

UFRJ – IM - DCC



Sistemas Operacionais I

Unidade IV Gerência de Memória Secundária



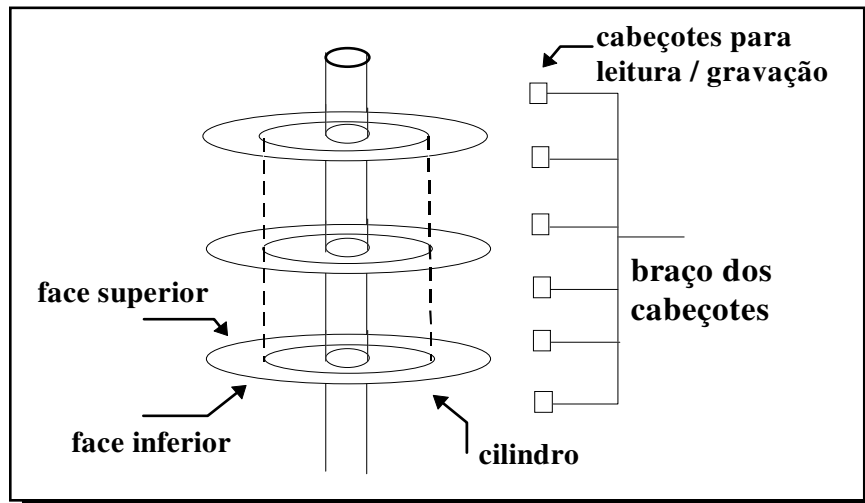
ORGANIZAÇÃO DA UNIDADE

- Gerência de Entrada e Saída
- Gerência de memória secundária
 - Estrutura Física
 - Escalonamento de E/S
 - RAID
- Sistema de arquivos

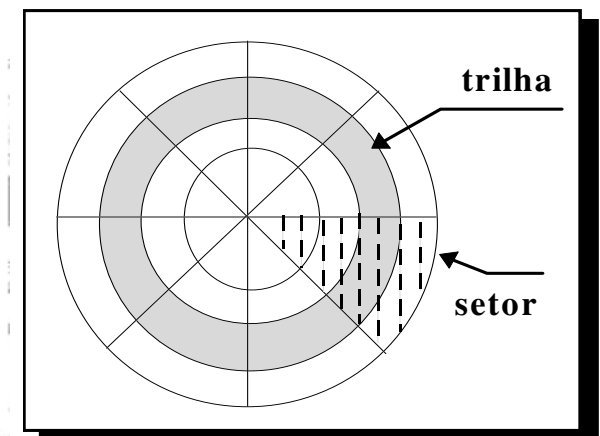
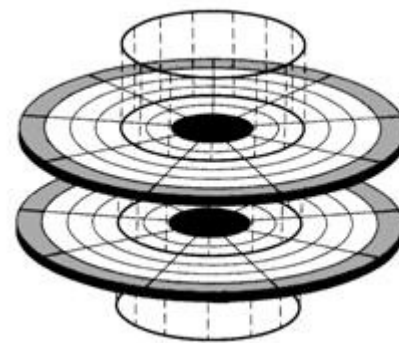
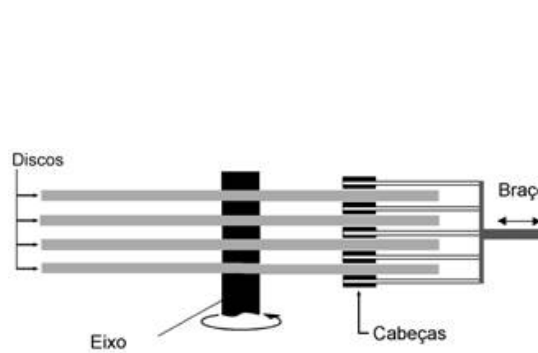


Gerência de Discos

Estrutura Física – Disco Magnético



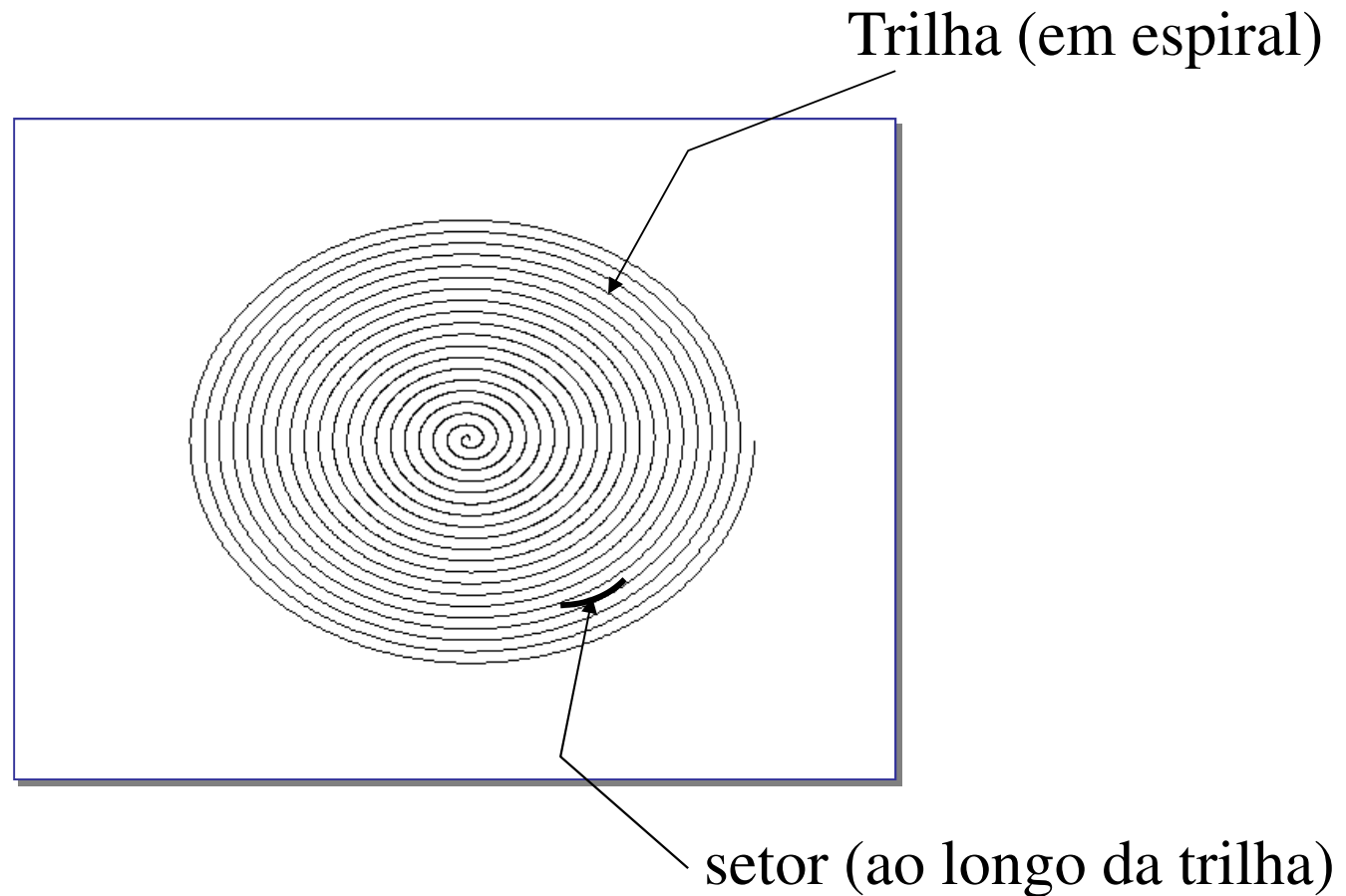
setor
cluster
trilha
cilindro





Gerência de Discos

Estrutura Física – Disco ótico





Gerência de Discos Desempenho

- Disco roda em velocidade constante
- Deve posicionar cabeça na trilha, no início do setor
seek time → tempo gasto para posicionar cabeça na trilha

$$T_s = n \cdot m + S, \quad n = \text{n}^\circ \text{ de trilhas}$$

m = constante (depende do disco)

S = Startup time

atraso rotacional médio (meia volta) → tempo gasto para posicionar setor (latência)

$$A = 1/2r, \quad r = \text{velocidade de rotação (em min.)}$$

transferência → tempo gasto para transferir b bytes

$$T_t = b/rN, \quad N = \text{qtd de bytes na trilha}$$

b = n° de bytes a serem transferidos

$$T = T_s + A + T_t$$



Gerência de Discos

Exemplo

Rotação=3600 rpm , seek time=20ms, 1 setor=512 bytes, 1 trilha=32 setores
Tempo de transferência de uma arquivo de 128 Kb para:

a) Organização Sequencial

Arquivo: 128 Kb => 256 setores, logo, ocupa $256/32 = 8$ trilhas

Para ler primeira trilha: seek = 20 ms atraso = $1/2r = 60/(2 \times 3600)$ s = 8,3 ms

leitura de 32 setores = $60/3600 = 16,7$ ms

$$T1 = 20 + 8,3 + 16,7 = 45 \text{ ms}$$

Próximas trilhas: seek ~ 0

$$T2 = 8,3 + 16,7 = 25 \text{ ms para cada trilha}$$

$$T = 45 + 7 \times 25 = 220 \text{ ms}$$

b) Randômico

Tempo para ler um setor: $T1 = 16,7/32 = 0,5$ ms

$$T = 256 \times (20 + 8,3 + 0,5) = 7373 \text{ ms}$$



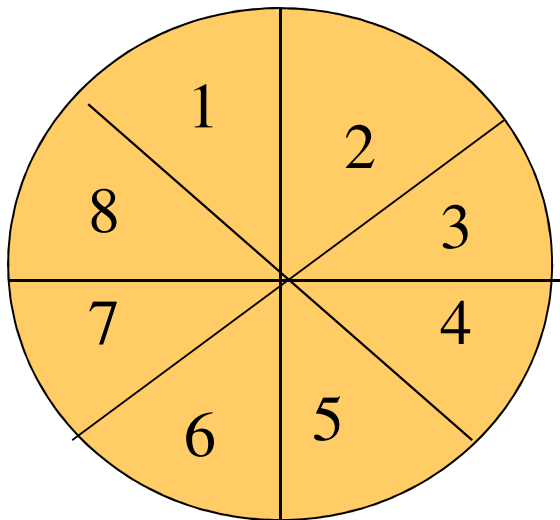
Motivação

- Após a leitura de um setor, há um intervalo para transferência do bloco lido.
- Haverá um atraso rotacional mesmo que os setores sejam consecutivos pois a cabeça de leitura/gravação já terá passado pelo início do setor.

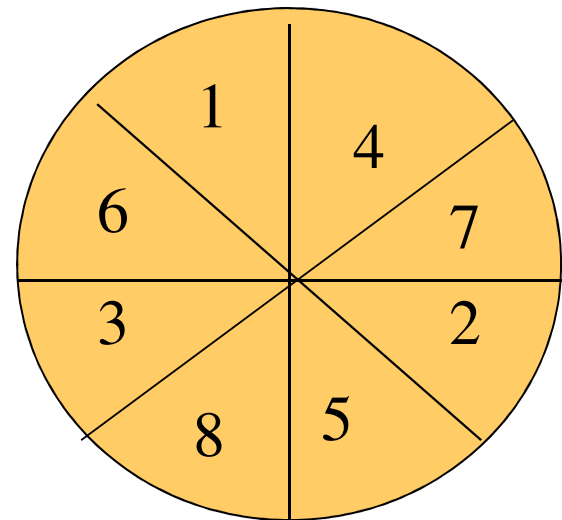
Solução
numeração dos setores de forma não contígua
(*interleaving*)



Gerência de Discos Exemplo



Fator = 0



Fator = 2



Baseado na fila de requisição

■ FIFO

- Mais simples
- Atendimento na ordem dos pedidos

■ Prioridade

- fora do controle do gerenciador

■ LIFO

- Diminui o movimento da cabeça de leitura em arquivos sequenciais



Escalonamento de disco

Baseado na localização do que foi requisitado

- **SSTF** (shortest service time first)
 - Fila é reordenada para atender as requisições de forma a minimizar o movimento da cabeça
 - Possibilidade de starvation
- **Scan** (elevador)
 - Variação do SSTF porém estipula uma direção preferencial
 - O sentido se inverte ao final da varredura
- **C-Scan**
 - Semelhante ao Scan porém com um sentido único



Baseado na localização segmentada

■ N-step-Scan

- Divide as requisições em filas de tamanho N
- As requisições de cada fila são processadas em conjunto
- Novas requisições são colocadas em novas filas
- Para valores grandes de N , é semelhante ao SCAN
- Para $N = 1$, equivale a FIFO

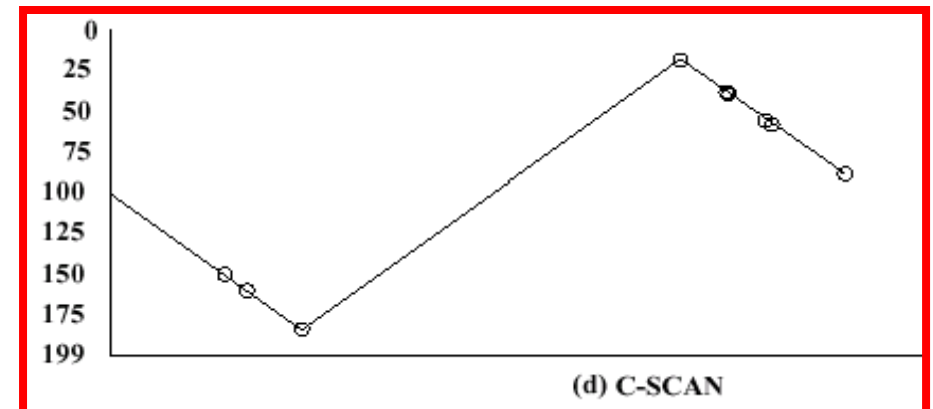
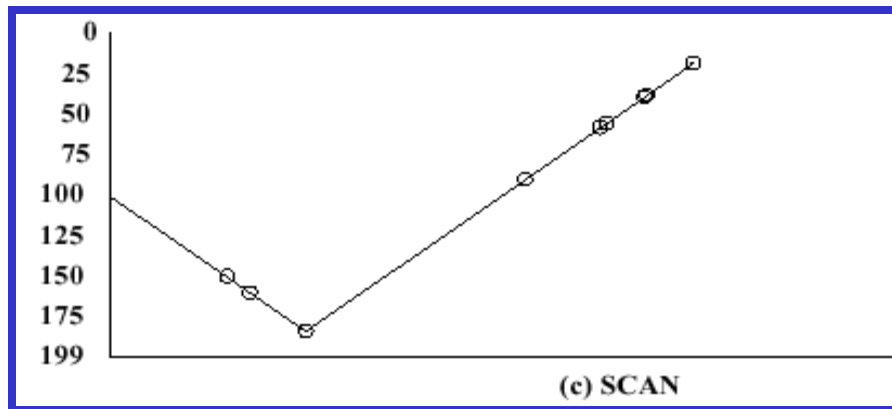
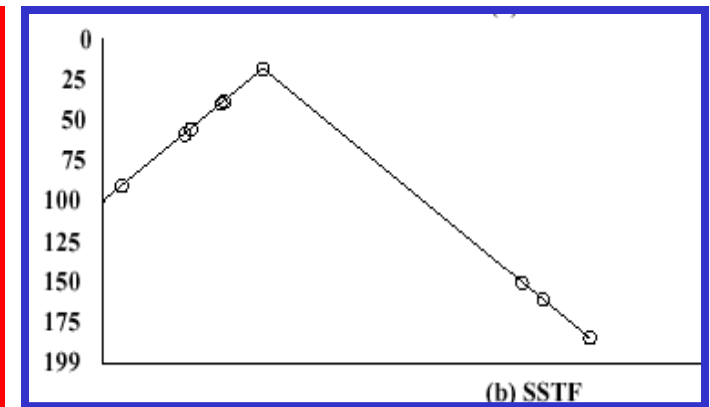
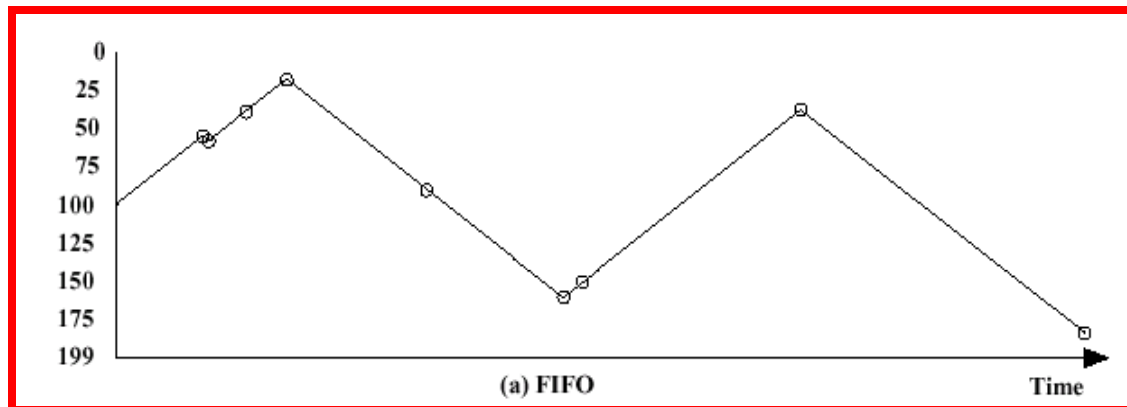
■ FSCAN

- Utiliza duas filas.
- Enquanto uma fila está sendo atendida, novas requisições são colocadas na outra fila.



Gerência de Discos Exemplo

Sequência iniciando na trilha 100: 55,58,39,18,90,160,150,38,184





Redundant Array of Independent Disks



RAID 5

- ▶ Optimiza a capacidade
- ▶ Tolerância a falhas
- ▶ Dados fraccionados



Redundant Array of Independent Disks

- Conjunto de discos que são vistos pelo S.O com uma única unidade lógica
- Os dados são distribuídos pelos discos
- Discos redundantes oferecem confiabilidade e a possibilidade de recuperação em caso de falhas
- São classificados em 7 níveis (0 até 6)



Redundant Array of Independent Disks

Tecnologia proposta em 1987 por Peterson, Gibson e Katz na Universidade de Berkeley – Califórnia.

- ✓ Nível 0: chamado de striping
- ✓ Nível 1: chamado mirroring, shadowing ou duplexing
- ✓ Nível 2: chamado striping with parity (obsoleto)
- ✓ Nível 3: chamado disk array with bit-interleaved data
- ✓ Nível 4: chamado disk array with block-interleaved data
- ✓ Nível 5: chamado disk array with block-interleaved distributed parity
- ✓ Nível 6: chamado disk array with block-interleaved distributed parity



Gerência de Discos RAID nível 0

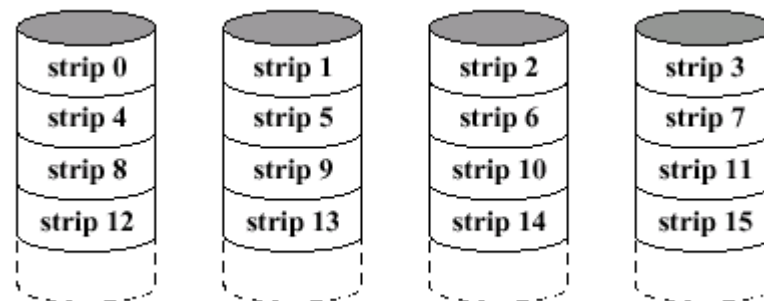
Conhecido como "Striping" ou "Fracionamento".

Os dados são divididos em pequenos segmentos e distribuídos entre os discos.

Não oferece tolerância a falhas, pois não existe redundância.

Usado para melhorar a performance do computador.

Muito usado em aplicações de CAD e tratamento de imagens e vídeos.

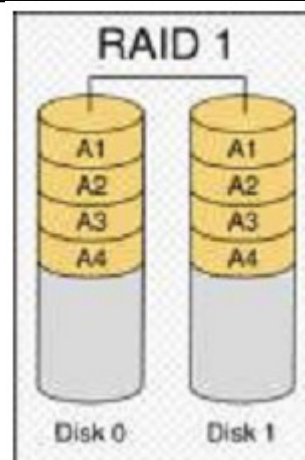




Gerência de Discos

RAID nível 1

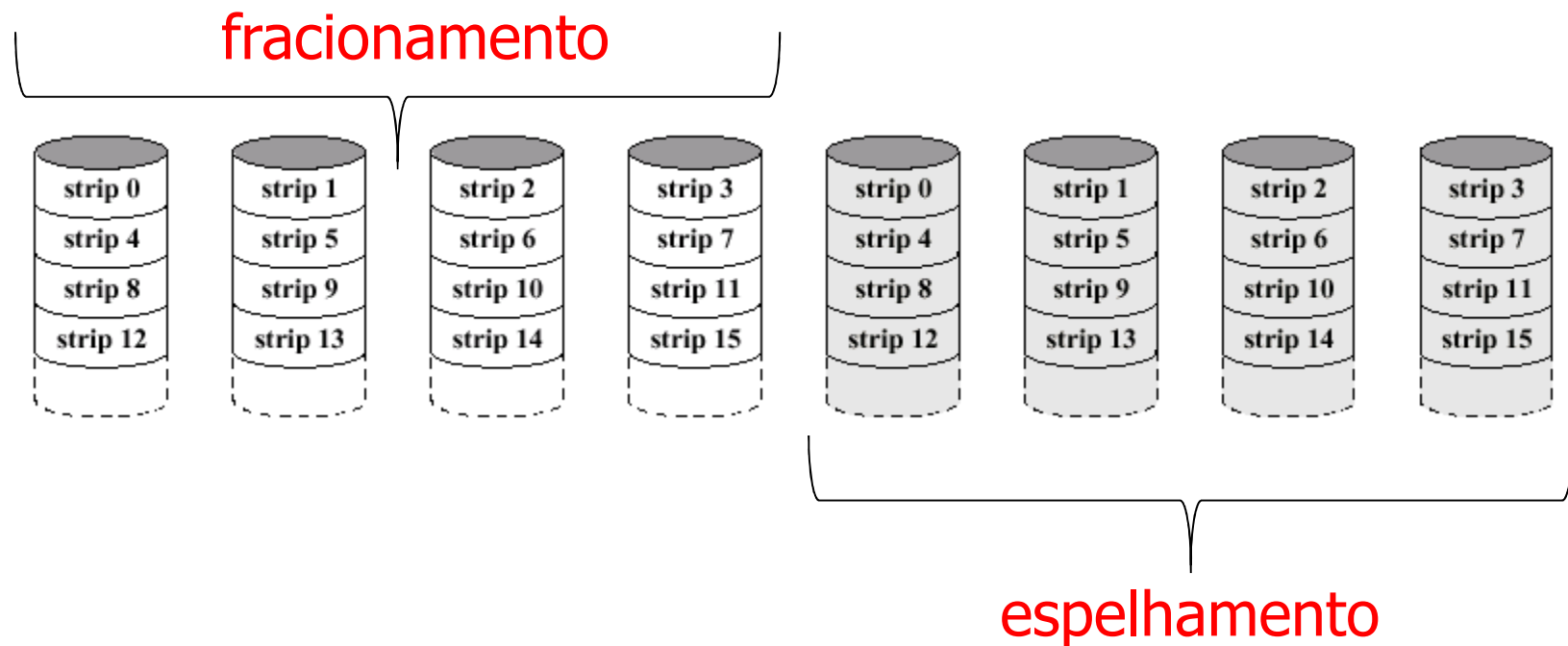
- ✓ Conhecido como "Mirroring" ou "Espelhamento".
- ✓ Funciona adicionando discos paralelos aos principais existentes no computador.
- ✓ Os discos que foram adicionados trabalham como uma cópia do primeiro.
 - ✓ Em caso de falha, a recuperação é imediata
- ✓ Gravação de dados é mais lenta, pois é realizada duas vezes, no entanto, a leitura dessas informações é mais rápida, pois pode-se acessar duas fontes.
- ✓ aplicação muito comum em servidores de arquivos.





Gerência de Discos RAID nível 0+1

Arquitetura que combina o RAID 0 com o RAID 1





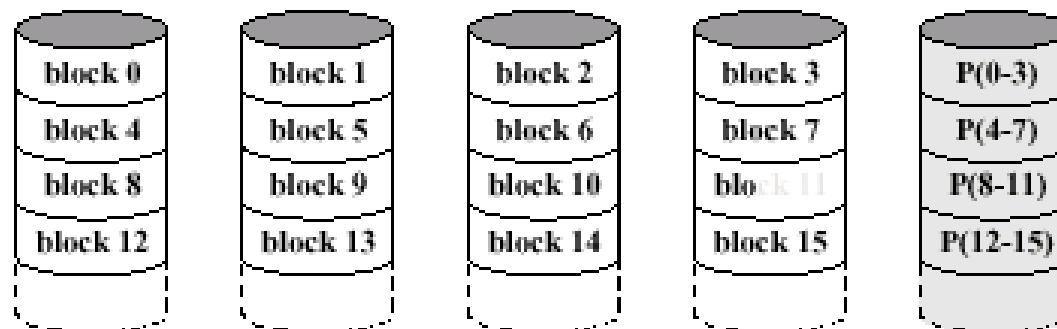
Gerência de Discos RAID nível 2/3/4

- ✓ Os dados são fracionados em diferentes discos – tipo RAID 0

RAID 2: paridade ECC (Error Correcting Code)

RAID 3: paridade a byte em disco extra

RAID 4: paridade de bloco em disco extra

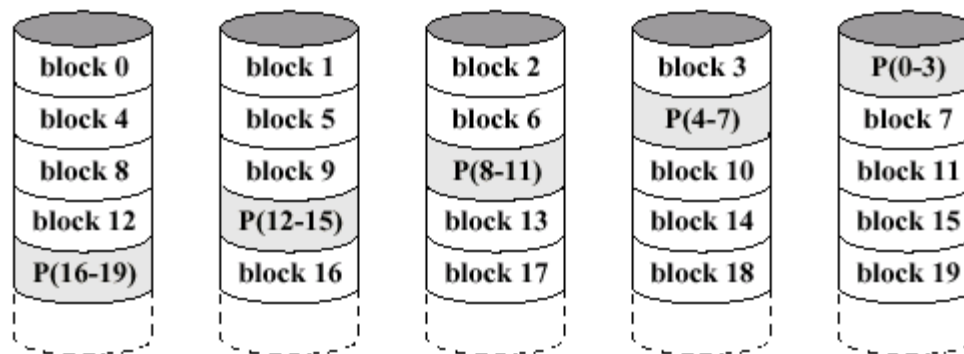


Ex. RAID 4



Gerência de Discos RAID nível 5

- ✓ Semelhante ao RAID nível 4 porém a paridade não fica armazenada em um único disco
- ✓ A gravação de dados é mais rápida, pois não é necessário acessar um disco de paridade a cada gravação.
- ✓ Mais utilizado e oferece resultados satisfatórios em aplicações não muito pesadas.





Gerência de Discos RAID nível 6

É a configuração “mais nova” e relativamente rara de ser achada, segue a mesma ideia do RAID 5 porém usa o dobro de bytes de paridade para recuperação.

É possível recuperar falha em até 2 unidades