

## SAUVEGARDE<sup>1</sup>

### I. Pourquoi :

Une entreprise doit sauvegarder ses données pour pallier à une éventuelle perte de données qui peut avoir trois causes :

- ✚ Perte de données suite à un problème applicatif ou utilisateur ;
- ✚ Panne matériel ;
- ✚ Destruction de la base (incendie...).

### II. Stratégies de sauvegarde

Suite à une perte de données, il faut restaurer la base et cela est plus ou moins facile, rapide et donc plus ou moins onéreux selon les stratégies de sauvegarde mises en œuvre par l'entreprise.

Il existe trois types de sauvegarde qui peuvent se compléter dans l'entreprise.

#### A. Sauvegarde intégrale

Toutes les données sont sauvegardées, le temps du traitement est donc long. C'est pourquoi ces sauvegardes sont espacées dans le temps

*Par exemple : tous les WE mais les données stratégiques comme la finance et gestion de production seront sauvegardées 2 fois par nuit, une sauvegarde des données en fin de journée et 1 après les batchs<sup>2</sup>).*

#### B. Sauvegarde incrémentale

Seules les données ajoutées, supprimées ou modifiées depuis la précédente sauvegarde sont sauvegardées. Cette sauvegarde est donc plus rapide.

#### C. Sauvegarde cumulative ou différentielle

On sauvegarde toutes les données depuis la dernière sauvegarde intégrale. Ainsi en cas de problème, il suffit de restaurer la base à partir de la sauvegarde intégrale et de la dernière sauvegarde cumulative.

### III. Les supports

#### A. Les types de support

##### 1. Les bandes magnétiques

Les sauvegardes peuvent être effectuées sur des bandes.

##### 2. Les disques

Les sauvegardes peuvent être effectuées sur un autre disque.

<sup>1</sup> Fiche 28 pages 95 à 98 Tome 2 Processus 10 Nathan

<sup>2</sup> Traitement par lot, quand le système est arrêté, en général la nuit et le week-end.

## B. Lieu de sauvegarde et durée de stockage

### 1. *Risque de destruction totale*


Pour accroître la sécurité des données, il vaut mieux stocker la sauvegarde sur un autre site soit en transportant les bandes magnétiques soit en utilisant un disque distant. Seule la dernière sauvegarde complète est ainsi stockée.

### 2. *Perte de données partielle suite à un problème applicatif ou utilisateur*


Pour restaurer les données telles qu'elles étaient à une date antérieure, on va chercher les données dans les sauvegardes incrémentales ou complètes. Ces sauvegardes sont gardées plus longtemps.

*Exemples :*

1. *Le vendredi midi les utilisateurs se rendent compte d'un problème et souhaitent récupérer les données telles qu'elles étaient le jeudi soir. Il faut donc effacer toutes les données, restaurer :*

 *la dernière sauvegarde complète (du WE) et toutes les sauvegardes incrémentales du lundi, mardi, mercredi et jeudi*

*ou*

 *la dernière sauvegarde complète (du WE) et la sauvegarde cumulative du jeudi.*

2. *Un utilisateur souhaite reprendre un document tel qu'il était 2 semaines auparavant. Il ira récupérer son document dans la sauvegarde incrémentale du jour.*
3. *Un utilisateur vient de supprimer par erreur un de ses dossiers de travail. Il va le récupérer dans la dernière sauvegarde incrémentale.*

## IV. RAID

### A. Présentation de la technologie RAID

La technologie RAID (acronyme de *Redundant Array of Inexpensive Disks*, parfois *Redundant Array of Independant Disks*, traduisez *Ensemble redondant de disques indépendants*) permet de constituer une unité de stockage à partir de plusieurs disques durs<sup>3</sup>. L'unité ainsi créée (appelée **grappe**) a donc une grande **tolérance de panne sans perte de données, ou bien une plus grande capacité d'écriture**. La répartition des données sur plusieurs disques durs permet donc d'augmenter la sécurité de ces données et la fiabilité des services associés.

Cette technologie a été mise au point à l'Université de Californie (Berkeley) en 1988. Elle consiste à constituer un disque de grosse capacité (donc coûteux) à l'aide de plus petits disques peu onéreux (c'est-à-dire dont le MTBF, *Mean Time Between Failure*, soit le temps moyen entre deux pannes, est faible).

Les disques assemblés selon la technologie RAID peuvent être utilisés de différentes façons, appelées **Niveaux RAID**.

---

<sup>3</sup> Disque dur : organe du PC servant à conserver les données de manière permanente.

## B. Niveaux RAID

L'université de Californie en a défini 4, auxquels ont été ajoutés les niveaux 0 et 6. Chacun d'entre eux décrit la manière de laquelle les données sont réparties sur les disques.

**Chacun de ces niveaux constitue un mode d'utilisation de la grappe, fonction:**

- **des performances**
- **du coût**
- **des accès disques.**

### 1. Niveau 0

Le niveau 0, appelé *stripping* (traduisez *entrelacement*) consiste à stocker les données en les répartissant sur l'ensemble des disques de la grappe, en les synchronisant. De cette façon, il n'y a pas de redondance, on ne peut donc pas véritablement parler de système RAID. En effet en cas de défaillance de l'un des disques, l'ensemble des données réparties sur les disques sera perdu. Toutefois, étant donné que chaque disque de la grappe a son propre contrôleur, cela constitue une solution offrant une **vitesse de transfert élevée**.

On parle de facteur d'entrelacement pour désigner la taille du fragment stocké sur chaque unité physique. Le débit de transfert moyen dépend de ce facteur (plus il est petit meilleur est le débit).

Le RAID 0 consiste ainsi en la juxtaposition logique de plusieurs disques durs physiques:

2 HD de 20 Go donneront un HD logique de 40Go,

1 HD de 10Go plus 1 de 27Go donneront un HD logique de 37Go.

### 2. Niveau 1

Le niveau 1 a pour but de **dupliquer l'information à stocker sur plusieurs disques**, on parle donc de *mirroring*, ou *shadowing* pour désigner ce procédé.

On obtient ainsi une plus grande sécurité des données, car si l'un des disques tombe en panne, les données sont sauvegardées sur l'autre. D'autre part, la **lecture peut être beaucoup plus rapide** lorsque les deux disques sont en fonctionnement. Enfin, étant donné que chaque disque possède son propre contrôleur, le serveur peut continuer à fonctionner même lorsque l'un des disques tombe en panne, au même titre qu'un camion pourra continuer à rouler si un de ses pneus crève, car il en a plusieurs sur chaque essieu...

En contrepartie la technologie RAID1 est très **onéreuse** étant donné que seule la moitié de la capacité de stockage n'est effectivement utilisée.

### 3. Niveau 2

Ce niveau RAID est désormais **obsolète**.

### 4. Niveau 3 et 4

Précurseurs du niveau 5

### 5. Niveau 5

Le niveau 5 combine sécurité des données et augmentation des performances en faisant intervenir la parité des informations sur les disques durs constituant la grappe. Il est très bien adapté aux serveurs de fichiers.

Chaque unité contient une partie des informations d'une autre. Si l'une d'entre elles tombe en panne, les données perdues sont reconstruites à partir des autres. Mais le Raid 5 ne tolère qu'une seule panne de disque.

Il améliore les taux de transferts en écriture et en lecture car les opérations s'effectuent simultanément sur les disques.

### 6. Niveau 6

Le niveau 6 a été ajouté aux niveaux définis par Berkeley. Il définit l'utilisation de 2 fonctions de parité, et donc leur stockage sur deux disques dédiés. **Ce niveau permet ainsi d'assurer la redondance en cas d'avarie simultanée de deux disques.**

## C. Comparaison

Les solutions RAID généralement retenues sont le RAID de niveau 1 et le RAID de niveau 5.

Le choix d'une solution RAID est lié à trois critères :

- **la sécurité** : RAID 1 et 5 offre tous les deux un niveau **de sécurité élevé**, toutefois **la méthode de reconstruction des disques varie** entre les deux solutions. En cas de panne du système, RAID 5 reconstruit le disque manquant à partir des informations stockées sur les autres disques, tandis que RAID 1 opère une copie disque à disque.
- **Les performances** : RAID 1 offre de meilleures performances que RAID 5 en lecture, mais souffre lors d'importantes opérations d'écriture
- **Le coût** : le coût est directement lié à la capacité de stockage devant être mise en oeuvre pour avoir une certaine capacité effective. La solution RAID 5 offre un volume utile représentant 80 à 90% du volume alloué (le reste servant évidemment au contrôle d'erreur). La solution RAID 1 n'offre par contre qu'un volume disponible représentant 50 % du volume total (étant donné que les informations sont dupliquées).

## D. Mise en place d'une solution RAID

Il existe trois façons différentes de mettre en place une solution RAID sur un serveur :

- **de façon logicielle** : il s'agit généralement d'un driver au niveau du système d'exploitation capable de créer un seul volume logique avec plusieurs disques (SCSI ou IDE).
- **de façon matérielle** :
  - **avec des matériels DASD (Direct Access Storage Device)** : il s'agit d'unités de stockage externes pourvues d'une alimentation propre. De plus ces matériels sont dotés de connecteurs permettant **l'échange de disques à chaud** (on dit généralement que ce type de disque est *hot swappable*). Ce matériel gère lui-même ses disques, si bien qu'il est reconnu comme un disque SCSI standard.



- **avec des contrôleurs de disques RAID** : il s'agit de cartes s'enfichant dans des slots PCI ou ISA et permettant de contrôler plusieurs disques durs.

### Plan

Sauvegarde

- I. Pourquoi :
- II. Stratégies de sauvegarde
  - A. Sauvegarde intégrale
  - B. Sauvegarde incrémentale
  - C. Sauvegarde cumulative ou différentielle
- III. Les supports
  - A. Les types de support
  - B. Lieu de sauvegarde et durée de stockage
- IV. RAID
  - A. Présentation de la technologie RAID
  - B. Niveaux RAID
  - C. Comparaison
  - D. Mise en place d'une solution RAID

**EXERCICE : GESTION DES SAUVEGARDES SUR RESEAU**

Vous venez de décrocher le poste d'adjoint au chef comptable d'une grande entreprise régionale. Ce service occupe 9 personnes et utilise le logiciel SAGE SAARI LIGNE 100 en réseau.

Le service comptable est organisé de la manière suivante :

FONCTIONS	NOMBRE	RÔLE
Chef comptable	1	Responsable de la comptabilité générale - Déclarations fiscales et sociales
Adjoint chef comptable	1	Assiste le chef comptable dans toutes ses missions - Situation trimestrielle
Comptable Paye	1	Travaux relatifs aux salaires
Comptable Fournisseurs	1	Suivi de la comptabilité fournisseurs
Comptable Clients	2	Suivi de la comptabilité clients
Trésorerie	1	Comptabilité des opérations de caisse et de banque
Assistant comptable	2	Assistent les comptables en fonction des besoins

Actuellement, la gestion des sauvegardes n'est pas réellement organisée faute de temps. Le dossier contenant les fiches comptables est copié en fin de mois sur le disque dur du serveur contenant le dossier partagé de SAGE.

**TRAVAIL A FAIRE**

**I. Dans le tableau d'analyse ci-dessous, vous indiquerez les droits d'accès de chaque membre du service comptable :**

FONCTIONS	STRUCTURE	JOURNAUX	TRAITEMENTS	ETATS
Chef comptable				
Adjoint chef comptable				
Comptable Paye				
Comptable Fournisseurs				
Comptable Clients				
Trésorerie				
Assistant comptable				

**II. Quels sont les risques et les inconvénients de la gestion actuelle des sauvegardes des fichiers comptables?**

**III. Vous réorganisez la gestion des sauvegardes. Dans une note adressée au chef comptable, vous exposerez les mesures que vous préconisez pour fiabiliser les sauvegardes en insistant notamment sur les quatre points suivants :**

1. Qui?
2. Quand?
3. Comment?
4. Sur quels supports?

**CORRECTION**

**I. Droits d'accès :**

Ce sont les droits d'accès souhaitable qui sont indiqués dans le tableau :

FONCTIONS	STRUCTURE	JOURNAUX	TRAITEMENTS	ETATS
<b>Chef comptable</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
<b>Adjoint chef comptable</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
<b>Comptable Paye</b>	Tiers	OD	Paye	paye
<b>Comptable Fournisseurs</b>	Tiers	ACHAT	Fournisseurs	Four nisseurs
<b>Comptable Clients</b>	Tiers	VENTES	Clients	Clients
<b>Trésorerie</b>		BANQUE CAISSE	Trésorerie	Trésorerie
<b>Assistant comptable</b>		X *	X *	X *

Pour les assistants, les droits seront accordés temporairement en fonction des besoins.

**II. Risques actuels en matière de gestion des sauvegardes :**

- ⚠ Risque 1 : Aucun responsable des sauvegardes n'est désigné.
- ⚠ Risque 2 : La procédure de sauvegarde est trop aléatoire
- ⚠ Risque 3 : Le support utilisé pour la sauvegarde n'est pas indéfectible.

**III. Nouvelle organisation proposée :**

<b>1</b>	<b>Qui ?</b>	L'adjoint au chef comptable
<b>2</b>	<b>Quand ?</b>	Une sauvegarde différentielle toutes les nuits Une sauvegarde complète chaque fin de semaine
<b>3</b>	<b>Comment ?</b>	Avec une procédure automatique grâce à un logiciel de sauvegarde
<b>4</b>	<b>Supports ?</b>	Utilisation de supports variés pour éviter les risques de destruction, par exemple rotation de 3 bandes magnétiques pour la sauvegarde journalière et un autre support pour la sauvegarde de fin de semaine ( disque dur ou CDROM réinscriptible).