

Discos

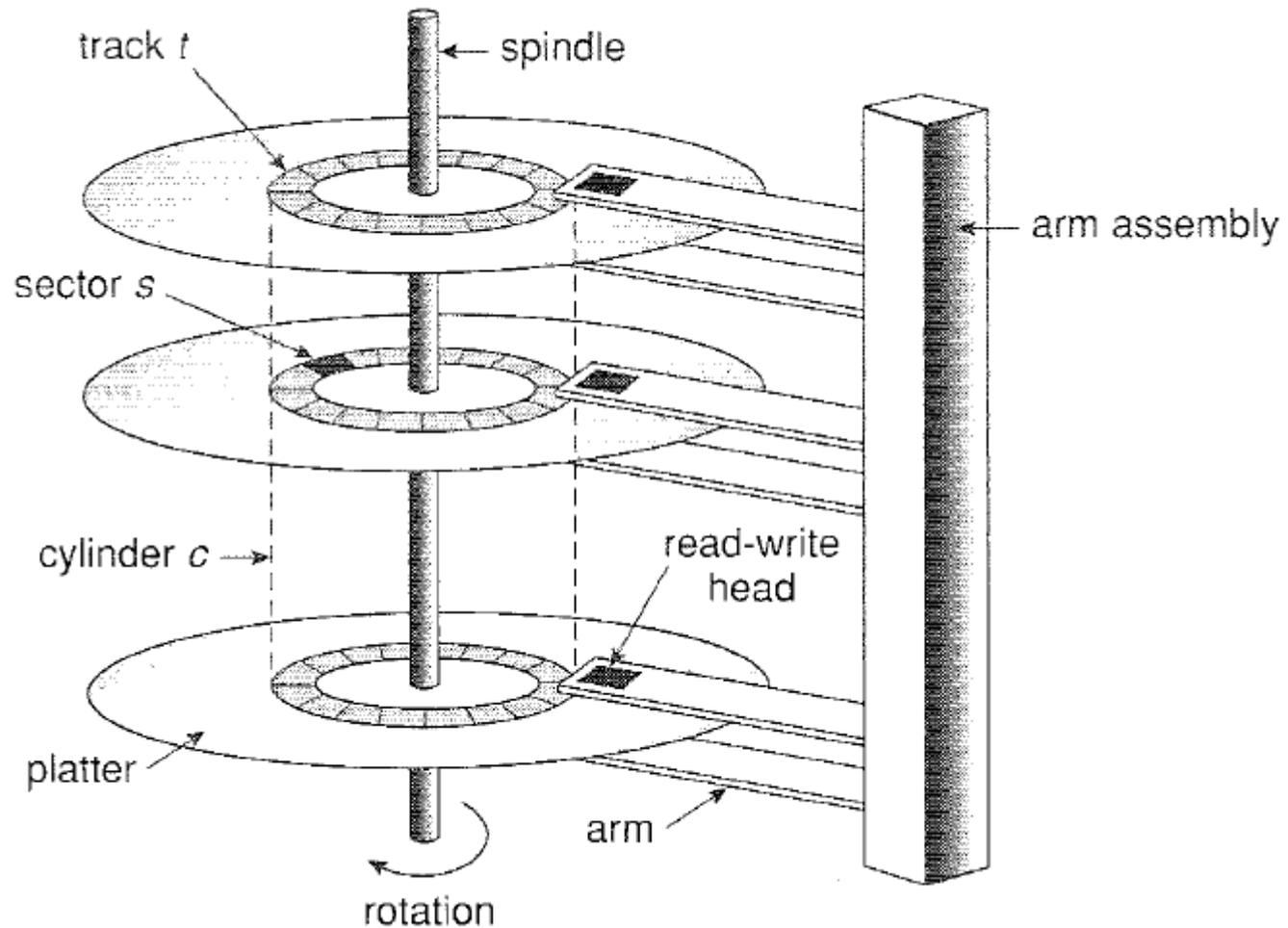
Hardware

Tadeu Ferreira Oliveira - tadeu.ferreira@ifrn.edu.br

Hardware

- Um ou vários discos
- Um ou vários braços
- Cada braço com uma cabeça de leitura
- Dividido logicamente em:
 - Cilindro
 - Trilha
 - Setor

Divisão Lógica



Evolução

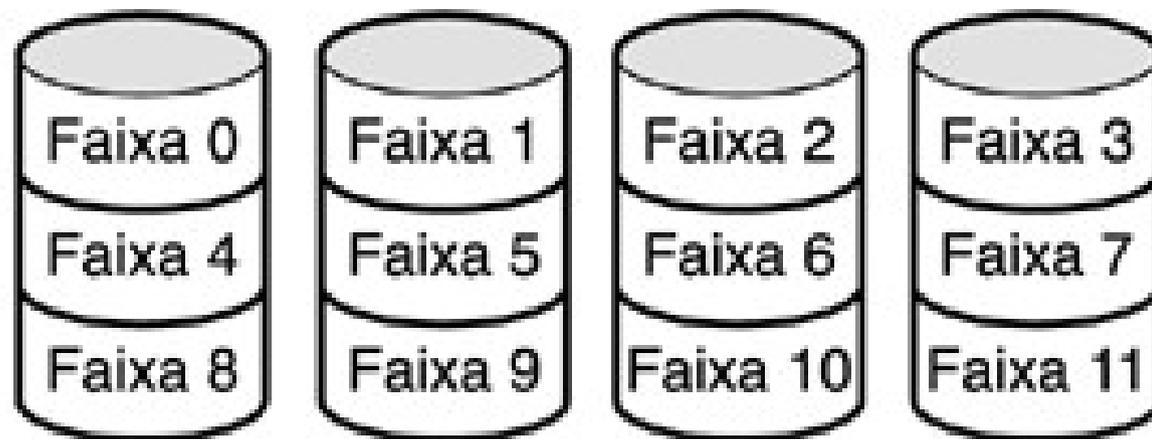
Parâmetro	Disco flexível IBM 360 KB	Disco rígido WD 18300
Número de cilindros	40	10 601
Trilhas por cilindro	2	12
Setores por trilha	9	281 (avg)
Setores por disco	720	35 742 000
Bytes por setor	512	512
Capacidade do disco	360 KB	18,3 GB
Tempo de posicionamento (cilindros adjacentes)	6 ms	0,8 ms
Tempo de posicionamento (caso médio)	77 ms	6,9 ms
Tempo de rotação	200 ms	8,33 ms
Tempo de pára/inicia do motor	250 ms	20 s
Tempo de transferência para um setor	22 ms	17 μ s

Redundant Array of Inexpensive Drivers (RAID)

- RAID 0
 - Divide em strips
 - Mais Performance
 - Menos segurança
- RAID 1
 - Backup de cada drive
 - Boa segurança
- RAID 2
 - Boa segurança
 - Codificação Hamming
- RAID 3
 - Um disco de paridade
- RAID 4
 - Paridade por faixa em um disco
- RAID 5
 - Paridade por faixa espalhada nos vários discos

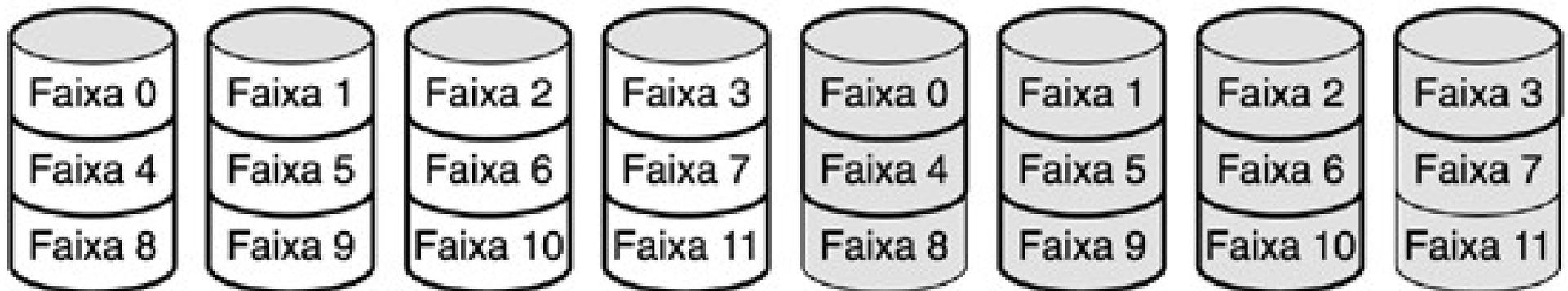
RAID 0

- Divide em Faixas (strips)
- Mais Performance Menos segurança



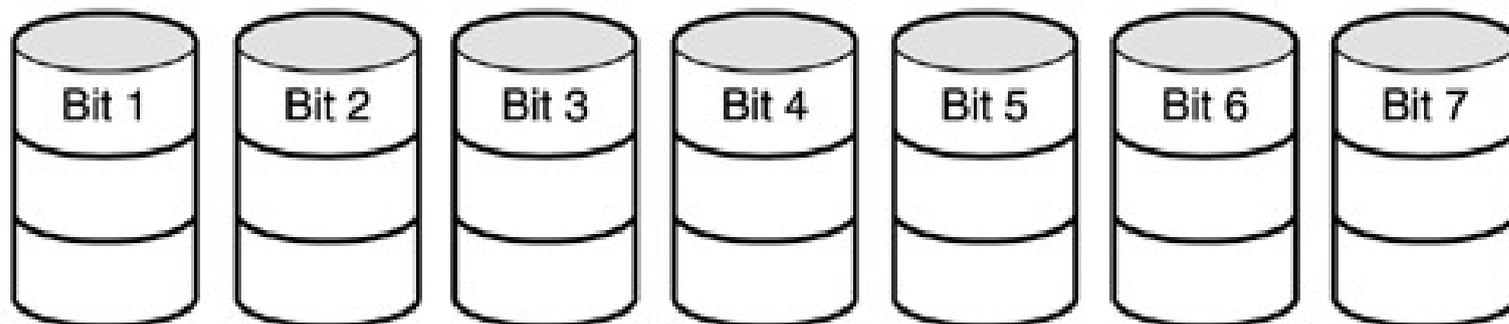
RAID 1

- Divide em Faixas
- Um backup para cada disco
- Escrita igual leitura até 2x mais rápido
- Boa Segurança



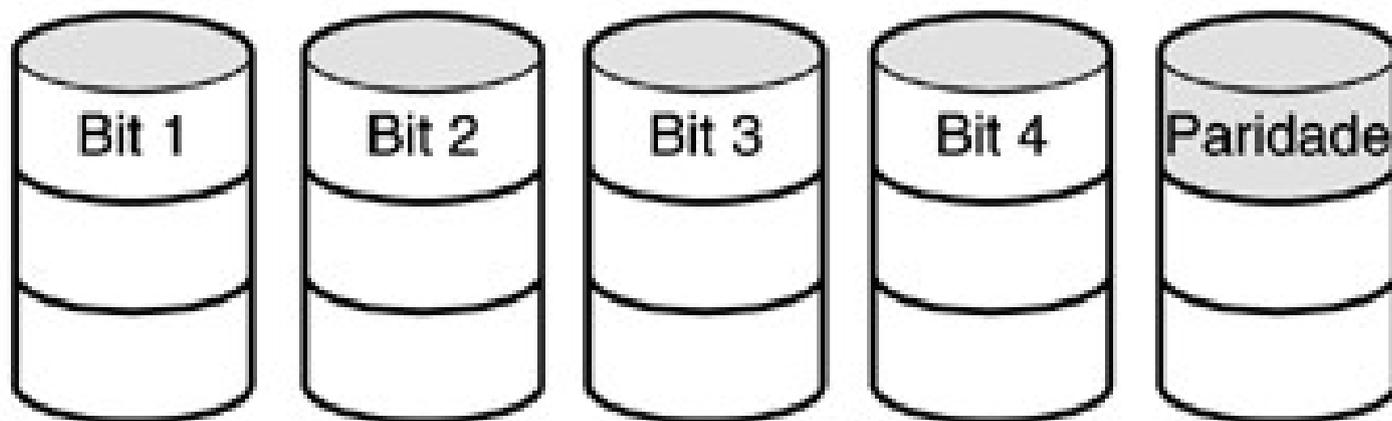
RAID 2

- Divide em byte ou palavra
- Usa a codificação Hamming
- Overhead excessivo



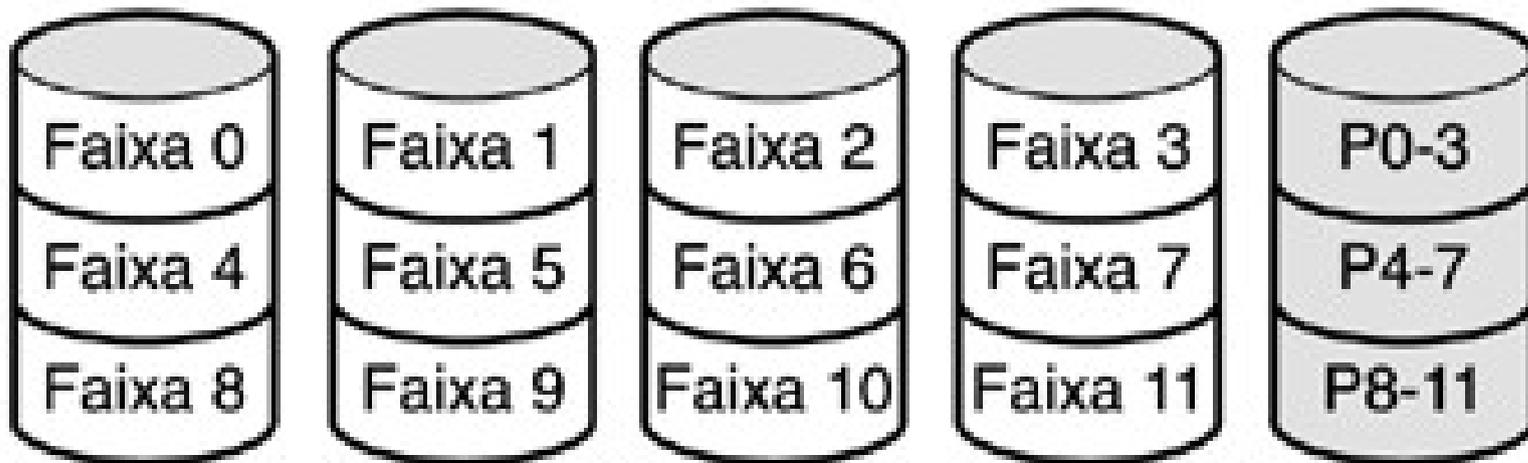
RAID 3

- Versão simplificada do RAID 2
- Com apenas um bit de paridade
- É possível recuperar-se da falha de um disco



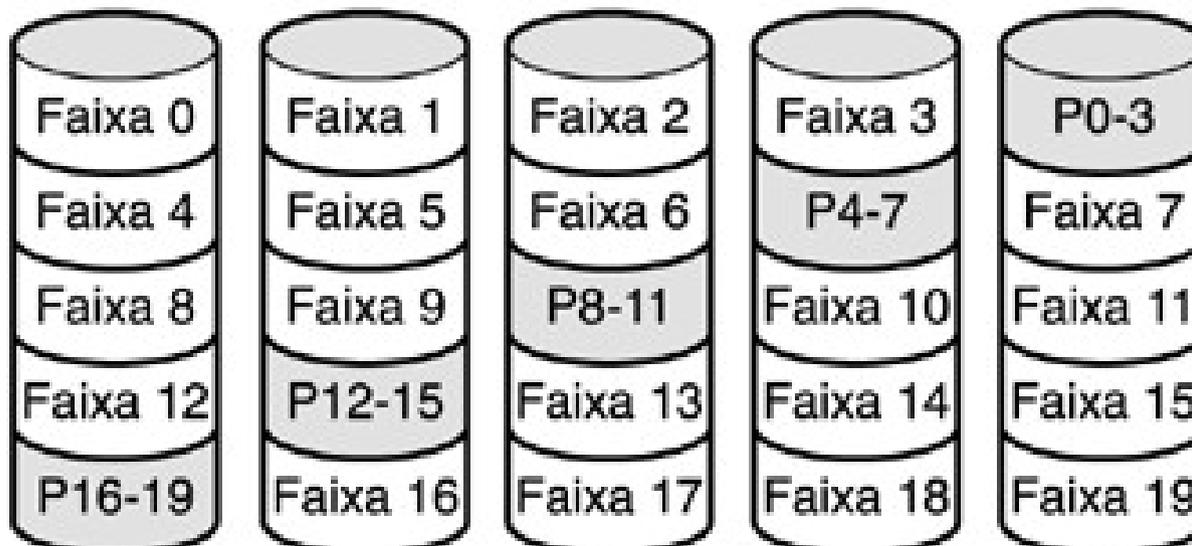
RAID 4

- Usa strips como o RAID 0
- Um drive apenas para paridade das faixas
- Se apenas uma faixa é mudada deve-se ler todas as outras para recalcular a paridade



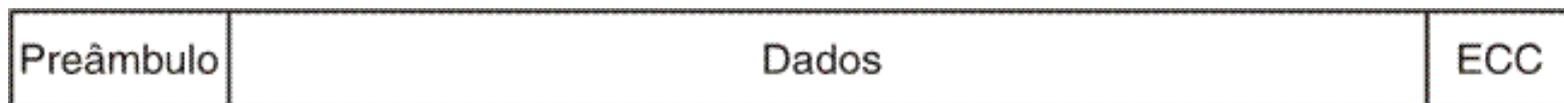
RAID 5

- A mesma idéia do RAID 4 porém a paridade é espalhada nos discos
- Melhora a performance
- A recuperação é mais complexa



Formatação do disco

- O preâmbulo ocupa espaço do disco
- Perde em média 20% por causa do ECC e preâmbulo
- Formatação de baixo nível
- Capacidade real X capacidade formatada



Boot

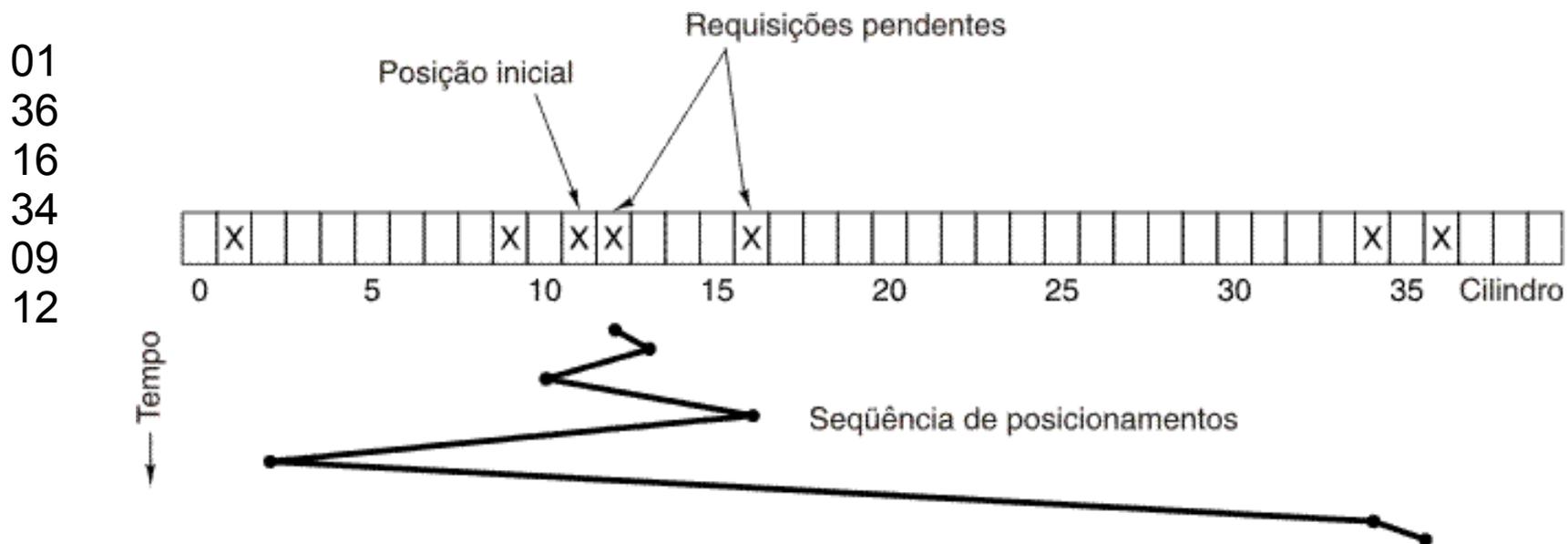
- BIOS inicia
- Busca o Setor zero do disco
- Lá deve encontrar o MBR (Master Boot Record)
- Este é executado e carrega o loader que inicia o sistema de arquivo do Sistema Operacional e carrega o kernel

Algoritmos de Escalonamento do Braço

- O tempo de leitura de um dado do disco é dividido em três etapas:
 - Tempo de Seek
 - Tempo de Rotação
 - Tempo de transferência do dado
- O tempo de Seek é o maior

Menor Seek Primeiro

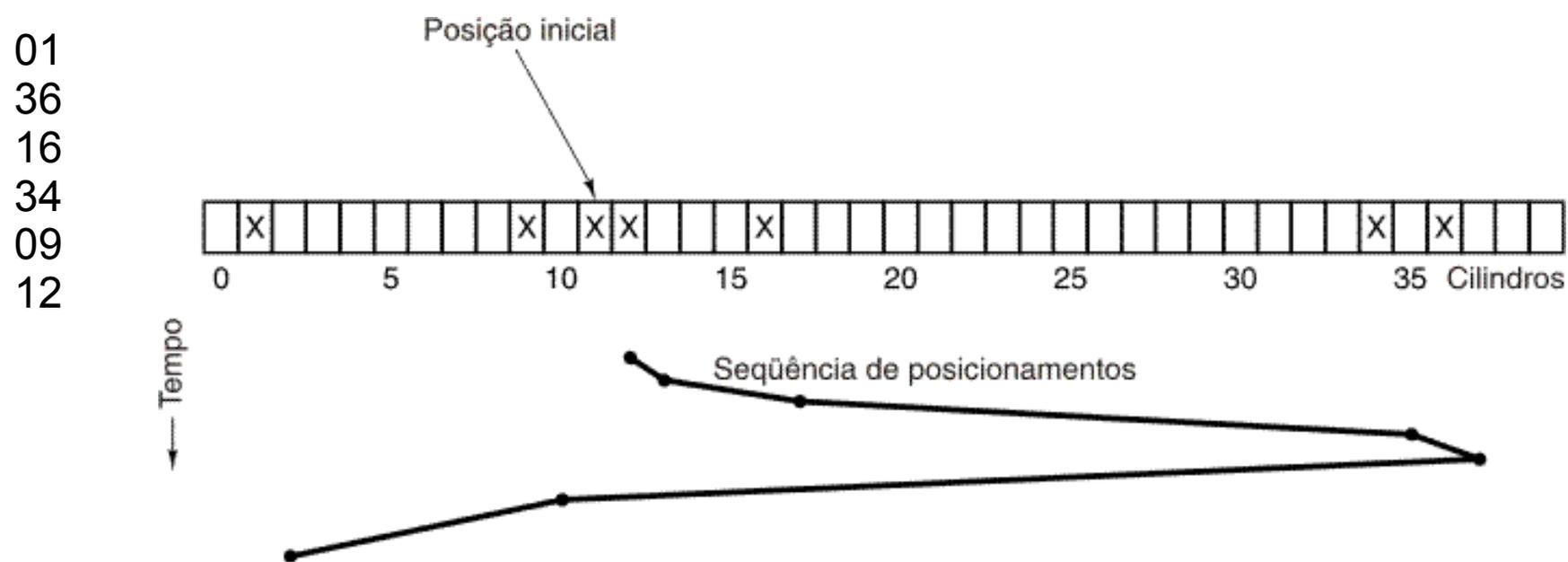
- Shortest Seek First (SSF)



O que acontece se tivermos muitos pedidos no centro do disco e poucos nas bordas?

Algoritmo do elevador

- Algoritmo do elevador
- A idéia é seguir sempre na mesma direção

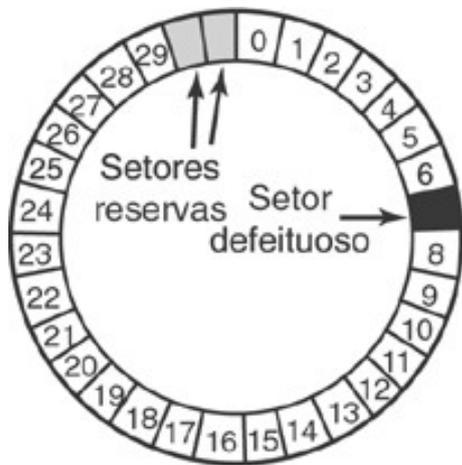


Algoritmo do elevador

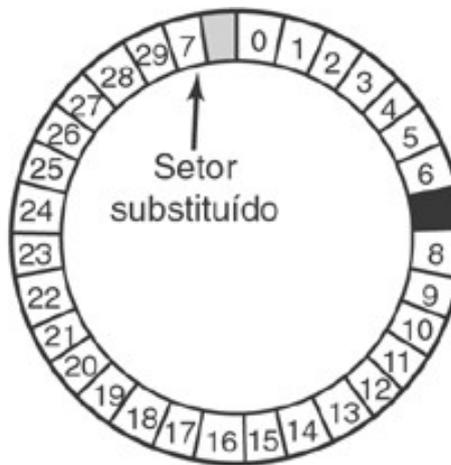
- No caso médio é pior que o SSF
- Oferece a certeza de que sempre atenderá a todos
- No pior caso um tempo de acesso constante de 2 vezes o número de cilindros
- A variação de (Teory, 1972) oferece uma pequena melhora no tempo de resposta
 - Ao chegar ao máximo mova o braço para o início do disco

Tratamento de Erro

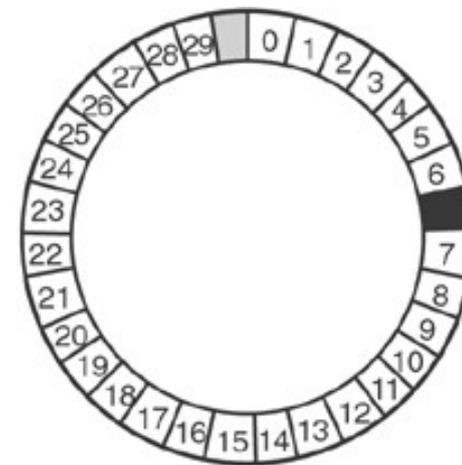
- Bad Sectors
 - Tratados pela controladora
 - Tratados pelo S.O.



(a)



(b)



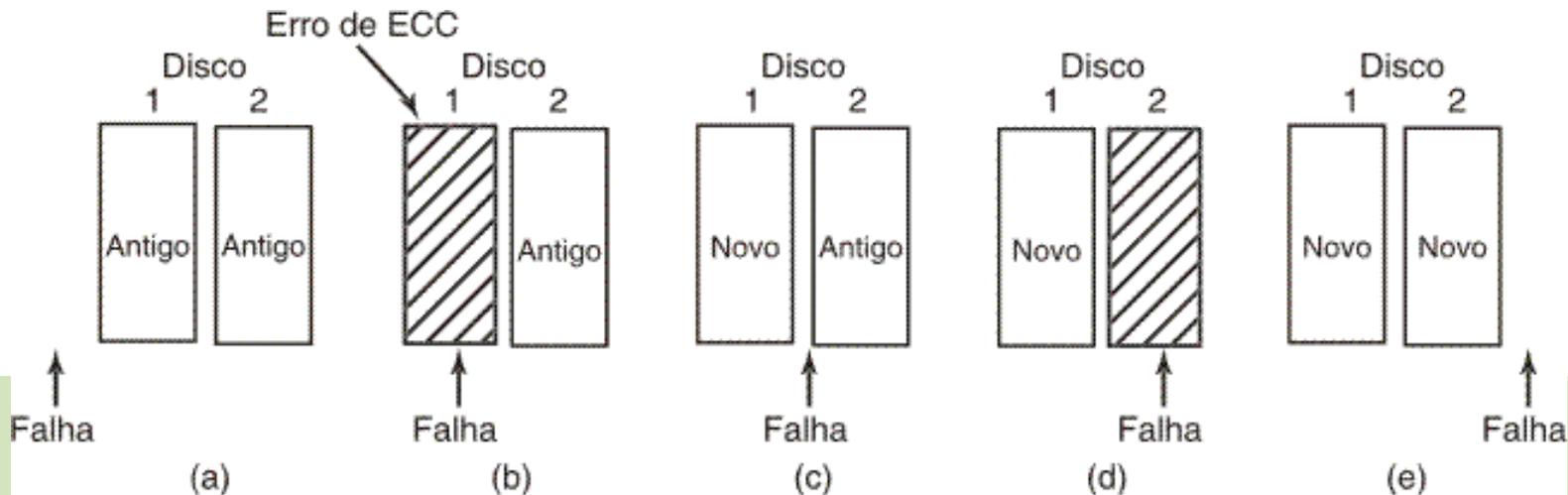
(c)

Tratamento de Erros

- Outro tipo de erro é o erro mecânico
- Erros de Seek
 - Tratados pela controladora
 - Ou pelo Driver
- Bugs na controladora
 - O driver deve ser capaz de perceber
 - Pode haver um comando para resetar a controladora

Stable Storage

- Usa dois discos para garantir leituras e gravações seguras
- Stable Writes
- Stable Reads
- Recuperação de Falha

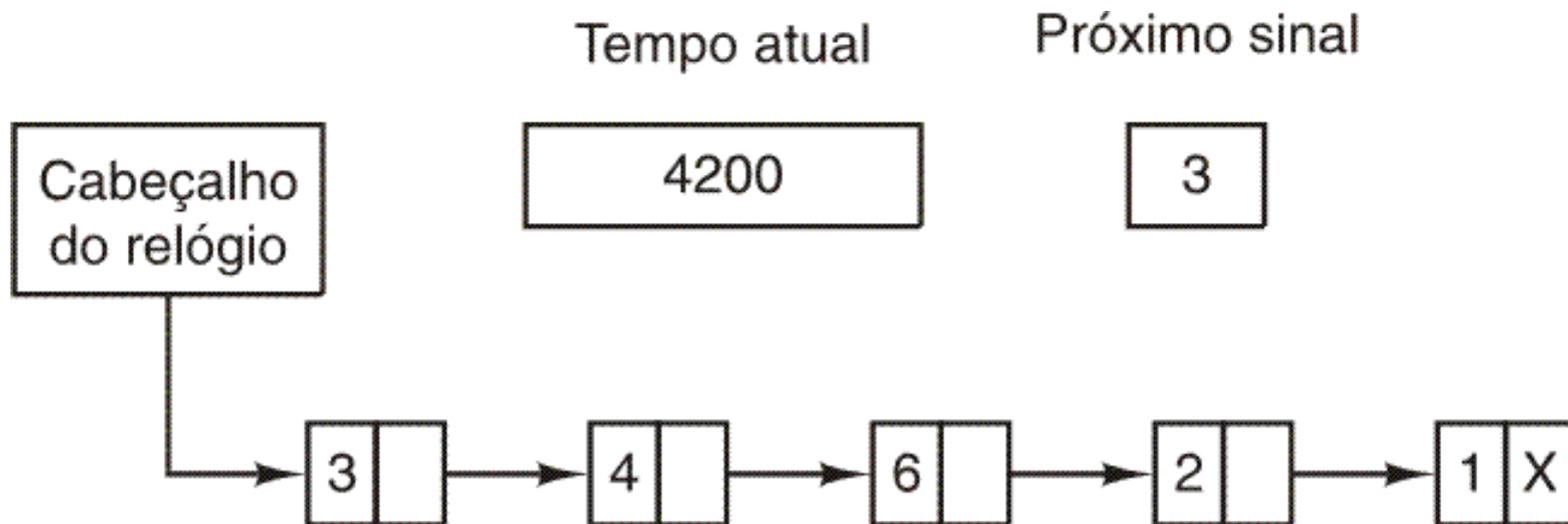


Relógios

- O hardware é basicamente uma peça de cristal de quartzo atravessada por uma corrente
- O cristal vibra em uma frequência conhecida
- Um registrador tem o numero ticks a contar
- Quando o registrador zerar gera uma interrupção para a CPU

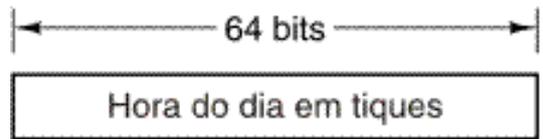
Relógios

- Podem ser programados para gerar uma interrupção a cada X ms

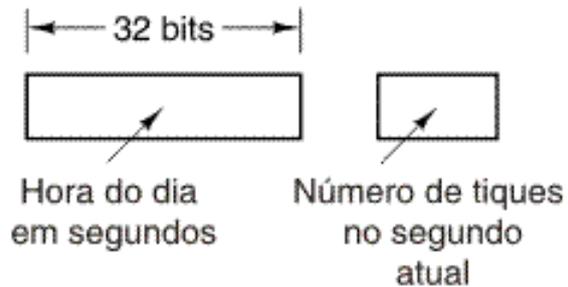


Relógios

- A hora do dia



(a)



(b)



(c)

Timers em Software

- Evitam tratar interrupções
- É testado sempre que o sistema vai do modo usuário para o modo kernel
- Pode ocorrer alguma imprecisão
- Para tempos da ordem de ms é aceitável
- Testes mostram que a chamadas de kernel ocorrem em média entre 2 e 18 micro-segundos