

SAUVEGARDE & RESTAURATION

Principes, outils et techniques de sauvegarde et de restauration des données

BTS SIO – Bloc n°2 : Administration des systèmes

Sommaire

- I. Quelques définitions : archivage, stockage et sauvegarde
- II. Les objectifs de la sauvegarde
- III. Stratégies de sauvegarde
- IV. Stratégies de sauvegarde : types/catégories de sauvegarde
- V. Stratégies de sauvegarde : le mécanisme des types de sauvegarde
- VI. Types de sauvegarde : un comparatif – avantages & inconvénients
- VII. Les problématiques de la sauvegarde
- VIII. Les solutions de sauvegarde
- IX. Sauvegarde : Les outils de stockage
- X. Sauvegarde : les supports de stockage

Quelques définitions : archivage, stockage et sauvegarde

- Archivage :
 - ensemble des actions, outils et méthodes utilisés
 - pour réunir, identifier, sélectionner, classer et conserver des contenus électroniques
 - sur un support sécurisé,
 - dans le but de les exploiter et de les rendre accessibles dans le temps : à titre de preuve (en cas d'obligations légales notamment ou de litiges) ou à titre informatif.

Quelques définitions : archivage, stockage et sauvegarde

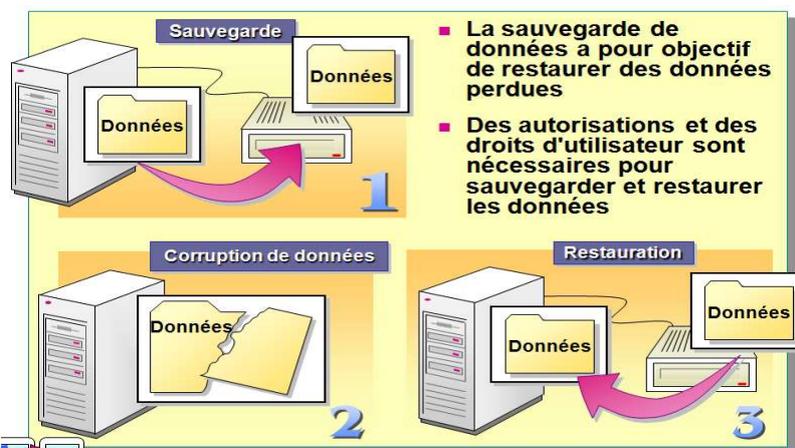
- L'archivage (suite):
 - Le contenu archivé est considéré comme figé et ne peut donc être modifié.
 - La durée de l'archivage est fonction de la valeur du contenu et porte le plus souvent sur du moyen ou long terme.

Quelques définitions : archivage, stockage et sauvegarde

- Le stockage:
 - actions, outils et méthodes permettant d'entreposer des contenus électroniques et servant de base au traitement ultérieur des contenus.

Quelques définitions : archivage, stockage et sauvegarde

- Le principe de la sauvegarde :



Quelques définitions : archivage, stockage et sauvegarde

- La sauvegarde :
 - actions, outils et méthodes destinés à **dupliquer des contenus électroniques dans un but sécuritaire.**
 - le but est **d'éviter leur perte en cas de dysfonctionnement du dispositif sur lequel ils sont enregistrés** (dans le cas de plans de reprise d'activité ou de continuité d'activité - PRA/PCA).

Quelques définitions : archivage, stockage et sauvegarde

- La sauvegarde (suite):
 - Le **contenu sauvegardé** n'est pas considéré comme figé et **peut donc être modifié ou remplacé.**
 - La **durée de la sauvegarde** est **fonction de sa périodicité** et porte le plus souvent sur du **court terme.**

Objectifs de la sauvegarde

- La sauvegarde permet de protéger les données contre :
 - erreurs de manipulation des utilisateurs,
 - défaut du système d'exploitation,
 - virus informatiques et logiciels malveillants,
 - dommages ou pannes du matériel (disques durs...),
 - suppression de fichiers (accidentelle ou non),
 - altération, modification ou suppression d'informations à mauvais escient,
 - vol de matériels, de logiciels, sabotages,
 - la perte de l'espace de stockage initial,
 - ou encore les catastrophes naturelles (incendie, inondation, ...).

Objectifs de la sauvegarde

- La sauvegarde des données est essentielle pour une entreprise, quelle que soit sa taille, à partir du moment où toute information importante est stockée sur l'infrastructure de son réseau.
- Il est rationnel de prendre en compte toutes les éventualités pouvant conduire à la perte partielle ou totale, temporaire ou définitive, des informations de l'entreprise.

Les stratégies de sauvegarde

- Le mécanisme de sauvegarde mis en œuvre doit impérativement :
 - être pensé de manière à assurer la pérennité et la récupération de l'ensemble des données critiques de l'organisation, quel que soit le sinistre subi,
 - sans perturber le fonctionnement du système d'information.

Les stratégies de sauvegarde

- Ainsi, le choix du mécanisme de sauvegarde doit faire l'objet d'une stratégie de sauvegarde, définissant :
 - les données à sauvegarder,
 - la fréquence et le mode de sauvegarde,
 - et d'un plan de reprise sur sinistre indiquant la démarche nécessaire pour rétablir le fonctionnement normal en cas d'incident.

Les stratégies de sauvegarde

- La stratégie de sauvegarde consiste à définir :
 - les données sauvegarder : la localisation de celles-ci, leur criticité et leur volumétrie ;
 - Sur quels supports ;
 - Quelles procédures appliquer ;
 - Quelles garanties sont attendues : le temps d'indisponibilité accepté et le récupération en cas de dommage majeur (plan de reprise d'activité);
 - Les intervenants et leur rôle respectif.

Les stratégies de sauvegarde

- La politique de sauvegarde implique de formaliser les objectifs de reprise en termes de :
 - **RTO** (Recovery Time Objective) : spécifie le délai maximum que l'entreprise tolère avant de reprendre son activité.
 - **RPO** (Recovery Point Objective) : désigne la durée maximum d'enregistrement des données qu'il est acceptable de perdre lors d'une avarie.

Stratégies de sauvegarde : types/catégories de sauvegarde

- On distingue habituellement les catégories de sauvegardes suivantes :
 - **Sauvegarde totale** : L'objectif de la sauvegarde totale (full backup) est de réaliser une copie conforme des données à sauvegarder sur un support séparé.
 - **Sauvegarde incrémentale (partielle)** : La sauvegarde incrémentale (en anglais incremental backup) consiste à copier tous les éléments modifiés depuis la sauvegarde précédente.

Stratégies de sauvegarde : types/catégories de sauvegarde

- On distingue habituellement les catégories de sauvegardes suivantes :
 - **Sauvegarde différentielle (partielle)** : La sauvegarde différentielle (en anglais differential backup) se focalise uniquement sur les fichiers modifiés depuis la dernière sauvegarde complète.

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de la sauvegarde complète :
 - Une copie de la totalité des données présentes sur les disques ou les partitions d'un serveur (défini comme client de sauvegarde) vers une unité de stockage (bandes magnétiques, cartouches, unités de disques).
 - *Une sauvegarde complète ne vérifie pas si le fichier a changé depuis la création de la dernière sauvegarde ; l'opération consiste à enregistrer aveuglément toutes les informations sur un support de sauvegarde, qu'elles aient changé ou non.*

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de la sauvegarde complète :
 - Une copie de la totalité des données présentes sur les disques ou les partitions d'un serveur (défini comme client de sauvegarde) vers une unité de stockage (bandes magnétiques, cartouches, unités de disques).
 - *Une sauvegarde complète ne vérifie pas si le fichier a changé depuis la création de la dernière sauvegarde ; l'opération consiste à enregistrer aveuglément toutes les informations sur un support de sauvegarde, qu'elles aient changé ou non.*

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de la sauvegarde incrémentale :
 - *vérifie si la date de modification est postérieure à la date à laquelle la sauvegarde a été effectuée, le fichier a été modifié et devra donc être sauvegardé. Si ce n'est pas le cas, le fichier n'a pas été modifié depuis la dernière sauvegarde et ne sera donc pas pris en compte lors de la sauvegarde actuelle.*

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de la sauvegarde différentielle :
 - *semblable à la sauvegarde incrémentale dans le sens où dans les deux cas, seuls les fichiers ayant été modifiés seront sauvegardés. Toutefois, les sauvegardes différentielles sont cumulatives.*
 - *en d'autres termes, avec une sauvegarde différentielle, dès l'instant où un fichier a été modifié, il sera toujours inclus dans toutes les versions suivantes des sauvegardes différentielles (jusqu'à la prochaine sauvegarde complète).*

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de différenciation des types de sauvegarde :
 - La solution mise en place est l'utilisation d'un **marqueur d'archivage**.
 - Chaque fichier possède ce marqueur d'archivage, qui est positionné à "vrai → 1" lorsque l'on crée ou modifie un fichier.
 - On peut interpréter cette position de la façon suivante: « je viens d'être modifié ou créé : je suis prêt à être archivé donc je positionne mon marqueur à vrai".
 - **Ce marqueur** est appelé aussi **attribut d'archivage** (ou **bit d'archivage**).

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de différenciation des types de sauvegarde :
 - **Sauvegarde complète** : tous les fichiers sont sauvegardés, indépendamment de la position du marqueur (vrai ou faux). Une fois le fichier archivé, celui-ci se voit attribuer la position de son marqueur (le bit d'archive) à "faux" (ou à "0").
 - **Sauvegarde incrémentale** : tous les fichiers dont le marqueur est à "vrai" sont sauvegardés. Une fois le fichier archivé, celui-ci se voit attribuer la position de son marqueur à "faux".

Stratégies de sauvegarde : le mécanisme des types de sauvegarde

- Le mécanisme de différenciation des types de sauvegarde :
 - **Sauvegarde différentielle** : Lors d'une sauvegarde différentielle, tous les fichiers dont le marqueur est à "vrai" sont sauvegardés. Une fois le fichier archivé, celui-ci garde la position de son marqueur tel qu'il l'avait avant la sauvegarde.
 - Ainsi, en fin d'exécution, **elle ne repositionne pas les attributs à 0.**

Types de sauvegarde : un comparatif - avantages & inconvénients

- Avantages et inconvénients des types de sauvegarde :

Types de sauvegarde	Avantages	Limites
Complète (ou Totale)	<ul style="list-style-type: none"> ▪ Restauration des données est facilitée car sur un seul support ▪ Image fidèle des données à un instant T 	<ul style="list-style-type: none"> ▪ Gros volumes de données ▪ Problème de lenteur ▪ Disponibilité des données car accès disques longs et intenses. ▪ Coût étant donnée la capacité nécessaire

Types de sauvegarde : un comparatif - avantages & inconvénients

- Avantages et inconvénients des types de sauvegarde (suite) :

Types de sauvegarde	Avantages	Limites
Incrémentale	<ul style="list-style-type: none"> ▪ Plus rapide que la totale ▪ Espace de stockage plus faible 	<ul style="list-style-type: none"> ▪ Posséder les sauvegardes précédentes pour reconstituer des données complètes.
Différentielle	<ul style="list-style-type: none"> ▪ Plus fiable que l'incrémentale car seule la complète est nécessaire pour reconstituer les données. 	<ul style="list-style-type: none"> ▪ Plus lente et plus coûteuse en espace de stockage que l'incrémentale.

Types de sauvegarde & restauration

- Chaque type de sauvegarde donne lieu à une procédure/démarche de restauration :
 - **Sauvegarde totale** : la restauration est simple, les fichiers et données sont faciles à trouver car stockés sur le même support. Cela facilite la copie des fichiers et données à partir du support.
 - **Sauvegarde différentielle** : une restauration demande seulement les supports de la dernière sauvegarde totale puis, la dernière sauvegarde différentielle.
 - **Sauvegarde incrémentale** : une restauration demande les supports de la dernière sauvegarde totale puis, la restauration de chaque sauvegarde incrémentale (les dernières).

Les problématiques actuelles de la sauvegarde

- L'architecture générale de la sauvegarde n'a quasiment pas changé.
- Par contre, les systèmes informatiques à sauvegarder ont profondément évolué du fait :
 - **de l'utilisation massive des postes nomades** : de plus en plus d'utilisateurs conservent des données sensibles sur leurs équipements nomades, ce qui nécessite de trouver des solutions de sauvegardes automatiques lors de la connexion des postes au réseau ;
 - **de la volumétrie croissante des données à sauvegarder** : ce qui a une conséquence sur les fenêtres de sauvegarde (Intervalle de temps durant lequel il est possible de maintenir un système en état dormant, pour permettre sa sauvegarde) ;

Les problématiques de la sauvegarde

- Par contre, les systèmes informatiques à sauvegarder ont profondément évolué du fait :
 - **de l'archivage** : les données courantes ne nécessitent généralement pas une conservation sur une durée supérieure à 3 mois. Cependant, certaines données imposent une conservation des données sur très long terme ;
 - **des machines virtuelles** : la mise en œuvre de la virtualisation serveur a modifié la typologie des données à sauvegarder. Une solution, bien adaptée aux machines virtuelles est de les sauvegarder dans leur intégralité, images disques et fichiers de configuration, ce qui implique de gérer le problème de la sauvegarde régulière de gros fichiers.

Les solutions de sauvegarde

- Un aperçu des solutions de sauvegarde :
 - **Sauvegarde sur bande** : les lecteurs de sauvegarde peuvent être autonomes et directement reliés à un serveur ou pilotés par une librairie (routine). Celles-ci sont des périphériques spécialisées comprenant plusieurs lecteurs, capables d'écrire et/ou lire sur plusieurs bandes simultanément.

Les solutions de sauvegarde

- Un aperçu des solutions de sauvegarde :
 - **La sauvegarde sur serveur** : il est possible de mettre en œuvre un premier niveau de protection des données au niveau du système de stockage centralisé, il suffit pour cela de disposer d'un espace disque supplémentaire.
 - **Les architectures à base de disques durs** - les architectures à base de disques durs ont pour avantages :
 - plus souples d'utilisation
 - plus performants en terme de temps d'accès et de nombre d'accès simultanés ;
 - rendus robustes grâce à des mécanismes de redondance.

Les solutions de sauvegarde

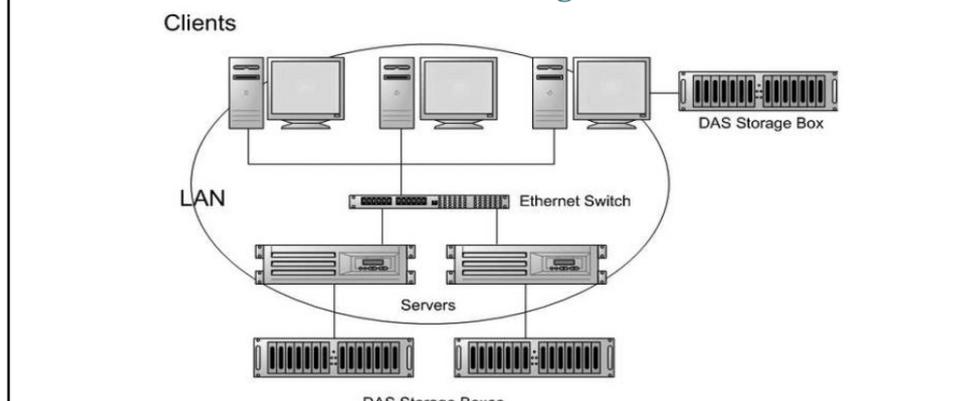
- Un aperçu des solutions de sauvegarde :
 - **Connexions aux baies :**
 - avant les systèmes « robotiques » et les lecteurs étaient souvent connectés au serveur en SCSI.
 - Pour les baies de disques actuelles, la connectique pourra être de 2 types : **directe ou réseau**.
 - En attachement direct, les baies supportent SAS (Serial Attached SCSI), iSCSI ou FC (Fibre Channel).
 - La connexion des baies par le réseau s'appuie essentiellement sur les protocoles iSCSI et NFS/CIFS.

Les solutions de sauvegarde

- Un aperçu des solutions de sauvegarde :
 - **La sauvegarde des postes nomades :**
 - une sauvegarde en continu, c'est un mécanisme qui sauvegarde les fichiers chaque fois qu'ils sont modifiés.
 - Le processus de sauvegarde est toujours actif : lorsqu'un fichier est modifié, une copie des données modifiées est mise dans un cache local.
 - Ce dernier contient donc toutes les versions de tous les fichiers modifiés.
 - Le cache local est ensuite transféré sur le serveur de sauvegarde dès que le poste nomade est connecté.
 - Ce transfert doit lui aussi être optimisé afin d'éviter une saturation de la bande passante.

Sauvegarde : les outils de stockage

- Les outils et méthodes de stockage des données lors d'une sauvegarde sont nombreux et divers :
 - **DAS (Direct Attached Storage)**

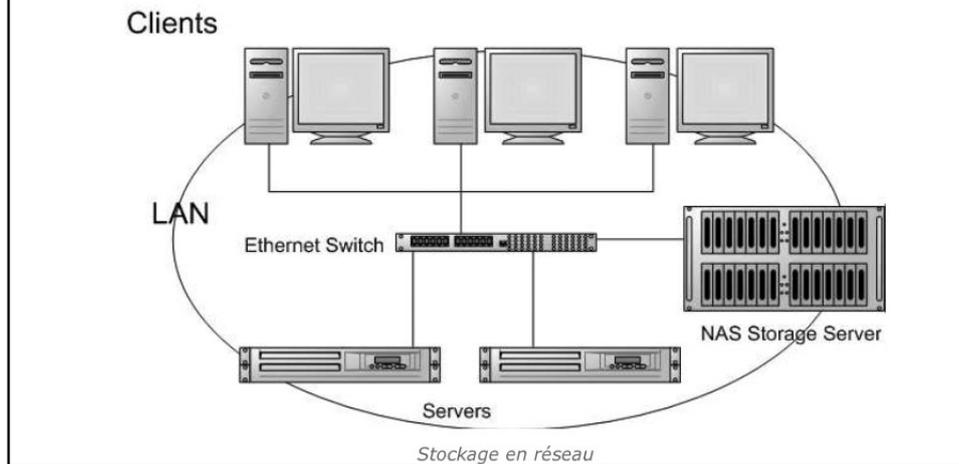


Sauvegarde : les outils de stockage

- **DAS (Direct Attached Storage) :**
 - le stockage en attachement direct consiste à relier un ou plusieurs sous systèmes de stockage à un serveur via un lien direct.
 - Le matériel servant au stockage est un périphérique attaché directement au serveur, généralement attaché en USB, Firewire, eSATA, SCSI, SAS ou Fibre Channel.
 - Il peut être constitué d'un simple disque dur USB ou une baie de stockage avec disques agrégés en RAID, créant ainsi une tolérance de panne sur disque tout en augmentant les performances I/O (Input/Output).

Sauvegarde : les outils de stockage

- **NAS (Network Attached Storage) :**



Sauvegarde : les outils de stockage

- **NAS (Network Attached Storage) :**

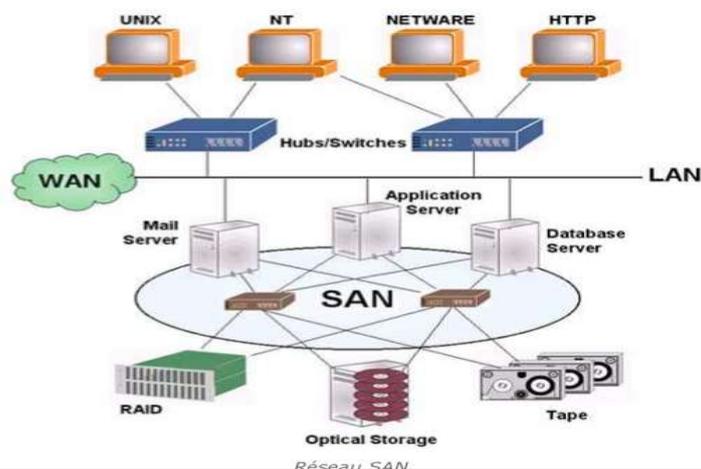
- Il s'agit d'une solution intégrée où l'entreprise concentre ses données sur un boîtier de stockage regroupant à la fois le matériel et le logiciel.
- Le NAS se construit sur un réseau IP dédié ou non au stockage.
- Le serveur NAS sert à partager les données dans des environnements réseaux hétérogènes et utilise une couche logicielle capable de gérer les droits d'accès aux données, tout en bénéficiant d'une souplesse d'administration.

Sauvegarde : les outils de stockage

- **NAS (Network Attached Storage) :**
 - Les protocoles mis en œuvre par les NAS sont standard : SMB/CIFS, NCP, NFS, FTP, HTTP.
 - Souvent, les serveurs NAS sont des serveurs avec des disques agrégés en niveaux de RAID sur des systèmes d'exploitation Windows (Windows Storage server) ou des dérivés de Linux.
 - Chaque constructeur y ajoute des fonctionnalités comme la gestion des quotas utilisateur et/ou disques, le filtrage de fichiers, Snapshot,...

Sauvegarde : les outils de stockage

- **SAN (Storage Area Network) :**



Sauvegarde : les outils de stockage

- **SAN (Storage Area Network) :**
 - Le SAN est un réseau purement dédié au stockage, soit en se basant sur TCP/IP par le biais de la technologie iSCSI, soit en utilisant la fibre optique par le biais du Fiber Channel.
 - Contrairement au NAS, le SAN organise un ensemble d'éléments indépendants.
 - Il est possible de regrouper différents supports de stockage (disques et bandes) et de les faire cohabiter sous un même logiciel d'administration.

Sauvegarde : les outils de stockage

- **SAN (Storage Area Network) :**
 - Le SAN est un ensemble de baies de stockage dont l'espace alloué est distribué en mode blocs, c'est-à-dire que chaque espace réservé sur cet espace sera dédié à un seul et unique serveur.

Sauvegarde : les outils de stockage

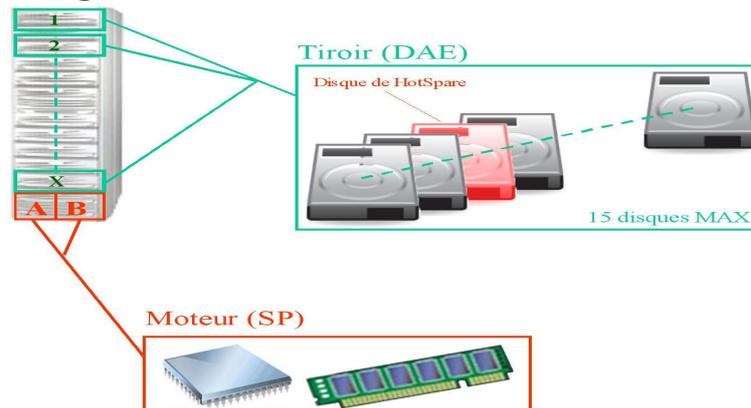
- Le SAN se distingue par les caractéristiques suivantes :
 - Un accès partagé à une ferme de stockage commune et à une ou plusieurs librairies de bandes pour la sauvegarde et la restauration.
 - Utilisation d'un réseau Fibre Channel distinct et spécifique au stockage ;
 - Gestion des transferts de données stockées entre les serveurs et les dispositifs de stockage sur le SAN, allégeant de ce fait la charge du LAN ;
 - Installation distante de sous-systèmes de disques durs et de librairies de bandes.

Sauvegarde : les outils de stockage

- On trouve généralement dans une architecture SAN:
 - Un ou plusieurs « **fabric**s » constitué de **switchs Fibre Channel** (appelés FC-SW)
 - La plupart des réseaux de stockage sont configurés en double-fabric (deux fabric parallèles indépendantes) : haute disponibilité par redondance des chemins d'accès, forte tolérance aux pannes, interventions de maintenance matérielle transparentes pour les applications...
 - Plusieurs baies qui peuvent être de type hétérogène :
 - l'espace disque de la baie est divisé en groupes RAID.
 - Ces groupes sont ensuite divisés en unités de stockage logiques appelées LUNs (Logical Unit Number – numéro d'identification d'une unité de stockage SAN).
 - Enfin, ces LUNs sont agrégés en meta- LUNs.;
 - Des serveurs Storage Processor ou SP, permettent d'administrer les baies de stockages (configuration des sauvegardes, gestion des groupes RAID, etc.) ;
 - Des serveurs d'application.

Sauvegarde : les outils de stockage

- Schéma représentant une baie de stockage et un Storage Processor :



Sauvegarde : les supports de stockage

- Les supports de stockage sont aussi divers et variés.
- Les **bandes magnétiques** :
 - constituent un support de stockage fiable, économique et à faible risque, utilisées généralement comme supports d'archivage - 19 nouvelles technologies de bandes correspondant à de nouveaux formats ont été mises sur le marché.
- Les **disques durs magnétiques** :
 - Deux connecteurs principaux existent pour brancher un disque dur : **IDE** (également appelé ATA ou PATA) et **SATA**. Les ports IDE limitent le refroidissement des composants (larges nappes). De plus, leur débit théorique est plus faible que celui des ports SATA. On trouve également des disques durs en **SCSI** (débits jusqu'à 64 Mo/s) ou **SAS** (jusqu'à 600 Mo/s).

Sauvegarde : les supports de stockage

- **Les supports optiques :**
 - Les lecteurs de médias de stockage amovibles optiques utilisent une source de lumière laser pour lire et/ou écrire des données sur le disque. Les CD (Compact Disc) et les DVD (Digital Versatile Disc) sont les deux principaux formats optiques.
 - La technologie Blue-ray fonctionne sur le même principe que le CD audio ou le DVD. Un faisceau laser (bleu) est projeté sur un disque, contenant des informations : la réflexion du rayon laser est traitée afin d'interpréter les informations numériques présentes sur le disque.

Sauvegarde : les supports de stockage

- **Le stockage holographique :**
 - Sur un disque holographique, l'information est inscrite sous forme d'hologramme au cœur d'un cristal
 - L'écriture des données nécessite l'emploi d'un laser, dont la lumière est divisée en deux faisceaux : un faisceau de référence et un faisceau objet. Le faisceau objet traverse un écran à cristaux liquides, sur lequel les données à stocker sont représentées (un pixel noir pour un 0, et un blanc pour un 1).
 - La rencontre avec le faisceau de référence crée un motif d'interférence, un hologramme. En modifiant l'angle du faisceau de référence, de nouvelles données peuvent être stockées dans le même volume du cristal.

Sauvegarde : les supports de stockage

- **Le stockage holographique :**
 - Pour l'instant, ce type de support holographique est réservé à l'archivage : les premiers disques ne sont pas réinscriptibles. Ce qui en limite leur utilisation.
 - Le stockage holographique se place néanmoins en concurrence directe avec des technologies optiques de type Blue-Ray, mais aussi avec des technologies de bandes.

Sauvegarde : les supports de stockage

- Tableau récapitulatif des supports de stockage traditionnels et holographique :

Stockage holographique	disques optiques Blu-Ray	Bandes	Disques durs d'entreprise
<ul style="list-style-type: none"> • Capacité : de 300 Go à 1,6 To • Taux de transfert : de 20 à 120 Mo/s • Temps de vie {pour archivage} : 50 ans • Prix au gigaoctet : 0,6 \$ (0,2 \$ à terme) • Fonction Worm : oui • Contact avec tête de lecture : non • Sécurité : chiffrement optique 	<ul style="list-style-type: none"> • Capacité : de 23 Go à 100Go • Taux de transfert : de 4 à 12 Mo/s • Temps de vie {pour archivage} : 50 ans • Prix au gigaoctet : de 1 à 2 \$ • Fonction Worm : non • Contact avec tête de lecture : non • Sécurité : non 	<ul style="list-style-type: none"> • Capacité : de 100 Go à 1To • Taux de transfert : de 20 à 120 Mo/s • Temps de vie {pour archivage} : de 7 à 10 ans • Prix au gigaoctet : de 1 à 25 \$ • Fonction Worm : oui • Contact avec tête de lecture : oui • Sécurité : oui, avec technologies dédiées 	<ul style="list-style-type: none"> • Capacité : de 18 Go à 1To • Taux de transfert : de 40 à 150 Mo/s • Temps de vie {pour archivage} : non communiqué • Prix au gigaoctet : moins de 3 \$ • Fonction Worm : non • Contact avec tête de lecture : oui • Sécurité : chiffrement possible

Sauvegarde : les supports de stockage

- Citez les critères de choix d'un support de stockage pour réaliser un dispositif de sauvegarde:
 - La capacité totale (native et disponible sur chaque support (Go ou To),
 - Les besoins en stockage exprimés par l'infrastructure,
 - Le coût du support (le coût total des cartouches nécessaires à un an de besoins en sauvegarde peut dépasser ce prix du lecteur),
 - Le taux de transfert/débit (Mo/s),
 - La durée de vie (années),
 - La compatibilité ascendante en fonction des modèles,
 - La standardisation du matériel,
 - Et la sécurité (chiffrement ou non) des données.
 - → WORM (Write Only Read Multiple)

SAUVEGARDE & RESTAURATION

FIN