

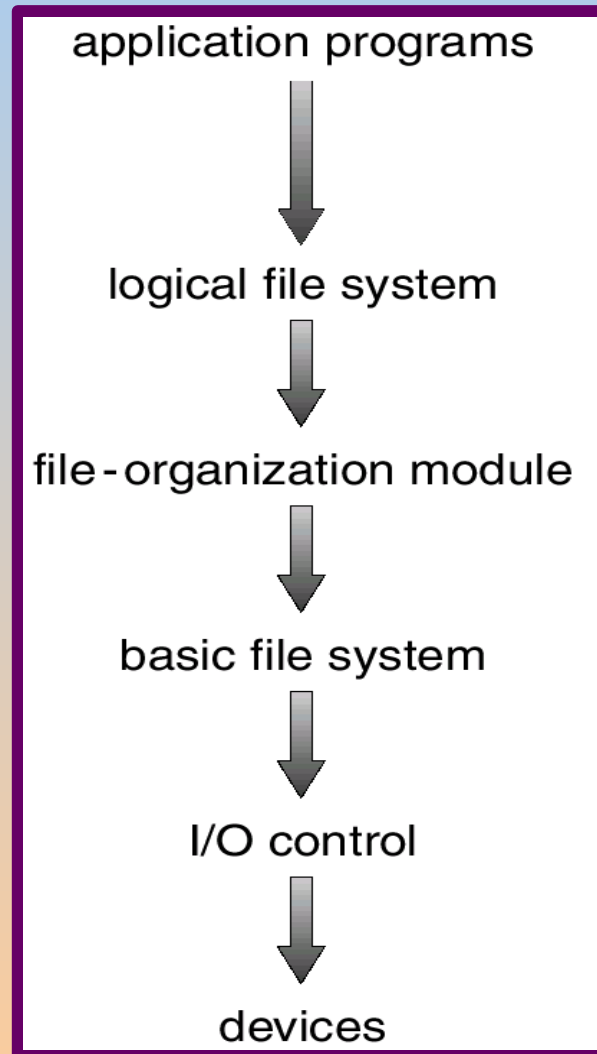
Implementação do Sistema de Arquivos

- Estrutura do sistema de arquivos
- Implementação do sistema de arquivos
- Implementação de diretórios
- Métodos de alocação
- Gerenciamento de espaço livre
- Eficiência e performance
- Recuperação de erros
- Sistemas de arquivos baseados em Log
- NFS

Estrutura de Sistemas de Arquivos

- Estrutura de arquivos
 - ◆ Unidade de armazenamento lógico
 - ◆ Coleção de informações relacionadas
- Sistemas de arquivos são armazenados em memória secundária (discos).
- Sistemas de arquivos são organizados, normalmente, em camadas.
- *Bloco de controle de arquivo* – Estrutura de armazenamento consistindo de informações sobre um arquivo.

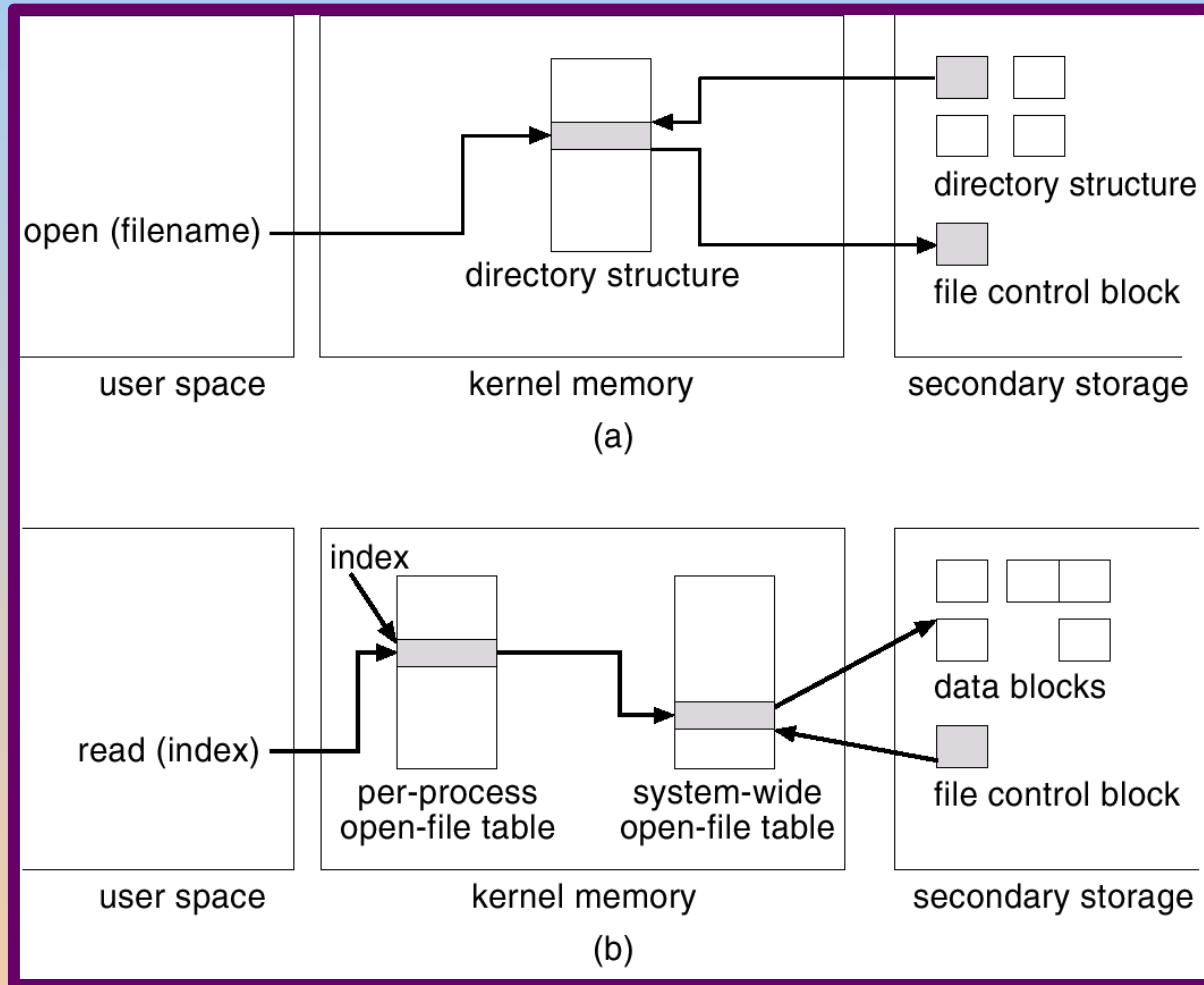
Sistema de arquivos organizado em camadas



Um bloco de controle típico de um arquivo

file permissions
file dates (create, access, write)
file owner, group, ACL
file size
file data blocks

Estruturas de sistema de arquivos em memória



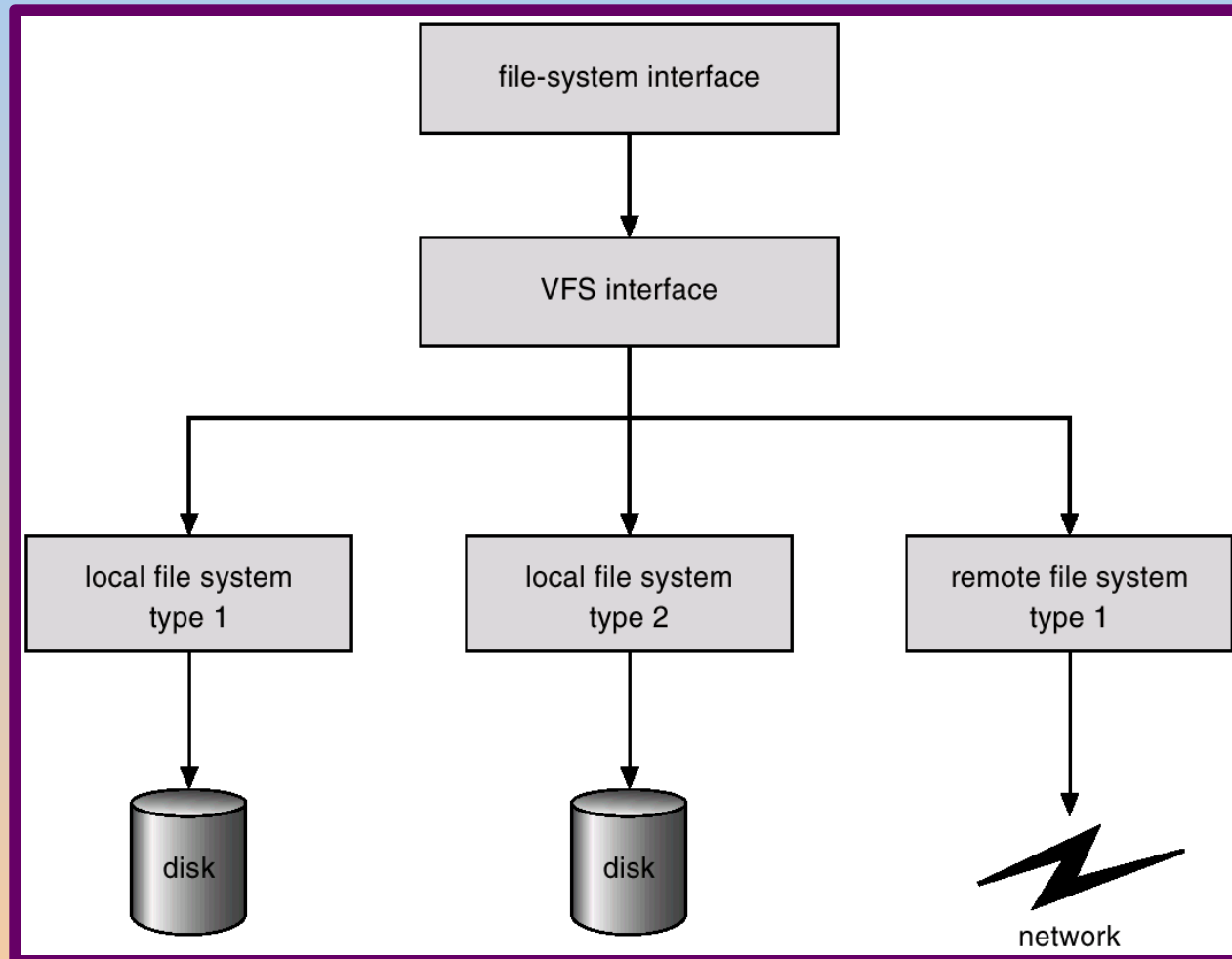
Abertura de arquivo

Leitura de arquivo

Sistema virtual de arquivos

- Virtual File Systems (VFS) disponibilizam uma maneira, baseada em orientação a objetos, de se implementar sistemas de arquivos.
- VFS permite que a interface de chamada de sistema (system call) possa ser usada por diferentes tipos de sistemas de arquivos.

Visão esquemática de um sistema virtual de arquivos



Implementação de diretórios

- Lista linear de nomes de arquivos com apontados para seus blocos de dados.
 - ◆ Simples de implementar
 - ◆ Busca muito custosa

- Tabela de hashing –
 - ◆ Diminui o tempo de busca por diretório
 - ◆ *Tamanho fixo*

Métodos de alocação

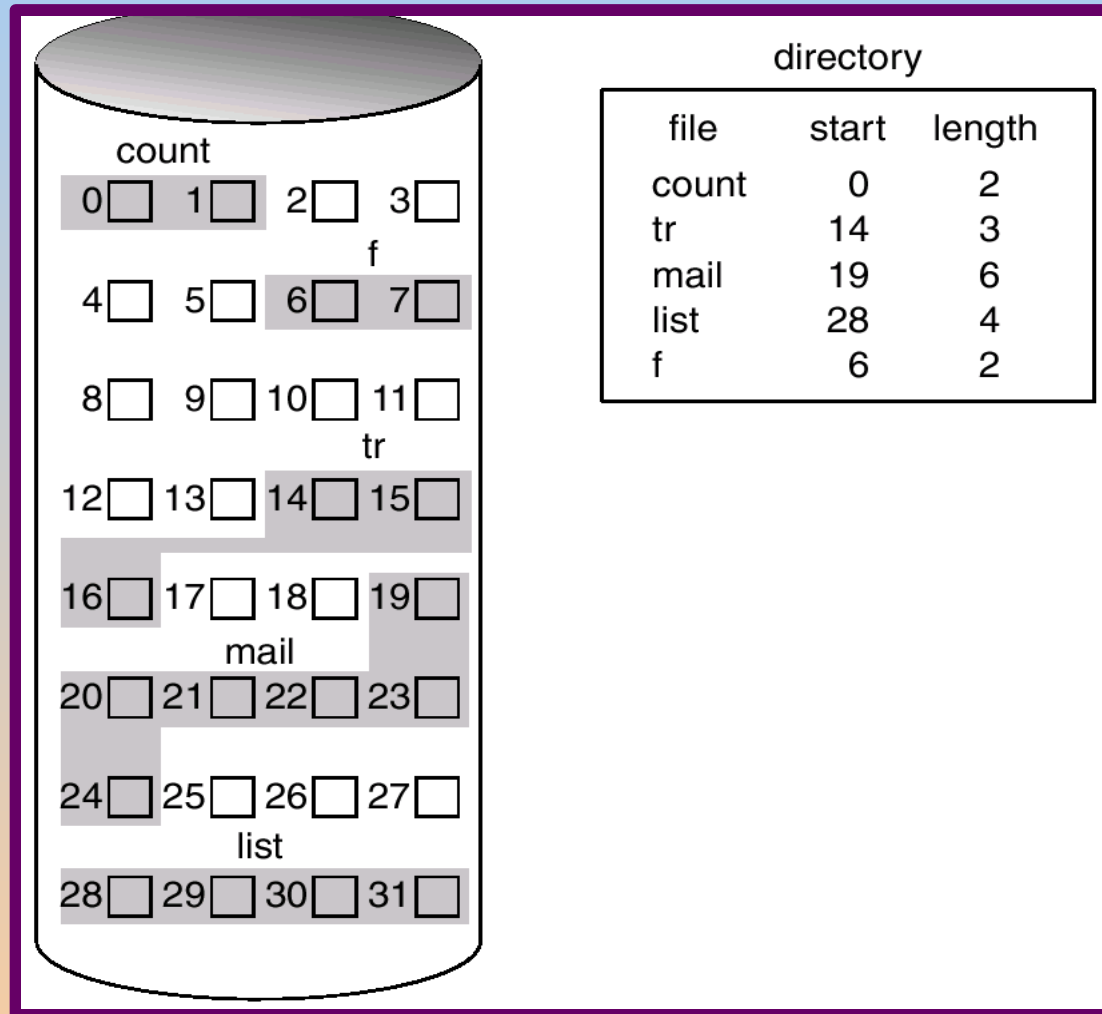
Um método de alocação diz respeito a como os blocos de um arquivo são alocados:

- Alocação contígua
- Alocação ligada
- Alocação indexada

Alocação contígua

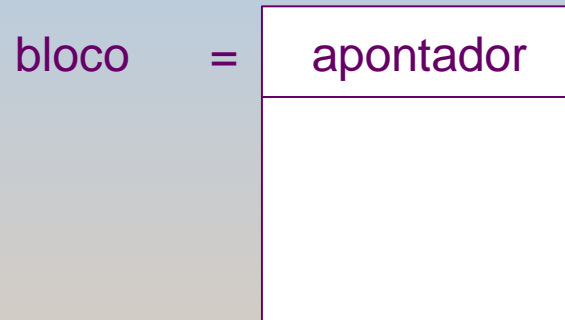
- Cada arquivo ocupa um conjunto contíguo de blocos no disco.
- Simples – somente a localização inicial (block #) e tamanho (número de blocos) são necessários.
- Acesso aleatório.
- Arquivos não podem crescer de tamanho.

Alocação contígua em um disco



Alocação ligada

- Cada arquivo é uma lista ligada de blocos de disco: blocos podem ser alocados em qualquer lugar do disco.

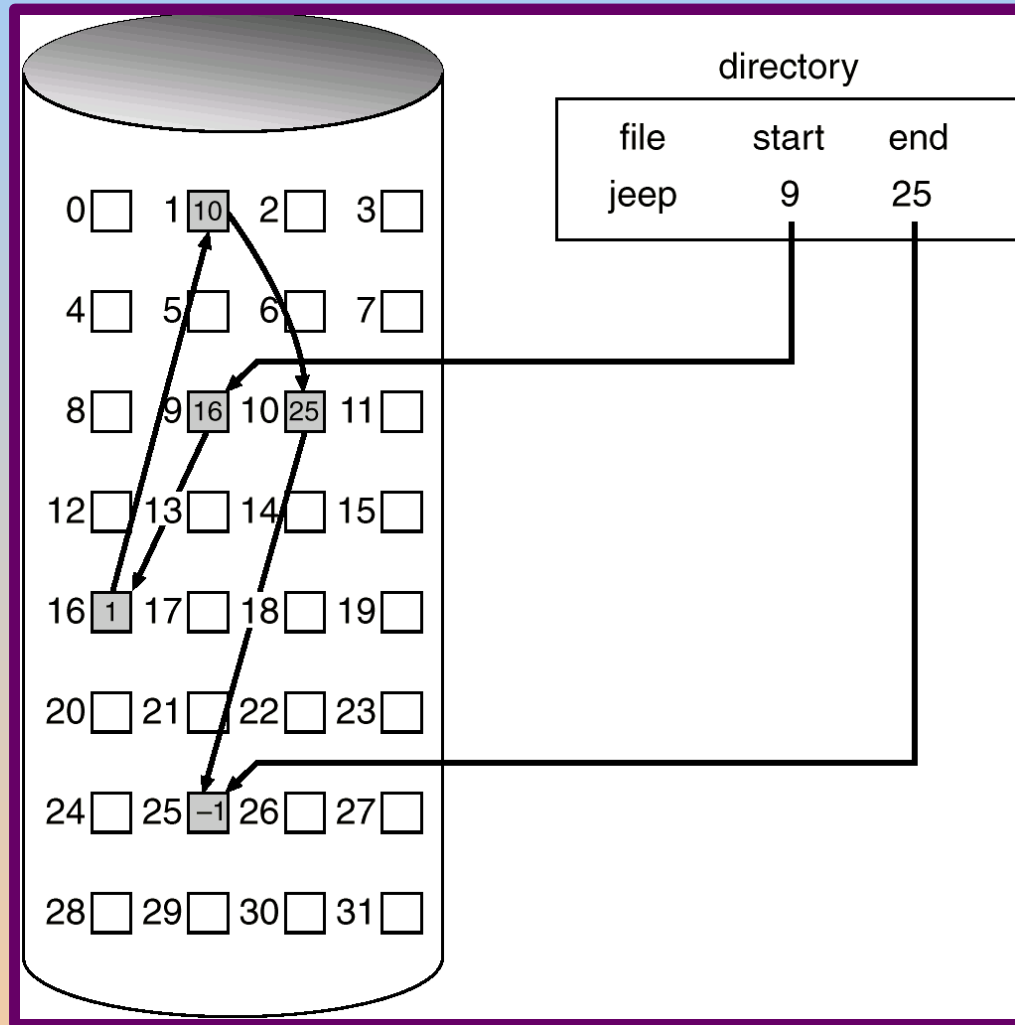


Alocação ligada (Cont.)

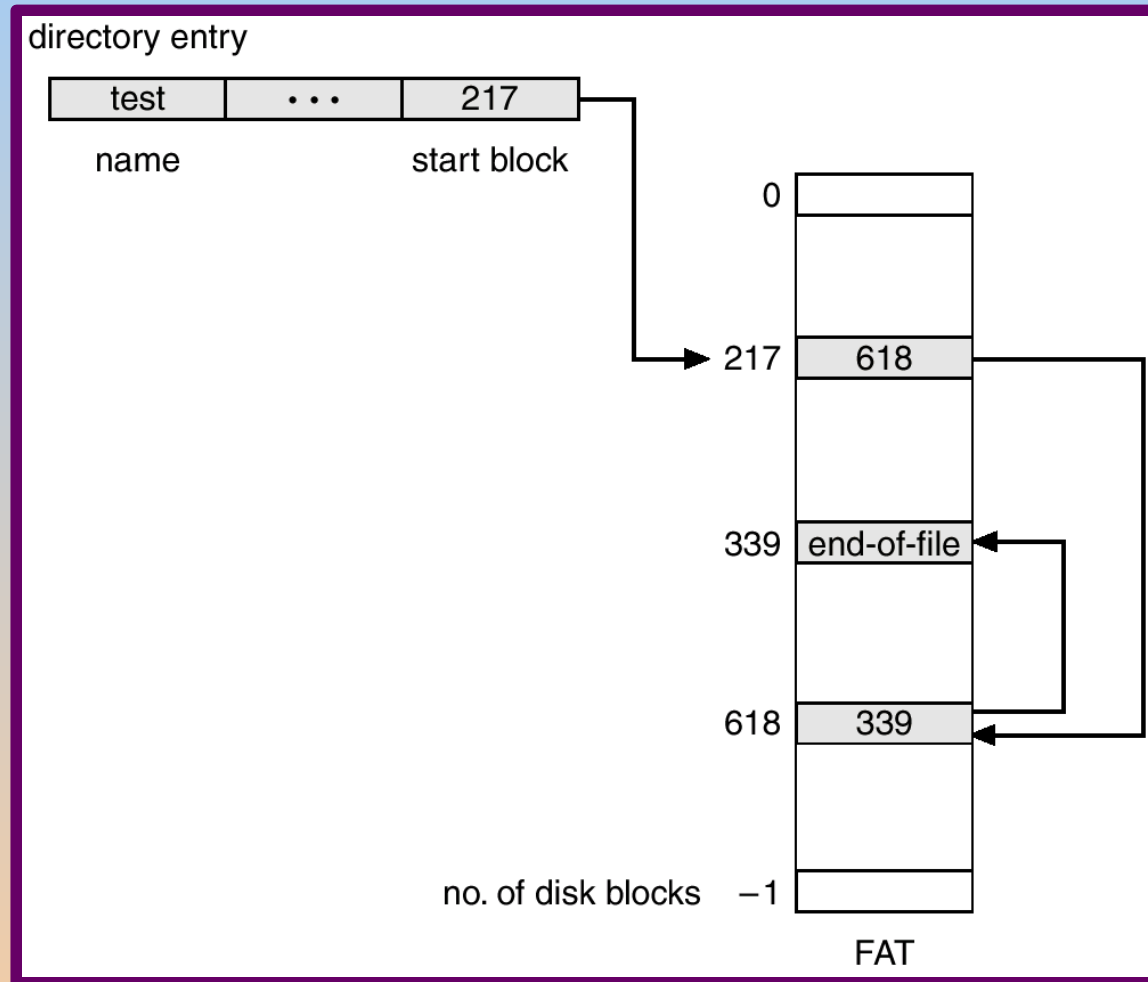
- Simples – Necessita somente do endereço inicial
- Não permite acesso aleatório

File-allocation table (FAT) – alocação de disco usada pelo MS-DOS.

Alocação ligada



Estrutura básica FAT



Alocação indexada

- Coloca todos os apontadores num único bloco. Chamado bloco de índice.
- Produz uma visão lógica.

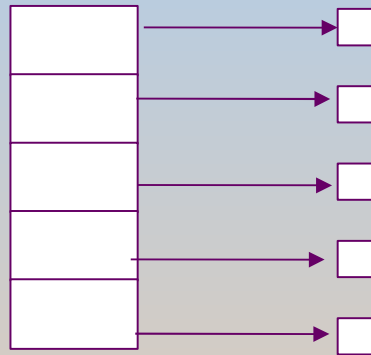
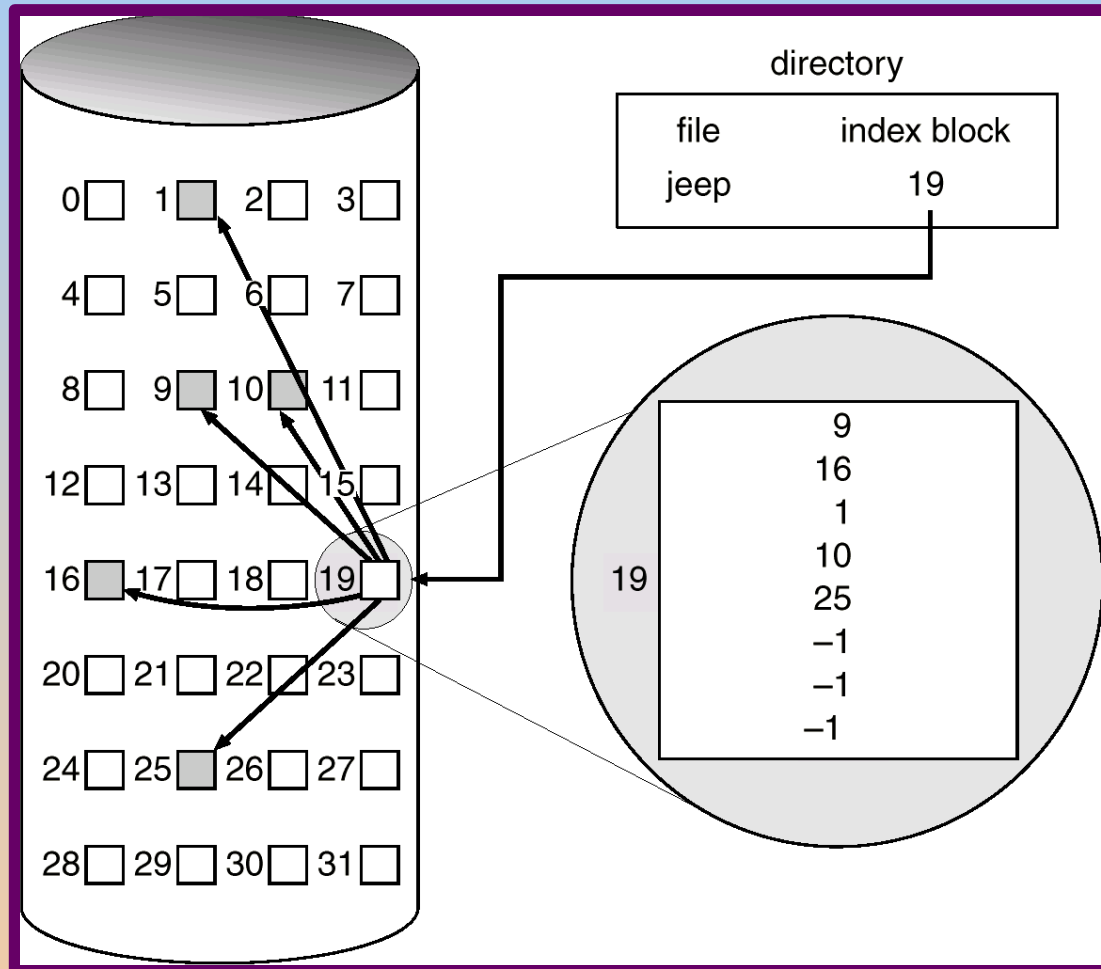
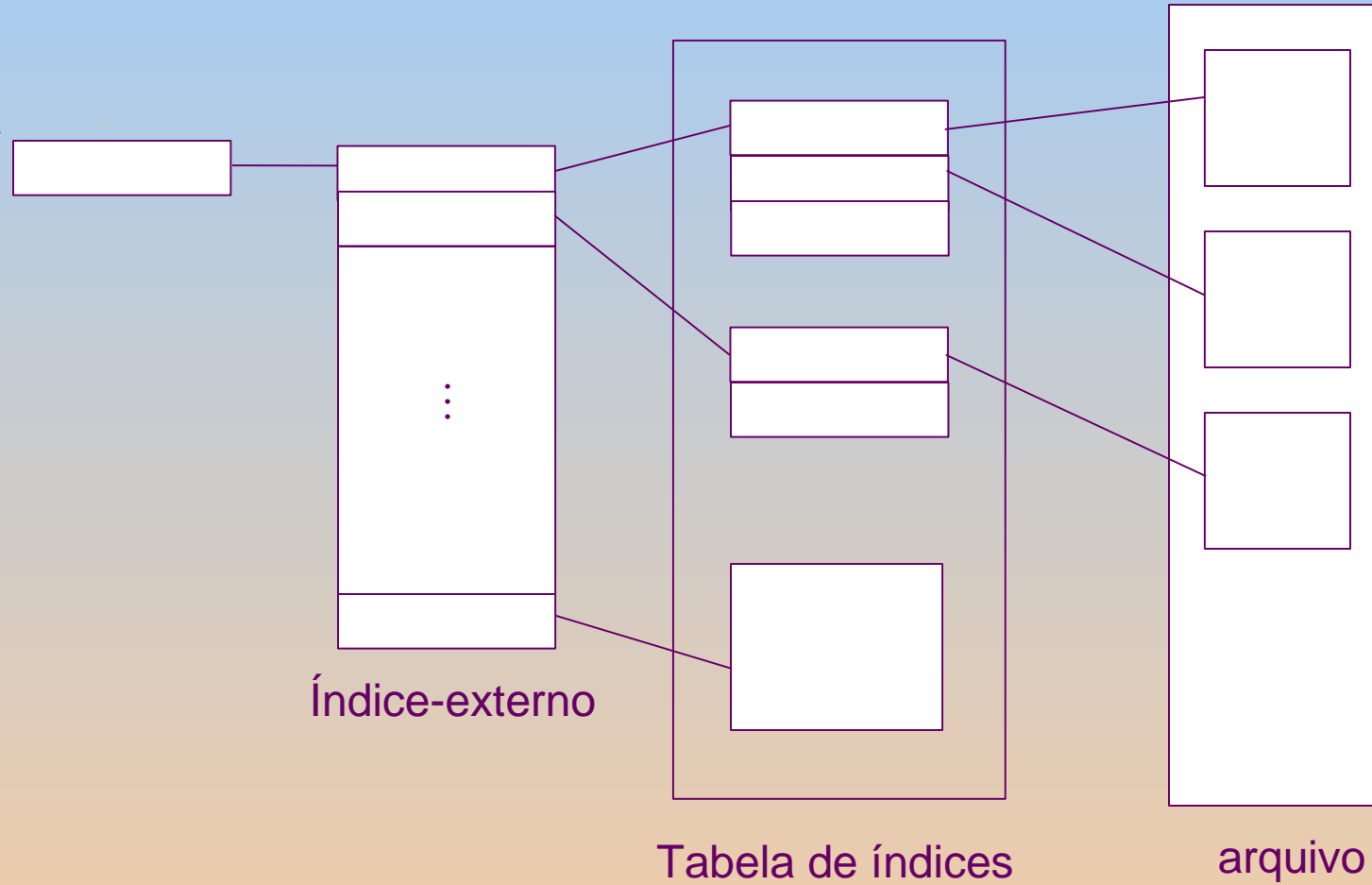


Tabela de índice

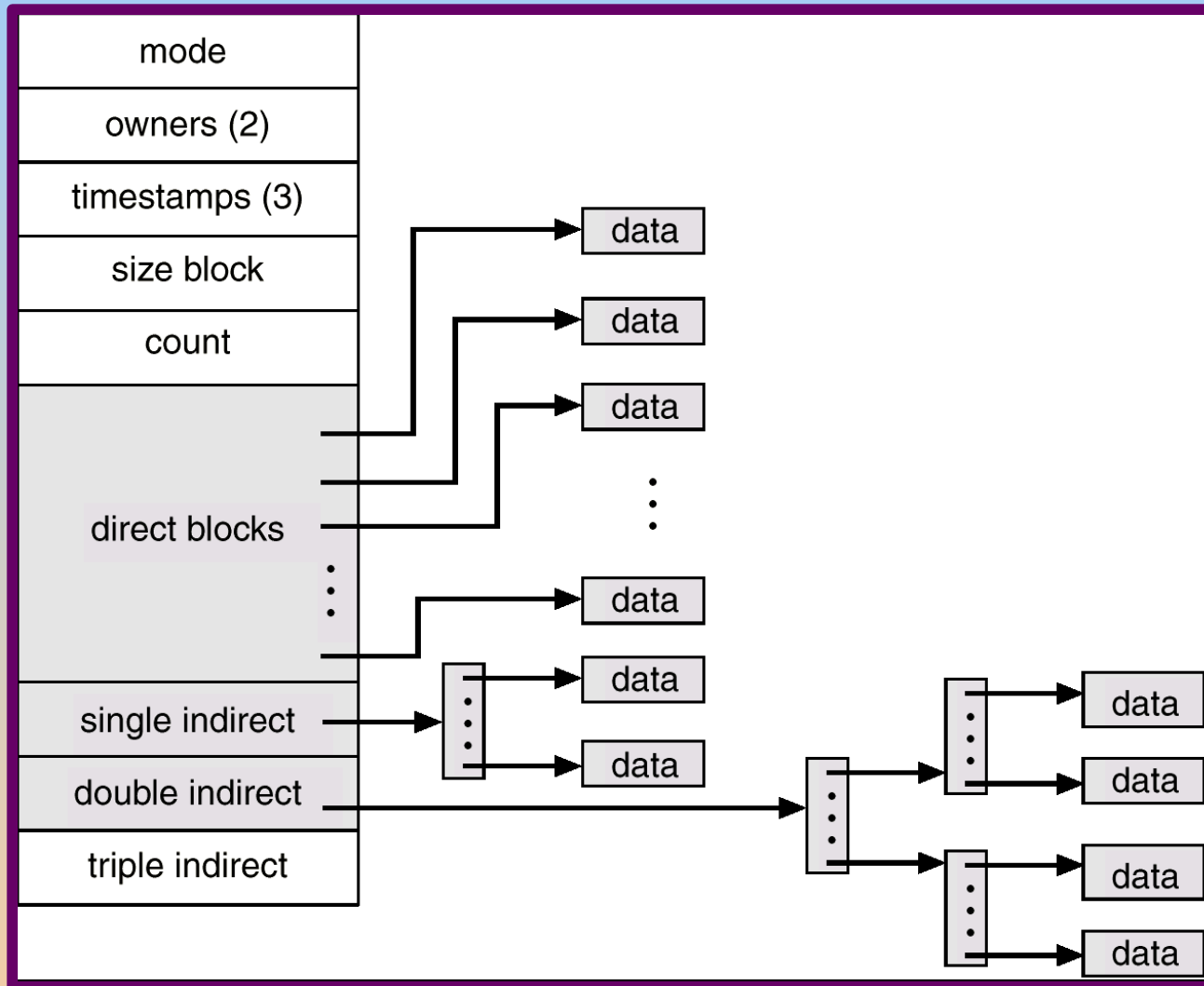
Exemplo de alocação indexada



Alocação indexada em dois níveis



Esquema combinado: UNIX (4K bytes por bloco)



Gerenciamento de espaço livre

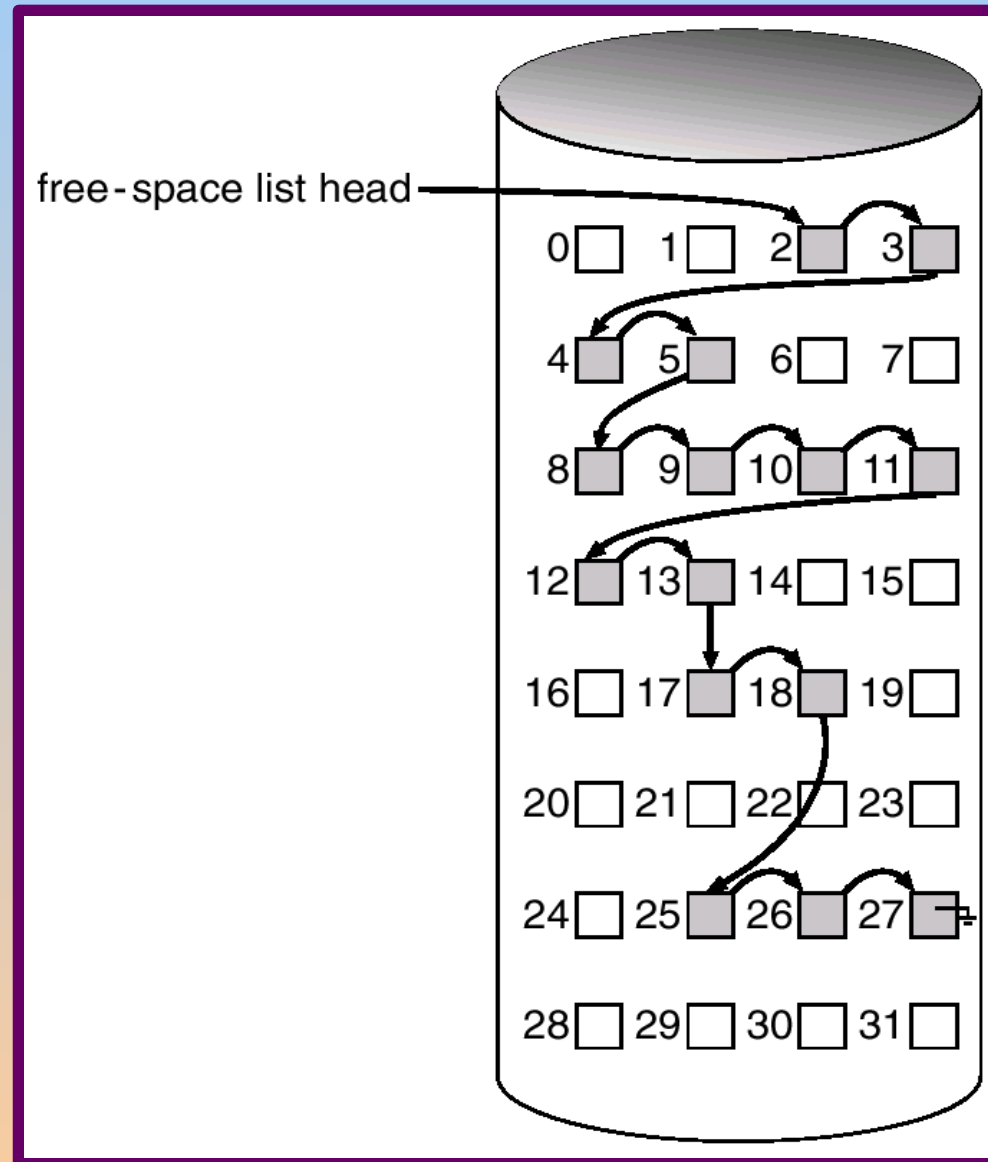
Solução I

■ Vetor de bits (n blocos)



$\text{bit}[i] = \begin{matrix} \square \\ \square \\ \square \end{matrix}$ $0 \Rightarrow \text{bloco}[i] \text{ livre}$
 $1 \Rightarrow \text{bloco}[i] \text{ ocupado}$

Solução II : Lista livre de blocos



Eficiência e performance

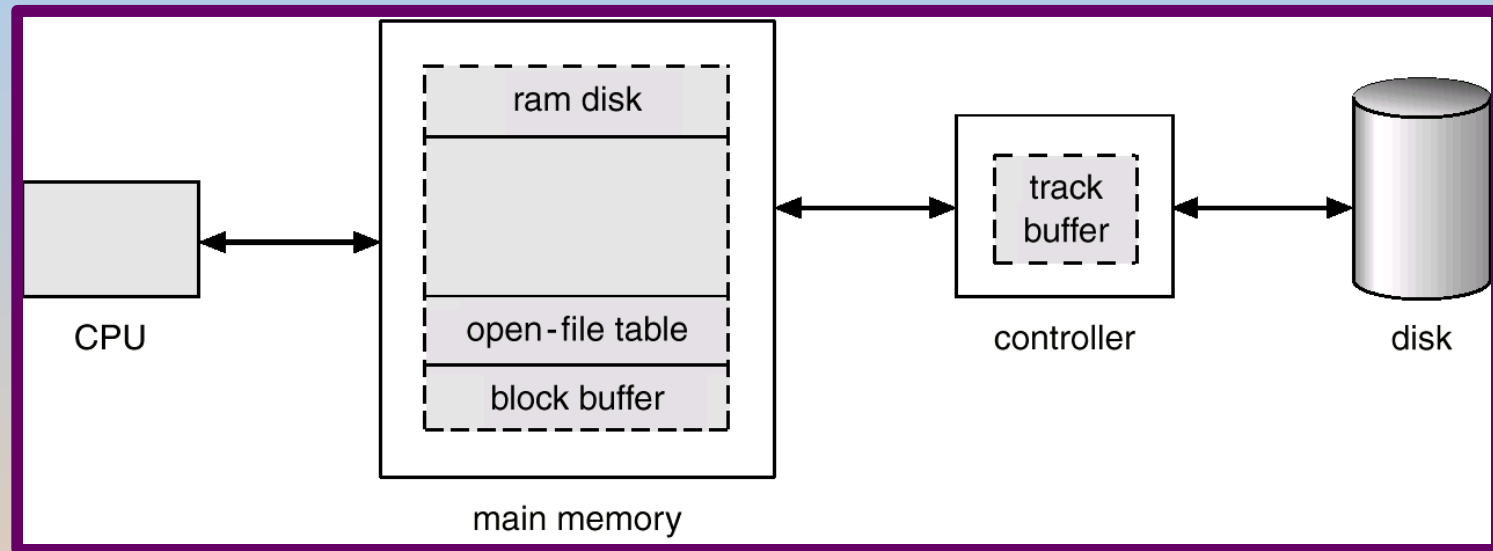
■ Eficiência e performance dependem de:

- ◆ Algoritmos de alocação de disco e diretórios
- ◆ Tipos de dados mantidos em cada entrada de diretório

■ Performance

- ◆ Cache de disco – separa uma área de memória para blocos usados frequentemente.
- ◆ Aumentar performance por seções dedicadas de memória, como disco virtual ou RAM disk.

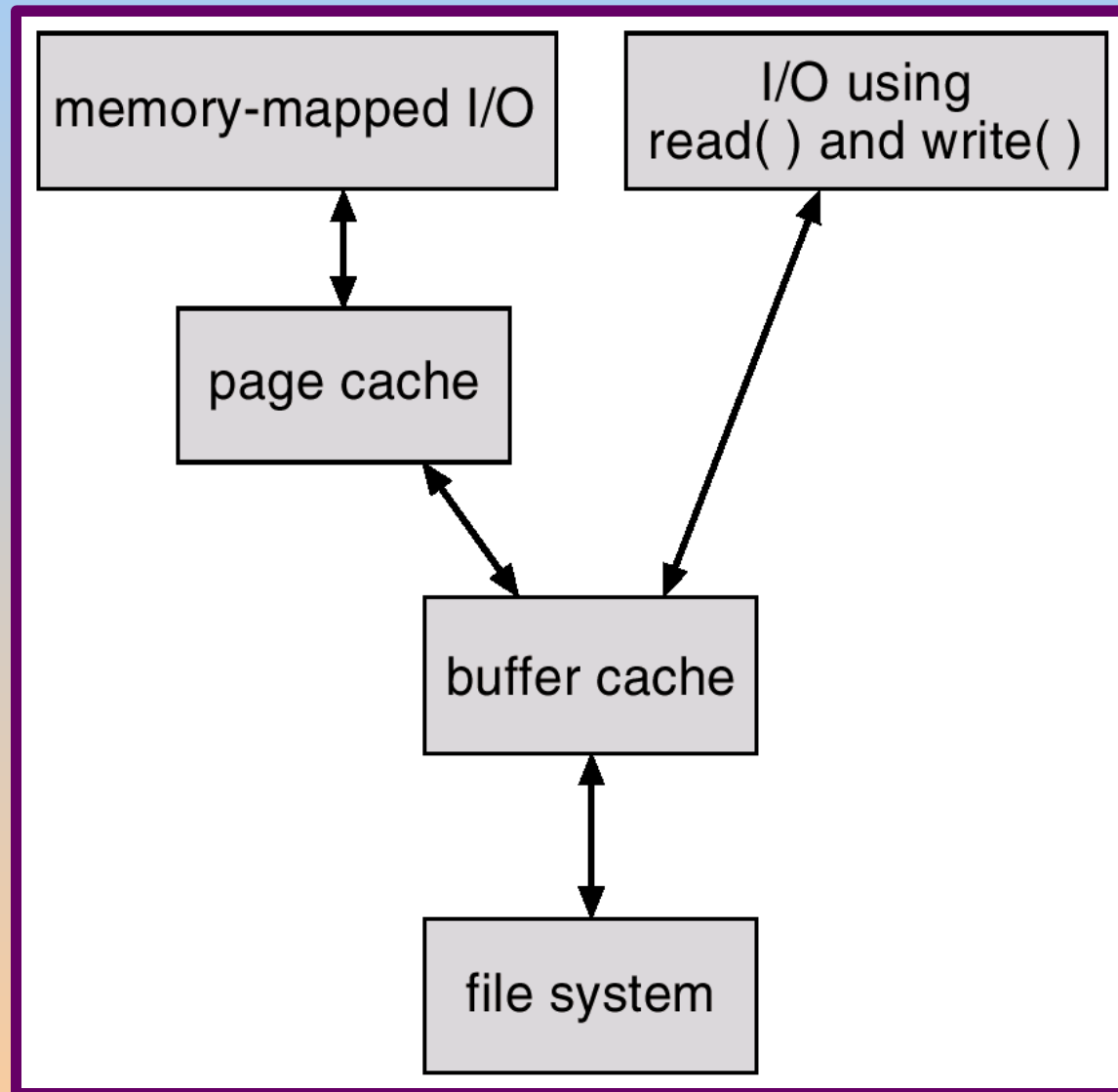
Várias localizações de cache de disco



Cache de páginas

- Um cache de páginas armazena páginas ao invés de blocos de disco, usando técnicas de memória virtual.
- E/S mapeada em memória usa cache de página.
- Rotina de E/S através do sistema de arquivos usa cache de buffer de disco (buffer cache).

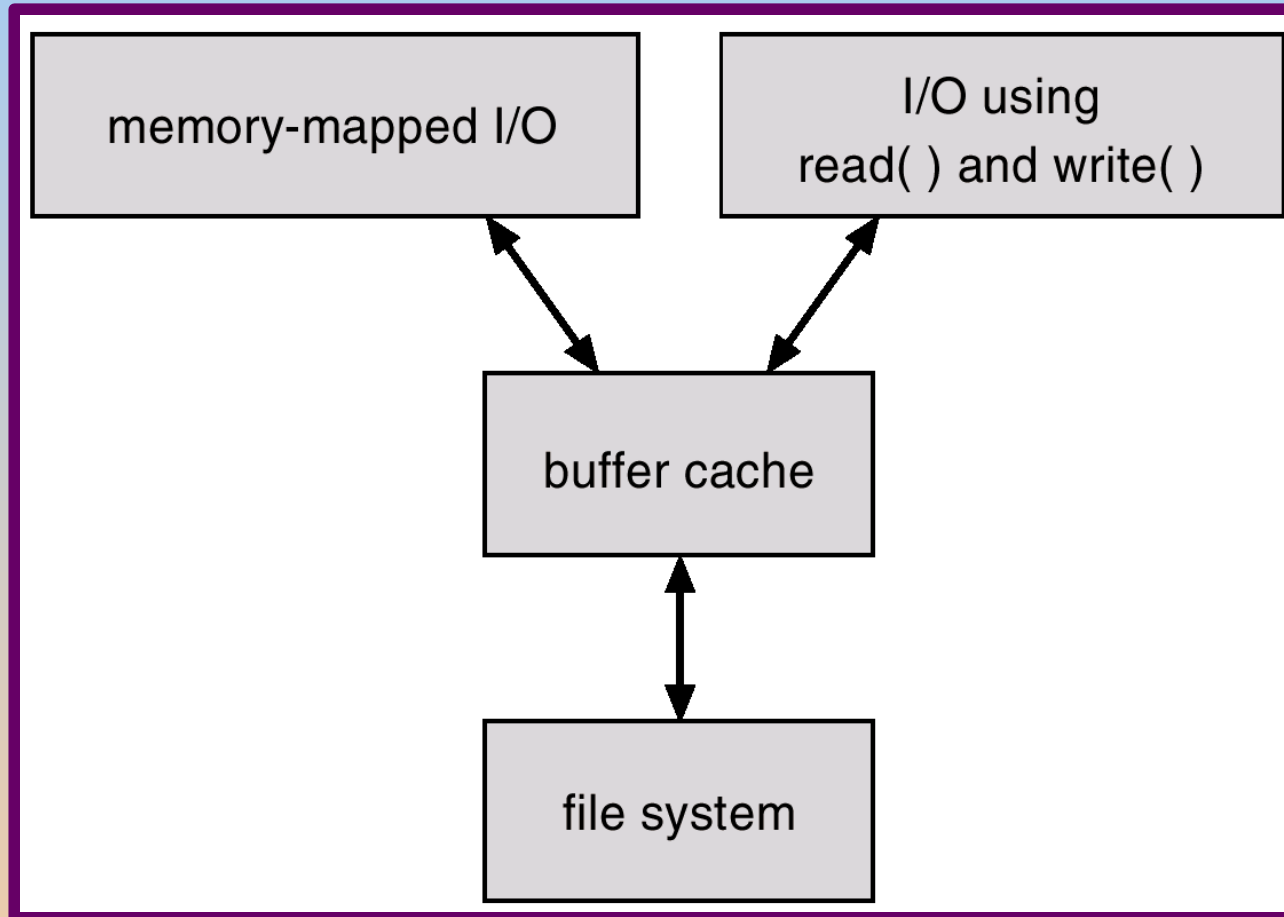
E/S sem cache unificado



Cache de buffer unificado

- Um sistema de cache de buffer unificado usa a mesma área de cache de página para páginas mapeadas em memória e E/S de sistema de arquivos ordinário.

E/S usando um cache de buffer unificado



Recuperação

- Verificação de consistência – compara dados na estrutura de diretório com os blocos de dados no disco e tenta tratar inconsistências.
- Usa programas de sistema para realizar backup de disco para outro dispositivo de armazenamento (disquete, fita magnética, etc).
- Restaura arquivos perdidos ou disco restaurando dados de backup.

Sistemas de arquivo baseados em Log

- Sistemas de arquivos baseados em log registram cada atualização do sistema de arquivos como uma transação.
- Todas as transações são escritas para um arquivo log. Uma transação é considerada confirmada (committed) uma vez que é escrita no log. Todavia, o sistema de arquivo pode não estar atualizado.
- As transações escritas no log são escritas assincronamente no sistema de arquivos. Quando o sistema de arquivos foi modificado, a transação é removida do log.
- Se o sistema de arquivos sofre avaria, todas as transações no log podem ainda ser realizadas.

Sun Network File System (NFS)

- Uma implementação e especificação de sistema de software para acessar arquivos remotos através de LANs (ou WANs).
- A implementação é parte dos sistemas Solaris e SunOS rodando em workstations SUN usando um protocolo de datagramas não-confiável(Protocolo UDP sobre IP e Ethernet).

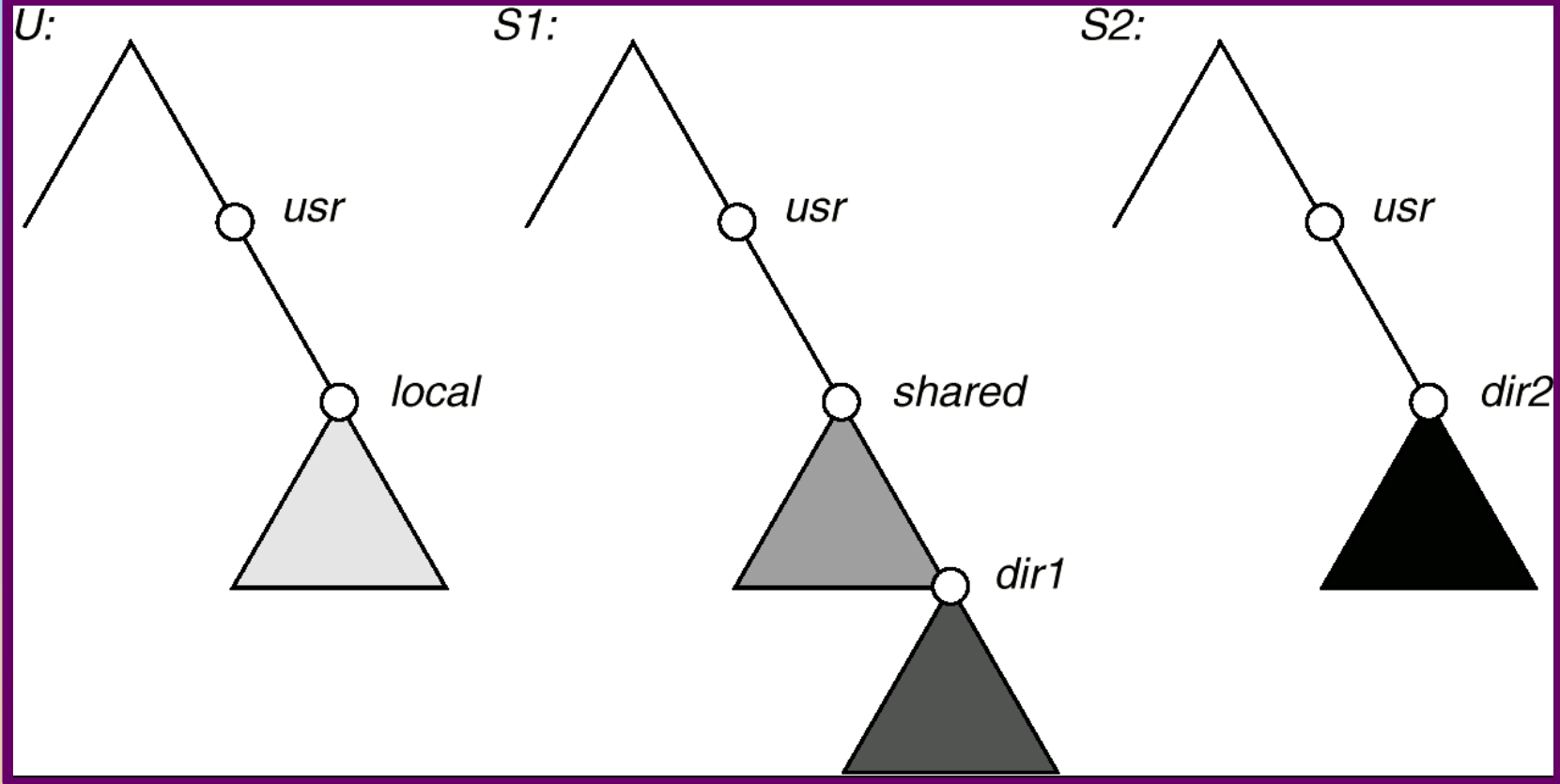
NFS (Cont.)

- Estações interconectadas são vistas como um conjunto de máquinas independentes com sistemas de arquivos independentes, que permite compartilhamento de entre arquivos de maneira transparente.
 - ◆ Um diretório remoto é montado sobre um diretório no sistema local de arquivos. O diretório remoto montado torna-se uma sub-árvore integral do sistema de arquivos locais, substituindo todas as sub-árvores descendentes do diretório local.
 - ◆ Especificação do diretório remoto para a operação de montagem é não-transparente: o nome do hospedeiro(host) do diretório remoto precisa ser disponibilizado. Arquivos no diretório remoto podem ser acessados de maneira transparente.
 - ◆ Potencialmente, qualquer sistema de arquivos (ou diretório dentro de um sistema de arquivos) pode ser montado remotamente no topo de qualquer diretório local.

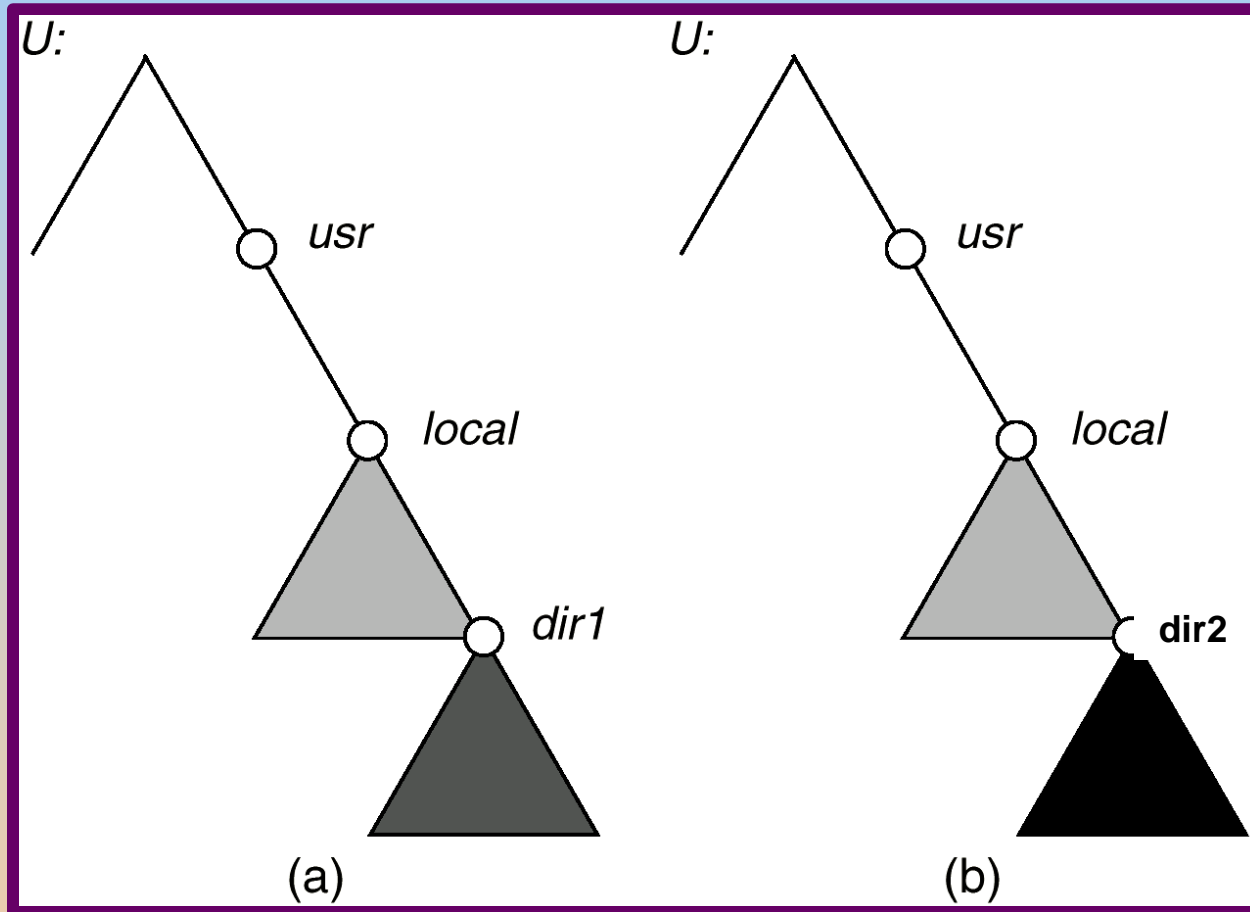
NFS (Cont.)

- NFS é projetado para operar em ambientes heterogêneos de diferentes máquinas, sistemas operacionais e arquiteturas de rede. Especificações NFS independem destes meios.
- Esta independência é alcançada através do uso de primitivas RPC construídas no topo de um protocolo de representação de dados externos (External Data Representation-XDR) usado entre duas interfaces independentes da implementação.
- A especificação NFS distingue entre os serviços disponibilizados por um mecanismo de montagem e os de acesso ao sistema remoto de arquivos.

Três implementações independentes de sistemas de arquivos



Montagem em NFS



Montagens

Montagens em cascata

Protocolo de montagem NFS

- Estabelece conexão lógica inicial entre servidor e cliente.
- Operação de montagem inclui nome do diretório remoto a ser montado e nome da máquina servidora que o contém:
 - ◆ Requisição de montagem é mapeada numa RPC e enviada para um servidor de montagem rodando na máquina servidora.
 - ◆ *Export list* – especifica os sistema de arquivo local exportáveis para montagem, juntamente com os nomes das máquinas que são permitidas montá-los.
- Se a requisição for autorizada em conformidade com a export list, o servidor retorna um manipulador de arquivo (file handle), que funciona como uma chave para futuros acessos.
- File handle – um identificador de sistema de arquivo e um número de inode para identificar o diretório montado dentro so sistema de arquivos exportado.
- A operação de montagem muda apenas a visão do usuário e não afeta o lado servidor.

Protocolo NFS

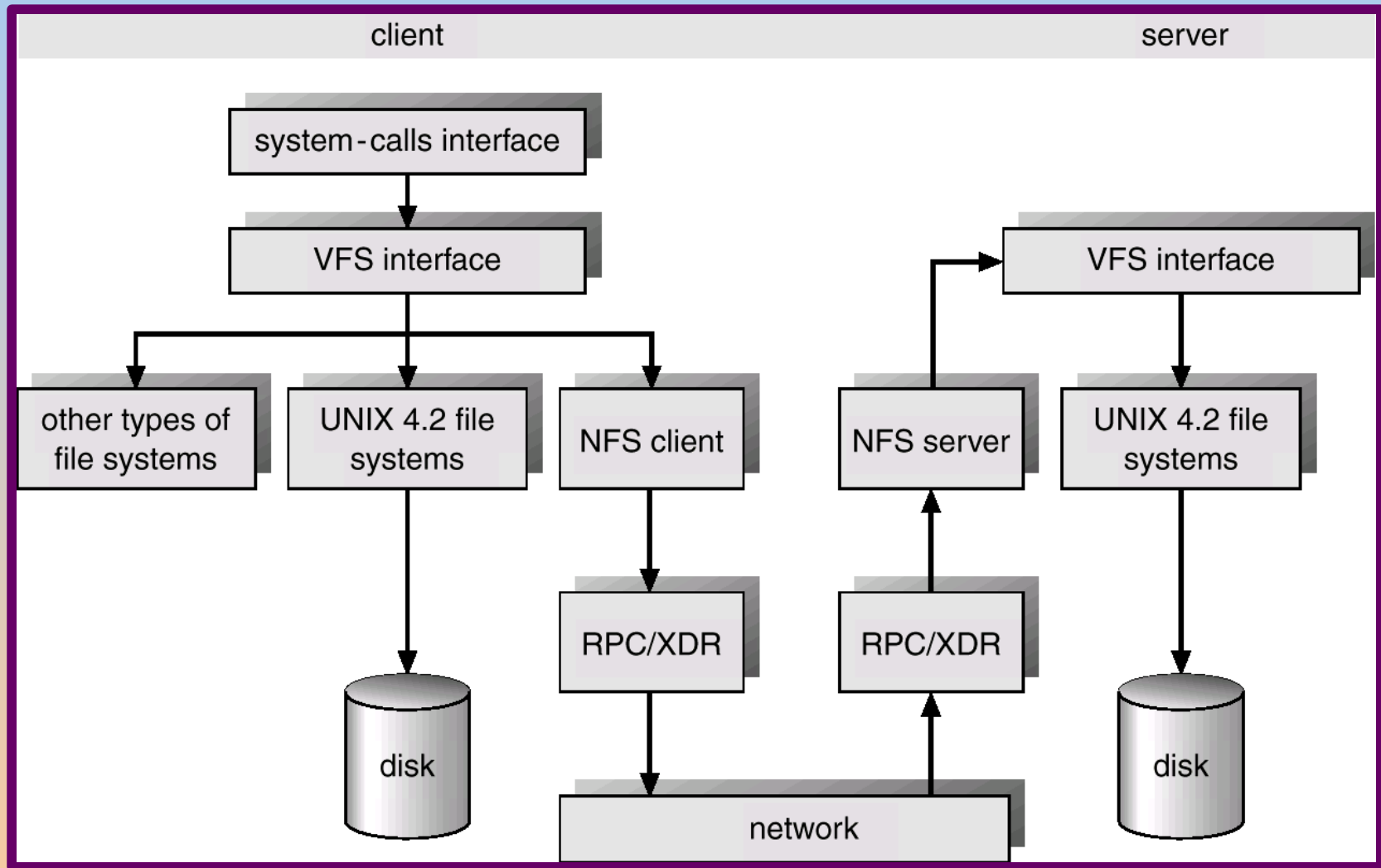
- Disponibiliza uma série de RPC's para operações para arquivos remotos:
 - ◆ Busca de arquivo dentro de um diretório
 - ◆ Leitura de um conjunto de entradas de diretório
 - ◆ Manipulação de links e diretórios
 - ◆ Acesso aos atributos de arquivos
 - ◆ Leitura e escrita de arquivos
- Servidores NFS não são baseados em estados: cada requisição precisa disponibilizar um conjunto completo de argumentos.
- Dados modificados precisam ser “comitados” no disco servidor antes dos resultados serem retornados ao cliente (perda das vantagens de cache).
- O protocolo NFS não disponibiliza mecanismos de controle de concorrência.

Três camadas mais importantes do UNIX

NFS

- Interface de sistema de arquivos UNIX (baseada em open, write, close e descritores de arquivo).
- *Camada Virtual File System (VFS)*– distingue entre arquivos locais e remotos. Arquivos locais são diferenciados pelos seus respectivos sistemas de arquivos.
 - ◆ O VFS ativa operações específicas de sistema de arquivos para manipular requisições locais de acordo com seus sistemas de arquivos.
 - ◆ Utiliza procedimentos do protocolo NFS para requisições remotas.
- Camada de serviço NFS – implementa o protocolo NFS.

Visão esquemática da arquitetura NFS



Tradução de nomes de caminho NFS

- Realizada através de quebra do caminho em nomes de componentes e realizando uma busca NFS separada para cada par de nome de componente e nó virtual de diretório (directory vnode).
- Para tornar a busca mais rápida, um esquema de cache de nome de diretório no lado cliente armazena vnodes para nomes de diretórios remotos.

Operações remotas NFS

- Quase existe uma correspondência um-para-um entre systems calls regulares do UNIX e RPCs de protocolo NFS, exceto na abertura e fechamento de arquivos.
- NFS adere ao paradigma de serviços remotos, mas emprega técnicas de buffering e cache para conseguir performance.
- Cache de blocos de arquivo – quando um arquivo é aberto, kernel verifica com o servidor remoto a necessidade de atualização de atributos já em cache. Blocos de arquivos em cache somente são usados se os correspondentes atributos estão atualizados.
- Cache de atributos de arquivo – o cache de atributos é atualizado toda vez que novos atributos chegam do servidor.
- Clientes não liberam blocos de escritas até o servidor confirmar que os dados foram escritos no disco.