

## Sistemas de Ficheiros

**1. Qual é o papel desempenhado pela memória de massa num sistema computacional? Indique que tipos de memória de massa estão normalmente presentes, ou podem ser ligados, a um computador portátil.**

A memória de massa tem função de armazenar aplicações e dados do sistema computacional, quando um sistema computacional é ligado só existe em memória principal, um programa, o bootloader que tem função de ler uma região específica da memória de massa que carrega o programa que implementa o ambiente de interação com o utilizador. Além disso, quase todos os programas durante a sua execução leem e alteram dados que estão armazenados na memória de massa. A um portátil podem ser ligada uma grande variedade de dispositivos memória de massa, alguns por exemplo HDD, leitores ópticos, pendrives USB, etc.

**2. Que requisitos são habitualmente impostos à memória de massa?**

A memória de massa não deve ser volátil (não perder informação), ter grande capacidade de armazenamento, providenciar acesso eficiente, manter a integridade da informação, permitir a partilha de informação (poder ser acedida de forma concorrente entre os múltiplos processos em execução).

**3. O que é um ficheiro? Explique a relevância deste conceito no âmbito da transferência de informação de e para a memória de massa.**

Ficheiro é um tipo abstracto de dados caracterizado por um conjunto de atributos, na qual podemos efectuar um conjunto de operações (leitura, escrita, etc), os ficheiros são importantes pois permitem uma interface simplificada para escrita e leitura de dados para a memória de massa.

**4. O que é um sistema de ficheiros? Porque é que é determinante que a parte do sistema de operação relacionada com a gestão da memória de massa apresente ao programador de aplicações uma interface uniforme de operações. Descreva as principais operações sobre ficheiros em Unix (consulte para isso o manual on-line)?**

Um sistema de ficheiros é a parte do sistema de operação responsável por implementar suporte a ficheiros e fornecendo um leque alargado de operações, ditas chamadas ao sistema, que estabelecem uma interface simples e segura de comunicação com a memória de massa para a sua manipulação. É importante que apresente ao programador uma interface uniforme de operações de forma a permitir interoperacionalidade, estabelecendo um meio comum de partilha de informação entre sistemas computacionais heterogéneos.

Em Unix as principais operações sobre ficheiros são criação, fecho, abertura, posicionamento, escrita e leitura.

**5. O que são os atributos de um ficheiro? Liste os mais comuns e justifique a sua razão de ser.**

Os atributos de um ficheiro, são um conjunto de atributos responsáveis por armazenar alguma informação de identificação e localização do ficheiro dentro do sistema bem como informação adicional útil ao utilizador. Os mais comuns são nome, localização, tipo, tamanho, pertença, protecção, etc que constitui o conjunto básico essencial à manipulação dos mesmos pelo utilizador.

**6. Um sistema de ficheiros contempla habitualmente diferentes tipos de ficheiros. Descreva os mais importantes e justifique a sua razão de ser.**

Os tipos de ficheiros mais importantes no sistemas de ficheiros tipicamente são: directorios, atalhos, ficheiros especiais e ficheiros regulares.

**7. Qual é o impacto que a existência de ficheiros de tipo directório introduz no acesso a um ficheiro concreto residente no sistema de ficheiros?. O que ocorreria se se implementasse num dispositivo de memória de massa com grande capacidade de armazenamento um sistema de ficheiros que não contemplasse a existência de directórios?**

Os ficheiros de tipo directorio permitem uma organização em arvore de todos os ficheiros do sistema, facilitando assim o acesso a ficheiros pelo utilizador. Caso se implementasse numa memoria de massa de grande dimensao um sistema de ficheiros sem suporte a ficheiros do tipo directorio o utilizador demoraria bastante tempo a procurar os ficheiros no dispositivo de memoria pois os mesmo encontrariam-se todos juntos sem qualquer tipo de organização.

**8. O que são atalhos? Apresente diferentes motivos para a inclusão deste tipo de ficheiros num sistema de ficheiros. Descreva de uma maneira clara e suficientemente detalhada uma situação concreta onde a existência de atalhos permite melhorar a qualidade de acesso à informação.**

Os atalhos tem a função de localizar outros ficheiros armazenndo o seu caminho, tornam-se uteis em casos como gerir versoes de aplicações, sendo possivel actualizar aplicacoes para versoes diferentes e manter o memso ponto de acesso para o utilizador (atalho), possibilitam acesso a ficheiros em sistemas de ficheiros distintos e sistemas de ficheiros remotos, e permitem acesso a zonas de dados protegida (pastas pessoais a qual so alguns users podem ter acesso por exemplo).

**9. Explique detalhadamente, socorrendo-se de um exemplo, como pode ser formada uma cadeia de atalhos. Porque é que tem que ser exigido que, ocorrendo uma cadeia de atalhos, ela tem que ser necessariamente linear? Como se pode impedir a ocorrência de cadeias circulares?**

Uma cadeia de atalhos pode ser formada indicando para cada atalho um novo atalho, esta cadeiras devem ser sempre lineares, pois caso sejam circulares ao tentar aceder ao atalho ele sera redirecionado num ciclo infinito para ele propio, bloqueando o sistema de acesso a ficheiros, uma forma de impedir isto sera garantir ao criar um novo atalho que o mesmo não apotente para ele propio ou em alternativa em runtime caso um atalho aponte para ele propio o ciclo é terminado mal seja detectada esta condição.

**10. Hoje em dia, é muito frequente os directórios dos sistemas de ficheiros incluírem duas entradas de directório com um significado bem definido: auto referência e referência ao directório imediatamente acima da hierarquia subjacente. Porque é que tal acontece? Justifique claramente a sua resposta.**

Um dos motivos para estas duas referencias e permitir facil referenciação do directorio actual e do directorio anterior permitindo uma interface facil tanto a nivel de navegação tanto como a referenciação relativa de ficheiros em programas por exemplo.

**11. O que são ficheiros regulares? É comum diferenciar os ficheiros regulares em ficheiros de texto e ficheiros binários, mas será que esta distinção é relevante para o sistema de operação? A quem é que de facto ela interessa? Porquê?**

Ficheiros regulares são os que contêm informação útil propriamente dita, o facto de serem ficheiros de texto ou binários não é revelante para o sistema operativo, sendo no entanto revelantes para o utilizador, isto devido ao facto de na realidade todos os ficheiros serem binários, no entanto os de texto têm uma codificação bem definida passível de serem lidos por aplicações desenhadas para esse efeito.

**12. O acesso ao conteúdo informativo dos ficheiros regulares é efectuado de dois modos fundamentalmente distintos: acesso sequencial e acesso aleatório. O que os distingue e a que tipos de ficheiros está cada um especialmente dirigido? Descreva uma situação concreta onde se mostra adequado aceder aleatoriamente a um ficheiro de texto. O que é que isso nos diz sobre a organização interna do conteúdo informativo?**

A diferença entre o acesso sequencial e o aleatório em ficheiros, é a forma como os dados são escritos no ficheiro, em escrita sequencial o ficheiro é escrito de forma sequencial este tipo é mais adequado a ficheiros de texto por exemplo, enquanto que na escrita aleatória o ficheiro é reescrito em posições aleatórias sendo necessário recorrer de forma intensiva à operação de posicionamento para definir posição onde a informação vai ser lida ou escrita, este tipo de acesso é mais indicado para ficheiros binários de aplicações por exemplo.

**13. O que é o directório de trabalho? Trata-se de algo endógeno ou exógeno ao sistema de ficheiros? Justifique claramente a sua resposta.**

O directório de trabalho é o directório em que nos situamos numa dada altura e que nesse momento serve como directório referência.

**14. A nomeação de um ficheiro particular num sistema de ficheiros organizado hierarquicamente pode ser realizada de duas maneiras distintas. Quais são elas? Ilustre a sua resposta apresentando exemplos válidos em Unix e Windows.**

Podemos referenciar um ficheiro a partir do seu caminho absoluto ou relativo, a partir do seu caminho relativo por exemplo em linux `./pasta/fic`, ou o seu caminho absoluto em windows `c:/pasta/pasta/fic`

**16. Em que consiste a arquitectura interna de um sistema de ficheiros? Que tipo de regiões, e qual a sua função, são normalmente consideradas?**

Designa a arquitectura interna de um sistema de ficheiros a organização do espaço de armazenamento interno do dispositivo, entendido como um array de blocos, de modo a fornecer a visão abstracta esperada pelo programador de aplicações. Normalmente são consideradas regiões de: cabeçalho, gestão do espaço, catalogação dos objectos e zona de dados.

**17. Ao especificar uma arquitectura interna para o sistema de ficheiros, o programador de sistemas tem que resolver dois problemas principais ao procurar organizar o espaço de armazenamento fornecido pelo dispositivo, entendido com um array de blocos, de modo a transmitir ao programador de aplicações a visão abstracta que ele espera. Quais são eles? Procure elaborar sobre as decisões que têm que ser tomadas.**

Especificar uma arquitectura interna significa, em primeiro lugar, isolar no array de blocos diferentes regiões a que se atribui um significado bem definido e, em segundo lugar, conceber estruturas de dados que descrevam de uma forma precisa e adequada a funcionalidade prescrita (a adequação procura enfatizar a existência de algoritmos para a sua manipulação eficiente).

**18. Em que consiste a gestão de espaço? Quais são as principais alternativas que são contempladas? Apresente vantagens e inconvenientes de cada uma delas.**

A gestão do espaço define o método usado para gerir dados na zona de dados. As alternativas consideradas são armazenamento contínuo, neste caso há aproveitamento quase completo do disco no entanto existem graves problemas de fragmentação após executar operações de apagamento, ou problemas relacionados com aumentar ficheiros. Segunda alternativa é armazenamento disperso neste caso quer o conteúdo informativo de cada ficheiro, quer o espaço livre são descritos pelo agregado de blocos que lhes estão associados, existem 3 alternativas de implementar este tipo de gestão: através de listas ligadas em que os nós são os próprios blocos, recorrendo a tabela de referência única, tabela de referência bit-map (apenas gere espaço livre), ou recorrendo a múltiplas tabelas de referência, esta implementação apesar de mais complexa elimina problemas relacionados com fragmentação da memória.

**19. Distinga fragmentação interna de fragmentação externa num sistema de ficheiros. Qual delas está sempre presente? Porquê? Como é que a outra é tida em conta em diferentes alternativas de gestão do espaço?**

Num caso de fragmentação interna apenas a última parte do bloco de conteúdo informativo de cada ficheiro é eventualmente desperdiçado, por outro lado em fragmentação externa o espaço livre deixa de estar concentrado numa só zona e aparece disperso por múltiplas regiões. Esta sempre presente a fragmentação interna, porque os ficheiros são escritos em bloco de tamanho bem definido e nem sempre ocupam a totalidade deste bloco.

**20. Duas soluções possíveis, para o problema da localização do conteúdo informativo dos ficheiros, são a existência de uma única tabela de referência ou de múltiplas tabelas de referência. Descreva estas duas abordagens e procure elucidar em que circunstâncias cada uma delas é a mais adequada.**

Numa tabela de referência única, esta está localizada num local pre-definido da zona de dados, e esta armazena ponteiro e descrição dos vários ficheiros no disco, no caso de múltiplas tabelas o que acontece que as tabelas podem referenciar outras de forma hierárquica, a primeira solução é melhor para casos de ficheiros de grande dimensão em que existem menos ficheiros e a lista é facilmente percorrida no entanto em caso de um grande número de ficheiros mais pequenos múltiplas tabelas terão melhor resultado.

**22. Uma forma muito eficiente para gestão do espaço livre é a existência de uma tabela de referência bitmap. Em que consiste? Descreva um inconveniente que lhe está associado.**

Uma tabela de bitmap tem tantas entradas quanto os blocos da zona de dados, as entradas desta tabela apenas tomam valor 1 se o bloco tiver livre e 0 em caso contrário, um inconveniente deste método é o facto de não contemplar qualquer ordenação de blocos, não podendo por isso ser aplicado à descrição do conteúdo de um ficheiro.

**24. Que vantagens advêm do armazenamento separado dos atributos dos ficheiros em estruturas de dados autónomas, os chamados nós-i?**

Separar os atributos dos ficheiros do seu conteúdo permite aumentar a eficiência de acesso aos dados, bem como simplificar o acesso à informação acerca dos ficheiros, com especial vantagem de que dentro do mesmo sistema de ficheiros a partilha de ficheiros é muito mais eficiente.

**25. Distinga ligações profundas (hard links) de ligações simbólicas? Que vantagens as primeiras apresentam sobre as segundas e quais são as suas limitações? Qual é o papel que o contador de referências desempenha neste contexto?**

Os hardlinks são ponteiros para os ficheiros que fazem parte do sistema de ficheiros enquanto que as ligações simbólicas, são elas próprias ficheiros, uma das vantagens relacionadas com a primeira é o facto de ser possível recuperar os ficheiros apagados já que o apagamento de um directorio passar a ser na verdade uma apagação de entrada de directorio, neste caso o contador de referências tem grande importância já que isto apenas será possível quando este contador tomar valor 0.

**26. Explique detalhadamente porque é que o acesso, para leitura ou escrita, a um bloco de um disco magnético deve ser evitado? Como é que o sistema de operação lida com este problema?**

Porque devido ao tempo de movimento de posicionamento da cabeça para leitura de um novo bloco tornam-se bastante lentas as leituras, por outro lado leituras de sectores sucessivos na pista são bastante rápidas. Para minimizar o overhead de acesso o sistema de operação actua de duas maneiras diferentes: scheduling do dispositivo (ordenação dos pedidos de acesso de modo a minimizar o tempo de pesquisa) e uso de cache (criar em memória principal uma região para armazenamento temporário de cópias dos frequentemente acedidos).

**27. O que é a cache de disco? Descreva a sua organização interna e o modo como opera.**

A cache de disco tem objectivo de criar em memória principal uma região para armazenamento temporário de cópias de sectores do dispositivo mais frequentemente acedidos para garantir um tempo de acesso da mesma ordem de grandeza da transferência de informação entre regiões de memória distintas. Sempre que seja necessário ler um bloco de dados do dispositivo, a cache é pesquisada a ver se o mesmo já está presente na cache, caso esteja presente o valor é lido da mesma caso contrário é transferido do disco para a cache e posteriormente é lido, em caso de escrita a cache é da mesma forma verificada pela presença desses dados caso já exista o valor de status e alterador e os dados são escritos, o tamanho da cache é limitado, sempre que não haja nós livres na cache o nó que não é acedido a mais tempo é selecionado para substituição, se o status deste não tiver sido marcado como alterado, antes de ser substituído os dados são actualizados no dispositivo e só após essa substituição se substitui os dados.