

## 1.2. Sistemes d'emmagatzematge redundant

Els sistemes d'emmagatzematge redundant s'implementen mitjançant el que s'anomena *RAID* (*redundant array of inexpensive disk*), en català:

‘conjunt de discos barats redundants’. Aquesta és una tecnologia que permet assolir alts graus de fiabilitat en l'emmagatzematge d'informació a partir d'un conjunt de discos durs de baix cost que ens podem trobar en qualsevol ordinador personal (PC).

Recentment, aquest terme s'ha revisat i s'ha canviat la paraula *inexpensive* per *independent*.

L'objectiu del **RAID** consisteix a dividir i replicar la informació entre diferents discos durs i, a part d'incrementar la fiabilitat de la transferència, també en pot augmentar la velocitat.

Quan múltiples discos físics formen part d'un RAID, el sistema operatiu els veu com un de sol.

Els sistemes RAID impliquen una quantitat de càlculs important quan es fan operacions de lectura/escriptura. En els sistemes més cars, hi ha una targeta controladora específica que s'encarrega de portar a terme aquests càlculs. Aquest cas es coneix amb el nom de **RAID de maquinari**.

En alguns casos, el mateix sistema operatiu o els controladors més simples necessiten el microprocessador central per ajudar a fer tots aquests càlculs, la qual cosa fa baixar el rendiment del sistema. Això es coneix com a **RAID de programari**.

Els controladors RAID més senzills sovint només ens ofereixen els nivells de RAID 0 i 1, que requereixen menys processament.

Els sistemes RAID, amb redundància de dades, poden continuar funcionant encara que un (o en algun cas més d'un) dels discos falli. Quan això succeeix, es reemplaça el disc defectuós per un de nou, el RAID es reconstrueix i el sistema continua funcionant normalment.

Alguns sistemes s'han d'apagar per canviar un disc defectuós i d'altres, suporten l'intercanvi en calent (*hot swapping*), que permet canviar un disc defectuós sense haver d'apagar el sistema. Aquest darrer sistema d'intercanvi en calent (*hot swapping*) es fa servir sovint en sistemes d'alta disponibilitat (*high availability*), que necessiten estar funcionant ininterrompudament.

Els sistemes RAID si bé ofereixen un nivell de seguretat més elevat que els que no ho són, no eliminen la necessitat de crear còpies de seguretat del sistema, perquè es poden perdre dades sense que el disc en el qual estan desades es faci malbé físicament. Per exemple, les dades podrien ser sobreescrites per un mal funcionament del sistema operatiu, o bé per un usuari amb intencions dolentes.

#### Grau de suport de RAID

Gairebé totes les plaques base actuals suporten RAID 0 i 1. Per configurar-los cal consultar el manual de la placa base i entrar a la BIOS del sistema abans d'instal·lar-hi el sistema operatiu.

### 1.2.1. Redundància

La redundància en els sistemes que utilitzen RAID s'aconsegueix o bé escrivint la mateixa informació en diversos discos (conegut com a *mirall* o *mirror*), o bé escrivint dades extra, com la paritat de les dades en algun dels discos que formen el RAID.

D'aquesta manera, s'aconsegueix que, malgrat que un dels discos del RAID falli, això no impliqui una pèrdua de dades.

Podem combinar els discos de diferents maneres, segons les nostres necessitats de velocitat, de capacitat i de protecció contra la pèrdua de dades. A aquestes combinacions, hi fem referència amb els diferents nivells de RAID.

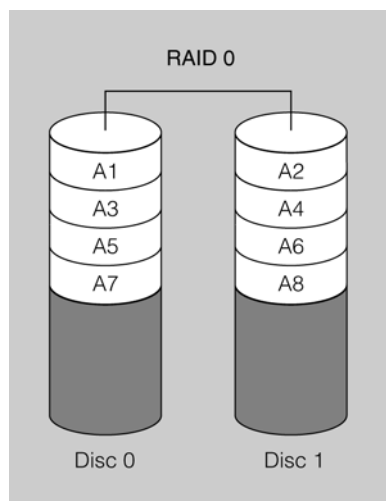
#### RAID 0

El RAID 0, conegut com a *stripping*, distribueix la informació entre diferents discos durs de manera que es guanya velocitat, però no es perd capacitat. Si un dels discos del RAID falla, aleshores es perdran totes les dades.

Aquest nivell de RAID només és aconsellable per a usuaris domèstics que no tenen dades crítiques.

Podem veure en la figura 3 que la informació es divideix en fragments, cadascun dels quals s'assigna a un dels discos. El nombre de fragments depèn del nombre de discos que formen el RAID. Aquesta manera de gravar la informació permet operacions de lectura/escriptura dels fragments de dades d'una manera simultània, de manera que s'aconsegueix més velocitat de transferència.

Figura 3. Esquema d'un sistema en RAID 0



L'inconvenient d'aquest sistema és la tolerància nul·la als errors en qualsevol dels discos.

Necessitem un mínim de dos discos durs per a implementar aquest tipus de RAID.

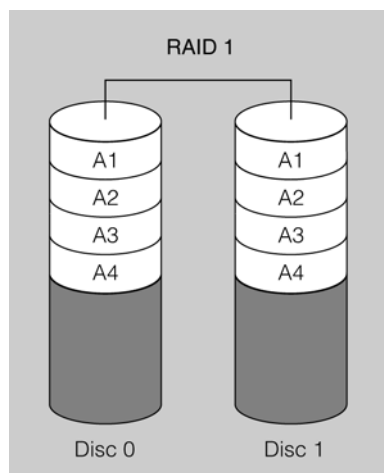
## RAID 1

El RAID 1, conegut com a *mirall*, o *mirror*, duplica la informació en cadascun dels discos que formen el RAID i comporta redundància total. Dos o més discos contenen la mateixa informació al mateix temps i no hi ha pèrdua d'informació encara que fallin tots els discos menys un. La capacitat és igual que la mida del disc més petit. Podem veure això en la figura 4.

Aquest nivell de RAID ens dóna una bona protecció de les dades a un cost baix, tot i que perdem capacitat. La velocitat no es veu afectada i seria semblant a la que tenim amb un sol disc dur.

Aquest sistema seria adequat per a petites empreses o usuaris particulars, que volen aconseguir seguretat en l'emmagatzematge amb un cost petit.

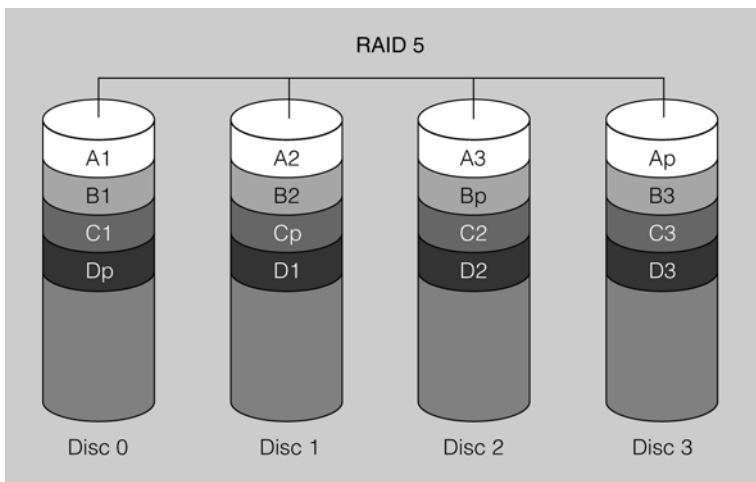
Figura 4. Esquema d'un sistema en RAID 1



## RAID 5

El RAID 5, conegut com a *stripping amb paritat*, combina tres o més discos de manera que es protegeix la informació davant la pèrdua d'un qualsevol dels discos (figura 5). La capacitat d'emmagatzematge és sempre l'equivalent a la suma de tots els discos menys un.

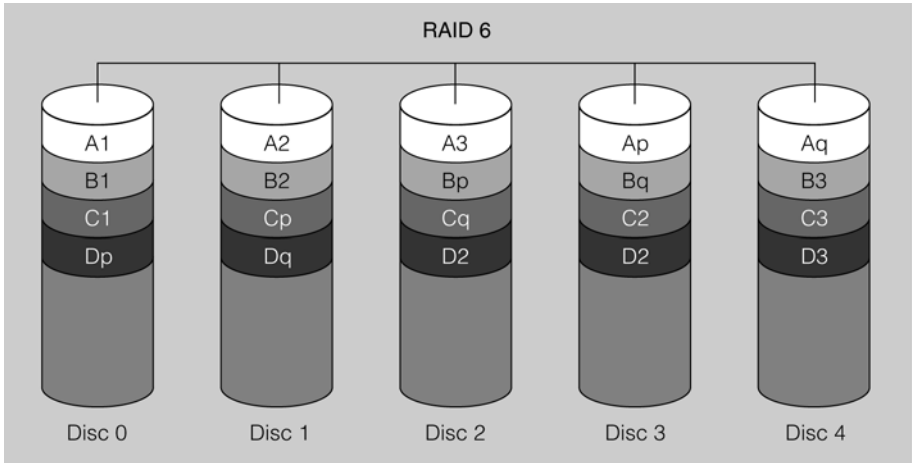
Figura 5. Esquema d'un sistema en RAID 5



# RAID 6

El RAID 6, conegut com a *stripping amb paritat doble*, és similar al RAID 5, però permet recuperar la informació encara que fallin dos dels discos (figura 6).

Figura 6. Sistema que fa servir un RAID 6

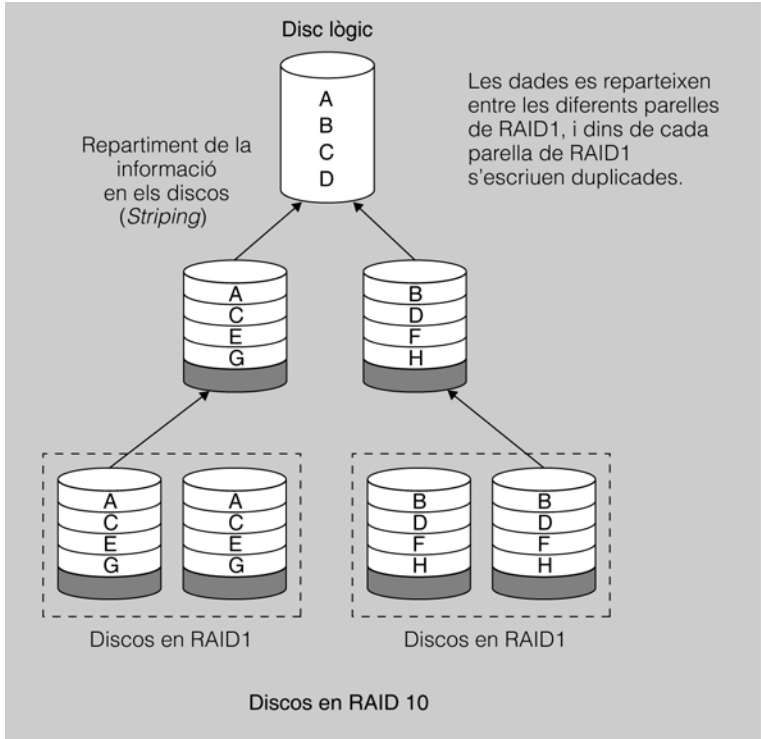


# RAID 1 + 0


El RAID 1 + 0 fa servir l'*stripping* i el mirall al mateix temps. És a dir, les dades es divideixen en dos discos per augmentar la velocitat, però al mateix temps estan duplicades en un altre disc, com a mínim, cosa que també millora la seguretat (figura 7).

Aquest sistema és el que permet assolir més rendiment, al mateix temps que tenim un grau de seguretat important.

Figura 7. Esquema d'un sistema que fa servir RAID 10 (o 1 + 0)



1.2.2. RAID basat en programari

Les implementacions de RAID basades en programari ens les proporcionen la majoria dels sistemes operatius actuals. S'estableix una capa addicional d'abstracció per sobre del programari que fa de controlador del disc, entre els discos lògics (RAID) i els discos físics. Els més populars són RAID 0, 1, 1 + 0 i 5. 

En la taula 1 podem veure el suport de diferents sistemes operatius per a diferents nivells de RAID.

Taula 1. Suport de diferents sistemes operatius a RAID de programari

Sistema operatiu	Suport de RAID de programari
Mac OS X Server	0, 1, 5 i 1 + 0
FreeBSD	0, 1, 3 i 5
GNU/Linux	0, 1, 4, 5 i 6
Microsoft	0, 1 i 5
NetBSD	0, 1, 4 i 5
OpenBSD	0 i 1
OpenSolaris / Solaris 10	0, 1 i 5

Els RAID de programari tenen avantatges i inconvenients respecte dels RAID basats en maquinari (targetes controladores específiques).

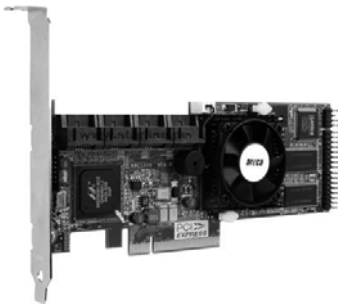
El principal avantatge que tenen és el cost d'implementació molt baix, ja que en l'actualitat gairebé totes les plaques base i sistemes operatius els suporten.

De totes maneres, el programari de RAID s'ha d'executar en un servidor connectat als dispositius d'emmagatzematge i el microprocessador del sistema ha de dedicar part del seu temps a executar el programari de RAID. Aquest temps és negligible per als sistemes en RAID 0 o 1, però pot ser important en sistemes en RAID amb paritat, o quan la quantitat de discos en el RAID és important. Això fa que en alguns casos el rendiment global del sistema disminueixi.

Altres possibles problemes són la congestió del bus de dades per transportar tota la informació de paritat i la dificultat per configurar el procés d'arrencada en cas que un dels discos falli. En aquests casos, cal la intervenció humana per s arreglar aquests problemes i fer que la màquina arrenqui un altre cop.

1.2.3. RAID basat en maquinari

Els RAID basats en maquinari no tenen els inconvenients dels RAID de programari, però són més cars, ja que requereixen targetes específiques,



Targeta controladora RAID, que conté quatre ports Serial ATA, un microprocessador Intel 80332 i 256 MB de memòria cau. Fa servir una interfície PCI-Express x4.

que inclouen un microprocessador dedicat al processament d'operacions d'entrada i de sortida. Sovint, aquestes targetes també inclouen una certa quantitat de memòria cau i permeten la connexió d'una quantitat important de discos durs, tot sovint superior a la que podríem connectar directament a una placa base.

#### 1.2.4. Possibles problemes i mètriques de fiabilitat d'un sistema RAID

Les principals causes d'errors en sistemes RAID, de les quals el fabricant ens pot informar per mesurar la fiabilitat del seu producte, són les següents:

**1) Errades correlacionals:** els sistemes de correcció d'errors en un sistema RAID, assumeixen que tots els discos són independents. Això permet fer un càlcul aproximat de la freqüència de les errades, i fer una tasca preventiva. Tot i que aquests errors són poc probables, aquesta tasca de prevenció és necessària per minimitzar la possibilitat de pèrdua d'informació. A la pràctica, els discos que formen el RAID tenen tots aproximadament la mateixa antiguitat i, com que la possibilitat d'errada mecànica en un disc augmenta segons l'antiguitat, això fa que la possibilitat d'un error d'un o més discos amb el temps augmenti, cosa que fa que la pèrdua d'informació sigui probable.

Això també depèn del tipus de disc; així, els dedicats a usos empresarials, del tipus SAS o canal de fibra (*fibre channel*), estan dissenyats per suportar un ús molt intensiu i optimitzat per ser utilitzats en RAID. En canvi, els discos més comuns que fem en els ordinadors de sobretaula són menys robustos, i fins i tot la seva utilització en RAID en podria escurçar la vida útil.

**2) Atomicitat:** quan s'escriu una informació en un sistema RAID, sovint s'escriu la mateixa informació en diferents discos. Però de vegades pot passar que quan múltiples discos estan sota la supervisió d'un sol processador d'entrada/sortida, s'escriguin diferents versions de la mateixa informació en diferents discos.

**3) Fiabilitat de la memòria cau d'escriptura (*write cache reliability*):** de vegades, els discos informen que s'han escrit les dades de seguida que són a la memòria cau, malgrat que encara no s'hagin escrit físicament sobre el disc. Si en aquest moment hi ha una pèrdua de corrent elèctric, o bé un error crític del sistema operatiu que fa que aquest deixi de respondre, en alguns sistemes de fitxers com ara FAT32 això pot provocar la pèrdua de les dades que eren a la memòria cau però encara no havien estat escrites en el disc. Per evitar-ho, de vegades la memòria cau del disc disposa d'una petita bateria que manté l'alimentació i resol el problema en la majoria de casos.

**4) Compatibilitat:** el format en què les dades es desen en un determinat RAID fa que de vegades no sigui possible llegir aquesta informació si connectem els discos a un controlador diferent. En el cas dels RAID del programari, això és menys probable que passi, ja que la implementació del RAID és independent del maquinari sobre el qual s'executa. A més a més, els discos que formen part d'un RAID 1 (mirall) poden ser llegits com un disc normal en una controladora normal, de manera que no és necessari un sistema RAID per recuperar-ne la informació.

**5) Altres problemes i virus informàtics:** tot i que el sistema RAID dona protecció davant de l'errada d'un disc, les dades encara poden ser destruïdes per un mal ús de l'operador de la màquina, el programari, el maquinari o la destrucció per culpa d'un virus informàtic. Els sistemes RAID ben dissenyats inclouen utilitats per fer còpies de seguretat, tot i que no permeten treballar-hi directament, i només permeten restaurar la informació.