

Universidade Federal de Pelotas

Instituto de Física e Matemática

Departamento de Informática

Bacharelado em Ciência da Computação

Arquitetura e Organização de Computadores II Aula 16

4. Memória virtual: Paginação, segmentação, fragmentação. Colocação de uma página na memória principal, faltas de página, escrita de uma página.

Prof. José Luís Güntzel

guntzel@ufpel.edu.br

www.ufpel.edu.br/~guntzel/AOC2/AOC2.html

Introdução

- Memória Principal (DRAM)
 - Tamanho limitado (máx, 1GB)
 - Tempo de acesso da ordem de nanossegundos
- Memória Secundária (disco magnético)
 - Grande capacidade de armazenamento (tipicamente, 80-160 GB)
 - Tempo de latência da ordem de milissegundos
- Pode não ser possível manter um programa inteiro na memória principal. Pior: deseja-se rodar vários programas "ao mesmo tempo"
 - Solução: manter apenas as partes ativas do(s) programa(s)

Memória Virtual

• Memória Virtual é a técnica que dá ao programador a ilusão de poder acessar rapidamente um grande espaço de endereçamento

Objetivos da técnica:

- Permitir que haja um meio seguro e eficiente de se compartilhar informações, armazenadas na memória, entre vários programas
- Minimizar os problemas causados aos programas pela existência de uma pequena quantidade de memória principal

Memória Virtual

- Os programas que compartilham a memória de determinada máquina mudam dinamicamente durante o processo de execução
- Cada programa deve ser compilado usando seu próprio espaço de endereçamento (ou seja, em uma região da memória acessível somente a esse programa)
- A técnica de memória virtual realiza a tradução do espaço de endereçamento de um programa para seus endereços reais

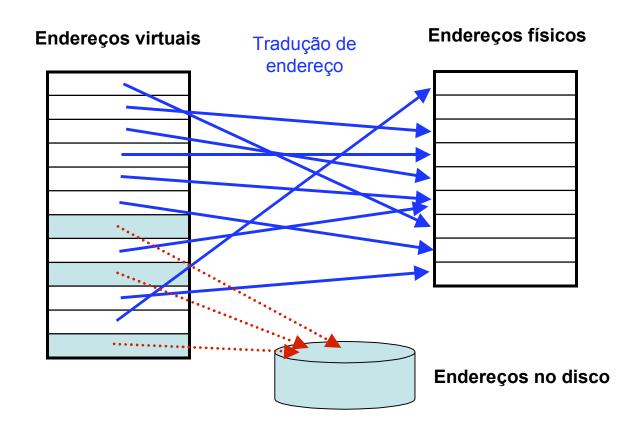
Memória Virtual

- A técnica de memória virtual permite que o tamanho de um único programa exceda a quantidade total de memória real disponível para sua execução
- A técnica de memória virtual gerencia automaticamente os dois níveis de hierarquia: memória principal (física) e memória secundária
- Antigamente,os programas eram divididos em pedaços.
 Os pedaços mutuamente exclusivos (overlays) eram identificados. Carga/exclusão de overlays da memória era realizada sob controle do próprio programa

Conceitos Envolvidos em Memória Virtual

- Página: é um bloco de tamanho fixo
- Falta de página: é uma falta no acesso à memória virtual
- Usando memória virtual: o processador sempre gera um endereço virtual, que é traduzido para um endereço real por meio da MMU (memory management unit), que é um sistema HW+SW
- O endereço real (também chamado de físico) é, então, usado para acessar a memória

Memória Virtual

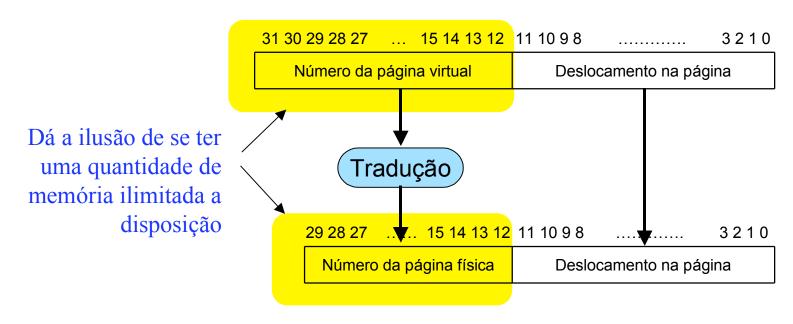


Memória Virtual e Relocação de Código

- A técnica de memória virtual simplifica a carga dos programas para execução, a partir da relocação
- A relocação mapeia os endereços virtuais usados por um determinado programa em endereços físicos (antes de tais endereços serem usados para acessar a memória)
- A técnica de relocação permite que um programa seja carregado em qualquer posição da memória principal
- Todos os sistemas de memória virtual atuais relocam os programas por meio de blocos de tamanho fixo, as páginas

Mapeamento de um endereço virtual em endereço físico

Endereço virtual



Endereço físico

Projeto de Memória Virtual

- A ocorrência de uma falta de página consome milhões de ciclos de relógio
- Este tempo é dominado pelo tempo para obter o primeiro endereço de uma página de tamanho típico

A penalidade imensa de uma falta de página conduz a decisões importantes, com relação ao projeto de sistemas de memória virtual...

Características dos Sistemas de Memória Virtual

- As páginas devem ser grandes o suficiente para amortizar o tempo de acesso muito grande (atualmente, de 16KB a 64KB)
- Uma técnica de colocação totalmente associativa tende a proporcionar uma baixa taxa de falta de página
- As faltas de página podem ser tratadas por software, em função do *overhead* introduzido ser pequeno, quando comparado ao tempo de acesso ao disco (permite o uso de algoritmos mais eficientes para a escolha de como colocar as páginas)
- Os sistemas de memória virtual usam write-back. (Writethrough não funciona bem porque as escritas são muito demoradas.)

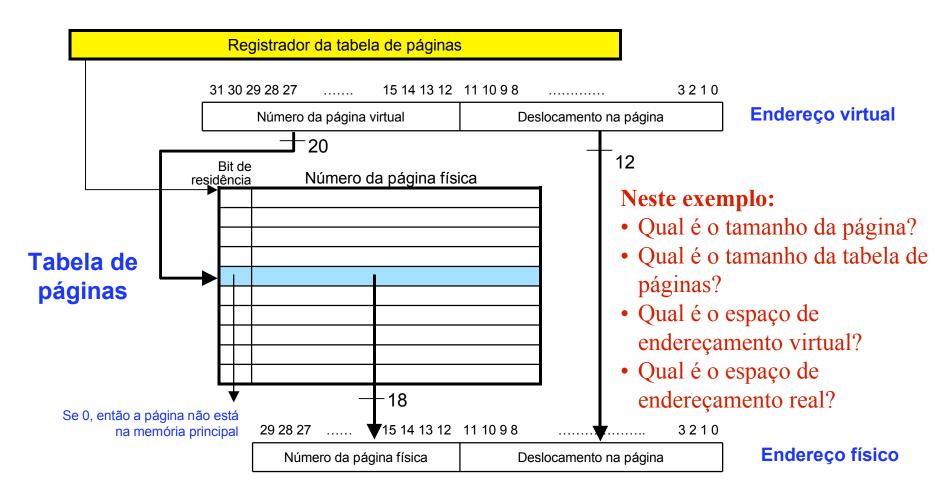
Colocação e Localização de uma Página na Memória

- A redução das faltas de página é o principal objetivo no projeto de um sistema de memória virtual
- Esta redução é obtenível através do uso de técnicas otimizadas para a colocação das páginas na memória principal
- Associatividade total é usada, atualmente:
 - (+) muita flexibilidade para substituição de páginas,
 - (-) dificuldade na localização das páginas

A Tabela de Páginas

- Permite a localização das páginas
- Indexa totalmente a memória principal
- Armazenada na memória principal
- É indexada com o número da página extraído do endereço virtual e contém o número da página física correspondente
- · Cada programa possui sua própria tabela de páginas
- Registrador da tabela de páginas: em HW, aponta para a posição inicial da tabela de páginas na memória principal
- Por hora, assumamos que a tabela de páginas está em uma posição fixa da memória principal, utilizando endereços contíguos

Tabela de Páginas



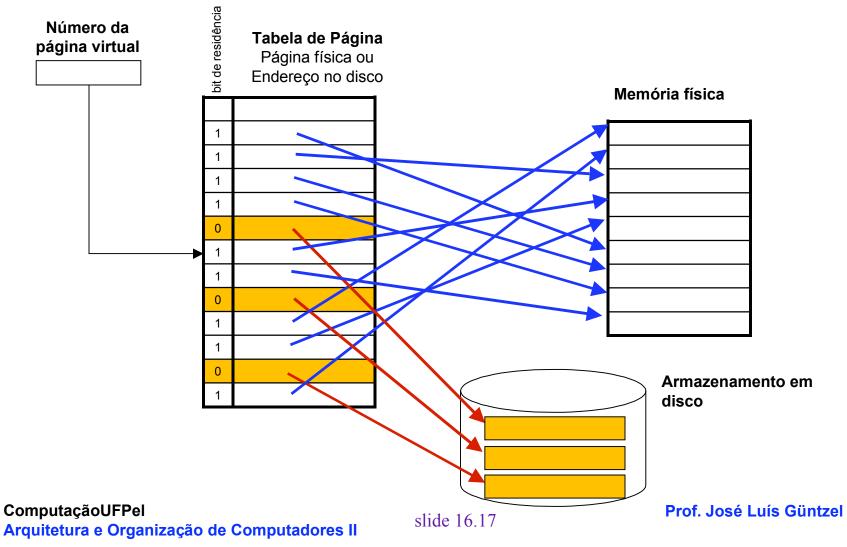
Faltas de Página

- Bit de residência = 0 indica falta de página
- Sistema operacional assume o controle, por meio do mecanismo de exceção
- O sistema operacional precisa
 - Encontrar a página faltante no nível hierárquico inferior (geralmente, no HD)
 - Decidir em que lugar da memória principal deve ser colocada a página requisitada
- O endereço virtual, por si só, não informa em que posição do HD está a página que gerou a falta de página

Faltas de Página

- O SO cria espaço em disco para todas as páginas virtuais de um processo, quando da criação do processo
- Neste momento, o SO também cria uma estrutura de dados para controlar onde cada página virtual está guardada no disco
- Tal estrutura pode ser parte da tabela de páginas, ou pode ser uma estrutura auxiliar, indexada da mesma maneira que a tabela de páginas

Tabela de Páginas



Escrita

- · A escrita no disco consome milhões de ciclos de relógio
- Esquema write-through não funciona
- Esquema write-back é usado: página é copiada para o disco no momento em que for substituída (nomenclatura: "copy-back")