

Aula 9

Organização de Computadores

Hierarquia de Memória
Memória Virtual

Profa. Débora Matos

Memória Virtual

A memória principal pode agir como uma “cache” para o armazenamento secundário.

Objetivo: permitir o compartilhamento seguro e eficiente da memória entre vários programas.

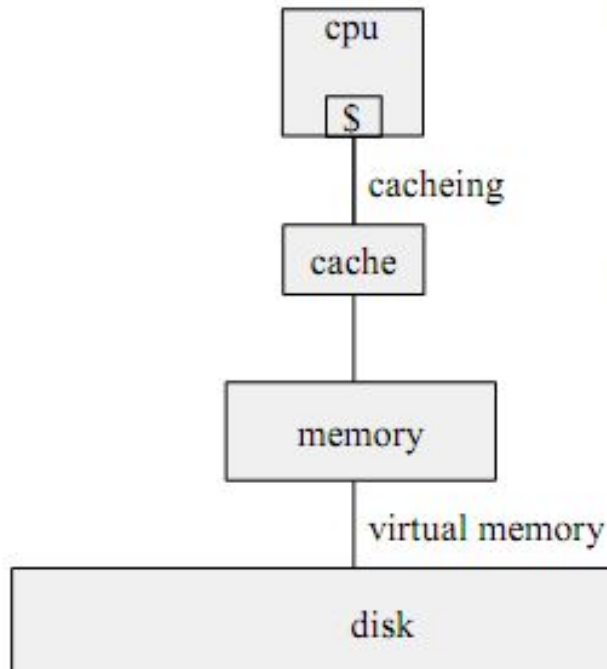
Quando um grupo de programas são executados ao mesmo tempo em um computador, a memória total exigida por todos os programas pode ser muito maior do que a quantidade de memória principal.

Memória Virtual

- Usando memória virtual, o computador endereça mais memória principal do que ele realmente possui e usa o disco rígido para conter o excedente;
- Esta área do disco é chamada de **arquivo de páginas**, pois contém porções da memória principal.

Memória Virtual

- De forma semelhante, a memória principal pode agir como uma “cache” para a memória secundária.
 - *Memória Virtual* : técnica que nos permite ver a memória principal como uma cache de grande capacidade de armazenamento
 - É apenas mais um nível na hierarquia de memórias



- mecanismo automático de gerência de memória, que traz automaticamente para a MP os blocos de informação (do disco) necessários
- usuário tem a impressão de trabalhar com uma memória única, do tamanho da memória secundária, mas com tempo de acesso próximo do tempo da MP

Memória Virtual

- mesma idéia da cache, mas com terminologia diferente

<u>cache</u>	<u>MV</u>
bloco	página (ou segmento)
cache miss	<i>page fault</i>
endereço	endereço virtual (ou lógico)
índice	endereço real (ou físico)

- endereço *virtual (lógico)*: gerado pelo programa
 - deve endereçar todo espaço em disco
 - maior número de bits
- endereço *real (físico)*: endereço na memória principal
 - menor número de bits

Memória Virtual (MV)

Por que MV é diferente das caches?

- Miss penalty é **MUITO** maior (milhões de ciclos)! Se informação não está na memória, está no disco!
- Logo:
 - *miss ratio* precisa ser bem menor do que em cache
 - alta penalidade do *miss* => necessário buscar blocos maiores em disco
 - princípio de localidade opera sobre blocos maiores de dados ou instruções e leva a *hit ratios* bem mais elevados
 - mapeamento associativo das páginas
 - *misses* são tratados por software (há tempo disponível)
 - técnica de escrita *write-through* não é uma opção. Usa-se *write-back*.

Memória Virtual

Cache -> bloco

Memória Virtual -> página

- Ao usar a memória virtual, o processador produz um endereço virtual, traduzido por uma combinação de HW e SW, para um endereço físico.
- O endereço é dividido em um número de página virtual e um offset de página.

Memória Virtual

- Memória virtual pode ser implementada por diferentes técnicas:
 - Paginação
 - Segmentação
 - Ambas

Paginação:

Processo de copiar uma página virtual do disco para um quadro de página na memória principal.

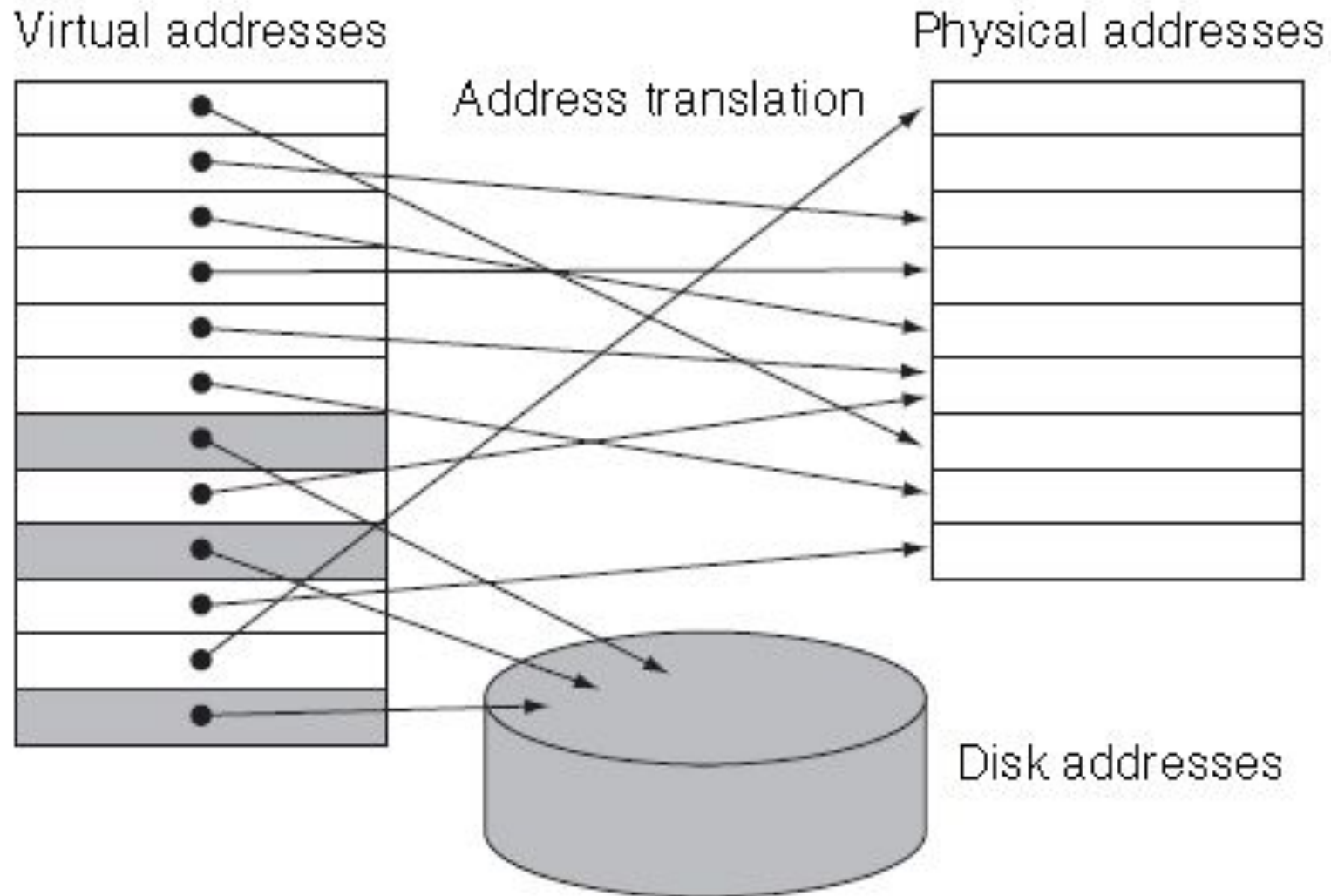
Memória Virtual

- A maneira mais comum de implementar memória virtual é usar **paginação**;
- **Paginação** é um método em que a memória principal é dividida em blocos de **tamanho fixo**;
- Os blocos de tamanho físico em que a memória principal é dividida são chamados de **quadros de páginas**.
- Sempre que a CPU gera um endereço este é em relação ao espaço de endereçamento virtual.
- **Fragmentação** é a área da memória que não consegue ser utilizada (ex. quando um processo pequeno aloca 1 quadro).

Memória Virtual

- Tanto a memória virtual quanto a memória física são desmembradas em páginas, cada página virtual é mapeada para uma página física;
- As páginas físicas podem ser compartilhadas, fazendo 2 endereços virtuais apontarem para o mesmo endereço físico.
- A memória virtual implementa a tradução do espaço de endereçamento em um programa para os endereços físicos;
- A tradução tem a função de proteger o espaço de endereçamento de um programa em relação aos outros.

Memória Virtual



Memória Virtual

- Relocação: os endereços virtuais são mapeados por um programa para diferentes endereços físicos antes de serem utilizados no acesso à memória.
- Dessa forma é possível carregar o programa em qualquer lugar na memória principal.
- Os programas são realocados como um conjunto de páginas de tamanho fixo;

Memória Virtual

Tempo de acesso

Tempo médio de acesso T_{ma} é dado por

$$T_{ma} = T_m + (1 - h) T_s$$

onde T_m = tempo de acesso à MP
 T_s = tempo de acesso ao disco
 h = hit ratio

p.ex. se $T_m = 20 \text{ ns}$, $T_s = 20 \text{ ms}$, $h = 0.9999$
então $T_{ma} = 2,02 \mu\text{s}$ (100 x maior do que T_m)

Paginação

- espaços de memória real e virtual divididos em blocos chamados de *páginas*
 - páginas tem tipicamente de 64 bytes a 4 Kbytes
- endereços virtuais e reais divididos em 2 campos
 - endereço da página
 - endereço da linha (ou palavra), dentro da página

Paginação

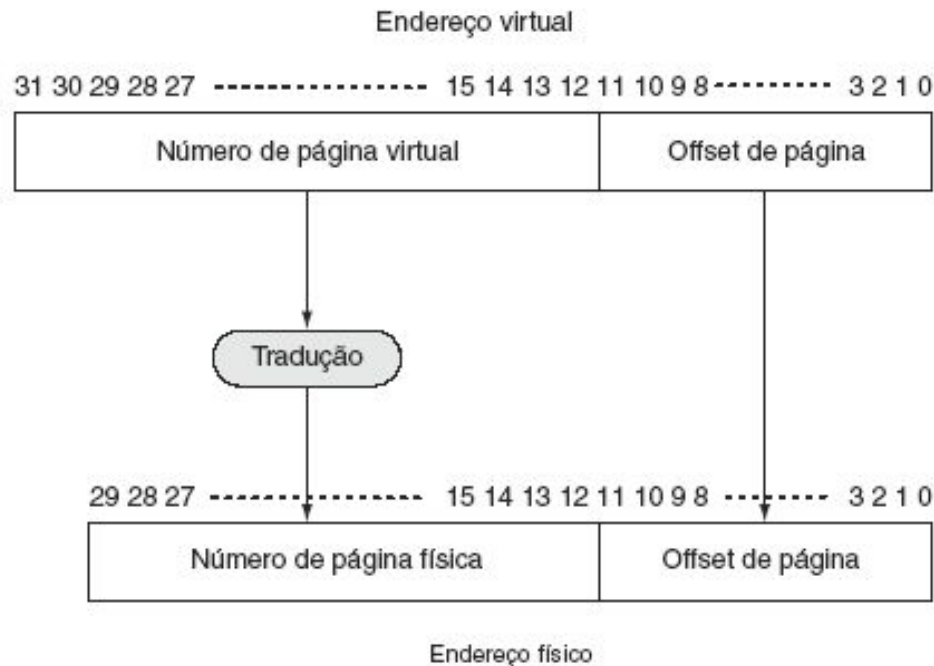
- O endereço onde as várias páginas residem são armazenadas em uma **tabela de páginas**.
- Cada processo possui a sua própria tabela de páginas;
- A tabela de páginas possui N linhas, onde N é o número de páginas virtuais do processo;
- Se existirem páginas do processo que não estão na memória principal, a tabela de páginas indica isto configurando o **bit válido** como 0.
- Cada entrada na tabela de páginas possui dois campos: **um bit válido e um número de quadro**.

Paginação

- **Bit de uso** da página – indica o uso da página. Se o bit permanece em zero (0), indica que a página não tem sido usada por um período de tempo e que ela pode ser enviada para o disco.
- Se o processo possui quadros livres na memória principal quando uma falha de páginas ocorre, a **nova página recuperada pode ser colocada em qualquer um dos quadros livres**;
- Se a memória alocada para o processo estiver cheia, uma **página vítima** deve ser selecionada.

Memória Virtual

- O endereço é dividido em número da página virtual e no deslocamento da página (offset)
 - O número de bits no campo de deslocamento (offset) da página determina o tamanho da página
 - O número de páginas endereçáveis via endereço virtual não precisa ser igual ao numero de páginas endereçáveis pelo endereço físico.



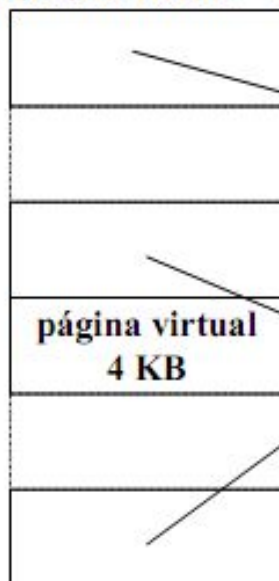
Memória Virtual

- A fim de indicar o local da tabela de páginas na memória, o HW possui um registrador que aponta para o início da tabela de páginas (registrador de tabela de páginas).

Exemplo

Paginação

memória virtual = 4 GB

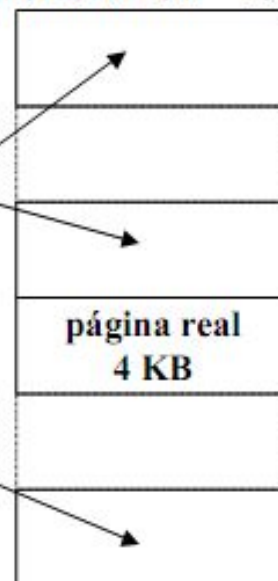


1 M páginas de 4 KB

nº página = 20 bits	
---------------------	--

endereço virtual = 32 bits

memória real = 256 MB

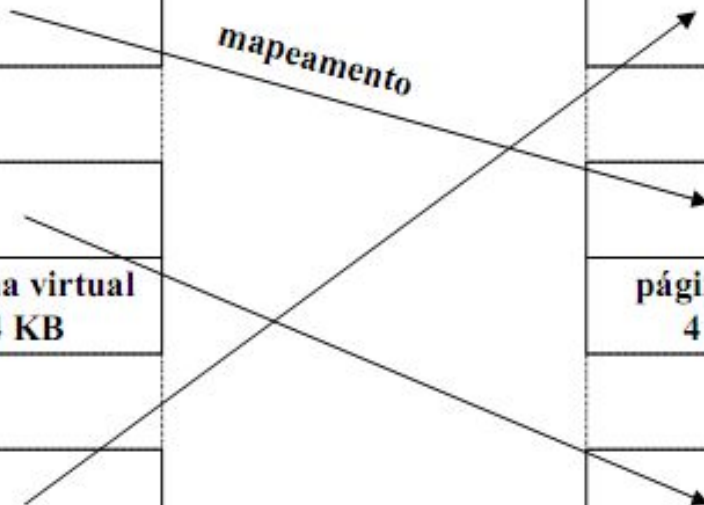


64 K páginas de 4 KB

nº página = 16 bits	
---------------------	--

endereço real = 28 bits

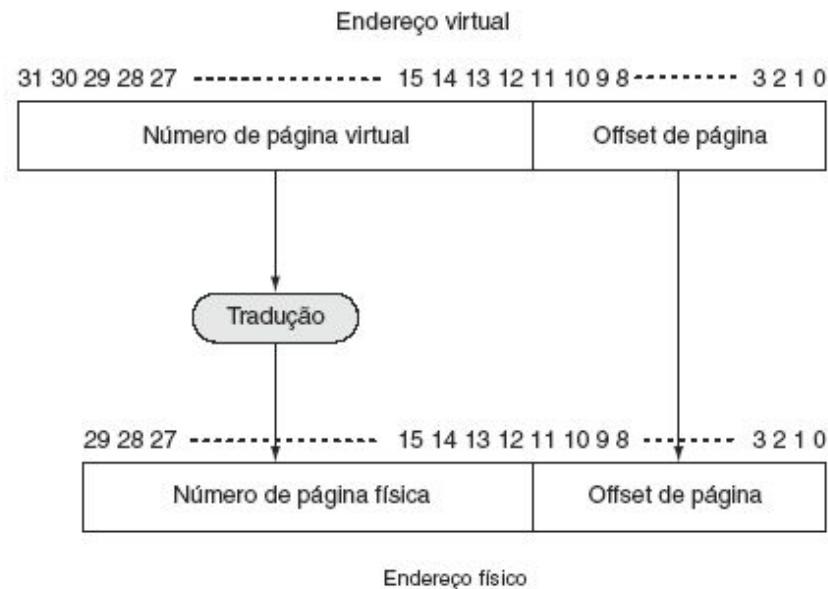
mapeamento



Páginas: blocos de memória virtual

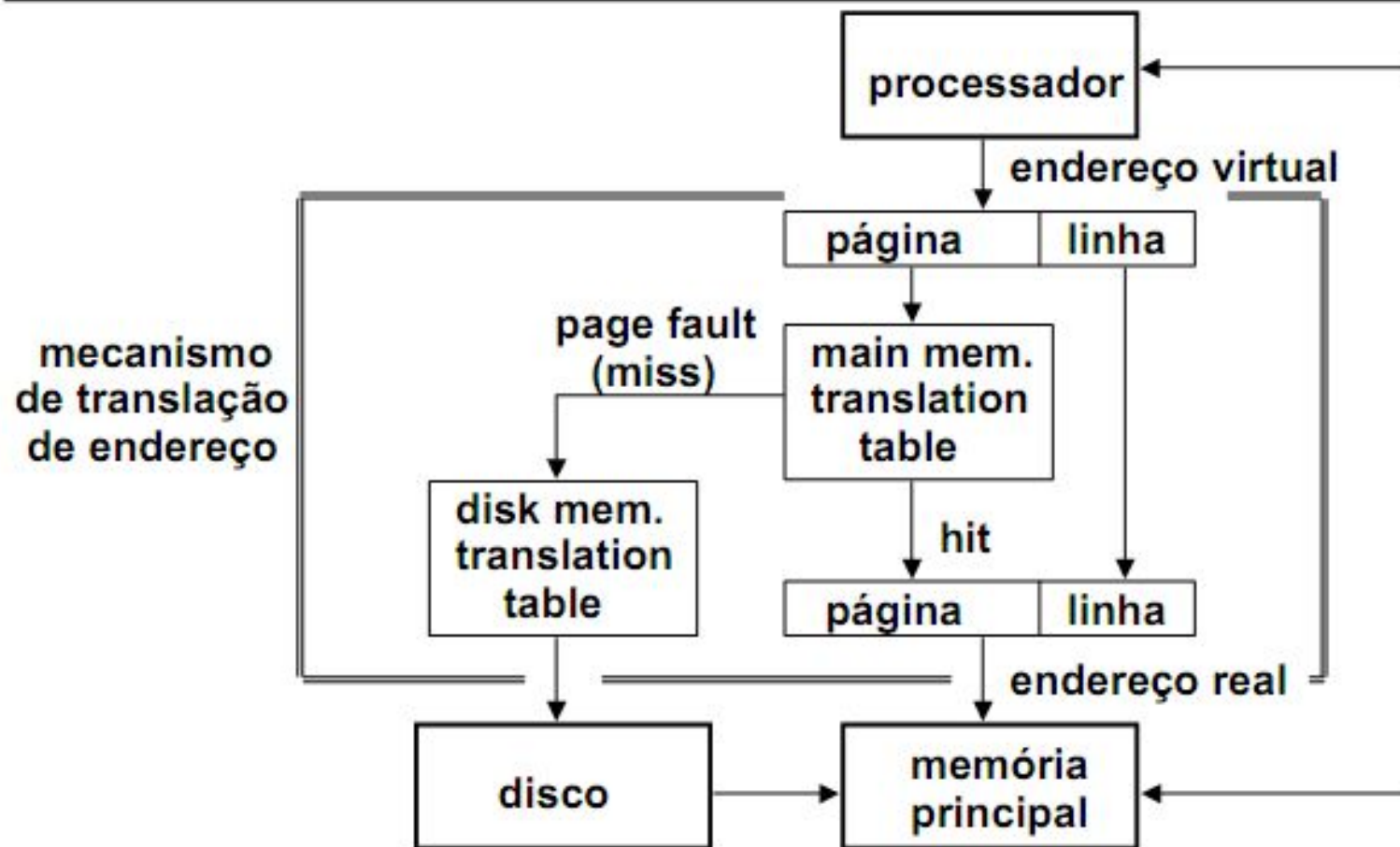
- Falhas de página: Os dados não estão na memória; recupere-os do disco
 - enorme penalidade de falha; portanto, as páginas devem ser bastante grandes (por exemplo, 4KB)
 - é importante reduzir as falhas de página (LRU vale a pena)
 - pode manipular as falhas no software em vez de no hardware
 - como a escrita direta é muito onerosa, usamos escrita adiada

Write-back



Penalidade
por falta de página:
consome milhões
de ciclos de clock

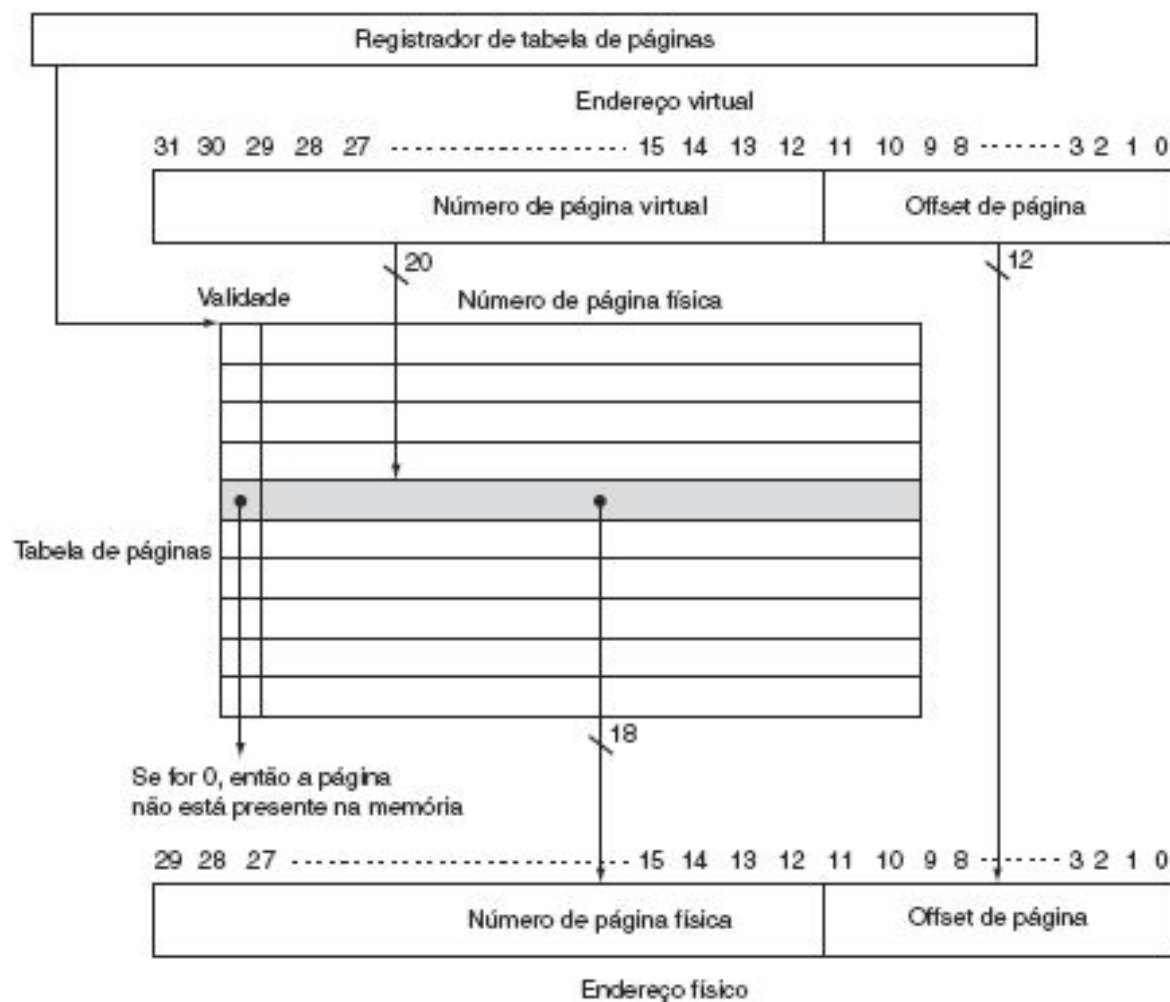
Paginação



Colocação e Localização de uma página na memória principal

- Diante da penalidade a ser paga por uma falta de página:
 - A redução de faltas por meio de otimização é um dos objetivos dos projetistas de sistemas da memória virtual.
- Tabela de Páginas
 - As páginas são colocadas por meio de uma tabela que indexa totalmente a memória principal.
 - Situa-se na memória principal.

Tabelas de página



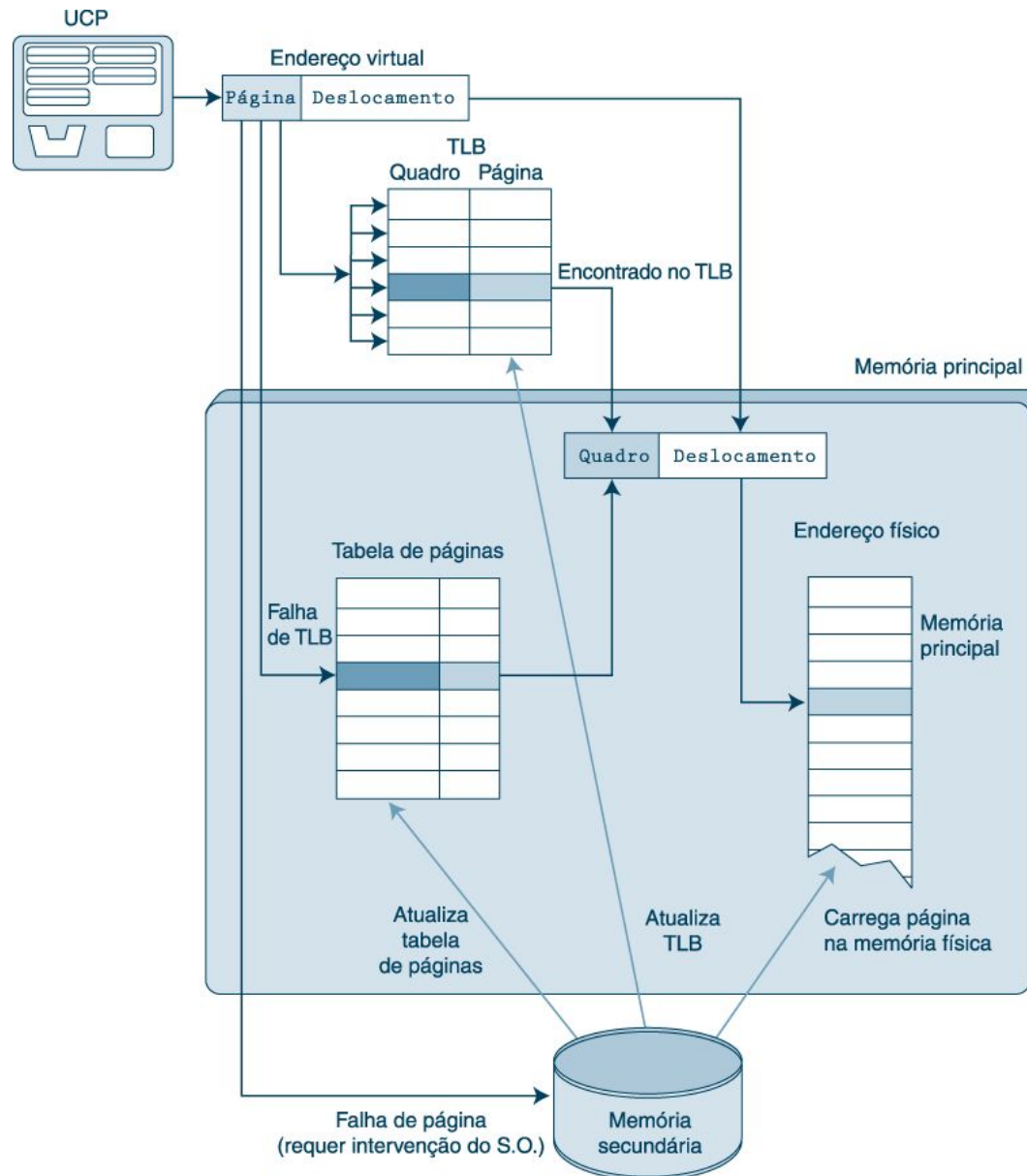
Melhoria no processo de tradução de endereço: TLB

- As referências as palavras nas páginas apresentam tanto localidade espacial quanto temporal.
- Máquinas modernas apresentam uma cache especial que armazena traduções mais recentes
- Esta cache, de uso exclusivo do processo de tradução de endereços é a TLB:
 - Translation lookaside buffer

Melhoria no processo de tradução de endereço: TLB

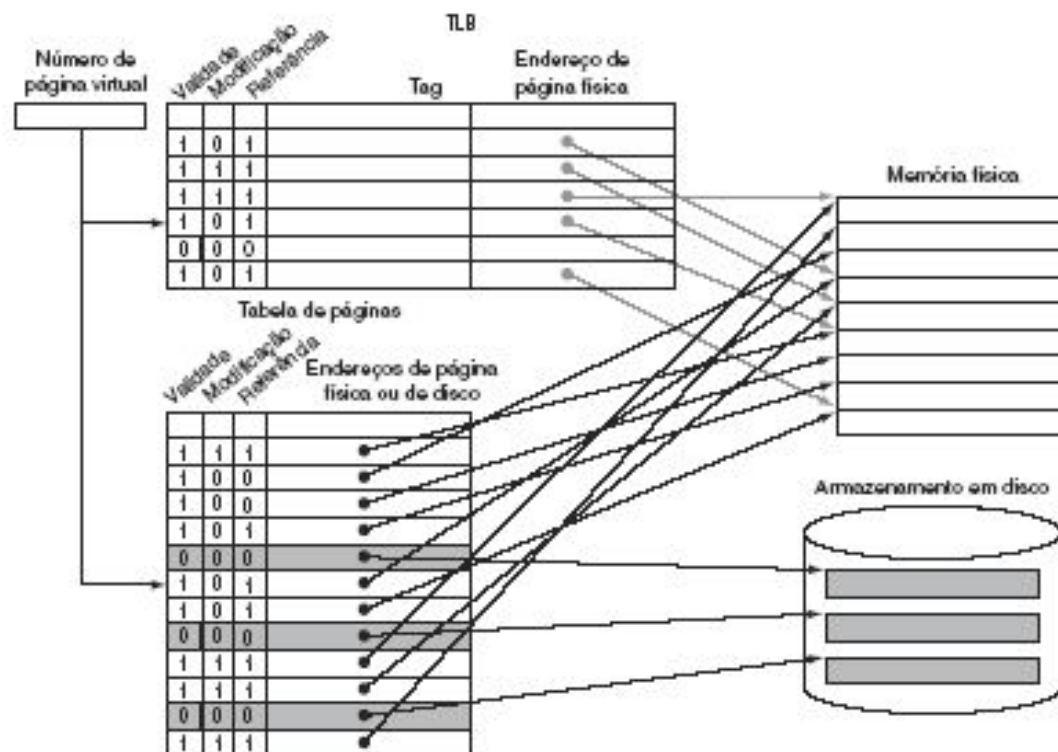
- As TLBs são utilizadas para acelerar a busca da tabela de páginas, armazenando os valores mais recentes das páginas pesquisadas na cache da tabela de páginas;
- Normalmente a TLB é implementada como cache associativa;

Exemplo de TLB sem Cache



Tornando a tradução de endereço rápida

- Um cache para traduções de endereço: translation-lookaside buffer (TLB)



Valores típicos: 16-512 entradas,
taxa de falhas: 0,01% – 1%
penalidade de falhas: 10-100 ciclos

TLBs e CACHES

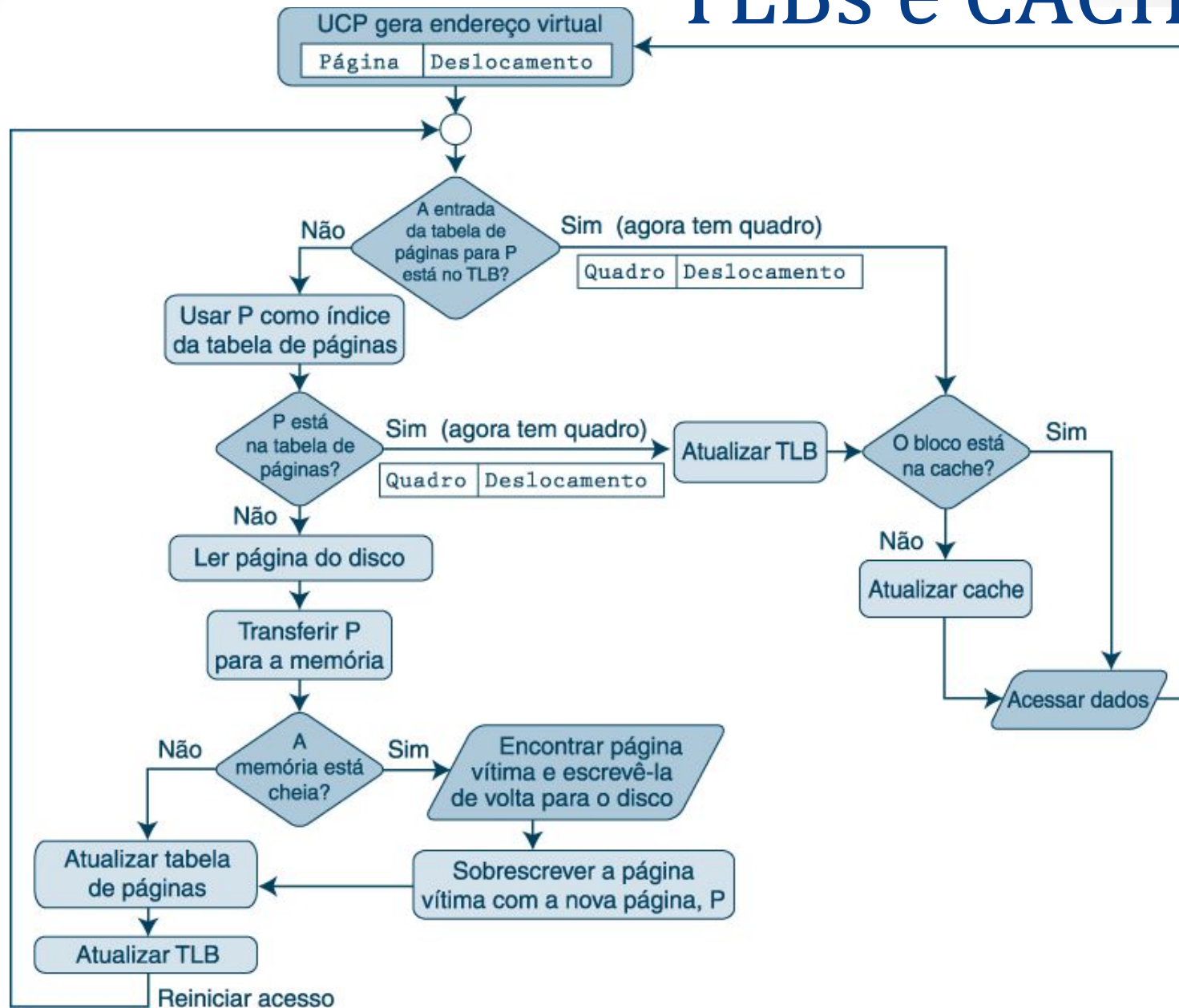
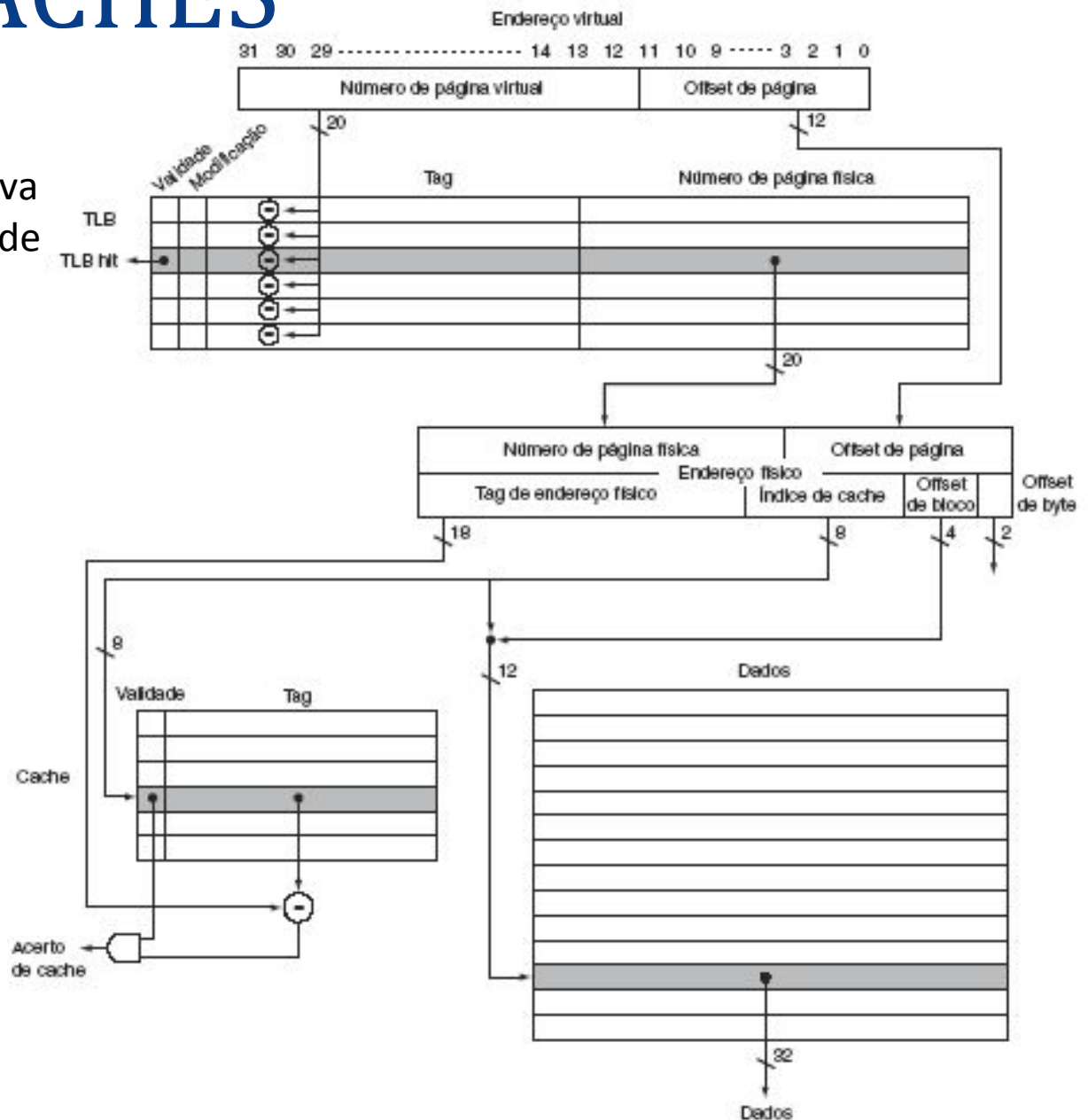


FIGURA 6.20 Reunindo tudo: TLB, tabela de páginas, cache e memória principal.

TLBs e CACHES

- TLB totalmente associativa
- Diagrama para operação de Leitura



Exemplo do Pentium

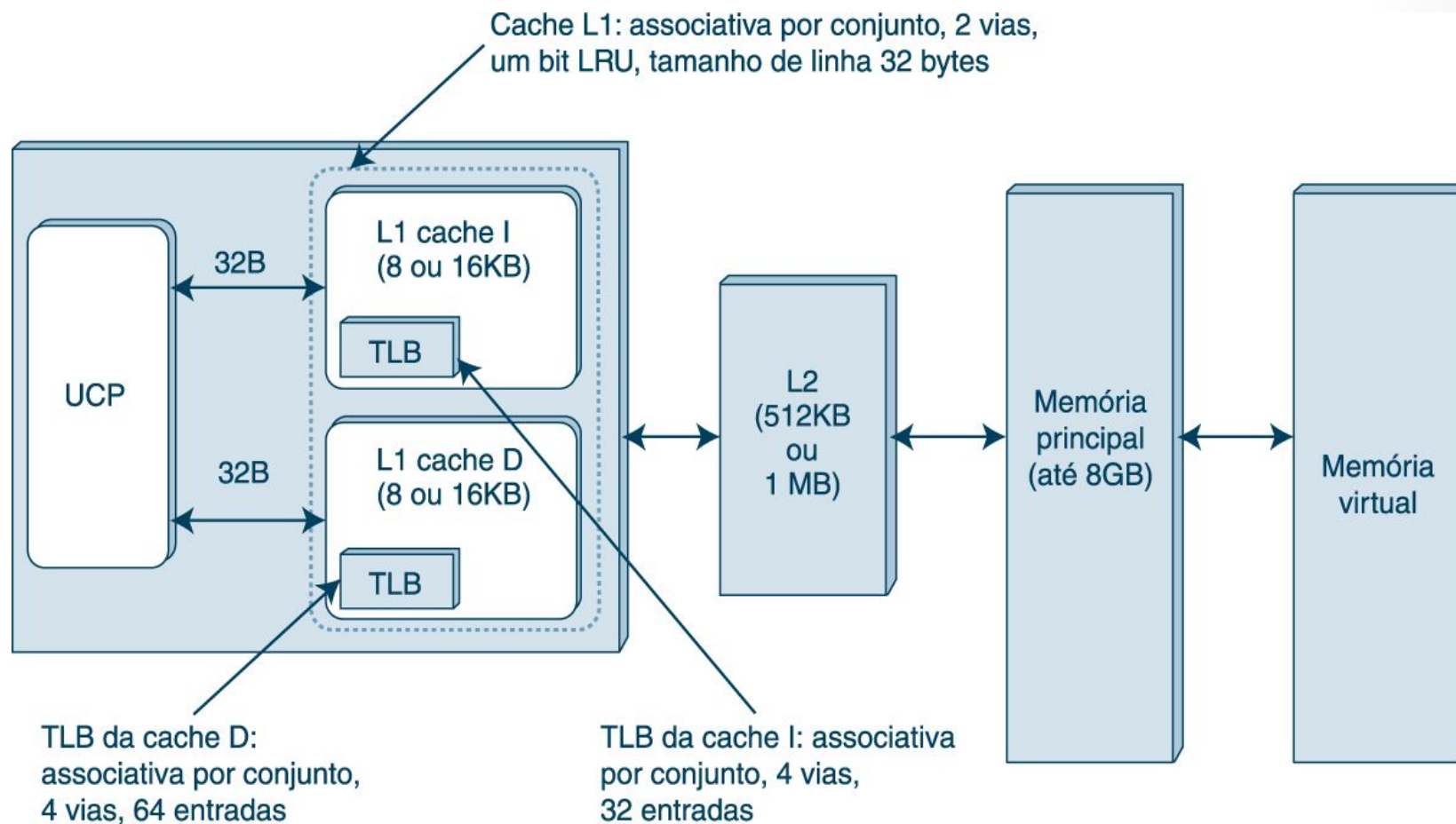


FIGURA 6.21 Hierarquia de memória do Pentium.

Exemplo:

Suponha um endereço de 8 bits no espaço virtual e memória física de 4 quadros de páginas (sem cache). E assuma que as páginas têm tamanho de 32 bytes. E o endereço da memória física tem 7 bits (4x32B). Suponha que algumas páginas foram trazidas para a memória principal. Suponha que o endereço virtual gerado é 13d. Qual o endereço físico do byte na memória principal?

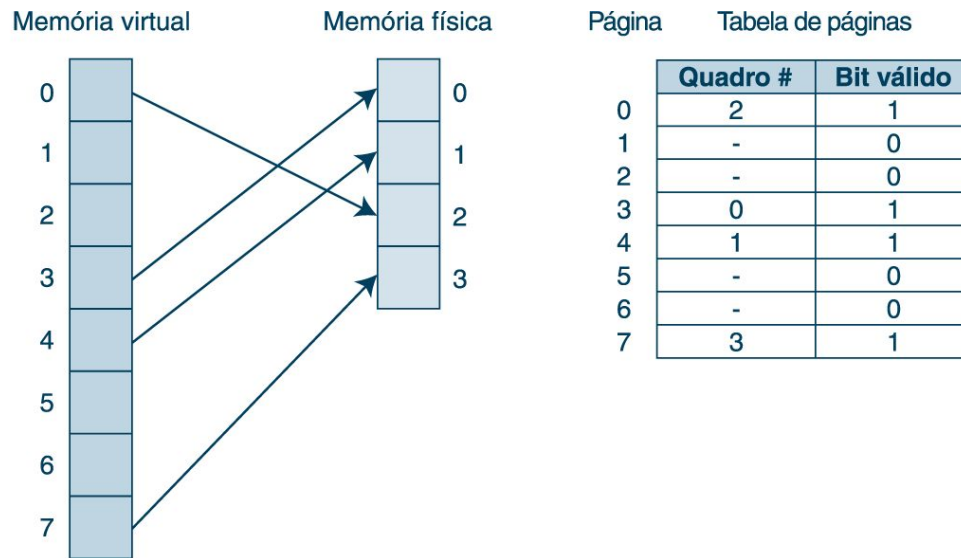


FIGURA 6.12 Estado atual usando paginação e tabela de páginas associada.

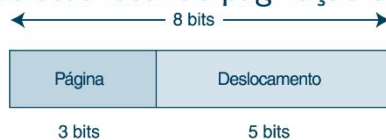


FIGURA 6.13 Formato para um endereço virtual de 8 bits com tamanho de página de $2^5 = 32$ palavras.

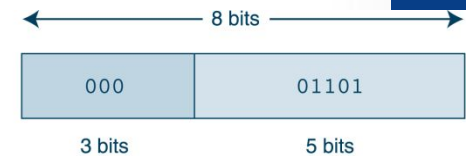


FIGURA 6.14 Formato para o endereço virtual $00001101_2 = 13_{10}$.

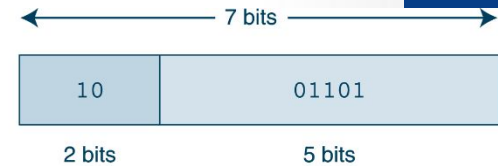


FIGURA 6.15 Formato para o endereço físico $1001101_2 = 77_{10}$.

Exemplo (cont.):

Suponha agora que o endereço virtual gerado é 127.
Qual o endereço físico do byte na memória principal?

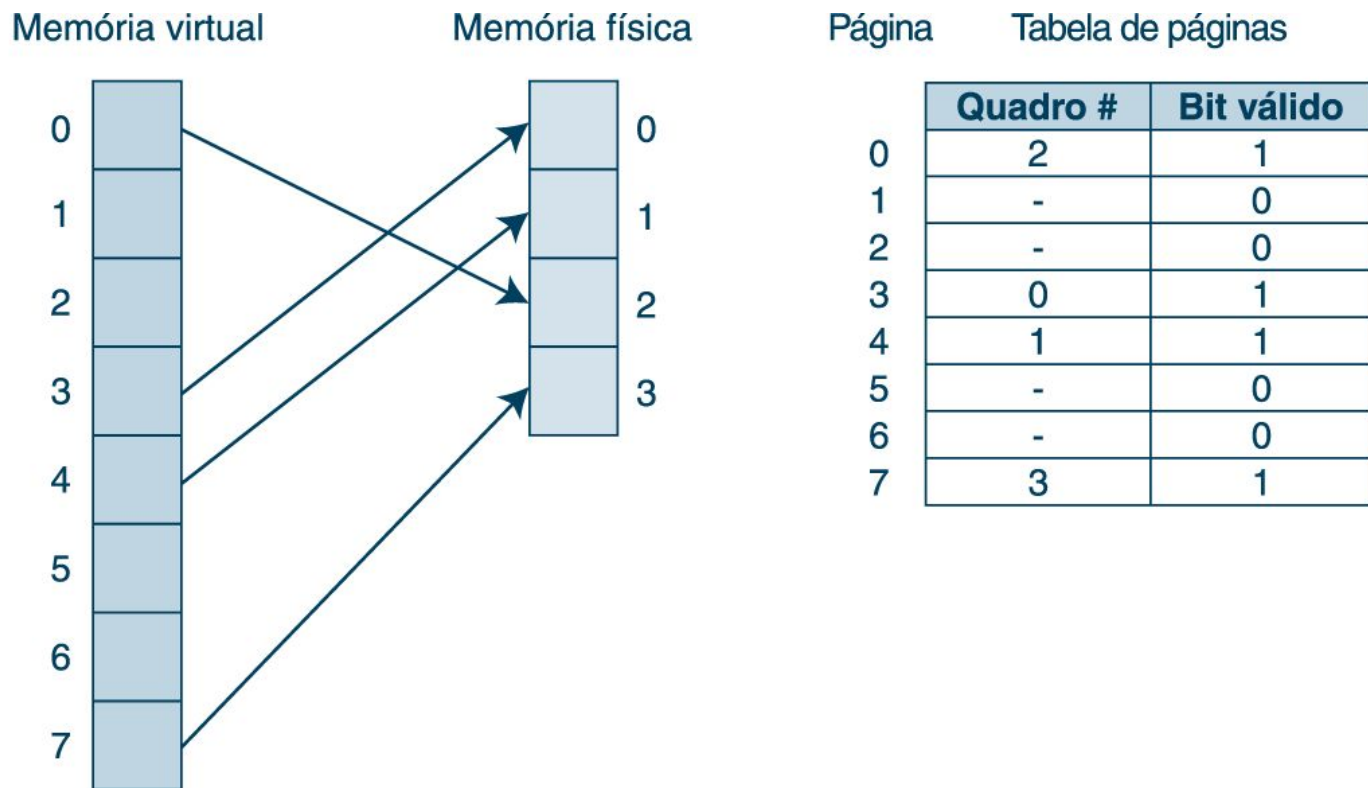


FIGURA 6.12 Estado atual usando paginação e tabela de páginas associada.

Segmentação

- O outro método de memória virtual é o por segmentação.
- Nesse caso, em vez de dividir o espaço de endereçamento virtual em páginas de tamanho fixo e o endereçamento físico em quadros de páginas de mesmo tamanho, o espaço de endereçamento virtual é dividido em segmentos;

Segmentação

- Na segmentação, o espaço de endereçamento virtual é dividido em unidades lógicas de tamanho variável.
- Quando um segmento precisa ser copiado para a memória física, o sistema operacional procura por uma porção de memória livre grande o suficiente para armazenar todo o segmento.