

Résumé

Cette fiche présente les systèmes RAID, les différents types (ou niveaux) de configurations existants, leurs avantages et inconvénients, ainsi que les éléments nécessaires à la mise en place de cette technologie.

Remarque préliminaire : Il convient de préciser que cette technologie, qui vise à augmenter la protection du système contre les pertes de données, n'est ordinairement installée que sur des machines critiques telles que des serveurs de fichiers, ou baies de disques durs utilisées dans les réseaux de stockage SAN (Storage Area Network).

Table des matières

- 1 Qu'est ce que le RAID ? →
- 2 Les types de RAID →
- 3 Les composants d'un Raid →



1 Qu'est ce que le RAID ?

RAID signifie « Redundant Array of Independent Disks » qui se traduit habituellement en français par « grappe de disques durs redondants ».

> Principe

En 1987 une publication, intitulée « A Case of Redundant Arrays of Inexpensive Disks (RAID) », rédigée par Patterson, Gibson et Katz de l'Université de Berkeley, en Californie, établissait le principe de base de la technologie RAID. Cela consiste à combiner plusieurs disques durs (un minimum de deux disques durs étant requis) en une grappe de disques durs apparaissant à la machine comme une unité logique de stockage unique. Pour les utilisateurs un seul disque dur est visible.

→ REMARQUE :

De nos jours « inexpensive » est souvent remplacé par « indépendant », car les prix des disques durs sont devenus abordables.

> Objectifs

Cette technologie a été mise au point principalement afin de :

- Augmenter le volume de stockage d'une machine en créant un seul espace logique de stockage correspondant pour les utilisateurs à un seul identifiant (Par exemple : Disque F:).
- Offrir une tolérance aux pannes de disques en fonction des différents niveaux de RAID décrits ci-dessous.
- Accroître les performances des machines.

2 Les types de RAID

Il existe sept types ou niveaux différents, de RAID 0 à RAID 6, qui peuvent en plus se combiner entre eux (RAID 0+1, RAID 1+0, etc.). Dans la suite, uniquement les types les plus couramment utilisés seront détaillés.

2.1 RAID 0 ou striping

Caractéristiques :

Le RAID 0, ou striping (entrelacement de données), écrit et lit les données séquentiellement sur les disques durs. C'est-à-dire qu'à l'écriture, chaque octet (Byte) est divisé en un nombre de morceaux égaux aux nombres de disques durs, qui, à la lecture sont reconstitués en l'octet original.

[→ suite](#)

Avantages :

Cette technologie augmente considérablement les performances du système en accédant à plusieurs disques durs en même temps et en réduisant ainsi le temps de transmission entre la mémoire et les disques durs. Par exemple sur un RAID 0 de n disques durs, chaque disque ne doit lire/écrire que $1/n$ des données, ce qui a pour effet de découpler les taux de transmissions, et donc d'accélérer les traitements.

Inconvénients :

Le RAID 0 ne permet pas de redondance des données et n'offre donc pas de protection contre la perte de données. En effet, la perte d'un disque dur entraîne la perte de l'entièreté des données.

Conclusion :

Etant donné le risque de perte de données, le RAID 0 n'est pas recommandé pour des applications stratégiques au sein d'une entreprise/administration ou à usage personnel. Mais pour des applications où le délai de traitement est l'élément critique (retouche d'image, vidéo,...) ce premier niveau de RAID est intéressant.

2.2 RAID 1 ou mirroring

Caractéristiques :

Le RAID 1, ou mirroring (mise en miroir de disque), écrit et lit les données simultanément sur tous les disques de la grappe. Les disques durs doivent être de capacité de stockage égale.

Avantages :

Cette technologie de mise en miroir offre une protection des données excellente (via une redondance parfaite) et garantit la sécurité des données en cas de défaillance d'un disque, mais n'apporte pas de gain de performance. En effet, tous les disques du système RAID sont à tout moment interchangeables, lors de la défaillance d'un disque le fonctionnement de la machine n'est nullement affecté. Une fois le disque défectueux remplacé, le contrôleur RAID entame automatiquement la synchronisation du nouveau disque avec les disques actifs.

Inconvénients :

Le principal inconvénient du RAID 1 réside dans le fait qu'il n'augmente pas la performance par rapport à un disque dur isolé. De plus, comme tous les disques sont miroirs l'un de l'autre, l'espace de stockage restera le même que pour un disque isolé.

Conclusion :

Ce type de RAID est recommandé pour les utilisateurs exigeant un maximum de fiabilité au niveau du stockage de données critiques sans besoins spécifiques de performances.

2.3 RAID 5

Caractéristiques :

Le RAID 5 associe le stripping (RAID 0) à un système de parité répartie. En clair, les données sont écrites séquentiellement sur un groupe de disques durs en utilisant un système basé sur les bits de parité (voir FicheBitsBytes) afin de pouvoir reconstituer les données en cas de perte d'un des disques durs. Les données de parité sont entrelacées à travers tous les disques durs. Un minimum de trois disques durs de taille égale est requis.

→ Exemple :

Sur la base de 3 disques dur de taille identique A, B et C, le système va enregistrer sur les disques A et B les données comme en mode RAID 0 (stripping) et sur le disque C les données de parité. Ainsi en cas de défaillance du disque A, les données qui y étaient accessibles le sont toujours avec les disques B et C. Il en va de même pour le disque B. Si le disque C tombe en panne, les informations se retrouvent sur A et B.

Avantages :

Le RAID 5 offrant une protection des données et de bonnes performances avec un overhead réduit pour la parité des données, c'est le plus populaire de tous. Le RAID 5 offre l'utilisation la plus efficace par rapport aux capacités des disques durs en assurant en plus la sécurisation des données.

Inconvénients :

On ne peut parler d'inconvénient, mais plutôt de reproche, lié à l'utilisation de la technique des bits de parité qui fait qu'une partie de l'espace de stockage ($1/n$, où n est le nombre de disques durs dans la grappe) n'est pas utilisée pour le stockage de données de production mais consommée par les données de contrôle (bits de parités).

Conclusion :

Ce type de RAID est adapté aux utilisations exigeantes alliant performance et criticité des données. C'est la technique conseillée pour tout serveur important ou critique (p.ex. serveur de fichiers, serveur Web, etc.).

[→ suite](#)

2.4 RAID 1 ou mirroring

Caractéristiques :

Le RAID 10 est en fait la combinaison du RAID 1 et du RAID 0 : on crée de multiples miroirs RAID 1 au-dessus desquels on ajoute un entrelacement (strip) RAID 0, un minimum de quatre disques durs de taille similaire est requis.

Avantages :

Cette technique permet de se prémunir contre de multiples défaillances disques, tant qu'il reste un disque par miroir RAID 1. Ainsi de très bonnes performances ainsi qu'un excellent niveau de protection des données est assuré.

Inconvénients :

Comme le RAID 1 cette technique est très grande consommatrice d'espace disque, en effet, elle partage l'espace utilisable en deux.

Conclusion :

Cette solution, qui est la plus onéreuse des solutions RAID, est plus particulièrement utilisée dans des cas où on exige des performances rapides et fiables avec un espace de stockage limité.

REMARQUE

Ce type n'est pas à confondre avec le RAID 0+1, ou de multiples entrelacements (RAID 0) sont mis en miroir RAID 1. Il est moins robuste que le RAID 1+0, car ne permet pas de gérer de multiples défaillances.

3

Les composants d'un RAID

Les composants générales d'une architecture RAID sont bien sûr les disques durs ainsi qu'un contrôleur RAID, dédié à la gestion de cette technologie. On parle de RAID logiciel « software RAID » dans le cas où le contrôleur est intégré dans le système d'exploitation et de RAID matériel « hardware RAID » dans le cas où la gestion est effectuée par une carte matériel dédiée.

Le choix au niveau du contrôleur logiciel ou matériel est aussi important que le choix du type de RAID à mettre en place.

► Le RAID logiciel :

Le RAID logiciel est la solution la moins coûteuse, mais peut générer de la charge supplémentaire sur le CPU (Central Processing Unit). Le RAID logiciel est, généralement, intégré dans le système d'exploitation, mais des logiciels dédiés existent. La plupart des systèmes d'exploitation limitent malheureusement le support au seul niveau RAID 0, certains supportant tout de même le RAID 5.

► Le RAID matériel :

Plus coûteux, le contrôleur RAID dédié supporte la technologie sans l'aide des ressources de la machine (CPU,...) et permet l'implémentation de tous les niveaux RAID. D'autres avantages sont :

→ Sécurité :

Un système d'exploitation qui devient instable n'aura aucune influence sur le bon fonctionnement du système RAID et donc indirectement sur la sécurité des données.

→ Performance :

Non seulement l'ajout d'un contrôleur limite la congestion du CPU, mais il élimine également le risque de congestion du bus de la machine. En effet, le bus n'est plus utilisé que pour les transferts entre les disques durs et la mémoire. De plus, la plupart des contrôleurs RAID sont équipés de mémoires caches permettant d'accélérer les traitements de requêtes vers ou en provenance des disques durs.