

Dispositivos de Entrada e Saída

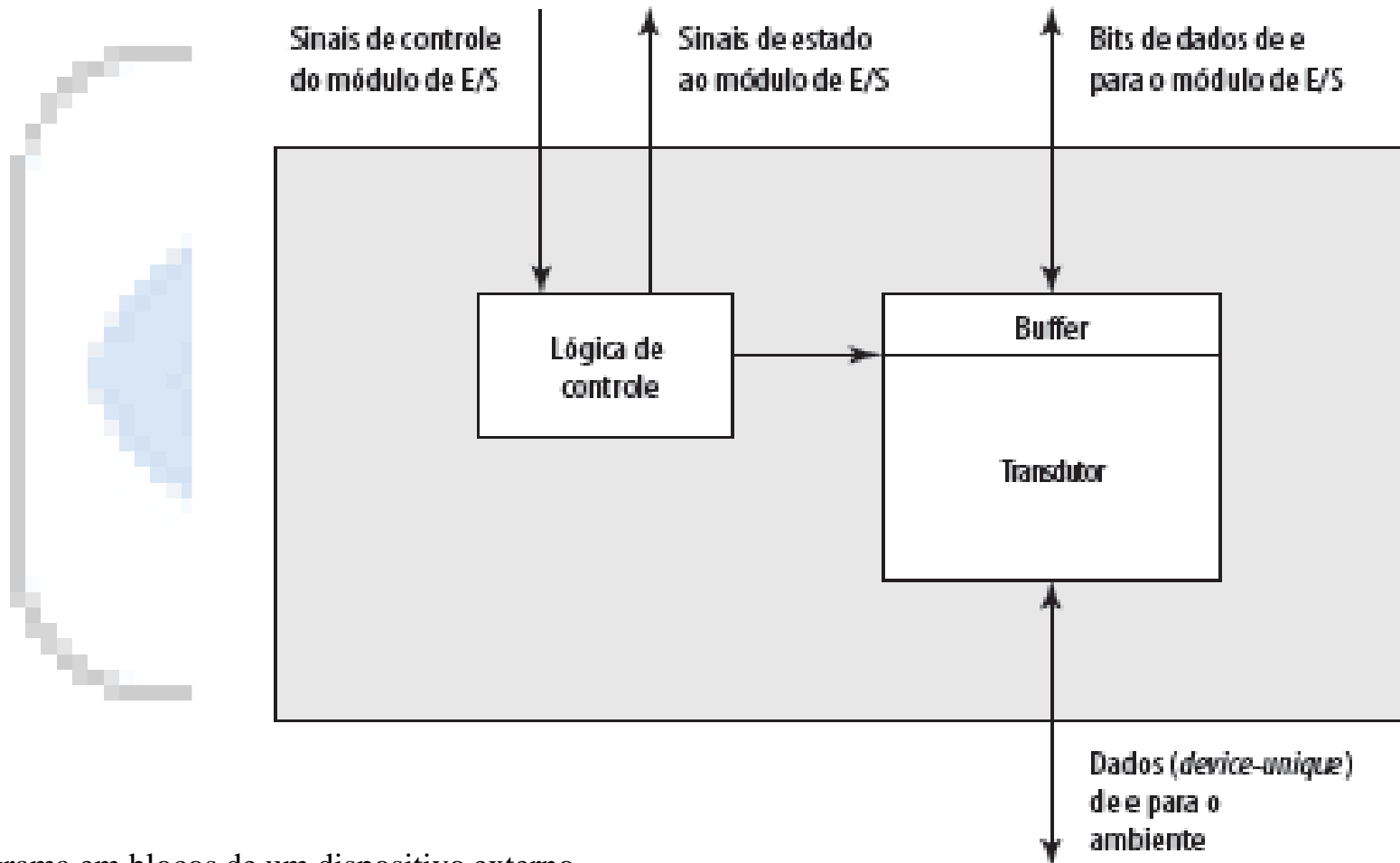


Figura 2: Diagrama em blocos de um dispositivo externo

Dispositivos de Entrada e Saída

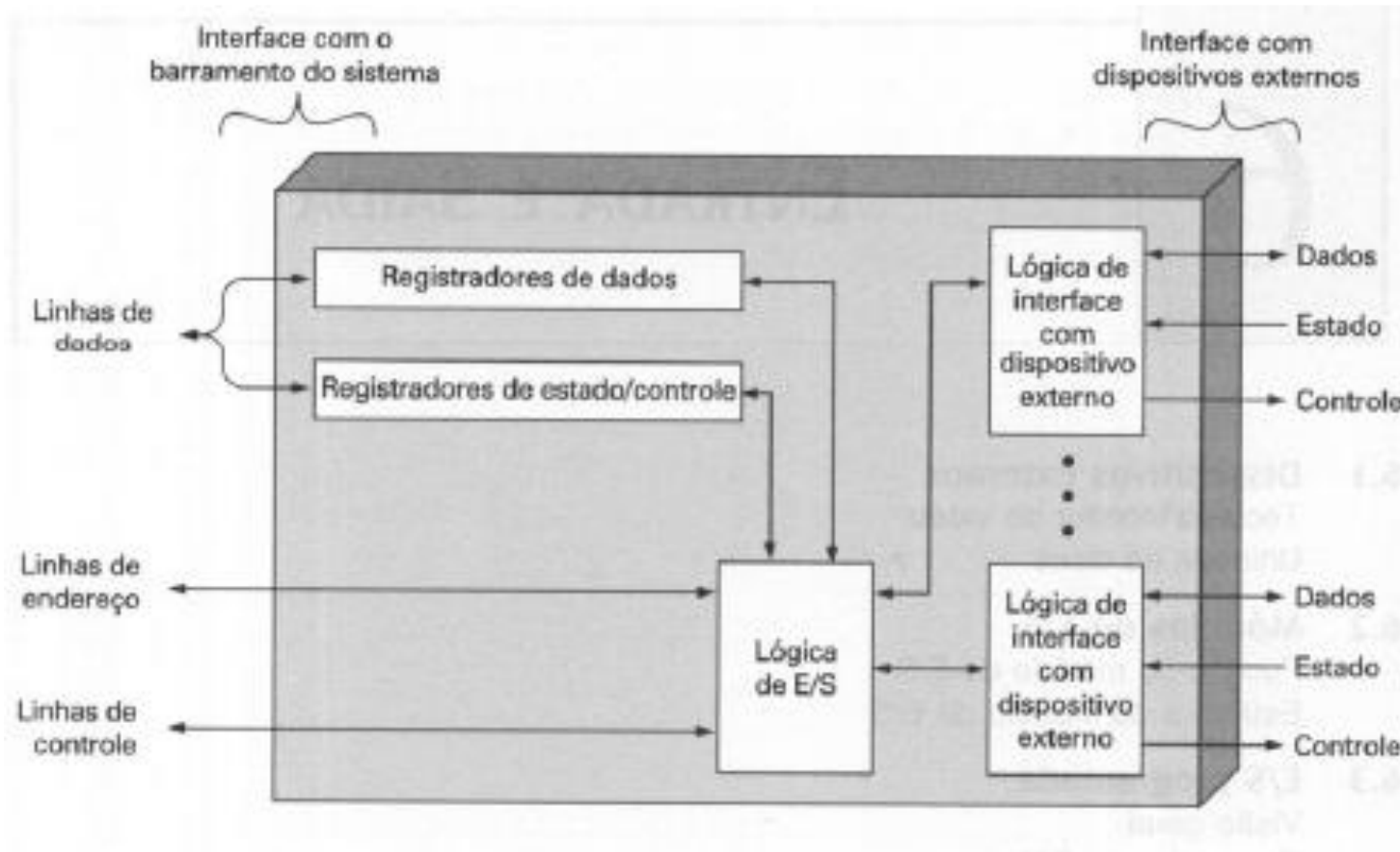


Figura 1: Arquitetura de E/S.

Dispositivos de Entrada e Saída

- A arquitetura de E/S constitui sua interface com o mundo exterior.
- É projetada para permitir um controle sistemático da interação com o mundo exterior e fornecer ao sistema operacional as informações de que ele necessita para gerenciar as atividades de E/S de maneira efetiva.

Dispositivos de Entrada e Saída

- Existem três técnicas principais de E/S:
 - **E/S programada:** é efetuada sob controle direto e contínuo do programa que requisitou a operação de E/S;
 - **E/S dirigida por interrupção:** o programa envia um comando de E/S e então continua a execução de instruções até que ocorra uma interrupção gerada pelo hardware de E/S, que sinaliza o término da operação de E/S requerida;
 - **Acesso direto à memória (DMA):** é controlada por um processador especializado de E/S, que se encarrega de transferir os blocos de dados.

E/S Programada

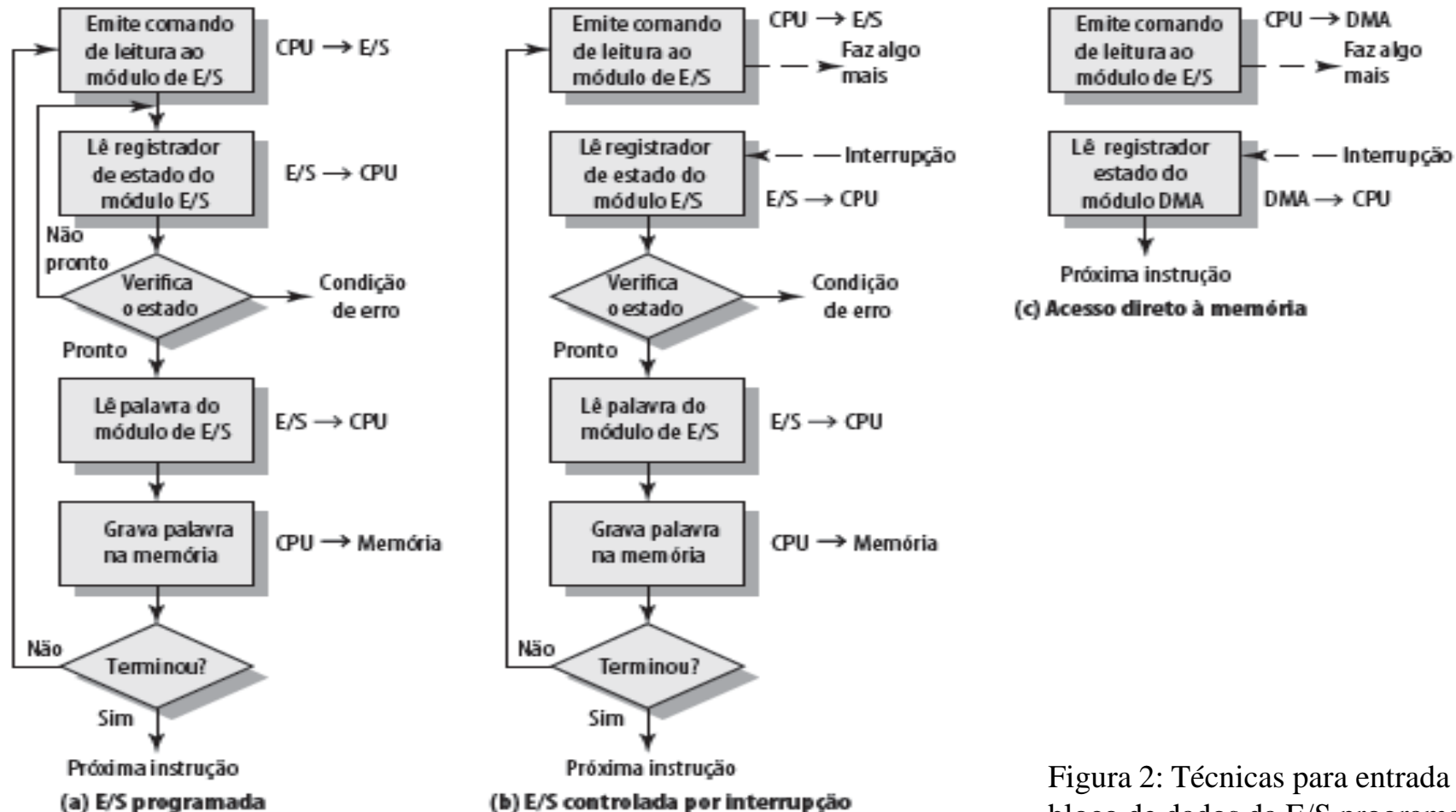


Figura 2: Técnicas para entrada de um bloco de dados da E/S programada.

E/S Programada

- CPU tem controle direto sobre E/S:
 - Conhecendo o estado.
 - Comandos de leitura/escrita.
 - Transferindo dados.
- CPU espera que módulo de E/S termine a operação.
- Desperdiça tempo de CPU.

Figura 2: Técnicas para entrada de um bloco de dados da E/S programada.

E/S Programada - Detalhe

- CPU solicita operação de E/S.
- Módulo de E/S realiza operação.
- Módulo de E/S define bits de estado.
- CPU verifica bits de estado periodicamente.
- Módulo de E/S não informa à CPU diretamente.
- Módulo de E/S não interrompe CPU.
- CPU pode esperar ou voltar mais tarde.

Figura 2: Técnicas para entrada de um bloco de dados da E/S programada.

E/S dirigida por interrupção

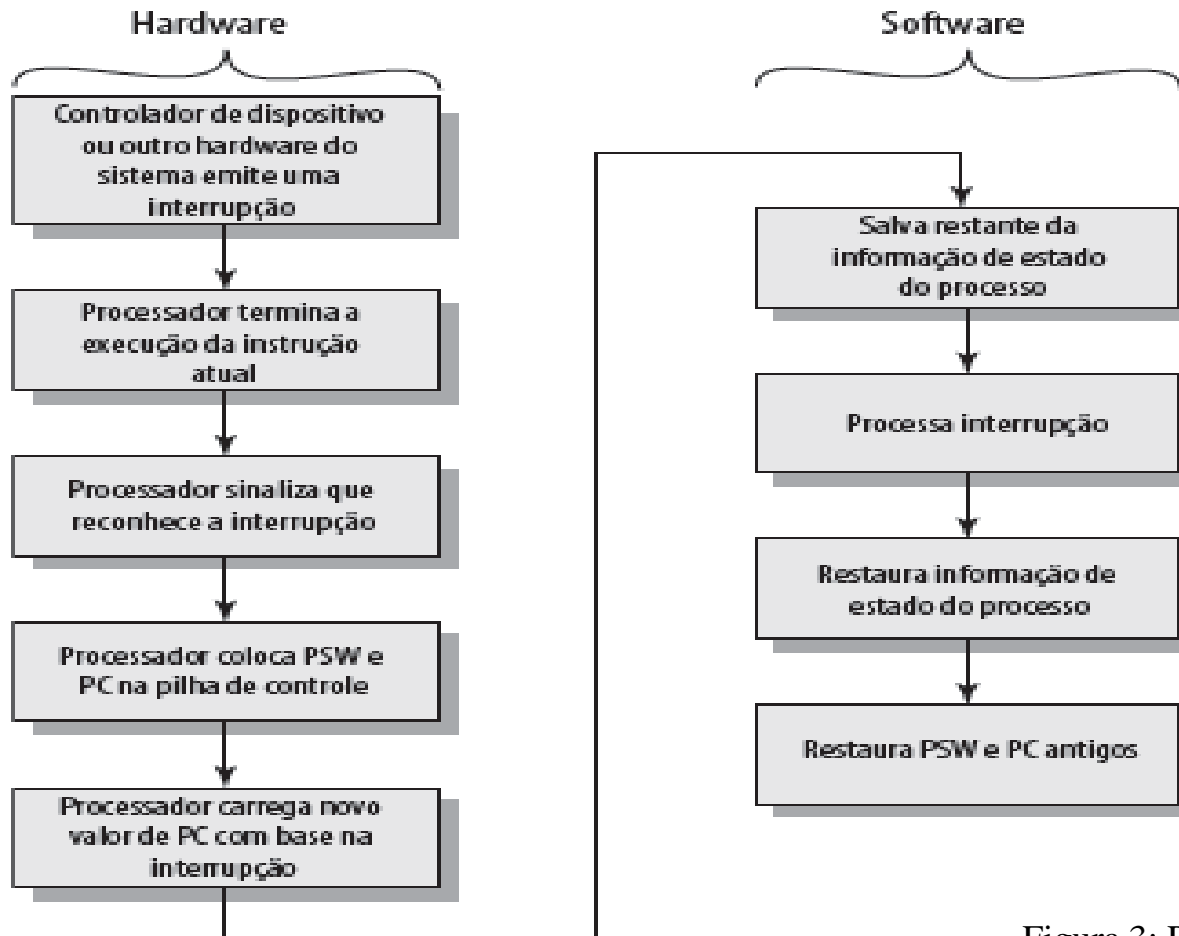


Figura 3: Processamento de interrupções simples

E/S dirigida por interrupção

- Contorna problema de espera da CPU.
- Sem verificação de dispositivo repetida da CPU.
- Módulo de E/S interrompe quando estiver pronto.

Figura 3: Processamento de interrupções simples

E/S dirigida por interrupção – Operação básica

- CPU emite comando de leitura.
- Módulo de E/S recebe dados do periférico enquanto CPU faz outro trabalho.
- Módulo de E/S interrompe CPU.
- CPU solicita dados.
- Módulo de E/S transfere dados.

Figura 3: Processamento de interrupções simples

DMA

- O DMA visa melhorar a performance geral do micro, permitindo que os periféricos transmitam dados diretamente para a memória, poupando o processador de mais esta tarefa.
- Função:
 - Módulo adicional (hardware) no barramento.
 - Controlador de DMA toma o comando da CPU para E/S.

DMA

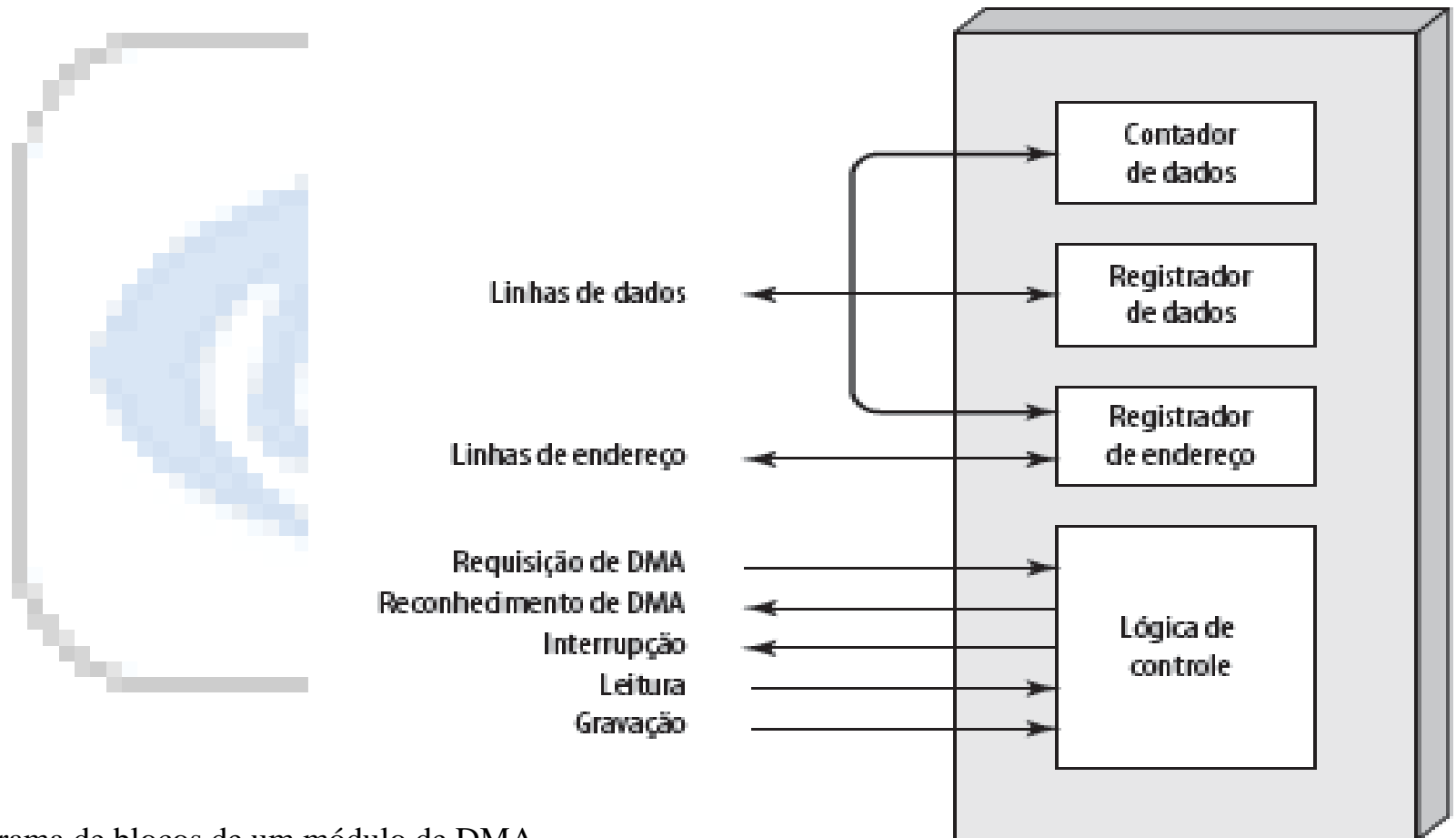
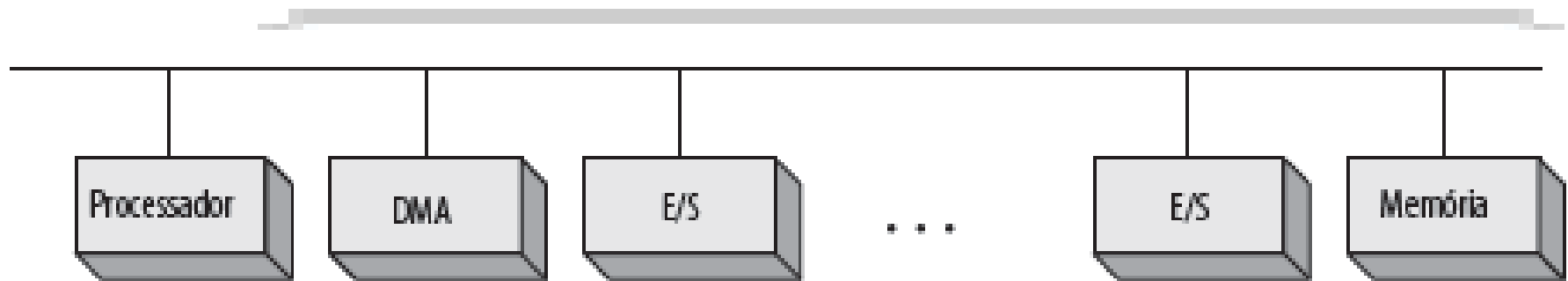


Figura 4: Diagrama de blocos de um módulo de DMA.

DMA - Operação

- CPU diz ao controlador de DMA:
 - Leitura/escrita.
 - Endereço do dispositivo.
 - Endereço inicial do bloco de memória para dados.
 - Quantidade de dados a serem transferidos.
- CPU prossegue com outro trabalho.
- Controlador de DMA lida com transferência.
- Controlador de DMA envia interrupção quando terminar.

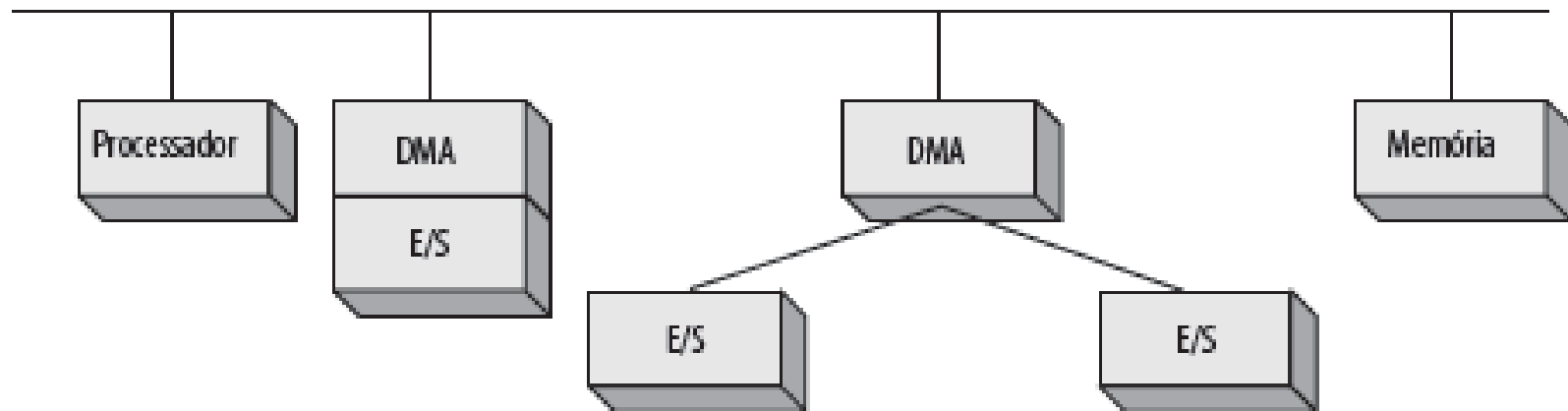
DMA - Configurações



(a) Único barramento, DMA separado

- Único barramento, controle de DMA separado.
- Cada transferência usa barramento duas vezes.
 - E/S para DMA, depois DMA para memória.
- CPU é suspensa duas vezes.

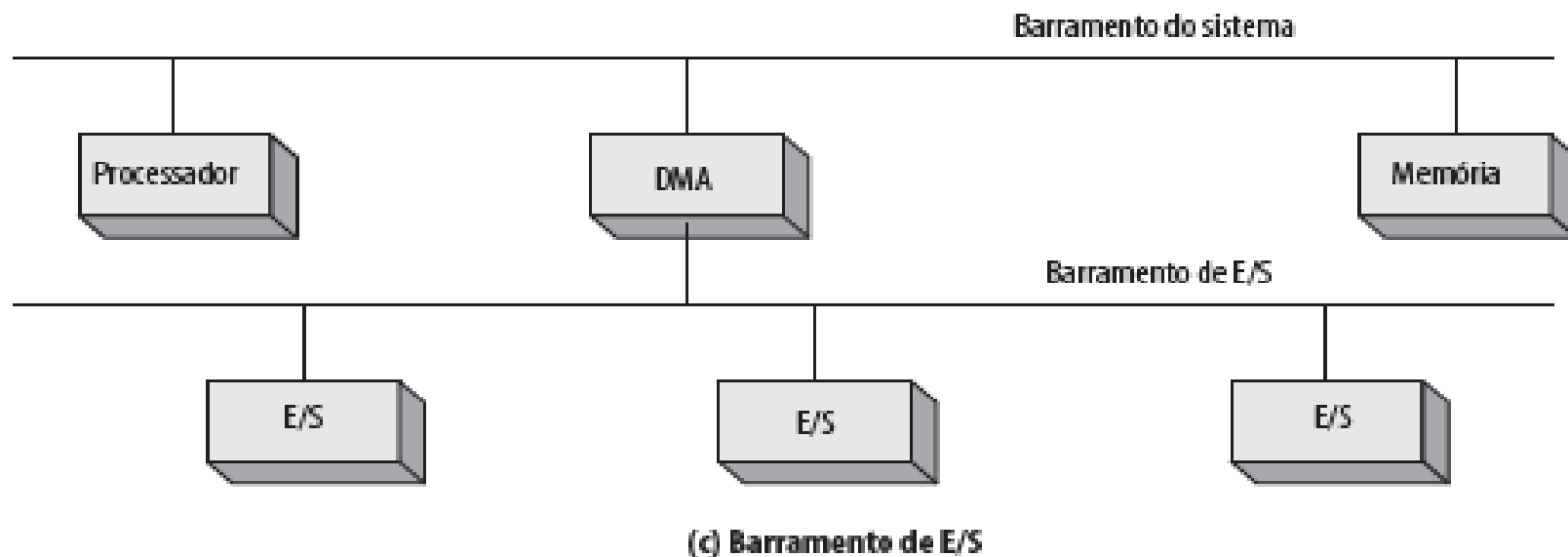
DMA - Configurações



(b) Único barramento, DMA-E/S integrados

- Único barramento, controlador de DMA integrado.
- Controlador pode aceitar mais de um dispositivo.
- Cada transferência usa barramento uma vez.
 - DMA para memória.
- CPU é suspensa uma vez.

DMA - Configurações



- Barramento de E/S separado.
- Barramento aceita todos dispositivos habilitados para DMA.
- Cada transferência usa barramento uma vez.
 - DMA para memória.
- CPU é suspensa uma vez.

Técnicas de E/S

	Sem Interrupções	Com Interrupções
Transferência entre memória e E/S por meio do processador	E/S programada	E/S dirigida por interrupção
Transferência direta entre memória e E/S		Acesso direto à memória (DMA)

Tabela 1: Técnicas de E/S

Dispositivos de Entrada e Saída

- Por que os periféricos não são diretamente conectados ao barramento de sistema?
 - Por existir uma grande variedade de periféricos, com diferentes mecanismos de operação.
 - Devido a taxa de transferência de dados dos periféricos ser, frequentemente, muito menor do que a taxa de transferência de dados da memória ou do processador
 - Devido aos periféricos usarem frequentemente formatos de dados e tamanhos de palavras diferentes dos usados no computador ao qual estão conectados.

Dispositivos de Entrada e Saída

- Por essas razões, é requerido um módulo de E/S, que deve desempenhar duas funções principais:
 - Fornecer uma interface com o processador e a memória, através do barramento do sistema ou do comutador central.
 - Permitir a interface com um ou mais dispositivos periféricos, através de conexões de dados adequadas.

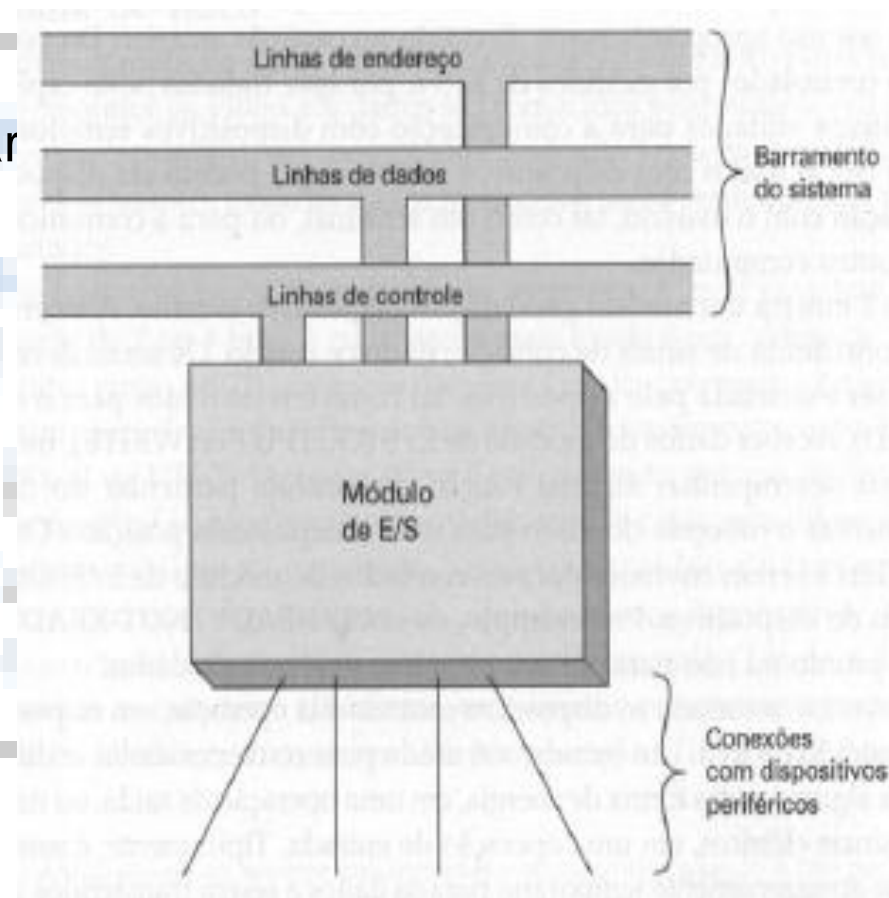


Figura 6: Modelo geral de um módulo de E/S.

Dispositivos de Entrada e Saída

- Funções dos módulo de E/S:
 - Controle e Temporização;
 - Comunicação com o processador;
 - Comunicação com os dispositivos;
 - Área de armazenamento temporário de dados;
 - Detecção de erros.

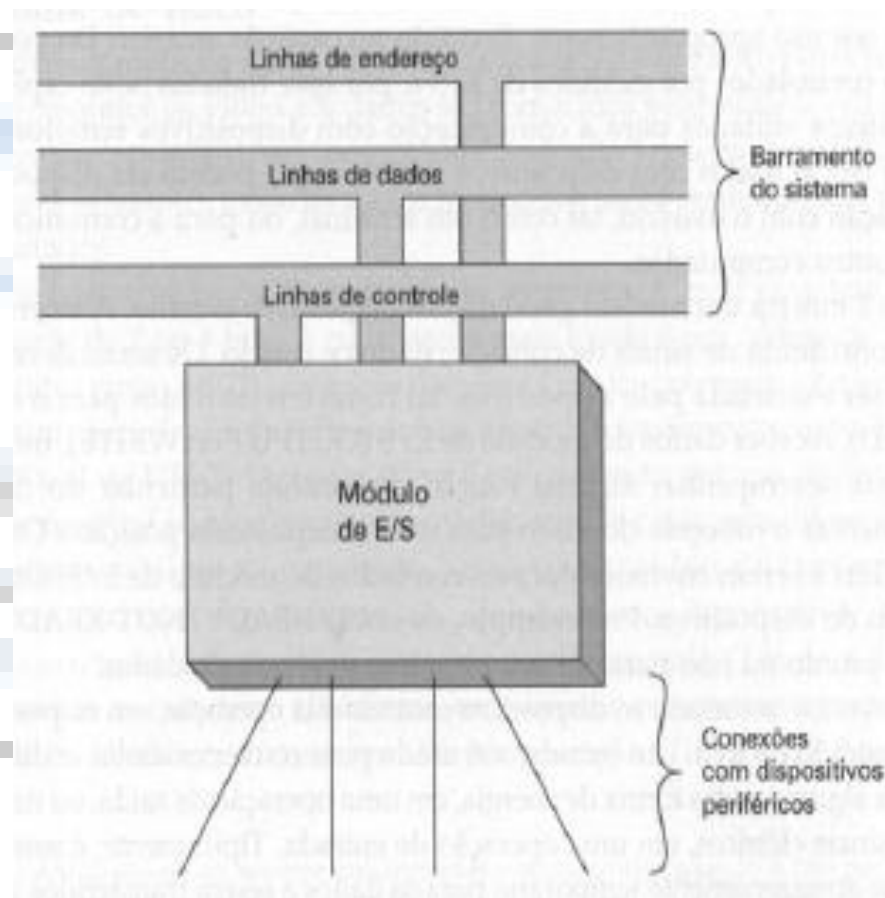


Figura 6: Modelo geral de um módulo de E/S.

RAID

- A melhoria no desempenho de memórias secundárias tem sido consideravelmente menor que a de processadores e da memória principal.
- Essa diferença faz dos sistemas de armazenamento de discos o principal foco de preocupação para melhoria do desempenhos global dos sistemas de computação.
- No caso do armazenamento de disco, essa idéia levou ao desenvolvimento de um agrupamento de discos que operam independentemente e em paralelo.

RAID

- RAID significa "Redundant Array of Inexpensive Disks", ou Agrupamento Redundante de Discos Independentes.
- Com diversos discos, diferentes requisições de E/S podem ser processadas em paralelo, desde que os dados requeridos residam em discos separados.
- Mais do que isso, uma única requisição de E/S poderá também ser executada em paralelo, se o bloco de dados a ser acessado for distribuído em vários discos.
- Com o uso de múltiplos discos, os dados podem ser organizados de diversas maneiras, podendo ser empregada alguma redundância para melhorar a confiabilidade.

RAID

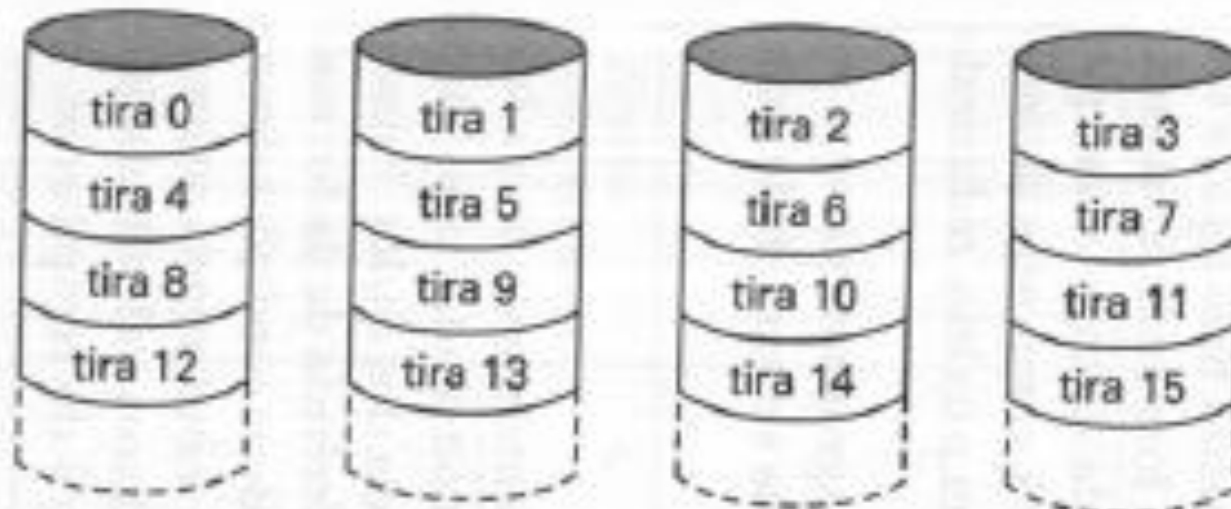
- O esquema de RAID consistem em níveis, o que não significa uma relação hierárquica, mas designam diferentes arquiteturas de projeto que compartilham três características comuns:
 - O RAID consistem em um agrupamento de unidades de disco físicos, visto pelo sistema operacional como uma única unidade de disco lógico;
 - Os dados são distribuídos pelas unidades de discos físicos do agrupamento;
 - A capacidade de armazenamento redundante é utilizada para armazenar informação de paridade, garantindo a recuperação dos dados em caso de falha em algum disco.

RAID – modos disponíveis

∴ RAID 0 (Striping) ∴

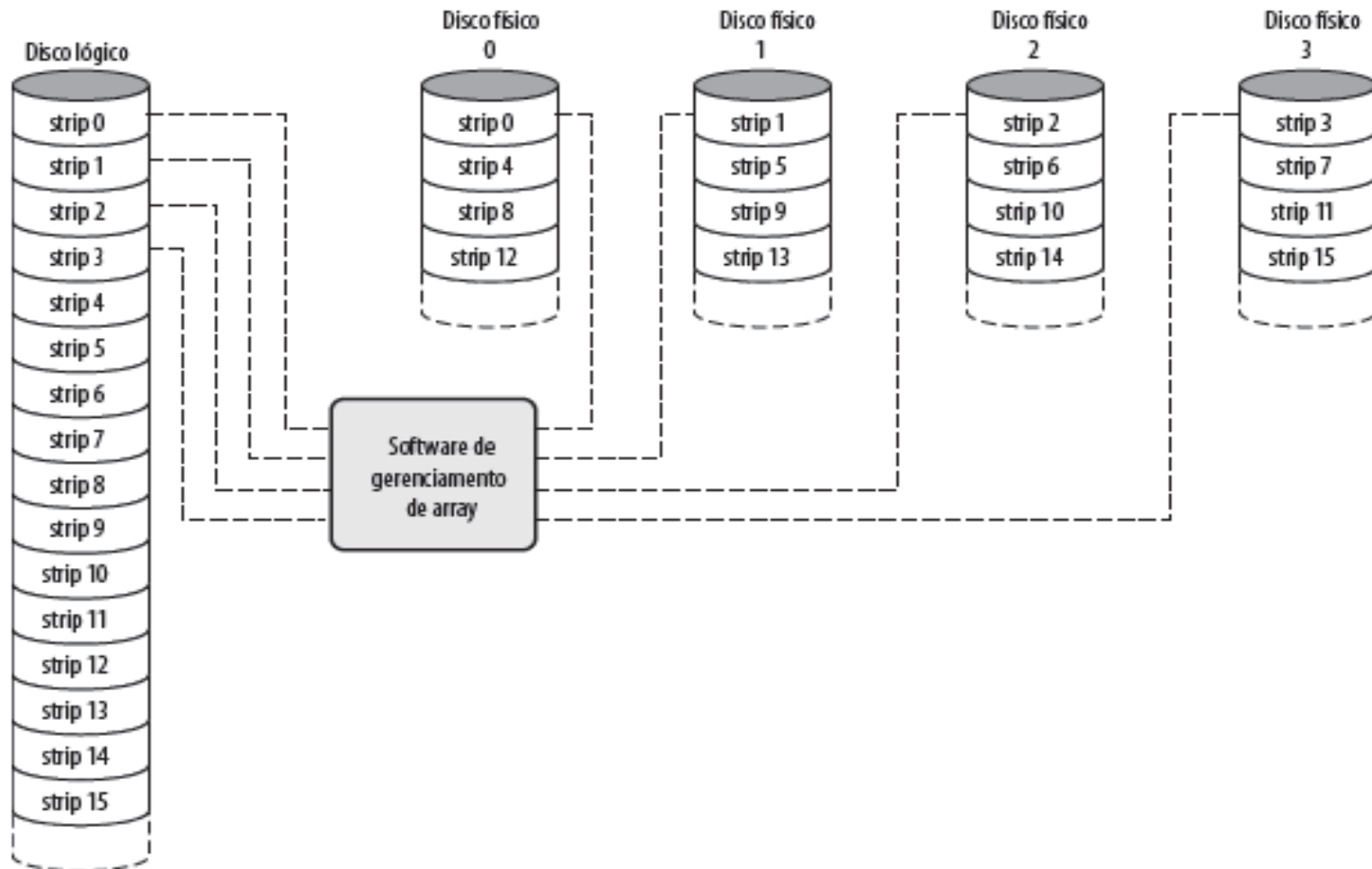
- É o modo que permite obter a melhor performance possível, sacrificando parte da confiabilidade. Todos os discos passam a ser acessados como se fossem um único drive.
- Ao serem gravados, os arquivos são fragmentados nos vários discos, permitindo que os fragmentos possam ser lidos/gravados ao mesmo tempo.
- O problema é que caso qualquer um dos HDs apresente problema, serão perdidos os dados armazenados em todos os HDs, já que qualquer arquivo torna-se inútil caso uma parte do código seja perdida.

RAID 0



a) RAID 0 (não-redundante)

Mapeamento de dados para RAID 0

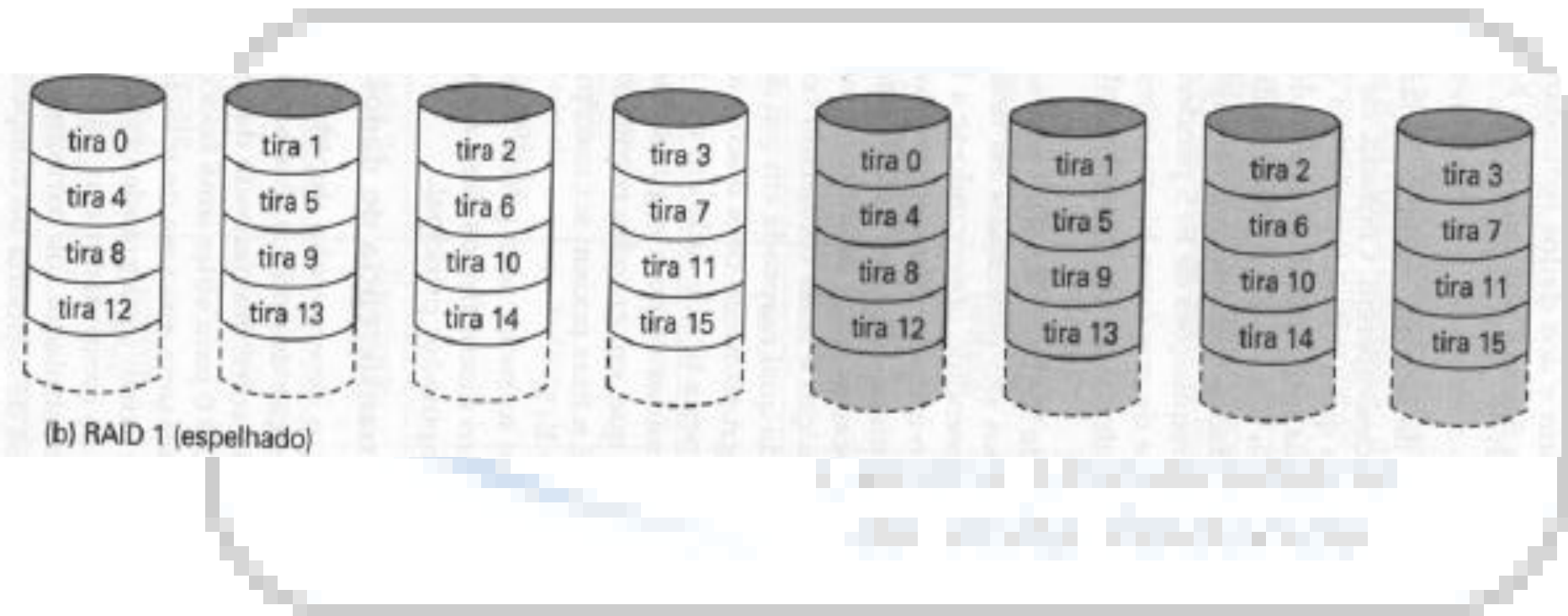


RAID – modos disponíveis

∴ RAID 1 (Mirroring) ∴

- É o famoso sistema de espelhamento, conseguido usando dois HDs.
- Um deles armazena dados, enquanto o seguinte armazena uma cópia fiel dos mesmos dados.
- Caso qualquer um dos HDs pare, ele é automaticamente substituído pelo seu "clone" e o sistema continua intacto. Na maioria das controladoras RAID SCSI é possível realizar a troca do HD defeituoso "a quente", com o micro ligado, recurso ainda não disponível nas controladoras RAID IDE.

RAID 1

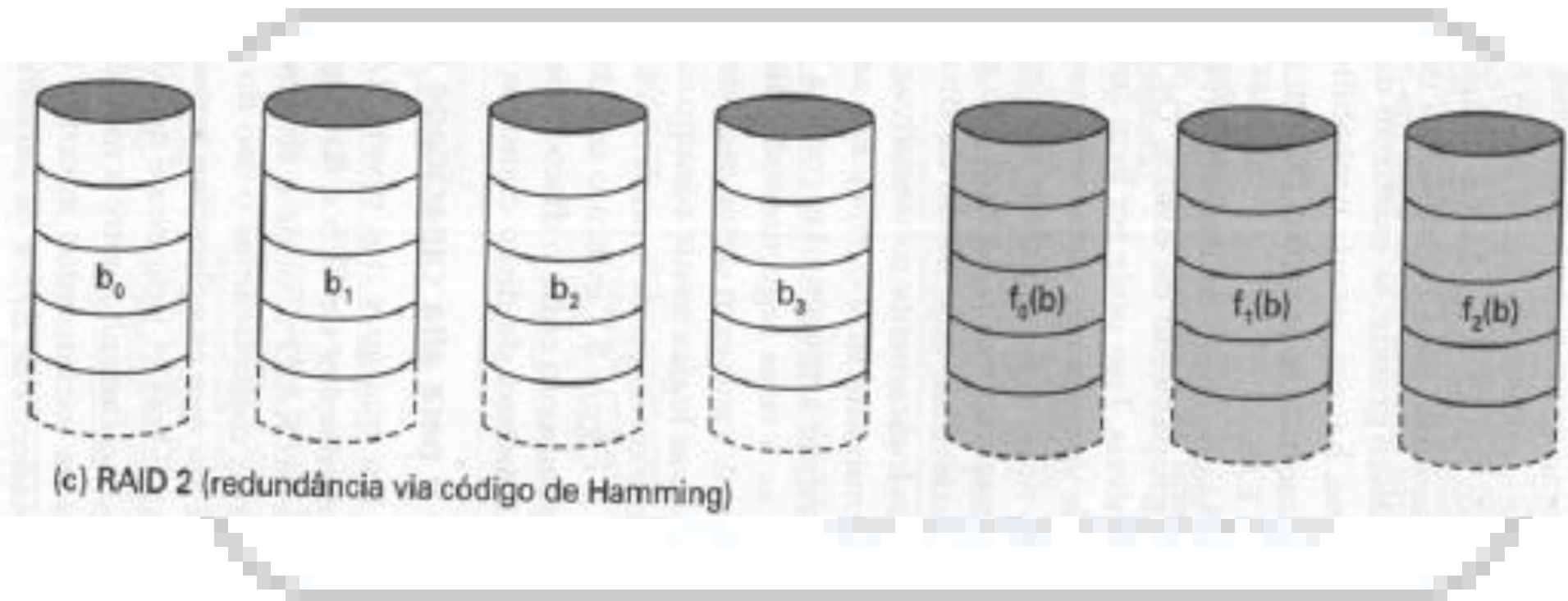


RAID – modos disponíveis

∴ RAID 2 ∴

- É um modo que não é mais utilizado.
- O RAID 2 consiste em embutir códigos de correção de erros em cada cluster de dados gravado. Porém, todos os HDs atuais já vem com sistemas de correção de erros embutidos, tornando o sistema obsoleto.

RAID 2

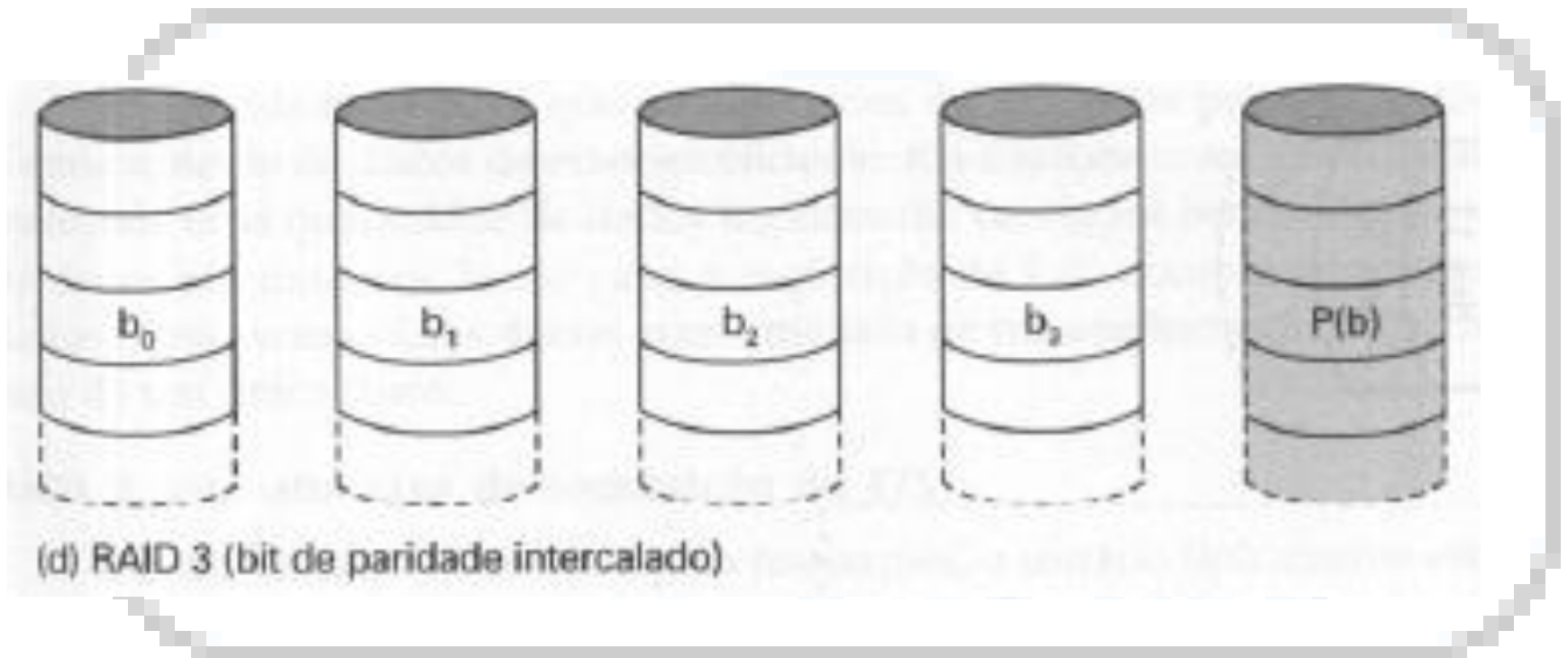


RAID – modos disponíveis

∴ RAID 3 ∴

- O RAID 3 usa um sistema de paridade para manter a integridade dos dados. Num sistema com 5 HDs, os 4 primeiros servirão para armazenar dados, enquanto o último armazenará os códigos de paridade.
- Nos 4 primeiros drives temos na verdade um sistema RAID 0, onde os dados são distribuídos entre os 4 HDs e a performance é multiplicada por 4. Porém, os códigos armazenados no 5º HD permitem recuperar os dados caso qualquer um dos 4 HDs pare.
- caso dois ou mais HDs apresentem problemas ao mesmo tempo, ou antes da controladora terminar a reconstrução dos dados, novamente perdem-se todos os dados de todos os HDs.

RAID 3

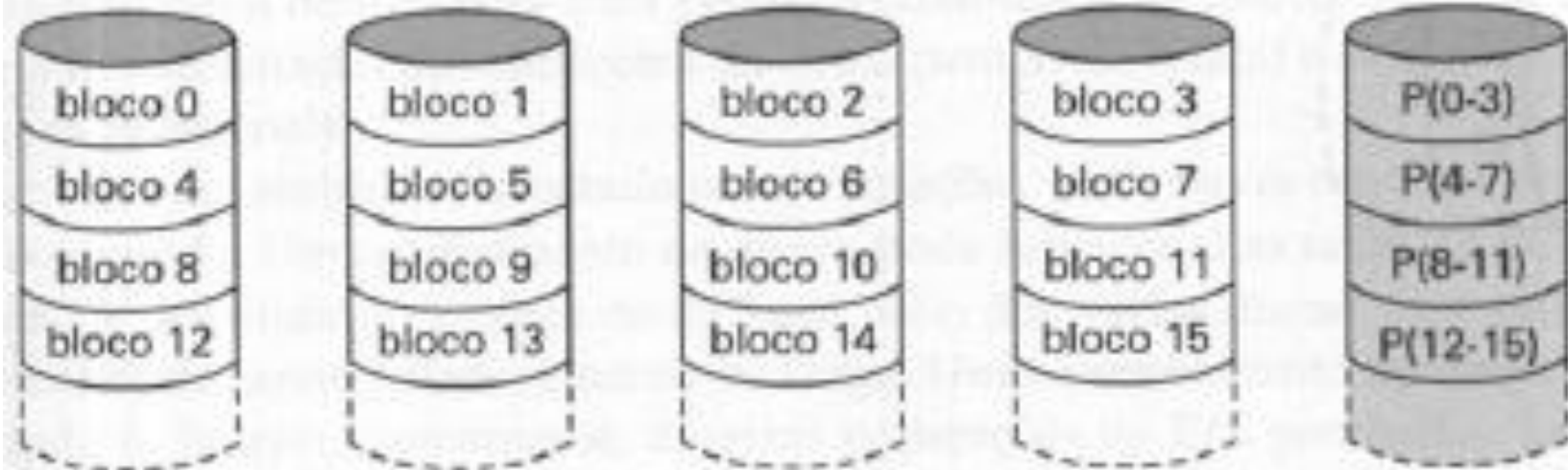


RAID – modos disponíveis

∴ RAID 4 ∴

- Este modo é parecido com o RAID 3, mas a forma como os dados são gravados nos demais discos é diferente.
- No RAID 4 os dados são divididos em blocos, pedaços bem maiores do que no RAID 3. Com isto, é possível ler vários arquivos ao mesmo tempo, o que é útil em algumas aplicações, porém o processo de gravação é bem mais lento que no RAID 3.
- O RAID 4 apresenta um bom desempenho em aplicações onde seja preciso ler uma grande quantidade de arquivos pequenos. No RAID 4 o tempo de reconstrução dos dados caso um dos HDs falhe é bem maior do que no RAID 3.

RAID 4



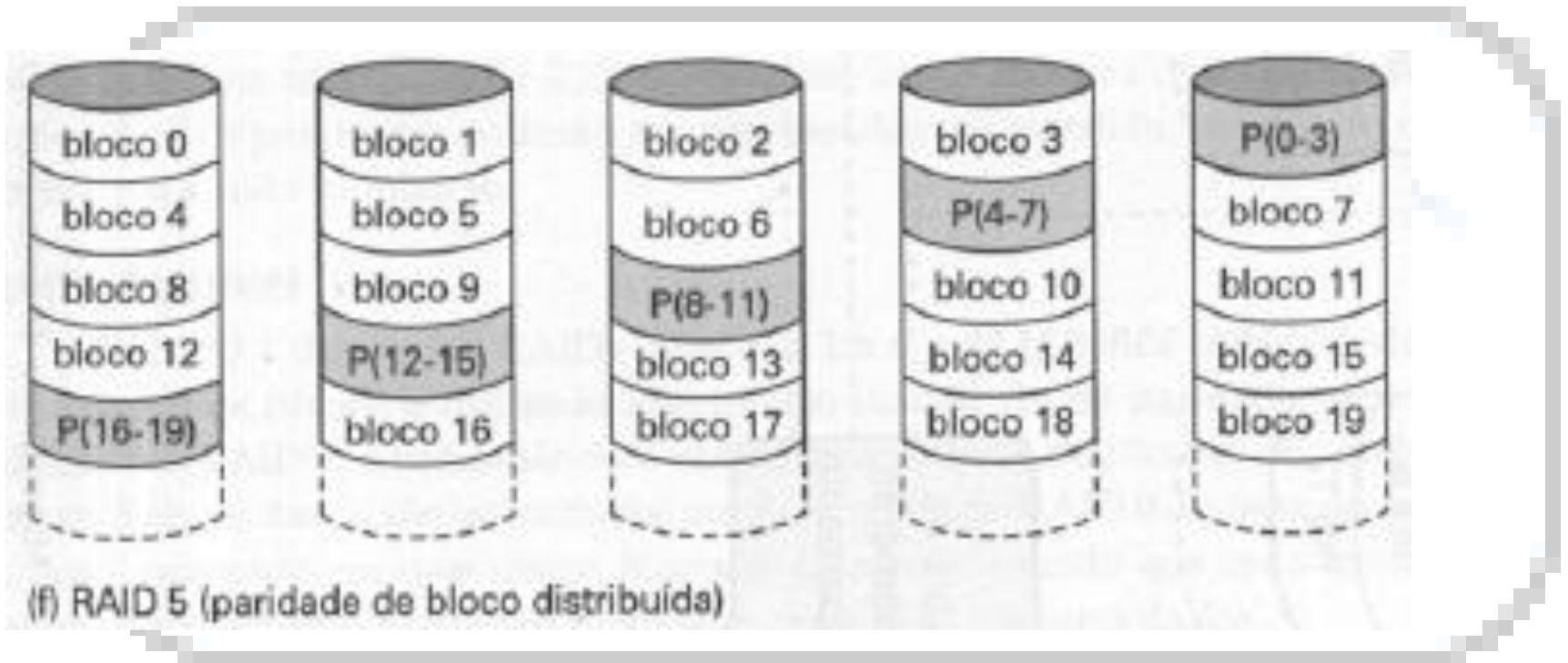
(e) RAID 4 (paridade de bloco)

RAID – modos disponíveis

∴ RAID 5 ∴

- É mais um sistema que baseia-se no uso de paridade para garantir a integridade dos dados caso um HD falhe.
- A diferença sobre o RAID 3 é que ao invés de dedicar um HD a esta tarefa, os dados de correção são espalhados entre os discos. A vantagem sobre o RAID 3 é alcançar taxas de leitura um pouco mais altas, pois será possível ler dados a partir de todos os HDs simultaneamente, entretanto as gravações de dados são um pouco mais lentas.
- O RAID 5 pode ser implementado a partir de 3 discos. Apesar dos dados de paridade serem espalhados pelos discos, o espaço esquivamente à um dos HDs é consumido por eles.

RAID 5

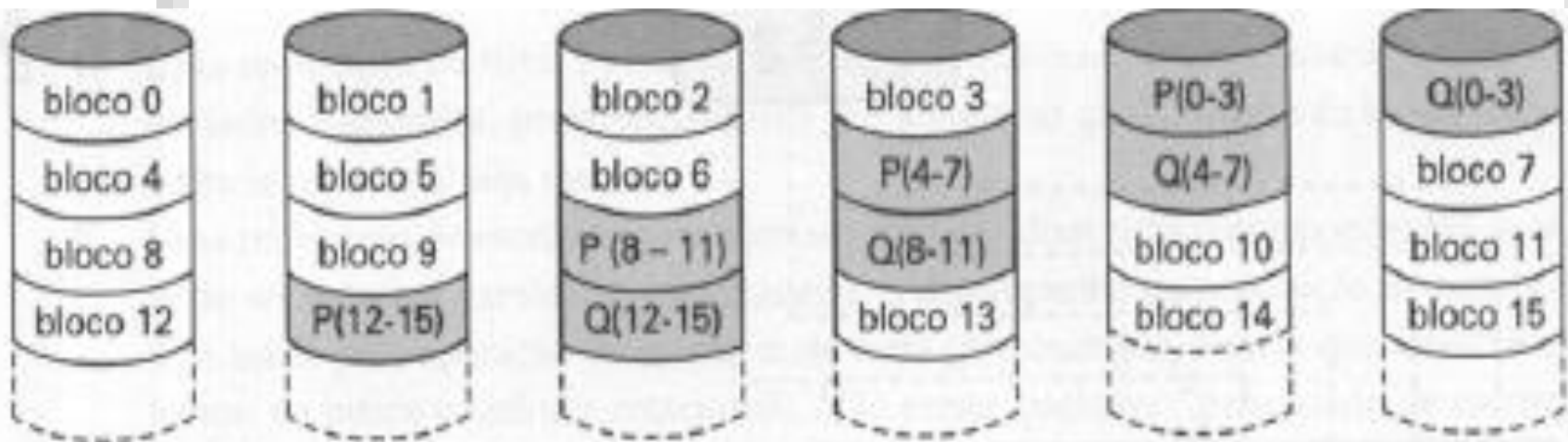


RAID – modos disponíveis

∴ RAID 6 ∴

- É um padrão relativamente novo, suportado por apenas algumas controladoras.
- É semelhante ao RAID 5, porém usa o dobro de bits de paridade, garantindo a integridade dos dados caso até 2 dos HDs falhem ao mesmo tempo.
- Ao usar 8 HDs de 20 GB cada um em RAID 6, teremos 120 GB de dados e 40 GB de paridade.

RAID 6



(g) RAID 6 (redundância dupla)

RAID – modos disponíveis

∴ RAID 10 ∴

- Este sistema combina características do RAID 0 e RAID 1, daí o nome.
- O RAID 10 pode ser implementando em sistemas com 4 discos ou mais, sendo obrigatório um número par (6, 8, etc.).
- Metade dos discos armazena dados e a outra metade armazena uma cópia. A metade que armazena dados é combinada, formando um sistema RAID 0, aumentando a performance, porém mantendo a confiabilidade, já que temos cópias de todos os dados. Usando 4 HDs de 20 GB em modo 10, teremos 40 GB de dados e o dobro de desempenho que em um HD sozinho.

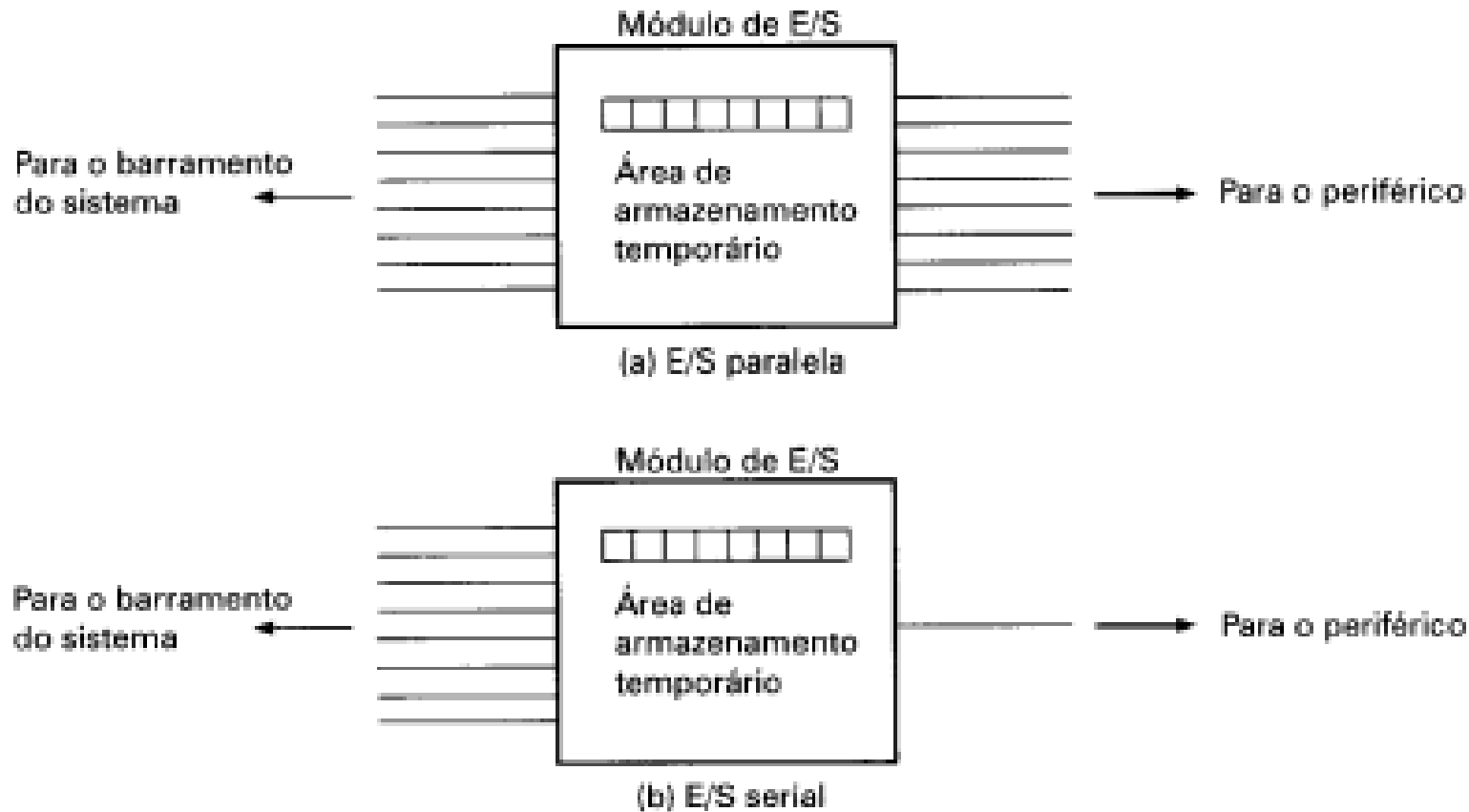
RAID – Resumo

Categoria	Nível	Descrição	Taxa de requisição de E/S <i>Leitura/escrita</i>	Taxa de transferência de dados <i>Leitura/escrita</i>	Aplicação típica
Intercalação de dados (striping)	0	Não-redundante	Tiras grandes: excelente	Tiras pequenas: excelente	Aplicações que requerem alto desempenho para dados não- críticos
Espelhamento	1	Espelhado	Bom/ razoável	Razoável/ razoável	Unidades de disco de sistema; arquivos críticos
Acesso Paralelo	2	Redundante via código	Pobre	Excelente	
	3	Paridade de bit intercalada	Pobre	Excelente	Aplicações com grandes requisições de E/S, como processamento de imagens e CAD
Acesso Independente	4	Paridade de bloco intercalada	Excelente/ razoável	Razoável/ pobre	
	5	Paridade de bit intercalada e distribuída	Excelente/ razoável	Razoável/ pobre	Buscas de dados, com altas taxas de requisição e grande volume de leituras
	6	Paridade de bloco dupla intercalada e distribuída	Excelente/ pobre	Razoável/ pobre	Aplicações que requerem grande disponibilidade de dados

Interface Externa

- Tanto a E/S paralela quanto E/S serial tem de interagir com o periférico. Em termos gerais, a interação em uma operação de escrita pode ser descrita como a seguir:
 - O módulo de E/S envia um sinal de controle pedindo permissão para enviar um dado;
 - O periférico reconhece a requisição;
 - O módulo de E/S transfere dados (uma palavra ou um bloco, dependendo do tipo de periférico);
 - O periférico sinaliza o recebimento dos dados.

E/S Paralela e E/S Serial



Questões

1. Qual a função do RAID? Explique o funcionamento.
2. Descreva os padrões de RAID
3. Cite e descreva as técnicas principais de E/S.
4. Por que os periféricos não são diretamente conectados ao barramento de sistema?
5. Explique o funcionamento dos módulos de E/S nas interfaces paralelas e seriais.
6. Cite e Explique as configurações de DMA