Implementando UFS Journaling en un ordenador de escritorio

Resumen

Un sistema de archivos con journaling utiliza un log para registrar todas las transacciones que tienen lugar en el sistema de archivos y conserva su integridad en caso de un fallo del sistema o un corte de alimentación. Aunque existe la posibilidad de perder cambios que no hayan sido guardados en los archivos, el journaling elimina casi por completo la posibilidad de que el sistema de archivos se dañe debido a un cierre repentino. También minimiza el tiempo que lleva verificar el sistema de archivos después de un fallo. Aunque el sistema de archivos UFS empleado por FreeBSD no implementa el journaling en si mismo, la nueva clase journal del framework GEOM en FreeBSD 7.*X* se puede utilizar para proporcionar un journaling independiente del sistema de archivos. Este artículo explica cómo implementar el journaling UFS en un ordenador típico de escritorio.

Tabla de contenidos

1. Introducción . . . 1
2. Entendiendo el journaling en FreeBSD . . . 2
3. Pasos durante la instalación de FreeBSD . . . 3
4. Configurando el journaling . . . 6
5. Resolución de problemas del journaling . . . 9
6. Lecturas adicionales . . . 11

1. Introducción

Si bien es cierto que los servidores profesionales, generalmente, están bien protegidos contra las paradas imprevistas, un ordenador típico de escritorio está a merced de problemas con la energía, reinicios accidentales y otros incidentes relacionados con el usuario que pueden terminar en paradas abruptas del sistema. Las actualizaciones de software, generalmente, protegen el sistema de archivos de manera eficiente en tales casos, aunque en la mayoría de ocasiones se requiere una larga verificación en segundo plano. En raras ocasiones, la corrupción del sistema de archivos llega a un punto en el cual se requiere la intervención del usuario y en el que podría producirse una perdida de datos.

La nueva función de journaling provista por GEOM puede ayudar en gran medida en esos escenarios, al eliminar, virtualmente, el tiempo requerido para la verificación del sistema de archivos, y garantizar que el sistema de archivos se restaure rápidamente a un estado consistente.

Este artículo describe un procedimiento para implementar UFS journaling en un escenario con un

ordenador típico de escritorio (un solo disco duro utilizado para el sistema operativo y los datos). Debe seguirse durante una nueva instalación de FreeBSD. Los pasos son lo suficientemente simples y no requieren de un uso demasiado complejo de la línea de comandos.

Después de leer este artículo, sabrás:

- Cómo reservar espacio para el journaling en una instalación nueva de FreeBSD.
- Cómo cargar y activar el módulo `geom_journal` (o añadir soporte para él en un kernel personalizado).
- Cómo hacer que tus sistemas de archivos ya existentes utilicen journaling, y qué opciones usar en el archivo /etc/fstab para montarlos.
- Cómo implementar journaling en nuevas (vacías) particiones.
- Cómo resolver problemas asociados con el journaling.

Antes de leer este artículo, deberías poder:

- Entender conceptos UNIX® y FreeBSD básicos.
- Estar familiarizado con el procedimiento de instalación de FreeBSD y la utilidad sysinstall.



El procedimiento descrito aquí está pensado para preparar una nueva instalación en la que todavía no se han almacenado datos reales del usuario en el disco. Aunque puedes modificar y ampliar este procedimiento a sistemas que ya están en producción, antes de hacerlo, debes realizar un *backup* de todos los datos que sean importantes. Jugar con discos y particiones a bajo nivel puede provocar errores fatales y pérdida de datos.

2. Entendiendo el journaling en FreeBSD

El journaling proporcionado por GEOM en FreeBSD 7.*X* no es específico de un sistema de ficheros (a diferencia de lo que ocurre en el sistema de ficheros ext3 de Linux®) pero funciona a nivel de bloque. Aunque esto significa que se puede aplicar a diferentes sistemas de ficheros, en FreeBSD 7.0-RELEASE sólo se puede aplicar en UFS2.

Esta funcionalidad se proporciona al cargar el módulo geom_journal.ko en el kernel (o añadirlo al compilar un kernel personalizado) y utilizando el comando `gjournal` para configurar los sistemas de archivos. En general, deberías añadir journal a los sistemas de archivos grandes, como /usr. Sin embargo, necesitarás (consulta la siguiente sección) reservar algo de espacio libre en el disco.

Cuando un sistema de archivos tiene journaling, se necesita algo de espacio en el disco para mantener el propio journaling. El espacio en disco que contiene los datos reales se conoce como *data provider*, mientras que el que contiene el journaling se conoce como *journal provider*. Los providers de data y journal deben estar en diferentes particiones cuando se hace journaling en una partición ya existente (que no esté vacía). Al hacer journaling en una partición nueva, tienes la opción de usar un solo provider para data y journal. En cualquier caso, el comando `gjournal` combina ambos providers para crear el sistema de archivos final, con el journaling. Por ejemplo:

- Quieres crear un journal para tu sistema de ficheros /usr, almacenado en /dev/ad0s1f (que ya contiene datos).
- Has reservado algo de espacio libre en la partición /dev/ad0s1g.
- Al usar `gjournal` se crea un nuevo dispositivo /dev/ad0s1f.journal donde /dev/ad0s1f es el data provider, y /dev/ad0s1g es el journal provider. Este nuevo dispositivo es usado para todas las operaciones de fichero siguientes.

La cantidad de espacio en disco que necesita reservar para el journal provider depende del uso del sistema de archivos y no del tamaño del data provider. Por ejemplo, un ordenador de oficina típico, con un journal provider de 1 GB para el sistema de archivos /usr será suficiente, mientras que un ordenador que haga un uso intensivo de E/S en el disco duro (por ejemplo, edición de video) podría necesitar más. Se producirá un kernel panic si el espacio del journal se agota antes de poder grabar los datos.



Es muy poco probable que los tamaños de journaling sugeridos aquí causen problemas con el uso de un ordenador de escritorio típico (como la navegación web, el procesamiento de textos y la reproducción de archivos multimedia). Si con tu trabajo se hace un uso intensivo del disco, usa la siguiente regla para una máxima fiabilidad: el tamaño de la RAM debe ajustarse al 30% del espacio del journal provider. Por ejemplo, si tu sistema tiene 1 GB de RAM, crea un journal provider de aproximadamente 3.3 GB. (Multiplica el tamaño total de tu RAM por 3.3 para obtener el tamaño del journal).

Para más información sobre journaling, por favor lee la página de manual de gjournal(8).

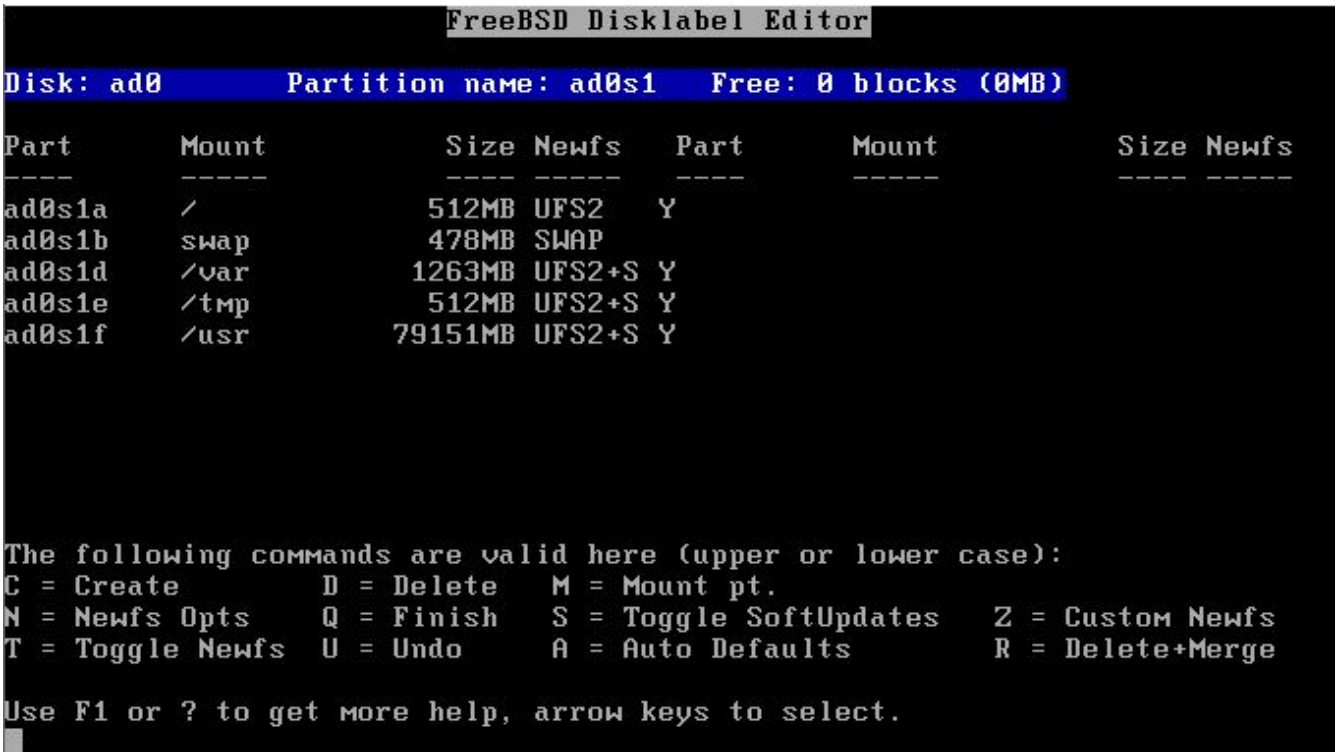
3. Pasos durante la instalación de FreeBSD

3.1. Reservando espacio para el journaling

Normalmente, un ordenador de escritorio típico tiene un disco duro donde se almacena el sistema operativo y los datos del usuario. Sin lugar a dudas, el esquema de partición predeterminado seleccionado por sysinstall es más o menos adecuado: un ordenador de escritorio no necesita una partición /var que tenga un gran tamaño, mientras que a /usr se le asigna la mayor parte del espacio en disco, ya que los datos del usuario y muchos paquetes están instalados en sus subdirectorios.

El particionamiento por defecto (el que se obtiene al presionar `A` en el editor de particiones de FreeBSD, llamado Disklabel) no deja ningún espacio sin asignar. Cada partición que tenga journal, necesita otra partición para el journal. Puesto que la partición más grande es /usr, tiene sentido hacer esta partición algo más pequeña para obtener el espacio necesario para el journaling.

En nuestro ejemplo se usa un disco de 80 GB. La siguiente captura de pantalla muestra las particiones por defecto creadas por Disklabel durante la instalación:



Si esto es más o menos lo que necesitas, los ajustes para el journaling son muy fáciles de hacer. Simplemente utiliza las teclas de desplazamiento para mover la fila resaltada a la partición /usr y presiona la tecla `D` para eliminarla.

Ahora, mueve la fila resaltada al nombre del disco en la parte superior de la pantalla y presiona la tecla `C` para crear una nueva partición para /usr. Esta nueva partición debe ser 1 GB más pequeña (si deseas añadir journaling en /usr solo), o 2 GB (si deseas añadir journaling tanto en /usr como en /var). Desde la ventana emergente, elige crear un sistema de archivos y escribe /usr como punto de montaje.



¿Deberías añadir journaling a la partición /var? Normalmente, el journaling tiene sentido en particiones que sean bastante grandes. Puedes decidir no añadir journaling a /var, hacerlo en un ordenador de escritorio no causará ningun daño. Si no se hace un uso intensivo del sistema de archivos (bastante probable para un ordenador de escritorio), es posible que prefieras asignar menos espacio en disco al journaling.

En nuestro ejemplo, utilizamos journaling tanto en /usr como en /var. Por supuesto puedes ajustar este procedimiento según tus propias necesidades.

Para mantener las cosas lo más sencillas posible, vamos a utilizar sysinstall para crear las particiones necesarias para el journaling. Sin embargo, durante la instalación, sysinstall insiste en pedirte un punto de montaje para cada partición que crees. En este punto, no tienes ningún punto de montaje para las particiones que mantendrán el journal, y en realidad, *ni siquiera las necesitas*. Estas particiones nunca van a ser montadas.

Para evitar estos problemas con sysinstall, vamos a crear las particiones de journal como particiones swap. La partición swap no se monta nunca, y sysinstall no tiene problemas para crear tantas particiones swap como sea necesario. Después del primer reinicio, será necesario editar el archivo /etc/fstab, y eliminar las entradas swap.

Para crear la partición swap, utiliza de nuevo las teclas de flechas para resaltar la parte superior de la pantalla de Disklabel, de forma que el nombre del disco en sí esté resaltado. Después presiona `N`, introduce el tamaño deseado (*1024M*) y selecciona "swap space" en el menú pop-up que aparecerá. Repite para cada journal que quieras crear. En nuestro ejemplo, creamos dos particiones para proporcionar el journal de /usr y /var. El resultado final se muestra en la siguiente captura de pantalla:



Cuando hayas terminado de crear las particiones, te sugerimos que anotes los nombres de las particiones y los puntos de montaje para que puedas consultar fácilmente esa información durante la fase de configuración. Esto ayudará a reducir los errores que puedan dañar tu instalación. La siguiente tabla muestra nuestras notas para la configuración de ejemplo:

Tabla 1. Particiones y journals

Partición	Punto de Montaje	Journal
ad0s1d	/var	ad0s1h
ad0s1f	/usr	ad0s1g

Continua con la instalación como lo harías normalmente. Sin embargo, te sugerimos que pospongas la instalación de software (packages) de terceros hasta que hayas configurado completamente el journaling.

3.2. Arrancando por primera vez

Tu sistema se iniciará como lo haría normalmente, pero deberás editar el archivo /etc/fstab para eliminar las particiones swap que creaste para los journals. Normalmente, la partición swap que usarás es la que tiene el sufijo "b" (por ejemplo, ad0s1b en nuestro ejemplo). Elimina el reto de particiones swap y reinicia para que FreeBSD deje de utilizarlas.

Cuando el sistema vuelva a arrancar, entonces estaremos listos para configurar el journaling.

4. Configurando el journaling

4.1. Ejecutando `gjournal`

Habiendo preparado ya las particiones necesarias, es bastante sencillo configurar el journaling. Necesitaremos cambiar al modo de usuario único así que haz login como `root` y teclea:

```
# shutdown now
```

Presiona `Enter` para obtener el shell por defecto. Necesitaremos desmontar las particiones a las que aplicaremos journaling, en nuestro ejemplo /usr y /var:

```
# umount /usr /var
```

Cargue el módulo requerido por el journaling:

```
# gjournal load
```

Ahora, utiliza tus notas para determinar qué partición se utilizará con cada journal. En nuestro ejemplo, /usr es ad0s1f y su journal será ad0s1g, mientras que /var es ad0s1d y utilizará el journal en ad0s1h. Se necesitarán los siguientes comandos:

```
# gjournal label ad0s1f ad0s1g
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.

# gjournal label ad0s1d ad0s1h
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
```



Si se utiliza el último sector de cualquier partición, `gjournal` devolverá un error. Tendrás que ejecutar el comando utilizando el flag `-f` para forzar una sobrescritura, es decir:

```
# gjournal label -f ad0s1d ad0s1h
```

Dado que se trata de una instalación nueva, es poco probable que se sobrescriba algo.

En este punto, se crean dos nuevos dispositivos, con los nombres ad0s1d.journal y ad0s1f.journal. Estos representan las particiones /var y /usr que tenemos que montar. Antes de realizar el montaje, debemos establecer la flag de journal y borrar la flag de Soft Updates:

```
# tunefs -J enable -n disable ad0s1d.journal
tunefs: gjournal set
tunefs: soft updates cleared

# tunefs -J enable -n disable ad0s1f.journal
tunefs: gjournal set
tunefs: soft updates cleared
```

Ahora, monta los nuevos dispositivos manualmente en sus respectivas ubicaciones (ten en cuenta que ahora podemos usar la opción de montaje `async`):

```
# mount -o async /dev/ad0s1d.journal /var
# mount -o async /dev/ad0s1f.journal /usr
```

Edita /etc/fstab y actualiza las entradas para /usr y /var:

```
/dev/ad0s1f.journal     /usr            ufs     rw,async      2       2
/dev/ad0s1d.journal     /var            ufs     rw,async      2       2
```



¡Asegúrate de que las entradas anteriores sean correctas, o tendrás problemas para arrancar después de reiniciar!

Por último, edita /boot/loader.conf y añade la siguiente línea para que el módulo gjournal(8) se cargue en cada arranque:

```
geom_journal_load="YES"
```

¡Felicidades! Tu sistema ahora está configurado para utilizar el journaling. Puedes escribir `exit` para volver al modo multiusuario o reiniciar para probar su configuración (recomendado). Durante el arranque verás mensajes como los siguientes:

```
ad0: 76293MB XEC XE800JD-00HBC0 08.02D08 at ata0-master SATA150
GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal ad0s1d clean.
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal ad0s1f clean.
```

Después de un apagado forzoso, el mensaje variará ligeramente, por ejemplo:

```
GEOM_JOURNAL: Journal ad0s1d consistent.
```

Normalmente esto significa que gjournal(8) ha utilizado la información del journal provider para devolver al sistema de ficheros a un estado consistente.

4.2. Añadiendo journaling a las nuevas particiones

Si bien el procedimiento anterior es necesario para las particiones con journaling que ya contengan datos, añadir journaling a una partición vacía es un poco más sencillo, ya que tanto los datos como el provider se pueden almacenar en la misma partición. Por ejemplo, supongamos que has instalado un nuevo disco, y que has creado una nueva partición, /dev/ad1s1d . Crear el journal sería tan simple como:

```
# gjournal label ad1s1d
```

El tamaño por defecto del journal será de 1 GB. Puedes ajustarlo utilizando la opción `-s`. El valor se puede introducir en bytes o añadiéndole `K`, `M` or `G` para representar Kilobytes, Megabytes o Gigabytes respectivamente. Date cuenta de que `gjournal` no te permitirá crear tamaños de journal que no sean apropiados por ser demasiado pequeños.

Por ejemplo, para crear un journal de 2 GB, puede usar el siguiente comando:

```
# gjournal label -s 2G ad1s1d
```

Puedes crear un sistema de archivos en tu nueva partición y habilitar el journaling utilizando la opción `-J`:

```
# newfs -J /dev/ad1s1d.journal
```

4.3. Añadiendo soporte de journaling en un kernel personalizado

Si no deseas cargar `geom_journal` como un módulo, puedes añadir la funcionalidad directamente a tu kernel. Edita el archivo de configuración del kernel personalizado y asegúrate de que incluyes estas dos líneas:

```
options UFS_GJOURNAL # Note: This is already in GENERIC

options GEOM_JOURNAL # You will have to add this one
```

Recompila e instala tu kernel siguiendo las instrucciones de instructions in the FreeBSD Handbook.

No te olvides de eliminar la entrada "load" apropiada de /boot/loader.conf si es que la has usado anteriormente.

5. Resolución de problemas del journaling

En la siguiente sección se analizan las preguntas más frecuentes relacionadas con los problemas relacionados con el journaling.

5.1. Durante los periodos de mucha actividad en el disco estoy teniendo un kernel panic. ¿Está relacionado con el journaling?

Es probable que el journaling se llene antes de tener la oportunidad de grabarse (flushed) el disco. Ten en cuenta que el tamaño del journal depende del uso, no del tamaño del provider de datos. Si la actividad del disco es alta, necesitas una partición más grande para el journal. Consulta la nota en la sección Entendiendo el journaling en FreeBSD.

5.2. Cometí algún error durante la configuración, y ahora no puedo arrancar. ¿Se puede arreglar esto de alguna manera?

Has olvidado (o te has equivocado al escribir) la entrada en /boot/loader.conf o hay errores en tu fichero /etc/fstab. Normalmente son fáciles de arreglar. Presiona `Enter` para obtener un shell por defecto en modo usuario único. Después localiza la raíz del problema:

```
# cat /boot/loader.conf
```

Si la entrada `geom_journal_load` no está o está mal escrita, no se crearán los dispositivos que utilizan journaling. Carga el módulo manualmente, monta todas las particiones y continúa con el arranque en modo multiusuario:

```
# gjournal load

ad0: 76293MB XEC XE800JD-00HBC0 08.02D08 at ata0-master SATA150
GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal ad0s1d clean.
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal ad0s1f clean.

# mount -a
# exit
(boot continues)
```

Si, por otro lado, esta entrada es correcta, echa un vistazo al archivo /etc/fstab. Probablemente

encontrarás una entrada que está mal escrita o directamente no está. En este caso, monta todas las particiones restantes a mano y continúa con el arranque en modo multiusuario.

5.3. ¿Puedo eliminar el journaling y volver a mi sistema de archivos estándar con Soft Updates?

Por supuesto. Utiliza el siguiente procedimiento, el cual revierte los cambios. Las particiones que has creado para los journal providers se pueden usar para otros fines, si así lo deseas.

Haz login como `root` y entra en modo usuario único:

```
# shutdown now
```

Desmonta las particiones que tengan journaling:

```
# umount /usr /var
```

Sincroniza los journals:

```
# gjournal sync
```

Para los journaling providers:

```
# gjournal stop ad0s1d.journal
# gjournal stop ad0s1f.journal
```

Borra los metadatos de journaling de datos los dispositivos que utilice:

```
# gjournal clear ad0s1d
# gjournal clear ad0s1f
# gjournal clear ad0s1g
# gjournal clear ad0s1h
```

Borra el flag de journaling del sistema de archivos y restaure el flag de Soft Updates:

```
# tunefs -J disable -n enable ad0s1d
tunefs: gjournal cleared
tunefs: soft updates set

# tunefs -J disable -n enable ad0s1f
tunefs: gjournal cleared
tunefs: soft updates set
```

Vuelve a montar los dispositivos antiguos a mano:

```
# mount -o rw /dev/ad0s1d /var
# mount -o rw /dev/ad0s1f /usr
```

Edita /etc/fstab y restáuralo a su estado original:

```
/dev/ad0s1f     /usr            ufs     rw      2       2
/dev/ad0s1d     /var            ufs     rw      2       2
```

Por último, edita /boot/loader.conf, elimina la entrada que carga el módulo `geom_journal` y reinicia.

6. Lecturas adicionales

El Journaling es una característica relativamente nueva en FreeBSD, y como tal, todavía no está muy bien documentada. Sin embargo, puedes encontrar útiles las siguientes referencias:

- Una nueva sección sobre journaling forma parte ahora del FreeBSD Handbook.
- This post in lista de correo de FreeBSD-CURRENT by gjournal(8)'s developer, Paweł Jakub Dawidek <pjd@FreeBSD.org>.
- This post in Lista de correo para preguntas generales sobre FreeBSD by Ivan Voras <ivoras@FreeBSD.org>.
- Las páginas del manual de gjournal(8) y de geom(8).