement semblables sauf que le premier travaille par octets et le second par blocs. Le RAID4 ne nécessite pas autant de synchronisme entre les disques. Le RAID3 tend donc à disparaître au profit du RAID4 qui offre des performances nettement supérieures.

Ces niveaux de RAID nécessitent une matrice de  $n$  disques (avec  $n \geq 3$ ). Les  $n - 1$  premiers disques contiennent les données tandis que le dernier disque stocke la parité (Voir section parité).

- Si le disque de parité tombe en panne, il est possible de reconstruire l'information de parité avec le contenu des autres disques de données.
- Si l'un des disques de données tombe en panne, il est possible de reconstruire l'information avec le contenu des disques de données restants et celui du disque de parité.

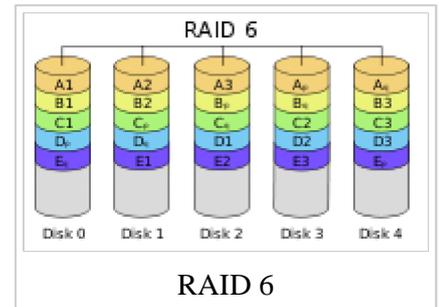
Il est important que le disque de parité soit de bonne qualité car il est à tout instant sollicité à l'écriture. Ce dernier point est une des limitations du RAID 3.

De même, si plus d'un disque vient à défaillir, il est impossible de remédier à la perte de données.



## RAID 6

Le RAID 6 est une évolution du RAID 5 qui accroît la sécurité en utilisant  $n$  informations redondantes au lieu d'une. Il peut donc résister à la défaillance de  $n$  disques. Les fondements mathématiques utilisés pour les informations de redondance du RAID 6 sont beaucoup plus complexes que pour le RAID 5 ; les implémentations de l'algorithme se limitent souvent à  $n = 2$  (soit la perte de 2 disques) de ce fait.



Des explications intéressantes mais un peu ardues sont disponibles dans la version anglaise de cet article<sup>5</sup> ainsi que dans *Les mathématiques du RAID 6*<sup>6</sup>.

Si la sécurité est plus grande, le coût en matériel est plus élevé et la vitesse est moindre. La puissance CPU nécessaire pour calculer les redondances et surtout pour reconstruire un volume défectueux est également nettement plus importante.

Les défauts majeurs sont :

- Les temps d'écriture sont longs à cause des calculs de redondance complexes.
- Le temps de reconstruction en cas de défaillance simultanée de 2 disques est extrêmement long.

Le RAID 6 était peu utilisé du fait de son surcoût. La récente envolée des capacités des disques ainsi que la vulgarisation de solution professionnelle à base de disque SATA a montré un intérêt nouveau dans l'utilisation du RAID 6, que ce soit par le biais de contrôleur Raid Hardware ou via du raid logiciel (Linux-2.6 intègre le RAID 6).

La capacité totale ( $C_t$ ), pour un système avec  $n$  disques, est de  $C_t = (n - 2) \times c$ . ( $c$  = capacité du plus petit des disques dur)

## RAIDn

Un système de stockage grevé de brevets appelé RAIDn<sup>7</sup> prétend permettre de dépasser largement les capacités du RAID 6.

## RAID DP

RAID DP (*Dual Parity*) ressemble au RAID6 à ceci près qu'en RAID DP les disques de parité sont fixes. Ce type de RAID est adopté en général sur les baies de stockage NAS.

## Les niveaux de RAID combinés

Fondamentalement, un niveau de RAID combiné est l'utilisation d'un concept de RAID classique sur des éléments constitutifs qui sont eux-mêmes le résultat d'un concept RAID classique. Le concept utilisé peut être le même ou

différent.

La syntaxe est encore un peu floue mais on peut généralement considérer que le premier chiffre indique le niveau de raid des "grappes" et que le second indique le niveau de raid global. Dans l'absolu rien n'empêche d'imaginer des RAID combinés à 3 étages ou plus mais cela reste pour l'instant plus du domaine de la théorie et de l'expérimentation.

Le nombre important (et croissant) de permutations possibles fait qu'il existe une multitude de raid combinés et nous n'en ferons pas l'inventaire. Nous pouvons cependant présenter les avantages et les faiblesses des plus courants.

Pour les calculs suivants, on utilise les variables suivantes :

- $G$  : nombre de disques par grappe ;
- $N$  : nombre de grappes ;
- $C$  : capacité d'un disque (tous les disques sont supposés identiques) ;
- $V$  : vitesse d'un disque.

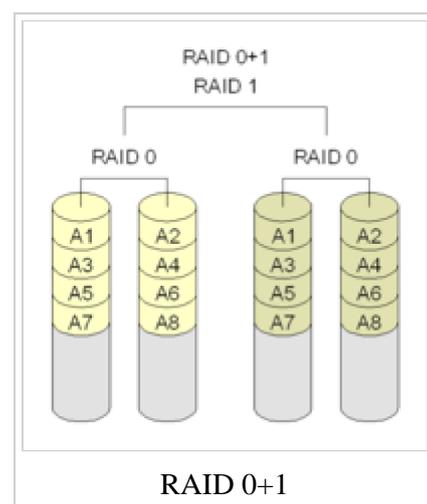
Les seuils de mise en défaut indiqués ci-dessous indiquent le nombre minimal de disques en panne pouvant entraîner une mise en défaut de l'ensemble du RAID (ie. en dessous de ce nombre de disques en panne le RAID ne peut pas être en défaut). En pratique il est possible qu'un RAID ayant plus que ce nombre de disques en panne fonctionne toujours mais il est recommandé de changer les disques défectueux le plus rapidement possible.

### le RAID 01 (ou RAID 0+1)

Il permet d'obtenir du *mirroring* rapide puisqu'il est basé sur des grappes en striping. Chaque grappe contenant au minimum 2 éléments, et un minimum de 2 grappes étant nécessaire, il faut au minimum 4 unités de stockage pour créer un volume RAID0+1.

La fiabilité est moyenne car un disque défectueux entraîne le défaut de toute la grappe qui le contient. Par ailleurs, cela allonge beaucoup le temps de reconstruction et dégrade les performances pendant la reconstruction. L'intérêt principal est que dans le cas d'un miroir à 3 grappes ou plus, le retrait volontaire d'une grappe entière permet d'avoir une sauvegarde "instantanée" sans perdre la redondance.

$$\begin{aligned} \text{Capacité totale : } C_t &= G \times C \\ \text{Vitesse maximale : } V_m &= G \times V \\ \text{Seuil de mise en défaut : } N &\text{ disques} \end{aligned}$$

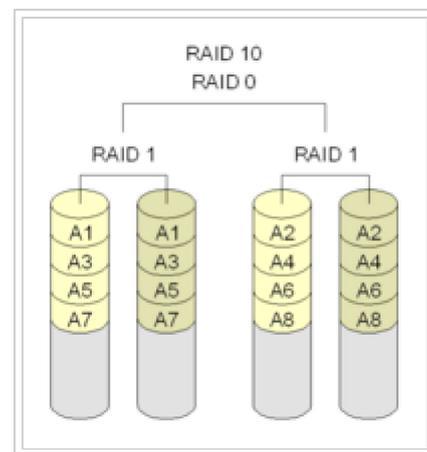


### RAID 10 (ou RAID 1+0)

Il permet d'obtenir un volume agrégé par bande fiable (puisque'il est basé sur des grappes répliquées). Chaque grappe contenant au minimum 2 éléments et un minimum de 2 grappes étant nécessaire, il faut au minimum 4 unités de stockage pour créer un volume RAID10.

Sa fiabilité est assez grande puisqu'il faut que tous les éléments d'une grappe soient défectueux pour entraîner un défaut global. La reconstruction est assez performante puisqu'elle ne mobilise que les disques d'une seule grappe et non la totalité.

$$\begin{aligned} \text{Capacité totale : } C_t &= N \times C \\ \text{Vitesse maximale : } V_m &= N \times V \\ \text{Seuil de mise en défaut : } G &\text{ disques} \end{aligned}$$



## RAID 05

RAID 10

Même principe que pour le raid 0+1 mais en employant du RAID5 pour la partie globale. Chaque grappe contenant au minimum 2 disques, et un minimum de 3 grappes étant nécessaire, il faut au minimum 6 unités de stockage pour créer un volume RAID05. Ce mode ne présente pas d'intérêt majeur par rapport à un RAID5 classique à  $N * G$  disques. Il est donc très peu utilisé.

$$\text{Capacité totale : } C_t = (N - 1) \times (G \times C)$$

$$\text{Vitesse maximale : } V_m = (N - 1) \times (G \times V) \text{ (cette formule néglige les temps de calcul de parité)}$$

Seuil de mise en défaut : 2 disques

## RAID 15

Il permet d'obtenir un volume agrégé par bandes avec redondance répartie très fiable (puisque'il est basé sur des grappes répliquées en miroir). Chaque grappe contenant au minimum 2 disques, et un minimum de 3 grappes étant nécessaire, il faut au minimum 6 unités de stockage pour créer un volume RAID15. Ce mode est très fiable puisqu'il faut que tous les disques de 2 grappes différentes cessent de fonctionner pour le mettre en défaut. Ce mode est cependant coûteux par rapport à la capacité obtenue.

$$\text{Capacité totale : } C_t = (N - 1) \times C$$

$$\text{Vitesse maximale : } V_m = (N - 1) \times V \text{ (cette formule néglige les temps de calcul de parité)}$$

Seuil de mise en défaut :  $2 \times G$  disques

## RAID 50

Il permet d'obtenir un volume agrégé par bandes basé sur du RAID 5. Chaque grappe contenant au minimum 3 disques, et un minimum de 2 grappes étant nécessaire, il faut au minimum 6 unités de stockage pour créer un volume RAID 50. Un des meilleurs compromis lorsque l'on cherche la rapidité sans pour autant vouloir trop dégrader la fiabilité. En effet, l'agrégat par bande (fragile) repose sur des grappes redondantes. Il suffit cependant que 2 disques d'une même grappe tombent en panne pour le mettre en défaut.

$$\text{Capacité totale : } C_t = N \times (G - 1) \times C$$

$$\text{Vitesse maximale : } V_m = N \times (G - 1) \times V \text{ (cette formule néglige les temps de calcul de parité)}$$

Seuil de mise en défaut : 2 disques

## RAID 51

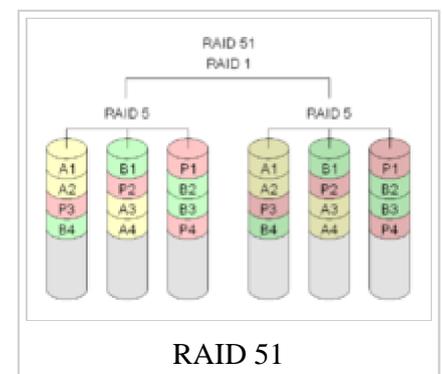
Il permet d'obtenir un volume répliqué basé sur des grappes en RAID5. Chaque grappe contenant au minimum 3 disques, et un minimum de 2 grappes étant nécessaire, il faut au minimum 6 unités de stockage pour créer un volume RAID51. C'est un mode coûteux (faible capacité au regard du nombre de disques).

$$\text{Capacité totale : } C_t = (G - 1) \times C$$

$$\text{Vitesse maximale en écriture : } V_{Mw} = (G - 1) \times V \text{ (cette formule néglige les temps de calcul de parité)}$$

$$\text{Vitesse maximale en lecture : } V_{Mr} = N \times (G - 1) \times V \text{ (cette formule théorique suppose une optimisation maximale qui n'est jamais atteinte)}$$

Seuil de mise en défaut :  $2 \times (N - 1)$  disques



## Les niveaux de RAID spéciaux

## RAID 5E

---

Le RAID 5E (E pour *Enhanced*, "amélioré") est une variante du RAID 5 dans laquelle des bandes de rechange sont réservées. Ces bandes sont réparties sur les disques de manière circulaire, comme pour la parité.

## RAID 5EE

---

Le RAID 5EE (E pour *Enhanced*, "amélioré") est une variante du RAID 5E qui offre de meilleurs temps de reconstruction. Contrairement au RAID 5E, l'espace de rechange réservé au sein de la grappe est distribué à travers tous les disques physiques.

## RAID 5DP

---

RAID DP (*Dual Parity*) ressemble au RAID 6 à ceci près qu'en RAID DP les disques de parité sont fixes. Ce type de RAID est adopté en général sur les baies de stockage NAS.

## RAID TP

---

Le RAID TP pour *Triple Parity RAID technology* a la même organisation que le RAID 6 mais utilise 3 codes de redondance. Ceci permet de continuer de fonctionner après la panne simultanée de 3 disques.

## Double parité

---

## RAID 1.5

---

Le RAID 1.5 est une évolution du RAID 1. Les accès en lecture ne se font non plus depuis un seul disque, mais en parallèle depuis tous les disques à la manière d'un RAID 0 ce qui apporte un gain sensible en performances, sans rien changer à la fiabilité.

## RAID 7, RAID ADG

---

L'évolution du RAID 3 permettant de fonctionner le tout de manière asynchrone. Il a été annoncé comme plus performant de 1,5x à 6x que n'importe quel autre RAID. Tout comme le RAID 3, un seul disque contient les parités de tous. Chaque disque a un contrôleur SCSI/SATA/Fibre\_Channel/... et le système est régi par une carte calculant la parité, gérant le cache et contrôlant les disques.

Cette version, développée par Storage Computer Corporation, est propriétaire et s'avère très coûteuse à mettre en place.

## RAID S

---

Le RAID S est un RAID 5 estampillé EMC<sup>2</sup> employé dans la gamme Symmetrix que l'on ne retrouve donc pas chez les autres constructeurs.

## Matrix RAID

---

L'*Intel Matrix Storage Technology* est disponible sur les chipsets ICH6R et ICH6RW. Il s'agit d'une solution RAID basée sur 4 ports Serial-ATA qui va permettre de mettre des disques durs en RAID de façon moins contraignante que les solutions actuelles. En effet les contrôleurs RAID actuels ne permettent avec deux disques identiques que de faire soit du RAID 0, soit du RAID 1. Bref, c'est l'un ou l'autre. Avec le Matrix RAID, Intel permet de combiner sur deux disques à la fois du RAID 0 et du RAID 1. Par exemple avec deux disques durs 250 Go, il est possible de créer un

RAID 0 (*striping*) de 100 Go (les 50 premiers Go de chaque disque) et un RAID1 (*mirroring*) de 200 Go avec les 400 Go restant. Avec les RAID traditionnels, un RAID 0 et un RAID 1 dans une même configuration nécessite l'usage de quatre disques. Même si c'est évident, ajoutons que l'ICH6R ne permet pas de faire du RAID 0+1 avec deux disques durs.

Aujourd'hui jusqu'à l'ICH8R/DO il est possible de faire un RAID-5 maximum avec 4 disques, depuis l'ICH9R/DO il est possible de faire un RAID-5 de 6 disques pour un maximum de capacité. En plus avec l'ICH 9/10 DO il est possible de rajouter un disque à une grappe existante.

## IBM ServeRAID 1E

Le Raid 1E est un système de mirroring par bandes permettant d'utiliser un nombre de disques supérieur à 2 pair ou impair. Les données sont découpées par bandes qui sont mirrorées (dupliquées) sur un des autres disques de la grappe. La capacité totale du volume logique est égal a la moitié de celle du volume physique.

## Sun RAID-Z

Le système ZFS de Sun intègre un schéma de redondance similaire au RAID 5 qui se nomme le RAID-Z. Le RAID-Z évite le « trou d'écriture » (*write hole*) du RAID 5 par une règle de copie-sur-écriture : plutôt que d'écrire par dessus des anciennes données avec de nouvelles, il écrit les nouvelles données dans un nouvel emplacement puis réécrit le pointeur vers les nouvelles données. Cela évite les opérations de lire-modifier-écrire pour des petits enregistrements en ne faisant que des écritures *full-strip*. Des petits blocs sont écrits en miroir au lieu d'être protégés en parité, ce qui est possible car le système de fichiers est conscient de la sous-structure de stockage et peut allouer de l'espace supplémentaire si nécessaire. Il existe également un RAID-Z2 qui utilise deux types de parités pour parvenir à un résultat similaire au RAID 6 : la possibilité de perdre jusqu'à deux disques sans perdre de données.

Le système RAID-Z est plus performant que le RAID 5. [réf. nécessaire]

## X-RAID (NetGear)

Le système X-RAID est une technologie développée par Netgear sur sa suite de produit ReadyNAS. Cette technologie permet suivant le nombre de disque de choisir entre les différents types de RAID et de passer de l'un à l'autre sans perte de données ou d'arrêt du service. X-RAID utilise principalement les RAID 1 et 5.

# Fiabilisation et réglages d'une installation RAID

## Choix du contrôleur

De nombreuses cartes mères proposent un contrôleur RAID intégré. Il est toujours préférable de choisir un contrôleur dédié à cette tâche. En plus de l'interface retenue, la mémoire cache du contrôleur dédié est un facteur important dans les performances du système RAID.

## Redondance de contrôleurs

Un paramètre à ne pas oublier dans la conception d'un RAID est le nombre d'interfaces contrôlant les disques (cartes IDE, SCSI, SATA ou SAS). En effet, le contrôleur RAID est un élément indispensable au fonctionnement de l'ensemble, s'il vient à défaillir, il entraîne l'indisponibilité de tous les éléments du RAID.

Si ce point est jugé crucial, il est possible de remédier à ce problème, en utilisant plus d'un contrôleur. Par exemple, dans le cas d'un RAID 1 logiciel simple reposant sur l'utilisation de disques IDE, il est conseillé de connecter chaque



Une carte RAID IDE Ultra 33

disque dur à un connecteur différent sur la carte mère. Dans les cas plus complexes, notamment lorsqu'il est fait usage d'une baie de disques externes, il est possible de rajouter une ou plusieurs baies redondantes, permettant non seulement de pallier la déficience de l'une d'entre elles, mais également de procéder à un équilibrage de charge (*load balancing*), ce qui améliore notablement les performances, notamment en évitant les goulots d'étranglement au niveau de la bande passante.

## Enfichage à chaud (hotplug/hotswap)

---

On parle abusivement de disques pouvant être enfichés à chaud (*hotplug/hotswap* en anglais), alors qu'en réalité, c'est la baie de disques du système ainsi que le contrôleur qui doivent être conçus de manière à permettre le retrait ou l'insertion de disques durs alors que le système est sous tension.

Cette fonctionnalité n'est pas disponible avec toutes les technologies :

- Bien qu'il n'y ait généralement pas de dommages physiques, les disques IDE ne gèrent pas cette fonctionnalité.
- Cette fonctionnalité est gérée par des disques SATA (sous réserve que le contrôleur le gère également).
- Cette fonctionnalité est gérée par des disques SCSI (sous réserve que le contrôleur le gère également) bien que le bus puisse être perturbé au moment de l'échange.

Cela permet :

- d'ajouter des disques de manière dynamique, de sorte qu'il soit possible de faire évoluer le système de stockage de données.
- de remplacer un matériel défectueux sans qu'il soit nécessaire d'interrompre le fonctionnement du système informatique.

L'utilisation de systèmes de connexion à chaud permet donc d'éviter l'indisponibilité durant une opération de maintenance.

## Disques de rechange (spare/hotspare)

---

Les disques de rechange (*spare/hotspare* en anglais) permettent de limiter la vulnérabilité d'une solution.

Un disque complémentaire est affecté à une unité RAID mais n'est pas utilisé au quotidien. Il est appelé disque de rechange. Lorsqu'un disque de la grappe vient à défaillir, le disque de rechange prend immédiatement et automatiquement son relais. Ce disque est alors reconstruit à partir des données présentes sur les autres disques, ce qui peut durer plusieurs heures en fonction de la quantité de données. Une fois le disque reconstruit, le système revient à un niveau optimal de sécurité et de performances.

Une fois le disque de rechange mis en service, il faut procéder à l'échange physique du disque en panne par un nouveau disque qui pourra jouer le rôle de nouveau disque de rechange.

## Cache

---

Le cache est une mémoire RAM rapide qui permet de stocker des informations à lire ou à écrire sur le RAID. Ce tampon a des performances très supérieures à celles des disques, il permet au système de lire ou d'écrire une rafale d'informations puis de passer à une autre tâche pendant que le contrôleur RAID se charge de "dépiler" les opérations à effectuer. Le cache a donc une influence positive très importante sur les performances globales du système.

Le cache de lecture ou « antémémoire de lecture » est une mémoire où sont pré-chargées les données dont le système va avoir besoin. Son fonctionnement est théoriquement très complexe dans la mesure où il apporte un gain de performance important dans le cas où le prédicat se révèle exact mais où il a tendance à dégrader les performances lorsque le prédicat est faux. Les accès aux données sont généralement de deux types : des lectures séquentielles de

données et des accès répétés à de mêmes données. La majorité des caches de lecture reposent donc d'une part sur la conservation en mémoire RAM des données récemment écrites ou lues et d'autre part sur une prélecture des données suivantes dans le cas où l'on vient d'effectuer une lecture séquentielle importante. La taille du cache de lecture et la pertinence du système de prédicat a un très grand impact sur les performances en lecture.

Le cache d'écriture est un tampon dans lesquelles sont placées les données à écrire sur les disques. Lorsqu'il n'est pas saturé, c'est-à-dire lorsque la quantité de données à écrire ne dépasse pas la taille du cache, le système peut envoyer les données à écrire très rapidement et effectuer d'autres tâches pendant que le contrôleur se charge d'écrire physiquement sur les disques. Le cache d'écriture accroît donc énormément les performances en écriture du système. Cependant, dans le cas d'écritures dont le volume dépasse largement la capacité du cache, le système devra bien attendre que le cache d'écriture se vide avant de pouvoir envoyer davantage de données. L'effet accélérateur du cache devient alors presque nul voire négatif sur certains contrôleurs.

Le cache d'écriture pose un problème de sécurité important. En effet, La mémoire cache étant volatile, les données écrites par le système mais pas encore physiquement inscrites sur les disques sont perdues en cas de rupture d'alimentation ou de défaillance. Il peut s'en suivre de très lourdes pertes ou corruptions de données. Dans certains cas graves l'intégrité du système de fichier peut être affectée et tout le volume peut alors devenir illisible. Pour pallier ce problème, on recommande de n'activer le cache d'écriture que si la machine est alimentée par l'intermédiaire d'un onduleur et que si le contrôleur RAID possède une batterie de backup (cette batterie permet de conserver le contenu du cache pendant 30 minutes à 6 heures selon les modèles de contrôleurs et la taille de la batterie).

Remarque : Seuls les véritables contrôleurs « RAID matériel » sont munis d'une mémoire cache. Pour les autres, des équivalents logiciels peuvent tenir lieu de cache.

## Granularité

---

La granularité ne s'applique qu'aux types de RAID employant des bandes de données. Il s'agit fondamentalement des niveaux RAID0 et RAID5 ainsi que leurs combinaisons avec d'autres niveaux comme par exemple le raid 10 ou le raid 15.

Il s'agit en fait de la taille des bandes de données. Cette taille configurable est généralement un multiple de 512 octets (taille d'un secteur). En pratique, elle varie le plus souvent de 16 Kio à 1 Mio. Sur la majorité des contrôleurs RAID, la valeur par défaut est de 64 Kio. Cette granularité par défaut est généralement le compromis qui permet d'obtenir les meilleures performances dans la majorité des cas. Cependant, dans des cas d'utilisation très précis, on peut être amené à la modifier.

Remarque : juger de la pertinence de ces modifications et de l'impact sur les performances nécessite une connaissance approfondie du fonctionnement interne des disques, du contrôleur RAID et du système de fichiers. Dans la mesure du possible, il est conseillé d'utiliser la valeur par défaut du contrôleur RAID.

## Les possibilités et les limites du RAID

Le RAID ne dispense pas d'effectuer des sauvegardes régulières. En effet, des défaillances à plusieurs disques sont plus fréquentes que l'on ne le croit (voir article RAID5 plus haut). De plus, des erreurs humaines (effacement/corruption de fichiers) finissent toujours par se produire.

### Ce que peut faire le RAID

---

- réduire les risques de pertes de données en cas de défaillance d'une unité de stockage
- améliorer les performances

### Ce que ne peut pas faire le RAID

---

- Protéger totalement des défaillances matérielles (éventualité de pannes successives de plusieurs disques ou du système RAID lui-même).
- Protéger les données des erreurs humaines (suppression accidentelle de fichiers).
- Protéger l'utilisateur des risques extérieurs au système (surcharge électrique qui grillerait l'ensemble des disques, incendie, vol, inondation, vandalisme).
- Protéger les données des virus et des spywares qui pourraient corrompre les données.

## Comment le RAID peut vous trahir

---

Le RAID a tendance à rendre les utilisateurs trop confiants. Cet excès de confiance en une technologie, somme toute très robuste et très fiable, peut entraîner des désastres. Multiplier le nombre de disques multiplie les risques de panne. De plus, les disques utilisés par une grappe RAID sont souvent de même type et de même âge. Ils auront donc une durée de vie similaire, c'est pourquoi il faut faire attention à utiliser des disques qui soient du même modèle mais pas de la même série. La complexité du système RAID ajoute des risques technologiques, de rares mais possibles dysfonctionnements du contrôleur RAID ou du logiciel, sont à craindre. Un disque défectueux peut aussi perturber le fonctionnement du contrôleur, logiquement ou électriquement ; ce qui peut causer la perte de plusieurs unités ; dans le cas d'un dysfonctionnement répété, la perte totale d'une grappe voire de toute l'unité de stockage est possible.

Il faut aussi garder à l'esprit que, l'utilisation d'une unité de stockage RAID n'apporte aucune protection contre :

- Les défaillances du système d'exploitation (intrinsèques ou dues à un problème de configuration ou d'un conflit de composants) ;
- Une destruction de données par dysfonctionnement d'un logiciel, virus ou malveillance ;
- Les risques physiques classiques (feu, inondation, vol, foudre et surtensions externes, surtensions internes à la machine, etc.), excepté pour les très onéreux miroirs distants (*remote mirroring*).

Il n'est donc pas raisonnable de faire totalement confiance à un système de stockage de données quel qu'il soit. Il convient alors d'effectuer des sauvegardes régulières la seule façon "sûre" de préserver une banque de données du vol ou de l'incendie étant d'en stocker une copie sur un support inerte dans un autre endroit sécurisé.