





# ¿Quién soy yo?



Alejandro Calderón Mateos  
**Profesor Ayudante**  
**Área de Arquitectura y Tecnología de los Computadores**

---



Tel: **+34 91 624 9497**



Fax: **+34 91 624 9129**



E-mail: acaldero  arcos.inf.uc3m.es

Home: <http://arcos.inf.uc3m.es/~acaldero>



Home2: **2.2 B 17 @ Leganés**



# ¿Quién soy yo?

**[http://gul.uc3m.es/wiki/doku.php?id=fortunes\\_de\\_la\\_uc3m](http://gul.uc3m.es/wiki/doku.php?id=fortunes_de_la_uc3m)**

Profesor: "¿Alguno ha compilado un kernel alguna vez?"

Solo levanta la mano un alumno

Profesor: "Bueno...¿quien tiene instalado Linux en casa?"

Levantan la mano 4 o 5

Profesor: "Vale. Al menos espero que tengais más vida social que yo."

-- Alejandro Calderón (Friki incomprendido)



# Sistemas RAID

- El sistema RAID fue propuesto por primera vez en 1988 por David A. Patterson, Garth A. Gibson y Randy H. Katz en la publicación “Un Caso para Conjuntos de Discos Redundantes Económicos (RAID)”.
- Permite gestionar múltiples discos:  
<http://es.wikipedia.org/wiki/RAID>



# wikiPeDIA: RAID

RAID - Wikipedia, la enciclopedia libre - Windows Internet Explorer

http://es.wikipedia.org/wiki/RAID

File Edit View Favorites Tools Help

Google raid Buscar PageRank 52 bloqueado(s) Opciones raid >

W RAID - Wikipedia, la enc... Google

Registrarse/Entrar

artículo discusión editar historial

## RAID

RAID (Redundant Array Of Independent/inexpensive Disks) es un término inglés que hace referencia a un **conjunto de discos redundantes independientes/baratos**. Este tipo de dispositivos se utilizan para aumentar la integridad de los datos en los [discos](#), mejorar la tolerancia a los fallos y errores y mejorar el rendimiento. En general permiten proveer discos virtuales de un tamaño mucho mayor al de los discos comúnmente disponibles. Inicialmente un sistema RAID era un conjunto de discos redundantes económicos.

Oficialmente los sistemas RAID se implementan en 7 configuraciones o niveles: RAID 0 a RAID 6. También existen combinaciones de niveles de RAID, las combinaciones más comunes son RAID 10 y RAID 0+1. Los sistemas RAID son comúnmente implementados con discos de la misma capacidad para todo el conjunto.

A nivel práctico y comercial, sólo los RAID impares, junto a las combinaciones de estos, se han impuesto en el mercado: RAID 1, 3, 5, 7, 10, y 0+1. Destacan por su aceptación sobre los demás el RAID 1, 5, 10, y 0+1.

Tabla de contenidos [esconder]

1 Hardware vs. Software

2 Niveles de RAID

2.1 RAID 0: Conjunto de discos divididos sin tolerancia a fallos (No Redundante)

2.1.1 Características y Ventajas

2.1.2 Desventajas

2.1.3 Aplicaciones Recomendadas

2.2 RAID 1: Espejo y Duplexing (Espejo)

2.2.1 Características

2.2.2 Ventajas

2.2.3 Desventajas

2.2.4 Aplicaciones Recomendadas

2.3 RAID 2: Código de Corrección de Error

2.4 RAID 3: Paridad de intervalo de bit (Paridad de Richard M. Price)

2.5 RAID 4: Unidad de paridad dedicada (Paridad de intervalo de bloque)

2.5.1 Características

Ir Buscar

herramientas

Lo que enlaza aquí Cambios en enlazadas Subir archivo Subir a Commons Páginas especiales Versión para imprimir Enlace permanente Citar este artículo

Internet 100%



# wikiPeDIA: RAID

**RAID (Redundant Array Of Independent/Inexpensive Disks)** es un término inglés que hace referencia a un conjunto de discos redundantes independientes/baratos. Este tipo de dispositivos se utilizan para aumentar la integridad de los datos en los **discos**, mejorar la tolerancia a los fallos y errores y mejorar el rendimiento.

En general permiten proveer discos virtuales de un tamaño mucho mayor al de los discos comúnmente disponibles.

Inicialmente un sistema RAID era un conjunto de discos redundantes económicos.

The screenshot shows a Mozilla Firefox window displaying the Spanish Wikipedia article on RAID. The title is "RAID - Wikipedia, la enciclopedia libre". The page content discusses RAID as a redundant array of independent/inexpensive disks, its purpose of increasing data integrity, and its historical nature as a collection of redundant economic disks. It also mentions the seven levels of RAID (RAID 0 to RAID 6) and their characteristics. The browser interface includes a toolbar, menu bar, address bar, and various extension icons at the bottom.



<http://www.cs.berkeley.edu/~pattrsn/Arch/RAID.jpg>



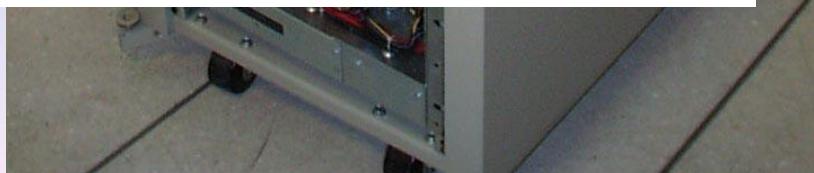
## Xserve RAID

Up to 7TB in 14 hot-swap drives

High-performance 2Gb Fibre Channel

Starting at just \$5999

<http://www.apple.com/xserve/raid/>





# Niveles RAID

□ Hay varios niveles de RAID:

- Niveles básicos:

- ✓ RAID 0
- ✓ RAID 1
- ✓ RAID 2
- ✓ RAID 3
- ✓ RAID 4
- ✓ RAID 5

[http://www.weethet.nl/dutch/hardware\\_raid.php](http://www.weethet.nl/dutch/hardware_raid.php)

<http://storage-system.fujitsu.com/jp/term/raid/>

- Combinaciones:

- ✓ RAID 10
- ✓ RAID 51
- ✓ Etc.

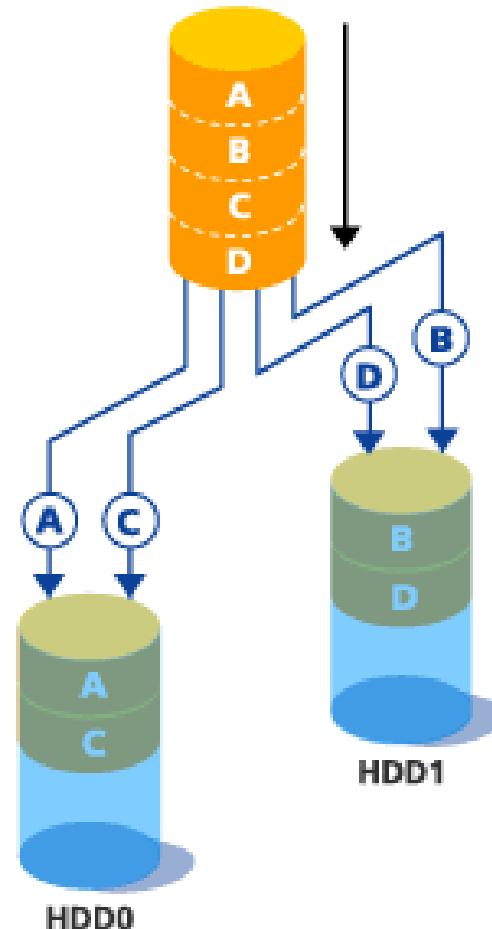


# RAID 0

- RAID 0 (*striping*)

- No ofrece tolerancia a fallos
- Gran velocidad en las operaciones de lectura y escritura
  - Si la controladora quiere

CPUからABCDというデータの書き込み指示



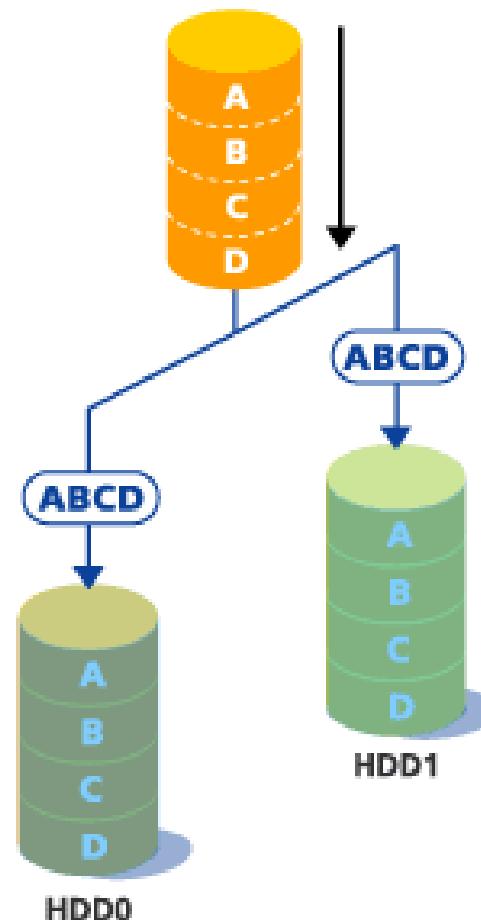
複数ディスクへ分散して書き込む  
ストライピング処理を行う



# RAID 1

- RAID 1 (*mirroring*)
- Tolerancia a 1 fallo
- Gran velocidad en las operaciones de lectura

CPUからABCDというデータの書き込み指示

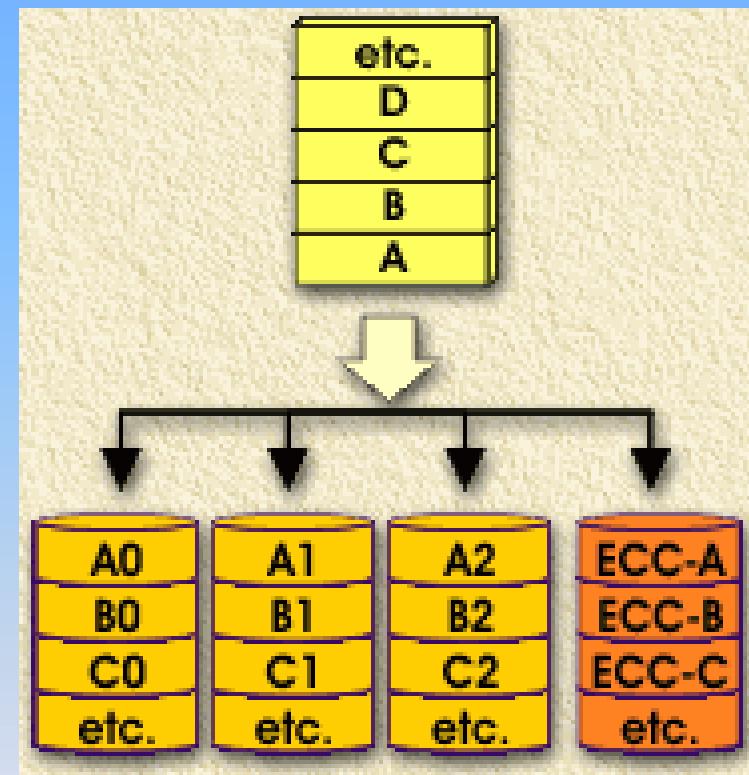


データを2台のディスクに同時に書き込む



# RAID 2

- ❑ RAID 2
- ❑ Detección de fallo
- ❑ Código Hamming

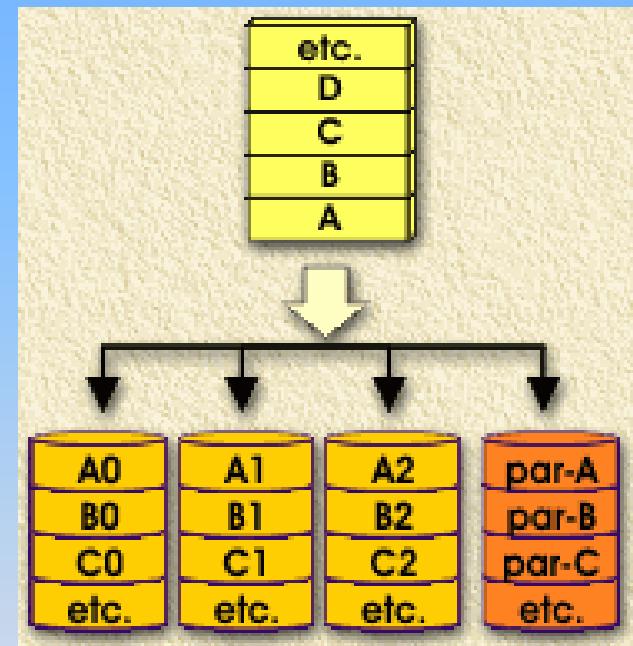




# RAID 3

- RAID 3 (*striping with dedicated parity, bit level*)

- Tolerancia a 1 fallo
- Paridad con cada bit de un bloque de disco (datos)

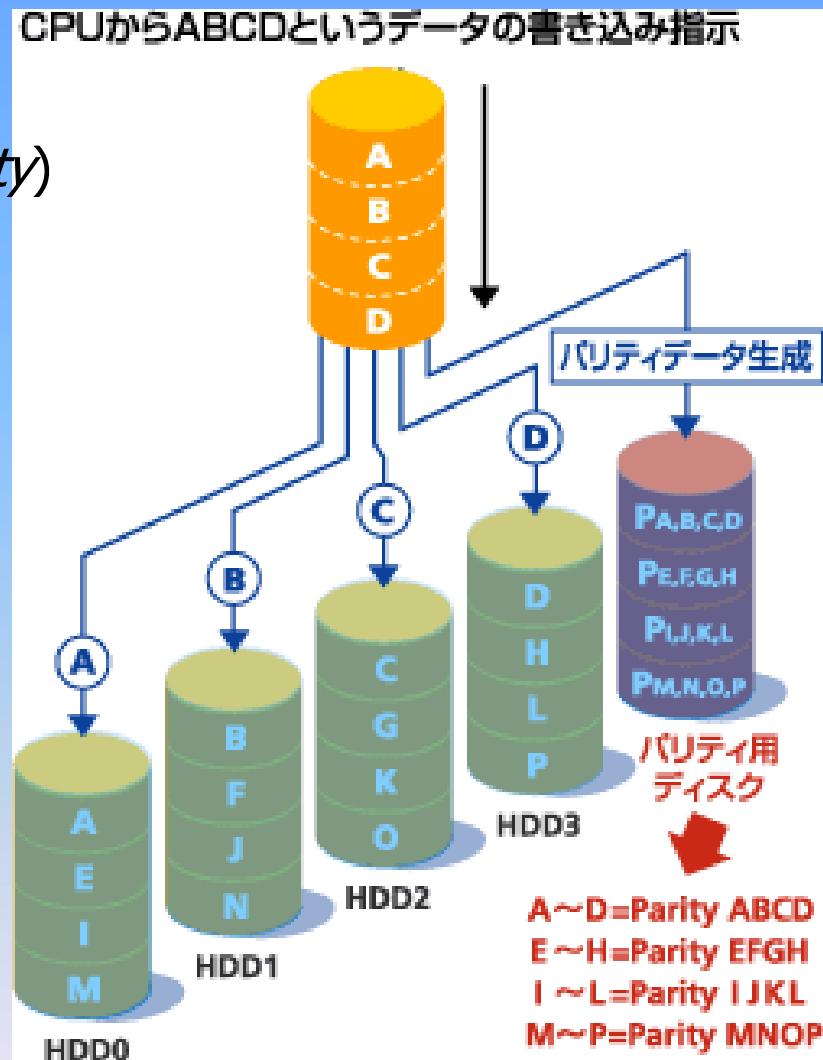




# RAID 4

## □ RAID 4 (*striping with dedicated parity*)

- Tolerancia a 1 fallo
- Paridad con cada bloque de disco (datos)
- Menor consumo de disco
- Más costoso el mantenimiento de la paridad
- Disco de paridad puede ser un cuello de botella en el rendimiento



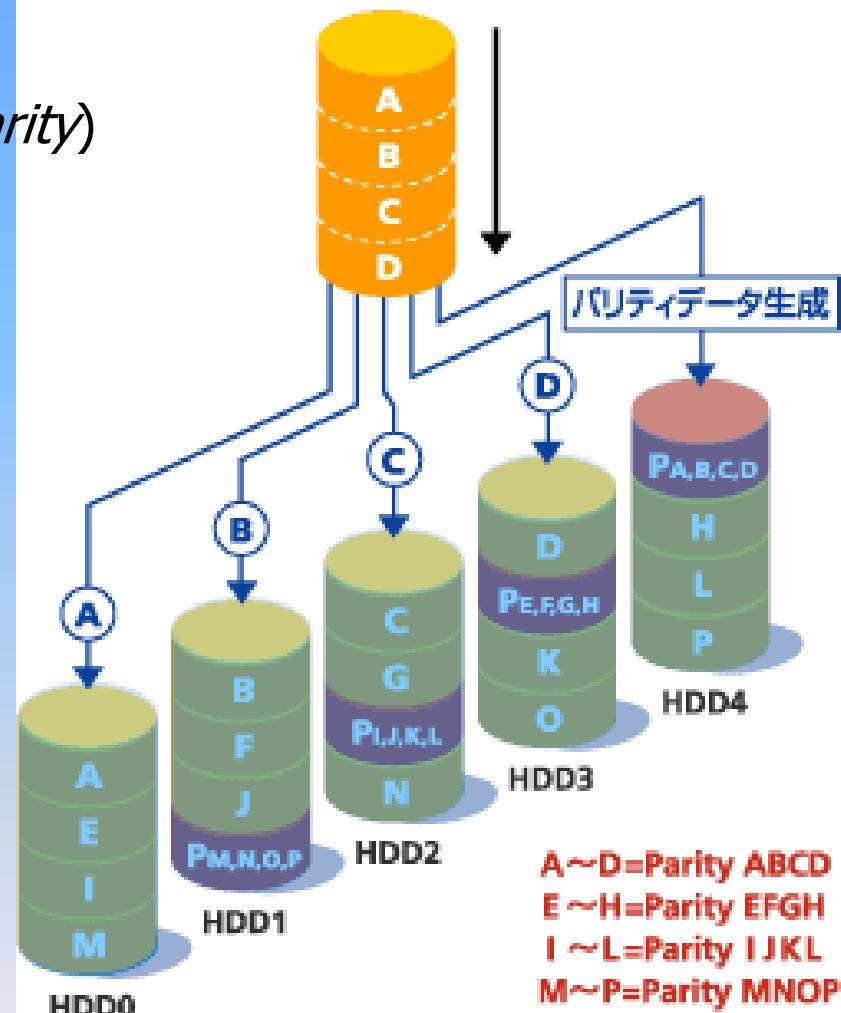


# RAID 5

## □ RAID 5 (*striping with distributed parity*)

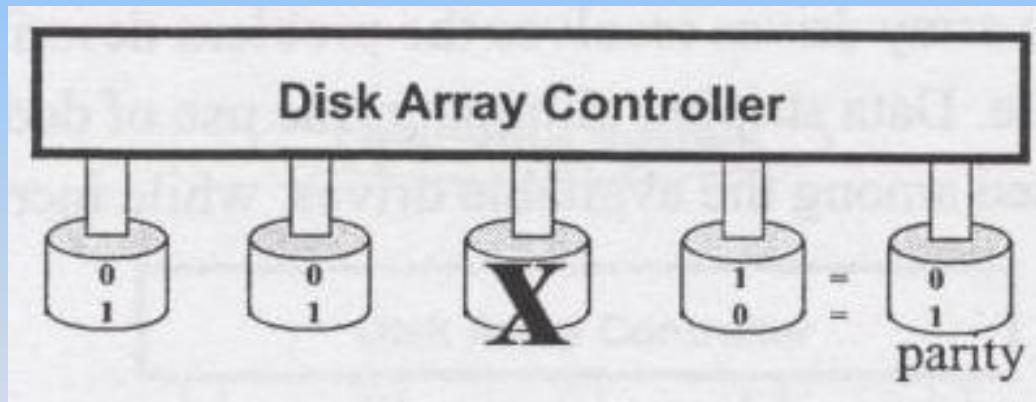
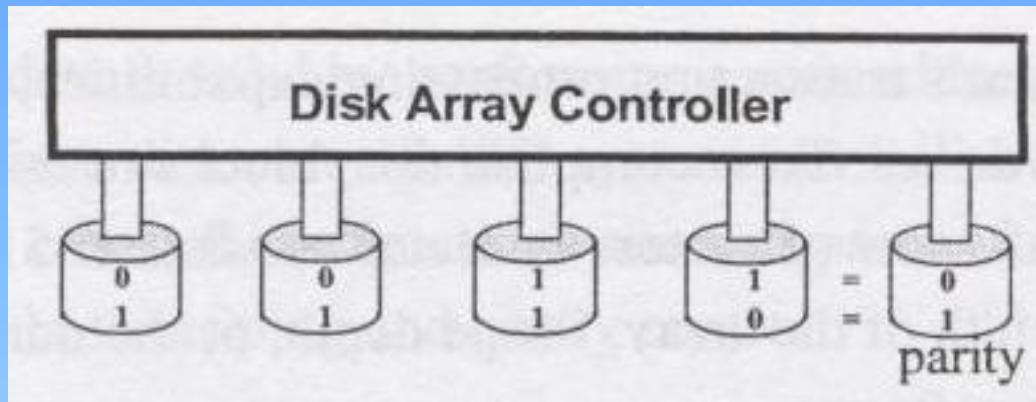
- Tolerancia a 1 fallo
- Como RAID 4 pero con bloques de paridad distribuidos entre todos los discos
  - Evita cuellos de botella

CPUからABCDというデータの書き込み指示





# Uso de la paridad

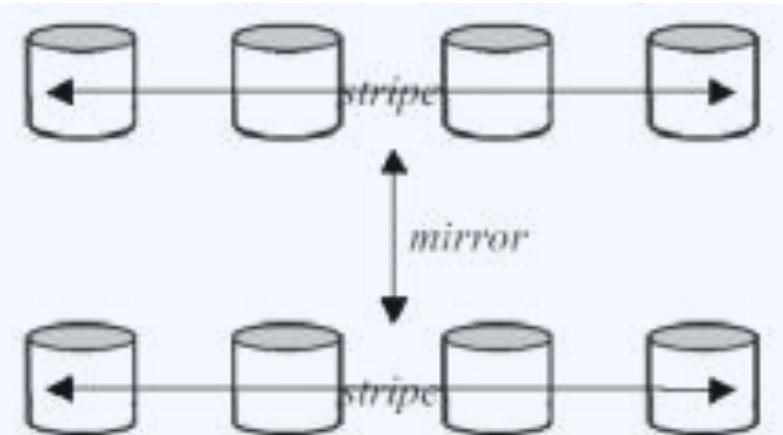


<http://www.commodore.ca/windows/raid5/raid5.htm>

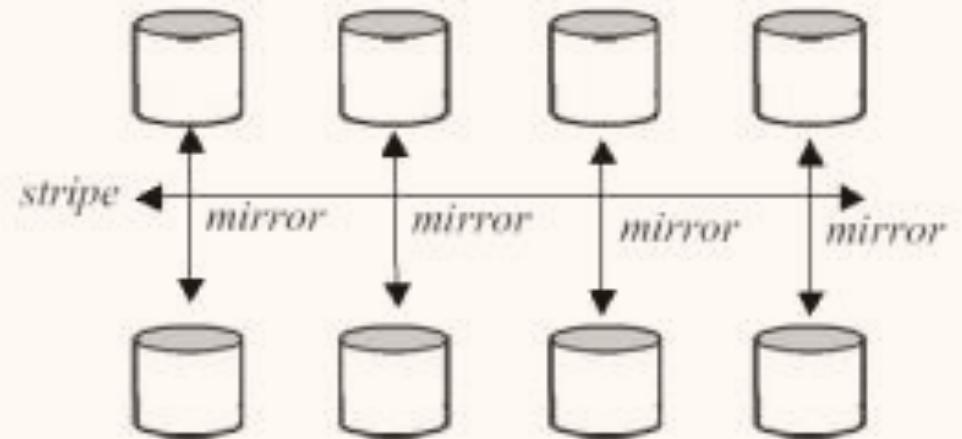


# Ejemplo de RAID combinado

RAID 0 + 1    ***(stripped and mirrored)***

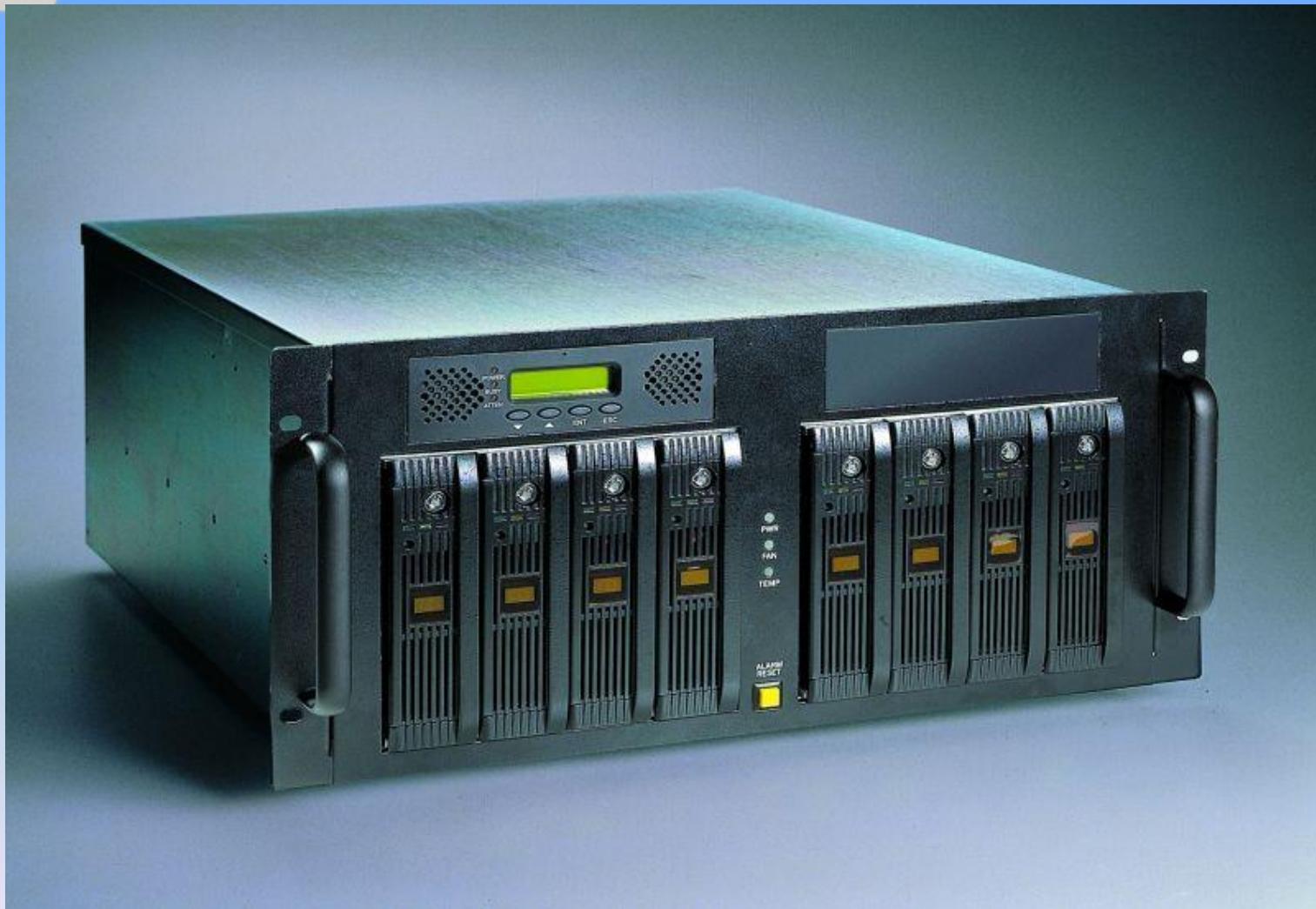


RAID 1 + 0    ***(mirrored and striped)***





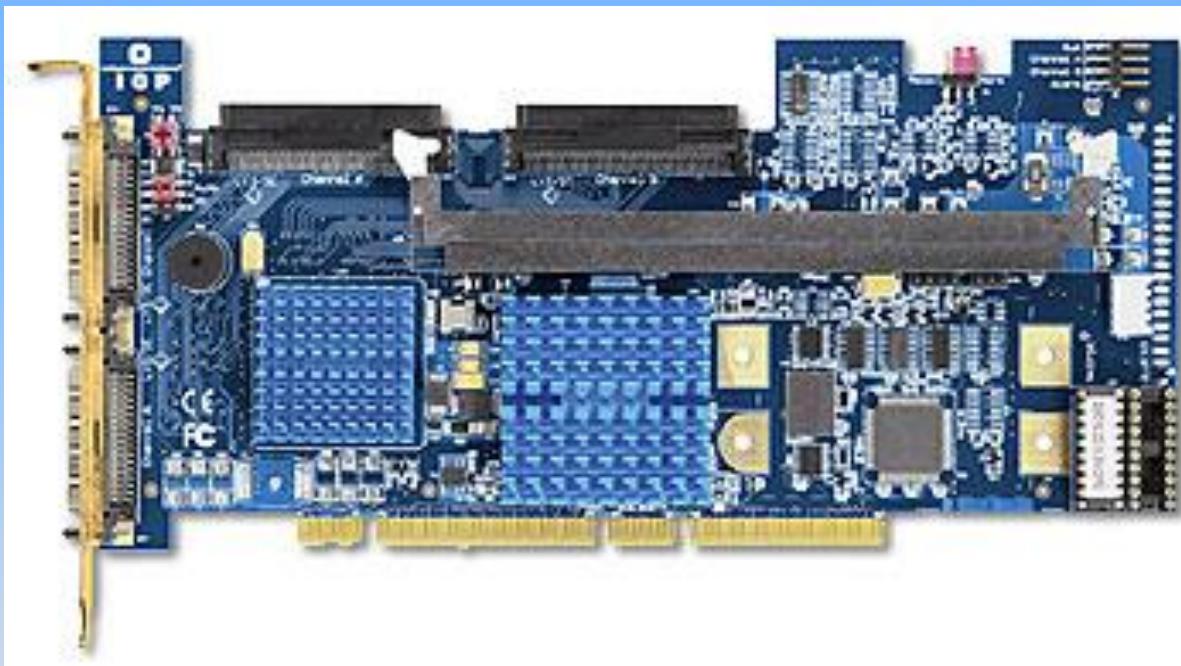
# Ingrediente1: DISCOS





## Ingrediente2: Controladora RAID

- Hardware



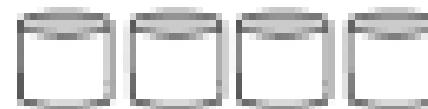
- Software



# Ingrediente3: ESQUEMAS RAID

<http://bebás.vlsm.org/v06/Kuliah/SistemOperasi/BUKU/img/c75-3-2-f1-RAID.png>

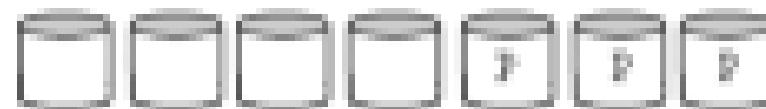
RAID 0: Non-redundant striping



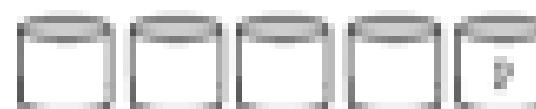
RAID 1: Mirrored disks



RAID 2: Memory-style ECC



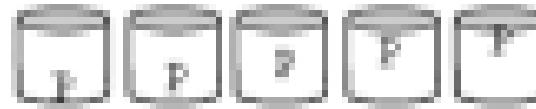
RAID 3: Bit-interleaved parity



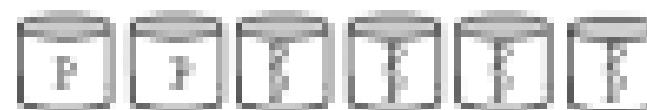
RAID 4: Block-interleaved parity



RAID 5: Block-interleaved distributed parity



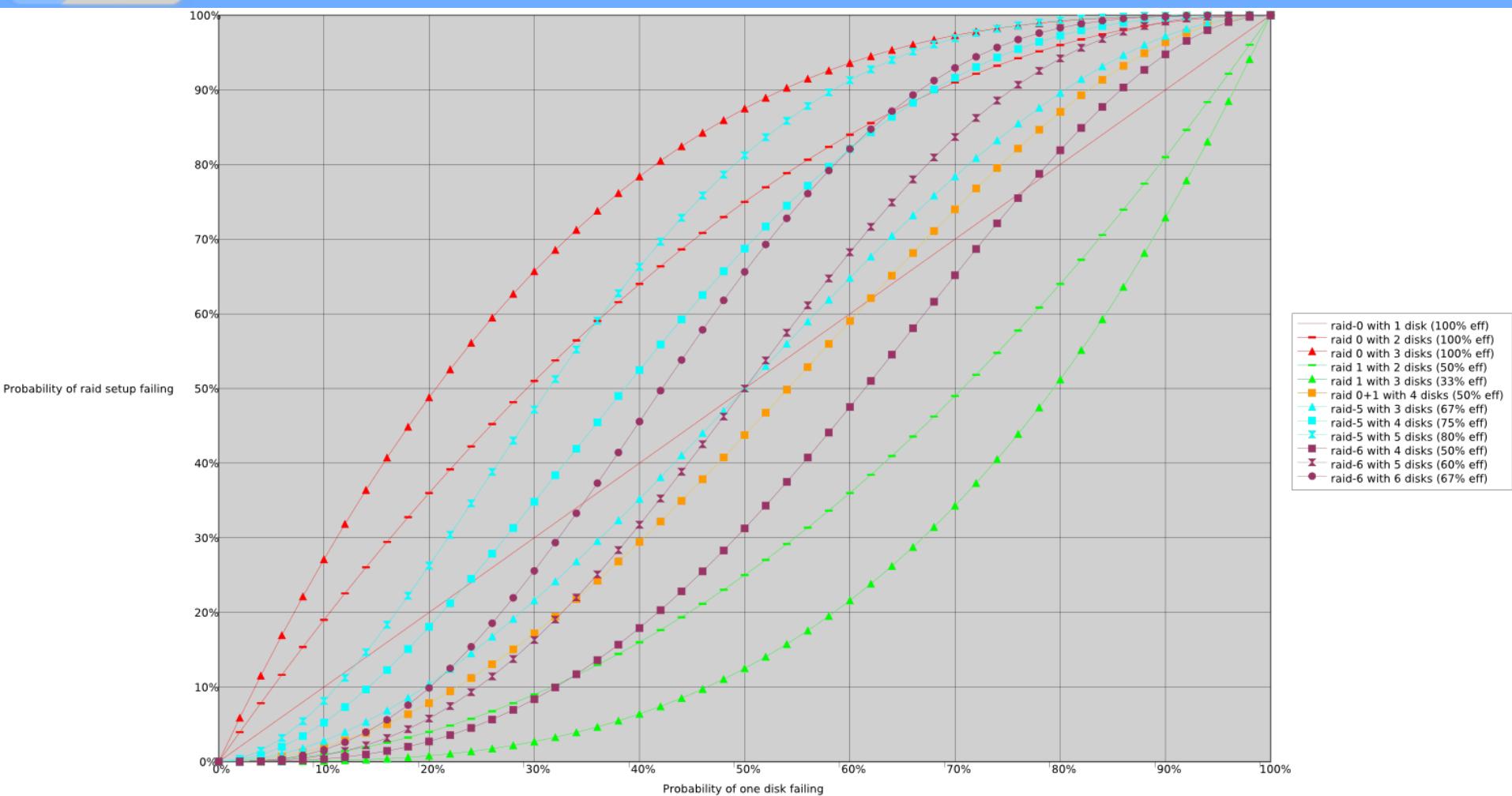
RAID 6: P+Q redundancy





# Selección de esquema RAID

<http://xkr47.outerspace.dyndns.org/tmp/raid-disk-failure-probabilities.png>





# Uso de RAID en Linux



- ❑ Loopback
- ❑ RAID con *mdadm*



# Sistema de ficheros *loopback*

- Un sistema de ficheros *loopback* es un fichero que **puede** actuar como un dispositivo de bloque y puede ser formateado con cualquier sistema de ficheros disponible antes de ser montado.
- Útil para diversas tareas:
  - Manipular imágenes de CD/DVD
  - Crear discos virtuales con los que **probar RAID**



# Sistema de ficheros loopback

1. Crearemos el fichero inicial vacío (>128 MB):

```
dd if=/dev/zero of=/tmp/vd-01 bs=1024 count=131072
```

2. Asociaremos el fichero al dispositivo de loopback:

```
losetup /dev/loop1 /tmp/vd-01
```

3. Usar el dispositivo loopback:

```
mkfs -t ext3 /dev/loop1  
mount /dev/loop1 /mnt  
mkdir /mnt/dir1  
umount /dev/loop1
```

4. Desasociar el dispositivo:

```
losetup -d /dev/loop1
```



# Sistema RAID (`mđadm`)

- MDAdm (*Multiple Device Administrator*) es la herramienta de administración de RAID software bajo Linux.
- Extiende la funcionalidad de RAIDtools:
  - Único programa
  - Puede funcionar sin fichero de configuración
  - Puede generar informes detallados
  - Puede funcionar como *daemon*, realizando seguimiento del RAID
- Ofrece apoyo para:
  - Linear
  - RAID 0
  - RAID 1
  - RAID 4 y RAID 5



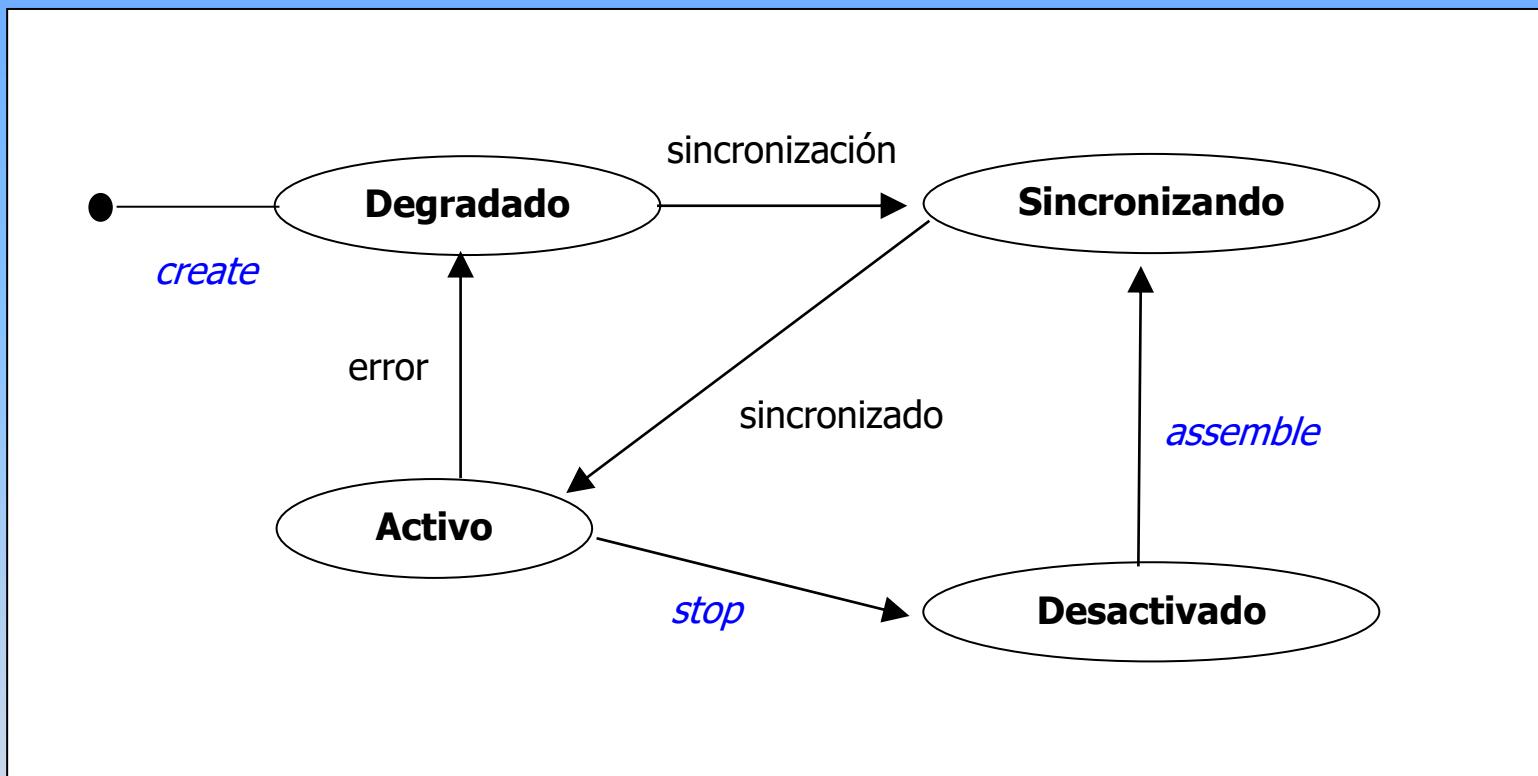
# Sistema RAID (młądm)

□ **Neil Brown** (<http://cgi.cse.unsw.edu.au/~neilb/>)





# Sistema RAID (mdadm)



*misc*: Permite realizar operaciones de forma individual



# Ejercicios prácticos

	Linear	R0	R1	R5
Crear Rx				
Listar los RAID				
Detalles del Rx				
Crear un sistema de ficheros y montarlo				
Espacio libre	+	+	$1/2$	$(n-1)/n$
Ancho banda	=	$*n$	$*n/2$	$*n$
Parar y arrancar				
Recuperación y discos en espera				



# Linear

- Crear un sistema *linear*:

```
# mdadm --create --verbose /dev/md5  
    --level=linear  
    --raid-devices=2 /dev/loop1 /dev/loop3
```



```
mdadm: chunk size defaults to 64K  
mdadm: array /dev/md5 started.
```





# Linear

## □ Consultar RAID en el sistema:

```
# cat /proc/mdstat
```



```
Personalities : [linear]
md5 : active linear loop3[1] loop1[0]
      262016 blocks 64k rounding
```





# Linear

- Consultar los detalles del */linear* creado:

```
# mdadm --detail /dev/md5
```



```
/dev/md5:  
      Version : 00.90.01  
Creation Time : Sat Mar 11 23:58:25 2006  
      Raid Level : linear  
      Array Size : 262016 (255.88 MiB 268.30 MB)  
      Raid Devices : 2  
      Total Devices : 2  
Preferred Minor : 5  
      Persistence : Superblock is persistent
```





# Linear

...

```
Update Time : Sat Mar 11 23:58:25 2006
          State : clean
Active Devices : 2
Working Devices : 2
Failed Devices : 0
Spare Devices : 0

Rounding : 64K
```



```
UUID : 6a08bb68:54e05845:53f1e21d:56ff9c18
Events : 0.1
```

Number	Major	Minor	RaidDevice	State	
0	7	1	0	active sync	/dev/loop1
1	7	3	1	active sync	/dev/loop3



# Linear

- Crear el sistema de ficheros sobre el *linear*:

```
# mkfs -t ext3 /dev/md5
```



```
mke2fs 1.38 (30-Jun-2005)
Etiqueta del sistema de ficheros=
Tipo de SO: Linux
Tamaño del bloque=1024 (bitácora=0)
Tamaño del fragmento=1024 (bitácora=0)
65536 nodos i, 262016 bloques
13100 bloques (5.00%) reservados para el super usuario
Primer bloque de datos=1
Maximum filesystem blocks=67371008
...
```





# Linear

...

32 bloques de grupos

8192 bloques por grupo, 8192 fragmentos por grupo

2048 nodos i por grupo

Respaldo del super bloque guardado en los bloques:

8193, 24577, 40961, 57345, 73729, 204801, 221185

Mientras se escriban las tablas de nodos i: terminado

Creando el fichero de transacciones (4096 bloques): hecho

Escribiendo superbloques y la información contable del sistema de  
ficheros: hecho

Este sistema de ficheros se revisará automáticamente cada 36 meses o  
180 días, lo que suceda primero. Utilice tune2fs -c o -i para  
cambiarlo.





# Linear

## □ Montar el sistema de ficheros

```
# mount /dev/md5 /mnt  
# mount
```



```
/dev/sda2 on / type ext3 (rw)  
/dev/proc on /proc type proc (rw)  
/dev/sys on /sys type sysfs (rw)  
/dev/devpts on /dev/pts type devpts (rw,gid=5,mode=620)  
/dev/shm on /dev/shm type tmpfs (rw)  
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)  
/proc on /var/named/chroot/proc type none (rw,bind)  
/dev/md5 on /mnt type ext3 (rw)
```





# Linear

## □ Consultar el espacio libre:

```
# df --si /mnt
```



S.ficheros	Tamaño	Usado	Disp	Uso%	Montado en
/dev/md5	260M	6,4M	241M	3%	/mnt



```
# df --si /mnt -h
```



S.ficheros	Tamaño	Usado	Disp	Uso%	Montado en
/dev/md5	248M	6,1M	229M	3%	/mnt





# Linear

- Consultar el ancho banda:

```
# hdparm -t /dev/md5
```



/dev/md5:

Timing xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx





# Linear

- Parar el *linear*:

```
# mdadm --stop /dev/md5
```





# Linear

- Arrancar el *linear*:

```
# mdadm --assemble /dev/md5 /dev/loop1 /dev/loop3
```



```
mdadm: /dev/md5 has been started with 2 drives.
```





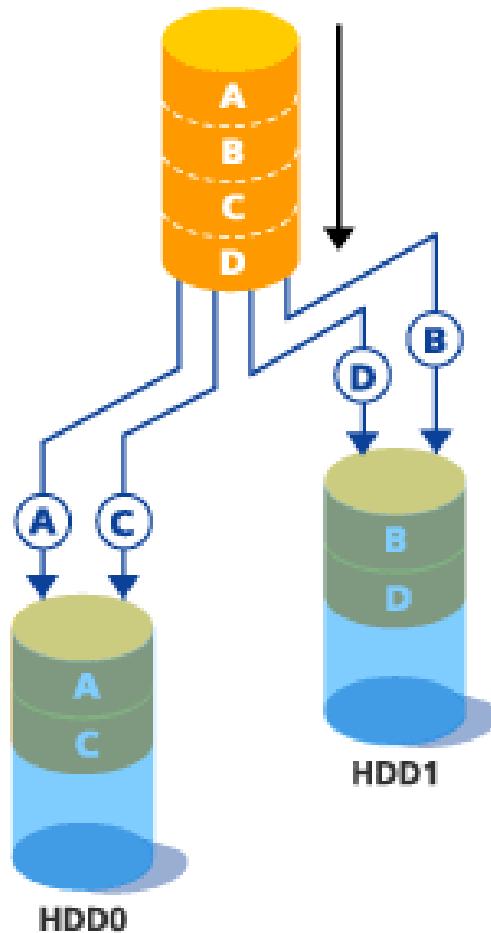
# Linear

- Crear un sistema lineal
- Consultar RAID en el sistema
- Consultar los detalles del *linear* creado
- Crear el sistema de ficheros sobre el *linear*
- Montar el sistema de ficheros
- Consultar el espacio libre
- Consultar el ancho banda
- Parar el *linear*
- Arrancar el *linear*



# RAID 0 (*stripping*)

CPUからABCDというデータの書き込み指示



複数ディスクへ分散して書き込む  
ストライピング処理を行う



# RAID 0

## □ Crear un sistema R0:

```
# mdadm --create --verbose /dev/md5  
    --level=0  
    --raid-devices=2 /dev/loop1 /dev/loop3
```



```
mdadm: chunk size defaults to 64K  
mdadm: /dev/loop1 appears to contain an ext2fs file system  
      size=262016K mtime=Thu Jan  1 01:00:00 1970  
mdadm: /dev/loop1 appears to be part of a raid array:  
      level=0 devices=2 ctime=Sun Apr  2 11:35:38 2006  
mdadm: /dev/loop3 appears to be part of a raid array:  
      level=0 devices=2 ctime=Sun Apr  2 11:35:38 2006  
Continue creating array? y  
mdadm: array /dev/md5 started.
```





# RAID 0

- Consultar RAID en el sistema:

```
# cat /proc/mdstat
```



```
Personalities : [linear] [raid0]
md5 : active raid0 loop3[1] loop1[0]
      262016 blocks 64k chunks
```





# RAID 0

- Consultar los detalles del RAID creado:

```
# mdadm --detail /dev/md5
```



```
/dev/md5:  
      Version : 00.90.01  
Creation Time : Sat Mar 11 23:58:25 2006  
      Raid Level : raid0  
      Array Size : 262016 (255.88 MiB 268.30 MB)  
      Raid Devices : 2  
      Total Devices : 2  
Preferred Minor : 5  
      Persistence : Superblock is persistent
```





# RAID 0

...

```
Update Time : Sat Mar 11 23:58:25 2006
          State : clean
Active Devices : 2
Working Devices : 2
Failed Devices : 0
Spare Devices : 0

Rounding : 64K
```



```
UUID : 6a08bb68:54e05845:53f1e21d:56ff9c18
Events : 0.11
```

Number	Major	Minor	RaidDevice	State	
0	7	1	0	active sync	/dev/loop1
1	7	3	1	active sync	/dev/loop3



# RAID 0

## □ Crear el sistema de ficheros sobre el R0:

```
# mkfs -t ext3 /dev/md5
```



```
mke2fs 1.38 (30-Jun-2005)
Etiqueta del sistema de ficheros=
Tipo de SO: Linux
Tamaño del bloque=1024 (bitácora=0)
Tamaño del fragmento=1024 (bitácora=0)
65536 nodos i, 262016 bloques
13100 bloques (5.00%) reservados para el super usuario
Primer bloque de datos=1
Maximum filesystem blocks=67371008
...
```





# RAIDO

...

32 bloques de grupos

8192 bloques por grupo, 8192 fragmentos por grupo

2048 nodos i por grupo

Respaldo del super bloque guardado en los bloques:

8193, 24577, 40961, 57345, 73729, 204801, 221185

Mientras se escriban las tablas de nodos i: terminado

Creando el fichero de transacciones (4096 bloques): hecho

Escribiendo superbloques y la información contable del sistema de  
ficheros: hecho

Este sistema de ficheros se revisará automáticamente cada 36 meses o  
180 días, lo que suceda primero. Utilice tune2fs -c o -i para  
cambiarlo.





# RAID 0

## □ Montar el sistema de ficheros

```
# mount /dev/md5 /mnt  
# mount
```



```
/dev/sda2 on / type ext3 (rw)  
/dev/proc on /proc type proc (rw)  
/dev/sys on /sys type sysfs (rw)  
/dev/devpts on /dev/pts type devpts (rw,gid=5,mode=620)  
/dev/shm on /dev/shm type tmpfs (rw)  
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)  
/proc on /var/named/chroot/proc type none (rw,bind)  
/dev/md5 on /mnt type ext3 (rw)
```





# RAID 0

- Consultar el espacio libre:

```
# df --si /mnt
```



S.ficheros	Tamaño	Usado	Disp	Uso%	Montado en
/dev/md5	260M	6,4M	241M	3%	/mnt





# RAID 0

- Consultar el ancho banda:

```
# hdparm -t /dev/md5
```



```
/dev/md5:
```

```
Timing xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```





# RAID 0

## □ Parar el R0:

```
# mdadm --stop /dev/md5
```





# RAID 0

## □ Arrancar el R0:

```
# mdadm --assemble /dev/md5 /dev/loop1 /dev/loop3
```



```
mdadm: /dev/md5 has been started with 2 drives.
```





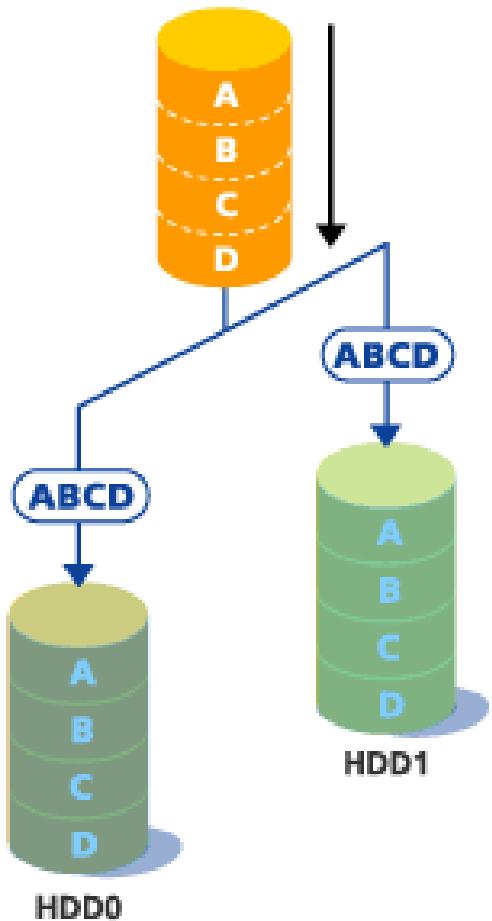
# RAID 0

- Crear un sistema lineal
- Consultar RAID en el sistema
- Consultar los detalles del R0 creado
- Crear el sistema de ficheros sobre el R0
- Montar el sistema de ficheros
- Consultar el espacio libre
- Consultar el ancho banda
- Parar el R0
- Arrancar el R0



# RAID 1 (*mirroring*)

CPUからABCDというデータの書き込み指示



データを2台のディスクに同時に書き込む



# RAID 1

## □ Crear un sistema R1:

```
# mdadm --create --verbose /dev/md5  
    --level=1  
    --raid-devices=2 /dev/loop1 /dev/loop3
```



```
mdadm: /dev/loop1 appears to contain an ext2fs file system  
      size=262016K mtime=Sun Apr  2 11:40:17 2006  
mdadm: /dev/loop1 appears to be part of a raid array:  
      level=0 devices=2 ctime=Sun Apr  2 11:36:06 2006  
mdadm: /dev/loop3 appears to be part of a raid array:  
      level=0 devices=2 ctime=Sun Apr  2 11:36:06 2006  
mdadm: size set to 131008K  
Continue creating array? y  
mdadm: array /dev/md5 started.
```





# RAID 1

## □ Consultar RAID en el sistema:

```
# cat /proc/mdstat
```



```
Personalities : [linear] [raid0] [raid1]
md5 : active raid1 loop3[1] loop1[0]
      131008 blocks [2/2] [UU]
      [=====>.....]  resync = 21.0% (28032/131008)
finish=1.6min speed=1038K/sec
```





# RAID 1

- Consultar RAID en el sistema (fin *sync*):

```
# cat /proc/mdstat
```



```
Personalities : [linear] [raid0] [raid1]
md5 : active raid1 loop3[1] loop1[0]
      131008 blocks [2/2] [UU]
```





# RAID 1

- Consultar los detalles del RAID creado:

```
# mdadm --detail /dev/md5
```



```
/dev/md5:  
      Version : 00.90.01  
Creation Time : Sun Apr  2 11:48:39 2006  
      Raid Level : raid1  
      Array Size : 131008 (127.94 MiB 134.15 MB)  
      Device Size : 131008 (127.94 MiB 134.15 MB)  
      Raid Devices : 2  
      Total Devices : 2  
Preferred Minor : 5  
      Persistence : Superblock is persistent
```





# RAID 1

...

```
Update Time : Sun Apr 2 11:50:50 2006
          State : clean
Active Devices : 2
Working Devices : 2
Failed Devices : 0
Spare Devices : 0
```



```
UUID : 7ad3a59c:dc909799:48fa13c7:d3606c1c
Events : 0.13
```

Number	Major	Minor	RaidDevice	State	
0	7	1	0	active sync	/dev/loop1
1	7	3	1	active sync	/dev/loop3



# RAID 1

## □ Crear el sistema de ficheros sobre el R1:

```
# mkfs -t ext3 /dev/md5
```



```
mke2fs 1.38 (30-Jun-2005)
Etiqueta del sistema de ficheros=
Tipo de SO: Linux
Tamaño del bloque=1024 (bitácora=0)
Tamaño del fragmento=1024 (bitácora=0)
32768 nodos i, 131008 bloques
6550 bloques (5.00%) reservados para el super usuario
Primer bloque de datos=1
Maximum filesystem blocks=67371008
...
```





# RAID 1

...

16 bloques de grupos

8192 bloques por grupo, 8192 fragmentos por grupo

2048 nodos i por grupo

Respaldo del super bloque guardado en los bloques:

8193, 24577, 40961, 57345, 73729

Mientras se escriban las tablas de nodos i: terminado

Creando el fichero de transacciones (4096 bloques): hecho

Escribiendo superbloques y la información contable del sistema de ficheros: hecho

Este sistema de ficheros se revisará automáticamente cada 22 meses o 180 días, lo que suceda primero. Utilice tune2fs -c o -i para cambiarlo.





# RAID 1

## □ Montar el sistema de ficheros

```
# mount /dev/md5 /mnt  
# mount
```



```
/dev/sda2 on / type ext3 (rw)  
/dev/proc on /proc type proc (rw)  
/dev/sys on /sys type sysfs (rw)  
/dev/devpts on /dev/pts type devpts (rw,gid=5,mode=620)  
/dev/shm on /dev/shm type tmpfs (rw)  
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)  
/proc on /var/named/chroot/proc type none (rw,bind)  
/dev/md5 on /mnt type ext3 (rw)
```





# RAID 1

- Consultar el espacio libre:

```
# df --si /mnt
```



S.ficheros	Tamaño	Usado	Disp	Uso%	Montado en
/dev/md5	130M	5,8M	118M	5%	/mnt





# RAID 1

- Consultar el ancho banda:

```
# hdparm -t /dev/md5
```



/dev/md5:

Timing xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx





# RAID 1

## □ Parar el R1:

```
# mdadm --stop /dev/md5
```





# RAID 1

## ❑ Arrancar el R1:

```
# mdadm --assemble /dev/md5 /dev/loop1 /dev/loop3
```



```
mdadm: /dev/md5 has been started with 2 drives.
```





# RAID 1

- Crear un sistema lineal
- Consultar RAID en el sistema
- Consultar los detalles del R1 creado
- Crear el sistema de ficheros sobre el R1
- Montar el sistema de ficheros
- Consultar el espacio libre
- Consultar el ancho banda
- Parar el R1
- Arrancar el R1



K

## □ Lista de módulos cargados:

```
# lsmod
```



raid0	12609	1
linear	9409	0
loop	20809	9
...		
ext3	133193	4
jbd	61785	1 ext3
raid1	21441	3
...		



# Marcar un dispositivo como que falla (sin martillazo :P)



```
# mdadm /dev/md5 --fail /dev/loop3
```



```
mdadm: set /dev/loop3 faulty in /dev/md5
```



```
# cat /proc/mdstat
```



```
md5 : active raid1 loop3[2] (F) loop1[0]  
      131008 blocks [2/1] [U_]
```





# quitar un dispositivo que falla

```
# mdadm --remove /dev/md5 /dev/loop3
```



```
mdadm: hot removed /dev/loop3
```



```
# cat /proc/mdstat
```



```
Personalities : [linear] [raid1]
md5 : active raid1 loop1[0]
      524224 blocks [2/1] [U_]
```





# añadir un dispositivo que NO falla

```
# mdadm --add /dev/md5 /dev/loop3
```



```
mdadm: hot added /dev/loop3
```



```
# cat /proc/mdstat
```



```
Personalities : [linear] [raid1]
md5 : active raid1 loop3[2] loop1[0]
      524224 blocks [2/1] [U_]
      [>.....]           recovery = 0.5% (3072/524224)
finish=5.6min speed=1536K/sec
```





# Seguimiento del RAID

- Es posible realizar un seguimiento mediante un demonio a dispositivos RAID, especificando el correo electrónico donde se enviarán las alertas.
- Por ejemplo:

```
mdadm --monitor --scan --daemonise  
--mail=root@localhost /dev/md0 /dev/md1 &
```

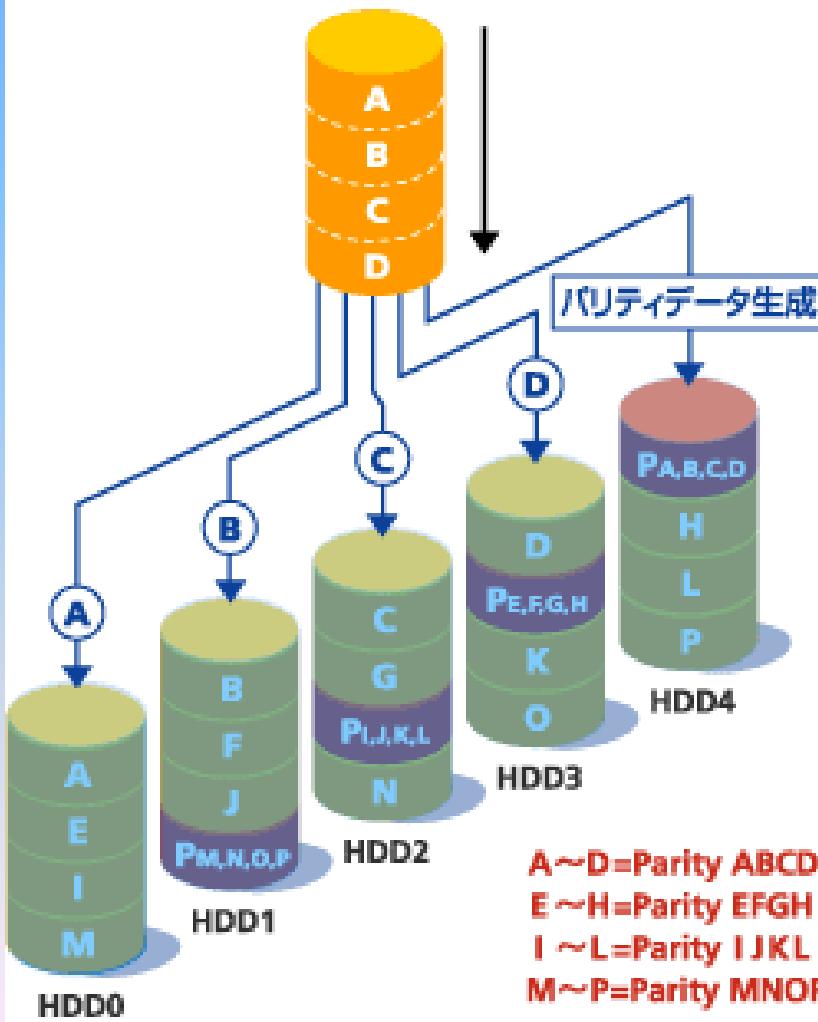


- Arranca mdadm como *daemon* que monitoriza dos RAID (*/dev/md0* y */dev/md1*) especificados y aquellos detectados (*--scan*).  
Las alertas se envían al correo *root@localhost.com*.



# RAID 5

CPUからABCDというデータの書き込み指示





# Crear un RAID 5

```
# mdadm --create --verbose /dev/md5  
    --level=5  
    --raid-devices=3 /dev/loop1 /dev/loop2 /dev/loop3
```



```
mdadm: layout defaults to left-symmetric  
mdadm: chunk size defaults to 64K  
mdadm: /dev/loop1 appears to contain an ext2fs file system  
      size=131008K mtime=Sun Apr  2 12:00:59 2006  
mdadm: /dev/loop1 appears to be part of a raid array:  
      level=1 devices=2 ctime=Sun Apr  2 11:48:39 2006  
mdadm: /dev/loop3 appears to contain an ext2fs file system  
      size=131008K mtime=Sun Apr  2 12:00:59 2006  
mdadm: /dev/loop3 appears to be part of a raid array:  
      level=1 devices=2 ctime=Sun Apr  2 11:48:39 2006  
mdadm: size set to 131008K  
Continue creating array? y  
mdadm: array /dev/md5 started.
```





# Consultar los RAID del sistema

```
# cat /proc/mdstat
```



```
md5 : active raid5 loop3[3] loop2[1] loop1[0]
      262016 blocks level 5, 64k chunk, algorithm 2 [3/2]
      [UU_]
      [=----->...] recovery = 86.7%
(113664/131008) finish=0.2min speed=1036K/sec

md1 : active raid1 sdd1[1] sdb1[0]
      244742144 blocks [2/2] [UU]

md0 : active raid1 sdc1[1] sda1[0]
      202900800 blocks [2/2] [UU]
```



# Consultar los RAID del sistema (fin sync)

```
# cat /proc/mdstat
```



```
md5 : active raid5 loop3[2] loop2[1] loop1[0]
      262016 blocks level 5, 64k chunk, algorithm 2 [3/3]
      [UUU]

md1 : active raid1 sdd1[1] sdb1[0]
      244742144 blocks [2/2] [UU]

md0 : active raid1 sdc1[1] sda1[0]
      202900800 blocks [2/2] [UU]
```



# Consultar los detalles del RAID 5 creado

```
# mdadm --detail /dev/md5
```



/dev/md5:

```
    Version : 00.90.01
Creation Time : Sun Apr  2 12:42:26 2006
    Raid Level : raid5
    Array Size : 262016 (255.88 MiB 268.30 MB)
    Device Size : 131008 (127.94 MiB 134.15 MB)
    Raid Devices : 3
    Total Devices : 3
Preferred Minor : 5
    Persistence : Superblock is persistent
```



...



# Consultar los detalles del RAID 5 creado (2)

...

Update Time : Sun Apr 2 12:44:37 2006  
State : clean

Active Devices : 3

Working Devices : 3

Failed Devices : 0

Spare Devices : 0

Layout : **left-symmetric**

Chunk Size : 64K



UUID : 57114263:aaebfdcd:db16773f:30eddf08

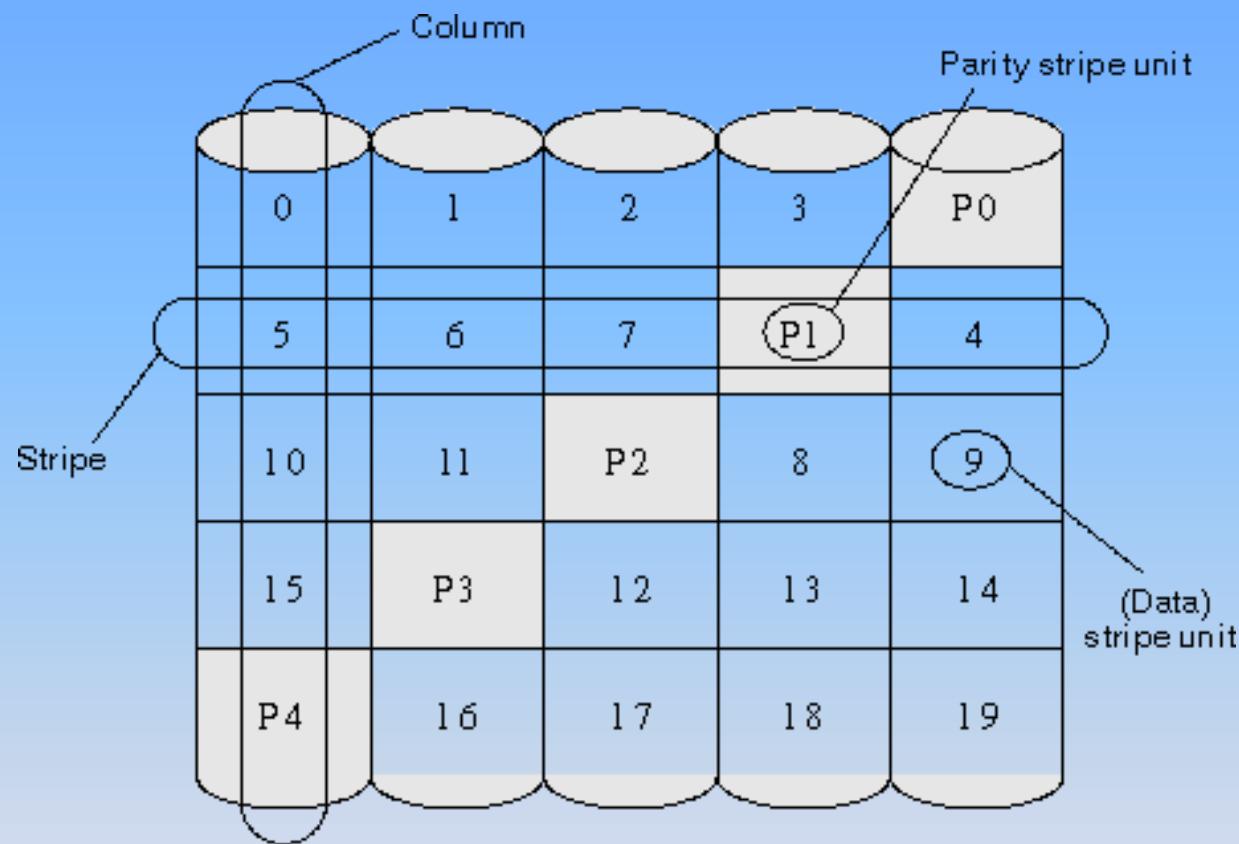
Events : 0.21

Number	Major	Minor	RaidDevice	State	
0	7	1	0	active sync	/dev/loop1
1	7	2	1	active sync	/dev/loop2
2	7	3	2	active sync	/dev/loop3





# Left-symmetric



<http://docs.sun.com/source/805-5706-10/805-5706-10.book>

# Crear un sistema de ficheros sobre /dev/md5

```
# mkfs -t ext3 /dev/md5
```



```
mke2fs 1.38 (30-Jun-2005)
Etiqueta del sistema de ficheros=
Tipo de SO: Linux
Tamaño del bloque=1024 (bitos/cora=0)
Tamaño del fragmento=1024 (bitos/cora=0)
65536 nodos i, 262016 bloques
13100 bloques (5.00%) reservados para el super usuario
Primer bloque de datos=1
Maximum filesystem blocks=67371008
32 bloque de grupos
8192 bloques por grupo, 8192 fragmentos por grupo
2048 nodos i por grupo
Respaldo del super bloque guardado en los bloques:
    8193, 24577, 40961, 57345, 73729, 204801, 221185
...
```





# Montar el sistema de ficheros

```
# mount /dev/md5 /mnt  
# mount
```



```
/dev/sda2 on / type ext3 (rw)  
/dev/proc on /proc type proc (rw)  
/dev/sys on /sys type sysfs (rw)  
/dev/devpts on /dev/pts type devpts (rw,gid=5,mode=620)  
/dev/shm on /dev/shm type tmpfs (rw)  
none on /proc/sys/fs/binfmt_misc type binfmt_misc (rw)  
/proc on /var/named/chroot/proc type none (rw,bind)  
/dev/md5 on /mnt type ext3 (rw)
```





# Consultar el espacio libre

```
# df --si /mnt
```



S.ficheros	Tamaño	Usado	Disp	Uso%	Montado en
/dev/md5	260M	6,4M	241M	3%	/mnt





# Consultar el ancho de banda

```
# hdparm -t /dev/md5
```



/dev/md5:

Timing xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx





# Detener el RAID 5

```
# mdadm --stop /dev/md5
```





# Arrancar de nuevo el RAID 5



```
# mdadm --assemble /dev/md5 /dev/loop1 /dev/loop2 /dev/loop3
```



```
mdadm: /dev/md5 has been started with 3 drives.
```



# RAID 5

- Crear un sistema lineal
- Consultar RAID en el sistema
- Consultar los detalles del R5 creado
- Crear el sistema de ficheros sobre el R5
- Montar el sistema de ficheros
- Consultar el espacio libre
- Consultar el ancho banda
- Parar el R5
- Arrancar el R5



# Sistema RAID (spare)

## □ Discos en espera (Hot Spare):

- Es posible definir un conjunto de **discos en espera (Spare disk)**
- Cuando falla un disco, se selecciona uno del conjunto para que lo sustituya en caliente (Hot), es decir sin parar el sistema





# Discos en espera (Hot Spare)

- Crea un RAID 5 (-l5) con 3 discos (-n3) y 2 discos en espera (-x2).

```
# mdadm -Cv /dev/md5 -l5 -n3 -x2  
-c128 /dev/loop{1,2,3,4,5}
```



```
mdadm: array /dev/md5 started.
```





# ¡Un sistema RAID necesita también de *backup*!

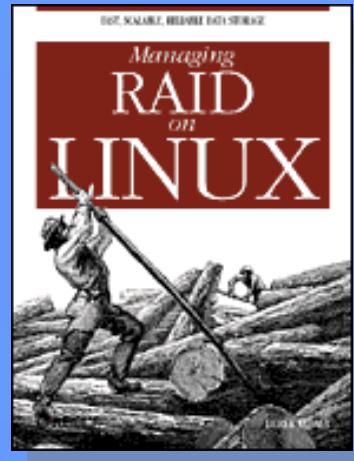
□ Independiente del modelo RAID elegido, es **IMPORTANTE realizar** una **copia de seguridad** periódica:

- Un sistema RAID no protege de fallos múltiples de disco.
- Un sistema de copia realizado con regularidad permite la recuperación de los datos perdidos por problemas no relacionados con fallos en discos, ellos incluye:
  - ✓ Errores humanos (`cat /dev/rdm0 > /dev/loop1`)
  - ✓ Errores hardware
  - ✓ Errores software



# Para más información...

- Libro ***Managing RAID on Linux***  
<http://www.oreilly.com/catalog/mraidlinux/>



- The Software-RAID HOWTO  
<http://www.tldp.org/HOWTO/Software-RAID-HOWTO.html#toc6>



# Sistemas RAID

**Grupo de Arquitectura y Sistemas Operativos (ARCOS@UC3M)**  
**Para el Grupo de Usuarios de Linux de la UC3M**

