

Lethe

Modificaciones realizadas para usar particiones como pseudo-tmpfs

Francisco Javier Díaz

Joaquín Ignacio Bogado García

Claudia Banchoff Tzancoff

Einar Felipe Lanfranco

{javierd,jbogado,cbanchoff,einar}@linti.unlp.edu.ar

LINTI - Facultad de Informática - Universidad Nacional de La Plata

La Plata, B1900ASD, ARGENTINA

Resumen—Lethe es una herramienta que permite congelar el contenido del sistema de archivos de sistemas basados en Debian GNU/Linux mediante la utilización de AUFS y tmpfs. Su funcionalidad es similar a la implementada por herramientas comerciales como Deep Freeze o Windows Steady State en ambientes Microsoft.

Este trabajo describe una serie de modificaciones realizadas al programa Lethe para ampliar la cantidad de cambios que puedan hacerse al sistema de archivos independientemente del tamaño de la memoria RAM, evitando así fallos en el sistema por agotamiento de recursos.

Palabras clave: Lethe, Lihuen, GNU/Linux, Debian, Aufs, Deep Freeze, Steady State

I. CONTEXTO

Este trabajo se realizó en el marco del proyecto Lihuen GNU/Linux [1]. El objetivo del proyecto es desarrollar y mantener una distribución de GNU/Linux, actualmente desarrollada por un equipo de docentes y alumnos del Laboratorio de Investigación de Nuevas Tecnologías Informáticas (LINTI) de la Universidad Nacional de La Plata. La distribución está orientada a entornos educativos y cuenta con un gran número de herramientas divididas en áreas como geografía, química, música, nivel inicial, matemáticas, lengua o astronomía. También se desarrolla paralelamente una sub-distribución de Lihuen que incluye herramientas para la mayoría de las materias que se dictan en la Facultad de Informática de la UNLP. Parte de los objetivos del proyecto cuentan la utilización, formación y difusión en materia de software libre.

II. INTRODUCCIÓN

En ambientes con equipos de uso público como laboratorios, salas de computación, cibercafé o academias, es deseable que el uso por parte de un usuario no afecte la configuración ni el software instalado en ese equipo.

La Facultad de Informática cuenta con tres salas de computación para alumnos, donde las máquinas poseen doble booteo: sistemas Microsoft Windows en sus versiones XP y 7 y Lihuen GNU/Linux, una distribución de GNU/Linux que se desarrolla en la Facultad de Informática

de la UNLP y se encuentra instalado en más de 80 equipos de dicha institución y en otras salas de computación de la Provincia de Buenos Aires. Estas salas tienen un uso académico y, en algunos casos, se requiere que los usuarios puedan instalar y probar herramientas durante las clases sin que estos cambios se vean reflejados en sesiones de otros usuarios.

Esta situación se puede generalizar a otros ambientes con equipos de uso público, como laboratorios, cibercafé o gabinetes de escuelas, en donde es deseable que las acciones de un usuario no afecten la configuración ni el software instalado en ese equipo.

Lethe[2] es un sistema que permite proteger el contenido del o los discos de una PC independientemente de los cambios que realice un usuario. Dicha protección se realiza mediante el congelamiento de las particiones del disco utilizando Aufs [3], está pensado y desarrollado para entornos Lihuen GNU/Linux, y en general, es aplicable en cualquier sistema derivado de Debian GNU/Linux [4].

III. EL FUNCIONAMIENTO DE LETHE

El desarrollo de Lethe se basa en una generalización del código de RootAUFS¹, escrito originalmente por Nicholas Schembri. RootAUFS fue pensado originalmente para reducir el impacto de los ciclos de escritura en las memorias de estado sólido (que reemplazan al disco rígido en algunas portátiles y *netbooks*) escribiendo los cambios realizados al sistema de archivos en memoria RAM en lugar de usar la memoria de estado sólido. Para esto, es necesario crear un sistema de archivos en memoria RAM llamado tmpfs y unificarlo con la raíz del sistema de archivos de la memoria de estado sólido mediante Aufs.

Aufs fue en un primer momento una reimplementación de Unionfs² versión 1.x y, aunque ahora se ha vuelto un proyecto totalmente independiente, conserva las características principales de Unionfs. Básicamente, es un sistema de archivos **unificable** y **apilable**; unificable

¹ <https://help.ubuntu.com/community/aufsRootFileSystemOnUsbFlash>

² <http://www.filesystems.org/project-unionfs.html>

porque el contenido de varios directorios (*branches* o ramificaciones) tendrá la apariencia de uno solo mientras que el contenido físico de éstos estará separado; apilable porque los *branches* se apilan uno sobre otro en el orden en que se van montando. Cada *branch* puede tener sus propios atributos; por ejemplo, si el *branch* es de sólo lectura o de lectura/escritura, si tiene control de errores o no, etc.

Lethe se aprovecha de la capacidad de Aufs para permitir el acceso a múltiples directorios a través de un único punto de montaje.

El sistema de archivos tmpfs, por su parte, es un sistema de archivos sumamente sencillo, que utiliza la memoria RAM (toda la disponible o parte de ésta, dependiendo de los parámetros pasados a la hora de montar la unidad) como si fuera un sistema de archivos en disco. Además, tmpfs comparte el espacio de la memoria RAM con otros sistemas tmpfs (por ejemplo *udev* y *shm*).

IV. EL PROBLEMA

El problema radica en que tmpfs no utiliza el espacio de almacenamiento de los discos si estos existen. No es posible crear un sistema de archivos de tamaño mayor que el total de memoria RAM disponible (generalmente de 2 a 4 GB). Por lo tanto, cuando la memoria se termina, por ejemplo debido a la copia de uno o varios archivos de gran tamaño (o incluso el borrado de varios archivos), el sistema se vuelve inestable debido a la falta de memoria, se torna inoperable y es necesario reiniciarlo.

V. Internals

Una parte importante en el funcionamiento de Lethe es la que crea el sistema de archivos temporal en la carpeta */rw*. Esta carpeta es la que posteriormente se une a la raíz del sistema de archivos mediante el empleo de Aufs y sobre la cual se hacen los cambios de todos los puntos de montaje que se congelan. La línea responsable de montar esta carpeta es la siguiente:

```
run_echo "mount -t tmpfs aufs-tmpfs /rw"
```

La función `run_echo` ejecuta el comando que recibe como argumento, realiza un manejo precario de errores y aborta la ejecución del *script* si el comando falla. Si el comando es exitoso, todo lo que se escriba en la carpeta */rw* será volátil ya que el sistema de archivos tmpfs usa exclusivamente memoria RAM como medio de almacenamiento (como ya se ha mencionado).

En este punto del arranque, el sistema raíz está montado en la variable `$rootmnt`. Luego se mueve esta raíz a */ro* con el comando siguiente:

```
run_echo "mount --move $rootmnt /ro"
```

De esta manera el sistema de archivos temporal queda montado en */rw* y el sistema de solo lectura queda montado en */ro*.

Posteriormente ambos sistemas de archivos se unifican con Aufs.

```
run_echo "mount -t aufs -o dirs=/rw:/ro=ro
aufs /aufs"
```

El parámetro `dirs` indica que tanto */rw* como */ro* van a ser accesibles a través de */aufs* y que */ro* además va a

ser de solo lectura (*/ro=ro*).

Debido a esto, ambos sistemas de archivos quedan ocultos en */aufs/rw* y */aufs/ro*. Para que puedan ser accedidos como un único sistema de archivos a través de */* es necesario mover */aufs* a */root* (la cual en posteriores *scripts* de la secuencia de arranque se transforma en */*).

```
run_echo "mount --move /aufs /root"
```

Básicamente³, esto permite que Lethe funcione, ya que cuando se crea un archivo en */* o cualquier otro subdirectorio, se crea en realidad en */aufs/rw/*. Por ejemplo, si ejecutamos `touch4 /home/lihuen/miarchivo`, además de ser accesible mediante `/home/lihuen/miarchivo`, será accesible mediante `/aufs/rw/home/lihuen/miarchivo`. Sin embargo, si tratamos de acceder al sistema de archivos original mediante */aufs/ro*, por ejemplo ejecutando `rm /aufs/ro/bin/bash` recibiremos un error indicando que el sistema de archivos es de solo lectura (aún teniendo privilegios de superusuario).

VI. LA SOLUCIÓN

Habiendo explicado el funcionamiento interno de Lethe, se describe a continuación una solución al problema del agotamiento de memoria producido por el uso del sistema de archivos tmpfs.

En lugar de utilizar la (a veces escasa) memoria RAM, podemos dedicar espacio en disco para escribir los cambios producidos en el sistema de archivos. La idea es tener una partición de tamaño suficiente (y mayor que la memoria RAM) de, digamos, 50 a 100 GB, sin contenido y formateada con algún sistema de archivos (por ejemplo *ext3*). Esta partición del disco debe tener la etiqueta 'lethe' para que el instalador la encuentre y configure la variable de entorno `$LETHE_PART` en `/etc/lethe/lethe.conf`. En cada reinicio del sistema, si Lethe está activo, borrará el contenido de esta partición y la utilizará para montar */rw* en lugar de montar */rw* con el sistema de archivos tmpfs.

La solución se implementó como sigue:

```
source ${rootmnt}/etc/lethe/lethe.conf
if [ -n "$LETHE_PART" ]; then
    run_echo "mount $LETHE_PART /rw"
    run_echo " rm -r -f /rw/*"
else
    run_echo "mount -t tmpfs aufs-tmpfs /rw"
fi
```

De esta manera, si la partición existe se utiliza como sistema de archivos temporal y se borra cada vez. Si la variable `$LETHE_PART` no está declarada, entonces se utiliza el sistema de archivos tmpfs como en el caso anterior.

³ Se han evitado algunos detalles de implementación como la desactivación de *AppArmor* para Ubuntu, correcciones a `/etc/mtab` y algunas tareas que se agregan a `rc.local` con el fin de simplificar la comprensión de este texto y no desviarnos del objetivo del trabajo.

⁴ `touch` es un programa que permite crear archivos rápidamente desde la consola, si bien su objetivo principal es cambiar la fecha de última modificación de un archivo y de ahí su nombre.

VII. VENTAJAS Y DESVENTAJAS

Esta solución tiene la desventaja de que incrementa el tiempo de arranque de la máquina debido a que durante el arranque se borra el contenido de la partición dedicada. Esto es necesario para simular el efecto de reiniciar en la memoria RAM. El tiempo de arranque no es constante entre reinicios ya que depende de la cantidad y el tamaño de los archivos generados en la sesión previa.

Otra desventaja es que el descubrimiento de la partición dedicada se realiza durante la instalación, es decir, una vez. Si la partición cambia será necesario cambiar a mano el archivo `/etc/lethe/lethe.conf` y apuntar la variable `$LETHE_PART` al dispositivo que corresponda.

Además es necesario cambiar el esquema de particionamiento para que Lethe haga uso exclusivo de una partición dedicada.

A pesar de lo arriba mencionado, es importante destacar que esta solución permite superar el límite de tamaño en tmpfs, restringido al tamaño de la memoria RAM, utilizando una partición dedicada tan grande como la partición utilizada. Además, si la partición no está disponible, el sistema usa el sistema tmpfs convencional basado en memoria RAM.

VIII. PRUEBAS REALIZADAS

Se realizaron pruebas exitosas sobre máquinas virtuales donde se generaron archivos nuevos de hasta 6 veces el tamaño de la memoria RAM sin notar degradaciones en el sistema. Además se realizó la instalación masiva de la nueva versión en una de las salas de PCs de la Facultad de Informática de la UNLP donde podrá monitorearse su desempeño durante los próximos 6 meses.

Ya con el paquete disponible para el público en general, se espera recibir reportes de usuarios de diferentes partes del mundo y sobre otros sistemas operativos además de Lihuen, como ya ha sucedido con lanzamientos anteriores.

IX. TRABAJO A FUTURO

Por el momento, la detección de la partición a usar como pseudo-tmpfs se realiza durante leyendo el archivo de configuración. Se piensa incluir la detección de la partición al vuelo, brindando la posibilidad de utilizar *pendrives* como pseudo-tmpfs.

Además se realizará una investigación acerca del uso de archivos como dispositivos de bloques. Es sabido que QEMU/KVM utiliza un sistema que permite crear discos virtuales en archivos que aumentan de tamaño en la medida que sea necesario. Esto evitaría tener que crear un archivo de, digamos, 50 GB, darle formato y borrarlo en cada reinicio. Además, haría innecesario el reparticionamiento del disco.

X. FORMACIÓN DE RECURSOS HUMANOS

Actualmente en el proyecto Lihuen GNU/Linux trabaja un grupo de docentes y alumnos de la Facultad de Informática de la UNLP, investigando sobre diversos temas relacionados con la adaptación de Lihuen a las escuelas

de la zona. Dentro de estas áreas se incluyen tópicos como accesibilidad, clientes livianos, herramientas para educación, sistemas operativos.

Entre las últimas asistencias a congresos podemos citar:

- Las Jornadas Regionales de Software Libre[5] realizadas en San Luis a finales de octubre de 2010, donde el grupo presentó varios trabajos. Entre los trabajos presentados, se incluyen un taller de programación de interfaces gráficas con PyGKT, una charla sobre el uso de robots en la enseñanza de programación en escuelas técnicas y una charla de Lethe donde se abarcaron los temas de funcionamiento e implementación.
- Las Octavas Jornadas de Software Libre[6] realizadas el 21 de octubre de 2010 en la Facultad de Informática de la Universidad Nacional de La Plata. Estas estuvieron organizadas por el grupo de trabajo del proyecto Lihuen y la Secretaría de Extensión de dicha facultad.

El grupo complementa la actividad de investigación realizando actividades de extensión universitaria. Por ejemplo durante el año 2010 y lo que va del presente año se visitaron varias escuelas de la zona, donde se brindaron charlas acerca de los beneficios del uso de software libre a alumnos que cursan los últimos años del secundario y se colaboró en la instalación de redes de clientes livianos.

REFERENCIAS Y ENLACES DE INTERÉS

- [1] Página principal - sitio oficial de lihuen. <http://lihuen.linti.unlp.edu.ar/index.php/Portada>.
- [2] Javier Díaz. Joaquín Bogado. Claudia Banchoff. Einar Lanfranco. Lethe: Congelador de particiones para lihuen gnu/linux. *Encuentro Chileno de Ciencias de la Computación, Santiago de Chile, Chile*, nov/2009.
- [3] Aufs @ sourceforge. <http://aufs.sourceforge.net/>.
- [4] Debian - el sistema operativo universal. <http://www.debian.org/>.
- [5] JRSL. <http://jornadasregionales.org/jrsl2010v2/schedule/index>.
- [6] Software libre - octavas jornadas de software libre. <http://sl.linti.unlp.edu.ar/octavas-jornadas-de-software-libre/>.