

## Administración: Backups

# RIESGOS LABORALES

Los datos siempre acaban perdiéndose en el peor momento, pero con una estrategia de copias de seguridad adecuada no tendremos problemas para restablecer los archivos perdidos.

**POR MARC ANDRÉ SELIG**

Debido a la amplia tipología de requerimientos en lo concerniente a pérdida de datos, a lo largo del tiempo han surgido un gran número de soluciones distintas. Cada tipo de solución tiene sus ventajas e inconvenientes. En el artículo de este mes de taller de administración, describiremos algunas técnicas y herramientas básicas de backups.

## Alternativas de backups

La cinta magnética ha sido durante mucho tiempo el medio de almacenamiento más habitual, y lo sigue siendo en redes con grandes cantidades de datos a almacenar. Las cintas son baratas en relación a su capacidad, pero su escasa velocidad es probablemente su mayor desventaja. Una solución basada en cintas y un jukebox suele ser la mejor para backups automatizados. Aún así, las cintas magnéticas suelen ser demasiado caras para entornos domésticos o de oficinas pequeñas.

Hoy día las soluciones basadas en CD, DVD, memorias flash o discos duros internos o externos son mucho más comunes. En entornos mayores, los administradores pueden usar también sistemas NAS (Network Attached

Storage) para añadir capacidad a sus discos duros.

Del mismo modo que existen distintos medios de almacenamiento, existen también diferentes estrategias de backup. En la mayoría de los casos, los administradores optan por backups incrementales, que guardan sólo los cambios ocurridos desde la última copia. Este método permite un ahorro considerable de espacio en los medios de almacenamiento, que se traduce en una mejora en cuanto al coste, y la convierte en la estrategia más utilizada.

La gran desventaja de los backups incrementales es que la restauración de los archivos consume mucho más tiempo que con un backup completo. Además de todo esto, es probable que los administradores tengan que estar cambiando el medio de almacenamiento si no disponen de un equipo jukebox. Una tercera variante es el backup diferencial, en el cual siempre se guarda los cambios respecto a la última copia completa. La Figura 1 ilustra los tres métodos.

## Off-line, on-line, "en caliente"

La elección del método de backup puede depender de las circunstancias en las que se recuperarán los datos. Si el fichero que necesita un usuario está en una cinta dentro de un armario, el proceso para recuperarlo puede llevar tiempo y esfuerzo.

Por contra, los backups on-line "en caliente" se guardan en dispositivos que soportan acceso automatizado 24 horas al día, 7 días a la semana. Este método ahorra tiempo, y normalmente también dinero. Sin embargo, los backups "en caliente" sólo protegen frente a daños del hardware. No tienen protección frente a errores del usuario o administrador, errores que propagaran al medio de almacenamiento tan pronto como se haga la copia. Esta es la razón por la cual la mayoría de los administradores no considera el backup en caliente como alternativa al convencional.

## Formatos

Los administradores no se ponen de acuerdo en los pros y contras de guardar los ficheros simples tal cual, o bien guardados dentro de archivos que los contengan, con un formato estructurado, metadatos y checksum.

Los ficheros simples se recuperan más rápidamente, y si el medio de almacenamiento sufre un fallo sólo quedará afectado ese archivo, mientras que si un archivo "contenedor" falla, seguramente se pierdan muchos ficheros.

Los archivos que contienen ficheros tienen algunas características que los backups de ficheros simples no tienen. Por ejemplo, pueden guardar información acerca del propietario, permisos de acceso y marcas de tiempo junto con el fichero propiamente dicho. Incluso podemos hacer copia de seguridad de disposi-



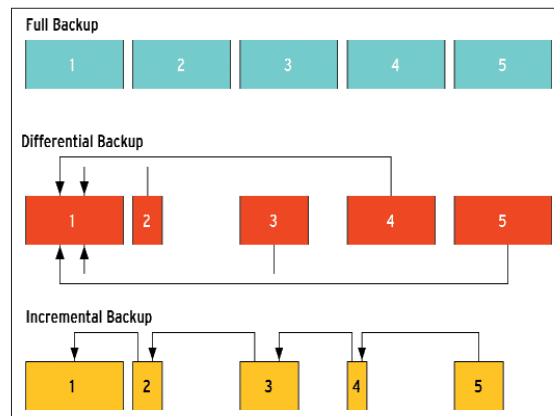
tivos especiales del directorio `/dev/`. Señalar también que las cintas magnéticas no están orientadas a almacenamiento de ficheros individuales.

Algunos programas, como `tar` y `cpio`, intentan un término medio. Si un archivo `cpio` resulta dañado, quedan afectados sólo los ficheros guardados en la zona dañada. El programa vuelve a sincronizar desde la siguiente marca de fin de fichero, y todos los ficheros que le siguen se recuperan sin problemas.

Al tratar este tema, también hemos de considerar la compresión y la encriptación de los ficheros. La característica de resincronización de `cpio` que hemos comentado sólo funciona para backups sin compresión. Si un error nos impide descomprimir un archivo contenedor, `cpio` no nos será de gran ayuda.

La popular herramienta `gzip` se cierra en caso de leer errores, por lo que es una mala elección para hacer copias de seguridad. (`zcat` al menos descomprime el archivo hasta el punto donde aparece el error de lectura). La alternativa, `bzip2`, comprime y descomprime los ficheros en bloques de 900 kbytes como mucho. Si sufrimos un error de lectura, probablemente sólo perdamos un bloque, salvándose de este modo los bloques anteriores y posteriores.

Los administradores también se encuentran ante un dilema similar con la



**Figura 1: Un backup completo guarda todos los ficheros, un backup diferencial guarda todos aquellos ficheros modificados desde el último backup completo y un backup incremental guarda todos los ficheros modificados desde el último backup, ya fuese este incremental o completo.**

encriptación de datos. La mayoría de los programas de encriptación usados en backups nos impedirán acceder a los datos si sufrimos un error. Una manera de solucionarlo podría ser comprimir y descomprimir cada fichero de manera individual dentro del archivo contenedor. La herramienta `afio` es una alternativa a `cpio` a este respecto, ya que puede manejar la encriptación individualizada de los ficheros.

## Backup con CD

Una solución backup con cinta magnética como Amanda (véase cuadro 1) puede ajustarse a entornos menores, pero sigue estando orientada preferente-

mente a grandes sistemas. Para usuarios domésticos, o pequeñas empresas puede que les sea más que suficiente un backup basado en CD o DVD. En comparación con las cintas magnéticas, los CDs y DVDs son extremadamente baratos y tienen un ciclo de vida bastante largo.

El listado 1 muestra un sencillo script para hacer backups, que utiliza `gpg` para encriptar los datos y guarda un checksum MD5. Si se pierde el CD, al menos no tenemos que preocuparnos de que accedan a los datos indebidamente. Podemos modificar el script para poder usar memoria flash, o un disco duro externo.

## El Camino a Seguir

Un sistema de backups es tan bueno como lo sean los datos guardados en el medio de almacenamiento. Y estos datos no tienen que ser necesariamente los que debían de haberse guardado. Por lo tanto tiene mucho sentido verificar los backups para comprobar que pueden leerse y que son los datos adecuados cada cierto tiempo.

Asimismo, deberíamos asegurarnos de que los usuarios son capaces de recuperar los datos por sí mismos. No hay nada peor que tener la necesidad de recuperar un backup configurado por otra persona hace mucho tiempo, y no ser capaces de hacerlo porque esta persona ya no está en la empresa.

El caso de pérdida total de los datos conlleva otras muchas consideraciones. Como el propio sistema operativo probablemente no estará disponible, tiene sentido prever un sistema de rescate. Este sistema podría arrancar desde un CD o disco duro externo y permitiría al administrador recuperar todos los datos desde allí. Por supuesto, este tipo de soluciones requiere de una buena planificación y algo de práctica. ■

### Listado 1: Script simple para Backup

```
01 #!/bin/sh
02
03 [ `id -u` -eq 0 ] || ( echo
    'Must be root to write a
    CD/DVD!' && exit )
04
05 TODAY=`date +%Y%m%d.%H%M`
06 MYKEY='0x598342d9'
07
08 umask 022
09 mkdir -p
    /tmp/root/backup-$TODAY
10
11 cd /
12 tar cf - etc home usr/local |
    \
13 gpg -v --homedir $HOME/.gnupg
    -e -r $MYKEY | \
14 tee
    /tmp/root/backup-$TODAY/backup
    -$TODAY.tar.gpg | \
15 md5sum -b
    >/tmp/root/backup-$TODAY/bac-
    kup-$TODAY.tar.gpg.md5
16
17 cd /tmp/root
18 mkisofs -r -pad -o backup.iso
19 backup-$TODAY
20 cdrecord -v -eject -multi
    dev=0,0,0 -driveropts= burn-
    proof -speed=24 -pad
    backup.iso
21
22 rm -rf backup-$TODAY
    backup.iso
```

### RECURSOS

- [1] Afio: <http://directory.fsf.org/sysadmin/backup/afio.html>
- [2] Amanda: <http://www.amanda.org>
- [3] Amanda para Windows: <http://sourceforge.net/projects/amanda-win32/>