

# Le coin de l'expert

ERWAN BARRET

## RAID : disques, débit et sécurité

*C'est en voulant réaliser des économies que des chercheurs ont assemblé plusieurs petits disques pour en constituer un gros. Le but : sécuriser les données à un coût acceptable. Depuis, le RAID a fait son chemin et reste incontournable.*

Attribué à l'université de Berkeley, le principe du RAID (*Redundant Array of Inexpensive Disks*, ou ensemble redondant de disques à bas prix) a eu tout le loisir d'évoluer en quinze ans de carrière. À l'époque où un disque dur de quelques centaines de mégaoctets valait le prix d'une voiture, assembler des disques qui en faisaient quelques dizaines avait un sens. Pourtant, même si les capacités de stockage et les performances ont évolué, le principe du RAID reste valable, pour la vitesse comme pour la sécurité.

Utilisé sur des serveurs de toutes sortes, le RAID s'est considérablement développé avec l'explosion du multimédia et du montage vidéo numérique. Les disques durs disponibles au moment où sont apparus les premiers systèmes de type Avid étaient tout simplement

incapables d'offrir les taux de transfert indispensables. Le RAID peut donc remplir deux missions : sécuriser les données critiques ou améliorer le débit de données. Il n'est en effet pas évident de cumuler les deux qualités. Mais dans les deux cas, on assemble plusieurs disques physiques que l'on transforme en un seul volume logique à tolérance de panne.

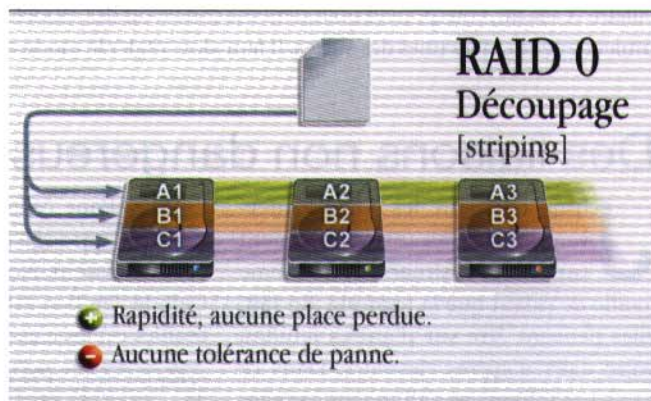
### DEUX GRANDES MÉTHODES

Le RAID repose sur deux grands principes : la redondance, qui consiste à dupliquer les données pour en avoir un double en cas de panne, et le découpage, un mode d'enregistrement qui répartit les segments de données sur plusieurs disques pour accélérer les chargements. Et quand des périphériques RAID répon-

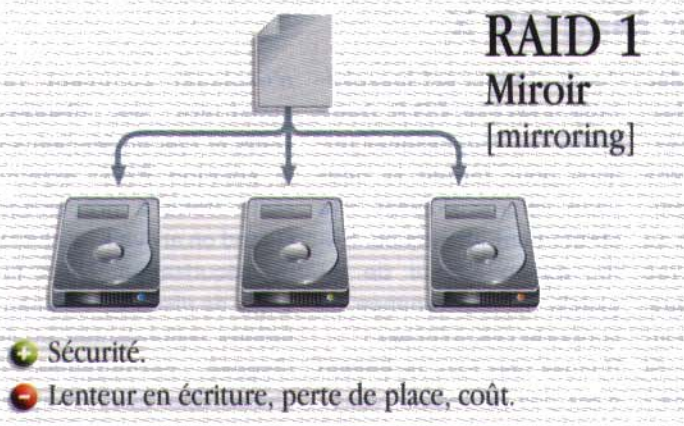
dent à un besoin d'accessibilité permanente des données, ils sont conçus pour autoriser leur remplacement s'ils présentent une avarie et sans mise hors tension de l'unité centrale qui les accueille. Cette opération peut même être automatisée à l'aide de composants de secours qui ne servent pas tant qu'une panne ne survient pas. Elle est accompagnée, le plus souvent, d'envois automatiques de messages d'alerte via le réseau. Les différentes combinaisons sont appelées "niveaux de RAID" et sont numé-

tées de 0 à 5. Certains de ces niveaux sont presque tombés en désuétude.

On trouve à la base de la plupart des schémas de RAID le principe du *striping*, littéralement "découpage en bandes". C'est ce que l'on nomme communément le RAID 0. Un chiffre bien porté car, si ce découpage réduit le temps d'accès aux données, c'est bien là son seul avantage. Il lui manque en effet la notion fondamentale qu'est la sécurité par duplication des données. Il a beau employer au mieux tout l'espace disque disponible, un



**Véritablement situé au degré zéro du RAID,** le découpage (*striping*) consiste à répartir les données sur plusieurs disques. En dépit de gains réels en temps d'accès et en débit, il souffre d'une vulnérabilité encore supérieure à celle de simples disques.



**On ne peut pas faire plus simple que le RAID 1.** Si les données sont importantes, il suffit de les dupliquer en temps réel. Une solution séduisante et efficace, mais coûteuse et finalement peu performante en phase d'écriture.

argument de poids joue contre lui : en cas de défaillance d'un seul des disques, c'est l'ensemble des données qui est perdu. Le RAID 0 s'avère donc statistiquement moins sûr que des disques classiques isolés.

## REDONDANCES ET PARITÉ

Pour équilibrer, il fallait un niveau de RAID en tout point opposé au RAID 0. Connus sous le nom de RAID 1, le *mirroring* permet tout simplement d'écrire les données au même moment sur plusieurs disques, afin de minimiser les risques de perte d'informations. Il peut être panaché avec le RAID 0 pour donner le RAID 0+1, assez fiable mais très onéreux, puisqu'il requiert au moins quatre disques et, de toute façon, deux fois plus de disques que d'espace exploitable. Mais cette combinaison ne constitue pas encore une sécurité suffisante pour les activités où les données sont vitales, ce qui a conduit au développement de schémas plus élaborés.

Souvent oublié, le RAID 2 met en œuvre la même technique de découpage que le RAID 0, mais procède aussi à des contrôles de parité pour assurer l'intégrité des données. Calculée au moment de l'écriture, ce qui induit un ralentissement, une somme de contrôles est affectée à chaque segment de données. Si la somme de contrôles calculée en lecture est différente de celle rattachée aux données lors de l'écriture, c'est qu'il y a corruption. Mais le rôle de la somme de contrôles va au-delà du diagnostic, puisqu'elle permet, en cas de panne, de recalculer un segment de données perdu. Les sommes de contrôles de tous les segments sont recopiées simultanément sur deux disques, ce qui fait du RAID 2 l'un des niveaux de RAID les plus sûrs. L'utilisation de données de parité a sur le miroir l'avantage de dupliquer les données tout en gâchant beaucoup moins de place.

Plus fréquemment rencontré, le RAID 3 reprend le même principe, mais réduit la

taille des segments pour gagner en rapidité, et dédie un seul disque au stockage des sommes de contrôles. Ce niveau de RAID, le plus rapide de tous, est couramment employé pour les fichiers très volumineux (vidéo, imageries satellite ou médicale). Moins utilisé car plus lent, le RAID 4 en est très proche, mais il est optimisé pour la lecture plutôt que pour l'écriture, ce qui le destine davantage à des applications de bases de données à haut niveau

de consultation.

Tous deux tendent à être délaissés au profit du RAID 5, qui répartit les données de parité sur tous les disques, à l'instar des blocs de données eux-mêmes. Non seulement la charge de travail est mieux partagée, mais l'écriture systématique des données de parité sur un disque différent de celui des données concernées offre des chances maximales de récupération en cas de panne. Sur les systèmes les plus évolués, aptes à écrire les données de parité en temps réel, le RAID 5 assure des performances proches de celles du RAID 3, tout en étant beaucoup plus sûr.

## MATÉRIEL OU LOGICIEL

Bien qu'il associe toujours matériel et logiciel, le RAID peut être envisagé sous deux formes. On parle de RAID logiciel par opposition au RAID matériel quand des équipements non spécifiques sont confiés à un logiciel spécialisé, chargé de répartir des données sur un groupe

de disques en fonction du niveau de RAID choisi. Très populaire sous Linux, et désormais proposé sous Mac OS X, le RAID logiciel est généralement cantonné aux niveaux de RAID 0 et 1 pour de simples questions de performances. Ensuite, on confie de préférence le travail à des périphériques spécialisés et plus ou moins intégrés. On notera toutefois l'existence de NXRAID pour Mac OS X, un logiciel de NYRO Technix ([www.nyro.com](http://www.nyro.com)) capable de gérer du RAID 5 sans carte contrôleur spécialisée.

Le matériel, qu'il s'agisse d'une simple carte contrôleur RAID ou d'une unité de sauvegarde en réseau, est principalement chargé des lourds calculs de parité. Il assume également une partie du surcroît de travail en écriture et prend en compte les contraintes de permutation à chaud des disques et alimentations.

La solution RAID idéale est le boîtier indépendant, aussi appelé SCSI to SCSI. Entièrement autonome, il possède sa propre alimentation, souvent redondante, et toute la ventilation nécessaire pour abriter plusieurs disques, généralement permutables à chaud. Une telle tour RAID, branchée sur une interface SCSI, est considérée par la machine comme un seul volume externe classique. Ce sont les cartes contrôleur situées dans la tour RAID qui prennent en charge toute la sécurisation.

Bien que la tendance évolue, le RAID est généralement associé aux disques SCSI pour des raisons de fiabilité et, dans une moindre mesure, de débit. Mais leurs concurrents ►

► à la norme IDE pourraient bien avoir les qualités de leurs défauts : un peu moins fiables, un peu plus lents et beaucoup moins chers, ils pourraient être tout simplement les *Inexpensive Disks* que le RAID est censé mettre en œuvre. On peut même installer des disques IDE sur des cartes contrôleur SCSI, à l'aide de cartes capables de les déguiser en disques SCSI. Ce "SCSI du pauvre" permet de recycler à moindres frais des cartes contrôleur RAID pour disques SCSI ou des tours RAID censées accueillir des disques SCSI, beaucoup plus onéreux.

### LE SCSI CONCURRENCÉ

Il est difficile de dire si le duel entre les deux grands standards de disques durera encore longtemps, mais le SCSI est attaqué sur plusieurs fronts. Les disques SCSI font progressivement place à des disques IDE, mais l'interface SCSI elle-même est confrontée à des concurrents très sérieux. Ainsi, l'interface FireWire, couplée à des disques IDE, comme dans les simples disques externes FireWire, offre beaucoup des avantages du SCSI : débit, fiabilité, montage externe et possibilité de chaîner plusieurs périphériques. Ce qui lui vaut d'être déjà proposée par certains constructeurs pour des solutions RAID d'entrée de gamme. Ce n'est sans doute qu'un début, puisque le FireWire à 800 Mbps est attendu pour très bientôt.

Les disques IDE équipent déjà de nombreux systèmes RAID, dont le Xserve RAID d'Apple, une unité de stockage destinée au Xserve

## Du RAID pour Mac OS X

Le RAID logiciel de Mac OS X permet de transformer un ensemble de disques en un unique volume RAID, en cochant simplement l'option appropriée au moment du formatage avec l'outil Utilitaire de disque. Mac OS X est logiquement limité au RAID 0 ou 1. Si l'intérêt du premier est discutable, les tests prouvent que le second est efficace. Comparée aux outils disponibles sous Linux, cette première mouture s'avère assez contraignante. Pour une raison inexplicée, elle ne s'applique qu'aux disques HFS+ et non aux traditionnels UFS d'UNIX. Un volume RAID créé par

l'outil d'Apple est inexploitable sous Mac OS 9. Autre limite, les deux disques utilisés pour former le nouveau volume doivent être strictement identiques. Enfin, aucun éditeur de logiciels de maintenance ne supporte ces volumes.

Encore une fois, la finition n'est pas au rendez-vous : l'interface en lignes de commande, censée être plus complète que l'Utilitaire de disque, peut en fait induire l'utilisateur en erreur en lui fournissant des données erronées. Manque d'informations encore en cas de problème : le RAID est activé, mais aucune documentation

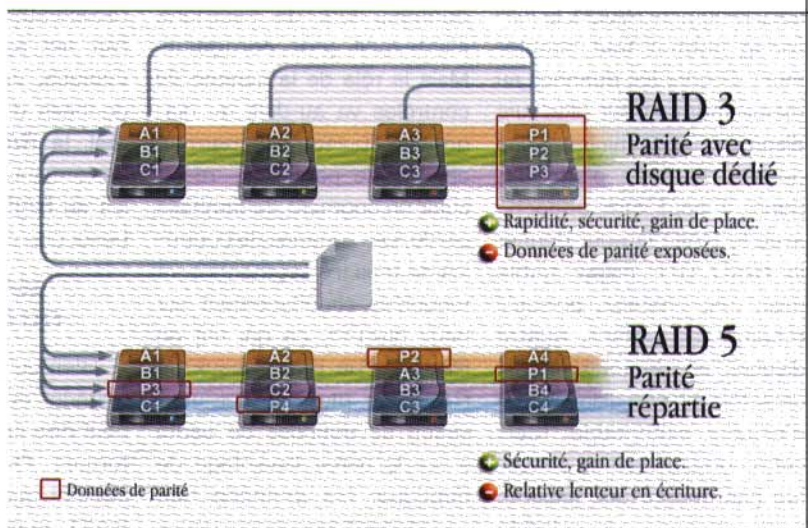
n'explique comment reconstruire ses disques. Dans le cas du Macintosh, la solution RAID externe reste donc à préconiser. Quand on arrive à ce niveau de sécurisation, il est logique de disposer d'alimentations redondantes et de disques permutables à chaud. Bien que l'on ait pu voir épisodiquement des mentions d'alimentations de ce type sur certains G4, la philosophie du RAID n'est pas encore la vocation du Macintosh, ni même du récent Xserve. Il faudra donc attendre son module RAID pour qu'Apple fasse véritablement son entrée dans ce milieu.

et dotée de quatorze disques ATA-100, une évolution de l'IDE. Ce standard au débit théorique de 100 Mo/s évoluera encore, puisque l'ATA-133 se répand déjà, et comble

en partie le fossé des performances le séparant du SCSI. Côté interface, Apple a délaissé le SCSI au profit du Fibre Channel pour relier son unité de stockage Xserve RAID à son serveur Xserve. Le Fibre Channel est une architecture mixte, à la croisée des réseaux et des périphériques, du cuivre et de

la fibre optique. Bien que très discret dans le milieu strictement informatique, il est employé depuis longtemps dans les domaines où le débit est vital, notamment la vidéo.

En choisissant cette interface pour le Xserve RAID, Apple ne fait qu'emboîter le pas à plusieurs de ses nouveaux et sérieux concurrents du monde RAID. ■



**En affectant aux segments de données un calcul de parité,** on s'assure de leur intégrité. C'est le principe des niveaux de RAID 3, 4 et 5. Les deux premiers réservent un disque aux sommes de contrôles, alors que le RAID 5 les répartit.