

La sauvegarde



Sommaire :

Introduction.....	4
La législation impose un archivage des données.....	4
Sauvegarde VS Archivage.....	4
La Sauvegarde.....	4
Objectifs :.....	4
Caractéristiques :.....	5
L'Archivage.....	5
Objectifs :.....	5
Caractéristiques :.....	5
La sauvegarde des données.....	6
La Procédure de Sauvegarde des Données.....	6
Les différents types de sauvegardes.....	6
1. Sauvegarde complète.....	6
2. Sauvegarde différentielle.....	7
3. Sauvegarde incrémentielle.....	7
Qu'est ce qu'un NAS ?.....	8
Avantages d'un NAS.....	8
Utilisation du NAS dans une stratégie de sauvegarde.....	9
Limites d'un NAS.....	9
⚠ Attention au temps de démarrage du NAS.....	9
Qu'est ce qu'un SAN ?.....	10
🔍 Définition d'un SAN (Storage Area Network).....	10
🛠 Fonctionnement.....	10
✅ Avantages.....	11
❌ Inconvénients.....	11
⚡ Interface SAS (Serial Attached SCSI).....	12
🔍 Qu'est-ce que le SAS ?.....	12
⚙️ Caractéristiques techniques.....	12
✅ Avantages.....	13
❌ Inconvénients.....	13
🆚 SAS vs SATA.....	13
⌚ Le RAID (Redundant Array of Independent Disks).....	14
🔍 Définition.....	14
⚙️ Fonctions principales du RAID.....	14
🔢 Les principaux niveaux de RAID.....	14
RAID 0 – "Stripping" (performance).....	14
RAID 1 – "Mirroring" (sécurité).....	15
RAID 5 – "Striping avec parité distribuée".....	15
RAID 6 – "Striping avec double parité".....	15
RAID 10 (ou 1+0) – "Stripping + Mirroring".....	16
RAID 50 (RAID 5+0).....	16
RAID 60 (RAID 6+0).....	16
Informations complémentaire du RAID.....	17

Introduction

Dans un environnement numérique de plus en plus central à l'activité des entreprises, la protection et la pérennité des données sont devenues essentielles. Cela passe par la mise en place de procédures de sauvegarde et d'archivage rigoureuses. Ces deux notions, bien que complémentaires, répondent à des objectifs distincts.

La législation impose un archivage des données

La législation impose une obligation d'archivage des données, notamment pour respecter les exigences légales, fiscales et réglementaires. Il est donc impératif de conserver certains documents pendant une durée déterminée, et dans des conditions garantissant leur intégrité, leur accessibilité et leur sécurité.

Sauvegarde VS Archivage

La Sauvegarde

La sauvegarde est une opération consistant à copier des données afin de pouvoir les restaurer en cas de perte, de suppression accidentelle, ou de défaillance du système.

Objectifs :

Restaurer les données en cas de panne ou d'incident

Limiter le risque de perte d'informations

Assurer une conservation à court terme

Caractéristiques :

Réalisée fréquemment (quotidienne, hebdomadaire, etc.)

Destinée à faire face aux sinistres

Stockée sur des supports variés : disques durs, serveurs, cloud, bandes magnétiques

L'Archivage

L'archivage est une opération visant à conserver durablement des documents ou des données, qu'ils soient sur support physique ou numérique, à des fins de consultation ultérieure ou de conformité réglementaire.

Objectifs :

Conserver l'information sur le long terme

Garantir la conformité avec les obligations légales (ex : réglementation fiscale)

Caractéristiques :

Conservation dans un format figé, sécurisé

Organisation par types de documents, durées légales

Accès restreint et contrôlé

La sauvegarde des données

La Procédure de Sauvegarde des Données

Une stratégie de sauvegarde efficace repose sur une **procédure claire** qui prévoit :

- **La périodicité** : quotidienne, hebdomadaire, mensuelle, selon la criticité des données.
- **Le type de sauvegarde** : complète, différentielle, incrémentielle.
- **Le nombre d'exemplaires** : plusieurs copies sont recommandées (règle du 3-2-1).
- **Le lieu de stockage** : sur site, hors site, dans le cloud.
- **Les personnes responsables** : désignation des salariés en charge de la mise en œuvre, du suivi et des tests de restauration.

Il est **essentiel de tester régulièrement** la restauration des sauvegardes afin de s'assurer que les données peuvent être récupérées rapidement et intégralement en cas d'incident.

Les différents types de sauvegardes

1. Sauvegarde complète

- **Définition** : copie **intégrale** de toutes les données sélectionnées.
- **Avantages** :
 - Restauration rapide et simple
 - Toutes les données sont centralisées dans un seul fichier
- **Inconvénients** :
 - Temps de sauvegarde long
 - Consomme beaucoup d'espace de stockage

2. Sauvegarde différentielle

- **Définition :** copie **uniquement des fichiers modifiés** depuis la dernière sauvegarde complète.
- **Avantages :**
 - Plus rapide que la sauvegarde complète
 - Moins d'espace utilisé
- **Inconvénients :**
 - La restauration nécessite la **dernière sauvegarde complète + la dernière sauvegarde différentielle**

3. Sauvegarde incrémentielle

- **Définition :** copie **uniquement des fichiers modifiés** depuis la dernière sauvegarde (**complète ou incrémentielle**).
- **Avantages :**
 - Très rapide à exécuter
 - Économie importante d'espace
- **Inconvénients :**
 - Restauration plus lente, car nécessite la **dernière sauvegarde complète + toutes les sauvegardes incrémentielles suivantes**

Il existe aussi des variantes ou techniques complémentaires, comme :

- **La sauvegarde miroir** (copie en temps réel, sans historique)
- **La sauvegarde continue** (sauvegarde automatisée dès qu'un fichier est modifié)
- **La sauvegarde dans le cloud** (hébergée à distance)

Qu'est ce qu'un NAS ?

Le **NAS** (Network Attached Storage, ou stockage en réseau) est un serveur de stockage connecté à un réseau local, permettant à plusieurs utilisateurs et systèmes d'accéder à des fichiers centralisés. Il se présente généralement sous la forme d'un boîtier contenant un ou plusieurs disques durs configurés en redondance (RAID).

Avantages d'un NAS

- **Accessibilité** : les données sont accessibles à tout moment depuis le réseau, voire à distance avec une configuration adaptée.
- **Sécurité** : possibilité de gérer les droits d'accès utilisateur, de chiffrer les données, et d'effectuer des sauvegardes automatiques.
- **Redondance** : la plupart des NAS utilisent des systèmes RAID, qui permettent de conserver les données même en cas de défaillance d'un disque dur.
- **Centralisation** : toutes les données sont stockées à un endroit unique, facilitant la gestion et la sauvegarde.
- **Automatisation** : les NAS peuvent être configurés pour effectuer des sauvegardes régulières, locales ou vers un service cloud.

Utilisation du NAS dans une stratégie de sauvegarde

Le NAS joue un rôle central dans une stratégie de sauvegarde efficace. Il peut servir de :

- **Destination de sauvegarde principale** pour les postes de travail, serveurs ou bases de données.
- **Relais de sauvegarde hors site**, en synchronisant son contenu avec un autre NAS situé dans un autre bâtiment ou dans le cloud.
- **Plateforme d'archivage**, en stockant des documents peu consultés, mais devant être conservés sur le long terme.

Limites d'un NAS

- **Coût initial** : selon les modèles et la capacité, un NAS peut représenter un investissement important.
- **Vulnérabilité au vol ou aux sinistres** : un NAS situé sur site peut être affecté par un incendie, un vol ou une inondation. D'où l'importance d'une sauvegarde complémentaire hors site.
- **Maintenance requise** : comme tout matériel, le NAS nécessite une surveillance, des mises à jour régulières et une gestion technique.

⚠️ Attention au temps de démarrage du NAS

Le démarrage d'un NAS peut prendre un temps considérable, surtout dans les contextes professionnels ou lorsqu'il est configuré avec des volumes RAID complexes. En moyenne :

- Un **NAS domestique** simple (2 baies, sans RAID ou en RAID 1) démarre en **2 à 5 minutes**.
- Un **NAS professionnel** avec 4 à 8 baies en **RAID 5 ou RAID 6** met généralement entre **8 et 20 minutes** à être pleinement opérationnel.
- En cas de **vérification de la parité RAID** (souvent automatique après un arrêt brutal ou une coupure), le processus peut durer **plusieurs heures** (parfois **6 à 12 h**), selon la taille des disques et la charge système.

Ce délai peut impacter les services réseau dépendant du NAS (partages de fichiers, sauvegardes automatiques, bases de données, etc.). Il est donc **fortement recommandé** d'utiliser une **alimentation onduleur (UPS)** pour éviter les redémarrages non planifiés, et de planifier les maintenances pendant des plages horaires creuses.

Qu'est ce qu'un SAN ?

Un **SAN** (Storage Area Network) est un **réseau spécialisé** dédié exclusivement au **stockage des données**. Contrairement au NAS, qui utilise des protocoles réseau classiques (comme SMB ou NFS) et se comporte comme un serveur de fichiers, le SAN fonctionne à un niveau plus bas, en exposant directement des blocs de données aux serveurs, un peu comme un disque dur branché localement.

Définition d'un SAN (Storage Area Network)

Un **SAN** est une infrastructure réseau à haute vitesse qui **connecte des serveurs à des baies de stockage** (matériel de stockage dédié), généralement via des protocoles comme **Fibre Channel (FC)**, **iSCSI**, ou **NVMe over Fabrics**.

Fonctionnement

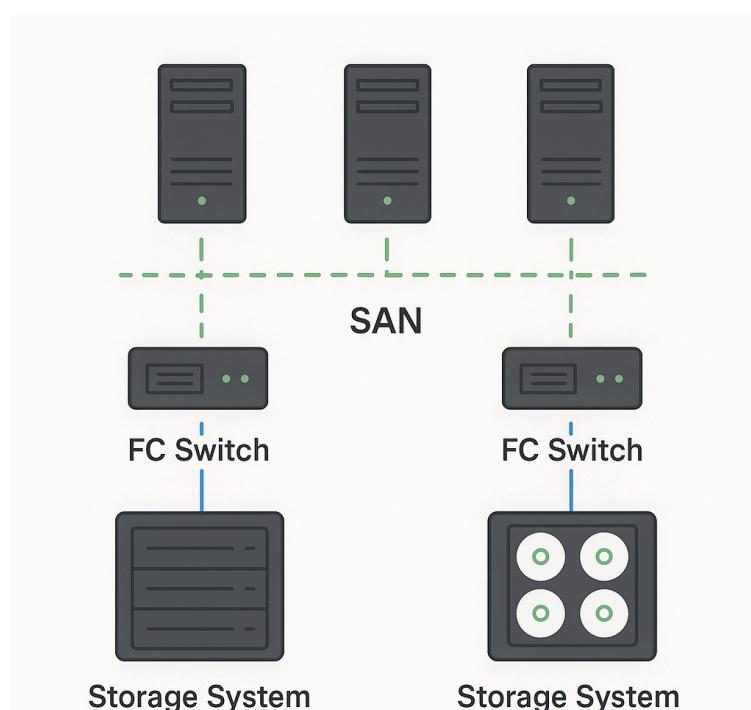
- Les serveurs accèdent au stockage via le SAN **comme s'il s'agissait de disques locaux**, mais en réalité, les données sont stockées dans des **baies de stockage centralisées**.
- Le SAN transporte des **blocs de données** (niveau bloc), contrairement au NAS qui transporte des **fichiers** (niveau fichier).
- Il est **complètement indépendant du réseau de données classique (LAN)**.

✓ Avantages

- **Très haute performance** : grâce à des connexions dédiées (souvent en 16, 32 ou 64 Gbps en Fibre Channel).
- **Faible latence** : essentiel pour les applications critiques comme les bases de données ou la virtualisation.
- **Grande évolutivité** : possibilité d'ajouter de nombreux disques, baies ou serveurs sans interruption de service.
- **Haute disponibilité** : souvent redondé avec des chemins multiples et des systèmes de tolérance aux pannes.
- **Mutualisation des ressources de stockage**

✗ Inconvénients

- **Coût élevé** : matériel spécialisé, commutateurs Fibre Channel, cartes HBA, etc.
- **Complexité de mise en œuvre** : requiert des compétences spécifiques en administration SAN.
- **Moins adapté aux PME** : sauf si les performances ou la haute disponibilité justifient l'investissement.





Interface SAS (Serial Attached SCSI)



Qu'est-ce que le SAS ?

Le **SAS (Serial Attached SCSI)** est une interface de communication en **point à point** conçue pour le transfert de données à haute vitesse entre un contrôleur (carte HBA ou RAID) et des périphériques de stockage, comme des **disques durs** ou des **SSD professionnels**.

Il est l'évolution du **SCSI parallèle**, offrant une plus grande fiabilité, de meilleures performances et une meilleure évolutivité.



Caractéristiques techniques

- **Débit** : les interfaces SAS actuelles peuvent atteindre des vitesses de transfert de :
 - **3 Gbit/s (SAS-1)**
 - **6 Gbit/s (SAS-2)**
 - **12 Gbit/s (SAS-3)**
 - **24 Gbit/s (SAS-4) (en cours de déploiement)**
- **Connexion point à point** : contrairement au SATA, le SAS peut se connecter à plusieurs disques via un **expander**, permettant une topologie en étoile.
- **Compatibilité SATA** : les connecteurs SAS peuvent gérer à la fois des disques SAS et **SATA**, bien que l'inverse ne soit pas possible (un contrôleur SATA ne gère pas les disques SAS).
- **Fiabilité professionnelle** : les disques SAS sont conçus pour des environnements 24/7 avec des cycles de lecture/écriture plus intensifs que les disques SATA.

✓ Avantages

- **Performance élevée** : particulièrement utile dans les serveurs, SAN et NAS haut de gamme.
- **Hot-swap** : permet le remplacement de disques à chaud sans arrêt du système.
- **Meilleure tolérance aux pannes** : avec des fonctions avancées comme la double connexion (multi-path).
- **Évolutivité** : possibilité de connecter jusqu'à plusieurs centaines de disques à un seul contrôleur via des expanders.

✗ Inconvénients

- **Coût plus élevé** que les solutions SATA (tant pour les disques que pour les contrôleurs).
- **Complexité d'installation** si utilisée avec des expanders ou dans un environnement SAN.

vs SAS vs SATA

Caractéristique	SAS	SATA
Débit	Jusqu'à 24 Gbit/s	Jusqu'à 6 Gbit/s
Fiabilité	Très élevée (pro)	Moyenne (grand public)
Compatibilité	Compatible SATA	Pas compatible SAS
Durée de vie (MTBF)	>1,5 million d'heures	~1 million d'heures
Environnement ciblé	Serveurs, data centers	PC, NAS domestique



Le RAID (Redundant Array of Independent Disks)



Définition

Le **RAID** est une technologie de virtualisation du stockage qui combine plusieurs disques durs physiques en un ou plusieurs volumes logiques dans le but d'**améliorer les performances, d'assurer la redondance, ou les deux**. Il est largement utilisé dans les serveurs, NAS, SAN et postes de travail exigeants.



Fonctions principales du RAID

- **Tolérance aux pannes** : certains niveaux de RAID permettent de continuer à fonctionner même si un ou plusieurs disques tombent en panne.
 - **Amélioration des performances** : la répartition des données peut accélérer les lectures et/ou les écritures.
 - **Sécurité des données** : bien que le RAID ne remplace pas une sauvegarde, il réduit le risque de perte de données liées à une défaillance matérielle.
-



Les principaux niveaux de RAID

Voici les versions les plus courantes, avec leurs avantages et inconvénients :

RAID 0 – "Stripping" (performance)

- **Principe** : données réparties sur plusieurs disques sans redondance.
- **Avantage** : très bonnes performances en lecture/écriture.
- **Inconvénient** : aucune tolérance aux pannes (si 1 disque tombe, tout est perdu).
- **Nombre de disques minimum** : 2

RAID 1 – "Mirroring" (sécurité)

- **Principe** : copie exacte des données sur 2 disques (miroir).
 - **Avantage** : tolérance aux pannes (1 disque peut tomber).
 - **Inconvénient** : capacité divisée par 2.
 - **Nombre de disques minimum** : 2
-

RAID 5 – "Striping avec parité distribuée"

- **Principe** : données réparties + parité (information de secours) sur tous les disques.
 - **Avantage** : bon compromis entre performance, sécurité et capacité.
 - **Inconvénient** : performances d'écriture un peu réduites ; tolère la perte d'un seul disque.
 - **Nombre de disques minimum** : Nombre voulu + 1
-

RAID 6 – "Striping avec double parité"

- **Principe** : identique au RAID 5, mais avec **deux disques de parité**.
- **Avantage** : peut supporter la perte de **deux disques** simultanément.
- **Inconvénient** : plus lent que RAID 5, moins d'espace utilisable.
- **Nombre de disques minimum** : 4

RAID 10 (ou 1+0) – "Stripping + Mirroring"

- **Principe** : combinaison de RAID 1 (miroir) et RAID 0 (répartition).
 - **Avantage** : très bonnes performances et redondance.
 - **Inconvénient** : coûteux (capacité effective = 50 %).
 - **Nombre de disques minimum** : 4 (paire de miroirs)
-

RAID 50 (RAID 5+0)

- **Principe** : plusieurs groupes RAID 5 combinés en RAID 0.
 - **Avantage** : performance + redondance, meilleure tolérance que RAID 5 seul.
 - **Inconvénient** : plus complexe, tolère un disque perdu par groupe RAID 5.
 - **Nombre de disques minimum** : 6
-

RAID 60 (RAID 6+0)

- **Principe** : plusieurs groupes RAID 6 combinés en RAID 0.
- **Avantage** : haute tolérance aux pannes (2 disques par groupe).
- **Inconvénient** : capacité réduite, coûts élevés.
- **Nombre de disques minimum** : 8

Informations complémentaire du RAID

Pour garantir un fonctionnement optimal d'un système RAID, certaines conditions doivent être respectées concernant les disques utilisés :

- **Disques de même taille** : Il est fortement recommandé d'utiliser des disques durs ou SSD ayant la même capacité. En effet, dans un ensemble RAID, la capacité totale sera limitée par le disque le plus petit. Par exemple, si un disque de 1 To est associé à un disque de 500 Go, l'espace utilisable sur chaque disque sera réduit à 500 Go, ce qui entraîne une perte de capacité.
- **Type et vitesse similaires** : Pour éviter les goulets d'étranglement, il est conseillé d'utiliser des disques ayant des caractéristiques proches en termes de vitesse de rotation (pour les HDD), de vitesse de lecture/écriture, et de technologie (HDD ou SSD). Cela permet d'assurer une homogénéité dans les performances du RAID.
- **Marques et modèles** : Bien que ce ne soit pas une obligation stricte, utiliser des disques identiques (même marque et modèle) peut réduire les risques de défaillances prématuées et faciliter la gestion du RAID.
- **Configuration matérielle ou logicielle adaptée** : Certains types de RAID nécessitent un contrôleur matériel spécifique, tandis que d'autres peuvent être configurés via le système d'exploitation (RAID logiciel). Le choix dépendra du niveau de performance et de fiabilité souhaité.

En respectant ces recommandations, on maximise la fiabilité, les performances, et la durée de vie globale du système RAID.