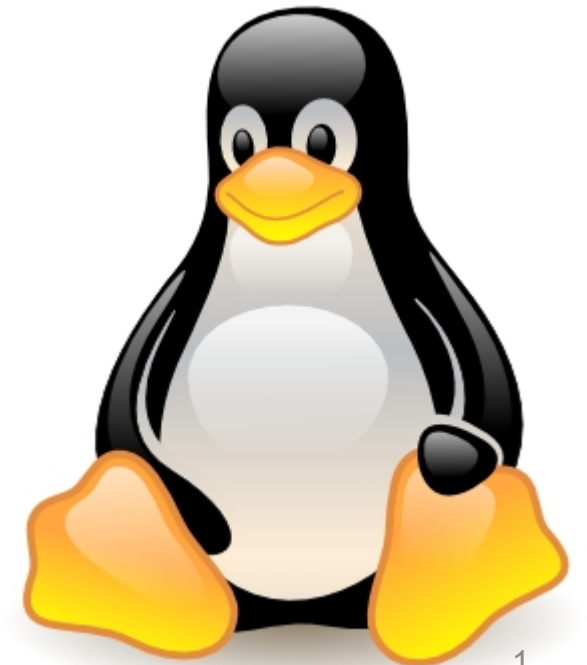


Sistemas Operacionais II

Sistemas de Arquivos



13.1 Introdução

■ Arquivos

- Coleção nomeada de dados manipulada como uma unidade.
- Residem em dispositivos de armazenamento secundário.

■ Os sistemas operacionais podem criar uma interface que facilita a navegação dos arquivos de um usuário.

- Os sistemas de arquivo podem proteger esses dados contra corrupção ou perda total decorrente de acidentes.
- Os sistemas que gerenciam grande quantidade de dados compartilhados podem se beneficiar dos bancos de dados como uma alternativa aos arquivos.



13.2 Hierarquia de dados

- **As informações são armazenadas em computadores segundo uma hierarquia de dados.**
- **O nível mais baixo da hierarquia de dados é composto de bits.**
 - Os padrões de bits representam todos os itens de dados de interesse nos sistemas de computador.

13.2 Hierarquia de dados

- **O nível seguinte da hierarquia de dados são os padrões de bits de tamanho fixos, como bytes, caracteres e palavras.**
 - Byte: em geral 8 bits.
 - Palavra: o número de bits que um processador pode operar por vez.
 - Os caracteres mapeiam bytes (ou grupos de bytes) para símbolos como letras, números, pontuação e novas linhas.
 - Os três conjuntos de caracteres mais populares em uso hoje são o ASCII, o EBCDIC e o Unicode.
 - Campo: um grupo de caracteres.
 - Registro: um grupo de campos.
 - Arquivo: um grupo de registros relacionados.



13.2 Hierarquia de dados

- **O nível mais alto da hierarquia de dados é um sistema de arquivos ou banco de dados.**
- **Volume é uma unidade de armazenamento de dados que pode conter vários arquivos.**

13.3 Arquivos

- **Arquivo: coleção nomeada de dados que pode ser manipulada como uma unidade por operações como:**
 - Abrir
 - Fechar
 - Criar
 - Destruir
 - Copiar
 - Renomear
 - Listar



13.3 Arquivos

- **Itens de dados individuais dentro de um arquivo podem ser manipulados por operações como:**
 - Ler
 - Escrever (gravar)
 - Atualizar
 - Inserir
 - Apagar
- **Os arquivos podem ser caracterizados por atributos como:**
 - Localização
 - Acessibilidade
 - Tipo
 - Volatilidade
 - Atividade
- **Os arquivos podem consistir em um ou mais registros.**



13.4 Sistemas de arquivo

■ **Sistemas de arquivo**

- Organizam arquivos e gerenciam o acesso aos dados.
- São responsáveis pelo gerenciamento dos arquivos, pelo gerenciamento do armazenamento auxiliar (secundário e terciário), pelos mecanismos de integridade do arquivo e pelos métodos de acesso.
- Preocupam-se primordialmente com o gerenciamento do espaço de armazenamento secundário, em especial com o armazenamento em disco.

13.4 Sistemas de arquivo

■ Características do sistema de arquivo:

- Devem exibir independência de dispositivos:
 - Os usuários devem poder referir-se a seus arquivos por nomes simbólicos em vez de ter de utilizar nomes de dispositivos físicos.
- Devem também oferecer o recurso de cópia de segurança e recuperação para evitar perdas acidentais ou danos maliciosos às informações.
- Podem ainda oferecer o recurso de criptografia e deciptação para tornar as informações úteis apenas ao público a que se destina.

13.4.1 Diretórios

■ **Diretórios:**

- Arquivos que contêm o nome e o local de outros arquivos no sistema de arquivos para organizar e localizar arquivos rapidamente.

■ **A entrada de diretório armazena informações como:**

- Nome do arquivo
- Local
- Tamanho
- Tipo
- Horário de acesso
- Horário de modificação e criação

13.4.1 Diretórios

Figura 13.1 Exemplo de conteúdo de diretório de arquivo.

<i>Diretório</i>	<i>Campo Descrição</i>
Nome	Série de caracteres que representa o nome do arquivo.
Localização	Bloco físico ou localização lógica do arquivo no sistema de arquivo (ou seja, um nome de caminho).
Tamanho	Número de bytes consumido pelo arquivo.
Tipo	Descrição do propósito do arquivo (por exemplo, arquivo de dados ou arquivo de diretório).
Horário de acesso	Última vez que o arquivo foi acessado
Horário de modificação	Última vez que o arquivo foi modificado.
Horário de criação	Quando o arquivo foi criado.



13.4.1 Diretórios

- **Sistema de arquivos de nível único (ou plano):**
 - É a organização mais simples de sistemas de arquivo.
 - Armazena todos os seus arquivos usando um único diretório.
 - Dois arquivos não podem ter o mesmo nome.
 - O sistema de arquivo tem de executar uma busca linear do conteúdo do diretório para localizar cada arquivo, o que pode prejudicar o desempenho.

13.4.1 Diretórios

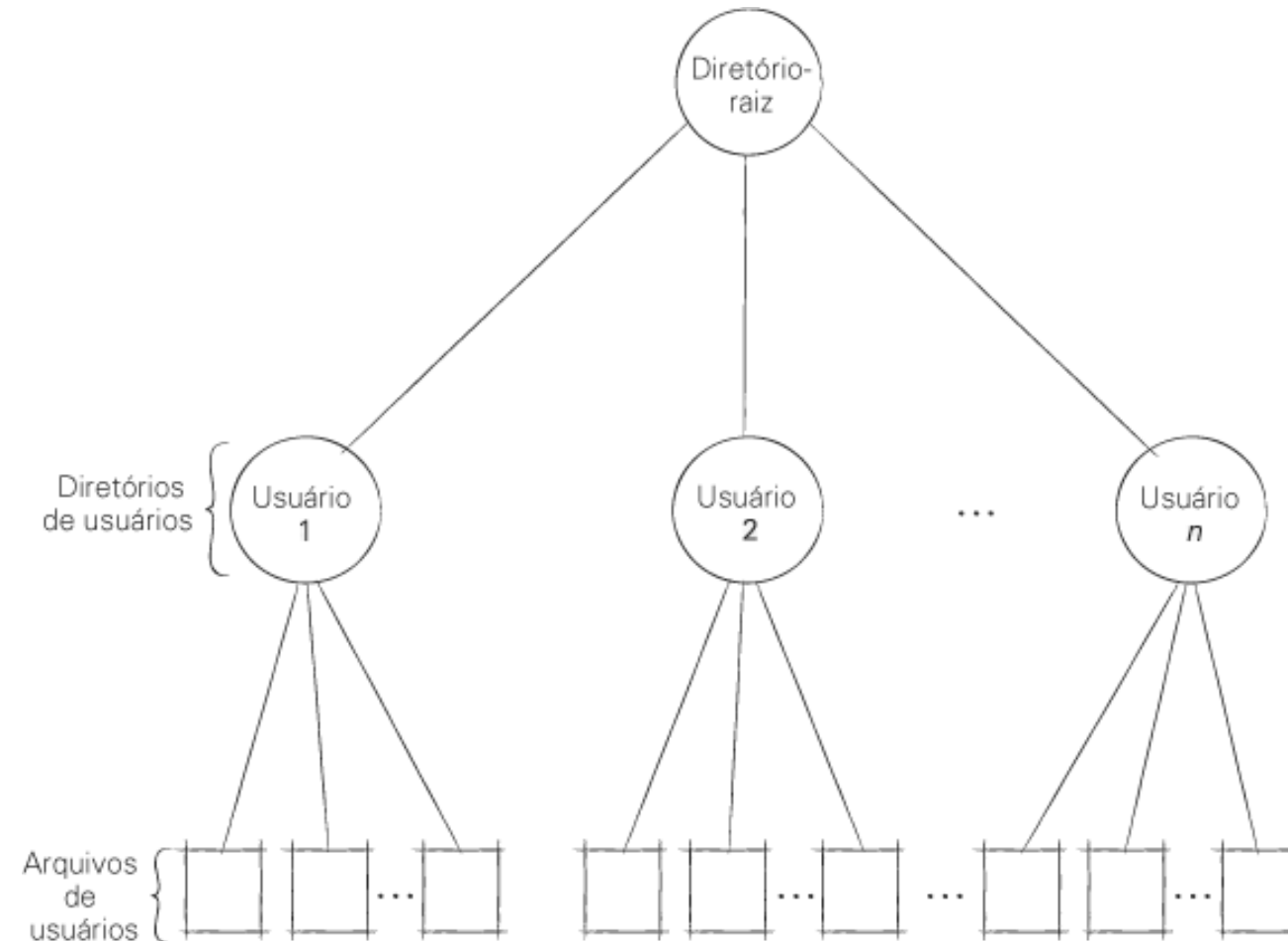
■ **Sistema de arquivo hierárquico:**

- Uma raiz indica em que lugar do dispositivo de armazenamento começa o diretório-raiz.
- O diretório-raiz aponta para os vários diretórios, cada um dos quais contém uma entrada para cada um de seus arquivos.
- O nome do arquivo tem de ser exclusivo apenas em um determinado diretório do usuário.
- O nome de um arquivo em geral é formado tal como o nome de caminho desde o diretório-raiz até o arquivo.



13.4.1 Diretórios

Figura 13.2 Sistema de arquivo hierárquico de dois níveis.



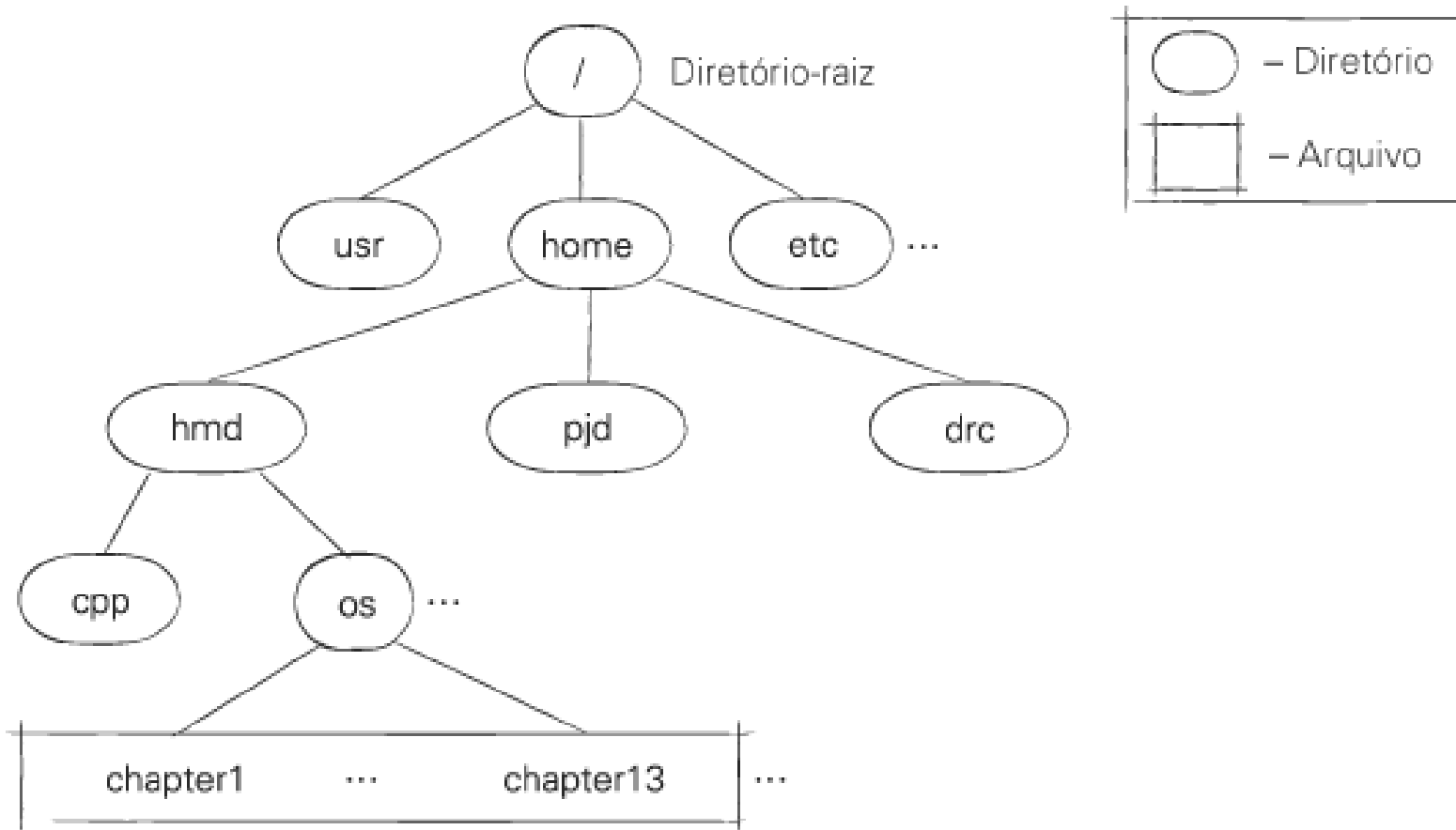
13.4.1 Diretórios

■ Diretório de trabalho

- Simplifica a navegação usando nomes de caminho.
- Permite que os usuários especifiquem um nome de caminho que não comece no diretório-raiz (isto é, um caminho relativo).
- Caminho absoluto (isto é, o caminho que começa na raiz) = diretório de trabalho + caminho relativo.

13.4.1 Diretórios

Figura 13.3 Exemplo de conteúdo de sistema de arquivo hierárquico.



13.4.1 Diretórios

- **Ligação: uma entrada de diretório que se refere a um arquivo de dados ou diretório que normalmente está localizado em um diretório diferente.**
 - Facilita o compartilhamento de dados e pode ajudar os usuários a acessar arquivos localizados em toda a estrutura de diretório de um sistema de arquivo.
 - Ligação flexível: entrada de diretório que contém o nome de caminho para outro arquivo.
 - Ligação estrita: entrada de diretório que especifica a localização de um arquivo (em geral um número de bloco) no dispositivo de armazenamento.



13.4.1 Diretórios

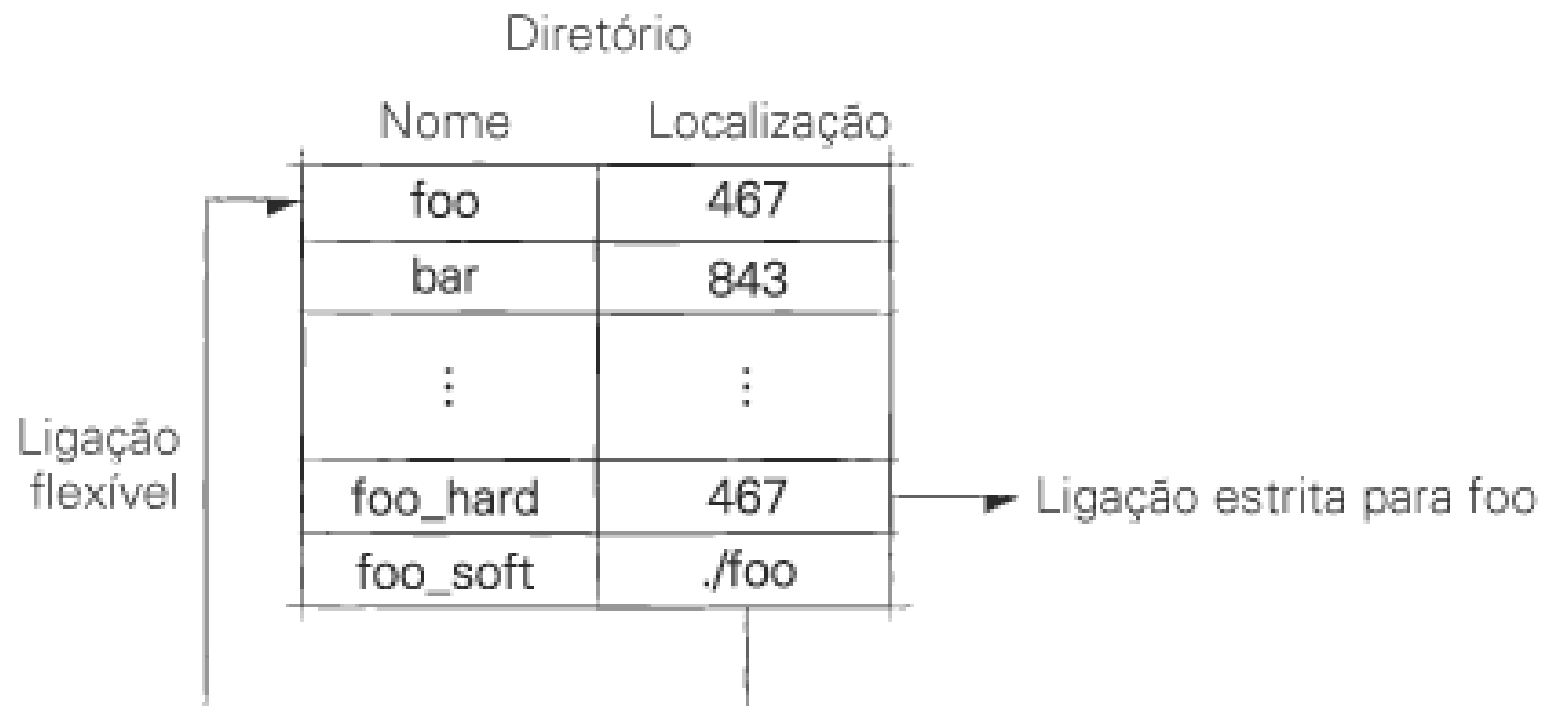
■ **Ligação (*continuação*)**

- Pelo fato de o disco rígido especificar a localização física de um arquivo, a ligação estrita referencia dados inválidos quando essa localização física do arquivo correspondente muda.
- Pelo fato de a ligação flexível armazenar a localização lógica do arquivo no sistema de arquivo, não exige atualização quando os dados do arquivo são movidos.
- Entretanto, se um usuário mover um arquivo para um diretório diferente ou renomear o arquivo, toda possível ligação flexível com esse arquivo não será mais válida.



13.4.1 Diretórios

Figura 13.4 Ligações de um sistema de arquivo.



13.4.2 Metadados

■ Metadados

- Informações que protegem a integridade do sistema de arquivo.
- Não podem ser modificados pelos usuários.

■ **Vários sistemas de arquivo criam um superbloco para armazenar informações críticas que protegem a integridade do sistema de arquivo.**

- Um superbloco pode conter:
 - O identificador de sistema de arquivo.
 - A localização dos blocos livres do dispositivo de armazenamento.
- Para diminuir o risco de perda de dados, a maioria dos sistemas de arquivo distribui cópias redundantes do superbloco pelo dispositivo de armazenamento.

13.4.2 Metadados

- **A operação de abertura de um arquivo retorna um descritor de arquivo.**
 - Um índice de inteiros não negativos dentro da tabela de arquivos abertos.
- **A partir desse ponto, o acesso ao arquivo é orientado pelo descritor de arquivo.**
- **Para permitir rápido acesso a informações específicas do arquivo, como permissões, a tabela de arquivos abertos em geral contém blocos de controle de arquivo, também chamados de atributos:**
 - Estruturas altamente dependentes do sistema que podem incluir o nome simbólico do arquivo, sua localização no armazenamento secundário, dados de controle de acesso e assim por diante.



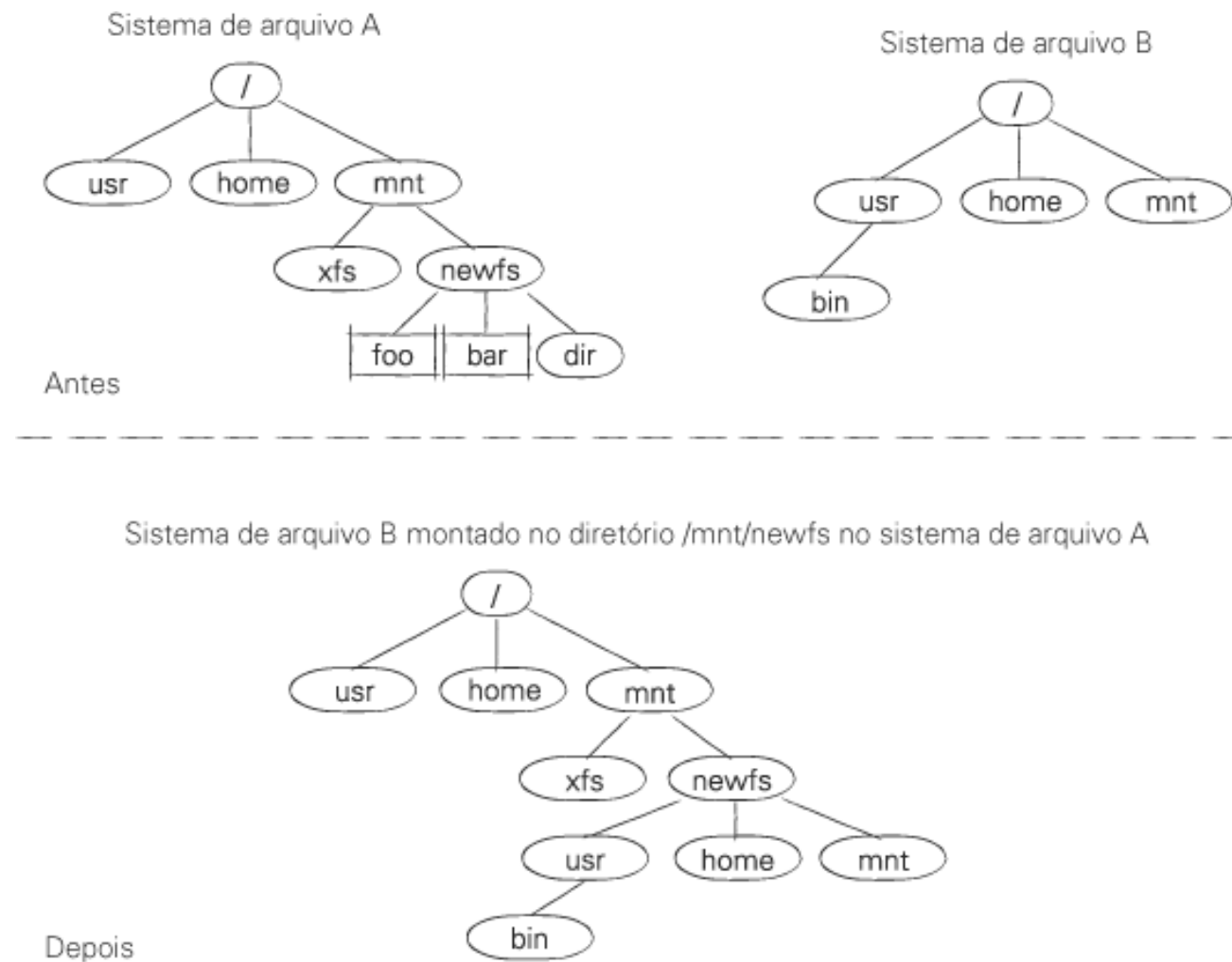
13.4.3 Montagem

- **Operação de montagem**
 - Associa vários sistemas de arquivo em um único espaço de nome de modo que possam ser referenciados de um único diretório-raiz.
 - Designa um diretório, denominado ponto de montagem, no sistema de arquivo nativo para a raiz do sistema de arquivo montado.
- **Os sistemas de arquivo gerenciam os diretórios montados por meio de tabelas de montagem:**
 - Contêm informações sobre a localização dos pontos de montagem e os dispositivos para os quais apontam.
- **Quando o sistema de arquivo nativo encontra um ponto de montagem, usa a tabela de montagem para determinar o dispositivo e o tipo de sistema de arquivo montado.**
- **Os usuários podem criar ligações flexíveis para sistemas de arquivo montados, mas não pode criar ligações estritas entre sistemas de arquivo.**



13.4.3 Montagem

Figura 13.5 Montagem de um sistema de arquivo.



13.6 Alocação de arquivos

■ Alocação de arquivos

- O problema de alocar e liberar espaço em armazenamento secundário é, de certa maneira, semelhante ao experimentado na alocação de memória principal sob multiprogramação de partição variável.
- Os sistemas de alocação contíguos em geral têm sido substituídos por sistemas de alocação não contíguos mais dinâmicos.
 - Os arquivos tendem a aumentar ou diminuir ao longo do tempo.
 - Os usuários raramente sabem com antecedência de que tamanho serão seus arquivos.

13.6.1 Alocação contígua de arquivos

■ Alocação contígua

- Coloca os dados dos arquivos em endereços contíguos no dispositivo de armazenamento.
- Vantagem
 - Registros lógicos sucessivos em geral estão fisicamente adjacentes uns aos outros.
- Desvantagens
 - Fragmentação externa.
 - Mau desempenho se os arquivos aumentarem ou diminuïrem ao longo do tempo.
 - Se um arquivo exceder o tamanho originalmente especificado e não houver nenhum bloco contíguo disponível, deverá ser transferido para uma nova área de tamanho adequado, o que exige operações adicionais de E/S.



13.6.2 Alocação de arquivo não contígua por lista encadeada

■ **Esquema de alocação de arquivo não contígua por lista encadeada por setor:**

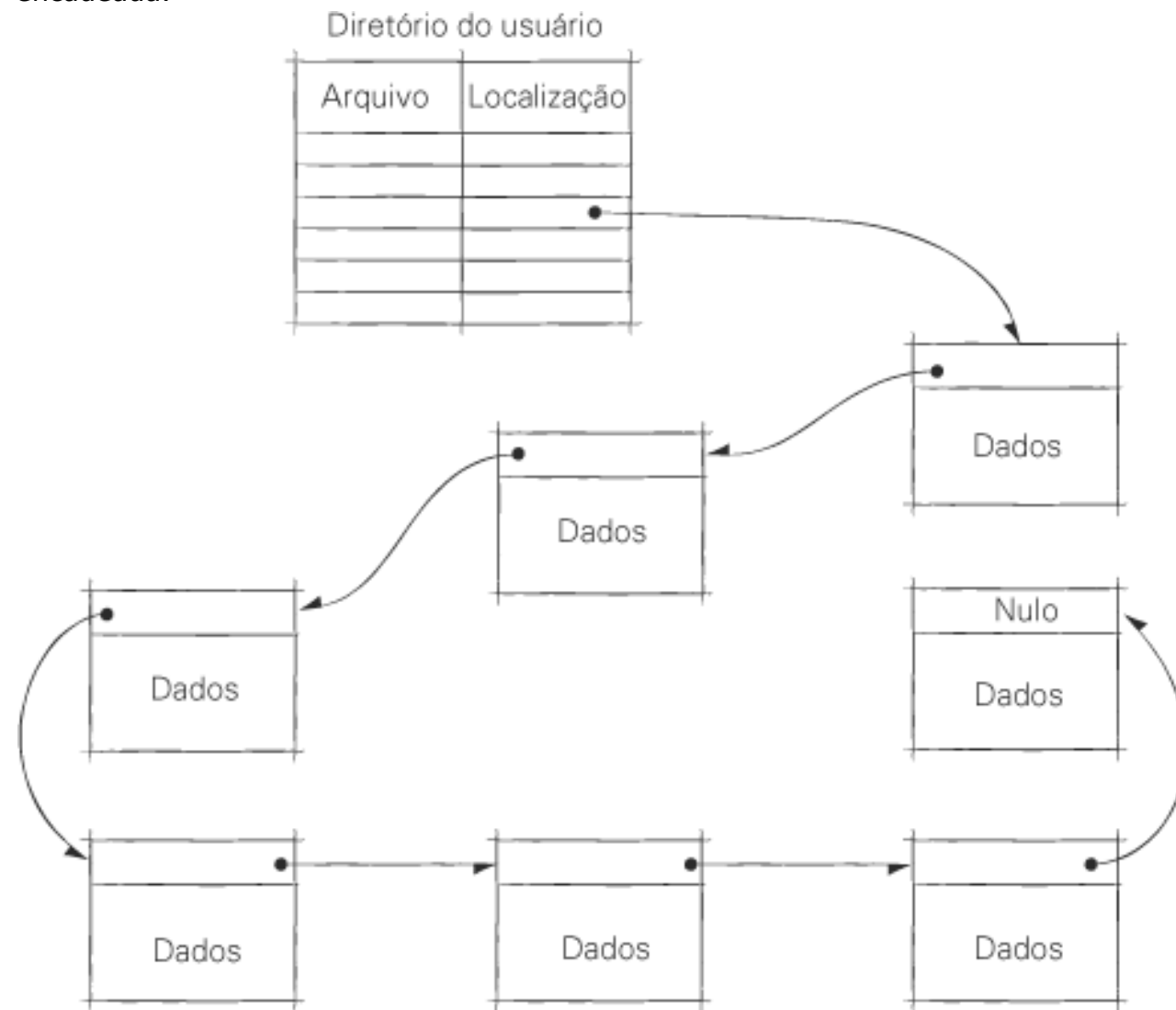
- Uma entrada de diretório aponta para o primeiro setor de um arquivo.
 - A porção de dados de um setor armazena o conteúdo do arquivo.
 - A porção do ponteiro aponta para o setor seguinte do arquivo.
- Os setores que pertencem a um arquivo comum formam uma lista encadeada.

13.6.2 Alocação de arquivo não contígua por lista encadeada

- **Ao executar a alocação de blocos, o sistema aloca blocos de setores contíguos (às vezes chamados de extensões).**
- **Encadeamento de blocos**
 - Entradas no diretório do usuário apontam para o primeiro bloco de cada arquivo.
 - Os blocos de arquivo contêm:
 - Um bloco de dados.
 - Um ponteiro para o bloco seguinte.

13.6.2 Alocação de arquivo não contígua por lista encadeada

Figura 13.6 Alocação não contígua de arquivo usando uma lista encadeada.



13.6.2 Alocação de arquivo não contígua por lista encadeada

- **Quando um registro for localizado**

- A cadeia tem de ser pesquisada desde o início.
- Se os blocos estiverem dispersos pelo dispositivo de armazenamento (o que é normal), o processo de pesquisa pode ser vagaroso, visto que a busca é feita bloco por bloco.

- **A inserção ou exclusão é feita modificando-se o ponteiro no bloco anterior.**

13.6.2 Alocação de arquivo não contígua por lista encadeada

■ Blocos grandes

- Podem provocar significativa fragmentação interna.
 - O arquivo não ocupa todo o espaço do bloco, desperdiçando espaço.

■ Blocos pequenos

- Podem fazer com que os dados de um arquivo sejam espalhados por vários blocos dispersos no dispositivo de armazenamento.
- Mau desempenho, visto que o dispositivo de armazenamento executa várias buscas para acessar todos os registros de um arquivo.

13.6.3 Alocação de arquivo tabular não contígua

■ Alocação de arquivo tabular não contígua

- Usa tabelas que armazenam ponteiros para blocos de arquivo.
 - Isso reduz o número de buscas demoradas necessárias para acessar um registro em particular.
- Entradas de diretório indicam o primeiro bloco de um arquivo.
- O número do bloco corrente é usado como um índice para a tabela de alocação de blocos para determinar a localização do bloco.
 - Se o bloco corrente for o último bloco do arquivo, a entrada da tabela de alocação de blocos será nula.

13.6.3 Alocação de arquivo tabular não contígua

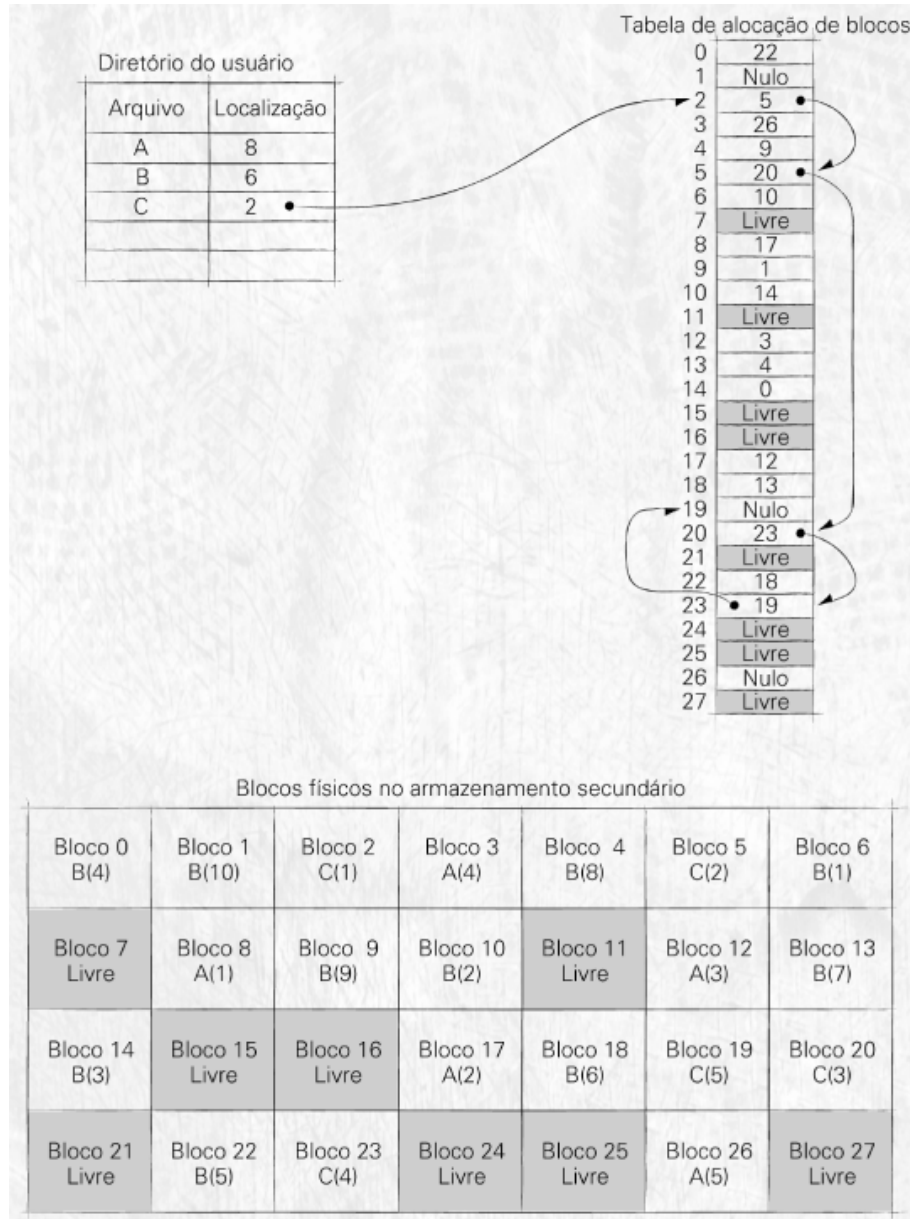


Figura 13.7 Alocação de arquivo tabular não contígua.

13.6.3 Alocação de arquivo tabular não contígua

- **Ponteiros que localizam dados de arquivo são armazenados em uma localização central.**
 - A tabela pode ser mantida em cache de modo que a cadeia de blocos que compõem um arquivo possa ser percorrida rapidamente.
 - Isso melhora os tempos de acesso.
- **Entretanto, para localizar o último registro de um arquivo:**
 - O sistema de arquivo talvez precise seguir vários ponteiros na tabela de alocação de blocos.
 - Isso poderia despendar um tempo significativo.
 - Mas ainda é relativamente rápido se a tabela estiver toda em cache.

13.6.3 Alocação de arquivo tabular não contígua

- **Quando um dispositivo de armazenamento contém muitos blocos:**
 - A tabela de alocação de blocos pode se tornar grande e fragmentada.
 - Isso diminui o desempenho do sistema de arquivo.
- **Uma implementação popular da alocação de arquivo tabular não contígua é o sistema de arquivos FAT da Microsoft.**

13.6.4 Alocação de arquivos não contígua indexada

■ **Alocação de arquivos não contígua indexada:**

- Todo arquivo dispõe de um bloco de índice ou de vários blocos de índice.
- Os blocos de índice contêm uma lista de ponteiros que apontam para os blocos de dados do arquivo.
- A entrada do diretório do arquivo aponta para seu bloco de índice, que pode reservar as últimas poucas entradas para armazenar ponteiros para mais blocos de índice, uma técnica denominada encadeamento.

13.6.4 Alocação de arquivos não contígua indexada

- **A principal vantagem do encadeamento de bloco de índice em relação às implementações de lista encadeada simples:**
 - A busca pode ocorrer nos próprios blocos de índice.
 - Os sistemas de arquivo em geral colocam os blocos de índice perto dos blocos de dados a que se referem, de modo que os blocos de dados possam ser acessados rapidamente depois que seu bloco de índice for carregado.

13.6.4 Alocação de arquivos não contígua indexada

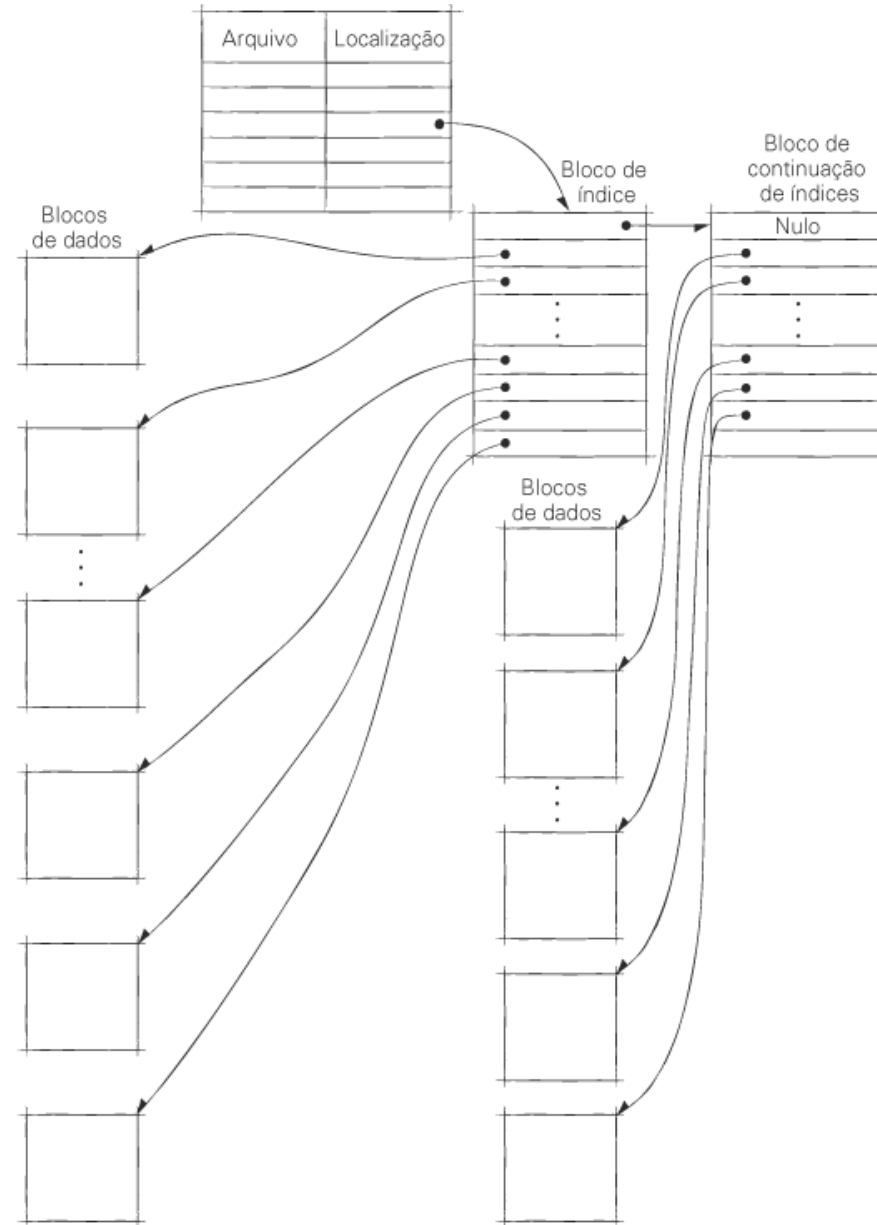


Figura 13.8 Encadeamento de blocos de índices.

13.6.4 Alocação de arquivos não contígua indexada

- **Os blocos de índice são chamados inodes (isto é, *index nodes* ou nós de índice) em sistemas operacionais baseados no UNIX.**

13.6.4 Alocação de arquivos não contígua indexada

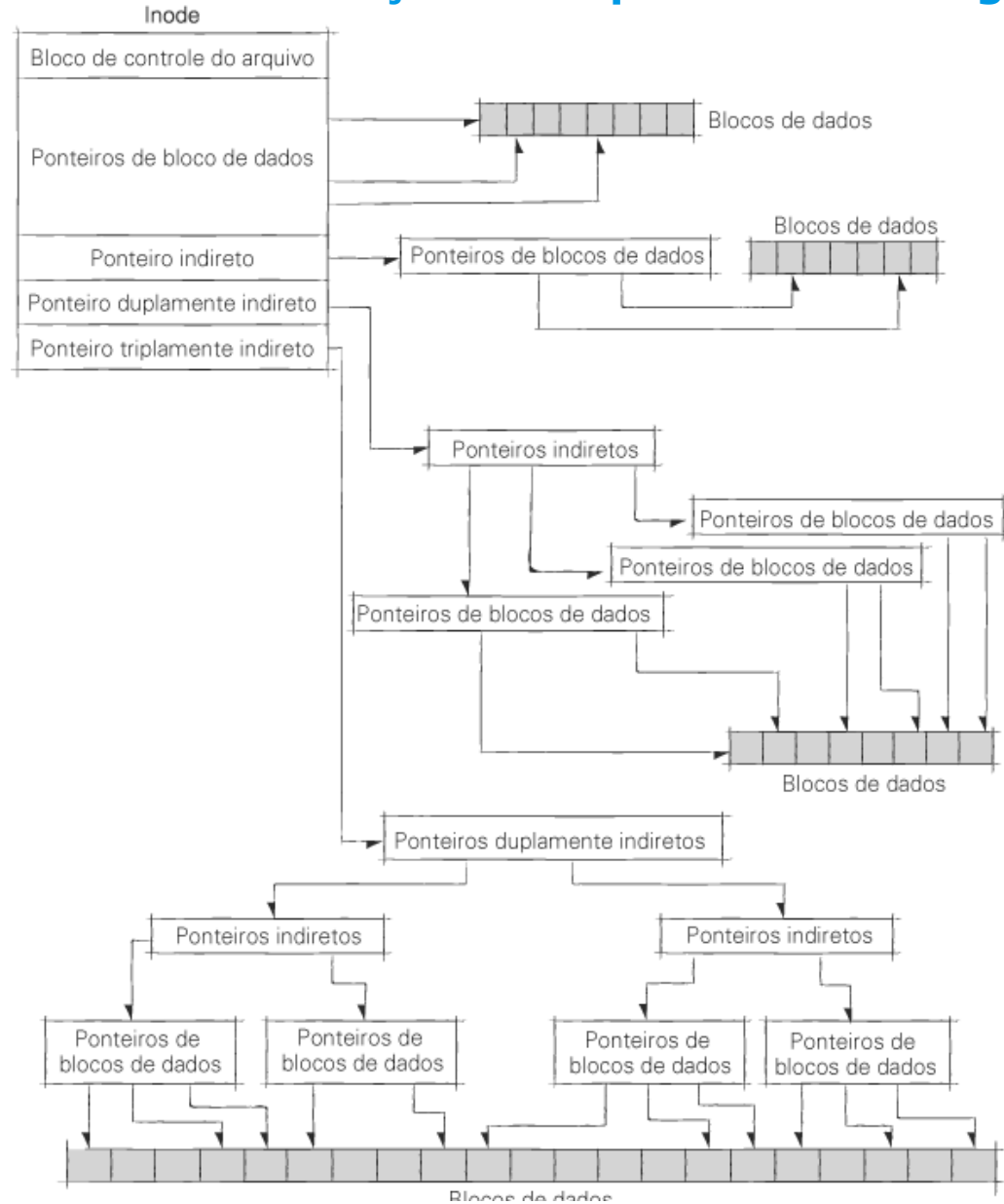


Figura 13.9 Estrutura do inode.

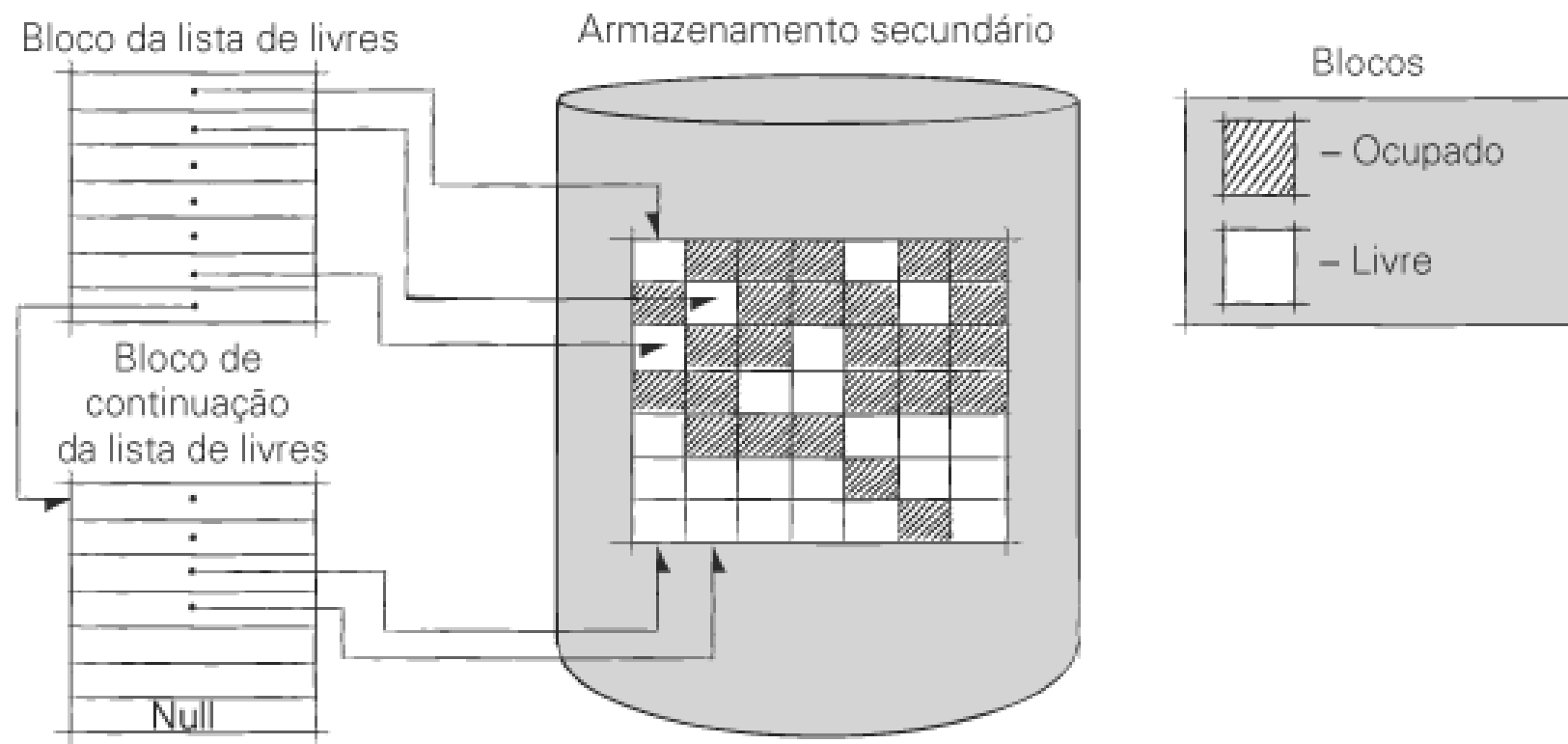
13.7 Gerenciamento de espaço livre

- **Alguns sistemas usam uma lista de livres para gerenciar o espaço livre do dispositivo de armazenamento.**
 - Lista de livres: lista encadeada de blocos que contêm a localização dos blocos livres.
 - Os blocos são alocados desde o começo da lista de livres.
 - Os blocos recém-liberados são anexados ao fim da lista.
- **Baixa sobrecarga para executar operações de manutenção da lista de livres.**
- **Os arquivos tendem a ser alocados em blocos não contíguos.**
 - Isso aumenta o tempo de acesso ao arquivo.



13.7 Gerenciamento de espaço livre

Figura 13.10 Gerenciamento de espaço livre usando uma lista de livres.

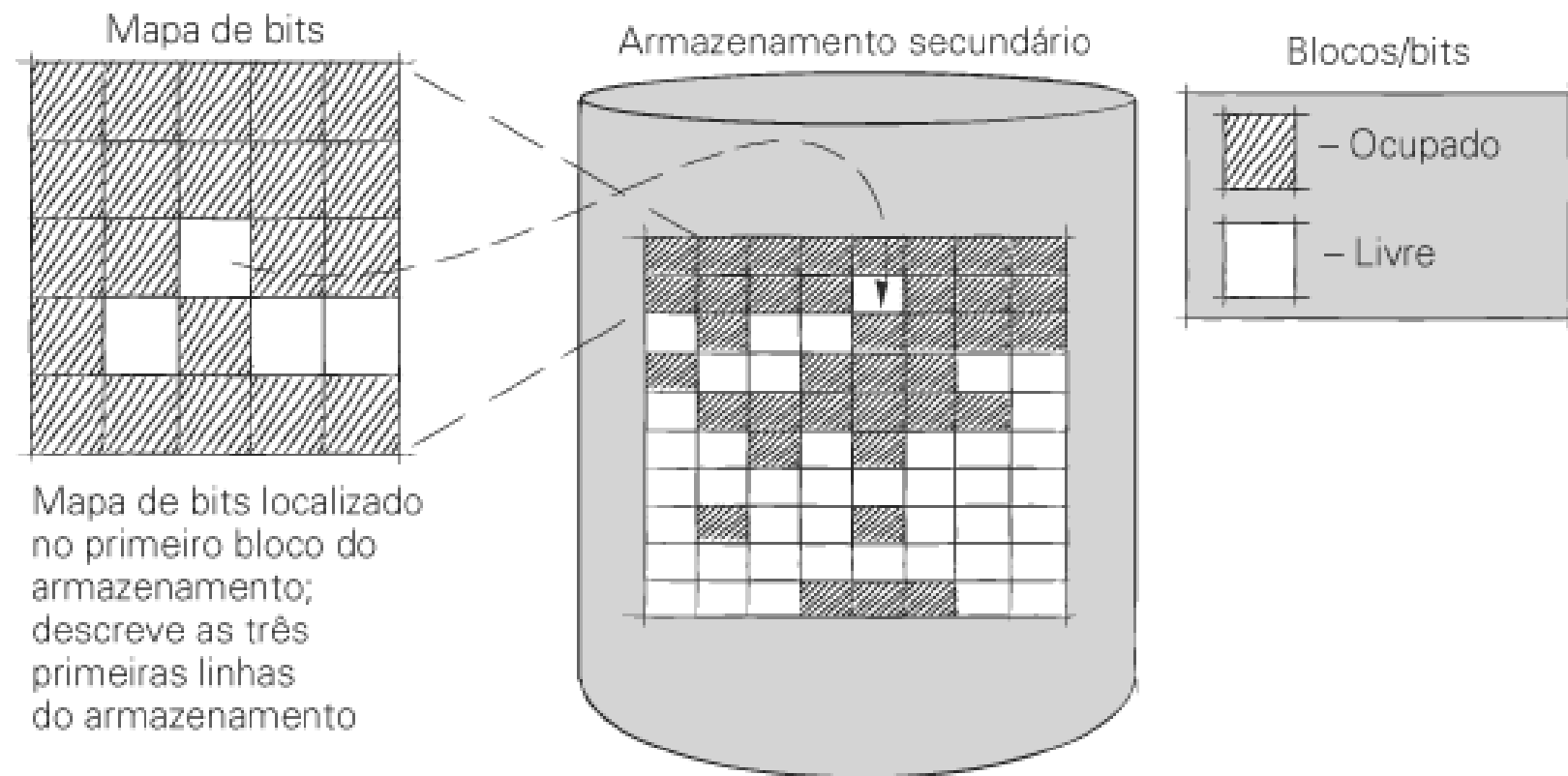


13.7 Gerenciamento de espaço livre

- **Um mapa de bits contém um bit para cada bloco na memória.**
 - *i*ésimo bit corresponde ao *i*ésimo bloco no dispositivo de armazenamento.
- **Vantagem do mapa de bits sobre a lista de livres:**
 - O sistema de arquivo consegue determinar rapidamente se existem blocos contíguos disponíveis em localizações específicas no armazenamento secundário.
- **Desvantagem do mapa de bits:**
 - O sistema de arquivo pode precisar pesquisar todo o mapa de bits para encontrar um bloco livre, caso em que a sobrecarga de execução é substancial.

13.7 Gerenciamento de espaço livre

Figura 13.11 Gerenciamento de espaço livre usando um mapa de bits.



13.8 Controle de acesso a arquivos

- **Os arquivos em geral são usados para armazenar dados sensíveis, como:**
 - Número de cartão de crédito
 - Senha
 - Número da previdência social
- **Portanto, devem contar com mecanismos para controlar o acesso dos usuários aos dados.**
 - Matriz de controle de acesso
 - Controle de acesso por classe de usuários

13.8.1 Matriz de controle de acesso

- **Matriz bidimensional de controle de acesso:**
 - A entrada a_{ij} é 1 se o usuário i tiver permissão de acesso ao arquivo j .
 - Do contrário, $a_{ij} = 0$.
- **Em uma instalação com grande número de usuários e grande quantidade de arquivos, essa matriz geralmente seria extensa e esparsa.**
- **Não é adequada para a maioria dos sistemas.**

13.8.1 Matriz de controle de acesso

Figura 13.12 Matriz de controle de acesso.

Usuário \ Arquivo										
	1	2	3	4	5	6	7	8	9	10
1	1	1	0	0	0	0	0	0	0	0
2	0	0	1	0	1	0	0	0	0	0
3	0	1	0	1	0	1	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0
5	1	1	1	1	1	1	1	1	1	1
6	0	0	0	0	0	1	1	0	0	0
7	1	0	0	0	0	0	0	0	0	1
8	1	0	0	0	0	0	0	0	0	0
9	1	1	1	1	0	0	0	0	1	1
10	1	1	0	0	1	1	0	0	0	1

13.8.2 Controle de acesso por classes de usuário

- **Uma técnica que requer um espaço consideravelmente menor é o controle de acesso de várias classes de usuários.**
- **Dentre as classes de usuários encontram-se:**
 - O proprietário do arquivo
 - Um usuário específico
 - Grupo
 - Projeto
 - Público
- **Dados de controle de acesso**
 - Podem ser armazenados como parte do bloco de controle de arquivo.
 - Em geral consomem uma quantidade insignificativa de espaço.



Referências Bibliográficas

- DEITEL, H. M.; DEITEL, P. J.; CHOFFNES, D. R.; Sistemas Operacionais: terceira edição. São Paulo: Pearson Prentice Hall, 2005.
Cap. 13.
- TANENBAUM, Andrew S.; Sistemas Operacionais Modernos . 3ed. São Paulo: Pearson Prentice Hall, 2010. Cap. 4.

