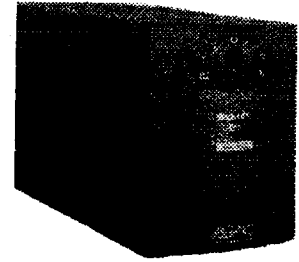


## CHAPITRE 7 : LES DISPOSITIFS MATERIELS DE TOLERANCE AUX PANNES

Les disques durs sont des supports de stockage d'une certaine fiabilité. Ils sont pourtant sujets aux pannes, et leur contenu peut être détérioré de manière irréversible.

Outre les dispositifs d'archivage et de sauvegarde vus précédemment, il existe des technologies matérielles permettant de protéger les données en cas de défaillance du système de stockage.

### 1 - Les onduleurs



#### Fonctions

Le courant électrique est distribué sous la forme d'un signal sinusoïdal qui, en théorie, est parfait et dont les caractéristiques sont constantes (230V, 50hertz). Mais de nombreux phénomènes peuvent perturber ce signal : surtensions, baisses de tensions, variations de fréquence ou altération de la courbe (distorsions harmoniques), voire coupures de l'alimentation. Ces perturbations risquent de nuire au fonctionnement d'un système informatique.

Les baisses de tensions peuvent être dues :

- à l'éloignement du site par rapport à la centrale électrique
- à une demande ponctuelle trop forte due à la mise en route d'équipements autres

Les hausses de tension proviennent de :

- l'arrêt brutal d'équipements puissants proches
- l'arrêt simultané de nombreux équipements dans le secteur

Les variations de fréquences et les distorsions harmoniques sont plus rares et dues à :

- des perturbations électromagnétiques
- la mise en marche d'un groupe électrogène

La fonction d'un onduleur est de garantir la stabilité du courant fourni au système informatique et de pallier temporairement aux coupures éventuelles.

#### Technologie

L'onduleur (dénomination usuelle adoptée pour "alimentation de secours à onduleur") transforme le courant alternatif fourni en courant continu et le stocke dans des batteries. Ce courant continu est ensuite re-transformé en courant alternatif (on le fait "onduler") dont les caractéristiques sont garanties pour alimenter un système informatique.

La puissance d'un onduleur est souvent donnée en VA (voltampère), qui est l'unité de mesure de la puissance émise, alors que le Watt est l'unité de mesure de la puissance consommée.

Un VA = 0,7 W. La puissance d'un onduleur doit donc être égale ou supérieure à la somme des puissances des équipements qu'il alimente (une UC: 200W+un écran: 100W+1lecteur externe: 30W exigent environ 400 à 500VA).

Une autre caractéristique est sa durée d'autonomie : pendant combien de temps peut-il fournir de l'énergie aux équipements auxquels il est relié ? (10 à 15 minutes pour les petits onduleurs). Cette durée devra être utilisée pour sauvegarder les données en cours de traitement et arrêter "proprement" les systèmes.

### **Off Line** (*passive stand by*)

Technologie la plus courante pour la micro-informatique et la moins chère. Les micros sont alimentés par le secteur EDF. L'onduleur surveille le courant fourni, mais ne le transforme pas. En cas de problème, il bascule sur ses batteries pour fournir un courant stabilisé. Le temps de commutation de quelques millisecondes (4 à 5) est très bien toléré par les micros, dont les condensateurs supportent jusqu'à 10ms de coupure électrique.

Avantages : technologie économique à cause du faible nombre de composants

Inconvénients : Puissance limitée (2000VA) due à la présence d'un transformateur basse fréquence dont le poids et le coût deviennent rédhibitoires au-delà de cette puissance.

Si les perturbations sont trop fréquentes, l'onduleur ne cesse de basculer entre le secteur et les batteries et risque de ne plus avoir le temps de les recharger.

### **Line interactive** (*active stand by*)

Il s'agit en fait d'une technologie Off line intelligente. L'onduleur est géré par un microprocesseur qui scrute le courant fourni et réagit de manière paramétrable aux variations constatées. Le logiciel de pilotage peut permettre par exemple de préciser la fourchette des tensions acceptables. En outre, le courant peut être filtré avant d'être délivré à l'ordinateur pour le débarrasser des parasites les plus importants.

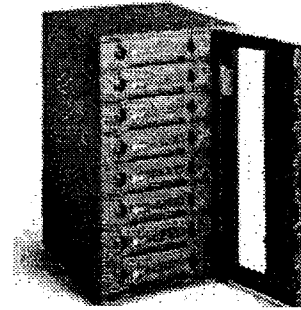
Les inconvénients sont les mêmes que pour la technologie Off Line, mais le prix plus élevé est justifié par une fiabilité plus grande et une plus grande souplesse d'utilisation.

### **On Line** (*continuous operation*)

Ce type d'onduleur alimente en permanence les équipements qu'il protège. Le courant fourni par le secteur est transformé en courant continu et stocké. Ce courant stocké est transformé en courant alternatif de qualité parfaite qui alimente les équipements. Il n'y a plus de temps de commutation, mais l'onduleur fonctionne en permanence et doit donc faire l'objet d'un entretien régulier et soigneux.

Une autre manière d'augmenter la fiabilité du système de stockage consiste à utiliser plusieurs disques en parallèles selon des procédés variables. Bien entendu, ces disques sont tributaires de leur alimentation électrique, et l'utilisation d'un onduleur garde son importance.

## II - La technologie RAID



**RAID** est un acronyme signifiant " **R**edundant **A**rray of **I**nexpensive (ou **I**ndépendent) **D**isk " ou ensemble redondant de disques indépendants.

Un ensemble (" Array ") de type RAID est une collection de disques agissant comme une unité unique de stockage qui supporte la tolérance de panne de disque sans perte de données et qui fonctionne de façon indépendante des autres sous-systèmes.

### Les différents niveaux de RAID

Un groupe de recherche de l'université californienne de BERKELEY a défini six niveaux de RAID. Chaque niveau correspond à la manière dont les données sont stockées sur les différents disques (un compromis entre coût, sécurité et vitesse). La compréhension des ces niveaux est importante car chaque niveau est optimisé pour une utilisation différente.

RAID niveau 0 : Mode " STRIPING " En mode RAID 0, la donnée à stocker est répartie sur différents disques synchronisés et aucune information de redondance n'est stockée, ce qui en résulte une vitesse de transfert importante. Néanmoins, la moindre panne de disque entraîne la perte irrémédiable de données. Le nombre minimum de disques requis, pour ce mode est de 2. Ce niveau de RAID est aussi appelé mode " STRIPING ".

RAID niveau 1 : Mode " Mirroring " et " Duplexing " En mode RAID 1, la donnée est intégralement dupliquer d'un disque sur un autre, d'où une redondance importante en cas de panne d'un disque ; néanmoins ce mode n'admet pas le " HOT SWAP " des disques. Les performances sont supérieures à un disque seul mais le coût du Méga octets est très élevé. Généralement ce mode est réalisé directement par le système d'exploitation de la machine " hôte " ou serveur, entraînant ainsi une charge non négligeable de la CPU. Ce mode requiert au minimum 2 disques et est communément appelé " MIRRORING " lorsqu'il est réalisé sur le même canal SCSI, et " DUPLEXING " lorsque 2 canaux SCSI sont utilisés.

RAID niveau 2 : Code ECC de HAMMING En mode RAID 2, chaque bit du mot de la donnée est stocké sur un ou plusieurs disques de données et le système génère et stocke sur un ou plusieurs disques ECC un code de correction d'erreur selon l'algorithme de Hamming. L'avantage de ce système est un taux de transfert possible très élevé et une correction d'erreur à la volée sans dégradation des performances. Néanmoins ce système est d'un coût élevé par rapport aux performances obtenues. D'autre part, aujourd'hui tous les périphériques SCSI possèdent en interne leur propre gestion de correction d'erreur, et de ce fait le RAID 2 n'est pratiquement plus utilisé actuellement.

RAID niveau 3 : Transfert parallèle avec gestion de parité En mode RAID 3, la donnée à stocker est répartie ( " STRIPED " ) en octets sur différents disques synchronisés de données et le système génère et stocke une parité sur un seul disque de parité. Il en résulte des taux de transfert très importants en lecture mais aussi en écriture. La panne d'un disque de données a très peu d'impact sur la performance et le nombre minimum de disques pour l'utilisation du mode 3 est de 3. La complexité du contrôleur est moyenne en mode RAID 3, néanmoins il est très difficile de le réaliser uniquement par logiciel. En réalité, très peu de constructeurs implémentent un réel RAID niveau 3 car en général la taille minimum d'une information stockée est de la taille d'un secteur ( typiquement 512 octets ).

RAID niveau 4 : Disques de données indépendants avec gestion de parité partagée En mode RAID 4, un block entier de données est stocké sur un seul disque et le système génère et stocke une parité de plusieurs blocks provenant de différents disques de données sur seul disque de parité. Les taux de transferts sont très importants en lecture de larges fichiers mais ce mode est peu performant en écriture. D'autre part, du fait du partage de la parité sur des blocks entiers de données, la reconstruction des données peut s'avérer difficile en cas de panne d'un disque. La panne d'un disque a un impact moyen sur les performances du système. Dans ce mode, les disques de données sont indépendants et ne sont pas lus en parallèle sauf dans le cas où les données proviennent de disques différents. La taille d'un block peut varier entre la taille d'un secteur et plusieurs Méga octets suivant le contrôleur utilisé et le coût du méga octets est relativement faible en mode 4.

RAID niveau 5 : Disques de données indépendants avec gestion de parité partagée et distribuée En mode RAID 5, les informations sont stockées de manière similaire au mode 4, par contre la parité est générée et est stockée de façon distribuée sur les disques de données. Les performances sont importantes pour la lecture de larges fichiers et bonnes pour la lecture de petits fichiers et en mode écriture. D'autre part, du fait du partage distribué de la parité sur des blocks entiers de données, la reconstruction des données peut s'avérer difficile en cas de panne d'un disque. La panne d'un disque a un impact moyen sur les performances du système. Dans ce mode, les disques de données sont indépendants et ne sont pas lus en parallèle sauf dans le cas où les données proviennent de disques différents. La taille d'un block peut varier entre la taille d'un secteur et plusieurs Méga octets suivant le contrôleur utilisé et le coût du méga octets est relativement faible en mode 5. . Ce mode requiert au minimum 3 disques.

RAID niveau 6 : Disques de données indépendants avec gestion de parité partagée, distribuée et répartie Le mode RAID 6 est similaire au mode 5 mais utilise plusieurs disques de parité. De ce fait, le mode 6 admet de perdre plus d'un disque de données et de continuer à fonctionner en mode dégradé. Malheureusement, le RAID 6 est peu disponible commercialement.

RAID niveau 7 : Transferts asynchrones optimisés pour importants débits de données En mode RAID 7, les informations à stocker sont réparties sur plusieurs disques de données avec un ou plusieurs disques de parité. Le mode 7 met en jeu une carte microprocesseur fonctionnant sous un noyau temps réel qui contrôle, toutes les opérations des canaux SCSI hôtes et des disques, le calcul de la parité, la gestion du cache ainsi que la surveillance des disques. Tous les transferts sont en mode asynchrone ce qui accroît de 1,5 à 6 fois les performances en écriture par rapport aux autres niveaux RAID. Ce mode supporte la perte de plusieurs disques suivant le nombre de disques de parité assignés, et d'autre part peut supporter théoriquement 12 canaux hôtes et 48 disques connectés. La plupart des contrôleurs de RAID 7 peuvent être paramétrés et administrés sur une liaison Ethernet à travers un agent SNMP.

Certains constructeurs de contrôleur RAID proposent des combinaisons de plusieurs niveaux de RAID, en voici quelques exemples :

RAID 5+1

ou

RAID niveau 1+0 : Combinaison du mode 0 (" STRIPING ") et du mode 1 (" MIRRORING ")

Ce mode à l'avantage de combiner les performances avec la sécurité.

RAID niveau 53 : Combinaison du mode 0 (" STRIPING ") et du mode 3

Ce mode accroît davantage encore la combinaison des performances avec la sécurité.