

## **RAID**

**R**edundant **A**rray of **I**nexpensive **D**iscs

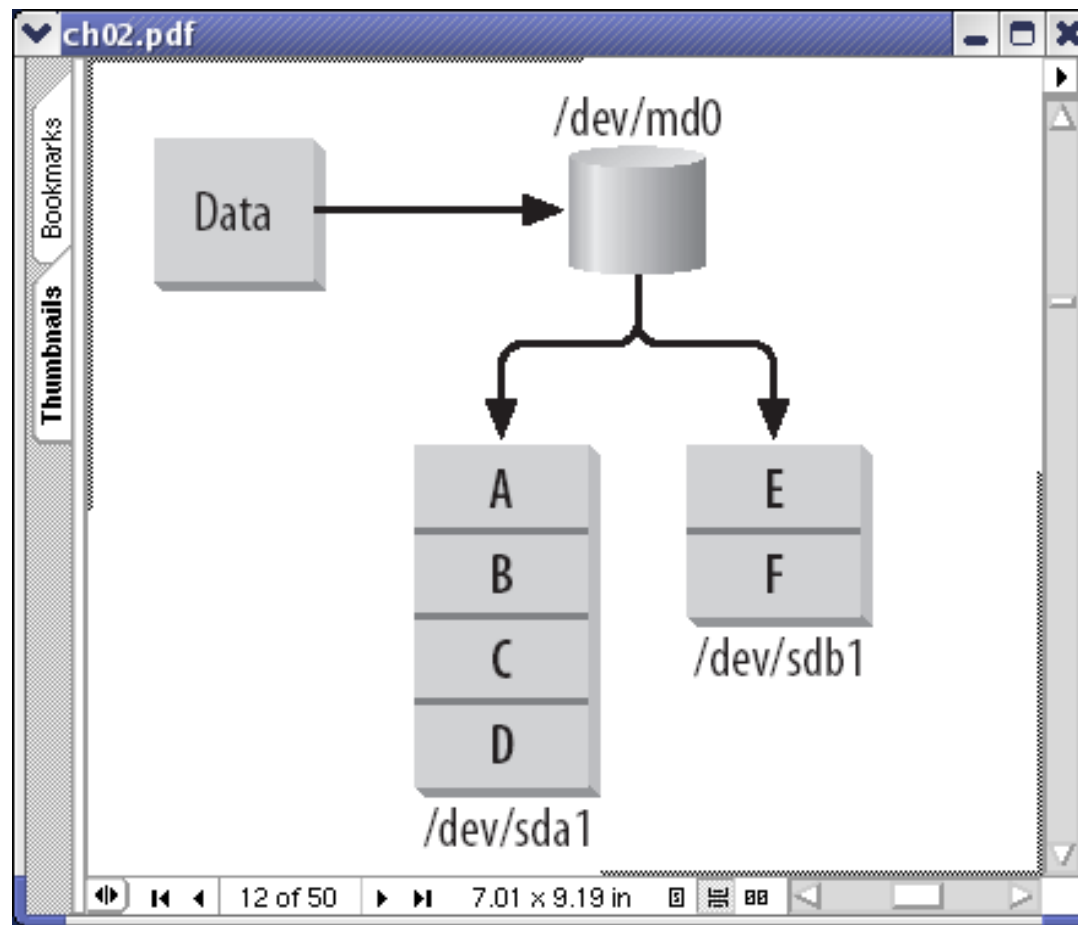
## **Objectiu**

**Utilitzar varis discs com un sol dispositiu**

**Incrementar el bandwidth del SF**

## **Problema Principal**

**Calen més discs. Es pot produir més fallades de disc.**



# RAIDO

## Dades: blocs

## Distribució de dades

Round-robin a tots els discs

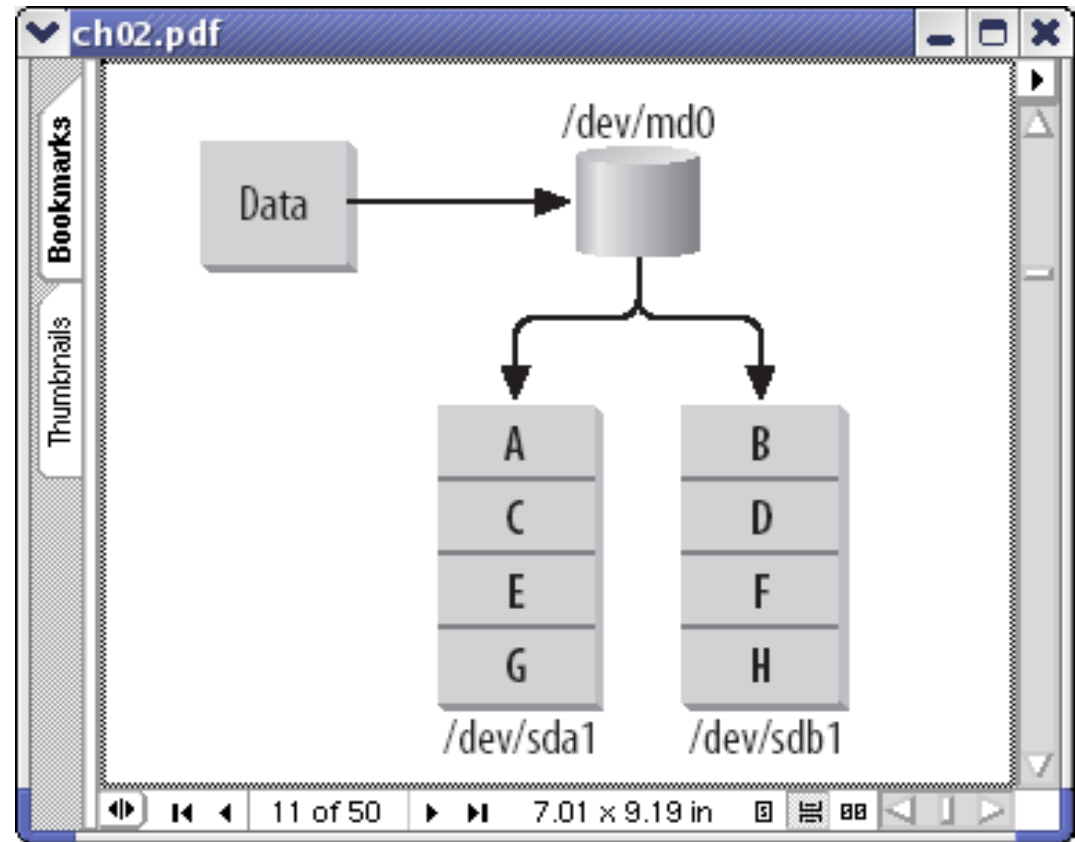
## Avantatges

Alt bandwidth

Gran capacitat

## Desavantatge

No és tolerant a cap fallada



# RAID1

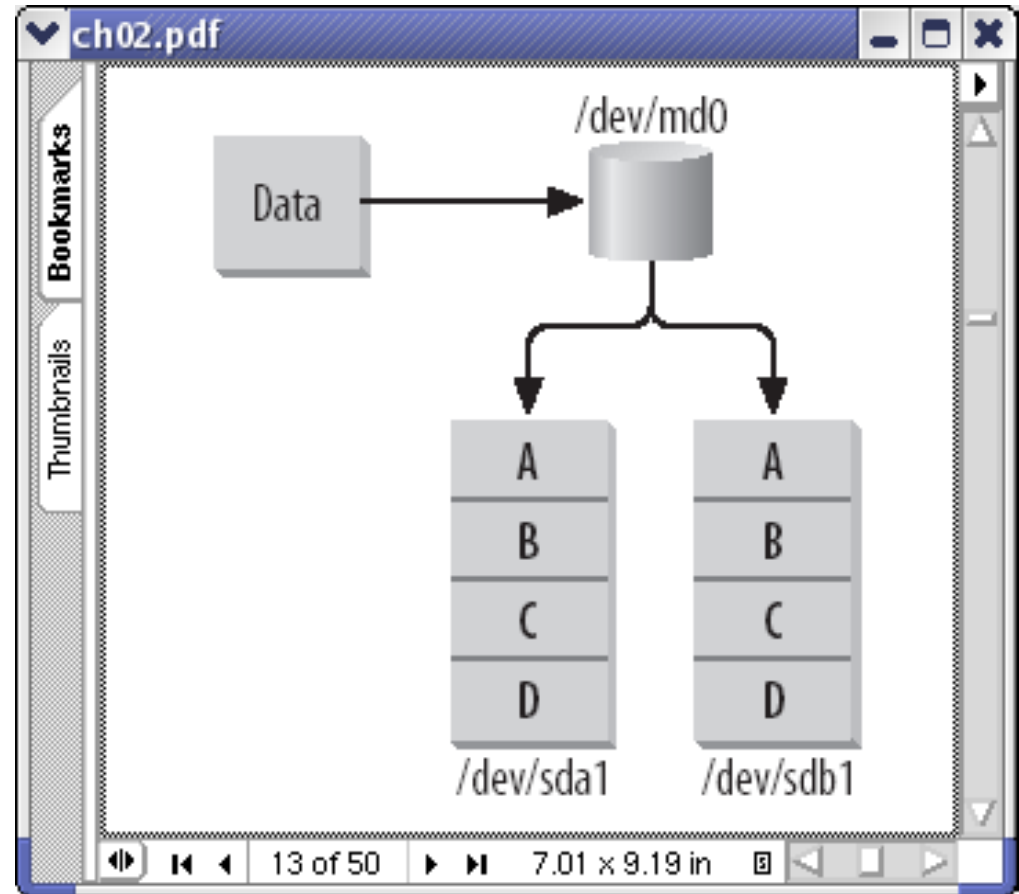
## Dades: blocs

## Distribució de dades

Duplicació de blocs en 2 particions

## Avantatges

- \* Tolerant a fallades. Si un disc falla, disposem de l'altre.
- \* Lectures gairebé el doble de ràpides.
- \* Escripcions igual que 1 sol disc



# RAID3 i RAID4

**Dades:** RAID3:bytes;

RAID4:blocs

## Distribució de dades

RAID0 + disc de paritat

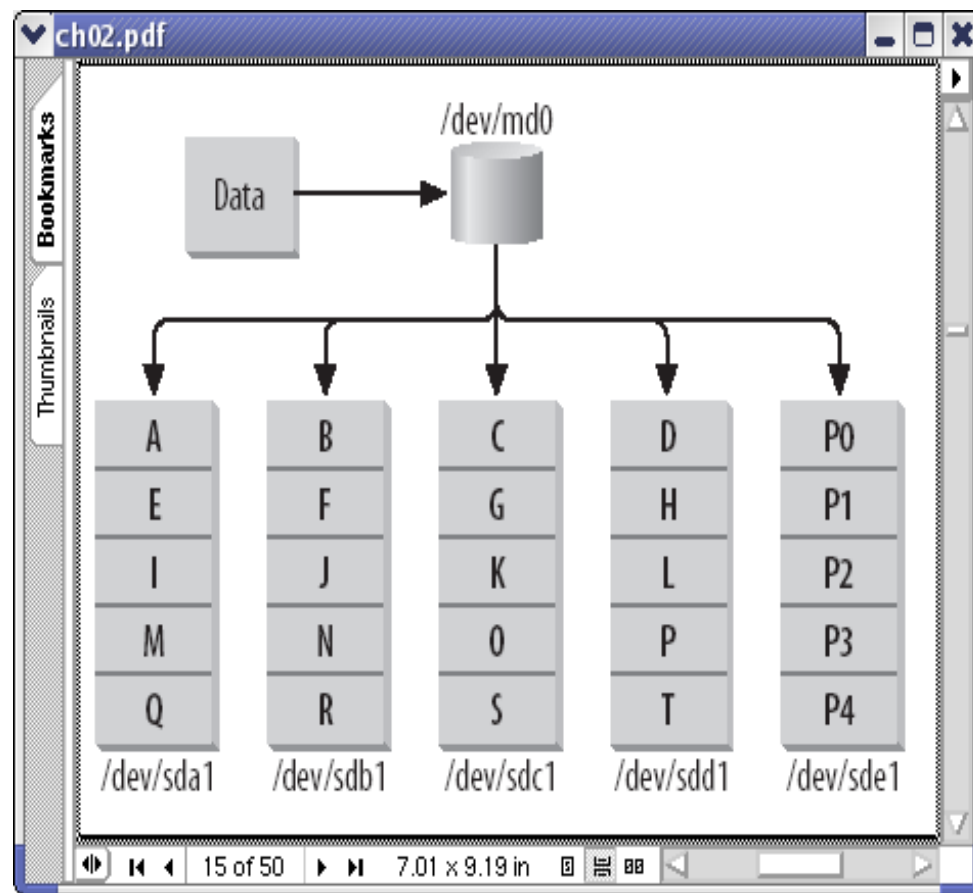
Paritat computada com **XOR**

**10100111 XOR 01110011**

**= 11010100**

**Avantatge Si falla un disc de dades es pot recuperar**

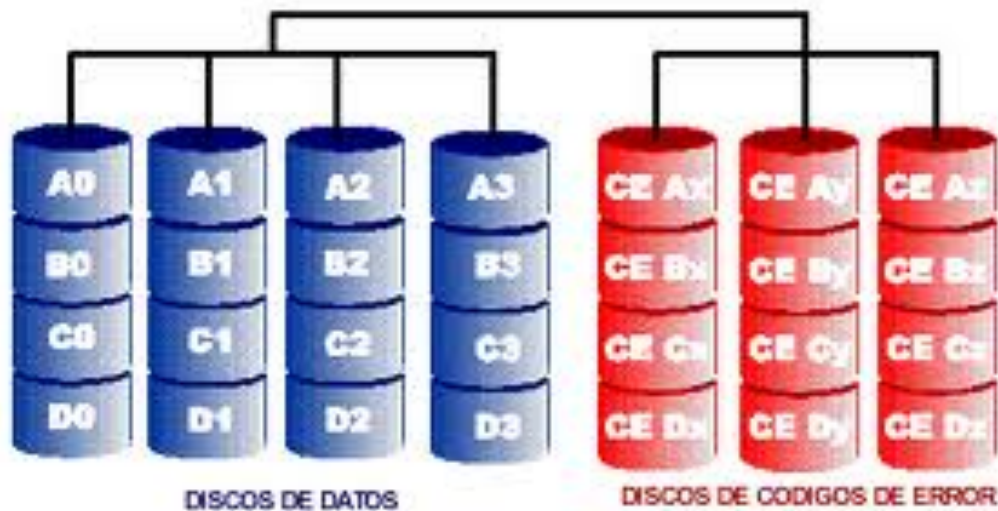
**Desavantatge el disc de paritat és un coll d'ampolla**



# RAID2

**Dades: bits**

**Informació adicional:** Codi Hamming: codi detector i corrector d'errors



**Exemple** Hamming (7,4)

7 bits de dades +

4 bits codi corrector

**Detecta i corretgeix 1 bit**

# HAMMING – algorisme (exemple)

Bits  $2^k$ : bits de paritat. Resta de bits: dades. Associant bits de paritat als bits de dades:  
• el bit de paritat (del nombre amb paritat) de la posició de les potències de 2 “ $k$ ” ( $k=2^n$ )  
comproba els bits en les posicions que tenguin al bit  $k$  en la seva representació binària.

Exemple: 1 : 3 (d1), 5 (d2), 7 (d4), 9 (d5) i 11 (d7)

2: 3, 6, 7, 10 i 11

Exemple Hamming de 7 bits (7,4): "0110101": (sp/ap: sense/amb paritat)

	1	2	3	4	5	6	7	8	9	10	11
	p <sub>1</sub>	p <sub>2</sub>	d <sub>1</sub>	p <sub>3</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	p <sub>4</sub>	d <sub>5</sub>	d <sub>6</sub>	d <sub>7</sub>
sp			0		1	1	0		1	0	1
p <sub>1</sub>	1		0		1		0		1		1
p <sub>2</sub>		0	0			1	0			0	1
p <sub>3</sub>				0	1	1	0				
p <sub>4</sub>								0	1	0	1
ap	1	0	0	0	1	1	0	0	1	0	1

**P1** = D1 xor D2 xor D4  
xor D5 xor D7

**P2** = D1 xor D3 xor D4  
xor D6 xor D7

**P3** = D2 xor D3 xor D4

**P4** = D5 xor D6 xor D7

# ***HAMMING comprovació sense errors***

	1	2	3	4	5	6	7	8	9	10	11	
	p <sub>1</sub>	p <sub>2</sub>	d <sub>1</sub>	p <sub>3</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	p <sub>4</sub>	d <sub>5</sub>	d <sub>6</sub>	d <sub>7</sub>	Bit comprovació
Paraula rebuda	1	0	0	0	1	1	0	0	1	0	1	
p <sub>1</sub>	<b>1</b>		0		1		0		1		1	0
p <sub>2</sub>		<b>0</b>	0			1	0			0	1	0
p <sub>3</sub>				<b>0</b>	1	1	0					0
p <sub>4</sub>								<b>0</b>	1	0	1	0

Bit de comprovació: 0 no error, 1 error



# HAMMING comprovació amb errors

	1	2	3	4	5	6	7	8	9	10	11	
	p <sub>1</sub>	p <sub>2</sub>	d <sub>1</sub>	p <sub>3</sub>	d <sub>2</sub>	d <sub>3</sub>	d <sub>4</sub>	p <sub>4</sub>	d <sub>5</sub>	d <sub>6</sub>	d <sub>7</sub>	Bit comprovació
Paraula rebuda	1	0	0	0	1	1	0	0	1	0	0	
p <sub>1</sub>	1		0		1		0		1		0	1
p <sub>2</sub>		0	0			1	0			0	0	1
p <sub>3</sub>				0	1	1	0					0
p <sub>4</sub>								0	1	0	0	1

	p <sub>4</sub>	p <sub>3</sub>	p <sub>2</sub>	p <sub>1</sub>	
Binari	1	0	1	1	
Decimal	8		2	1	= 11

Per tant hi ha un error al bit 11 (d<sub>7</sub>)

# RAID5

**Dades: blocs**

**Distribució de dades**

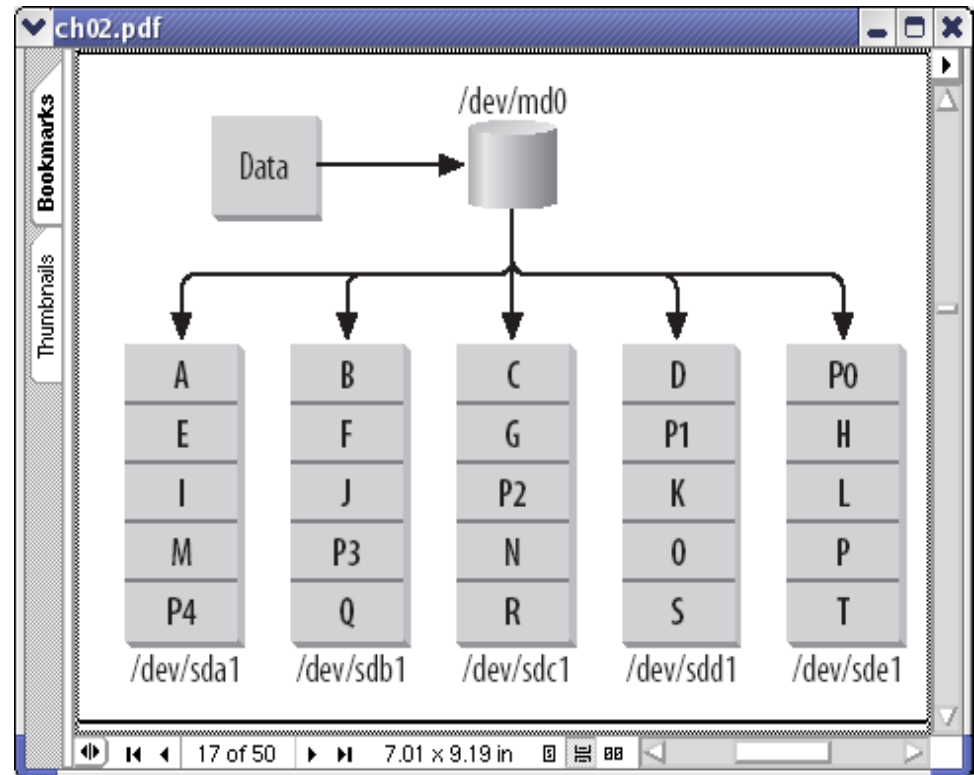
RAID4 + **interleaved parity**

**Avantatge**

Bon rendiment  
(Similar a RAID0)

**Desavantatge**

Sol permet el fallo d'un disc



# RAID6 – P+Q

**Dades: blocs**

**2 Blocs de paritat**

**Permet el fallo de 2 discs**

**Desavantatge**

Més discs de paritat

Disc

0	1	2	3	4	5
0	1	2	3	Q0	P0
5	6	7	Q1	P1	4
10	11	Q2	P2	8	9
15	Q3	P3	12	13	14
Q4	P4	16	17	18	19
P5	20	21	22	23	Q5
24	25	26	27	Q6	P6

# ***RAID6 – Two-Dimensional Matrix***

**Discs organitzats com una matriu de 2-dimensions**

**Parity computed per**

Cada columna

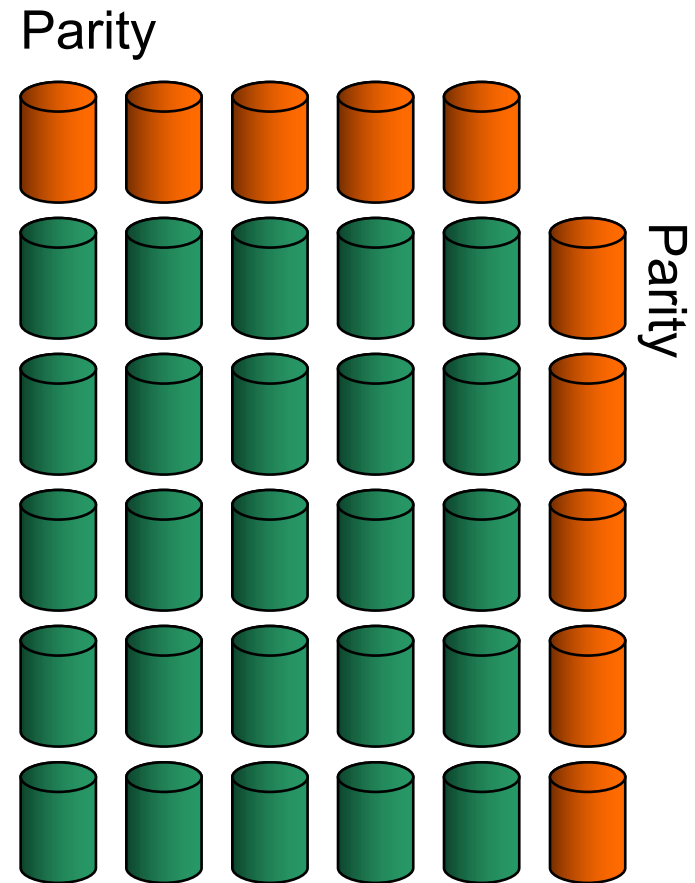
Cada fila

**Avantatge**

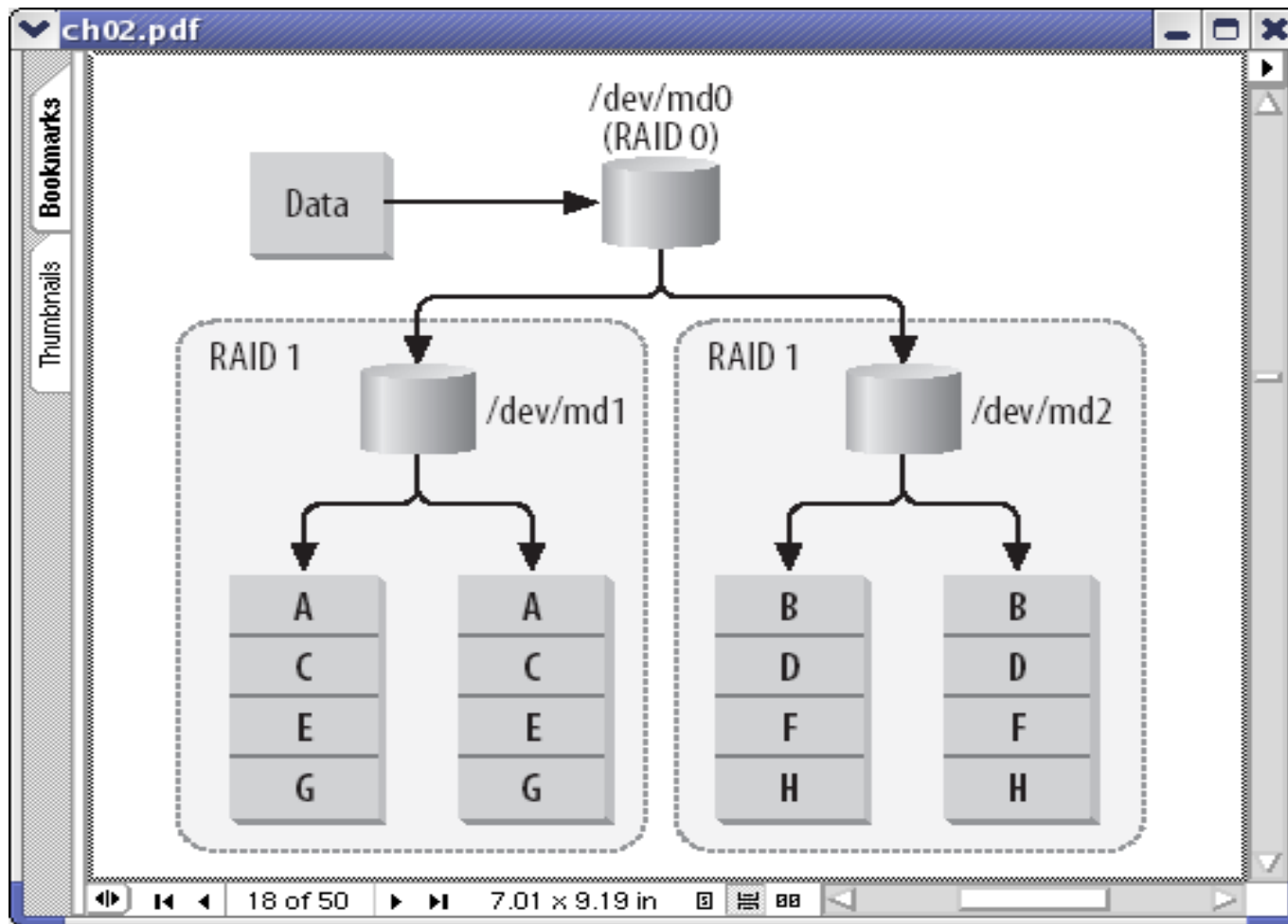
Permet la fallada de 2 discs

**Desavantatge**

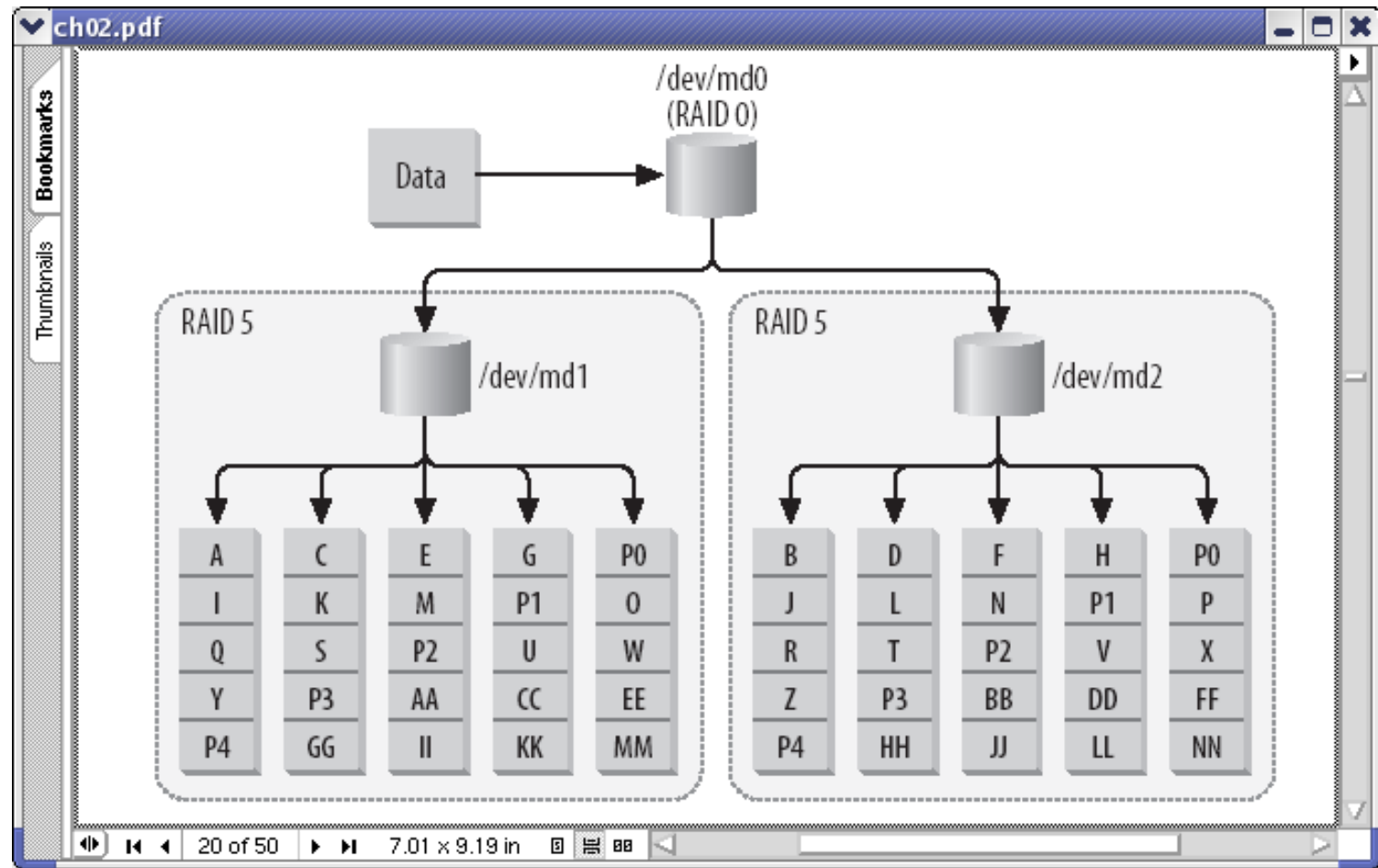
Més blocs de paritat



# RAID10



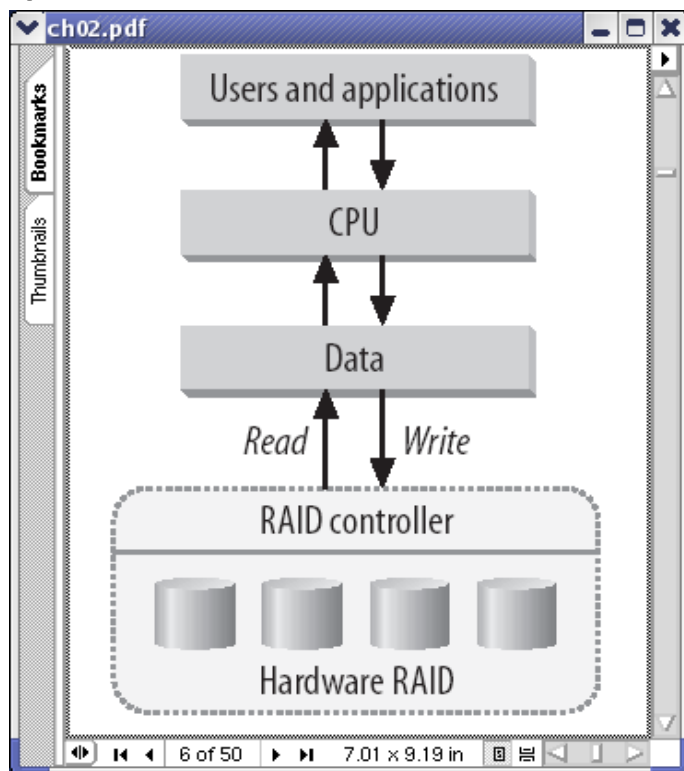
# RAID50



# ***RAID Implementacions***

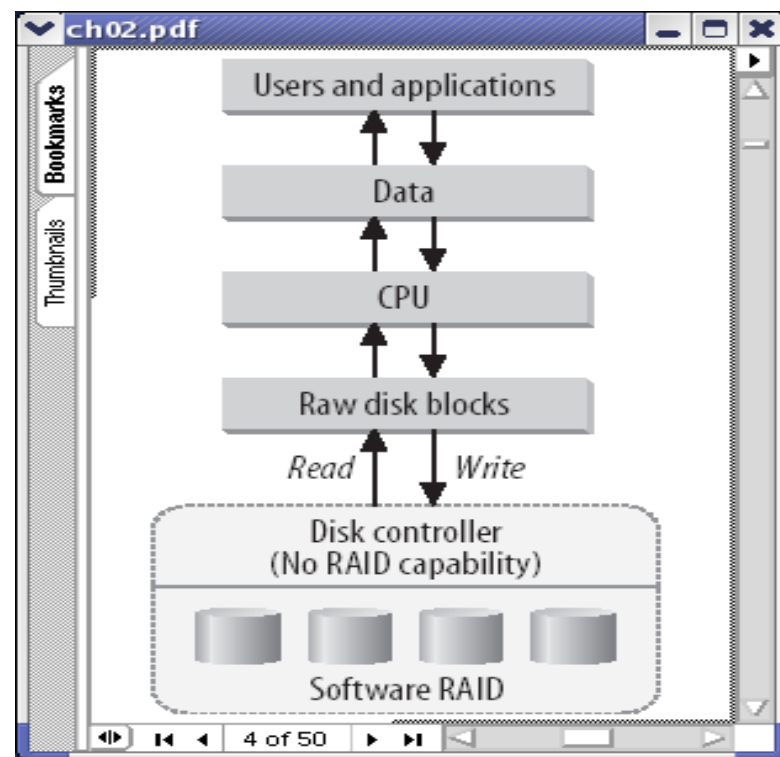
## **Hardware RAID**

- Gestió realitzada per un  $\mu$ -controlador addicional



## **Software RAID**

- Gestió realitzada per la CPU



# ***Software RAIDs en Linux***

**Linux Supports the following Software Raids:**

- **Linear, Raid0, Raid1, Raid4 and Raid 5**

**Paquet Raid:**

**\$ apt-get install mdadm**

**Per carregar el mòdul Raid:**

**\$ modprobe raid1**

o bé

**\$ modprobe raid0**



# *Creating RAID partitions with fdisk*

**# fdisk /dev/sda**

**m:** help

**d:** delete partition

**n:** create new partition

**p:** primary

**e:** extended // extended = logical1 +...+ logicaln

**p:** print partition table

**l:** list partition types

**t:** change partition type (**fd**: Raid auto-detect)

**w:** save & exit

**w:** exit without saving

# ***Creació d'un RAID en Linux***

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

```
Personalities :
```

```
Event: 0
```

```
unused devices: <none>
```

```
# mdadm --create /dev/md0 --level=raid5 --raid-devices=5  
    /dev/sdb1 /dev/sdc1 /dev/sdd1 /dev/sde1 /dev/sdf1
```

```
Continue creating array? yes
```

```
mdadm: array /dev/md0 started.
```

# ***Creació d'un RAID en Linux***

**\$ cat /proc/mdstat**

**Personalities : [raid5]**

**md0 : active raid5 sdb1[0] sdc1[1] sdd1[2] sde1[3] sdf1[4]**

**586075008 blocks level 5, 64k chunk (mida sector),  
algorithm 2 unused devices: <none>**

**Ara sol cal crear i montar el sf del raid**

# *Creació d'un RAID en Linux*

## Creant el sf:

```
# mkfs.ext4 /dev/md0
```

## Montant el RAID:

1. En `/etc/fstab` tenim una línia de la forma:

```
/dev/md0 punt_montatge tipus_sf defaults,user 0 0
```

2. `# mount punt_de_montatge`

o bé `# mount /dev/md0`

o bé reiniciar i que es monti automàticament

# ***Obtenint informació d'un RAID en Linux***

```
$ sudo mdadm --query /dev/md0
```

```
/dev/md0: 558.92GiB raid5 5 devices, 0 spares. Use mdadm --  
detail for more detail.
```

```
$ mdadm --query /dev/sdb1
```

```
/dev/sdb1: is not an md array
```

```
/dev/sdb1: device 0 in 5 device active raid5 /dev/md0. Use  
mdadm --examine for more detail
```

# ***Obtenint informació d'un RAID en Linux***

```
# mdadm --examine /dev/sdb1
```

**Raid Level : raid5**

**Device Size : 195358336 (186.31 GiB 200.05 GB)**

**Raid Devices : 5**

**Update Time : Sun Jan 21 22:23:05 2007**

**Chunk Size : 64K**

<b>Major</b>	<b>Minor</b>	<b>RaidDevice</b>	<b>State</b>
<b>8</b>	<b>17</b>	<b>0</b>	<b>active sync /dev/sdb1</b>
<b>8</b>	<b>33</b>	<b>1</b>	<b>active sync /dev/sdc1</b>