Sauvegarde versus Archivage: sauvegarde

- sauvegarde: c'est dupliquer et externaliser les données pour se protéger :
  - des erreurs humaines
  - des crash matériel
  - des erreurs logicielles
  - des sinistres (incendie, inondation, ...)
  - de la malveillance (vol, virus, ...)
- la durée de vie d'un jeu de sauvegarde est limitée dans le temps (quelques mois au plus)
- Sauvegarde des données / du système

Sauvegarde versus Archivage: sauvegarde

- on sauvegarde certaines données pour se protéger de certains risques
  - faire la liste des risques dont il faut se protéger
  - adapter le cahier des charges des sauvegardes et des autres mesures en fonction des ces risques
  - les sauvegardes ne sont qu'un des éléments permettant de garantir la continuité du service à côté d'autres : disques raid, redondance des serveurs, ...

Sauvegarde versus Archivage: archivage

- l'archivage, c'est le stockage hors ligne de données peu utilisées
  - le temps d'accès a souvent peu d'importance
  - on peut avoir une hiérarchie de mode de stockage (du plus rapide au moins rapide (bande à aller chercher) suivant la fréquence de l'utilisation des données concernées)
- la durée de vie d'une archive se compte au minimum en années, voire en décennies
- il faut donc penser à l'obsolescence des matériels, des supports, des logiciels, des standards utilisés

Solutions qui ont prouvé leur inefficacité

- Faire faire les sauvegardes par les utilisateurs : pour dégager sa responsabilité mais aucune garantie qu'elles seront faites
- Sauvegardes sur des supports peu fiables (disquettes, DAT, ...)
- Sauvegarde en écriture seule : il faut valider sauvegarde et procédures de restauration
- Sauvegarde d'un système en cours d'exécution
- Un seul support de sauvegarde
- Sauver sur une partition du même disque
- Pas de sauvegarde hors site (incendie, ...)

Sauvegardes : procédure/planification

- Planifier les sauvegardes, tester leur réalisation
  - il faut avoir l'assurance que les sauvegarde prévues ont eu lieu
  - les procédures de sauvegardes doit être écrites
  - procédures testées
  - procédures et planification validées par les utilisateurs/propriétaires des données

Sauvegardes : planification

- choix des données à sauvegarder :
  - données des utilisateurs (y compris : boîtes aux lettres, profil, ...), fichiers de configuration, ..;
    - retrouver un système en état suppose de réinstaller le système puis de restaurer les données
    - volumétrie plus faible
  - Système entier avec procédure de redémarrage:
    - restauration directe et rapide d'un système opérationnel
    - volumétrie plus importante
- périodicité, choix des données ont un impact fort sur la volumétrie et donc sur le coût =>compromis

Problèmes liés à la volumétrie:

- Coût
- Charge réseau
- Durée des sauvegardes
- Indisponibilité des serveurs/applications dans le cas où le logiciel de sauvegarde impose l'arrêt des logiciels.

Restauration

- procédures de restauration
  - écrites
  - au résultat validé par les propriétaires des données (pour ne pas en oublier)
  - **testées régulièrement en grandeur réelle** de façon à garantir :
    - que toutes les données pertinentes ont été sauvegardées
    - d'être capable de tout restaurer correctement

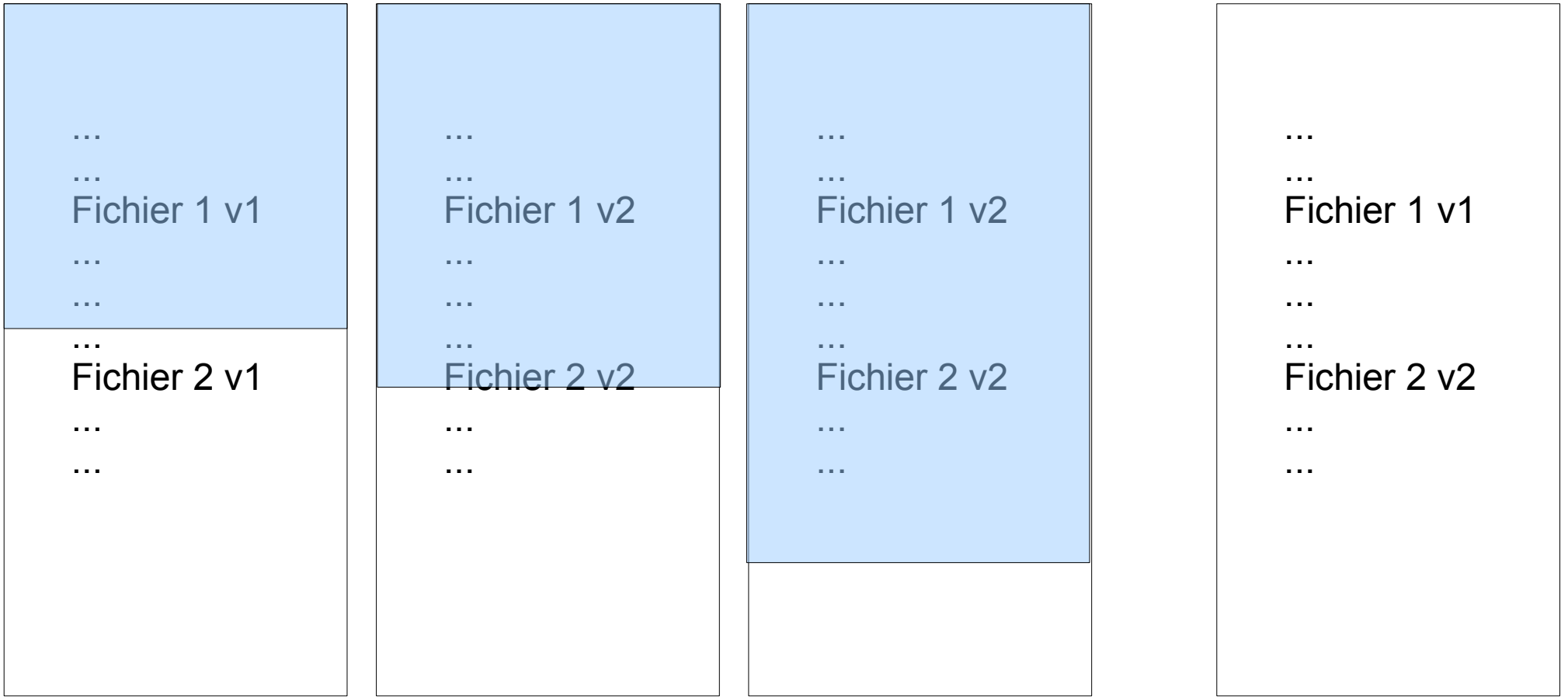
Fiablisation

- utiliser des média fiables
  - éviter disquettes, CD RW, DAT, ..et leur préférer CD R (bof), DLT, AIT, ...
  - varier les marques de media pour palier un défaut de fabrication dans une série, ...)
- gérer le vieillissement des média de sauvegarde/archivage (reprise sur des média récents, ...)

Fiabilisation (2)

- fiabiliser l'environnement des media de sauvegarde:
  - inondation, incendie, vol (coffre ignifugé, ...)
  - sauvegarde hors site (penser à la confidentialité des données, au risque de vol, à l'interception des données, ...)
- indexer les données pour s'y retrouver dans le volume total des jeux de sauvegardes
  - indexer les données
  - étiqueter les media

Sauvegarde live: problématique



Si l'application modifie plusieurs fichiers durant la sauvegarde, les versions sauvées peuvent ne pas être cohérentes.

Sauvegarde

Sauvegarde live: solutions

- Arrêt de l'application pendant la sauvegarde:
  - garantit de plus qu'aucun verrou n'empêchera l'accès à un fichier
- Instantané (Snapshot): image des blocs du sgf, en cas d'effacement de fichier, les blocs ne sont pas réalloués tant que la sauvegarde n'est pas finie
  - Suppose un support dans le sgf
  - Proposé par les NAS, par afs, LVM, UFS2 (FreeBSD, ...), ...
  - Application supplémentaire : proposer aux utilisateurs une image des états antérieurs de leur compte à moindre coût.

Sauvegarde live: snapshot

- Là, on mettra une illustration du fonctionnement des snapshot
- Fait en live au tableau.

Sauvegardes incrémentales/différentielles

- Sauvegarde incrémentales : fichiers créés ou modifiés depuis la sauvegarde précédente
  - Diminue le volume à sauver
  - Restauration nécessite toutes les sauvegardes, restaure toutes les versions d'un même fichier
  - Peu adapté si la totalité des fichiers changent
- Sauvegarde différentielle : fichiers créés/modifiés depuis la dernière sauvegarde de référence
  - Diminue le volume à sauver
  - Plus de volume qu'en incrémental
  - Restauration nécessite la sauvegarde de référence et la dernière sauvegarde différentielle

Exemple:

- Comparer la taille des sauvegardes et les procédures de restauration dans les cas suivants : (ST: sauvegarde totale, I: sauvegarde incrémentale, D: sauvegarde Différentielle)
- Cas 1) ST, I1, I2, I3, I4, I5, I5
- Cas 2) ST, D1, D2, D3, D4, D5
- Cas 3) ST, I1, I2, I3, D1, I4, I5

Dump: un outil de sauvegarde sous unix

- Dump est un outil unix
- Il est efficace (travail directement au niveau du sgf)
- Sauvegarde non portable d'un unix à un autre (lié au sgf)
- Niveau (0 à 9) permettant un gestion très souples des sauvegarde incrémentales/differentielles:
  - Une sauvegarde niveau *n* sauvegarde tous les fichiers modifiés depuis la dernière sauvegarde de niveau *n-1*

Syslog

- syslogd: daemon chargé de gérer les journaux d'une machine
  - journaux: /var/log/*.log (en général)
- peut gérer les journaux d'hôtes distants
  - option « -r » à positionner explicitement
  - rfc 3164: BSD Syslog protocol
  - udp port 514
  - supporté par de nombreux type d'équipement réseau: un standard incontournable

Syslog

- sécurité:
  - pas d'authentification, de filtrage des sources,
  - pas de chiffrement des informations
  - udp: non connecté, pas d'assurance de délivrance

Syslog

- gestion des journaux:
  - gaffe classique: un disque plein à cause de journaux accumulés
  - outils de gestion des journaux : logrotate, newsyslog, ...: compresser, déplacer, effacer, ...

Syslog

- analyse des journaux:
  - pour détecter un problème et/ou en déterminer les causes **après coup**
  - pour alerter d'un problème **en cours**
  - des rapport d'analyse de journaux trop long ne sont pas (plus) lus. Il faut :
    - réagir rapidement aux choses graves
    - extraire les informations pertinentes de la masse d'information
  - Deux types d'outils
    - outils d'analyse de journaux: logcheck, logsurfer, swatch, sec, ...
    - via un ids: système de détection d'intrusion

syslog-ng:

- configuration plus souple
- classement des messages par leur contenu, par l'hôte d'origine
- meilleure redirection des messages sur le réseau
- possibilité de chroot
- peut utiliser UDP et TCP
- chiffrement et authentification du trafic réseau
- portable
- export des journaux vers un sgbd

configuration: syslog.conf

- facilité.niveau<tab>action
- facilité: type de service source

Action
fichier
terminal
pipe |
@machineDistante
utilisateur1,utilisateur2,...
*

Niveau	
emerg (panic)	Situations de panique.
alert	Situations urgentes.
crit	Situations critiques.
err (error)	Erreurs.
warning (warn)	Messages de WARNING.
notice	Messages divers.
info	Messages d'informations.
debug	Débogage.

Facilités	
kern	Le noyau.
user	Process des utilisateurs.
mail	Système de courrier.
daemon	Démons systèmes.
auth	Authentification.
lpr	Système de spooling d'imprimante.
news	Usenet.
uucp	UUCP.
cron	Démon cron.
mark	Messages generes a intervals réguliers.
local0-7	Huit niveaux de messages locaux.
syslog	Messages internes a syslogd.
authpriv	Messages privés auth.
*	Toutes les facilités sauf mark.

Syslog : demo

- lister le syslog d'un système existant
- lister un journal de /var/log, montrer les entrées "MARK" insérées par syslogd
- tester son comportement avec la commande logger
  - logger -p mail.crit "boîte au lettre en feu :-) »
  - logger -p news.err "pas de nouvelles, bonne nouvelle"
  - comparer l'effet avec le contenu de syslog.conf et notamment que le message est stocké si son niveau est supérieur ou égal à celui de la règle
- le modifier en y insérant une entrée
- tester l'entrée insérée avec logger

Bibliographie sur la supervision et sur syslog

- « unix, guide de l'administrateur » de Nemeth, Snyder & Al, Campus press
- « MISC No 22 » (revue): superviser sa sécurité
- Ntsyslog: http://ntsyslog.sourceforge.net/
- http://www.linux-kheops.com/line/html/line/line-dec1996/datas/syslog.htm
-

comptes utilisateurs: création

- uid
- modifier /etc/passwd & Co
- mot de passe
- dossier personnel
- fichier d'initialisation dans $HOME
- donner les bons droit au dossier perso (chgrp, chown)
- déclarer l'utilisateur dans les services usuels (mail, ...)
- tester le compte

comptes utilisateurs

- structure d'un fichier /etc/passwd
- passwd: pour changer son mot de passe
- shadows passwords: /etc/shadow
- commande d'administration :
  - dépend du système d'exploitation
  - exemples:
    - useradd/adduser
    - userdel

groupes

- /etc/group
- chaque utilisateur a un groupe initial (/etc/passwd) et des groupes secondaires (/etc/group)
- groups: liste les groupes de l'utilisateur
- groupes sous BSD:
  - l'utilisateur appartient à tous les groupes
  - création de dossier/fichier: groupe du dossier père
  - gestion des groupes: pw (création/suppression, ajout d'utilisateurs, ...)

groupes

- groupe sous SysV et Linux
  - l'utilisateur appartient à un instant donné à un seul groupe => newgrp pour changer de groupe
  - création de dossier/fichier: groupe du dossier père ou groupe de l'utilisateur (Linux, autorisé par SysV)
  - gestion des groupes
    - groupadd, groupmod, groupdel: ajout/suppression de groupes
    - usermod -G group,... login: ajoute login au(x) groupe(s)

planification de tâches: cron et atd

- cron: tâches planifiées régulières
- atd: exécution unique
- cron et arrêt systèmes/chgt d'heures
- commande crontab:
  - crontab -l : lister
  - crontab -r : supprimer
  - crontab -e : modifier
- dossier daily, monthly, ...: (dépend de l'OS)

format du fichier crontab

- régles communes:
  - # en début de ligne indique un commentaire
  - les champs sont séparés par des espaces
  - les espaces de la commandes sont laissés inchangés. commande exécutée par sh
  - dans la commande, % indique un saut de ligne
  - contenu des champs :
    - *, entier, entier-entier, des entiers/intervalles séparés par des virgules
- crontab utilisateur :

```
minute heure jourDuMois jourDeLaSemaine commande
```

- crontab sytème (souvent : /etc/crontab)

```
minute heure jourDuMois jourDeLaSemaine utilisateur
  commande
```

crontab: exemples

- commandes valides :

```
echo date courante: `date` >> /tmp/test
mutt -s "coucou Pascal" petit@shayol.org % coucou
  % courrier de test
find / -xdev -name core -atime +7 -exec /bin/rm
  -f {} \;
```

- spec de temps valides:

```
*0 * * * : toutes les 10 mn
10 2 * * : tous les jours à 2h10
0 23 * 0 : tous les dimanches à 23h00
0 20-23,0-7,10,12,14,16,18 * * * : toutes les
  heures entre 20h00 et 7h00 puis toutes les deux
  heures
```

cron : sécurité

- contrôle d'accès :
  - cron.allow: seuls utilisateurs habilités à programmer des tâches
  - cron.deny: seuls utilisateur NON autorisés à programmer des tâches (suppose l'absence de cron.allow)
  - si ni cron.(allow|deny): seul root y a droit
- contrôle d'accès réalisé par la commande crontab
  - => les fichiers crontab doivent avoir les bons droits