

Organização de Computadores

Prof. Robson de Souza

Memória Secundária

Seja qual for o tamanho da memória principal, ela sempre será muito pequena. As pessoas sempre querem armazenar mais informações do que a memória pode conter, ainda mais porque, à medida que a tecnologia melhora, elas começam a pensar em armazenar coisas que antes não era viável. A memória secundária se faz necessária porque armazenar uma grande quantidade de informações em memória principal, cache e registradores tornaria esse processo extremamente caro.

Hierarquias de memória

A solução tradicional para armazenar grandes quantidades de dados é uma hierarquia de memória, como mostra a figura abaixo:



Fonte: (Tanenbaum e Austin, 2013)

No topo, estão os registradores da CPU, que podem ser acessados à velocidade total da CPU. Em seguida, vem a memória cache, que está na faixa de 32 KB a alguns megabytes. A memória vem logo após, hoje com tamanhos que vão de 1 GB para sistemas muito básicos até centenas de gigabytes na extremidade mais alta. Depois, vêm os discos magnéticos e SSDs. Por fim, temos fitas magnéticas e discos ópticos para armazenagem de arquivos.

À medida que descemos na hierarquia, três parâmetros aumentam. Primeiro, o tempo de acesso fica maior. Os registradores da CPU podem ser acessados em um nanossegundo ou menos. Memórias cache demoram um pequeno múltiplo dos registradores da CPU. Acessos à memória principal normalmente levam 10 nanossegundos. Agora, vem uma grande lacuna, porque tempos de acesso a discos são no mínimo 10 vezes mais lentos para discos em estado sólido e centenas de vezes mais lentos para discos magnéticos.

Segundo, a capacidade de armazenagem aumenta à medida que descemos na hierarquia. Registradores de CPU são bons para, talvez, 128 bytes, caches para algumas dezenas de megabytes, memórias principais para alguns gigabytes, discos em estado sólido para centenas de gigabytes e discos magnéticos para terabytes.

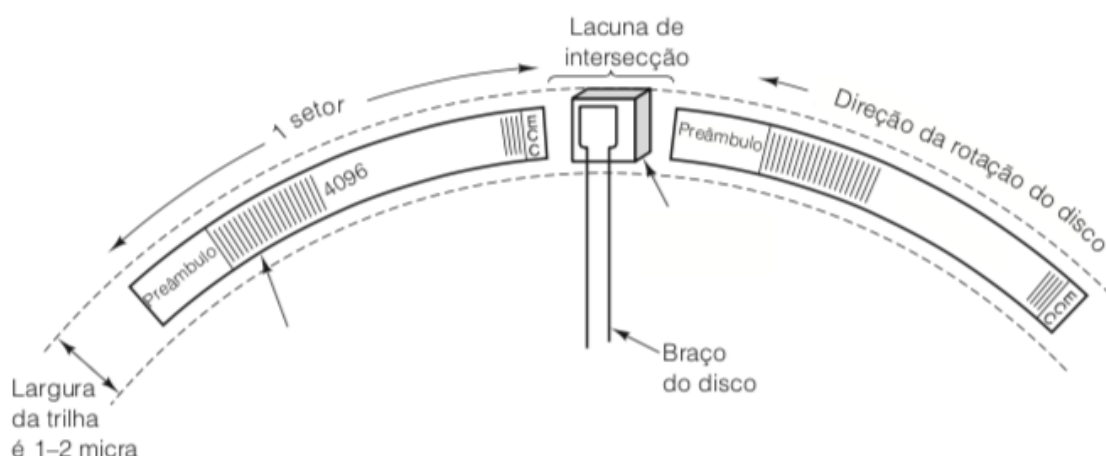
Terceiro, o número de bits por dólar gasto aumenta descendo a hierarquia. Embora os preços atuais mudem com rapidez, a memória principal é medida em dólares/megabyte, o disco em estado sólido em dólares/gigabyte e a armazenagem em disco magnético em centavos/gigabyte.

Discos magnéticos

Um disco magnético é composto de um ou mais pratos de alumínio com um revestimento magnetizável. No início, esses pratos tinham até 50 cm de diâmetro, mas agora têm normalmente de 3 a 9 cm, e discos para notebooks já estão com menos de 3 cm e continuam encolhendo.

Um cabeçote de disco que contém uma bobina de indução flutua logo acima da superfície, apoiado sobre um colchão de ar. Quando uma corrente positiva ou negativa passa pelo cabeçote, ele magnetiza a superfície logo abaixo dele, alinhando as partículas magnéticas para a esquerda ou para a direita, dependendo da polaridade da corrente.

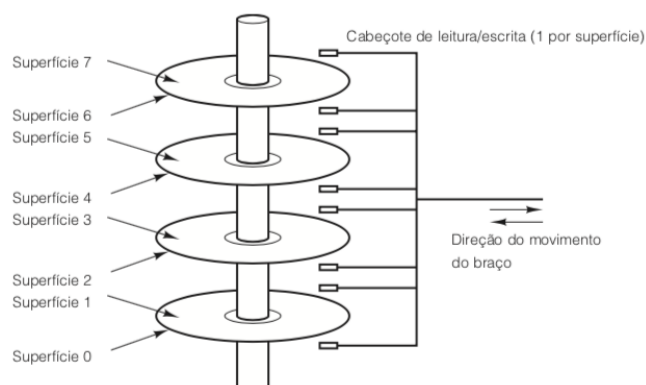
Quando o cabeçote passa sobre uma área magnetizada, uma corrente positiva ou negativa é induzida nele, o que possibilita a leitura dos bits armazenados antes. Assim, à medida que o prato gira sob o cabeçote, uma corrente de bits pode ser escrita e mais tarde lida. A geometria de uma trilha de disco é mostrada na Figura abaixo:



Fonte: (Tanenbaum e Austin, 2013)

Todos os discos têm braços móveis que podem se mover para dentro e para fora a diferentes distâncias radiais da haste ao redor da qual o prato gira. A cada distância radial pode ser escrita uma trilha diferente. A largura de uma trilha depende da largura do cabeçote e da precisão com que ele pode ser posicionado radialmente.

Existem discos que são compostos de vários pratos empilhados na vertical. Cada superfície tem seu próprio braço e cabeçote. Os braços são agrupados de modo que todos se movimentem para diferentes posições radiais ao mesmo tempo. O conjunto de trilhas em uma dada posição radial é denominado cilindro. A figura abaixo mostra como é esse tipo de disco.



Fonte: (Tanenbaum e Austin, 2013)

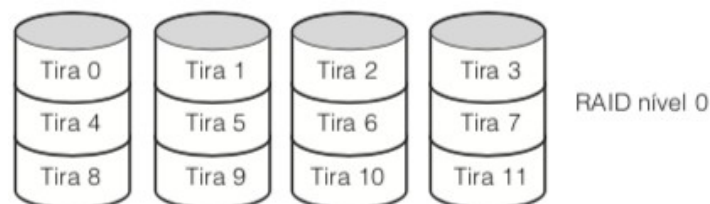
RAID

Como vimos, muitas vezes é usado processamento paralelo para acelerar o desempenho da CPU. Ao longo dos anos, ocorreu a várias pessoas que a E/S paralela também poderia ser uma boa ideia. Em seu artigo de 1988, Patterson et al. sugeriram seis organizações específicas de disco que poderiam ser usadas para melhorar o desempenho, a confiabilidade do disco, ou ambos (Patterson et al., 1988). Essas ideias logo foram adotadas pela indústria e deram origem a uma nova classe de dispositivos de E/S, denominados **RAID**. Patterson et al. definiram RAID como Redundant array of Inexpensive disks (Arranjo Redundante de Discos Baratos), mas a indústria redefiniu o I como “independente” em vez de barato (Inexpensive).

A ideia fundamental de um RAID é instalar uma caixa cheia de discos próxima ao computador, em geral um grande servidor, substituir a placa do controlador de disco por um controlador RAID, copiar os dados para o RAID e então continuar a execução normal. Em outras palavras, um RAID deveria parecer um HD comum para o sistema operacional, mas ter melhor desempenho e melhor confiabilidade.

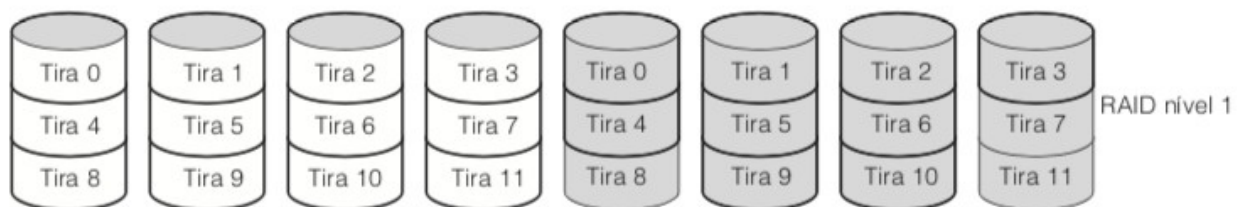
Além de parecerem um disco único para o software, há uma propriedade comum a todos os RAIDs, que é a distribuição dos dados pelos drives para permitir operação paralela. Patterson et al. definiram vários esquemas diferentes para fazer isso e, agora, eles são conhecidos como RAID nível 0 até RAID nível 5.

O **RAID nível 0** consiste em ver o disco virtual simulado pelo RAID como se fosse dividido em tiras de k setores cada. A organização RAID nível 0 escreve tiras consecutivas nos drives por alternância circular, como demonstrado na Figura abaixo:



Fonte: (Tanenbaum e Austin, 2013)

O **RAID nível 1** é um RAID verdadeiro. Ele duplica todos os discos, portanto, há quatro discos primários e quatro de backup. Para uma escrita, cada tira é escrita duas vezes. Para uma leitura, qualquer das duas cópias pode ser usada, distribuindo a carga por mais drives. Por conseguinte, o desempenho da escrita não é melhor do que o de um único drive, mas o de leitura pode ser duas vezes melhor. A tolerância a falhas é excelente: se um drive falhar, basta usar a outra cópia em seu lugar. A recuperação consiste na simples instalação de um novo drive e em copiar todo o drive de backup para ele. A figura abaixo mostra esse RAID:



Fonte: (Tanenbaum e Austin, 2013)

Ao contrário dos níveis 0 e 1, que trabalham com tiras de setores, o **RAID nível 2** trabalha por palavra, possivelmente até por byte. A ideia aqui é escrever bit a bit em discos diferentes.

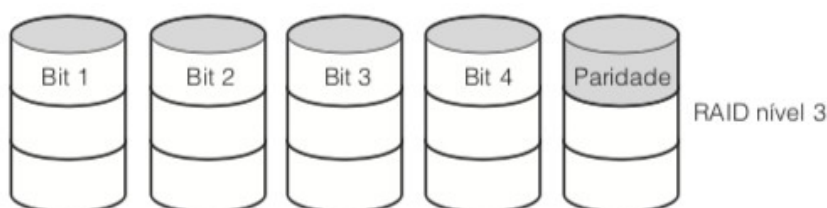
Como exemplo da eficiência disso, o computador Thinking Machine CM-2 usava esse esquema, pegando palavras de 32 bits de dados e adicionando 6 bits de paridade para formar uma palavra de Hamming de 38 bits, mais um bit extra para paridade de palavra, e distribuindo cada palavra em 39 drives de disco. O rendimento total era imenso porque em um tempo de setor ele podia escrever o equivalente a 32 setores de dados. Além disso, perder um drive não causava problemas, porque essa perda equivaleria a perder 1 bit em

cada palavra de 39 bits lida, algo que o código de Hamming poderia manipular facilmente. Existem duas desvantagens principais em utilizar RAID nível 2, primeiro que esse esquema requer que as rotações de todos os drives sejam sincronizadas, segundo, esse esquema exige muito do controlador. A figura abaixo mostra o RAID nível 2:



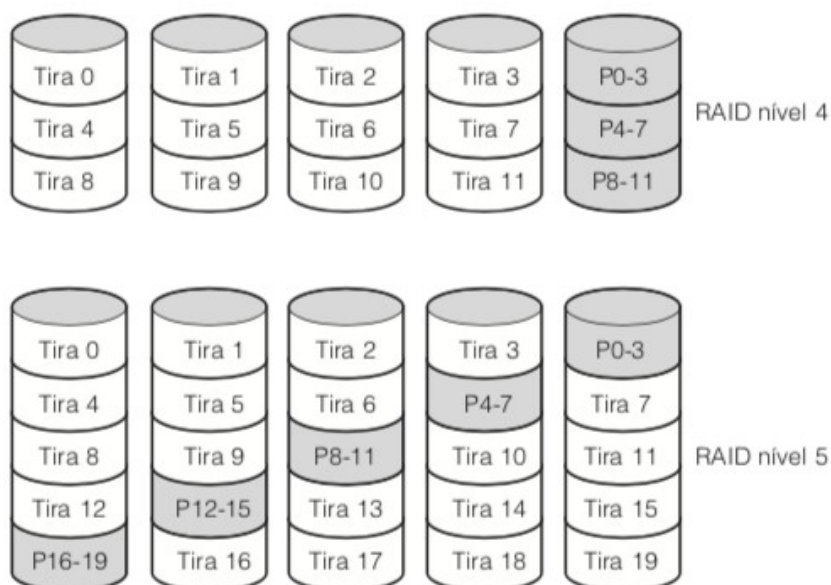
Fonte: (Tanenbaum e Austin, 2013)

O **RAID nível 3** é uma versão simplificada do RAID nível 2. Nesse arranjo, um único bit de paridade é computado para cada palavra de dados e escrito em um drive de paridade. A figura abaixo mostra esse esquema:



Fonte: (Tanenbaum e Austin, 2013)

RAIDs níveis 4 e 5 de novo trabalham com tiras, e não com palavras individuais com paridade, não requerem drives sincronizados. O RAID nível 4 é como o RAID nível 0, com paridade tira por tira escrita em um drive extra. O RAID nível 5 segue uma lógica semelhante, mas trabalha distribuindo os bits de paridade uniformemente por todos os drives, por alternância circular, conforme mostra a Figura abaixo:



Fonte: (Tanenbaum e Austin, 2013)

Discos em Estado Sólido (SSDs)

Os discos em estado sólido (SSDs – Solid State Disks) são discos feitos de memória flash não volátil. São uma alternativa de alta velocidade às tecnologias tradicionais em disco magnético. Os SSDs foram

inventados por causa de uma falha que ocorre em transistores chamada de “injeção de portadora quente”, isso significa que uma carga elétrica é embutida dentro de um transistor que funcionava, deixando-o em um estado onde fica permanentemente ligado ou desligado. Isso ocorre devido ao desgaste que os transistores sofrem com o uso permanente.

Embora em geral considerado sentença de morte para um transistor, Fujio Masuoka, enquanto trabalhava para a Toshiba, descobriu um modo de aproveitar esse mecanismo de falha para criar uma nova memória não volátil. No início da década de 1980, ele inventou a primeira memória flash.

Os discos flash são compostos de muitas células de memória flash em estado sólido. As células da memória flash são feitas de um único transistor flash especial. Embutido no transistor há uma porta flutuante que pode ser carregada e descarregada usando altas voltagens.

Para programar uma célula de bit flash, uma alta tensão (no mundo dos computadores, 12v é uma alta tensão) é aplicada à porta de controle, que acelera o processo de injeção de portadora quente na porta flutuante. Os elétrons são embutidos na porta flutuante, que coloca uma carga negativa interna no transistor flash. A carga negativa embutida aumenta a tensão necessária para ligar o transistor flash e, testando se o canal liga ou não com uma tensão alta ou baixa, é possível determinar se a porta flutuante está carregada ou não, resultando em um valor 0 ou 1 para a célula flash. A carga embutida permanece no transistor, mesmo que o sistema perca a alimentação, tornando a célula de memória flash não volátil.

Os SSDs são bem mais rápidos que os discos magnéticos. Enquanto um disco magnético típico pode acessar dados em até 100 MB/s, um SSD pode operar duas a três vezes mais rápido, dependendo do fabricante, até mais. E como o dispositivo não possui partes móveis, ele é muito adequado para uso em notebooks, onde trepidações e movimentos não afetarão sua capacidade de acessar dados.

A desvantagem dos SSDs, em comparação com discos magnéticos, é o seu custo. Além disso, SSDs possuem uma taxa de falha maior que discos magnéticos. Uma célula flash típica pode ser escrita somente por cerca de 100 mil vezes antes que não funcione mais. O processo de injetar elétrons na porta flutuante a danifica aos poucos, bem como seus isoladores ao redor, até que não funcione mais. Para aumentar o tempo de vida dos SSDs, é usada uma técnica denominada **nivelamento de desgaste**, para espalhar as escritas por todas as células flash no disco. Toda vez que um novo bloco de disco é escrito, o bloco de destino é reatribuído a um novo bloco do SSD, que não foi escrito recentemente.

Referências bibliográficas:

TANENBAUM, Andrew S. Organização Estruturada de Computadores, 2007, 5ª Edição.

TANENBAUM, Andrew S. AUSTIN, Todd; Organização Estruturada de Computadores, 2013, 6ª Edição.