

UFFS

Ciência da Computação

Sistemas Operacionais

Prof. Marco Aurélio Spohn

Entrada/Saída

- 5.1 Princípios do hardware de E/S
- 5.2 Princípios do software de E/S
- 5.3 Camadas do software de E/S
- 5.4 Discos
- 5.5 Relógios
- 5.6 Terminais com base em caracteres
- 5.7 Interfaces gráficas do usuário
- 5.8 Terminais de rede
- 5.9 Gerenciamento de energia

Princípios do Hardware de E/S

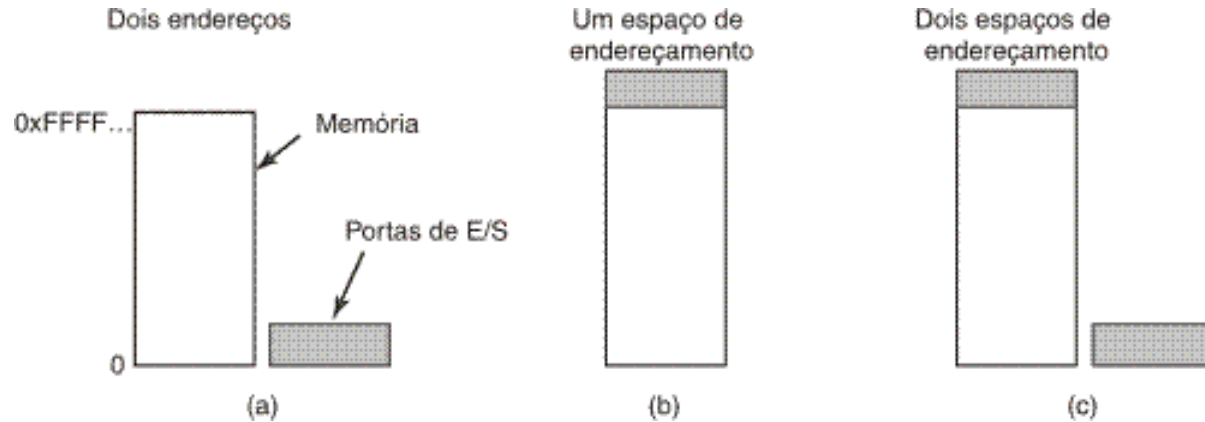
Dispositivo	Taxa de dados
Teclado	10 bytes/s
Mouse	100 bytes/s
Modem 56 K	7 KB/s
Canal telefônico	8 KB/s
Linhas ISDN dual	16 KB/s
Impressora a laser	100 KB/s
Scanner	400 KB/s
Ethernet clássica	1,25 MB/s
USB (<i>universal serial bus</i> — barramento serial universal)	1,5 MB/s
Câmara de vídeo digital	4 MB/s
Disco IDE	5 MB/s
CD-ROM 40x	6 MB/s
Ethernet rápida	12,5 MB/s
Barramento ISA	16,7 MB/s
Disco EIDE (ATA-2)	16,7 MB/s
FireWire (IEEE 1394)	50 MB/s
Monitor XGA	60 MB/s
Rede SONET OC-12	78 MB/s
Disco SCSI Ultra 2	80 MB/s
Ethernet Gigabit	125 MB/s
Dispositivo de Fita Ultrium	320 MB/s
Barramento PCI	528 MB/s
Barramento da Sun Gigaplane XB	20 GB/s

- Taxas de dados típicas de dispositivos, redes e barramentos

Controladores de Dispositivos

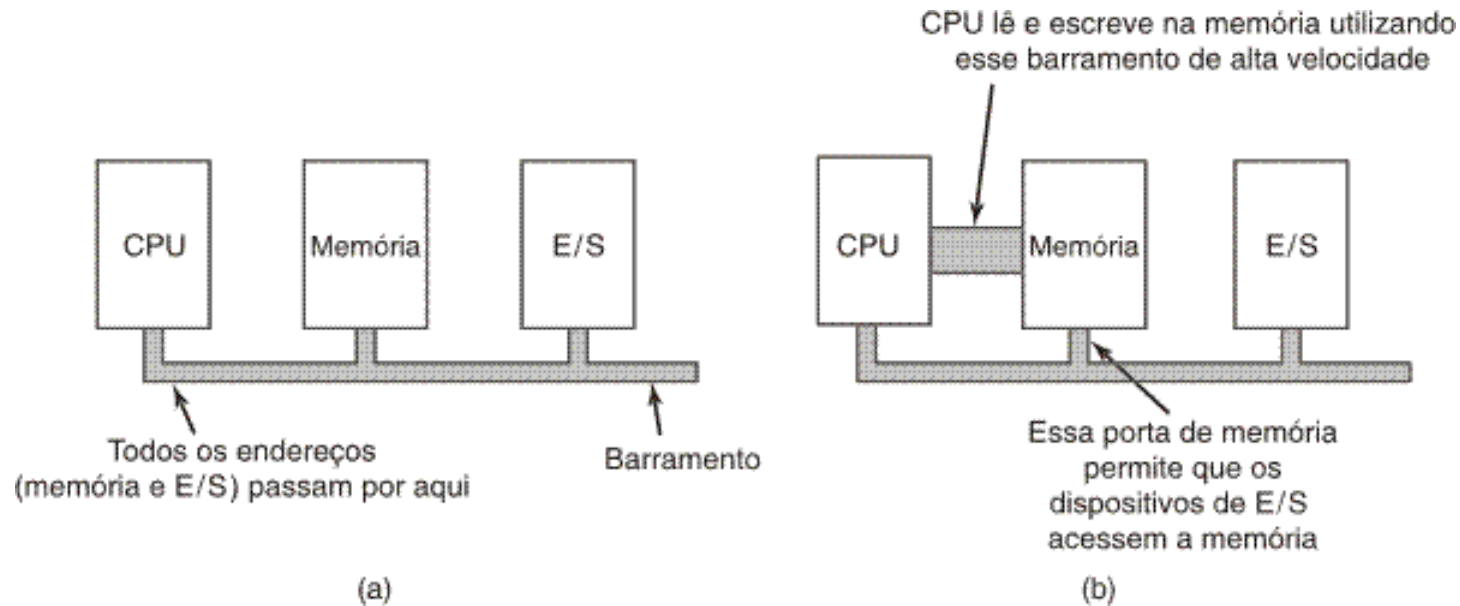
- Componentes de dispositivos de E/S
 - *mecânico*
 - *eletrônico*
- O componente eletrônico é o controlador do dispositivo
 - *pode ser capaz de tratar múltiplos dispositivos*
- Tarefas do controlador
 - *converter fluxo serial de bits em bloco de bytes*
 - *executar toda correção de erro necessária*
 - *tornar o bloco disponível para ser copiado para a memória principal*

E/S mapeada na memória (1)



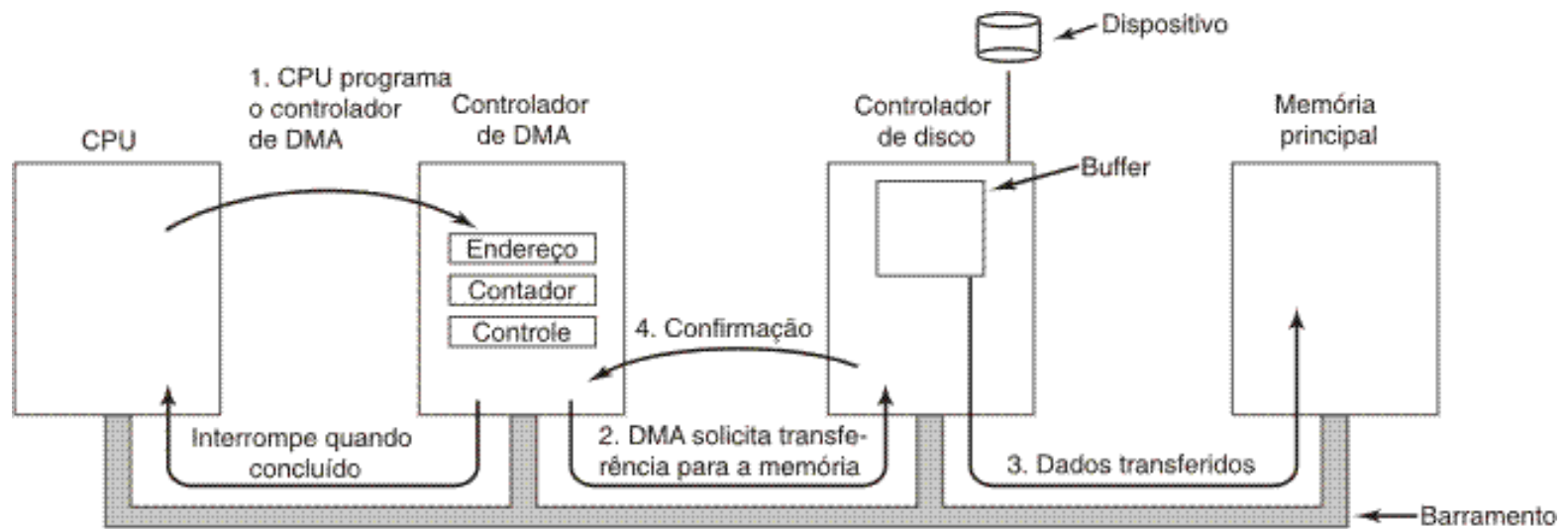
- a) Espaços de memória e E/S separados
- b) E/S mapeada na memória
- c) Híbrido

E/S mapeada na memória (2)



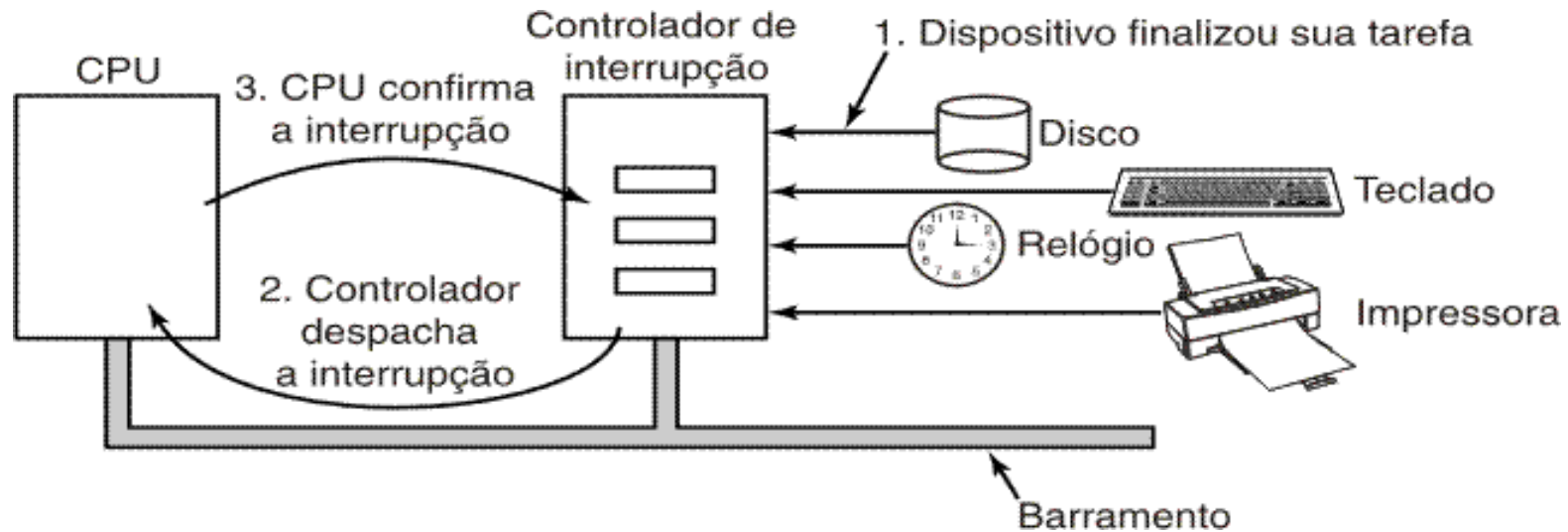
- (a) Arquitetura com barramento único
- (b) Arquitetura com barramento dual (acesso à memória, geralmente, é muito mais rápido do que acesso a outros dispositivos de E/S)

Acesso Direto à Memória (DMA)



- Operação de uma transferência com DMA

Interrupções Revisitadas



- Como ocorre uma interrupção: Conexões entre dispositivos e controlador de interrupção usam linhas de interrupção no barramento em vez de fios dedicados. **Somente após tratar a interrupção** a CPU confirma a interrupção, evitando assim que outras interrupções interrompam a CPU (i.e., após o ack, o controlador de interrupção está liberado para enviar outro sinal de interrupção).

Princípios do Software de E/S

Objetivos do Software de E/S (1)

- Independência de dispositivo
 - *Programas podem acessar qualquer dispositivo de E/S sem especificar previamente qual (disquete, disco rígido ou CD-ROM)*
- Nomeação uniforme (pelo nome do caminho)
 - *Nome de um arquivo ou dispositivo pode ser uma cadeia de caracteres ou um número inteiro que é independente do dispositivo*
- Tratamento de erro
 - *Trata o mais próximo possível do hardware*

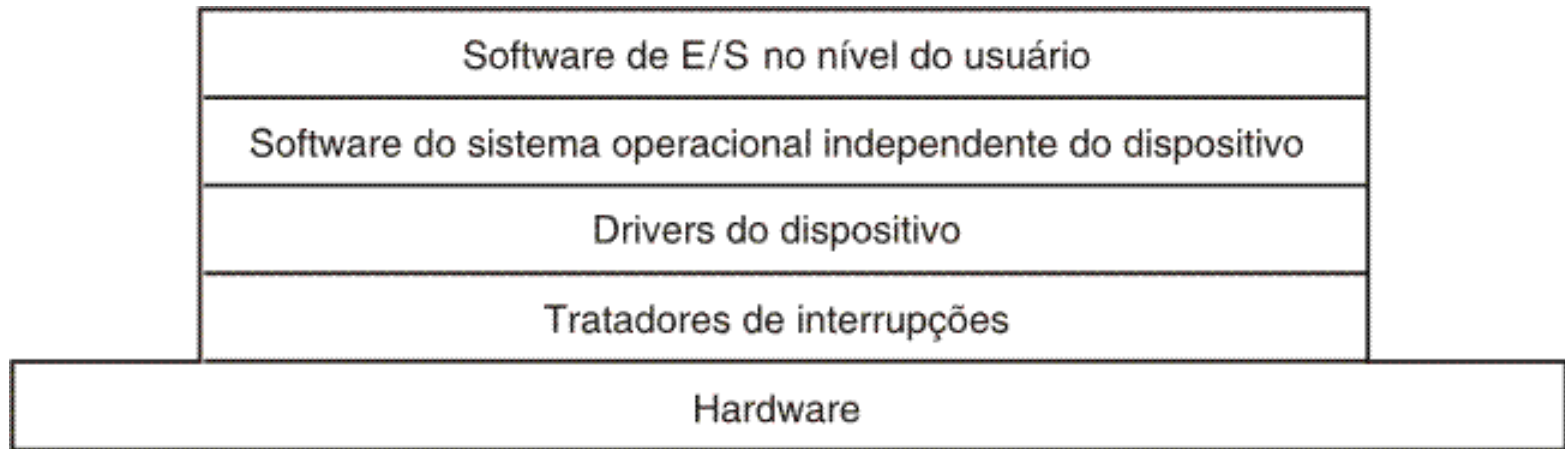
Objetivos do Software de E/S (2)

- Transferências **Síncronas** vs. **Assíncronas**
 - transferências bloqueantes (síncronas) vs. orientadas a interrupção (assíncronas):
 - *Para o desenvolvedor, operações bloqueantes ficam mais fáceis de serem tratadas. Todavia, de fato, a maioria são assíncronas mas o SO cria a ilusão que elas são síncronas.*
 - utilização de buffer para armazenamento temporário
 - dados provenientes de um dispositivo muitas vezes não podem ser armazenados diretamente em seu destino final (e.g., tráfego de rede, streaming de áudio/vídeo (taxa chegada vs reprodução))

Objetivos do Software de E/S (2)

- Dispositivos Compartilháveis vs. Dedicados
 - *discos são compartilháveis*
 - *unidades de fita não são*

Camadas do Software de E/S

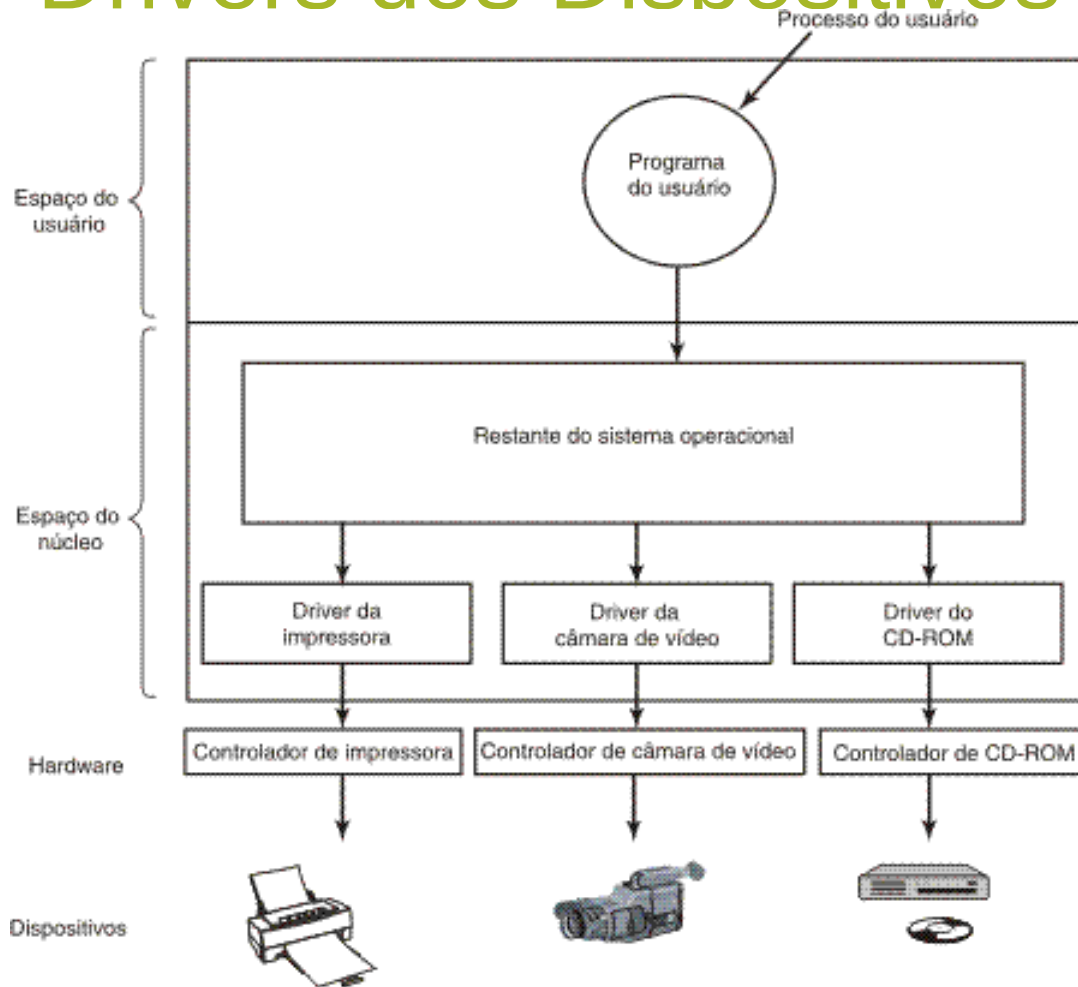


- Camadas do sistema de software de E/S

Tratadores de Interrupção

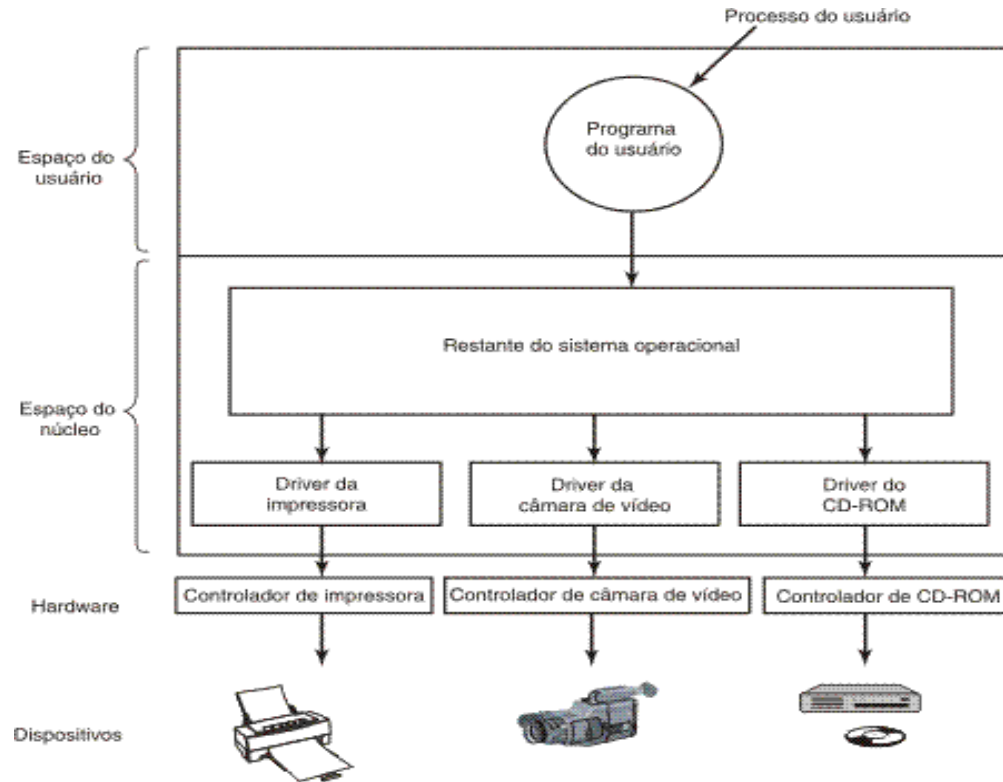
- As interrupções devem ser escondidas o máximo possível
 - *uma forma de fazer isso é bloqueando o driver que iniciou uma operação de E/S até que uma interrupção notifique que a E/S foi completada*
- Rotina de tratamento de interrupção cumpre sua tarefa
 - *e então desbloqueia o driver que a chamou*

Drivers dos Dispositivos



- Posição lógica dos drivers dos dispositivos
- A comunicação entre os drivers e os controladores de dispositivos é feita por meio do barramento

Drivers dos Dispositivos



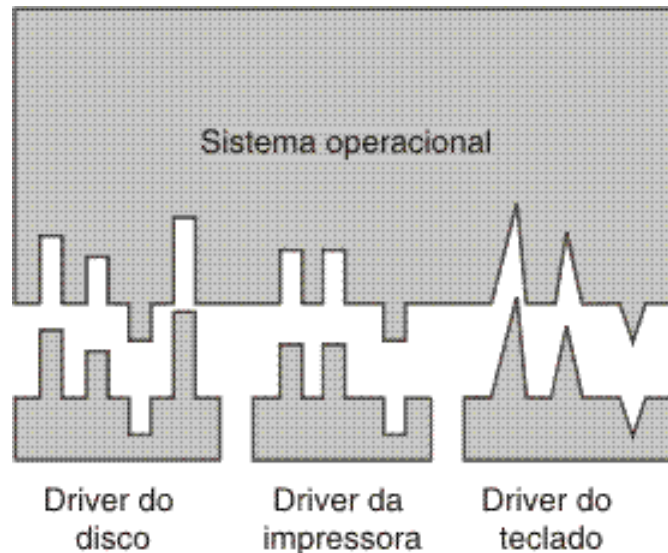
- Dispositivos de bloco e de caractere
- Drivers têm de ser **reentrantes**: pode ser chamado uma segunda vez antes que a primeira chamada tenha sido concluída.

Software de E/S Independente de Dispositivo (1)

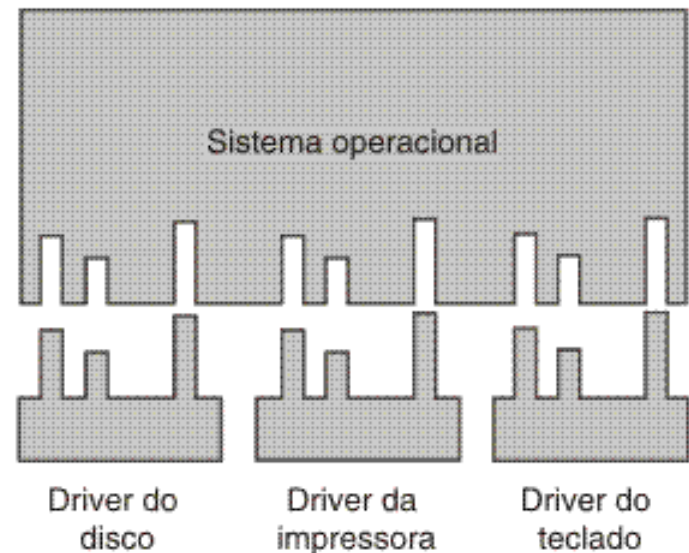
- Interface uniforme para os drivers dos dispositivos
- Armazenamento em buffer
- Relatório dos erros
- Alocação e liberação de dispositivos dedicados
- Fornecimento de tamanho de bloco independente de dispositivo
- **Funções do software de E/S independente de dispositivo**

• ‘

Software de E/S Independente de Dispositivo (2)



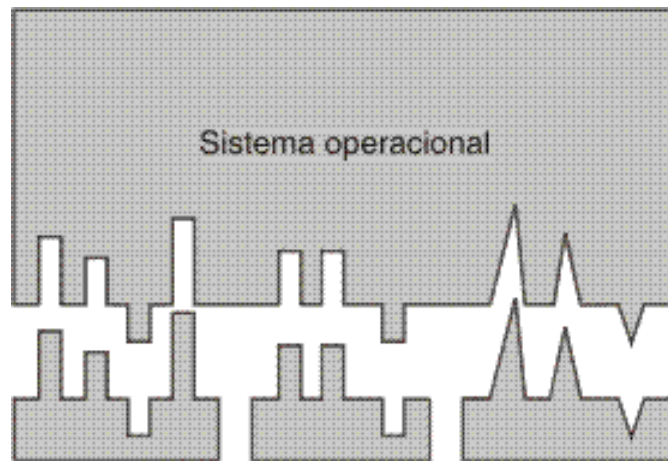
(a)



(b)

- (a) Sem uma interface-padrão do driver: cada driver de dispositivo apresenta uma interface diferente para o S.O. (i.e., as chamadas que o SO pode fazer a cada driver difere de driver para driver).

Software de E/S Independente de Dispositivo (2)

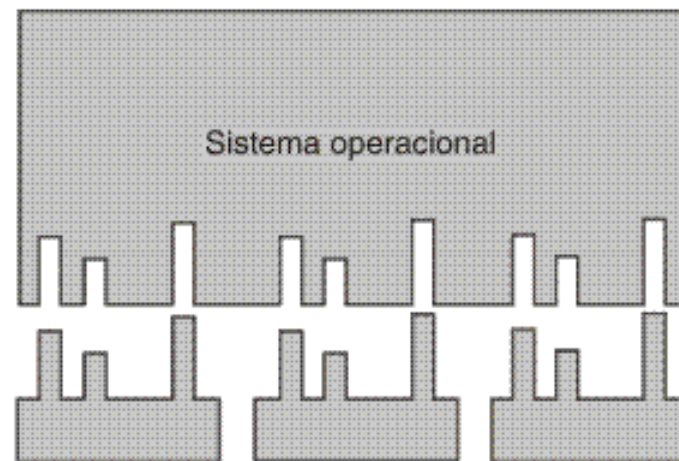


Driver do disco

Driver da impressora

Driver do teclado

(a)



Driver do disco

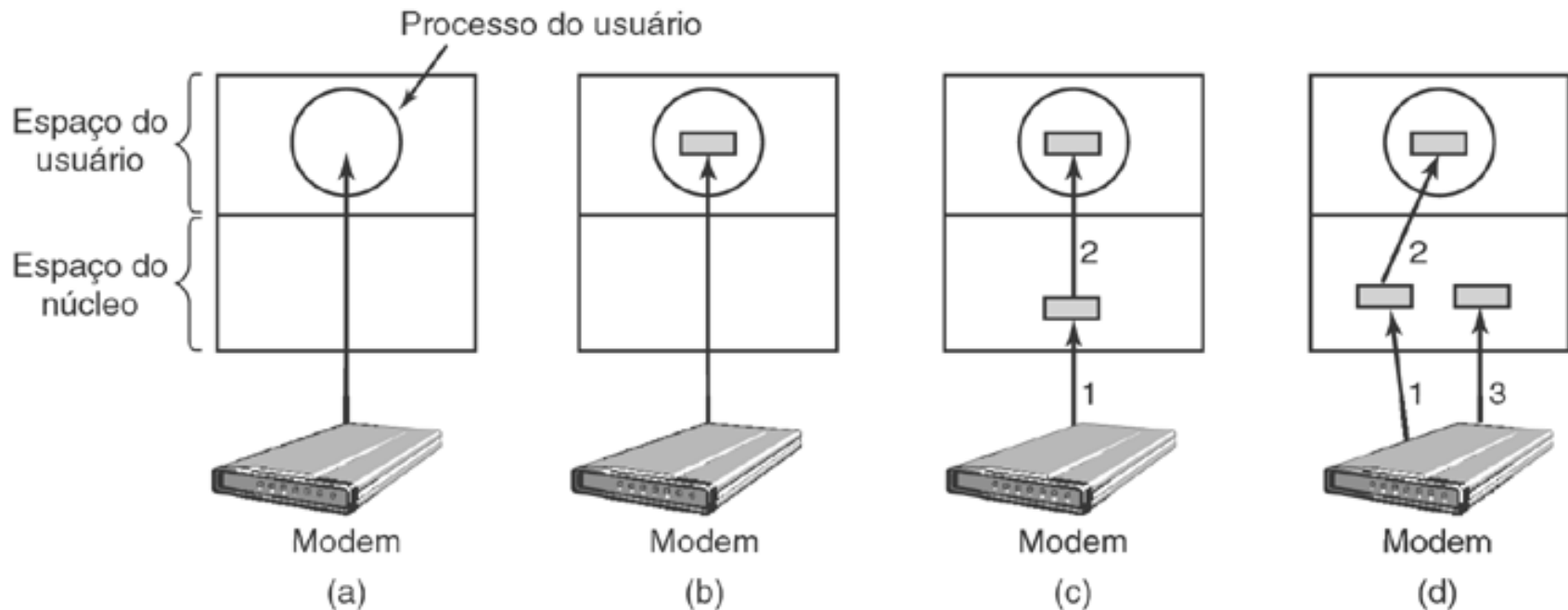
Driver da impressora

Driver do teclado

(b)

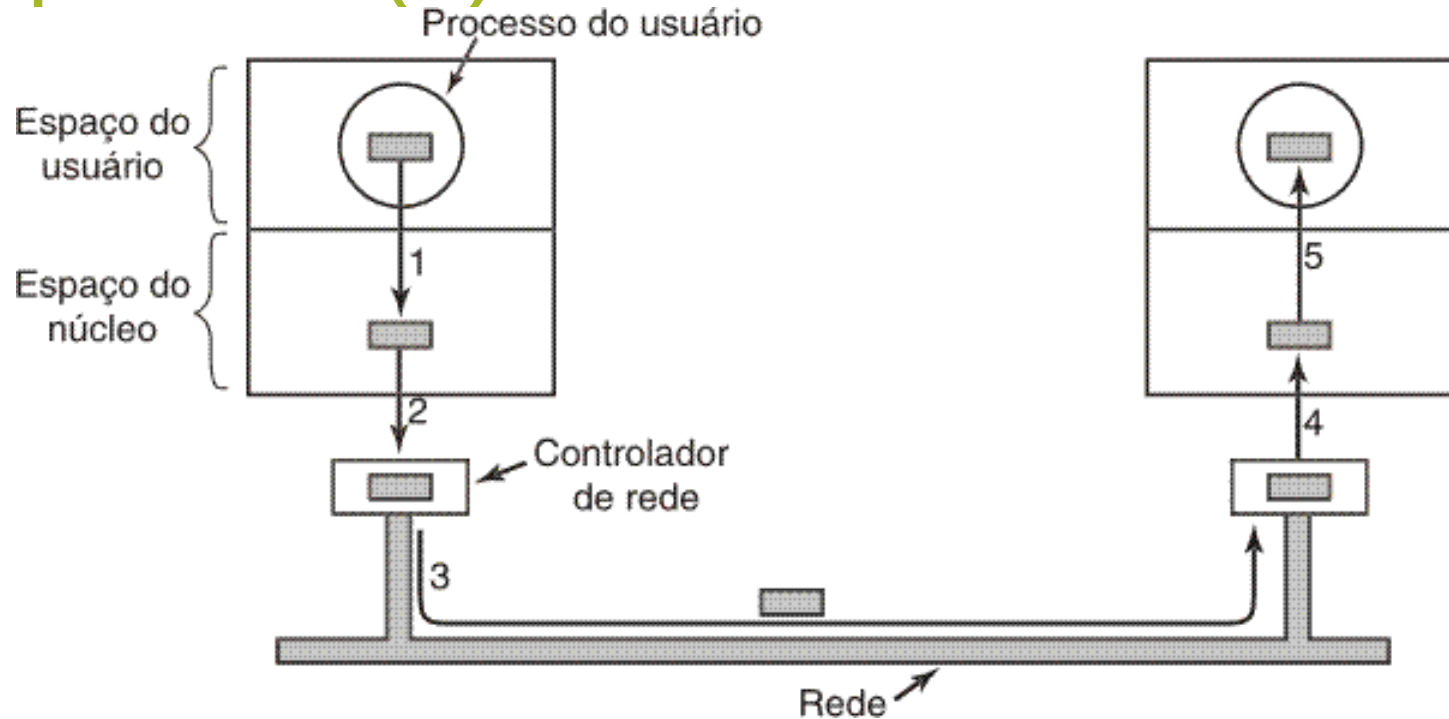
- (b) Com uma interface-padrão do driver: acoplar um novo driver fica fácil, desde que ele atenda a interface padrão.

Software de E/S Independente de Dispositivo (3): uso de *buffers*



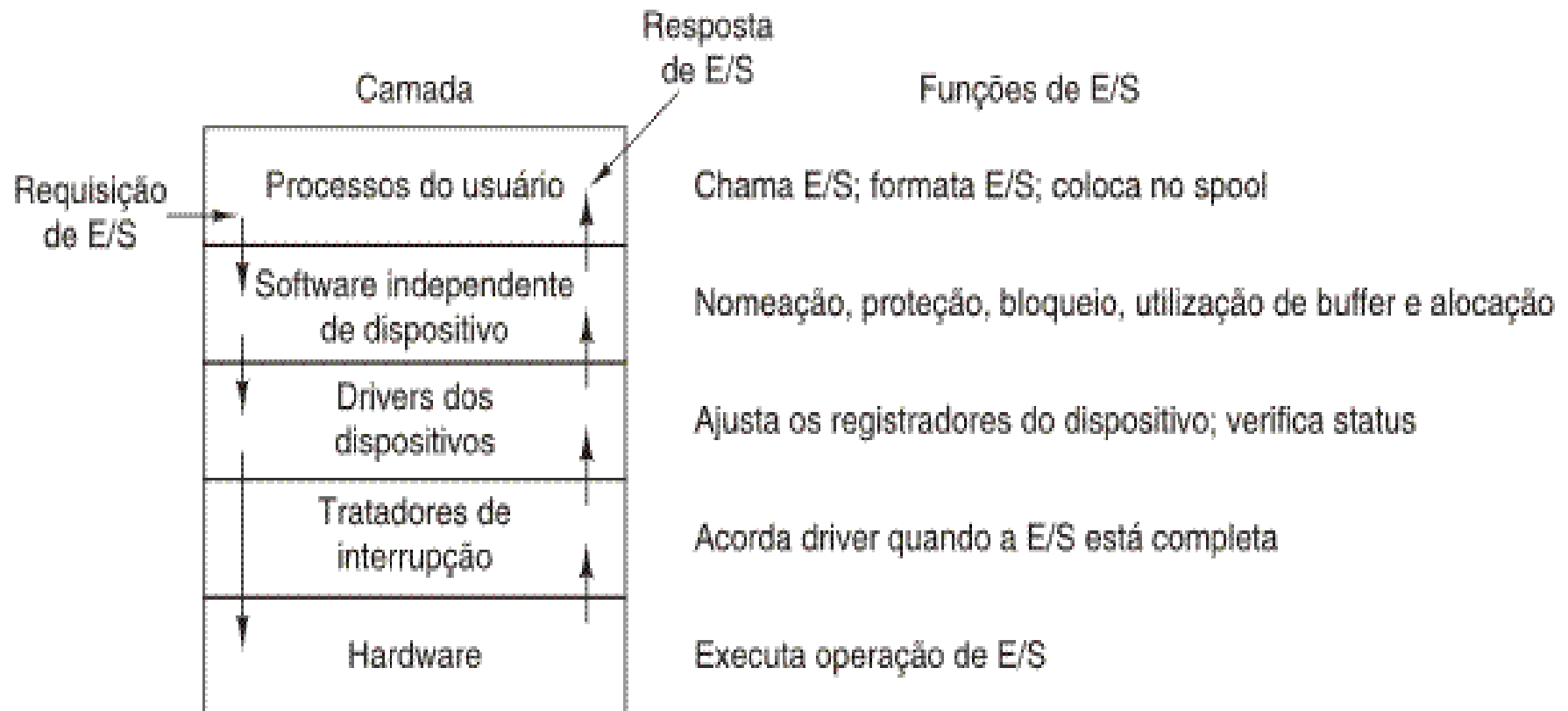
- a) Entrada sem utilização de buffer
- b) Utilização de buffer no espaço do usuário
- c) Utilização de buffer no núcleo seguido de cópia para o espaço do usuário
- d) Utilização de buffer duplo no núcleo

Software de E/S Independente de Dispositivo (4)



- A operação em rede pode envolver muitas cópias de um pacote: controlador de rede precisa de seu próprio buffer para evitar que uma leitura direta do buffer fique fora da cadência de transmissão devido a interrupções durante leitura.

Software de E/S no Espaço do Usuário



- Camadas do sistema de E/S e as principais funções de cada camada

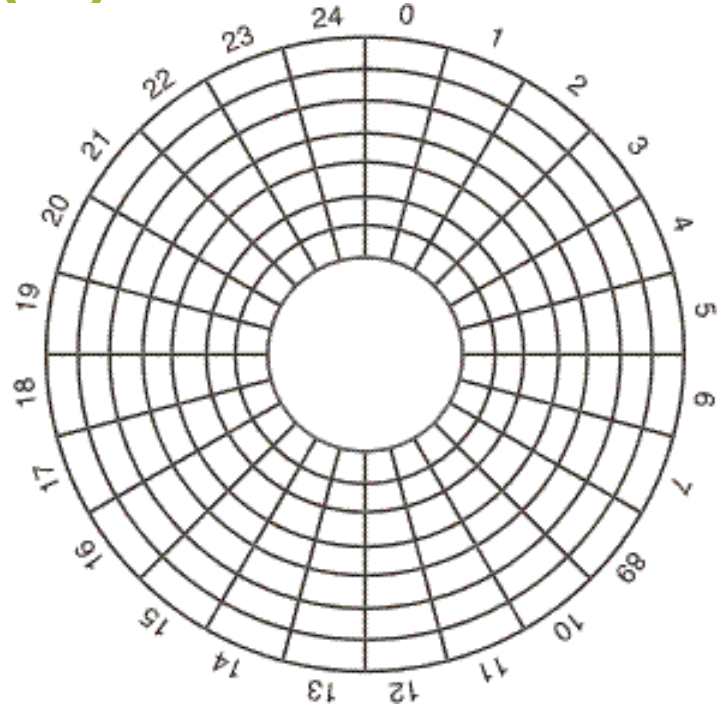
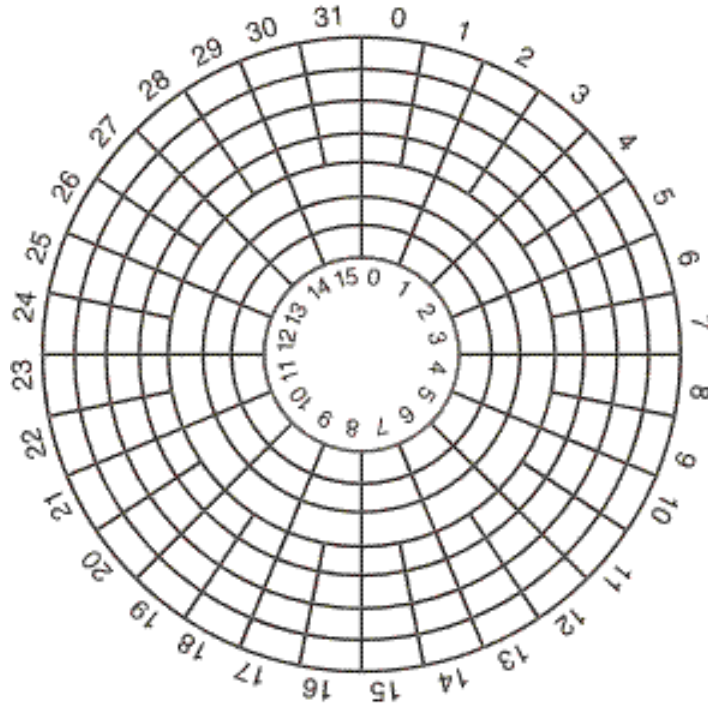
Discos

Hardware do Disco (1)

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

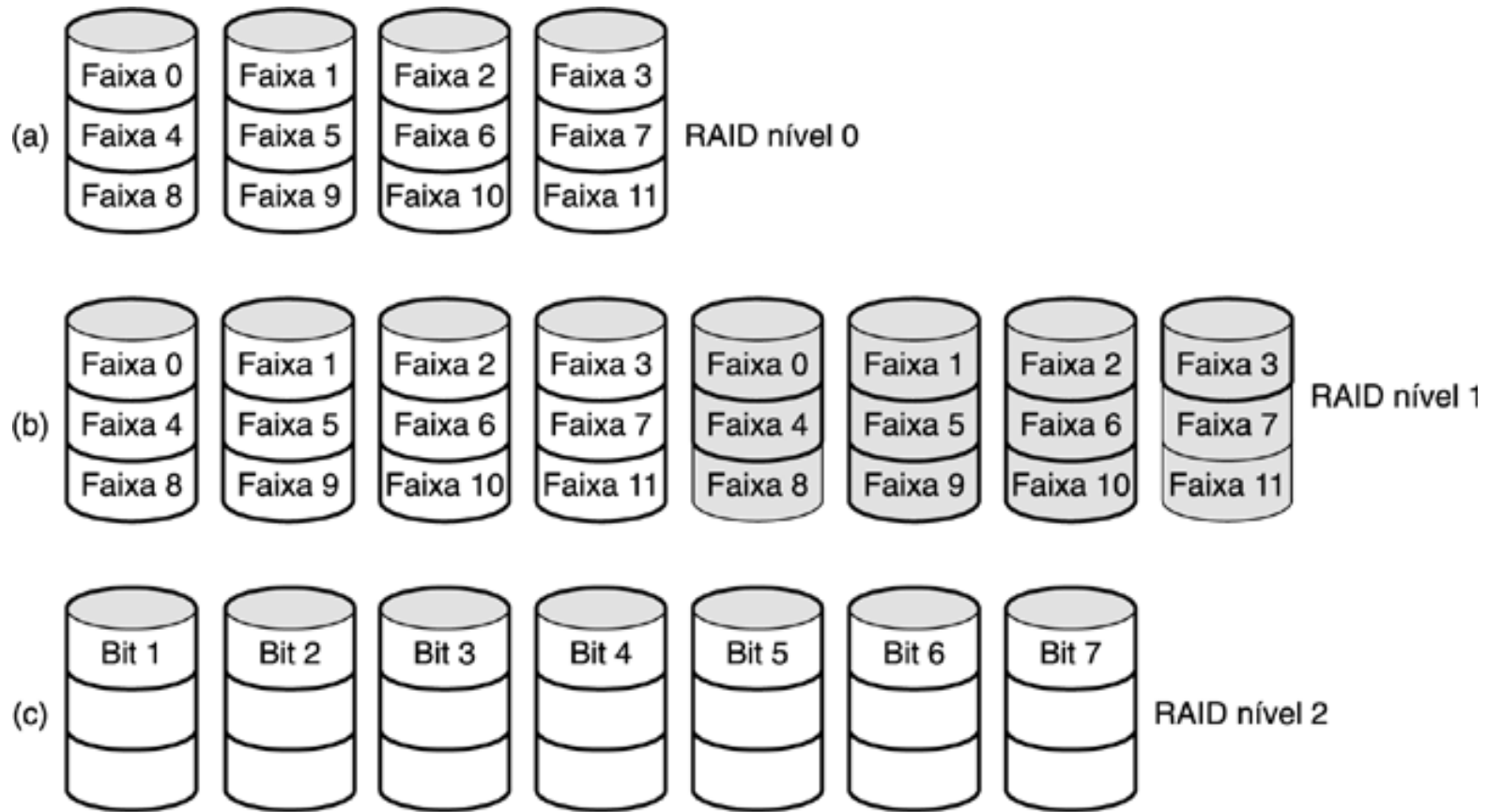
- Parâmetros de disco para o disco flexível original do IBM PC e o disco rígido da Western Digital WD 18300

Hardware do Disco (2)



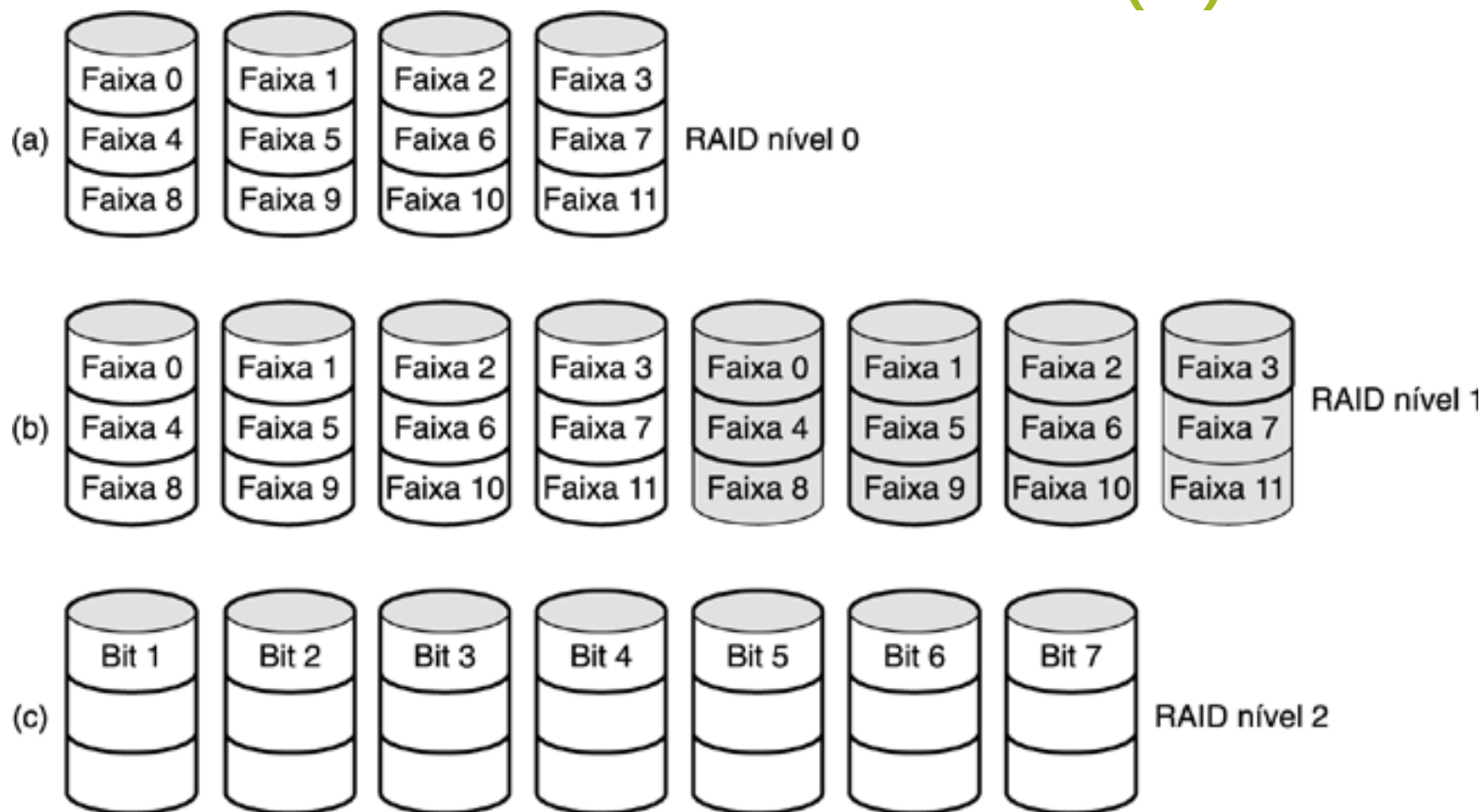
- Geometria física (fig. Esq.) de um disco com duas zonas (uma zona com 16 setores/trilha, outra com 32 setores/trilha).
- Uma possível geometria virtual para esse disco
- Com logical block addressing (LBA), setores são enumerados consecutivamente (iniciando em zero), sem considerar a geometria do disco.

Hardware do Disco (3)



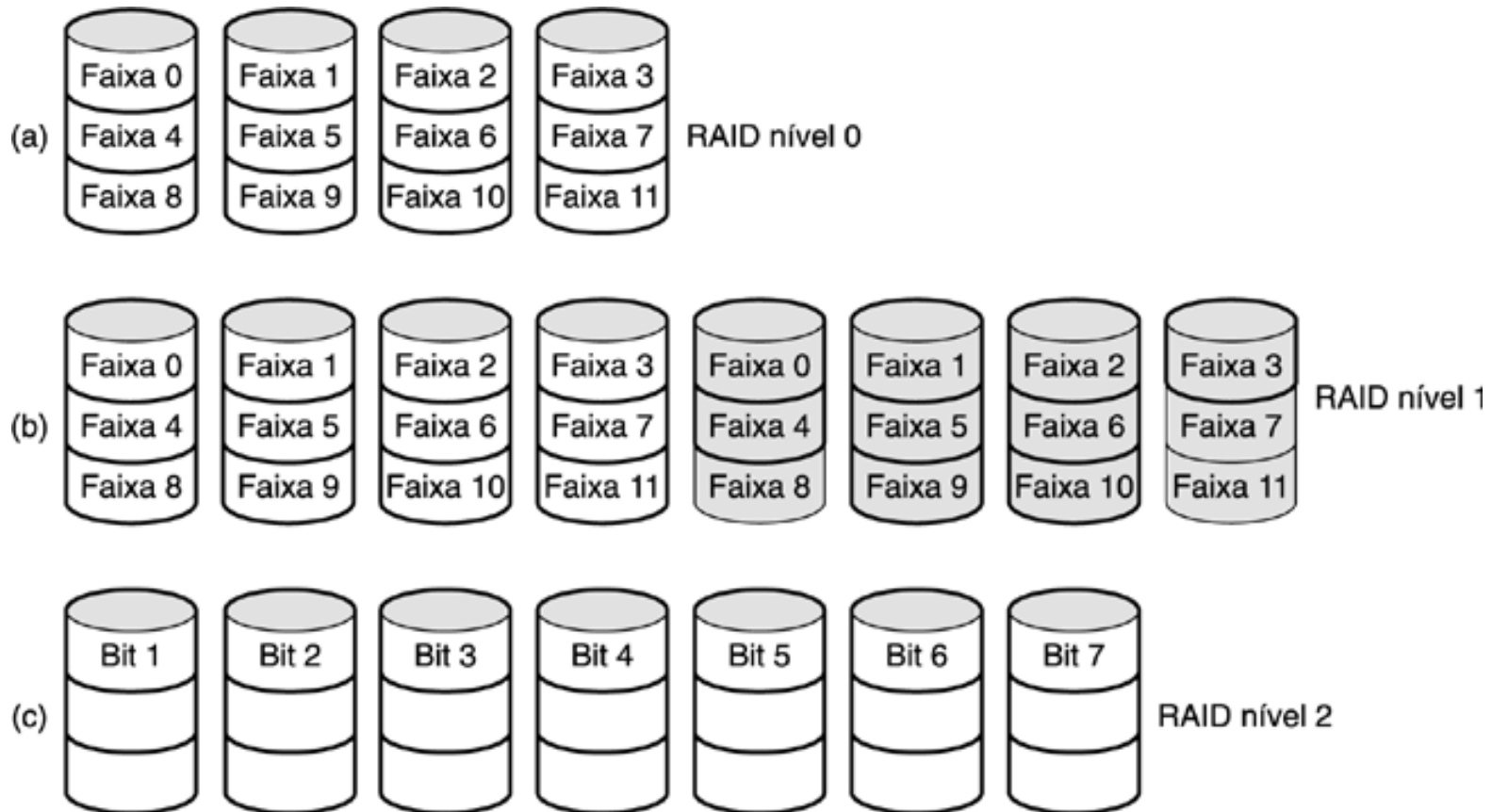
- RAID (redundant array of inexpensive disks): arranjo redundante de discos baratos.

Hardware do Disco (3)



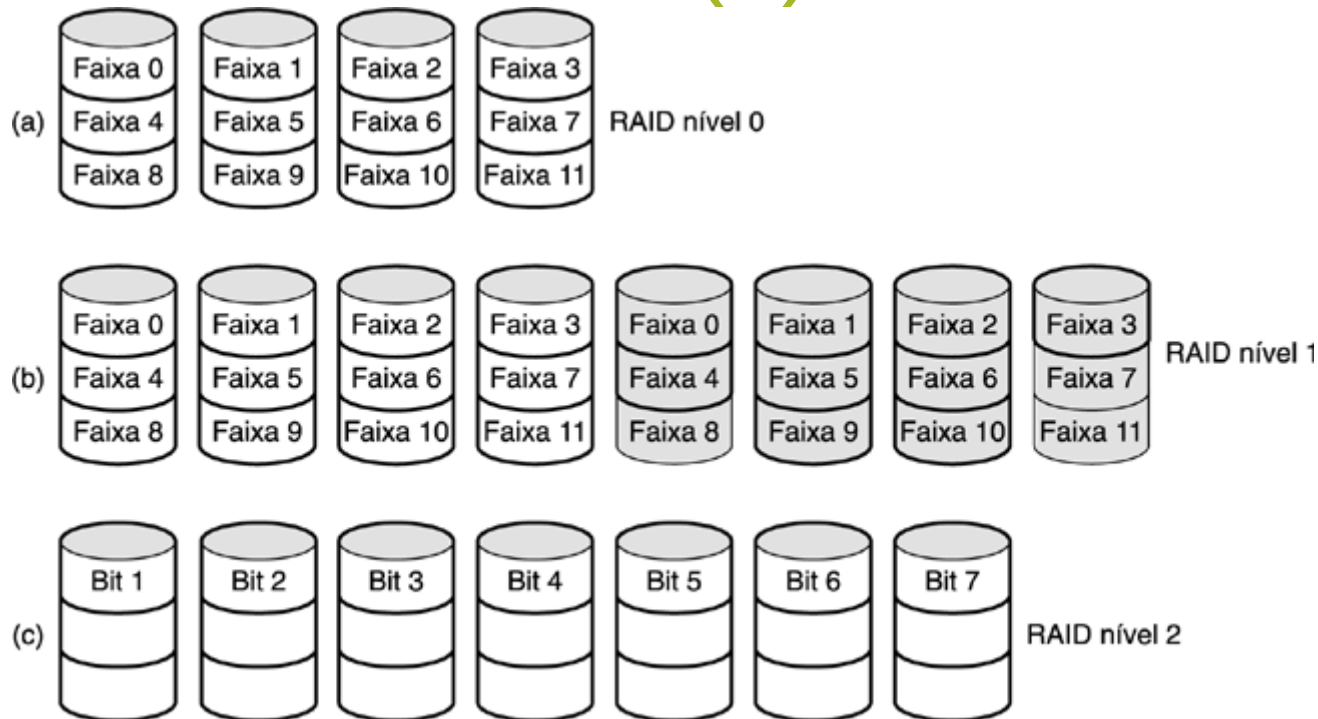
- RAID 0 (striping): cada faixa pode ser um múltiplo de setores; explora paralelismo no acesso (mais rápido); no entanto, **não é de fato um RAID, pois não há redundância.**

Hardware do Disco (3)



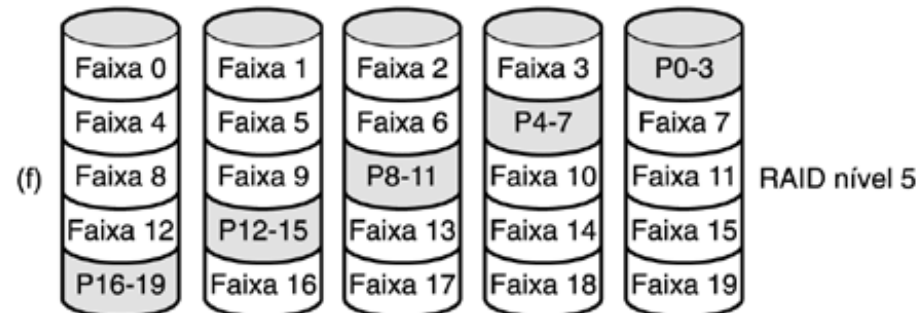
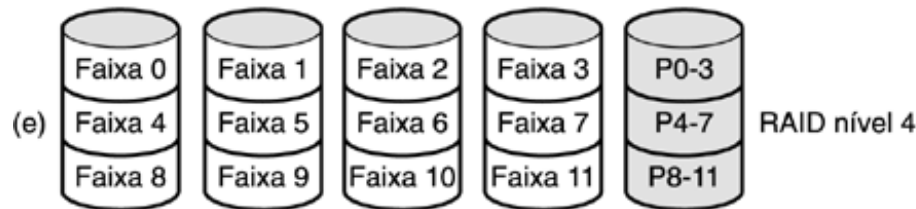
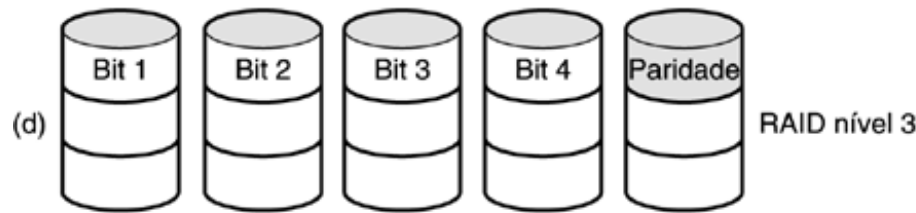
- RAID 1 (b): Cada disco é duplicado; apresenta redundância e melhora no desempenho de acesso.

Hardware do Disco (3)



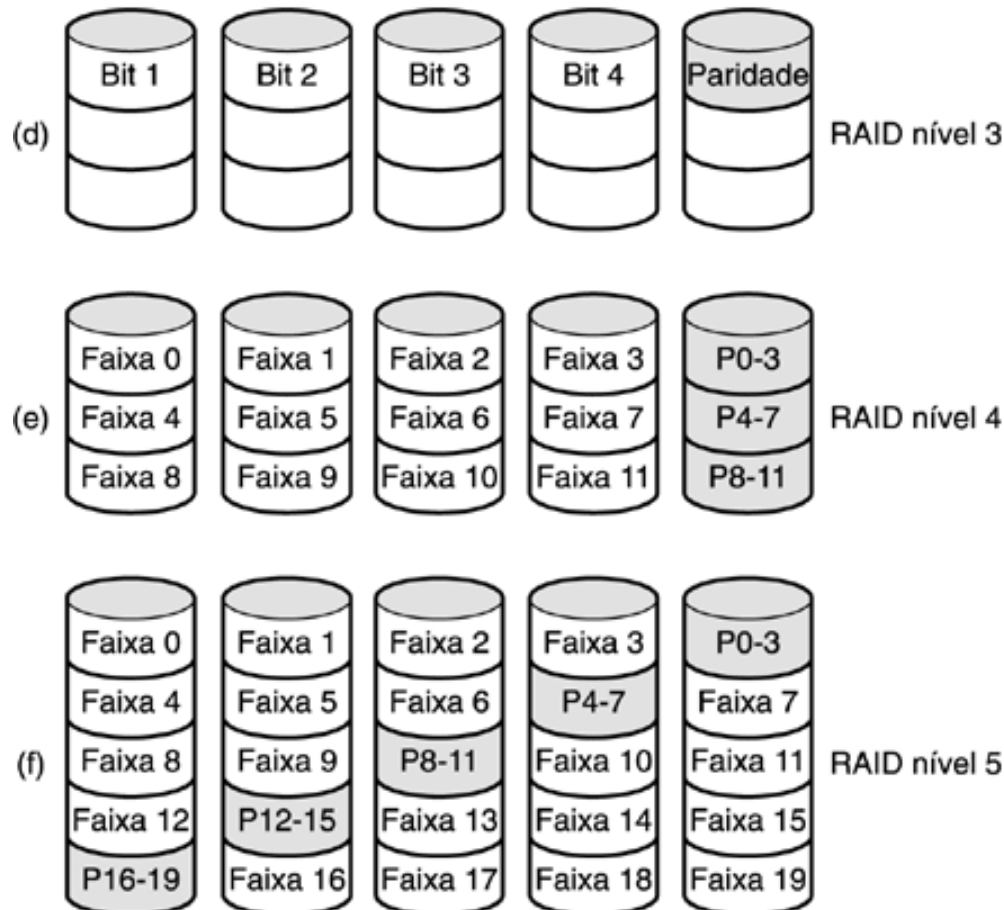
- Bits de paridade (e.g., 4 bits dados + 3 bits código de Hamming): falha de um disco pode ser facilmente recuperada
- RAID 2 (c): palavras (bytes ou nibbles) são quebradas, bit-a-bit, e gravados separadamente em discos distintos
- Discos devem estar SINCRONIZADOS!!

Hardware do Disco (4)



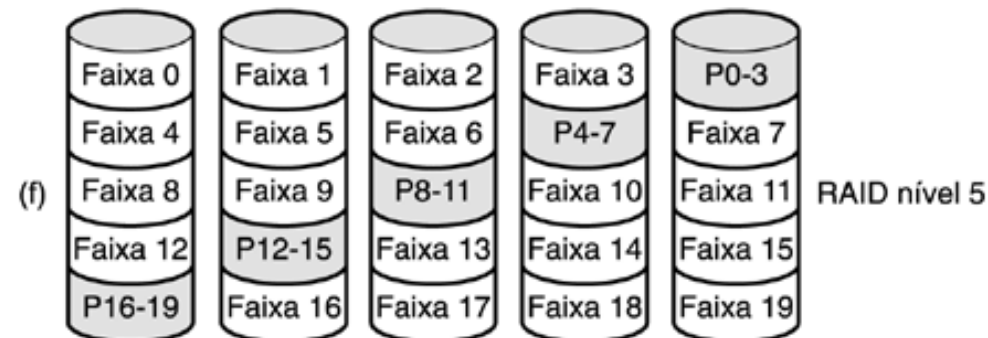
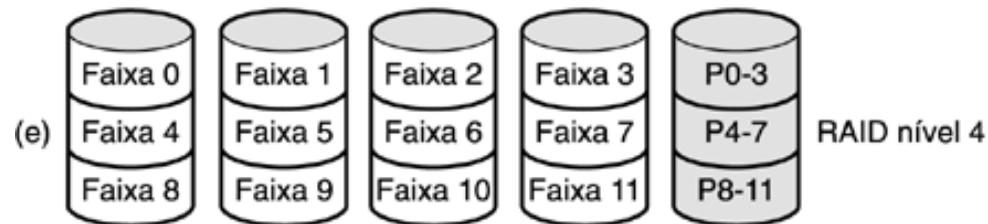
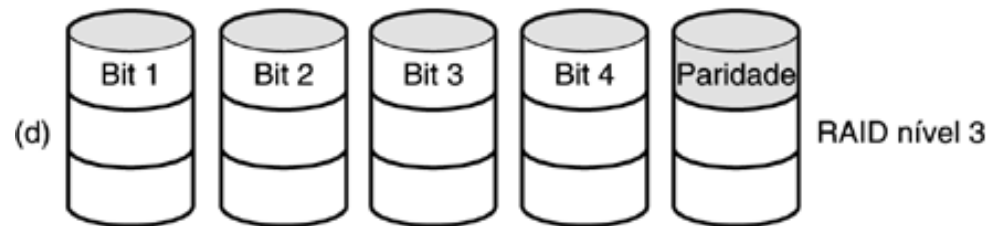
- RAID 3 (d): versão simplificada do RAID 2 com apenas um bit de paridade; **permite a recuperação no caso de falha de UM disco**, mas não em caso de erros aleatórios; discos também precisam estar sincronizados

Hardware do Disco (4)



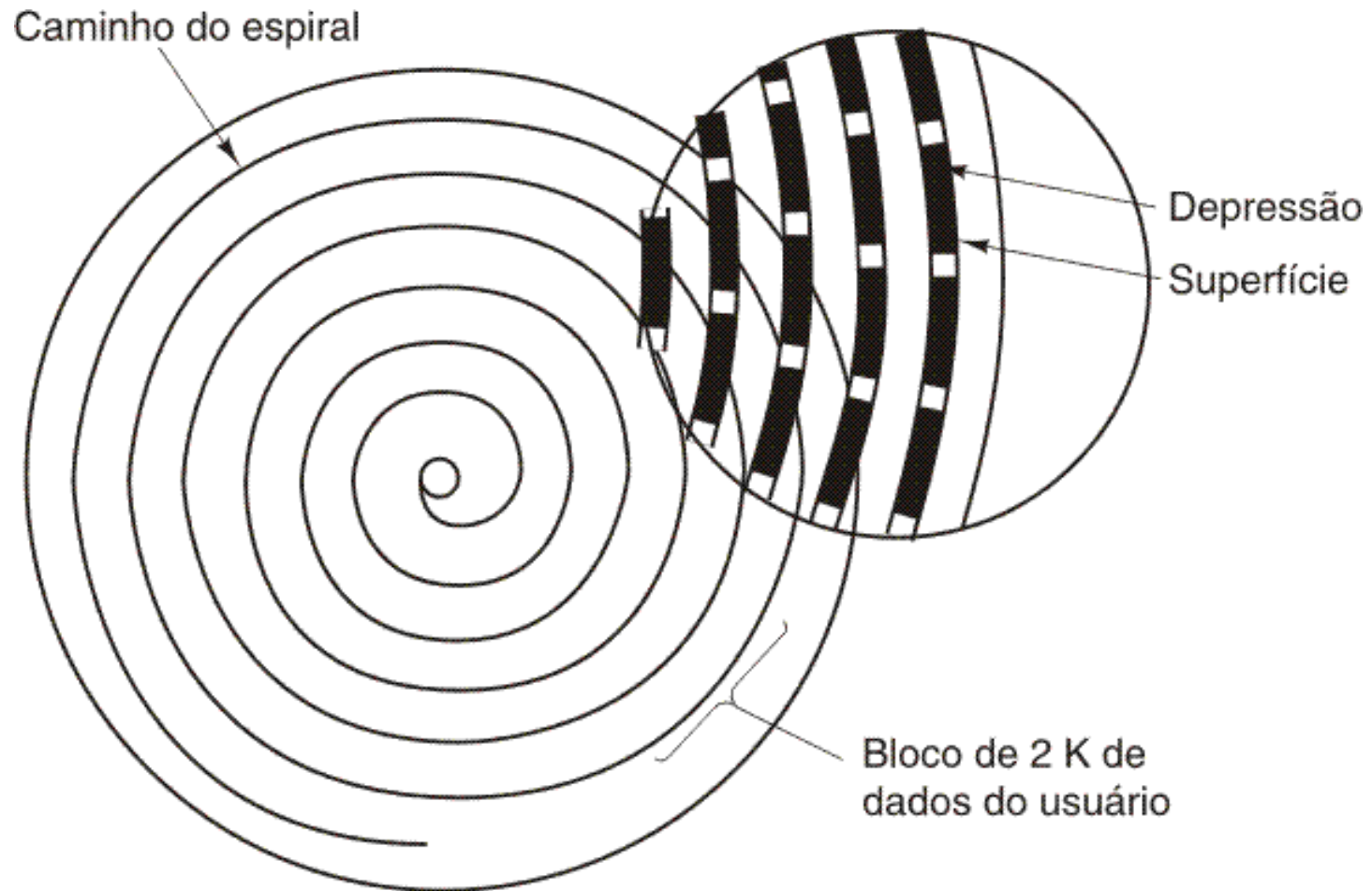
- RAID 4 (e): similar ao RAID 0 (faixas=setores) mas com disco de paridade; alterações em uma das faixas requer reler todas as demais faixas para recalcular paridade.

Hardware do Disco (4)



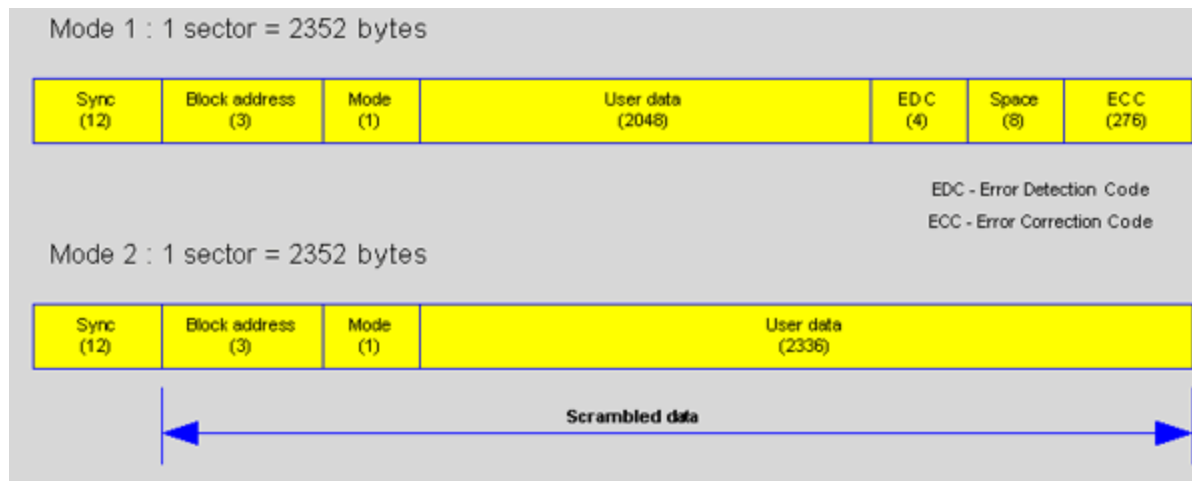
- RAID 5 (f): no RAID 4, o disco de paridade pode se tornar um gargalo; para contornar esse problema, distribui-se o conteúdo de paridade por todos os discos; recuperação é complexa.

Hardware do Disco (5)



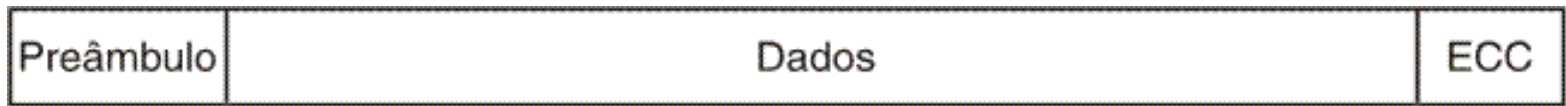
- Estrutura de gravação de um CD

Hardware do Disco (6)



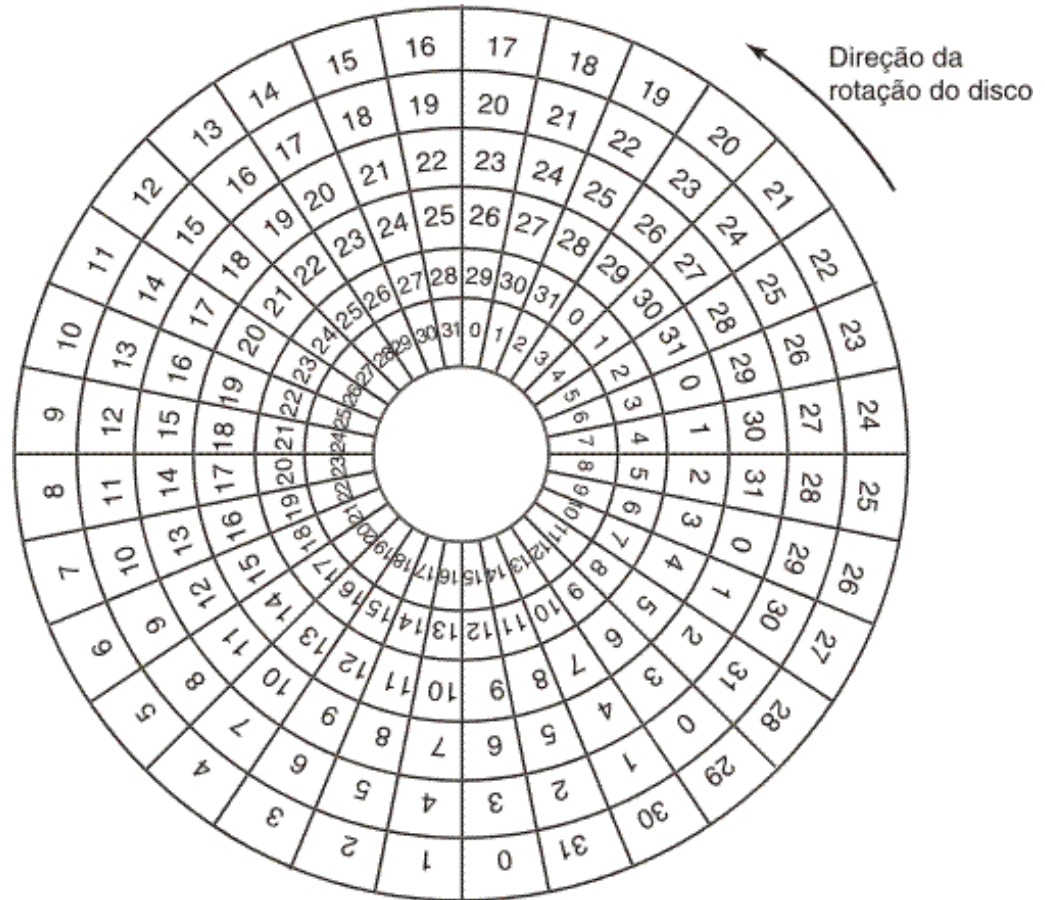
- Formato de um setor em um CD-ROM

Formatação de Disco (1)



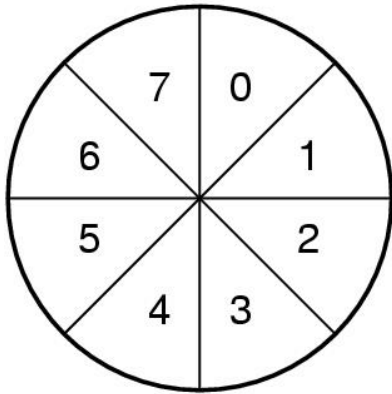
- Um setor do disco: o preâmbulo contém um padrão binário para reconhecer o início do setor, bem como o número do cilindro e do próprio setor.

Formatação de Disco (2)

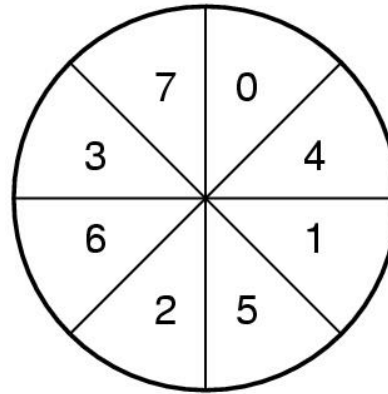


Uma ilustração da torção cilíndrica: permite leitura consecutiva de uma trilha para outra

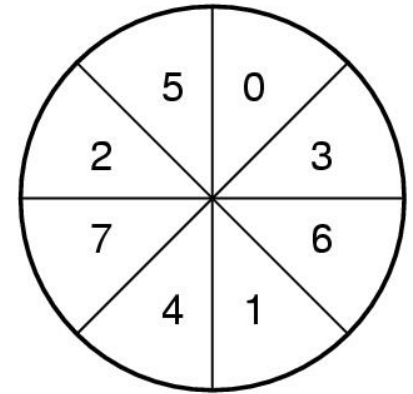
Formatação de Disco (3)



(a)



(b)



(c)

a) Sem entrelaçamento

b) Entrelaçamento simples: fornece ao controlador um certo “descanso” entre setores consecutivos, permitindo tempo para a cópia do conteúdo no buffer do disco para a memória principal

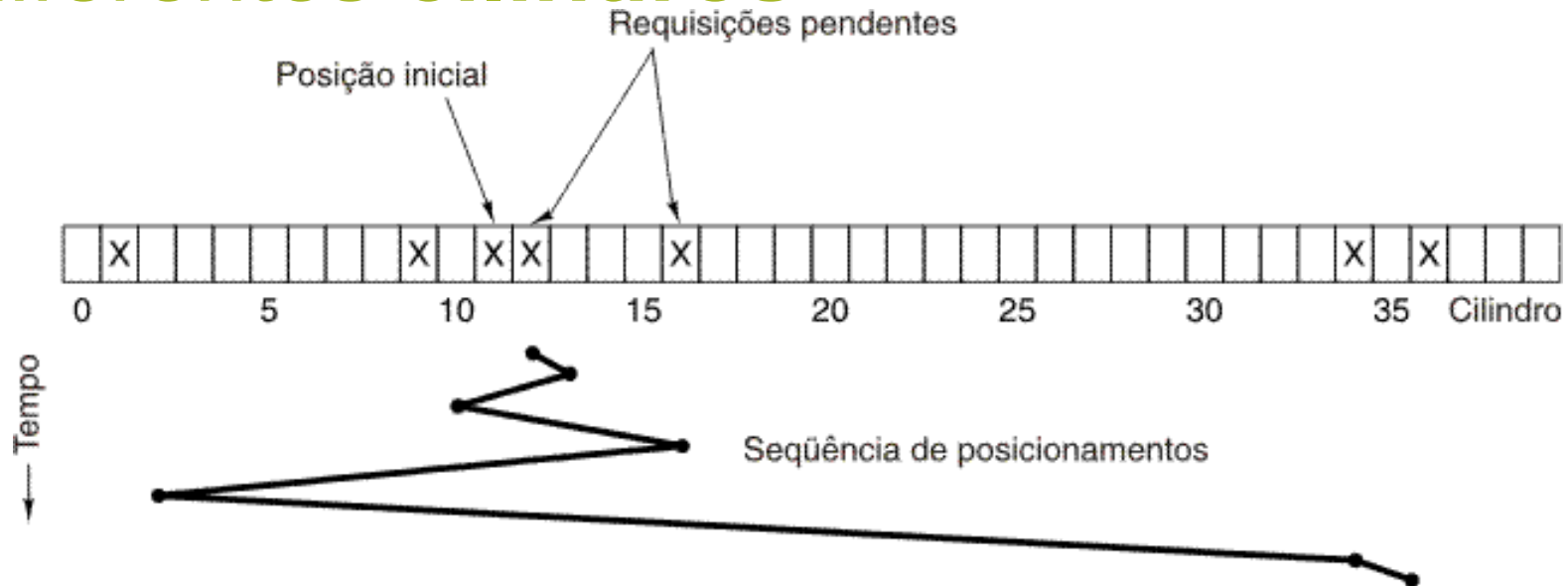
c) Entrelaçamento duplo: caso o processo de cópia seja muito lento

d) Copiar toda uma trilha para o buffer evita utilizar entrelaçamento

Algoritmos de Escalonamento de Braço de Disco (1)

- Tempo necessário para ler ou escrever um bloco de disco é determinado por 3 fatores
 1. tempo de posicionamento (*seek time*)
 2. atraso de rotação
 3. tempo de transferência do dado
- Tempo de posicionamento domina
- Checagem de erro é feita por controladores

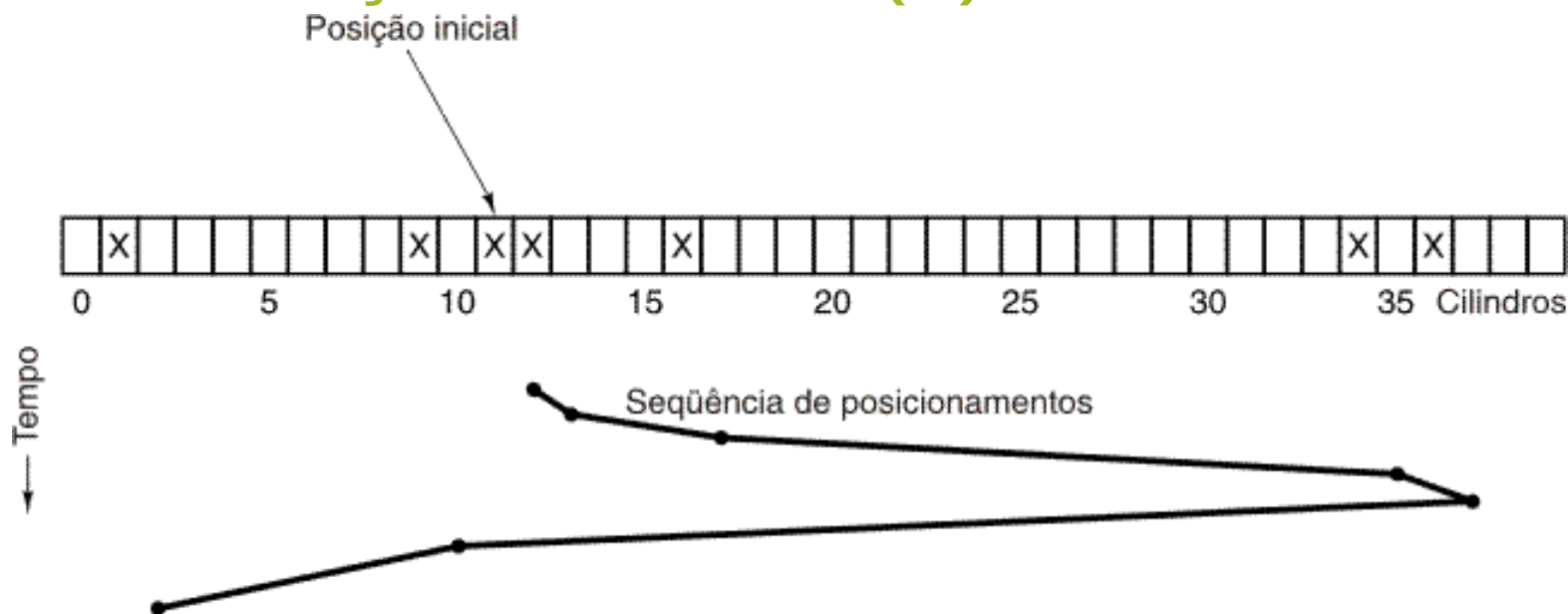
Algoritmos de Escalonamento de Braço de Disco (2): requisições em diferentes cilindros



Algoritmo de escalonamento de disco **Posicionamento**

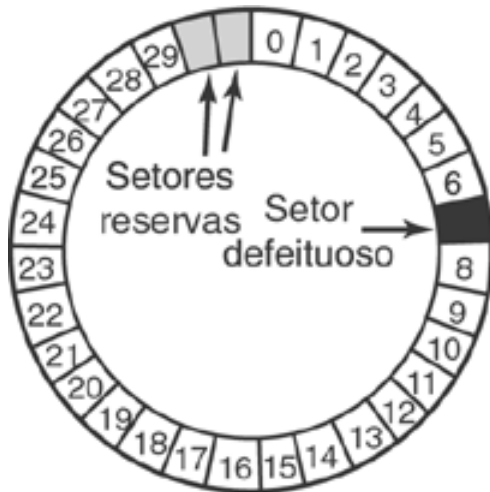
- **Mais Curto Primeiro** (Shortest Seek First): atende primeiro a requisição mais próxima da posição (cilindro) atual. Nesse exemplo a sequência seria: 12, 9, 16, 1, 34 e 36

Algoritmos de Escalonamento de Braço de Disco (3)

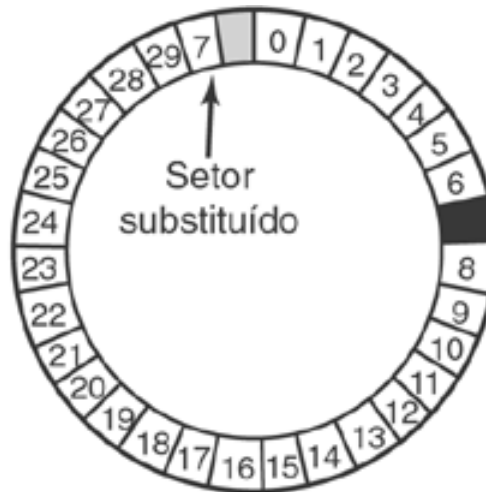


- O **algoritmo do elevador** para o escalonamento das requisições do disco: atende primeiro todas as requisições pendentes na direção atual (“sobe” ou “desce”) de movimento

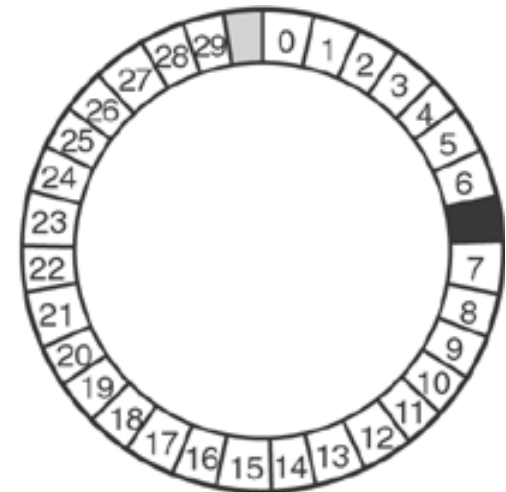
Tratamento de Erro



(a)



(b)



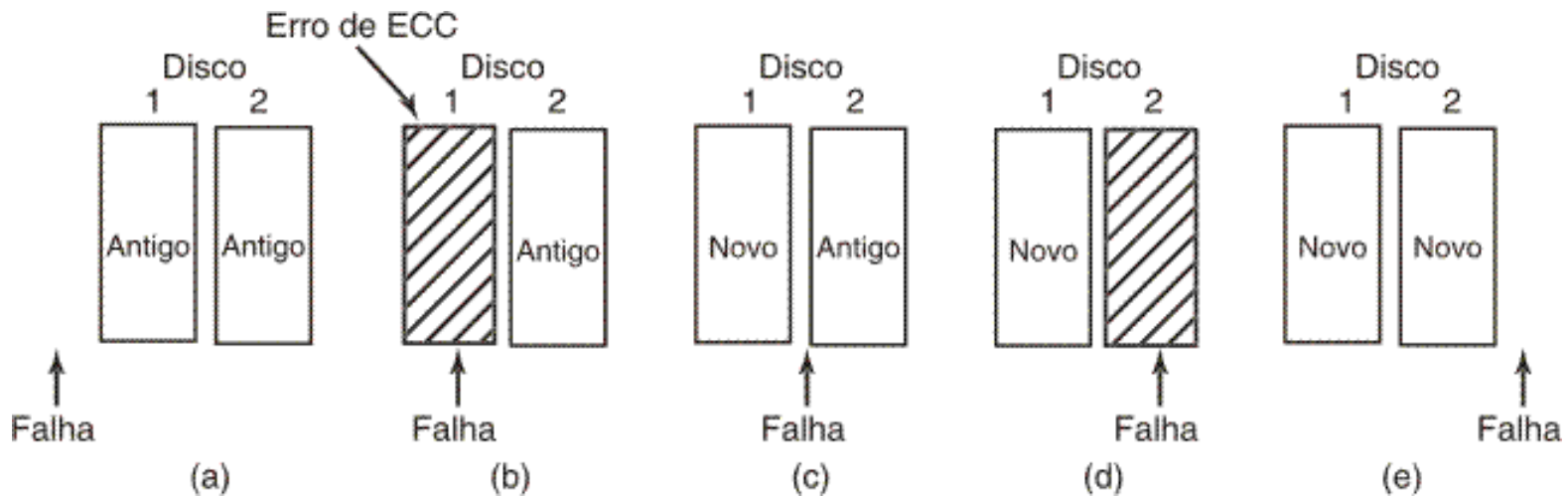
(c)

- a) Uma trilha de disco com um setor defeituoso
- b) Substituindo um setor reserva por um setor defeituoso
- c) Deslocando todos os setores para pular o setor defeituoso

Armazenamento estável: utiliza dois discos

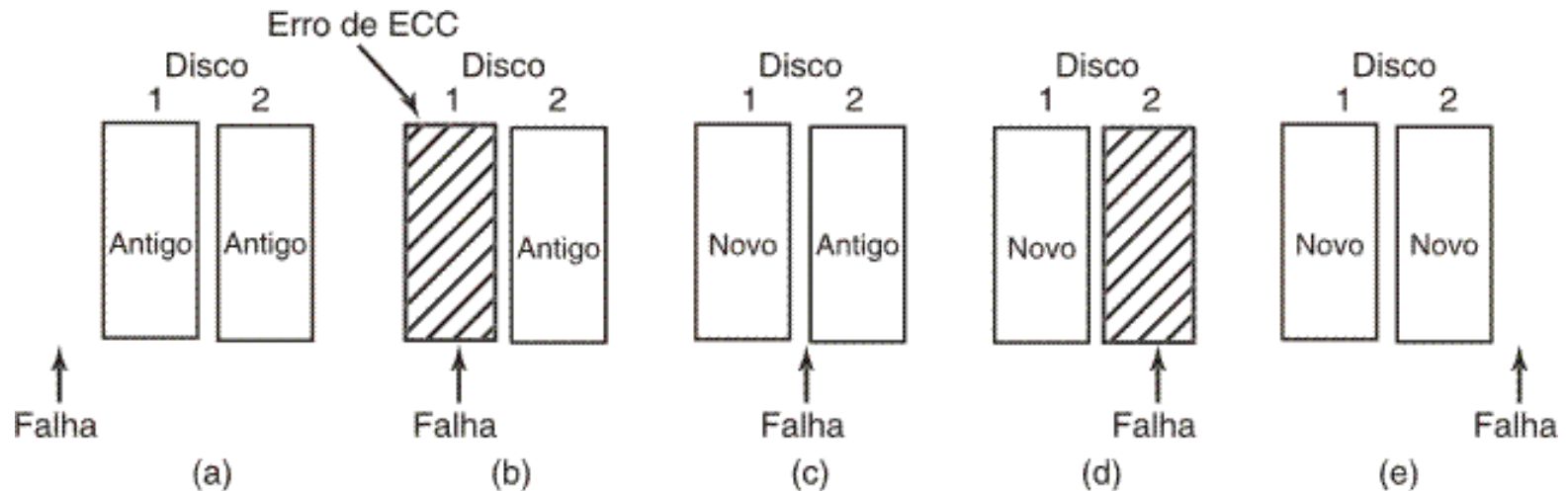
- Operação de escrita estável (**protege contra a inserção de dados corrompidos - RAID não protege contra isso**): ao escrever em um setor, pode ocorrer erros devido a uma falha da CPU ou do disco. Utilizando-se técnicas de ECC pode-se identificar a maioria dos erros. Sistemas RAID não identificam dados corrompidos durante a escrita (veja, mesmo que tenha bits de paridade, estes serão computados sim mas com o conteúdo com erro!!!)

Armazenamento estável: utiliza dois discos



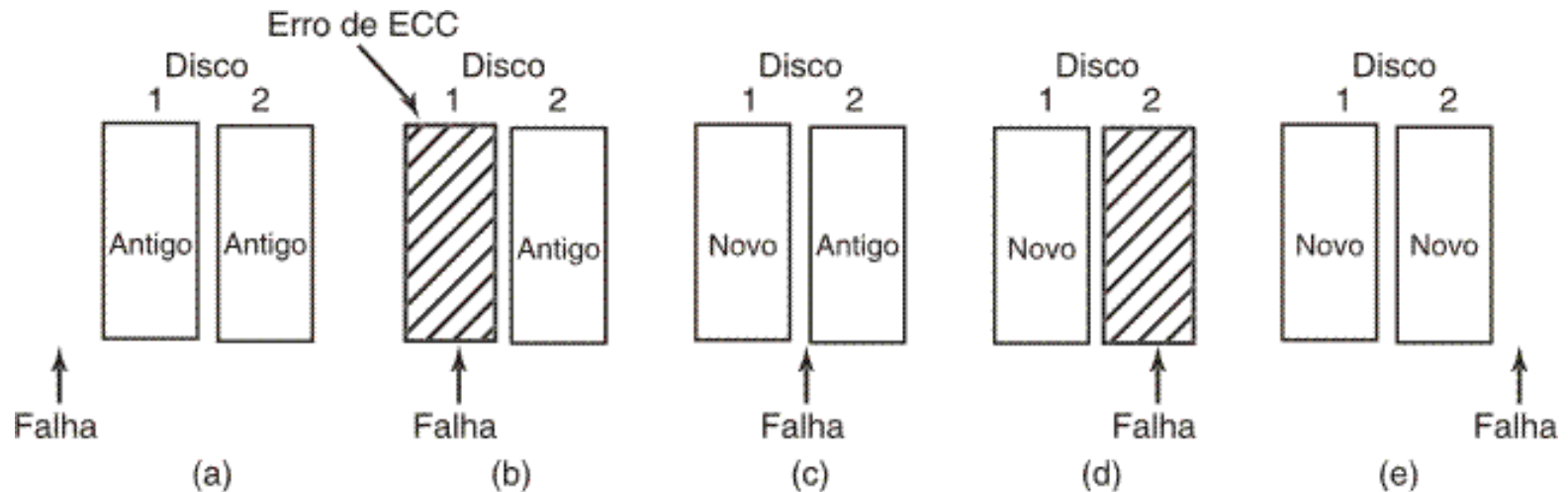
- Operação de escrita estável: **ou o disco escreve corretamente o dado ou não escreve nada, deixando os dados existentes intactos.**
- A seguir, análise de cada um dos possíveis cenários de falhas.

Armazenamento estável: utiliza dois discos



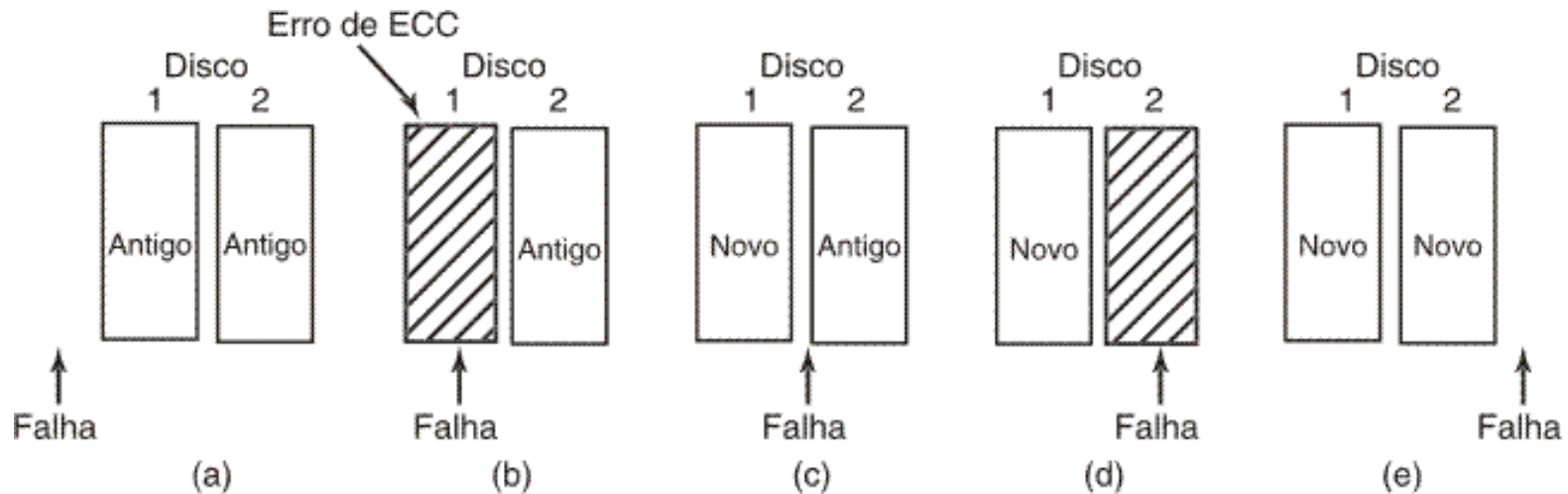
- Análise da influência das falhas nas escritas estáveis: (b) falha da CPU ocorre durante escrita no disco 1: cópia antiga, estável, é recuperada para o disco 1 via cópia correta no disco 2.

Armazenamento estável: utiliza dois discos



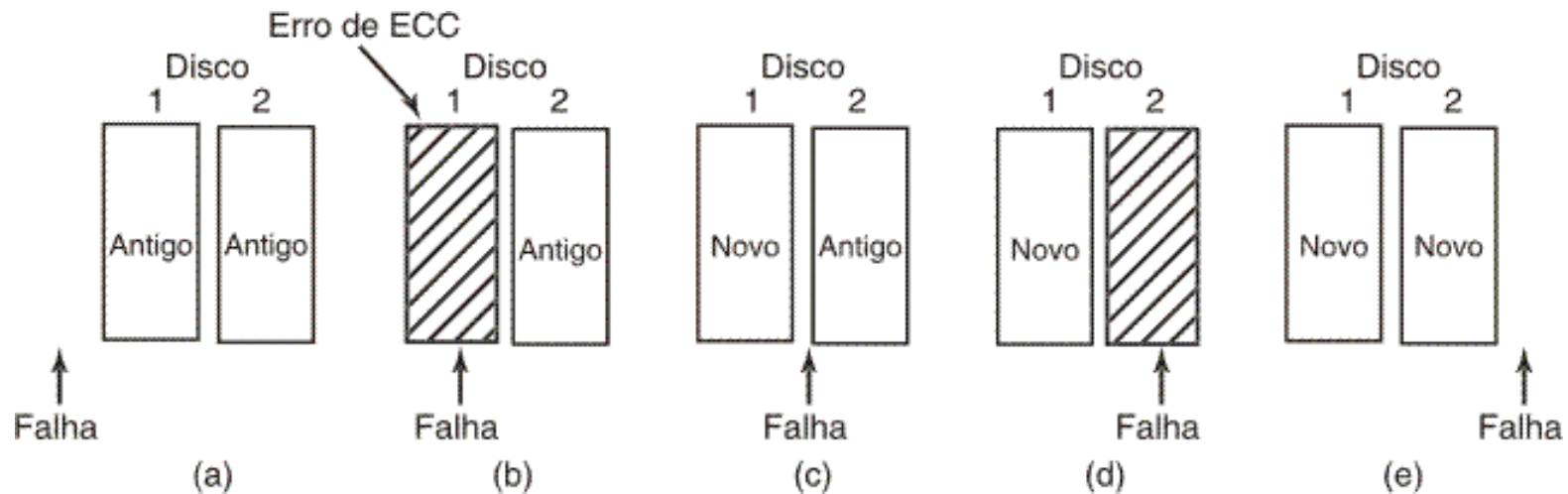
- Análise da influência das falhas nas escritas estáveis: (c) **falha da cpu** ocorre após escrita na unidade 1: não há a necessidade de desfazer a primeira escrita (nova e correta), bastando copiar da unidade 1 para a unidade 2.

Armazenamento estável: utiliza dois discos



- Análise da influência das falhas nas escritas estáveis: (d) falha da cpu ocorre durante a escrita na unidade 2: não há a necessidade de desfazer a primeira escrita (nova e correta), bastando copiar da unidade 1 para a unidade 2.

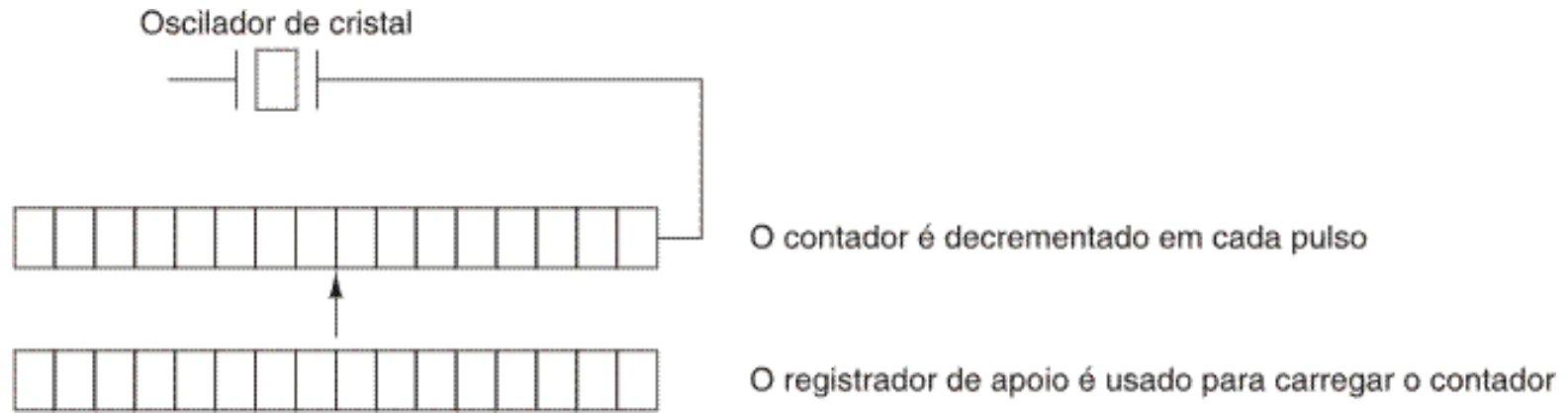
Armazenamento estável: utiliza dois discos



Análise da influência das falhas nas escritas estáveis: (e) falha da cpu acontece após ambas as escritas terem ocorrido com sucesso: nesse caso o programa de recuperação percebe que ambas as cópias são iguais e nada precisa ser feito.

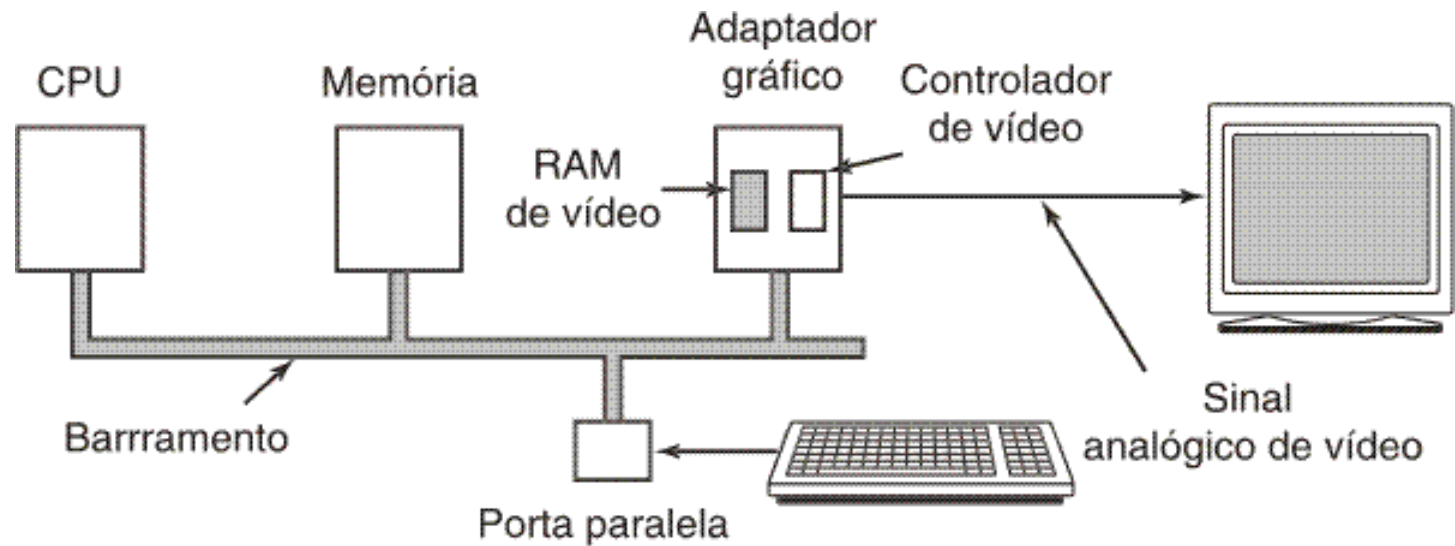
Relógios

Hardware do Relógio



- Um relógio programável: define os tiques de relógio; o valor do contador é ajustado para corresponder ao tempo de interrupção desejado.

Hardware de Vídeo



- Vídeos mapeados na memória
- driver escreve diretamente na RAM de vídeo do monitor

Software de Entrada

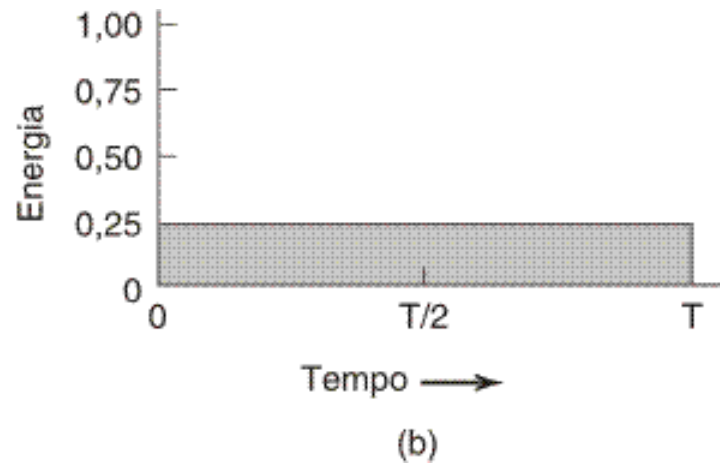
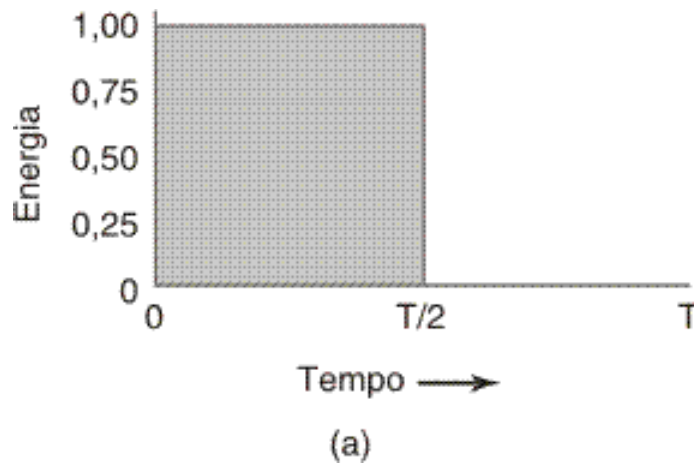
- Driver de teclado entrega um número
 - *driver* converte para caracteres (usa uma tabela ASCII)
- Exceções, adaptações necessárias para outras linguagens
 - muitos SOs fornecem mapas de teclas ou páginas de códigos carregáveis

Gerenciamento de Energia (1)

Dispositivo	Li et al. (1994)	Lorch e Smith (1998)
Monitor de vídeo	68%	39%
CPU	12%	18%
Disco rígido	20%	12%
Modem		6%
Som		2%
Memória	0,5%	1%
Outros		22%

Consumo de energia de várias partes de um computador notebook

Gerenciamento de Energia (2)



- Execução em velocidade máxima do relógio
- Cortando a voltagem pela metade
 - corta a velocidade do relógio também pela metade,
 - consumo de energia cai para 4 vezes menos

Gerenciamento de Energia (3)

- Dizer aos programas para usar menos energia
 - pode significar experiências mais pobres para o usuário
- Exemplos
 - muda de saída colorida para preto e branco
 - reconhecimento de fala reduz vocabulário
 - menos resolução ou detalhe em uma imagem