

# SISTEMAS DE ARQUIVOS

**Raí E. T. de Freitas**  
**[raiemmanuel50.50@gmail.com](mailto:raiemmanuel50.50@gmail.com)**

# INTRODUÇÃO

# INTRODUÇÃO

- Toda aplicação precisa armazenar e recuperar dados
- Um processo tem uma quantidade limitada de informações na RAM
  - Algumas aplicações usam pouca ram
    - Jogos antigos
      - <https://jovemnerd.com.br/noticias/games/doom-plataformas-inusitadas>
    - Firmware de um microondas
  - Algumas aplicações usam muita ram
    - Jogos atuais em 3d com ray tracing
      - GTA VI
    - Software CAD 3D

# INTRODUÇÃO

- Um problema dessa abordagem é que a RAM é volátil
  - Após o processo ser concluído, os dados são perdidos
- Muitas aplicações necessitam de armazenamento permanente
  - Banco de dados
    - ACID
      - D - durabilidade
    - Após uma transação ser confirmada, seus dados devem permanecer eternamente (ou até serem alterados) mesmo em caso de falha ou queda de energia
      - Backup, log e afins
    - Os dados devem permanecer no sistema por semanas, meses, anos, ...
    - Imagine perder os dados de um banco apenas porque o servidor reiniciou, falha no sistema, processo terminou de executar ...

# INTRODUÇÃO

- O terceiro problema é a necessidade de múltiplos processos utilizarem partes de uma informação ao mesmo tempo
  - Considere uma lista telefônica no espaço de endereçamento de um processo
    - Apenas o próprio processo pode acessar a informação
      - O compartilhamento deve ser explícito
    - A informação deve-se tornar independente de processo

# INTRODUÇÃO

- Três requisitos para armazenamento de informação por um longo prazo
  - Deve ser possível armazenar uma grande quantidade de informações
  - As informações devem sobreviver ao término de execução dos processos
  - Múltiplos processos devem poder acessar a informação ao mesmo tempo
- Discos magnéticos foram usados por décadas para atingir esses objetivos do armazenamento de longo prazo ->
  - SSD tem se tornado muito popular
    - Não possuem partes mecânicas que podem quebrar
    - Possuem um rápido acesso aleatório
  - Complementa ou substitui os HDs
- Fitas e discos ópticos não são tão usados mais
  - Desempenho inferior
  - Backup

# INTRODUÇÃO



# INTRODUÇÃO

- Considere que o hd/ssd é uma sequência linear de blocos que permite duas operações
  - Ler o bloco k;
  - Escrever no bloco k;
- Como encontrar as informações nesse array?
- Como impedir que um usuário leia os dados de outros?
  - DCL banco de dados
- Como saber quais blocos estão livres?
  - Lembram de mapa de bits e lista encadeada?



# INTRODUÇÃO

- O SO abstrai o processador ao criar a abstração chamada de processo
- O SO abstrai a memória RAM ao criar a memória e endereçamento virtual
- Como abstrair os dados do disco?
  - Abstração chamada de arquivo
    - .txt, .docx, .png, .mp3, ...
- Essas três abstrações juntas são as principais do SO
  - Processo
  - Memória virtual
  - Arquivos

# INTRODUÇÃO

- O que são arquivos?
  - Unidades lógicas de informação criadas pelos processos
    - Processo do Word
      - .docx
    - Processo do Photoshop
      - .png, .jpg
  - Arquivologia
  - Um disco consegue ter milhares ou milhões de arquivos
  - Cada arquivo é independente entre si
  - Considere um disco de 1TB. Quantas músicas de 5MB caberiam?
- Processos podem criar ou ler arquivos existentes ->
  - <https://hub.asimov.academy/tutorial/como-criar-e-deletar-arquivos-em-python/>



```
# Importando o módulo Path do pathlib
from pathlib import Path

# Definindo o caminho e o nome do arquivo
caminho_arquivo = Path('arquivo.txt')

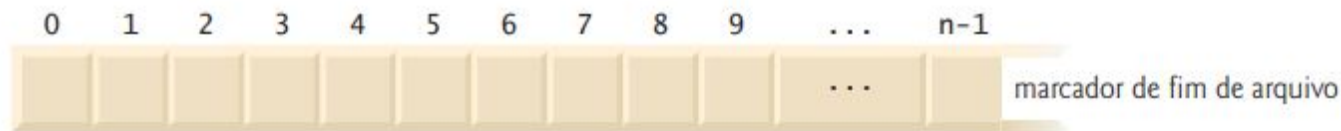
# Criando e abrindo o arquivo para escrita
with open(caminho_arquivo, 'w') as arquivo:
    arquivo.write('Olá, mundo Python!')
```

# INTRODUÇÃO

- A persistência das informações dos arquivos é independente do ciclo de vida do processo
  - O término do processo não deve apagar as informações
  - As informações devem ser persistentes
    - Serializable Java
  - As informações devem ser apagadas apenas quando o seu proprietário deliberadamente as apagar
- Dados armazenados em variáveis e arrays são temporários
  - Os dados são perdidos quando saem do escopo ou o programa termina a sua execução
- Para retenção de dados no longo prazo, os computadores usam arquivos
  - Dispositivos de armazenamento secundário
    - HD, SSD, Pen-drive

# INTRODUÇÃO

- Os dados mantidos nas mídias secundárias são chamados de persistentes
  - Continuam existindo para além da execução do programa
  - arquivos.java
- O Java vê os arquivos como um fluxo de bytes
  - Cada SO tem um mecanismo próprio de marcar o fim de arquivo
  - `\cr` e `\lf` - carriage return e line feed



---

**Figura 15.1** | Visualização do Java de um arquivo de  $n$  bytes.

# INTRODUÇÃO

- Serialização de dados
  - Ato de gravar (serializar) ou ler (deserializar) dados em algum tipo de armazenamento (arquivo no disco, ssd, banco de dados) ou transmitidos pela rede
- `serializacao.java`

# INTRODUÇÃO

- Arquivos são gerenciados pelo sistema operacional
  - Como são estruturados, nomeados, acessados, protegidos e afins
- A parte do sistema operacional que lida com arquivos é o sistema de arquivos

# INTRODUÇÃO

- Sob a ótica do usuário, o aspecto mais importante do sistema de arquivos é a forma como os arquivos aparecem
  - Forma de nomeação, proteção, operações permitidas, afins



# ARQUIVOS SOB A ÓTICA DO USUÁRIO

- Nomeação de arquivo
  - Um arquivo é um mecanismo de abstração
  - Mecanismo para armazenar dados no disco e depois recuperá-las
  - O ideal é isolar o usuário dos detalhes de como e onde as informações estão armazenadas e como os disco funcionam internamente
  - O processo é obrigado a nomear o arquivo criado
    - Após o término do processo, o arquivo continua persistente no disco
    - Outro processo pode acessar essas informações por meio desse nome
- Cada sistema operacional define suas regras de nomeação de arquivos
  - Caracteres do alfabeto
  - Caracteres especiais
  - Alguns são proibidos

# ARQUIVOS SOB A ÓTICA DO USUÁRIO

- Cada sistema operacional define suas regras de nomeação de arquivos
  - Tamanho do arquivo
    - 255 caracteres
  - Case sensitive
    - Diferencia maiúsculas e minúsculas
    - Unix é case sensitive
    - MS-DOS não é case sensitive
    - casa.txt, Casa.txt e CASA.txt
  - Separação com .
    - programa.c, casa.jpg, hino.mp3
      - A segunda parte é chamada de extensão
    - Unix e derivados não impõem o uso de extensões

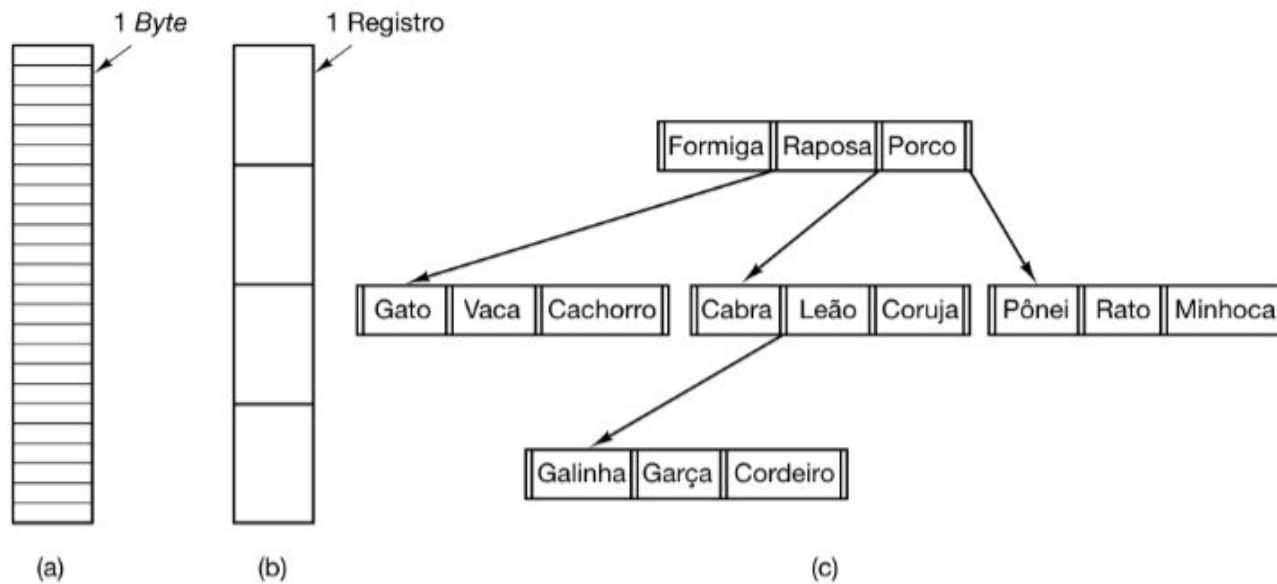
# ARQUIVOS SOB A ÓTICA DO USUÁRIO

Extensão	Significado
.bak	Cópia de segurança
.c	Código-fonte de programa em C
.gif	Imagem no formato <i>graphical interchange format</i>
.html	Documento em HTML da <i>world wide web</i>
.jpg	Imagem codificada segundo padrões JPEG
.mp3	Música codificada no formato MPEG (camada 3)
.mpg	Filme codificado no padrão MPEG
.o	Arquivo-objeto (gerado por compilador, ainda não ligado)
.pdf	Arquivo no formato PDF ( <i>portable document file</i> )
.ps	Arquivo PostScript
.tex	Entrada para o programa de formatação TEX
.txt	Arquivo de texto
.zip	Arquivo compactado

**Figura 4.1** Algumas extensões comuns de arquivos.

# ESTRUTURA DE ARQUIVOS

- Arquivos podem ser estruturados de três maneiras



**Figura 4.2** Três tipos de arquivos. (a) Sequência de *bytes*. (b) Sequência de registros. (c) Árvore.

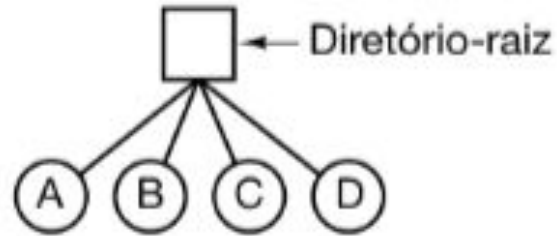
# DIRETÓRIOS

- Para controlar e organizar os arquivos, o sistema de arquivos mantém diretórios ou pastas
  - Internamente são arquivos também, porém especiais
- Sistemas de diretórios em nível único e hierárquico

# SISTEMAS DE DIRETÓRIOS EM NÍVEL ÚNICO

- Um único diretório contendo todos os arquivos
  - Diretório-raiz
- Implementado nos primeiros computadores pessoais
  - Havia apenas um usuário
- Um diretório com 4 arquivos ->
- Quais as vantagens desse modo de organização?
  - Simplicidade
  - Facilidade em encontrar os arquivos rapidamente
    - Só há um local para procurar
- Ainda utilizado em sistemas embarcados simples
  - Câmeras digitais

# SISTEMAS DE DIRETÓRIOS EM NÍVEL ÚNICO



**Figura 4.7** Um sistema de diretório em nível único contendo quatro arquivos.

# SISTEMAS DE DIRETÓRIOS HIERÁRQUICO

- Seria difícil usar diretórios únicos nos sistemas modernos
  - Excesso de arquivos
  - Ideal seria o agrupamento de arquivos relacionados em um mesmo local
  - Fotos de um álbum, conjunto de arquivos fonte de um programa, entre outros
- Uma solução é a criação de uma hierarquia
  - Árvore de diretórios
  - Não há limites de diretórios
    - Pode-se criar quantos agrupamentos lógicos forem necessários
    - Organizar seus arquivos de forma natural



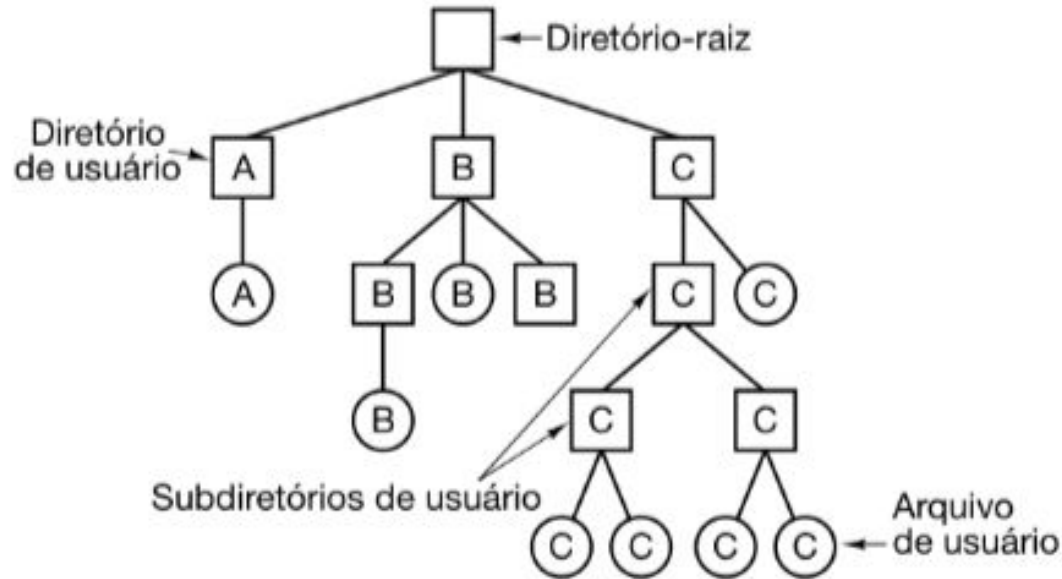
# SISTEMAS DE DIRETÓRIOS HIERÁRQUICO

Windows  
UNIX  
MULTICS

\usr\ast\caixapostal  
/usr/ast/caixapostal  
>usr>ast>caixapostal

- E se diversos usuários compartilharem um servidor de arquivos?
  - Cada usuário pode ter seu diretório-raiz privado ->
  - A capacidade de criação arbitrária de diretórios permite uma estruturação complexa e poderosa
  - Usado pelos sistemas de arquivos modernos
- Como organizar os nomes dos diretórios e arquivos?
  - Nome de caminho absoluto
    - Consiste no caminho do diretório raiz até o arquivo
    - /disciplinas/sistemas operacionais/sistemas de arquivos/aula.pdf
    - Sempre começam no diretório raiz e são únicos
    - Unix usa /
    - Windows usa \

# SISTEMAS DE DIRETÓRIOS HIERÁRQUICO



**Figura 4.8** Um sistema hierárquico de diretórios.

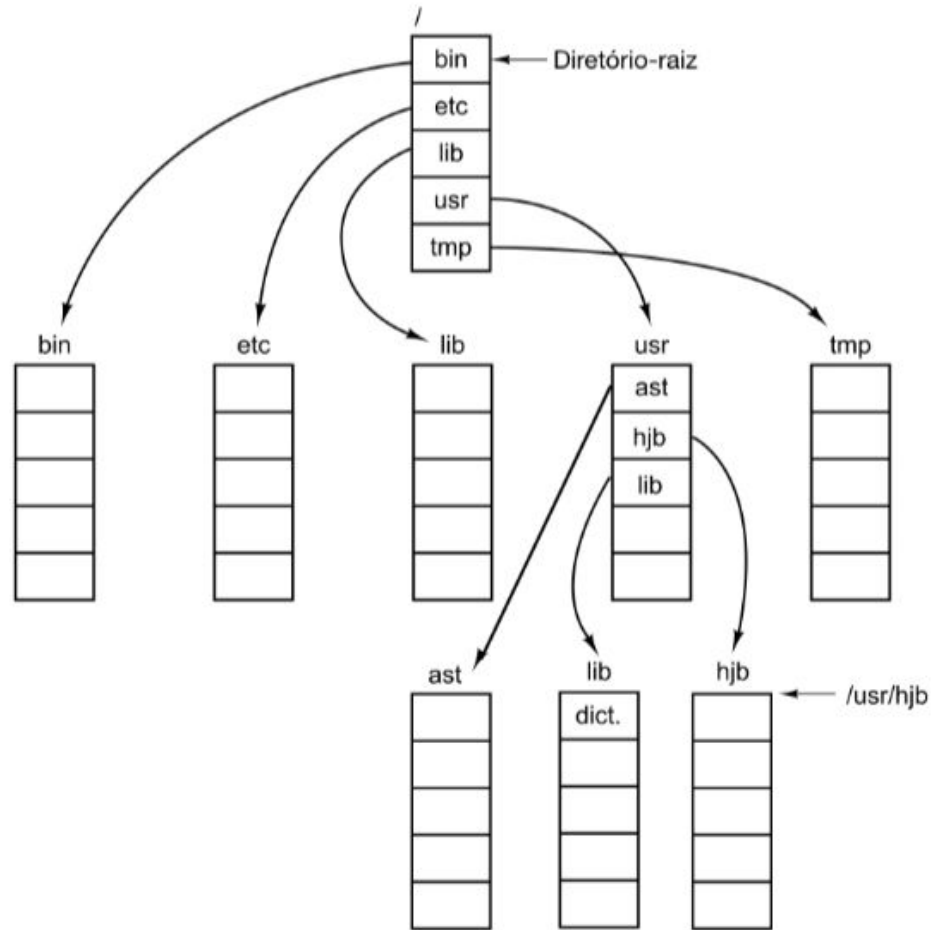
# SISTEMAS DE DIRETÓRIOS HIERÁRQUICO

Windows  
UNIX  
MULTICS

\usr\ast\caixapostal  
/usr/ast/caixapostal  
>usr>ast>caixapostal

- Como organizar os nomes dos diretórios e arquivos?
  - Nome de caminho absoluto
    - Sempre que o caminho começar com o separador (/ , \ ou >) significa que é um caminho absoluto
  - Nome de caminho relativo
    - Usado em conjunto da ideia de diretório atual
    - Define-se um diretório de trabalho
      - Os caminhos partem desse diretório
    - Considere o absoluto /user/desktop/imagens/foto.jpg
      - O diretório de trabalho /user/desktop
      - Pode-se referenciar o arquivo relativamente com imagens/foto.jpg
    - . e .. referem-se ao diretório atual e o diretório pai

# SISTEMAS DE DIRETÓRIOS HIERÁRQUICO



**Figura 4.9** Uma árvore de diretórios UNIX.

# OPERAÇÕES COM DIRETÓRIOS

- Create

- Criação de um diretório
- Inicialmente vazio
  - Lembre do . e ..
- mkdir

- Delete

- Remoção de um diretório
- Apenas diretórios vazios podem ser removidos
  - Apenas . e .. é considerado vazio

- Opendir

- Ler diretório
- Listar os arquivos do diretório
- Antes da leitura é necessário a abertura

# OPERAÇÕES COM DIRETÓRIOS

- `Closedir`
  - Fechar o diretório após usado
- `Link`
  - Permite um arquivo aparecer vários diretórios
  - Faz referência a um arquivo previamente existente

# IMPLEMENTAÇÃO DE SISTEMAS DE ARQUIVOS

- Visão do usuário preocupa-se apenas com a nomeação de arquivos, operações, árvores de diretórios
- Como os arquivos e diretórios estão armazenados?
- Como o espaço em disco é gerenciado?
- Como tudo funciona eficientemente?

# ESQUEMA DO SISTEMAS DE ARQUIVOS

- O sistema de arquivos é armazenado em disco
- O disco pode ser dividido em diversas partições lógicas
  - Cada partição pode ter autonomia de possuir um diferente sistema de arquivos
    - Gerenciador de disco windows



# IMPLEMENTAÇÃO DE ARQUIVOS

- Como controlar quais blocos do disco possuem qual arquivo?
- Alocação contígua
  - Modo mais simples
  - Armazena cada arquivo como uma sequência contígua de blocos do disco
  - Em disco com blocos de 1kB, um arquivo de 50kb ocuparia 50 espaços consecutivos
    - Em blocos de 2kB?
  - Blocos inteiro são usados
    - Desperdício pode surgir
  - Fácil implementação
    - Gerencia apenas endereço em disco do primeiro bloco e o tamanho
    - Fácil leitura
      - Todas as partes estão contíguas
      - Uma única leitura consegue ler tudo

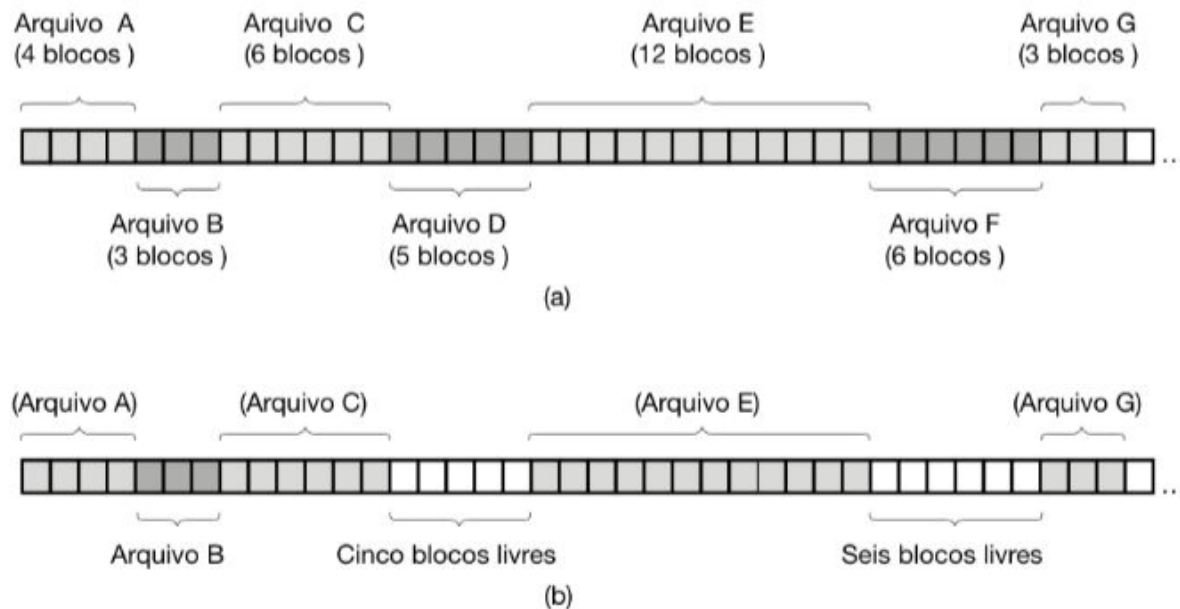
# IMPLEMENTAÇÃO DE ARQUIVOS

- Como controlar quais blocos do disco possuem qual arquivo?
- Alocação contígua
  - Fragmentação com o tempo
    - Arquivos removidos liberam o espaço entre blocos alocados
  -

# IMPLEMENTAÇÃO DE ARQUIVOS

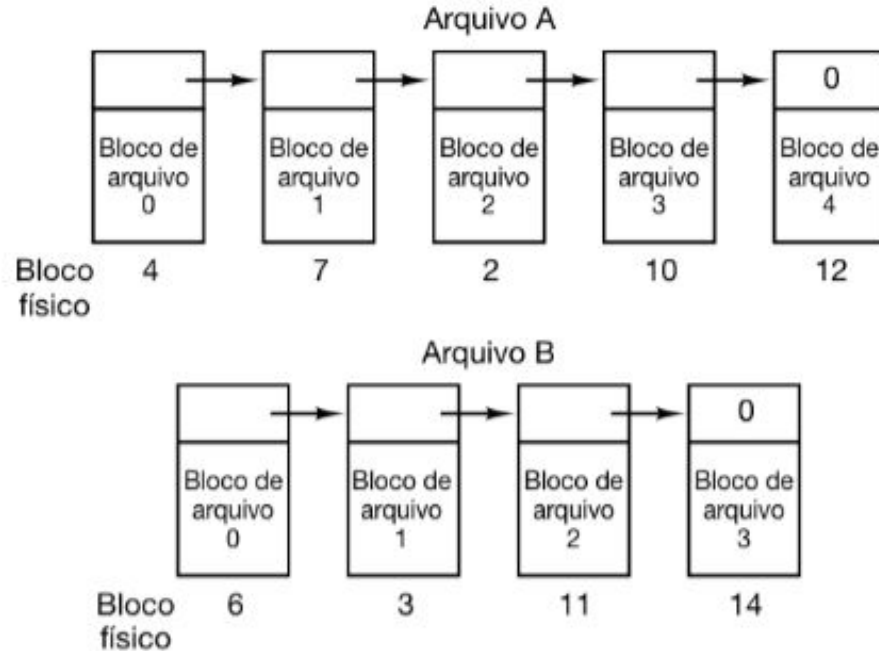
- Como controlar quais blocos do disco possuem qual arquivo?
- Alocação por lista encadeada
  - Lista encadeada de blocos do disco
  - Não há fragmentação externa
    - Somente interna se os dados não completarem o bloco inteiro
  - Endereço do primeiro bloco
  - Leitura mais lenta em relação ao contíguo simples
    - Lê um bloco de cada vez
  - Ponteiros ocupam espaço

# IMPLEMENTAÇÃO DE ARQUIVOS



**Figura 4.12** (a) Alocação contígua de espaço de disco para sete arquivos. (b) O estado do disco após os arquivos *D* e *F* terem sido removidos.

# IMPLEMENTAÇÃO DE ARQUIVOS



**Figura 4.13** Armazenando um arquivo como uma lista encadeada de blocos de disco.

1 - Em um sistema de arquivos de um sistema operacional, o nível mais alto da estrutura de diretórios é chamado de diretório:

- A) Raiz.
- B) Origem.
- C) Pai.
- D) Vértice.

2 - Dentro da arquitetura de um sistema operacional, o módulo do sistema de arquivos é crucial para a organização, o armazenamento e o acesso aos dados. Considerando suas responsabilidades fundamentais, qual das seguintes descrições NÃO se alinha com as funções primárias desse módulo?

- A) Aloca recursos de processamento para facilitar a execução de aplicações.
- B) Mantém a integridade dos dados e a segurança através de permissões e criptografia.
- C) Define protocolos para a nomenclatura, o armazenamento e a proteção de arquivos.
- D) Supervisiona a organização, o armazenamento seguro e o acesso aos dados em dispositivos de armazenamento.

3 - Suponha que você está participando de um workshop universitário sobre sistemas operacionais, o instrutor destaca a importância dos sistemas de arquivos. Durante uma sessão interativa, ele utiliza um exemplo prático para ilustrar a organização eficiente de dados. Indique qual é a função principal de um sistema de arquivos em um sistema operacional.

- A) Gerenciar a velocidade do processador.
- B) Armazenar e organizar dados em arquivos e diretórios.
- C) Controlar a temperatura do hardware.
- D) Executar operações matemáticas complexas.
- E) Monitorar a largura de banda da rede.



4 - A arquitetura de um sistema operacional refere-se à maneira como os componentes do sistema, como processador, memória e dispositivos de entrada/saída, estão organizados e interagem entre si. Com base nessas informações, julgue o item seguinte. Apenas um tipo de sistema de arquivos pode ser utilizado em um único disco rígido, e não é possível haver múltiplas partições com diferentes sistemas de arquivos no mesmo disco.

Certo ou Errado?

<https://www.qconcursos.com/questoes-de-concursos/questoes/779cba78-70?q=Q30096>