

O hacemos copias de respaldo externas o perderemos los datos

GUARDAESPALDAS

¿Quién necesita atacantes teniendo administradores de sistemas? Vemos por qué copiar datos no significa lo mismo que tener los datos respaldados. **Por Kurt Seifried**

La ironía aparece cuando menos te lo esperas, esta vez ha sido en forma de desastre después de haber borrado el equivalente de un mes de datos. Mientras trasteábamos unas copias de respaldo durante la preparación de este artículo, se perdió */var/* casi por completo, así como la totalidad del directorio */home/*. El desastre no hubiese sido tal de haber guardado las copias de respaldo diarias en otro sitio distinto de */home/backups/*. ¡Ups!

Hacer Copias No Significa Tenerlas

Si los datos no están disponibles, o si los sistemas que los procesarán y servirán no están disponibles, tenemos un problema. El servidor web que introduce este artículo pasó a ser inservible tras perder */var/* y */home/*. Lo único que hace ahora es mostrar errores 404 y poco más. Para estar seguros de que los datos estarán disponibles, hay que hacer copias de respaldo. Parece simple, ¿verdad? En realidad, la mayoría de nosotros lo hacemos mal y, a pesar de que nos tomamos la molestia de crear copias de seguridad, lo único que hacemos es copiar datos a ubicaciones igual de vulnerables.

El caso que exponemos es el del típico error que consiste en almacenar las copias de seguridad en el mismo sistema que se

está respaldando y, para empeorar las cosas, en un directorio de acceso frecuente. Aunque tampoco habría importado demasiado esto último. Debido a que el servidor sólo dispone de un único dispositivo de almacenamiento, basta con un solo fallo del disco para perder completamente los datos, independientemente de cuántas copias de respaldo se hagan del sistema localmente y de cómo se hagan. Incluso instalando un segundo disco en la máquina, sigue habiendo bastantes probabilidades de que a causa de un solo evento (un controlador defectuoso, un atacante borrando el sistema deliberadamente, un fuego, algún líquido derramado, una fuente de alimentación descontrolada, un robo, etc.) se produzca la pérdida de los datos en todos los discos.

¿Cómo Son las Copias Reales?

Hay tres elementos principales implicados en la elaboración de verdaderas copias de seguridad.

Primero: Hay que estar seguros de que los datos se han copiado realmente. Son demasiados los sistemas que escriben datos, en CDs, DVDs o cintas, de forma inadecuada, siendo éstos inútiles al no ser recuperables. Idealmente, habría que comprobar cada una de las copias hechas pero, como no es

demasiado práctico, al menos hay que hacer comprobaciones aleatorias para poder estar seguros de que los datos son realmente recuperables.

Segundo: Se deben ubicar las copias de respaldo lejos del sitio objeto de la copia, con un acceso de lectura tan restringido como sea posible. Esto no significa necesariamente que deban estar en una ubicación física distinta (algo que por otro lado siempre es buena idea), pero su ubicación tiene que estar al menos lo suficientemente separada como para que un evento o fallo individual, como el formato de un array de discos o la pérdida de un servidor, no dé al traste a la vez con las copias de seguridad y los datos originales. Un ejemplo perfecto (aparte del *faux pas* descrito al inicio de este artículo) es el del sitio web *AVSIM Online*, que perdió 13 años de datos por culpa de un solo ataque [1]. Según las informaciones publicadas, *AVSIM Online* contaba con dos servidores que se copiaban datos el uno al otro en un intento de respaldar dichos datos mutuamente. Como decíamos antes, muchos de nosotros sólo copiamos datos, sin llegar a hacer verdaderas copias de seguridad. En este caso, el atacante accedió a ambos servidores, ya que eran prácticamente idénticos, eliminando de ambos tanto los datos originales como las copias de seguridad. *AVSIM Online* perdió su sitio web, sus correos, las librerías de archivos, los foros, etc., de un tirón, y probablemente nunca recupere sus datos. En nuestro caso tuvimos suerte, ya que sólo se eliminaron los datos y los archivos de registro de un mes, por lo que sólo ha habido que esperar un mes para volver a recopilarlos ... menos mal que no se trataba de la información financiera de nadie.

Tercero: Hay que estar seguros de no eliminar archivos de la copia a menos que se esté absolutamente convencido de que dichos archivos no serán necesarios. Por este motivo, RAID no es una solución de backup. Aunque haya varios dispositivos en una configuración RAID, de modo que la pérdida de uno o incluso de varios dispositi-

vos no provoque la pérdida de los datos, siempre cabe la posibilidad de que se pierdan por otro motivo, como su eliminación accidental (*rm*, *mkfs*, etc.) o su alteración por otros medios (*cat foo > bar*).

Sacarlos del Sistema

Por suerte, casi cualquier programa dedicado a la realización de copias de seguridad permite obtener datos desde un cliente y almacenarlos en un servidor dedicado, un array de discos, una cinta, un DVD, etc. Hay opciones excelentes para Linux: como Amanda [2], que suele estar incluida en casi todas las distribuciones; o BackupPC [3] y Bacula [4], ambas tratadas en Linux Magazine [5] [6]. Aunque no vamos a exponer los detalles aquí, basta con decir que son todas muy potentes, que cuentan con un montón de ajustes y que, en definitiva, son capaces de realizar copias de respaldo de los datos si se configuran correctamente. Una solución rápida para salir del paso puede ser *rsync* [7] (*yum install rsync*, *apt-get install rsync*, etc.).

El Problema de rsync

Rsync fue diseñado para mantener sincronizados conjuntos de archivos de gran tamaño entre sistemas o directorios. Como tal, es una buena solución de backups “para pobres” o herramienta de asistencia. Se pueden hacer copias mediante *tar* y *mysqldump*, colocando los archivos en directorios con una marca de tiempo (ahora veremos por qué). Para usar *rsync*, simplemente creamos el archivo *rsync.conf* bajo */etc/*:

```
uid = backups
gid = backups
use chroot = yes
[backups]
    path = /backups/
    read only = yes
```

Luego, lo habilitamos en *inetd* o *xinetd*. En el lado del cliente, usaremos un comando al

Los Principios de la Seguridad

Los tres principios de la seguridad son: Disponibilidad, Integridad y Confidencialidad (conocidos también como la triada AIC: *Availability*, *Integrity* y *Confidentiality*). A modo de resumen, hay que garantizar que se haga lo que hay que hacer, que los datos no han sido manipulados por un eventual atacante, y que dichos datos se mantienen confidenciales.

estilo de *rsync -a 10.1.2.3::backups/* /mis-backups/* para copiar los contenidos del directorio */backups/* del servidor remoto *10.1.2.3* al directorio local */misbackups/*. Prescindimos de la opción *-delete* para evitar males mayores, impidiendo así que *rsync* elimine los archivos locales que ya no existen en el punto remoto. ¿Qué podría ir mal? Si se eliminase un archivo en el servidor, aún tendríamos de la copia local, ¿no? Sí, pero si un atacante sobrescribiese un archivo, o si se diese un error en el script encargado de realizar las copias de respaldo (por ejemplo *cat 0 > algun_archivo*), la copia local se sobrescribiría también. Dicho de otro modo, adios a los datos. La solución pasa por realizar copias de seguridad incrementales, motivo por el cual es preferible mantener las copias en directorios cuyos nombres sean la fecha en que se realizaron y sincronizar solamente ese directorio:

```
rsync -a 10.1.2.3::backups/`date +%Y-%m-%d` /mis-backups/
```

En el peor de los casos se perderían los datos de ese día, pero no las copias de respaldo anteriores, ya que éstas se encuentran en directorios separados del servidor local en los que no escribiría *rsync*.

¿Sabemos Dónde Están las Copias?

Bien, ya tenemos backups diarios sacados del servidor a otra máquina, esperemos que lo suficientemente segura. Los programas automatizados fallan demasiado a menudo, las direcciones IP y los nombres de host también suelen cambiar, la configuración de *rsync* puede sufrir modificaciones, la partición de disco local dedicada a albergar los backups puede llenarse, quién sabe qué más puede pasar. La última pieza en el rompecabezas de las copias de respaldo es la notificación automatizada informando sobre si el backup ha tenido éxito o no. Los programas especializados en la realización de copias de seguridad mencionados anteriormente (*Amanda*, *Bacula*, etc.) soportan todas las notificaciones, pero ¿cómo podemos enviar notificaciones en el sistema que hemos montado con *rsync*? Hay una solución elegante y simple: Basta con añadir la siguiente línea al script encargado de hacer las copias de seguridad:

```
ls -la /mis-backups/`date +%Y-%m-%d` | mail -s
```

“notificación del backup diario” direccion@example.com

Este comando listará los contenidos del nuevo directorio y enviará los resultados al comando *mail*, que enviará un correo que contendrá todos los archivos y sus tamaños. De hecho, podemos ir más allá y listar los contenidos de los archivos comprimidos mediante el comando *tar -t*. En la mayoría de los casos, al ejecutar *rsync* con la opción *-v* (verbose), se generará una salida como ésta:

```
receiving file list ... done
2009-05-27/
2009-05-27/home.tar.bz2
2009-05-27/etc.tar.bz2
```

Nótese que si se ejecutan los comandos desde *crontab*, la salida se enviará por correo automáticamente al usuario que los ejecuta.

Probando las Copias Realizadas

Se podría decir que es el paso más importante a la hora de hacer copias de respaldo. Primero hay que hacerse con los backups y descomprimirlos, cargando la cinta en el sistema o haciendo lo que se haría en caso de tener que restaurar verdaderamente los datos; de no ser así, ¿cómo podríamos estar seguros de que los datos nos servirán? Cuando lo hayamos comprobado, ya podremos dormir tranquilos sabiendo que ni los desastres naturales ni los administradores de sistemas manazas arruinarán nuestro día. ■

RECURSOS

- [1] “Avsim.com Hacked – Total Data Loss” por James Anderson, Tech Fragments, 15 de Mayo de 2009, http://techfragments.com/news/769/Tech/Avsim-com_Hacked_-_Total_Data_Loss.html
- [2] Amanda: <http://www.amanda.org/>
- [3] BackupPC: <http://backuppc.sourceforge.net/>
- [4] Bacula: <http://www.bacula.org/>
- [5] “BackupPC” por David Nalley, Linux Magazine, Número 44, <http://www.linux-magazine.es/issue/44/064-068BackupPCLM44.pdf>
- [6] “Bacula” por Jens-Christoph Brendel y Marc Schöchlin, Linux Magazine, Número 10, <http://www.linux-magazine.es/issue/10/Bacula.pdf>
- [7] *rsync*: <http://samba.anu.edu.au/rsync/>