
Requisitos

Los conocimientos necesarios para la certificación LPIC-1:

- Nociones de base sobre los dispositivos, las particiones y los sistemas de archivos.
- Edición de archivos.
- Comandos de gestión de directorios y archivos.

Objetivos

Al final de este capítulo, deberá poder:

- Configurar e implementar software RAID.
- Administrar los diferentes tipos de discos.
- Gestionar los dispositivos iSCSI.
- Configurar y administrar LVM (*Logical Volume Manager*).

A. Gestión de los dispositivos de almacenamiento

Este tema está dividido en tres partes con pesos diferentes.

1. Configuración de discos RAID

Peso	3
Objetivos	Configurar e implementar software RAID. Esto incluye los niveles de RAID 0, 1 y 5.

a. Competencias principales

- Archivos de configuración y herramientas de gestión del software RAID.

b. Elementos empleados

- `mdadm.conf`
- `mdadm`
- `/proc/mdstat`
- Partición de tipo `0xFD`

2. Optimizar el acceso a los dispositivos de almacenamiento

Peso	2
Objetivos	Configurar el núcleo para administrar diferentes tipos de discos. Conocer las herramientas de software para listar y modificar la configuración de los dispositivos iSCSI.

a. Competencias principales

- Herramientas y comandos para configurar el DMA para los dispositivos IDE, incluyendo ATAPI y SATA.
- Herramientas y comandos para configurar discos SSD (*Solid State Drive*), incluyendo AHCI y NVMe.
- Herramientas y comandos para configurar o analizar los recursos del sistema (por ejemplo, las interrupciones).
- Conocimientos básicos del comando `sdparm` y su uso.
- Herramientas y comandos de gestión de dispositivos iSCSI.
- Conocimientos básicos de SAN, incluyendo los protocolos específicos (AoE, FCoE).

b. Elementos empleados

- `hdparm`, `sdparm`
- `nvme`
- `tune2fs`
- `fstrim`
- `sysctl`
- `/dev/hd*`, `/dev/sd*`, `/dev/nvme*`

- `iscsiadm, scsi_id, iscsid y iscsid.conf`
- `WWID, WWN, n° LUN`

3. Logical Volume Manager

Peso	3
Objetivos	Crear y suprimir volúmenes lógicos, grupos de volúmenes y volúmenes físicos. Este objetivo incluye las instantáneas (<i>snapshots</i>) y el redimensionamiento de los volúmenes lógicos.

a. Competencias principales

- Herramientas de la suite LVM.
- Redimensionar, renombrar, crear, suprimir volúmenes lógicos, grupos de volúmenes y volúmenes físicos.
- Crear y mantener instantáneas (*snapshots*).
- Activar grupos de volúmenes.

b. Elementos empleados

- `/sbin/pv*`
- `/sbin/lv*`
- `/sbin/vg*`
- `mount`
- `/dev/mapper/`
- `lvm.conf`

B. Configuración de los discos RAID

La tecnología RAID (*Redundant Array of Independent Disks*) permite combinar diferentes dispositivos para que sean vistos como un solo espacio de almacenamiento para las aplicaciones. De esta manera se puede mejorar el tiempo de acceso y/o la fiabilidad de los dispositivos de almacenamiento. Las diferentes técnicas empleadas se definen con respecto al nivel de RAID usado. Los niveles más corrientes son RAID 0, 1 y 5; los que se estudiarán en el marco de la certificación.

Se puede administrar el RAID en modo hardware, con los controladores de discos especializados o en modo software, desde el sistema operativo.

Linux implementa un piloto de gestión de software del RAID, el piloto `md` (*Multiple Device piloto*), que gestiona los niveles más corrientes de RAID, 0, 1 y 5, y que debe ser estudiado en el marco de la certificación.

☞ Otras soluciones pueden implementarse para gestionar el software RAID en Linux: RAID LVM o RAID directamente soportado por el gestor de sistema de archivos ZFS o Btrfs.

1. Los principales niveles de RAID

a. RAID 0

RAID 0 (agregación de bandas, *striping*) combina distintos discos en un solo conjunto. Los bloques de datos se reparten en bandas de tamaño idéntico que están repartidas uniformemente en los diferentes discos. Las operaciones de entrada/salida pueden ser, por lo tanto, muy rápidas, ya que los controladores de los discos las pueden efectuar de manera simultánea.

Sin embargo, la fiabilidad del conjunto es bastante baja ya que al perder un disco perderá el conjunto de los datos. No hay redundancia en los datos almacenados y la coherencia de los volúmenes lógicos se destruye en el caso de fallo en un disco.

El espacio de almacenamiento útil de un conjunto RAID 0 es igual a la capacidad útil del más pequeño de los discos multiplicado por el número de discos que lo componen, ya que no hay redundancia de datos y que las bandas de datos están repartidas de manera uniforme en los discos (cada disco tiene que tener el mismo número de bandas).

Ventajas:

- Rapidez de lectura y escritura del conjunto de los bloques.
- Uso óptimo del espacio de los discos, siempre y cuando los discos sean del mismo tamaño.

Inconvenientes:

- No hay redundancia de datos, por lo tanto, no hay tolerancia frente a fallos.
- La pérdida de un disco compromete el conjunto de los datos almacenados, la fiabilidad del conjunto es igual a la fiabilidad del menos fiable de los discos utilizados.

b. RAID 1

RAID 1 (discos en espejo, *mirroring*) combina distintos discos en un solo conjunto. Cada bloque de datos útil está escrito en cada uno de los discos. Esta redundancia asegura una fiabilidad excelente al conjunto, mayor cuanto mayor sea el número de discos. Mientras quede un disco operativo, los datos estarán intactos y mientras el controlador de ese disco funcione, esos datos seguirán estando accesibles.

Las operaciones de lectura pueden ser más rápidas, porque los controladores las pueden efectuar simultáneamente.

El espacio de almacenamiento útil del conjunto RAID 1 es igual a la capacidad útil del más pequeño de los discos.

Ventajas:

- Excelente tolerancia frente a fallos, proporcional al número de discos combinados (y al número de controladores de discos para la accesibilidad).
- Buen rendimiento en lectura.

Inconvenientes:

- El espacio de disco necesario es al menos dos veces el tamaño del espacio de disco útil.
- Puede haber un impacto en el rendimiento en escritura, incluso si en general las escrituras se hacen simultáneamente en los diferentes discos.

c. RAID 5

RAID 5 (agregación en bandas con paridad) combina al menos tres discos en un solo conjunto. Los bloques de datos están repartidos en bandas de tamaño idéntico, repartidas uniformemente en los diferentes discos, excepto en uno de ellos. Para cada conjunto de bandas, una banda de paridad estará calculada y escrita en el disco restante. La ubicación del conjunto de paridad estará repartida alternativamente en los discos.

En caso de pérdida de una banda de datos, la banda de paridad permitirá reconstituirla, asegurando la tolerancia frente a fallos. Pero este mecanismo solamente es eficaz en el caso en que solamente haya un disco que no esté accesible. Si hubiera dos discos o más que no estuvieran accesibles, habría pérdida de datos en el conjunto de los discos. Mientras que un disco no esté operativo, no habrá tolerancia frente a fallos para las nuevas escrituras. Es por esto por lo que el conjunto de discos en RAID 5 integra generalmente un disco de emergencia (*spare disk*), que solo se utiliza para reemplazar un disco defectuoso.

Cuando el disco defectuoso haya sido reparado o reemplazado, hay que reconstruir el conjunto RAID 5 reconstituyendo los datos y las bandas de paridad para escribirlos en el disco de reemplazo.

Las operaciones de lectura pueden ser muy rápidas, ya que las efectúan diferentes controladores de discos. Las escrituras pueden ser lentas, a causa del cálculo y de la escritura de la banda de paridad.

El espacio de almacenamiento útil de un conjunto RAID 5 es igual a la capacidad del más pequeño de sus discos, multiplicada por el número de discos que lo componen, menos 1 a causa de las bandas de paridad y menos 2 si hay un disco de emergencia (*spare*).

Ventajas:

- Tolerancia frente a fallos limitada a un disco. Mientras haya un disco en fallo, ya no habrá más tolerancia frente a fallos, excepto en el caso de que exista un disco de emergencia.

Inconvenientes:

- Una parte del espacio del disco no puede ser utilizado para los datos.
- El rendimiento en escritura puede verse impactado por el cálculo de la paridad.

2. Configuración del RAID

El piloto `md` es un módulo del núcleo que implementa el software RAID en un conjunto de dispositivos de almacenamiento, discos duros completos y/o particiones de discos duros.

El comando `mdadm` permite configurar volúmenes RAID y administrarlos. Forma parte del paquete `mdadm`.

a. Creación de un volumen RAID

Un volumen RAID está compuesto por distintos espacios de almacenamiento que pueden ser discos duros enteros o particiones de disco duros.

La creación de un volumen RAID se hace con la opción `-C` del comando `mdadm`. Hay que especificar el nombre del nuevo volumen o su número, el nivel de RAID que se quiera implementar y la lista de los espacios de almacenamiento que se le va a destinar.

📖 El archivo de configuración del comando, generalmente `/etc/mdadm/mdadm.conf`, es facultativo en las versiones recientes y no se crea durante la instalación del paquete.

Sintaxis

```
mdadm -C ArchivoEspecialVol -l|--level=Nivel -n|--raid-devices=NúmeroDevRaid
[ -x|--spare-devices=NúmEmergencia ] ArchivoEspecial1 ... ArchivoEspecialN
```

Principales parámetros

-C ArchivoEspecialVol	Archivo especial del volumen RAID creado.
-l --level=Nivel	Nivel de RAID.
-n --raid-devices=NúmeroDevRaid	Número de espacios de almacenamiento activos.
-x --spare-devices=NúmEmergencia	Número de espacios de almacenamiento de emergencia.
ArchivoEspecial1 ... ArchivoEspecialN	Espacios de almacenamiento.

Descripción

La opción `-C` crea un nuevo volumen RAID. El archivo especial que se le asociará, `ArchivoEspecialVol`, se encuentra generalmente bajo la forma `/dev/mdX`, donde `X` es un número, pero esto no es obligatorio.

La opción `-n` especifica el número de espacios de almacenamiento que utilizará el volumen. Debe ser igual o superior al número de elementos de la lista `ArchivoEspecial1...`, `ArchivoEspecialN` menos el número de espacios de almacenamiento de emergencia, indicado con la opción `-x`.

Una vez el volumen RAID creado, este puede ser usado inmediatamente. Es visto como un dispositivo en modo bloque, podemos por lo tanto crear un sistema de archivos o hacer de él un volumen físico LVM.

Ejemplos

Se utilizan dos particiones de discos duros, `/dev/sdb` y `/dev/sdc`, para crear un volumen RAID de nivel 1 (espejo). Como las particiones contienen sistemas de archivos y son de tamaños diferentes, el comando muestra una advertencia y solicita una confirmación:

```
mdadm -C /dev/md0 -l 1 -n 2 /dev/sdb1 /dev/sdc1
mdadm: partition table exists on /dev/sdb1
mdadm: partition table exists on /dev/sdb1 but will be lost or
      meaningless after creating array
mdadm: Note: this array has metadata at the start and
      may not be suitable as a boot device. If you plan to
      store '/boot' on this device please ensure that
      your boot-loader understands md/v1.x metadata, or use
      --metadata=0.90
mdadm: partition table exists on /dev/sdc1
mdadm: partition table exists on /dev/sdc1 but will be lost or
      meaningless after creating array
Continue creating array? y
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md0 started.
```

Se ha creado el volumen RAID:

```
ls -l /dev/md0
```

```
brw-rw---- 1 root disk 9, 0 11 agosto 13:49 /dev/md0
```

El archivo /dev/md0 es un archivo especial de bloques.

El comando `blkid` muestra información de los dos espacios de almacenamiento que componen el volumen RAID:

```
blkid /dev/sdb1 /dev/sdc1
```

```
/dev/sdb1: UUID="c8b9feb6-f72c-4aa3-08c1-99684961cd64" UUID_SUB="d635126c-d86b-44ca-3545-8a68feb4c033" LABEL="srvdebian:0" TYPE="linux_raid_member" PARTUUID="2cab24c8-01"
```

```
/dev/sdc1: UUID="c8b9feb6-f72c-4aa3-08c1-99684961cd64" UUID_SUB="dc72ae20-2586-6536-e0fe-4460ab87385c" LABEL="srvdebian:0" TYPE="linux_raid_member" PARTUUID="2cab241f-01"
```

Las dos particiones son de tipo RAID Linux y han recibido una etiqueta `srvdebian:0`, `srvdebian` es el nombre de la máquina.

b. Tipo de partición RAID

Cuando se crea una partición en un disco duro, se le puede atribuir un tipo bajo la forma de un valor numérico predefinido. Para las particiones que usan el modo de particionamiento tradicional en el mundo del PC, llamado MBR (*Master Boot Record*), los tipos principales son:

0x82	Linux swap
0x83	Linux
0xFD	Linux RAID auto
0xE8	LUKS (cifrado)
0x07	NTFS
0x0C	FAT32

Para las particiones que usan el modo de particionamiento más reciente y general, GPT (*GUID Partition Table*, con *GUID= Globally Unique Identifier*), el tipo `0xFD00` está reservado para las particiones RAID Linux.

En las antiguas versiones de RAID Linux, el tipo de partición `0xFD` (Linux RAID auto) servía en algunos casos (en particular para la carga inicial del núcleo) para detectar las particiones que componen los conjuntos RAID. Hoy día, este tipo de partición ya no es necesaria para que se activen los volúmenes RAID.

Se puede, sin embargo, fijar con este tipo las particiones usadas en los volúmenes RAID, para reconocerlas cuando se usen herramientas de gestión de particiones u algunas otras herramientas.

Para cambiar el tipo de una partición existente, se pueden usar distintos comandos.

Ejemplo

Para cambiar el tipo de la partición MBR `/dev/sdc1` a RAID:

```
sfdisk --id /dev/sdc 1 fd
```

o la sintaxis más reciente:

```
sfdisk --part-type /dev/sdb 1 fd
```

o con el comando interactivo `fdisk`:

```
fdisk /dev/sdc
```

```
t
```

```
Partition number (1,2, default 2): 1
```

```
Partition type (type L to list all types): fd
```

c. Estado de un volumen RAID

La opción `-D` del comando `mdadm` muestra las características y el estado de un volumen RAID.

Sintaxis

```
mdadm -D|--detail ArchivoEspecialVol
```

Descripción

El comando muestra todas las características del volumen RAID indicado: su nivel de RAID, su estado así como el de sus componentes y el estado de estos.

Ejemplo

Características y estado del volumen RAID nivel 1 `/dev/md0`:

```
mdadm -D /dev/md0
```

```
/dev/md0:
```

```
    Version : 1.2
    Creation Time : Thu Aug 11 13:49:06 2022
      Raid Level : raid1
      Array Size : 30026688 (28.64 GiB 30.75 GB)
    Used Dev Size : 30026688 (28.64 GiB 30.75 GB)
      Raid Devices : 2
    Total Devices : 2
      Persistence : Superblock is persistent

    Update Time : Thu Aug 11 14:03:44 2022
      State: clean, resyncing
    Active Devices: 2
    Working Devices: 2
    Failed Devices: 0
    Spare Devices: 0
```

```
Consistency Policy: resync
```

```
Resync Status: 25% complete
```