

Automatisation des sauvegardes “système”

Francis COUGARD

Introduction

Le besoin de systèmes de sauvegardes client-serveur automatisées se fait de plus en plus sentir dans les exploitations informatiques et s'étend aux sauvegardes dites “système” (sauvegardes permettant de recharger un système d'exploitation).

Le but est de pouvoir recharger en un temps minimal un serveur AIX en cas de crash partiel (changement nécessaire de disques internes) ou complet (changement complet du serveur par un serveur d'architecture analogue).

Les critères conduisant à une solution automatisée des sauvegardes système sont les suivants :

- Réduction des interventions en salle machine (montage-démontage de cassettes).
- Utilisation croissante de solutions de sauvegarde client-serveur, comme TSM, avec bibliothèque de stockage centralisée.
- Existence de réseaux rapides (ATM, FDDI, Ethernet 100 Mb et 1 Gb) permettant d'envisager des sauvegardes de volumes importants (environ 1 Go) par réseau en quelques minutes.

La solution standard préconisée par IBM est l'utilisation d'un serveur AIX NIM (*Network Installation Management*), système AIX installé sur le réseau local dans lequel se trouvent les systèmes à sauvegarder, qui sera capable, à tout moment, de recharger le système d'exploitation d'un serveur AIX par téléchargement réseau.

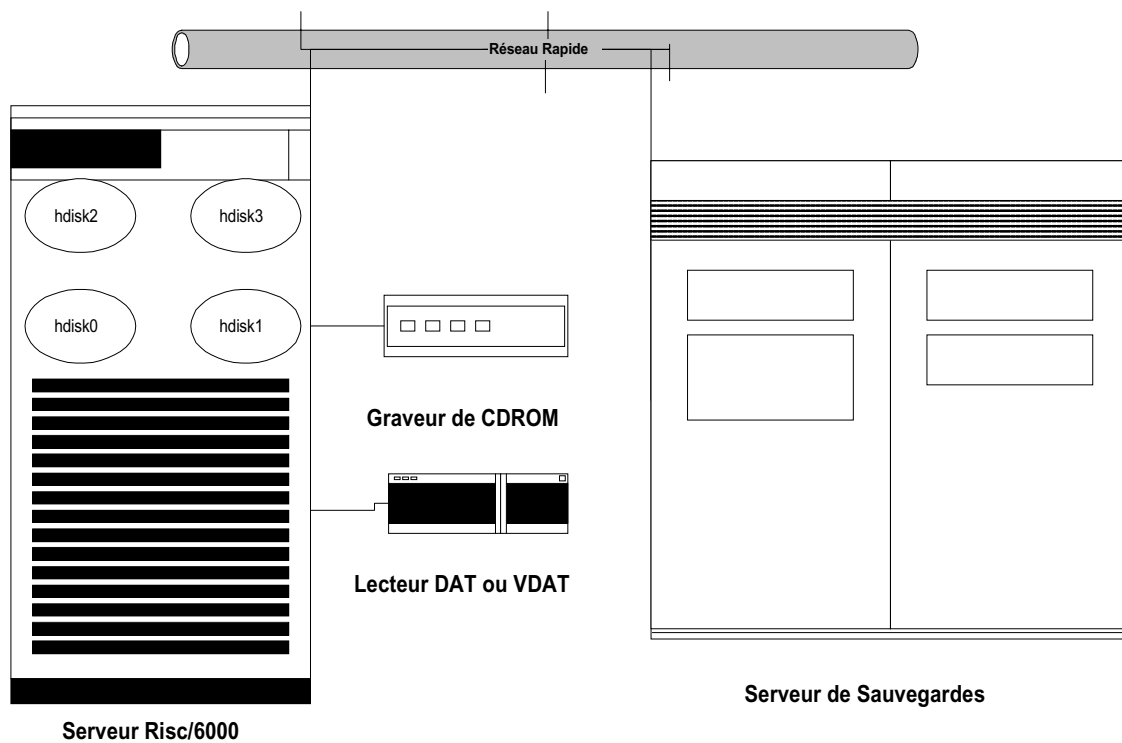
Toutefois, cette solution n'est économiquement envisageable que lorsque le nombre de serveurs à sauvegarder dépasse deux ou trois systèmes. C'est pourquoi nous vous proposons une solution alternative utilisant :

- la commande “ alt_disk_install ”, apparue en AIX 4.3.3, permettant de créer des noyaux AIX alternés.
- la commande “ mkcd ”, permettant de réaliser des sauvegardes “système” amorçables sur CDROM.

Configuration requise

- Un serveur Risc System/6000 équipé de :
 - 4 disques internes de taille ≥ 2 Go
 - AIX 4.3.3
 - Une carte réseau permettant un lien TCP/IP rapide vers le serveur de sauvegardes et sa robotique.
 - Un lecteur, interne ou externe, DAT / VDAT / DLT
 - Un graveur de CDROM externe en option.

Schéma de configuration



Configuration des unités de disques

- Le système d'exploitation est installé sur les disques “ **hdisk0** ” et “ **hdisk2** ” :
 - Ces deux disques constituent le groupe de volumes “ **rootvg** ”

- On fait du *mirroring* sur le groupe de volumes “ **rootvg** ” :
 - ◆ chaque volume logique dispose d’une copie logique sur le disque hdisk0 ET d’une copie logique sur le disque hdisk2
 - ◆ SAUF le volume de *dump* qui est dupliqué (hd7 dump primaire sur hdisk0 et hd71 dump secondaire sur hdisk2)
- Un groupe de volumes “ **backupvg** ” est constitué par le disque hdisk1. Il contient uniquement le système de fichiers “ **/backup** ” monté automatiquement lors du *boot* du système.
- Le disque hdisk3 n’est pas affecté à un groupe de volumes.

Planification des sauvegardes

- Chaque soir, à 20h00, une sauvegarde système est lancée et créée dans le système de fichiers **/backup** puis, ensuite, externalisée (dans notre exemple, le produit de sauvegarde utilisé est TSM) :
 - script : **/backup/mksysb.sh**
 - trace d’exécution : **/backup/mksysb.log**

```
#!/bin/ksh
# Ce script crée un backup système sur disque
# et génère la sauvegarde du système de fichier /backup
# via tsm sur système de stockage externe
#
exec 1> /backup/mksysb.log 2>&1
rm /backup/mksysb.image
date
echo "Phase de sauvegarde locale"
/usr/bin/mksysb -i /backup/mksysb.image
cp /image.data /backup
date
echo "Phase d'externalisation de la sauvegarde"
/usr/lpp/adsm/bin/dsmc incr /backup
```

- Chaque vendredi, à 21h00, un noyau AIX est créé à partir de l’image de sauvegarde créée précédemment sur le disque hdisk3. Ce noyau sert à recharger le noyau opérationnel en cas de problème.
Le fichier “ **image.data** ”, servant à générer le noyau alterné, est expurgé des informations concernant le *mirroring* afin de permettre une installation de système sur un seul disque :

- Script : **/backup/minisys.sh**
- Trace d'exécution : **/backup/minisys.log**

```
#!/bin/ksh
# Ce script crée un noyau alterné sur hdisk3 sans mirroring
# à partir d'une image de backup créée précédemment
#
exec 1> /backup/minisys.log 2>&1
cat /backup/image.data | sed -e 's/COPIES= 2/COPIES= 1/' \
    -e 's/hdisk0 hdisk1/hdisk0/' \
    > /backup/image.data.mini
/usr/sbin/alt_disk_install -Pall -B -d /backup/mksysb.image \
    -i /backup/image.data.mini hdisk3
/usr/sbin/alt_disk_install -X
/usr/bin/bootlist -m normal hdisk0 hdisk2
```

Remarque :

Ce mini-noyau n'est utilisable que si le niveau d'AIX du noyau opérationnel à reconstituer en cas de problème est de niveau inférieur ou égal au niveau exact d'AIX du mini-noyau (les commandes “ **lspp -L bos.mp** ” ou “ **lspp -L bos.up** ” permettent d'en connaître le niveau exact).

Il est donc conseillé de régénérer un mini-noyau après toute mise à niveau d'AIX sur le noyau opérationnel.

Sauvegarde optionnelle du mini-noyau sur CDRom

Note : Une sauvegarde sur lecteur dat, vdat ou dlt local reste possible.

Cette sauvegarde permet de reconstituer un système opérationnel, en cas de sinistre incluant la destruction des données du disque hdisk3 contenant le mini-noyau.

Elle doit se faire hors période d'exploitation.

Cette sauvegarde n'est pas automatisée.

Il est conseillé de l'effectuer à chaque changement de niveau de maintenance sous AIX.

Elle n'est possible que si le disque hdisk3 dispose de 1,5 à 2 Go d'espace libre.

Il est donc nécessaire de disposer :

- d'un disque hdisk3 de 4,5 Go,
- d'un lecteur de CDRom, de type SCSI externe, sélectionné dans la liste préconisée par IBM (dans notre cas, le lecteur testé est un Yamaha™ 4416SX),
- du câblage SCSI externe adéquat.

- Faire un *boot* sur `hdisk3` :

```
bootlist -m normal hdisk3  
shutdown -Fr
```

- Après le *boot* :

- Purger le système de manière à ce que sa sauvegarde tienne sur un seul CDROM :
 - ◆ Désinstaller les composants d'AIX non nécessaires (X11, documentation en ligne, ...).
 - ◆ Supprimer les fichiers de trace de taille trop importante.
 - ◆ Supprimer les systèmes de fichiers optionnels non utiles dans `rootvg`.

Cette opération de “nettoyage” doit être faite en pensant à conserver :

- ◆ Un noyau AIX opérationnel avec sa configuration réseau TCP/IP.
- ◆ L'agent de sauvegarde (TSM dans notre cas).

- Installer le logiciel du domaine public “`cdrecord`” :

`cdrecord 1.9` peut être trouvé au format “installp” sur le site :

<http://ftp.univie.ac.at/aix/>

Il suffit ensuite d'installer ce *filesset* (après décompression)

et

d'exécuter les deux commandes suivantes afin que soit possible la création d'un CDROM *bootable* pour le *backup* système :

```
ln -s /usr/samples/oem_cdwriters/mkrr_fs_gnu /usr/sbin/mkrr_fs  
ln -s /usr/samples/oem_cdwriters/burn_cd_gnu /usr/sbin/burn_cd
```

- Lancer la sauvegarde sur CDROM :

```
mkcd -d /dev/cd1 (si le graveur est détecté en tant que /dev/cd1)
```

Cette sauvegarde est longue car la gravure du CDROM est faite à la vitesse x1, quelle que soient les performances du graveur SCSI (prévoir de une à deux heures de traitement).

Cette sauvegarde nécessite un espace disque disponible important sur `hdisk3` (1,5 Go) car elle se compose des phases suivantes :

- ◆ création d'une image *bootable* sur disque,
- ◆ création d'une structure de système de fichiers contenant l'image complète,
- ◆ création d'un fichier image ISO9660 avant gravure,
- ◆ gravure.

- Redémarrer le système sur le noyau opérationnel :

```
bootlist -m normal hdisk0 hdisk2
shutdown -Fr
```

Remarque :

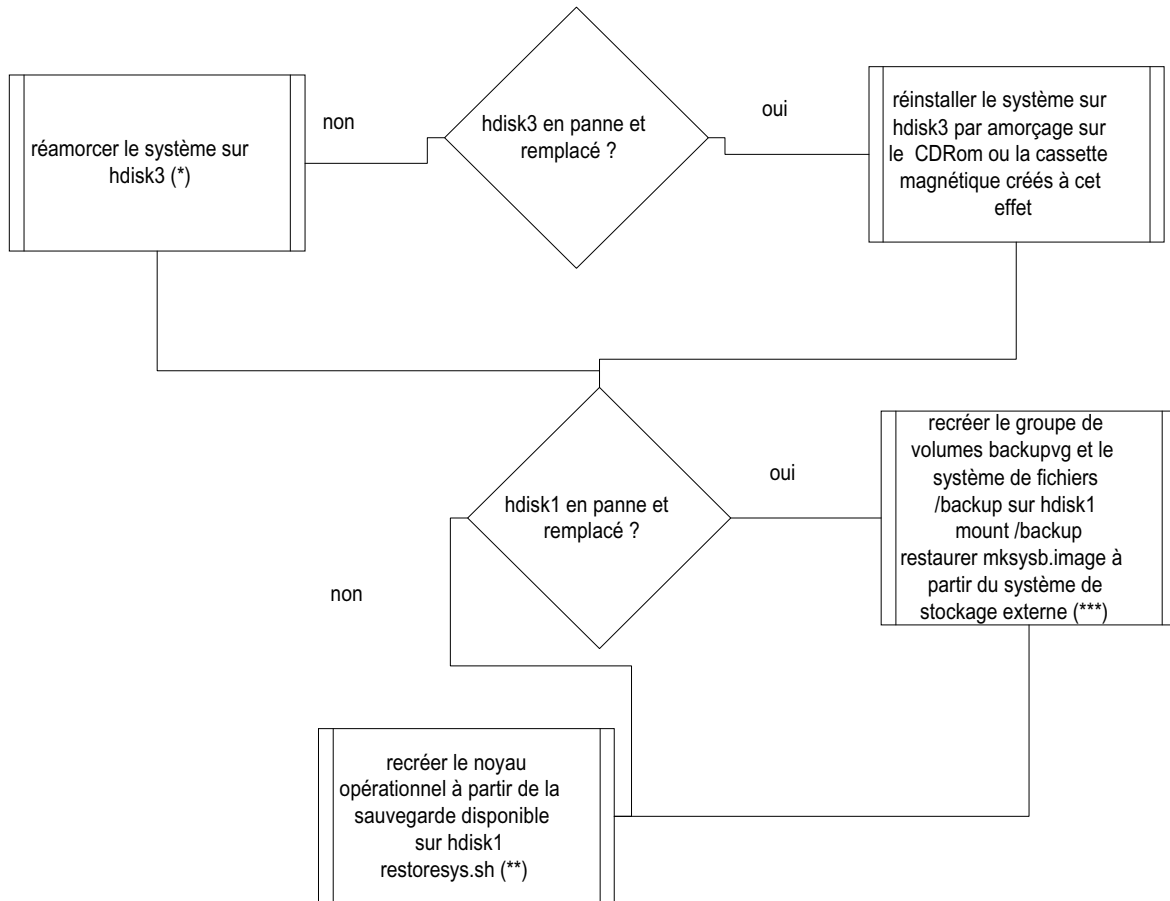
Ce CDRom ne peut être utilisé que pour un rechargement de système d'architecture analogue au système source (" **bootinfo -T** " renseigne sur ce type d'architecture).

Stratégie de restauration

Nous disposons, à tout moment durant la période d'exploitation :

- D'une sauvegarde système datant de la veille à 20 heures :
 - ◆ sur le disque hdisk1
 - ◆ ou, le cas échéant, sur un système de stockage externe.
- D'un mini-noyau, créé le vendredi précédent, opérationnel sur :
 - ◆ hdisk3
 - ◆ ou, le cas échéant, sur :
 - ◇ CDRom amorçable,
 - ◇ sur cassette amorçable 4mm / 8mm ou DLT.
- Supposons que, malgré le *mirroring*, une panne rende les disques système hdisk0 et hdisk2 indisponibles (le cas du remplacement complet du serveur par un serveur de même architecture correspond strictement au cas où tous les disques ont été remplacés).
- Supposons également que tous les disques défectueux aient été changés au préalable et remplacés par des disques vierges.

Les actions de la figure suivante sont alors applicables selon les hypothèses :



- (*) Certains anciens systèmes (rs6k à bus mca) ne permettent pas le choix du disque de *boot* en *stand-alone*.
- (**) Procédure explicitée page suivante et suivie d'un réamorçage du système complet.
- (***) Procédure explicitée page suivante dans le cas de TSM.

Comment recréer le noyau opérationnel

- Faire un *boot* sur le mini-noyau puis utiliser le *script* de restauration :
 - Script : **`/backup/restoresys.sh`**
 - Trace d'exécution : **`/backup/restoresys.log`**

```
#!/bin/ksh
# Ce script restaure un noyau complet et opérationnel à partir de la
# dernière sauvegarde présente sur hdisk1
#
#
exec 1> /backup/restoresys.log 2>&1
/usr/sbin/alt_disk_install -X old_rootvg
/usr/sbin/alt_disk_install -Pall -B -d /backup/mksysb.image \
                        hdisk0 hdisk1
/usr/sbin/alt_disk_install -X
/usr/bin/bootlist -m normal hdisk0 hdisk2
```

Remarque :

Après avoir fait un *reboot* du noyau opérationnel sur hdisk0, il est nécessaire de nettoyer, dans l'ODM (*Object Database Management*), toute trace du mini-noyau par la commande :

alt_disk_install -X old_rootvg

Comment retrouver la dernière image système valide sur le système de stockage externe

Il s'agit du cas où le disque hdisk1 a été remplacé :

- Recréer le groupe de volumes “ **backupvg** ” sur le disque hdisk1.
- Recréer le système de fichiers “ **/backup** ” dans ce groupe de volumes et le monter.
- Passer la commande de restauration adaptée au logiciel de sauvegarde utilisé.
Dans notre cas (TSM) : **dsmc restore /backup**

Conclusion

Cette méthode met en œuvre des mécanismes plus simples que NIM.

Elle s'avère utile dans la cas où on ne peut accepter le rechargement d'un système à partir d'une sauvegarde ancienne (système subissant des mises à niveau ou des changements de paramétrage fréquents).

Une variante simplifiée est envisageable si l'on dispose de 3 disques, au lieu de 4, sans *mirroring* de rootvg (clonage du noyau).

Dans le cas d'un sinistre partiel (présence d'un mini-noyau opérationnel et d'une image de *backup* récente sur un disque local), la méthode décrite précédemment peut s'avérer très rapide. Toutefois, NIM est recommandé dans un environnement multi-machines où le retour à une situation opérationnelle est plus rapide dans le cas d'un sinistre complet. ■