

# Mise en place d'un serveur de sauvegarde à moindre coût

Yann MORÈRE

29 septembre 2004

## Résumé

Ce document propose la mise en place d'un serveur de sauvegarde simple, à l'aide d'outils libres et de quelques disques durs. Ce n'est pas effectivement une sauvegarde au sens pur du terme, mais cela permet une synchronisation automatique des fichiers à partir de postes Windows et Linux sur une machine distante avec une connexion sécurisée, sans l'intervention de l'utilisateur d'une part, et d'autre part un «*mirroring*» sur deux autres disques sur cette même machine distante.

## 1 Installation et utilisation de cwrsrc sous windows

L'utilitaire **rsync** va permettre de réaliser les sauvegardes sur le serveur. Cette sauvegarde sera intelligente car la commande ne fait que synchroniser les fichiers entre la machine et le serveur de sauvegarde. Seule la première sauvegarde transférera l'intégralité des données. Par la suite, seuls les fichiers modifiés entre deux sauvegardes seront transférés.

Cet utilitaire est disponible en standard sur les systèmes Unix, mais n'existe pas sur les systèmes Windows. Il faut pour cela, soit utiliser les outils Cygwin en installant l'environnement, soit en récupérant seulement l'utilitaire **rsync**.

Ce dernier est disponible sur le site <http://www.itefix.no/phpws/index.php>, il s'agit de cwrsrc qui se présente sous la forme d'un installable.

Vous pouvez aussi télécharger le paquet suivant qui contient tout ce qui est nécessaire pour une installation sous windows (script d'installation, **scp**, exemple de fichier de configuration de sauvegarde) **lcwrsrc.zip**. En tant qu'administrateur, installer le logiciel, en gardant les options par défaut.

Dans une fenêtre de commande (démarrer->exécuter-> taper cmd puis ok) on va générer les fichiers nécessaires à la connexion.

### 1.1 Génération de la clé DSA pour authentification

```
c:\cwrsrc\ssh-keygen -t dsa -f c:\cwrsrc\id_dsa
```

Un jeu de clé publique et privée est généré et est copié dans le répertoire **cwrsrc**.

Si aucun fichier n'est spécifié, les clés sont copiées dans le répertoire **/cygdrive/c/Documents and Settings/user/.ssh/** et **/cygdrive/c/Documents and Settings/user/.ssh/id\_dsa.pub** par défaut.

Afin de mieux sécuriser le système il est possible de donner une «*passphrase*» lors de la création des clés.

### 1.2 Création de l'utilisateur de sauvegarde sur le serveur (sous linux)

sur le serveur Unix :

```
adduser morere #ajoute l'utilisateur morere
smbpasswd -a morere # ajoute l'utilisateur samba morere
su morere
ssh-keygen -t dsa
```

La dernière commande génère les clés du côté serveur et crée le répertoire **.ssh**.

### 1.3 Connexion automatique sur le serveur avec ssh

On va copier la clé publique du client sur le serveur dans le répertoire de `user/.ssh` afin de réaliser une connexion automatique.

La commande `scp`, n'existe pas non plus disponible en standard sous windows. Putty fournit la commande nécessaire par `pscp` disponible sur <http://www.chiark.greenend.org.uk/~sgtatham/putty/>.

*Attention, la version 0.52 pose des problèmes avec XP, j'ai utilisé la version 0.55.*

```
pscp "C:\Documents and Settings\user\.ssh\id_dsa.pub"
morere@1974.57.140.212:~/.ssh/authorized_keys
```

Un test de connexion est réalisé par la commande

```
ssh morere@194.57.140.212
```

ou dans le cas ou on spécifie le nom de fichier de clé :

```
pscp "C:\cwrsync\id_dsa.pub" morere@1974.57.140.212:~/.ssh/authorized_keys
```

Un test de connexion est réalisé par la commande

```
ssh -i c:\cwrsync\id_dsa morere@194.57.140.212
```

et la connexion devrait être automatique. L'option `-i` permet de spécifier ou `ssh` doit aller chercher la clé privée pour valider la connexion.

Dans le cas ou plusieurs machines se connectent sur le même compte pour faire des sauvegardes, il faut concaténer à la main les clés publiques dans le fichier `$HOME/.ssh/authorized_keys` sur le serveur.

### 1.4 Sauvegarde par rsync

Maintenant il faut envoyer par `rsync` les fichiers sur le serveur.

Il existe un problème avec les noms de fichiers longs et les `C:\Documents and Settings\user` par exemples, que `rsync` refuse.

Une première astuce est de se placer dans le répertoire qui convient, et d'appeler la commande `rsync` avec son chemin. Voici un exemple de fichier «sauv.bat» que l'on place dans le répertoire `.ssh` de l'utilisateur.

```
@ECHO OFF
cd C:\Documents and Settings\user
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -i c:\cwrsync\id_dsa -l morere"
"Bureau" 194.57.140.212:backup
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -i c:\cwrsync\id_dsa -l morere"
"Favoris" 194.57.140.212:backup
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -i c:\cwrsync\id_dsa -l morere" "Mes
documents" 194.57.140.212:backup
```

Une autre astuce, beaucoup plus propre est de passer par l'utilisation de «/cygdrive/c/» qui remplace «c:\» et permet d'utiliser des chemins absolus.

```
@ECHO OFF
echo utilisation de /cygdrive/c/Documents and Settings/user/
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -i c:\cwrsync\id_dsa -l morere"
"/cygdrive/c/Documents and Settings/user/Bureau" 194.57.140.212:backup
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -l morere"
"/cygdrive/c/Documents and Settings/user/Favoris" 194.57.140.212:backup
c:\cwrsync\rsync -av --rsh="c:\cwrsync\ssh -l morere"
"/cygdrive/c/Documents and Settings/user/Mes documents" 194.57.140.212:backup
```

Ces actions permettent de sauver les répertoires `Bureau`, `Favoris` et `Mes Documents` sur le serveur de sauvegarde dans le répertoire `backup`.

`Rsync` possède de nombreuses options listées ci-dessous

-v, --verbose	increase verbosity
-q, --quiet	decrease verbosity
-c, --checksum	always checksum
-a, --archive	archive mode, equivalent to -rlptgoD
-r, --recursive	recurse into directories
-R, --relative	use relative path names
--no-relative	turn off --relative
--no-implied-dirs	don't send implied dirs with -R
-b, --backup	make backups (see --suffix & --backup-dir)
--backup-dir	make backups into this directory
--suffix=SUFFIX	backup suffix (default ~ w/o --backup-dir)
-u, --update	update only (don't overwrite newer files)
-l, --links	copy symlinks as symlinks
-L, --copy-links	copy the referent of all symlinks
--copy-unsafe-links	copy the referent of "unsafe" symlinks
--safe-links	ignore "unsafe" symlinks
-H, --hard-links	preserve hard links
-p, --perms	preserve permissions
-o, --owner	preserve owner (root only)
-g, --group	preserve group
-D, --devices	preserve devices (root only)
-t, --times	preserve times
-S, --sparse	handle sparse files efficiently
-n, --dry-run	show what would have been transferred
-W, --whole-file	copy whole files, no incremental checks
--no-whole-file	turn off --whole-file
-x, --one-file-system	don't cross filesystem boundaries
-B, --block-size=SIZE	checksum blocking size (default 700)
-e, --rsh=COMMAND	specify the remote shell
--rsync-path=PATH	specify path to rsync on the remote machine
--existing	only update files that already exist
--ignore-existing	ignore files that already exist on receiver
--delete	delete files that don't exist on sender
--delete-excluded	also delete excluded files on receiver
--delete-after	receiver deletes after transfer, not before
--ignore-errors	delete even if there are I/O errors
--max-delete=NUM	don't delete more than NUM files
--partial	keep partially transferred files
--force	force deletion of dirs even if not empty
--numeric-ids	don't map uid/gid values by user/group name
--timeout=TIME	set I/O timeout in seconds
-I, --ignore-times	turn off mod time & file size quick check
--size-only	ignore mod time for quick check (use size)
--modify-window=NUM	compare mod times with reduced accuracy
-T --temp-dir=DIR	create temporary files in directory DIR
--compare-dest=DIR	also compare received files relative to DIR
--link-dest=DIR	create hardlinks to DIR for unchanged files
-P	equivalent to --partial --progress
-z, --compress	compress file data
-C, --cvs-exclude	auto ignore files in the same way CVS does
--exclude=PATTERN	exclude files matching PATTERN
--exclude-from=FILE	exclude patterns listed in FILE
--include=PATTERN	don't exclude files matching PATTERN
--include-from=FILE	don't exclude patterns listed in FILE
--files-from=FILE	read FILE for list of source-file names
-0 --from0	all file lists are delimited by nulls
--version	print version number
--daemon	run as an rsync daemon

```

--no-detach          do not detach from the parent
--address=ADDRESS    bind to the specified address
--config=FILE         specify alternate rsyncd.conf file
--port=PORT          specify alternate rsyncd port number
--blocking-io        use blocking I/O for the remote shell
--no-blocking-io     turn off --blocking-io
--stats              give some file transfer stats
--progress           show progress during transfer
--log-format=FORMAT   log file transfers using specified format
--password-file=FILE  get password from FILE
--bwlimit=KBPS        limit I/O bandwidth, KBytes per second
--write-batch=PREFIX write batch fileset starting with PREFIX
--read-batch=PREFIX   read batch fileset starting with PREFIX
-h, --help           show this help screen

```

Ce qu'il faut retenir c'est que `-a` remplace `-rlptgoD` qui est à peu près tout ce dont on a besoin. On peut ajouter `-v` pour avoir le mode verbeux.

`-a, --archive`

Ceci est équivalent à `-rlptgoD`. C'est un moyen rapide de dire que vous voulez les sous répertoire en préservant tout.

Pour aller plus loin, il faudrait stocker dans un fichier les répertoires à sauver et les répertoires à exclure. Ceci peut être fait avec les options

`--exclude=PATTERN`

This option allows you to selectively exclude certain files from the list of files to be transferred. This is most useful in combination with a recursive transfer.

You may use as many `--exclude` options on the command line as you like to build up the list of files to exclude.

See the EXCLUDE PATTERNS section for detailed information on this option.

`--exclude-from=FILE`

This option is similar to the `--exclude` option, but instead it adds all exclude patterns listed in the file FILE to the exclude list. Blank lines in FILE and lines starting with `';`' or `'#'` are ignored. If FILE is `-` the list will be read from standard input.

`--include=PATTERN`

This option tells rsync to not exclude the specified pattern of filenames. This is useful as it allows you to build up quite complex exclude/include rules.

See the EXCLUDE PATTERNS section for detailed information on this option.

`--include-from=FILE`

This specifies a list of include patterns from a file. If FILE is `-` the list will be read from standard input.

`--files-from=FILE`

Using this option allows you to specify the exact list of files to transfer (as read from the specified FILE or `-` for stdin). It also tweaks the default behavior of rsync to make transferring just the specified files and directories easier. For instance, the `--relative` option is enabled by default when this option is used (use `--no-relative` if you want to turn that off), all directories specified in the list are created on the destination (rather than being noisily skipped without `-r`), and the `-a` (`--archive`) option's behavior does not imply `-r` (`--recursive`) -- specify it explicitly, if you want it.

The file names that are read from the FILE are all relative to the source dir -- any leading slashes are removed and no `".."` references are allowed to go higher than the source dir. For example, take this command:

```
rsync -a --files-from=/tmp/foo /usr remote:/backup
```

If /tmp/foo contains the string "bin" (or even "/bin"), the /usr/bin directory will be created as /backup/bin on the remote host (but the contents of the /usr/bin dir would not be sent unless you specified -r or the names were explicitly listed in /tmp/foo). Also keep in mind that the effect of the (enabled by default) --relative option is to duplicate only the path info that is read from the file -- it does not force the duplication of the source-spec path (/usr in this case).

In addition, the --files-from file can be read from the remote host instead of the local host if you specify a "host:" in front of the file (the host must match one end of the transfer). As a short-cut, you can specify just a prefix of ":" to mean "use the remote end of the transfer". For example:

```
rsync -a --files-from=:/path/file-list src:/ /tmp/copy
```

This would copy all the files specified in the /path/file-list file that was located on the remote "src" host.

Il est possible de faire plus générique, en plaçant dans des fichiers la liste des choses que l'on veut sauver et ce que l'on veut éviter.

Dans notre cas, on va sauver les répertoires /cygdrive/c/Documents and Settings/user/, c'est à dire les données de l'utilisateur user en évitant de sauver les fichiers inutiles comme Application Data. fichier sauv.bat

@ECHO OFF

```
c:\cwrsync\rsync -av --include-from="c:\cwrsync\include.txt" --exclude-from="c:\cwrsync\exclude.txt" --rsh="c:\cwrsync\ssh -l yannsauv" "/cygdrive/c/Documents and Settings/Administrateur/" 194.57.140.212:backup_admin
```

```
yann@tuxpowered:~/temp/cwrsync$ more include.txt
```

Favoris

Mes documents

Bureau

```
yann@tuxpowered:~/temp/cwrsync$ more exclude.txt
```

.ssh

Application Data

Local Settings

Modèles

ntuser.ini

Voisinage réseau

Recent

SendTo

Voisinage d'impression

ntuser.dat.LOG

NTUSER.DAT

```
yann@tuxpowered:~/temp/cwrsync$
```

voici un autre exemple plus récent : fichier sauv.bat

@ECHO OFF

```
rem Sauvegarde du rep Documents and settings
```

```
c:\cwrsync\rsync -av --include-from="c:\cwrsync\includec.txt" --exclude-from="c:\cwrsync\excludec.txt"
```

```
rem Sauvegarde du rep utilisateur
```

```
c:\cwrsync\rsync -av --include-from="c:\cwrsync\included.txt" --exclude-from="c:\cwrsync\excluded.txt"
```

fichier excludec.txt

.ssh

Local Settings

Modèles

ntuser.ini

```

NTUSER.DAT
NTUSER.DAT.LOG
Voisinage réseau
Voisinage d'impression
SendTo
Menu Démarrer
Recent

```

fichier excluded.txt

```

HTML
devis_commandes
doc_materiel
fond d'ecran
icônes
image_skel
sites_web
temp

```

Afin de faire une installation *quasi* automatique sur les postes clients, il est possible d'utiliser le petit script suivant (fichier `install_user.bat`) :

```

@ECHO OFF
echo Installation pour utilisateur %1
echo Génération des clés
c:\cwrsync\ssh-keygen -t dsa -f c:\cwrsync\id_dsa
echo Copie des clés sur le serveur
C:\cwrsync\scp "c:\cwrsync\id_dsa.pub" %1@194.57.140.212:.ssh/authorized_keys
echo Test de connexion automatique
c:\cwrsync\ssh -i c:\cwrsync\id_dsa %1@194.57.140.212

```

L'utilisation est assez simple, il suffit d'appeler le script avec le login de l'utilisateur à créer :

```
install_user morere
```

Il suffit alors de répondre aux questions posées par les différents programmes.

De même, si vous avez deux disques durs locaux, il est possible de réaliser une sauvegarde synchronisée, toujours en utilisant `cwrsync`.

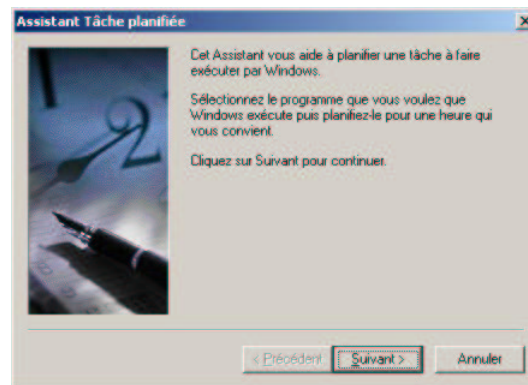
Le fichier `sauv_local.bat` permet de faire cela :

```

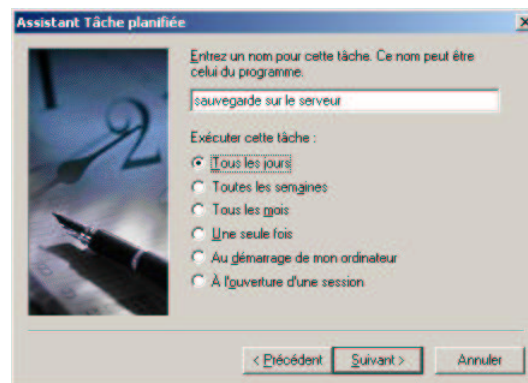
@ECHO OFF
echo Sauvegarde du rep Documents and settings
c:\cwrsync\rsync -av --include-from="c:\cwrsync\includec.txt"
--exclude-from="c:\cwrsync\excludec.txt" "/cygdrive/c/Documents and
Settings/yann/" "/cygdrive/h/utilisateurs/yann_windows/documents_and_settings"
echo Sauvegarde du rep utilisateur
c:\cwrsync\rsync -av --include-from="c:\cwrsync\included.txt"
--exclude-from="c:\cwrsync\excluded.txt" "/cygdrive/d/utilisateurs/yann/"
"/cygdrive/h/utilisateurs/yann_windows/"

```

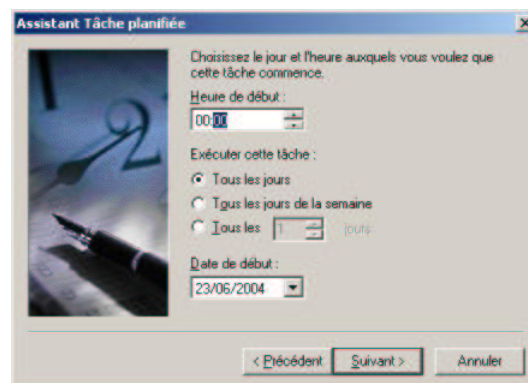
Ensuite il est possible de réaliser cela de manière automatique, grâce au «Tâches Planifiées sous windows». Menu Démarrer->Paramètres->Panneau de configuration...



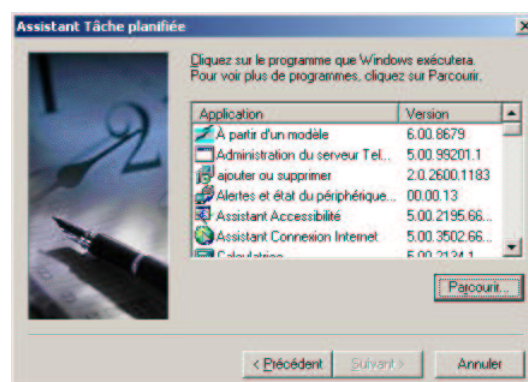
Ensuite double click sur Tâche planifiée, puis parcourir



Choisir le nom de la tâche et la fréquence d'exécution puis suivant

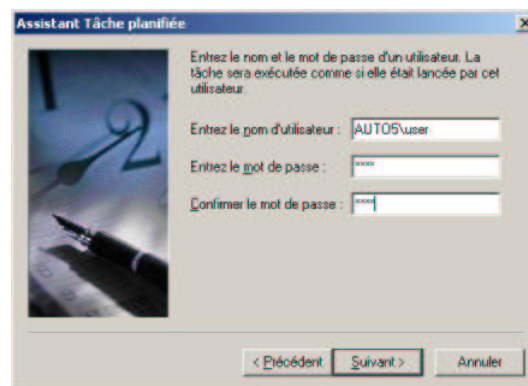
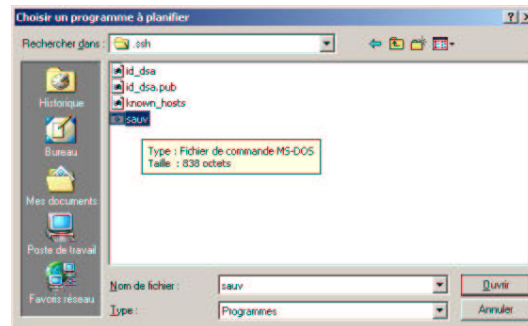


choisir l'heure de démarrage

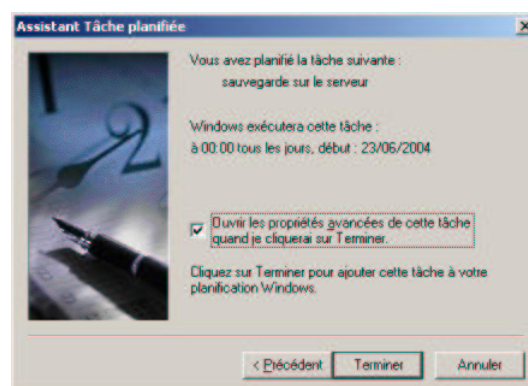




Choisir le programme qui doit être lancé par le planificateur. Dans notre cas il s'agira d'un fichier **bat** qui contiendra les commandes de sauvegarde.

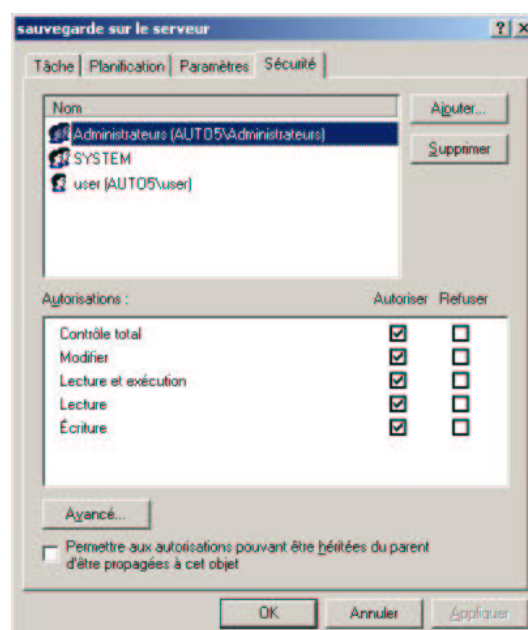
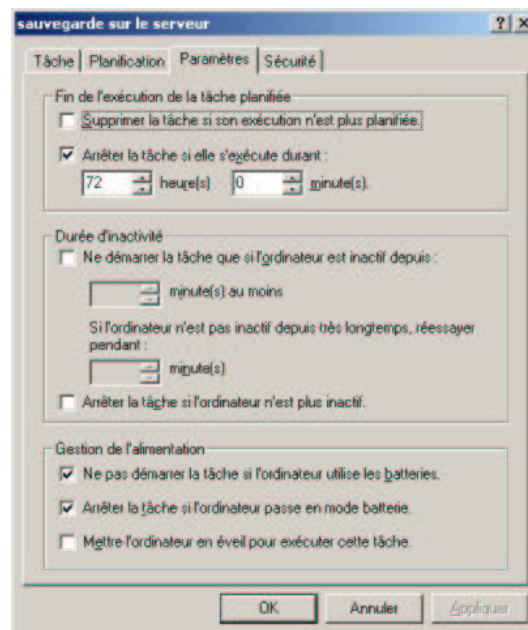


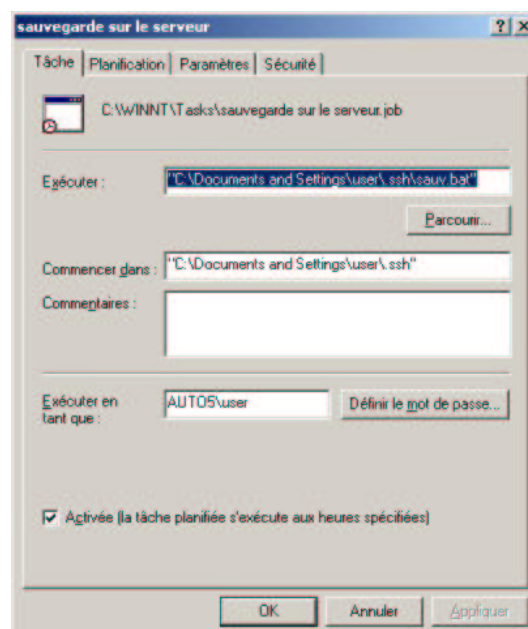
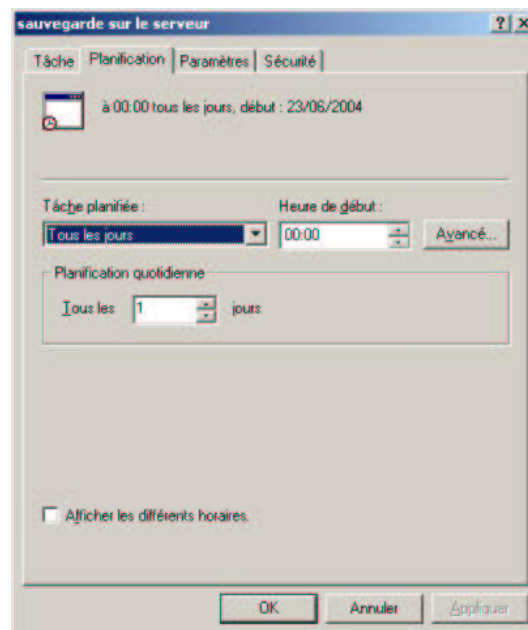
Choisir l'utilisateur sous lequel va s'exécuter la sauvegarde



Visualiser les propriétés avancées







La tâche est maintenant planifiée.

## 1.5 Important : Limitation

Afin que cette tâche puisse être lancée par le planificateur, il est nécessaire que l'utilisateur soit loggé sur la machine (sinon les connexions réseau ne sont pas activées et le **rsync** via **ssh** ne peut pas fonctionner). Ainsi, si l'utilisateur doit quitter son bureau et que la sauvegarde est programmée à 00h00 tous les jours, il doit verrouiller sa station (touche Windows-L ou démarrer->verrouiller) et non se déconnecter. Une piste pour résoudre ce problème, est de créer un service Windows qui fera le travail.

## 2 Restauration des données

Afin de simplifier le problème de la restauration des données, c'est un serveur samba qui a été choisi. Ce choix donne aussi de la souplesse aux utilisateurs, qui pourront choisir les fichiers à restaurer sur leur machine.

Bien sur il est tout à fait possible de restaurer les données de l'utilisateur d'un seul coup en utilisant le script `restaure.bat` suivant :

```
@ECHO OFF
```

```
rem Sauvegarde du rep Documents and settings
```

```
c:\cwrsync\rsync -av --rsh="ssh -i c:\cwrsync\id_dsa -l morere" 194.57.140.212:windows/Documents_and
```

```
rem Sauvegarde du rep utilisateur
```

```
c:\cwrsync\rsync -av --rsh="ssh -i c:\cwrsync\id_dsa -l morere" 194.57.140.212:windows/utilisateurs
```

Il suffit d'invertir l'ordre de la source et de la destination par rapport à la sauvegarde.

## 2.1 Configuration de Samba

Pour cela je vous renvoie à la page très bien faite de Yves Agostini <http://www.crium.univ-metz.fr/docs/system/samba/>.

mon fichier de configuration est le suivant `/etc/samba/smb.conf`

```
[global]
```

```
workgroup = LASC
```

```
netbios name = SERVEUR_KSTORE
```

```
server string = Serveur Samba du LASC
```

```
#optionnel netbios aliases = SERVEUR_KSTORE
```

```
os level = 99
```

```
local master = yes
```

```
printcap name = /etc/printcap
```

```
load printers = yes
```

```
socket options = TCP_NODELAY
```

```
#security = share
```

```
#encrypt passwords = no
```

```
security = user
```

```
encrypt passwords = yes
```

```
smb passwd file = /etc/samba/smbpasswd
```

```
guest account = nobody
```

```
hosts allow = 194.57.140. 127.
```

```
[public]
```

```
comment = Public Stuff
```

```
path = /home/samba
```

```
public = yes
```

```
writable = no
```

```
printable = no
```

```
browseable = yes
```

```
[homes]
```

```
comment = Home Directories
```

```
browseable = no
```

```
writable = yes
```

```
[printers]
```

```
comment = All Printers
```

```
path = /var/spool/samba
```

```
browseable = no
```

```
# Set public = yes to allow user 'guest account' to print
```

```
public = yes
```

```
writable = no
```

```
printable = yes
```

l'ajout d'un utilisateur samba sur le serveur se fait par la commande

```
smbpasswd -a login
```

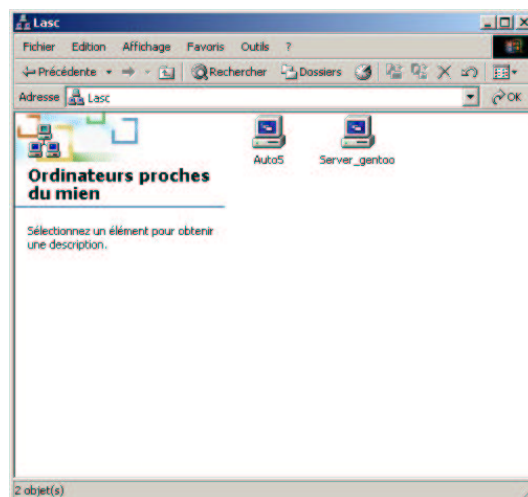
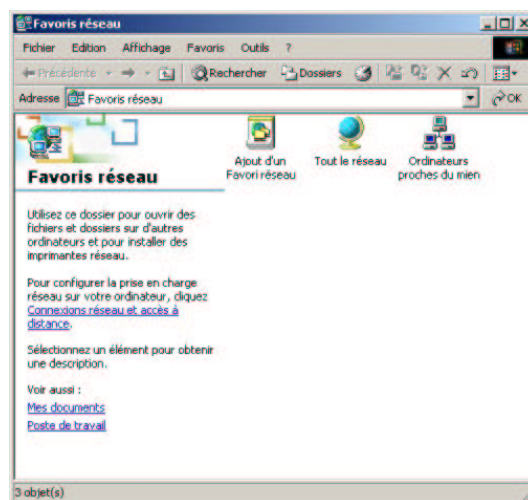
le système demande alors un mot de passe et une confirmation de mot de passe.

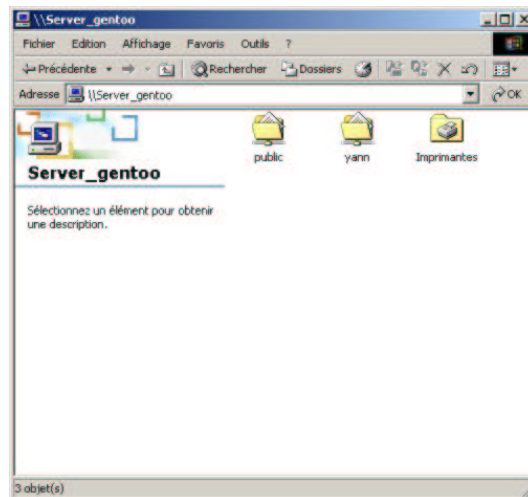
Il est possible d'utiliser pour samba les mêmes mots de passe que pour accéder à la machines windows. Dans ce cas, l'accès au serveur Samba sera quasiment transparent pour l'utilisateur (pas de login et pas de mot de passe).

Si les mots de passe sont différents, ce qui est préférable pour la sécurité, l'utilisateur, lors de la connexion sera invité à saisir un login et un mot de passe.

## 2.2 Connexion Du coté client

Par l'intermédiaire du voisinage réseau, l'utilisateur se connecte sur le serveur Samba de son groupe de travail, par l'intermédiaire de son login et mot de passe.





Un fois connecté sur le serveur, il ne reste plus à l'utilisateur de copier du serveur et coller en local les répertoire ou fichier qu'il veut restaurer.

Il est vraiment difficile de faire moins compliquer.

### 3 Installation et utilisation de rsync sous Linux/Unix

Sous Unix/Linux, la commande **rsync** est disponible. Si elle n'est pas installée par défaut, installez le paquetage correspondant. Pour notre debian un simple `apt-get install rsync` suffit en tant que **root**. Ensuite comme sous windows, il faut générer le jeu de clé publique/privée pour l'utilisateur. La syntaxe de commande est la même :

```
ssh-keygen -t dsa
```

Cette commande génère les clés de l'utilisateur et les place dans le répertoire `$HOME/.ssh`.

Il faut ensuite copier la clé publique sur le serveur.

```
scp ~/.ssh/id_dsa.pub morere@194.57.140.212:~/.ssh/authorized_keys
```

Un test de connexion est réalisé par la commande

```
ssh morere@194.57.140.212
```

Il suffit ensuite d'utiliser la commande **rsync** dans un script shell de la manière suivante (fichier `sauvegarde.sh`) :

```
#!/bin/bash
rsync -av --exclude-from=$HOME/sauvegarde/exclude.txt --rsh="ssh -l morere"
$HOME 194.57.140.212:linux
```

avec son fichier `exclude`

```
temp
GNUstep
```

Ensuite on insert une ligne dans la `crontab` par `crontab -u user -e`

```
0 2 * * * /home/user/sauvegarde/sauvegarde.sh
```

et voila le tour est joué, on sauvegarde le répertoire `home` sur le serveur dans le répertoire `linux` tous les jour à 2h00.

## 4 Sauvegarde des données du serveur *Mirroring*

Afin de limiter au maximum les coûts, la lourde tâche de sauvegarde sera assurée par deux disques durs. Le but est alors de sauver alternativement le contenu du répertoire `/home` du troisième disque sur l'un des deux disques de sauvegarde.

Pour cela on utilise 3 disques de 120Go pour le stockage et un disque de 3.2Go pour le système (Une debian Sarge (testing)).

```
morere@kstore:~$ df
Filesystem      1K-blocks      Used Available Use% Mounted on
/dev/hda2        2845540    1529672   1171324  57% /
tmpfs            128332         0     128332   0% /dev/shm
/dev/hde1       118176996   1796260  110377696   2% /home
/dev/hdg1       118176996   1794960  110378996   2% /sauvegarde1
/dev/hdh1       118176996     32828  112141128   1% /sauvegarde2
morere@kstore:~$
```

Un des disques de 120Go contient les répertoires `home` des utilisateurs et les deux autres servent à faire les sauvegardes synchronisées un jour sur deux (ou moins suivant le choix que l'on fait).

Dans mon cas, tous les jours impairs, les répertoires `home` sont synchronisés sur le disque `/sauvegarde1`, et tous les jours pairs sur le disque `/sauvegarde2`.

Ceci est réalisé par une simple commande insérée dans le `cron` par `crontab -e` en tant que `root`.

```
0 0 * * 0,2,4,6 rsync -a --exclude-from=/root/exclude /home/ /sauvegarde1/
0 0 * * 1,3,5 rsync -a --exclude-from=/root/exclude /home/ /sauvegarde2/
```

```
kstore:~# more exlude
/home/webCDcreator/
/home/lost+found/
/home/ftp/
kstore:~#
```

Voilà, pour l'instant le serveur est opérationnel pour notre laboratoire. La notion de quotas n'a pas été prise en compte. En effet les disques sont pour l'instant surdimensionnés pour nos besoins.

De même la suite du projet concerne l'écriture d'un programme C/GTK+ pour la configuration simplifiée et la création automatique des fichiers `sauv.bat`, `exclude.txt` etc.

Les questions et commentaires sont les bienvenus à l'adresse suivante : [morere@lasc.univ-metz.fr](mailto:morere@lasc.univ-metz.fr)