

Opciones RAID para el Escritorio Linux

EASY RAIDER

Linux ofrece varias opciones para conseguir la promesa RAID de accesos rápidos al disco duro y de seguridad de datos.

POR MARCUS NASAREK

En los últimos 15 años, la capacidad de los discos duros ha crecido por un factor de al menos mil. Hace 15 años un disco duro típico tenía una capacidad de unos 300 a 500 Mbytes y resultaban tan caros como los discos actuales de 300 a 500 Gbytes. Como actualmente son bastante comunes los PCs con dos o más discos duros, los usuarios domésticos pueden fácilmente permitirse el lujo de la redundancia de datos y de mayores rendimientos, que hasta hace unos años sólo se podían permitir las empresas en sus servidores. La tecnología que hace todo esto posible se conoce como RAID.

Introducción a RAID

RAID [1] fue diseñada hace unos 20 años por los estudiantes de postgrado de Berkeley, David Patterson, Garth Gibson y Randy Katz. En ese tiempo, RAID era la respuesta a un problema difícil: si se necesitaba una gran capacidad de almacenamiento había que elegir entre un disco grande, que era fiable pero caro, o un conjunto de discos pequeños, mucho menos fiables pero baratos.

Uno de los problemas con los discos pequeños era que el esfuerzo de gestión que conllevaba el reemplazo de un disco grande por varios pequeños era demasiado grande. RAID era un esquema para direccionar un array de múltiples discos como una única unidad. La solución propuesta también contaba con el riesgo de la pérdida de datos en caso de que uno de los discos del array

fallara. El documento que publicaron se titulaba “A Case for Redundant Array of Inexpensive Disks (RAID)”. Cuando algún tiempo después el precio de los discos grandes dejó de ser un problema, la palabra *Inexpensive* (Barato) se sustituyó por *Independent* (Independiente).

En el escenario RAID más sencillo pueden agruparse un número de discos para formar una única unidad y presentar esa unidad a la capa de aplicación como un único disco lógico. Esto es lo que los empollones denominan JBOD, “Just a Bundle of Disks” (Un conjunto de discos). Pero RAID puede hacer mucho más. Realiza las funciones de una capa de gestión entre el sistema de ficheros y el hardware permitiendo explotar las características físicas del array de discos. Distintas alternativas RAID ofrecen un rango completo de soluciones para mejorar el rendimiento o mitigar el riesgo de un fallo en el disco.

Las distintas alternativas RAID se denominan Niveles RAID, aunque el término nivel no implica jerarquía; al contrario, los distintos niveles RAID son realmente diseños independientes. Los siete niveles básicos se conocen como RAID 0 a RAID 7. Véase el cuadro titulado “Niveles RAID” para saber más sobre estas alternativas RAID.

Los niveles RAID 0 y 1 son los más importantes para los usuarios domésticos. El nivel 0 es también conocido como “striping” y proporciona el máximo rendimiento posible. El nivel 1 suministra un alto nivel de redundancia y se conoce como “mirroring”. Para

los usuarios que no puedan decidir si necesitan más rendimiento o más redundancia hay un nivel RAID especial conocido como RAID 1.5. Sin embargo, este nivel es una combinación: hacen falta al menos cuatro discos para poder combinar las propiedades de los niveles 0 y 1. RAID 1.5 combina el nivel 0 y el 1 con sólo dos discos, presentando los discos RAID de forma separada a la capa de aplicación. (Figura 1).

Rendimiento con Striping

Striping (RAID 0) implica la unión conjunta de los discos y su direccionamiento en bloques. Cuando una aplicación necesita escribir datos, la controladora RAID distribuye la carga sobre todos los discos.

Desde un punto de vista técnico la controladora simplemente vuelca los datos en la caché del disco duro y a continuación va al siguiente disco. Hasta que la controladora dialoga de nuevo con el primer disco, éste ha tenido tiempo suficiente para finalizar la tarea, debiendo tener los resultados ya listos para ser recogidos. Si se utilizan dos discos, los tiempos de acceso deberían ser teóricamente el doble de rápidos.

Desde luego, RAID 0 no distribuye los datos a nivel de caracteres; por el contrario, utiliza bloques de datos con cierto grado de granularidad. La probabilidad de un fallo se incrementa en función del número de discos que se tengan en el array. Por ello, esta técnica se recomienda en sistemas donde la velocidad es mucho más importante que la redundancia.

La Seguridad de las Imágenes

Mirroring (RAID 1) es exactamente lo contrario a la técnica anterior. Direcciona el disco en series y escribe los datos en cada disco del array; dicho de otro modo, RAID 1 realiza una imagen de los datos en el otro disco. Sin embargo, el almacenaje de copias idénticas de datos en cada disco significa que la capacidad total del array no va a ser mayor que la de un disco individualmente. Si se utilizan dos discos de 200 GBytes, la capacidad total va a ser de 200 GBytes en vez de 400 Gbytes. Mientras que el tiempo de escritura no es diferente del tiempo de escritura en un único disco, los accesos de lectura son mucho más rápidos, ya que la controladora puede leer desde los dos discos en paralelo.

La gran ventaja que RAID 1 ofrece es que, si uno de los discos falla, los datos aún estarán disponibles en el otro disco. Sin embargo, esto sólo es aplicable a los defectos físicos: los errores lógicos, como los borrados

no deseados, las modificaciones o la sobreescritura de los archivos del sistema de ficheros se reflejarán en el otro disco en tiempo real. Dicho de otro modo, RAID 1 no protege de los errores humanos o de los errores software.

La Combinación: RAID 1.5

Para obtener los beneficios de los niveles RAID 0 y 1 al mismo tiempo, se necesita un PC con al menos cuatro discos duros para una solución que los combine RAID 10 (véase el cuadro “Niveles RAID”). Esto no

representa una opción para la mayoría de los usuarios. Algunos fabricantes, siendo conscientes de que la mayoría de los usuarios no desean utilizar cuatro discos, ofrecen una solución de compromiso que tan sólo hace uso de dos discos.

RAID 1.5, que es como se conoce esta solución, divide cada uno de los dos discos en dos áreas. La controladora combina una de estas áreas de cada disco para proporcionar un array RAID 1, mientras que combina las otras dos para proporcionar un array RAID 0. Esta solución se muestra en la Figura 1.

DFI y Epox fueron los primeros fabricantes de placas madre que presentaron esta variante RAID, sin embargo, los vendedores fallaron en su propuesta; ambas placas tan sólo implementaban arrays RAID 1. Los beneficios sólo se obtuvieron en mejoras en las velocidades de lectura. Finalmente, Intel apareció con la controladora ICH6R I/O para los nuevos discos Serial ATA presentando la “Intel Matrix RAID”, que proporciona el soporte RAID 1.5 de forma correcta.

Niveles RAID

Además de los conocidos niveles RAID 0 y 1, existen una gran variedad de diseños RAID. El trabajo original de Patterson, Gibson y Katz cubrió los niveles del 1 al 5. Los niveles 0 y 6 se añadieron posteriormente, ya que fueron soluciones propietarias. Los niveles del 2 al 4, y 6 y 7 son menos populares y han llegado a ser más o menos insignificantes. Las combinaciones de otros diseños han llevado a los niveles conocidos como Nivel 0+1, 10, 30, 15, 50, 51, 55 y RAID-Z.

RAID 0 agrupa los discos y distribuye la carga sobre los discos. Esto hace que se mejoren considerablemente las velocidades de acceso. La capacidad total del array es igual a la suma de las capacidades de los discos. El riesgo de un fallo es bastante alto, ya que este nivel no ofrece ningún tipo de redundancia.

RAID 1 escribe datos idénticos simultáneamente a todos los discos del array. La capacidad final del array es igual a la capacidad de un disco simple. Las velocidades de lectura son normalmente más rápidas que con un único disco y los accesos de escritura son casi de la misma velocidad.

RAID 2 era bastante común en los antiguos mainframes, pero actualmente apenas se utiliza. Un array RAID 2 necesitaba al menos 10 discos. Con un control de errores sofisticado permitía a los usuarios detectar tanto los errores en los discos como los errores de escritura. Suponiendo que se tenían 10 discos, la tasa del tiempo de lectura y escritura era de 1 a 8 comparado con un único disco.

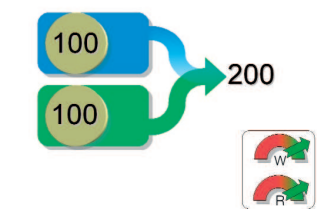
Añadiéndole a los discos duros operaciones a nivel de bits sobre los datos y almacenando los resultados se podían restaurar los datos perdidos cuando un disco fallaba mediante referenciado a los

datos existentes y a los resultados de la suma. El resultado de la suma se denomina paridad. RAID 3 almacena la paridad de los datos para el array en un único disco. Como el disco de la paridad se usa mucho más que los discos de datos, posee una tendencia natural a fallar primero.

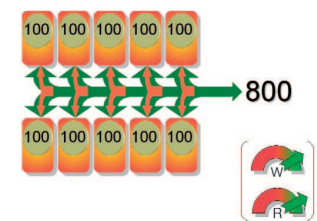
RAID 4 representa una pequeña modificación sobre RAID 3: donde RAID 3 utilizaba ristas de bytes, RAID 4 procesa los bloques completos de datos. Como resultado, RAID 4 puede manejar ficheros pequeños más eficientemente, mientras que RAID 3 sólo lo hacía con ficheros contiguos. Como RAID 3, RAID 4 utiliza un disco separado para la paridad.

RAID 5 es la versión más barata de almacenamiento de datos redundante. Suponiendo al menos un array de tres discos se obtiene el 66 por ciento del espacio para almacenar los datos frente al 50 por ciento proporcionado por RAID 1. Cuanto más discos se añadan, mejores serán los resultados. RAID 5 distribuye los datos y la información sobre la paridad en todos los discos del array, por lo que todos los discos tienen el mismo uso. Como contrapartida, la reconstrucción del RAID tras un fallo lleva más tiempo que con RAID 1.

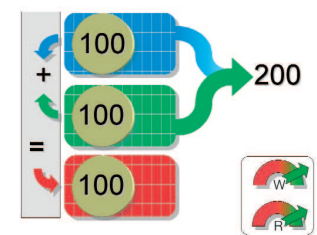
Dependiendo de la aplicación, pueden combinarse los niveles básicos de RAID más o menos de forma arbitraria. RAID 10 implementa un array RAID 0 usando dos arrays RAID 1. RAID 1 proporciona redundancia para una mayor seguridad de los datos, mientras que RAID 0 añade rendimiento. Si se tienen al menos seis discos, una combinación de RAID 5 y 0 (RAID 50) es incluso más eficiente.



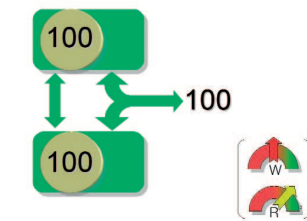
RAID 0: Más velocidad debido a las operaciones en paralelo.



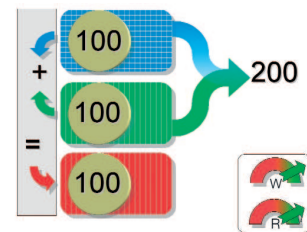
RAID 2: Seguridad en números.



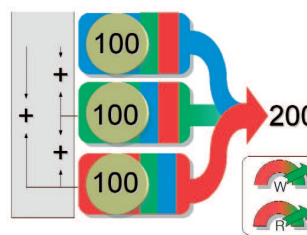
RAID 4: Una mejor solución de RAID 3 para ficheros pequeños.



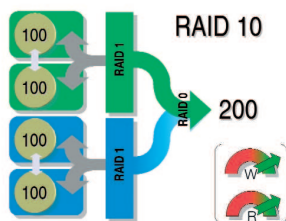
RAID 1: Redundancia a través de la duplicación de discos.



RAID 3: Striping con paridad.



RAID 5: Redundancia y rendimiento de lectura.



RAIDs combinados

Software vs. Hardware

Los diversos diseños RAID se implementaron originalmente como soluciones software. Como el esfuerzo computacional de la solución puramente software era demasiado pesada para las CPUs del momento, no tardó en aparecer la primera tarjeta controladora RAID que fue presentada para proporcionar una solución más potente basada en hardware. Esta clase de implementación se conoce como RAID hardware.

Aparecieron CPUs más potentes, con lo que era más fácil implementar RAID solamente basado en software. El software RAID de bajo presupuesto llegó a ser común en las máquinas domésticas que ejecutaban Linux. Los arrays RAID software normalmente no son tan rápidos como los basados en hardware. La carga de la CPU es normalmente apreciable en las operaciones diarias. Los beneficios de los accesos paralelos son difíciles de explotar con los controladores software.

RAID empezó a popularizarse entre los usuarios domésticos cuando los fabricantes empezaron a añadir chips controladores RAID en sus placas. Hoy en día, la mayoría de las placas modernas, especialmente las que poseen conectores para los discos rápidos SATA, poseen alguna clase de soporte RAID. Los chips integrados en la mayoría de las placas madre tan sólo gestionan parte del proceso, dejando que los drivers compartan el trabajo. Y hacer que esto funcione no es algo que sea siempre trivial.

Explorando RAID

La solución RAID basada en software hace tiempo que fue presentada en Linux. Ahora

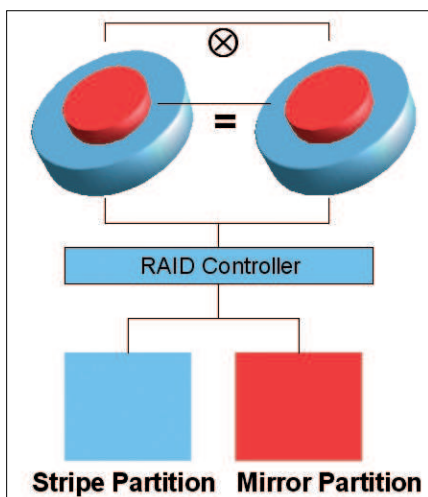


Figura 1: RAID Nivel 1.5 combina las características de RAID 0 y RAID 1

los ordenadores poseen potencia suficiente para procesar la carga adicional de gestionar el array de discos casi transparentemente, y los beneficios del rendimiento de una solución basada en la placa madre pueden ser marginales, ya que recaen en el uso de controladores software.

Para dar los primeros pasos con RAID, es perfectamente viable la instalación de una variante basada en software. Para ello se necesitan los privilegios de root en la máquina en cuestión junto con al menos dos particiones libres y la herramienta Mdmadm [2]. Hay que asegurarse de que las particiones no contengan datos, ya que se borrarán los sistemas de ficheros de estas particiones en el transcurso de nuestro experimento.

Incluso se puede utilizar una memoria USB para nuestros primeros ensayos RAID; sólo habrá que particionar la memoria con una herramienta como Gparted, Qtparted o Cfdisk. El experimento implica los siguientes pasos:

- Crear al menos dos particiones de igual tamaño.
- Usar *mdadm* para combinar las particiones y crear un área RAID.
- Formatear el dispositivo RAID con un sistema de ficheros.
- Acceder a la nueva partición RAID.
- Utilizar *mdadm* para monitorizar el array RAID.

El siguiente ejemplo utiliza un dispositivo USB referenciado como */dev/sdc* y dos particiones referenciadas como */dev/sdc1* y */dev/sdc2*. Para contruir un array RAID 1, hay que introducir el siguiente comando *mdadm*:

```
mdadm --create --verbose
/dev/md0 --level=1
--raid-devices=2
2 /dev/sdc1 /dev/sdc2
```

La opción *--create* crea el RAID; *--verbose* proporciona información más detallada

```

marcus@ubuntu:~/Desktop$ sudo mdadm --detail /dev/md0
/dev/md0:
  Version : 00.90.01
  Creation Time : Sat Feb 11 19:40:18 2006
  Raid Level : raid1
  Array Size : 505920 (494.06 MiB 518.06 MB)
  Device Size : 505920 (494.06 MiB 518.06 MB)
  Raid Devices : 2
  Total Devices : 2
  Preferred Minor : 0
  Persistence : Superblock is persistent

  Update Time : Sat Feb 11 19:48:08 2006
  State : active
  Active Devices : 2
  Working Devices : 2
  Failed Devices : 0
  Spare Devices : 0

  UUID : 68f6c658:22f4814c:aa9f8e65:49667a06
  Events : 0.6

   Number   Major   Minor   RaidDevice State
    -----   -----   -----   -----   -----
     0         8        33         0   active sync  /dev/.static/dev/sdc1
     1         8        34         1   active sync  /dev/.static/dev/sdc2

```

Figura 2: Mdmadm te muestra información sobre el estado de los dispositivos RAID.

sobre el proceso; */dev/md0* es el nombre del dispositivo RAID resultante y *--level=1* establece el nivel de redundancia al nivel RAID 1. El último parámetro pasa el número de particiones RAID y sus nombres.

Antes de poder utilizar el array RAID recién creado hay que formatear (utilizando ReiserFS en nuestro ejemplo) y montar el sistema de ficheros resultante:

```
mkreiserfs /dev/md0
mkdir /media/RAID
mount /dev/md0 /media/RAID
```

Ahora se puede utilizar el directorio */media/RAID* como cualquier otro disco. Sin embargo, en el fondo, los datos se escribirán en dos particiones para proporcionar la redundancia. El comando *mdadm -detail /dev/md0* presenta información del estado del array RAID. La Figura 2 muestra la salida de este comando.

Conclusiones

RAID ofrece diferentes beneficios para los usuarios domésticos. Con él se obtienen mayor redundancia de datos y accesos más rápidos al disco. La herramienta Mdmadm ayuda a implementar cualquiera de los niveles RAID presentados en este artículo. ■

Recursos

- [1] RAID en Wikipedia: <http://en.wikipedia.org/wiki/RAID>
- [2] Mdmadm: <http://cgi.cse.unsw.edu.au/~neilb/mdadm>