Laboratório de Pesquisa em Redes e Multimídia

Sistemas Operacionais

Gerência de Memória - Algoritmos de substituição de páginas

Universidade Federal do Espírito Santo
Departamento de Informática

Introdução

- Quando ocorre um *Page Fault*, o S.O. deve escolher que página remover para abrir espaço em memória.
- Se a página foi alterada (bit *Modified* setado) é preciso salvá-la em disco. Se não foi, basta sobrescrevê-la.
- É melhor não escolher para remoção uma página que é usada freqüentemente, pois ela pode ter que voltar para a memória logo.
- Troca ótima de página
  - Substituir a página para a qual falta mais tempo até ser necessária novamente
    - Marcar p/ cada página, quantas instruções faltam p/ que ela seja referenciada
  - Solução ótima, mas inviável!

Algoritmo NRU – *Not Recently Used* (1)

- Ou seja, algoritmo de substituição da página "não usada recentemente"
- Na maioria dos computadores com memória virtual, as entradas nas tabelas de páginas têm 2 bits de status
  - *Reference* bit (R) ; *Modified* bit (M)
- Algoritmo
  - Qdo o processo é iniciado, os bits R e M das páginas são zerados
  - Bits são sempre alterados quando a página é referenciada/modificada
  - Periodicamente o bit R é zerado (por exemplo, a cada tique de clock)
  - Quando acontece um *Page fault*, o S.O. inspeciona todas as páginas que encontram-se na memória e as separa em categorias...

Algoritmo NRU – *Not Recently Used* (2)

Isso pode ocorrer???

- Páginas são classificadas
  - Classe 0: *Not referenced, not modified* (R=0 , M=0)
  - Classe 1: *Not referenced, modified* (R=0 , M=1)
  - Classe 2: *referenced, not modified* (R=1 , M=0)
  - Classe 3: *referenced, modified* (R=1 , M=1)
- O S.O. remove uma das páginas (aleatoriamente) da classe mais baixa não vazia.
- Vantagens
  - Algoritmo fácil de entender e implementar
  - Desempenho adequado

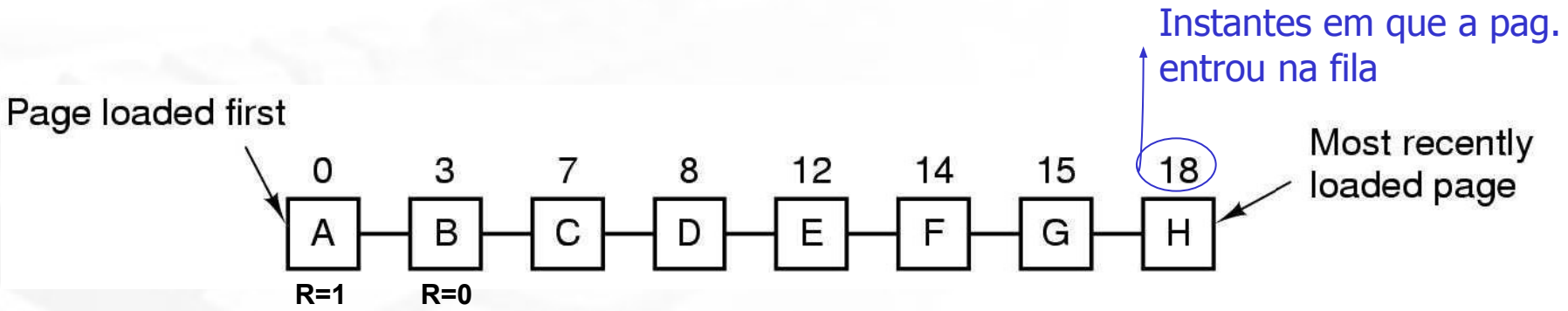
Algoritmo FIFO

- Mantém-se uma lista encadeada de páginas ordenada pela chegada das páginas à memória.
- Quando ocorre um *Page Fault*, a página no início da lista (que é a mais antiga) é a escolhida para a troca
- Vantagem:
  - Baixo custo
- Desvantagem:
  - A página mais antiga pode ser também uma página usada muito freqüentemente.
- Não empregado!

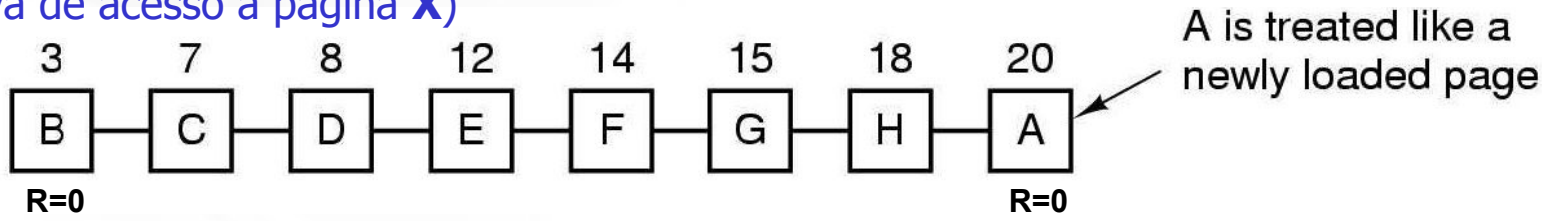
Algoritimo SC - Segunda Chance (1)

- Tenta melhorar o FIFO
- Cada página tem um bit R (referenciada)
- Antes de remover a página mais antiga (cabeça da fila), seu bit R é verificado
  - Se R=0, a página é substituída (a página referenciada ocupará o seu lugar na memória)
  - Se R=1, a página vai para fim da fila, como se houvesse sido carregada agora e seu bit é setado para 0
    - Verifica-se a página que virou "cabeça" da fila
- Se todas as páginas tiverem seu bit R=1, haverá uma volta completa

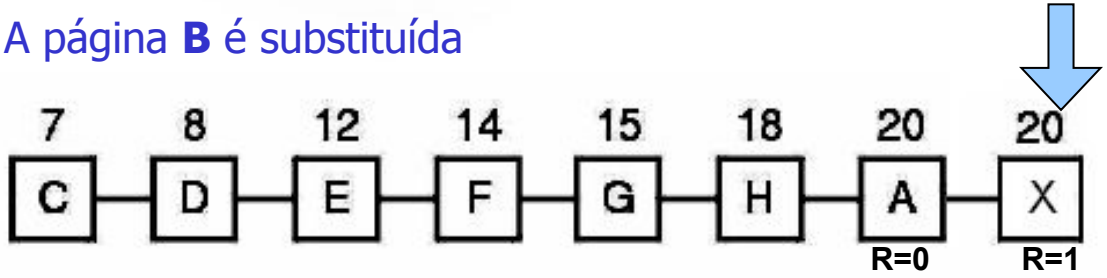
Algoritimo SC - Segunda Chance (2)



No instante 20 ocorre um Page Fault
(tentativa de acesso à página **X**)

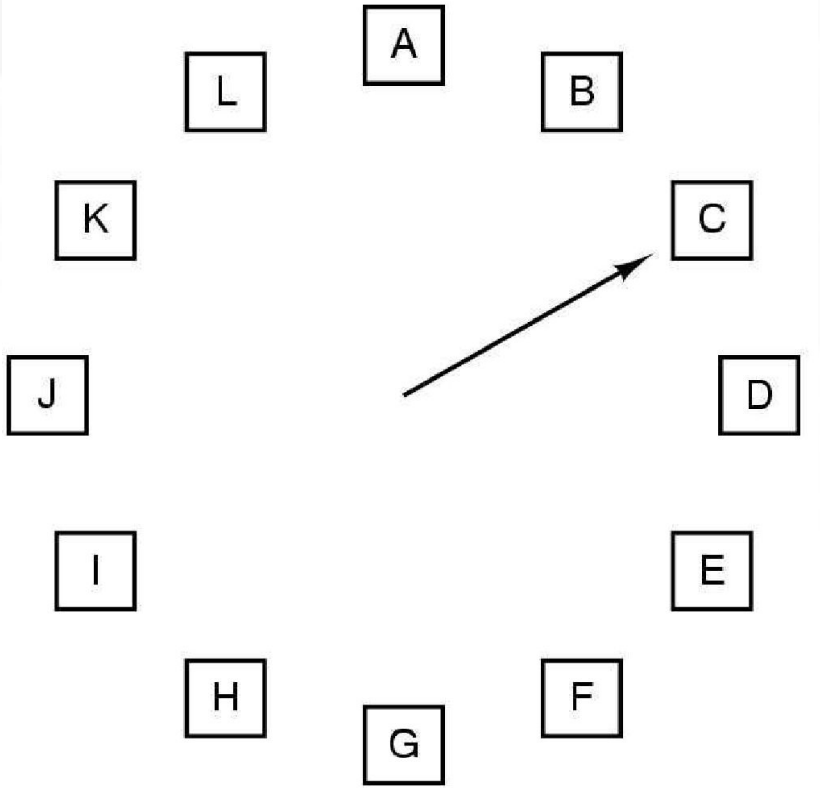


A página **B** é substituída



Algoritmo do Relógio

- Visa melhorar o desempenho do algoritmo SC, diferenciando apenas na implementação da fila



- O ponteiro sempre aponta para a página mais antiga
- Na ocorrência de um Page fault
  - Se o bit R desta página for 0, ela é substituída, e o ponteiro "roda" uma casa
  - Se R=1, R é resetado e o ponteiro avança para a próxima página até encontrar uma página com R=0

Algoritmo LRU – *Least Recently Used* (1)

Ou MRU – Menos Recentemente Usada

- Assume que as páginas usadas recentemente voltarão a ser usadas em breve
  - Substitui páginas que estão há **mais tempo sem uso.**
  - **PRINCÍPIO DA LOCALIDADE TEMPORAL**
    - Uma página acessada mais recentemente, tem mais chances de ser acessada novamente

Algoritmo LRU – *Least Recently Used*

Ou MRU – Menos Recentemente Usada

- Na **teoria**, para implementá-lo completamente, deveria-se manter uma lista encadeada de todas as páginas que estão na memória (muito custoso!)
  - página usada mais recentemente vai para o início da lista;
    - lista é reordenada a cada referência a memória
  - qdo há Page Fault, escolhe-se a última página da fila

NA PRÁTICA, há diferentes formas de implementá-lo!

Algoritmo LRU – *Least Recently Used*

Possíveis Implementações

- Auxílio de Hardware
  - Contador incrementado a cada instrução
  - Matriz mapeando memória física
- Simulando LRU em Software
  - LFU
  - Aging

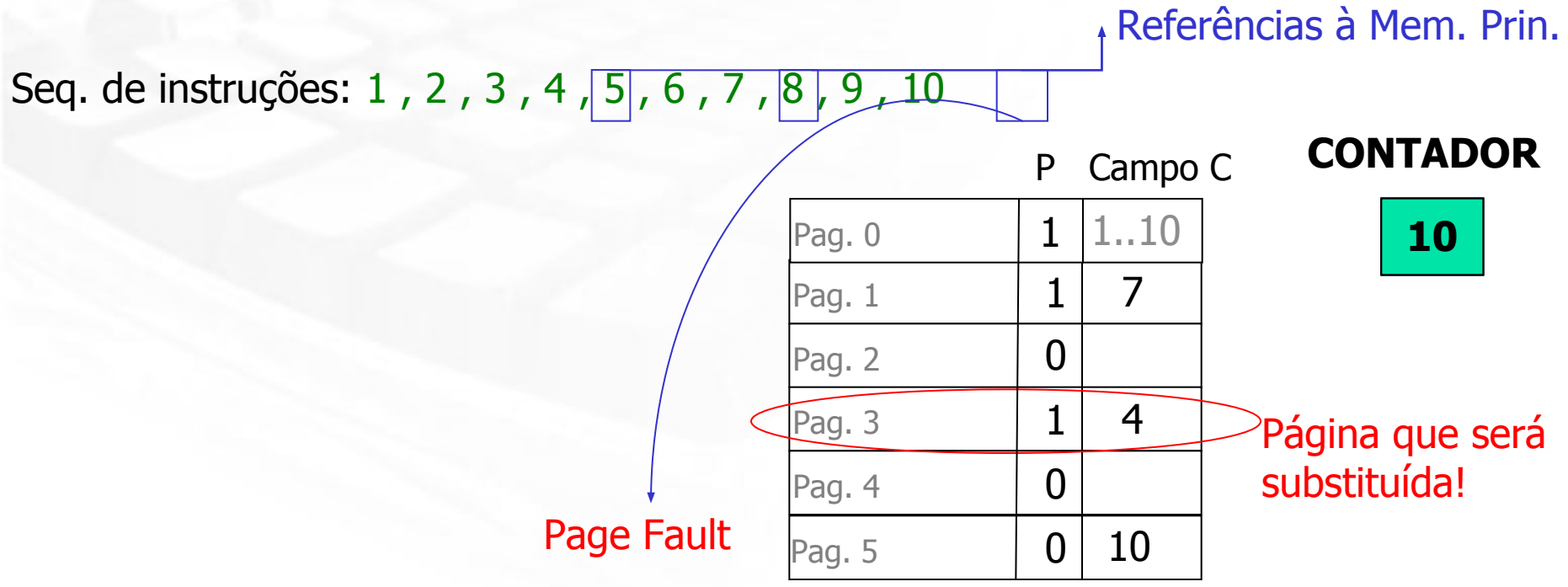
Algoritmo LRU – *Least Recently Used*

Em Hardware (1)

- Uma solução simples: manter uma idade para cada página.
  - Usar um contador C de 64 bits incrementado a cada instrução (em hardware)
  - Cada entrada da tabela de páginas deve ter um campo extra para armazenar o valor do contador
  - A cada referência à memória o valor corrente de C é armazenado na entrada da tabela de páginas na posição correspondente à página referenciada
  - Quando ocorre um Page Fault, a tabela de páginas é examinada, a entrada cujo campo C é de menor valor é a escolhida
  - Substitui página com o menor valor no campo do contador (maior idade)

Algoritmo LRU – *Least Recently Used*

Em Hardware (2)



Algoritmo LRU – *Least Recently Used*

Em Hardware (3)

- LRU usando matrizes
  - HW especial que mantém uma matriz **n x n**, onde **n** é o número de molduras
  - Inicialmente todos os bits da matriz são 0
  - Sempre que a moldura k é referenciada, o hardware seta todos os bits da linha k para 1, e depois zera todos os bits da coluna k para zero
  - Deste modo, a qualquer instante a **linha** com o **menor valor binário** é a menos recentemente usada

Algoritmo LRU – *Least Recently Used*

Em Hardware (4)

- LRU usando matrizes (cont.)

Página na moldura 0

	0	1	2	3
0	0	1	1	1
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0

Página na moldura 1

0	1	2	3
0	0	1	1
1	0	1	1
0	0	0	0
0	0	0	0

Página na moldura 2

0	1	2	3
0	0	0	1
1	0	0	1
1	1	0	1
0	0	0	0

Página na moldura 3

0	1	2	3
0	0	0	0
1	0	0	0
1	1	0	0
1	1	1	0

Página na moldura 2

0	1	2	3
0	0	0	0
1	0	0	0
1	1	0	1
1	1	0	0

Página na moldura 1

0	0	0	0
1	0	1	1
1	0	0	1
1	0	0	0

Página na moldura 0

0	1	1	1
0	0	1	1
0	0	0	1
0	0	0	0

Página na moldura 3

0	1	1	0
0	0	1	0
0	0	0	0
1	1	1	0

Página na moldura 2

0	1	0	0
0	0	0	0
1	1	0	1
1	1	0	0

Página na moldura 3

0	1	0	0
0	0	0	0
1	1	0	0
1	1	1	0

Algoritmo LRU – *Least Recently Used*

Simulando em Software (1)

- Problema das abordagens em HW
  - Dependem de um HW especial
  - Procurar uma solução em SW
- Simulando LRU em Software
  - **Algoritmo NFU** - Not Frequently Used
    - Um contador por página na memória
    - A cada tick, o S.O. percorre todas as páginas na memória e soma o bit R (0 ou 1) de cada página ao seu respectivo contador
    - Na ocorrência de *Page Fault*, a página c/ o menor contador é substituída
    - Problema: o algoritmo nunca esquece (reseta) o contador

Algoritmo LRU – *Least Recently Used*
Simulando em Software (2)

- **Algoritmo Aging**
  - Após cada período/TICK
    - Desloca o contador de 1 bit p/ a direita
    - Soma R ao bit mais significativo do contador
      - Feitas as somas, os bits R de cada página/frame são "zerados"

Página	Bits R para páginas 0-5 em t0	Bits R para páginas 0-5 em t1	Bits R para páginas 0-5 em t2	Bits R para páginas 0-5 em t3	Bits R para páginas 0-5 em t4
	1 0 1 0 1 1	1 1 0 0 1 0	1 1 0 1 0 1	1 0 0 0 1 0	0 1 1 0 0 0
0	10000000	11000000	11100000	11110000	01111000
1	00000000	10000000	11000000	01100000	10110000
2	10000000	01000000	00100000	00100000	10001000
3	00000000	00000000	10000000	01000000	00100000
4	10000000	11000000	01100000	10110000	01011000
5	10000000	01000000	10100000	01010000	00101000

Aspectos de Projeto

- Já vimos alguns algoritmos de substituição de páginas...
- Mas para estudarmos um dos mais utilizados, devemos entender alguns **aspectos de projeto** do sistema de paginação!

Políticas de Busca de Páginas de um Processo

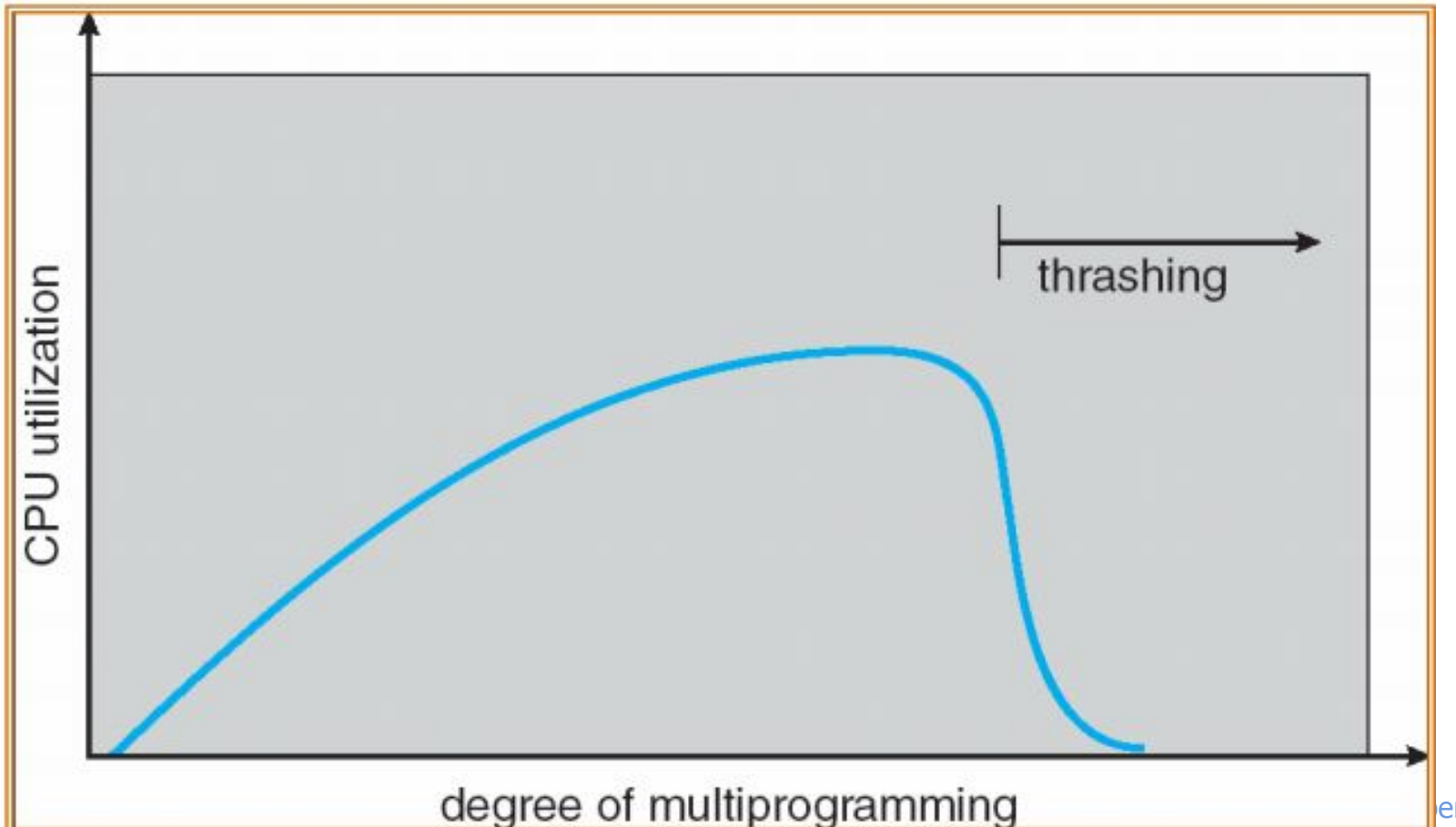
- Determina em que instante uma página deve ser trazida para memória principal
  - O objetivo é minimizar o número de faltas de página
- **Paginação por demanda**
  - No modo mais puro de paginação, os processos são iniciados sem qualquer página na memória
  - Quando a CPU tenta buscar a 1a instrução, há um *Page fault*, forçando o S.O. a carregar a página na MP
  - À medida que *Page faults* vão ocorrendo, as demais páginas são carregadas
- **Pré-paginação** (normalmente considera o ***Working Set***)

Working Set (1)

- O conjunto de páginas que um processo está atualmente usando é denominado ***Working Set*** (espaço de trabalho)
- Verifica-se que, para intervalos de tempo razoáveis, o espaço de trabalho de um processo mantém-se constante e menor que o seu espaço de endereçamento
- Se todo o *Working Set* está presente na memória, o processo será executado com poucas *Page Fault* até passar para a próxima fase do programa, quando o *Working Set* é atualizado
  - Ex: Compilador de dois passos
- Se vários processos tiverem menos páginas em memória que o seu espaço de trabalho, o sistema pode entrar em colapso (***trashing!!!***)

Working Set (2)

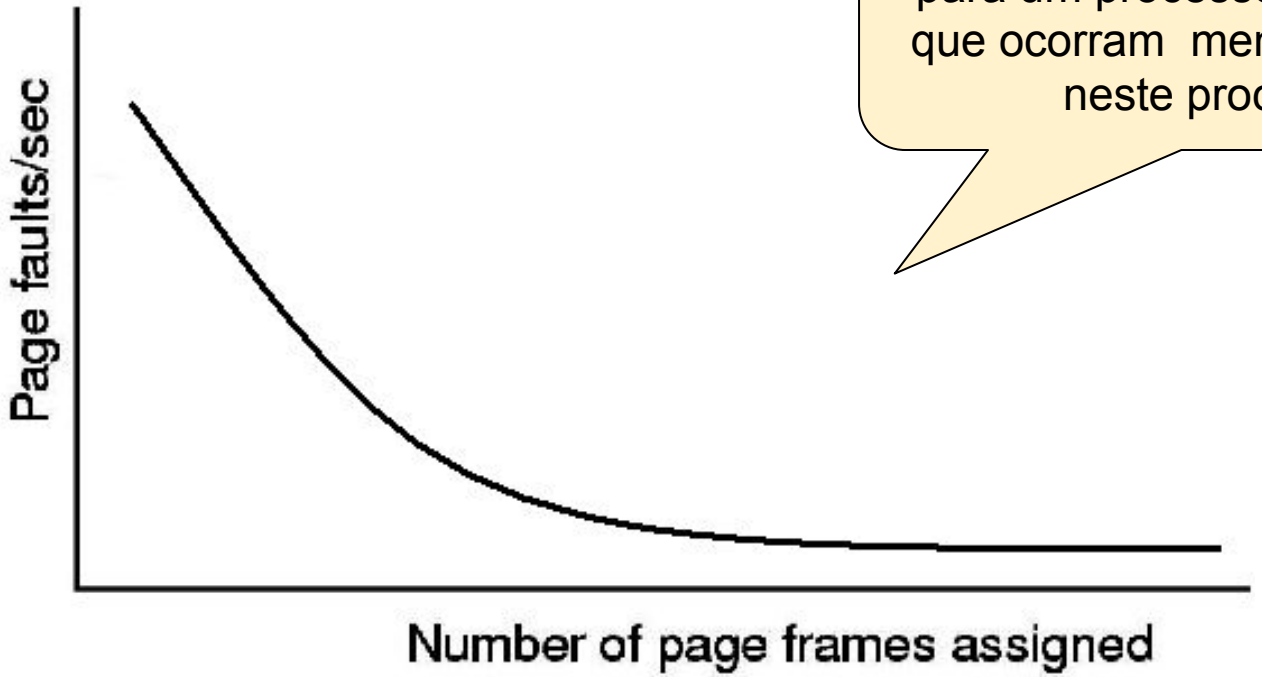
- *Trashing!!*



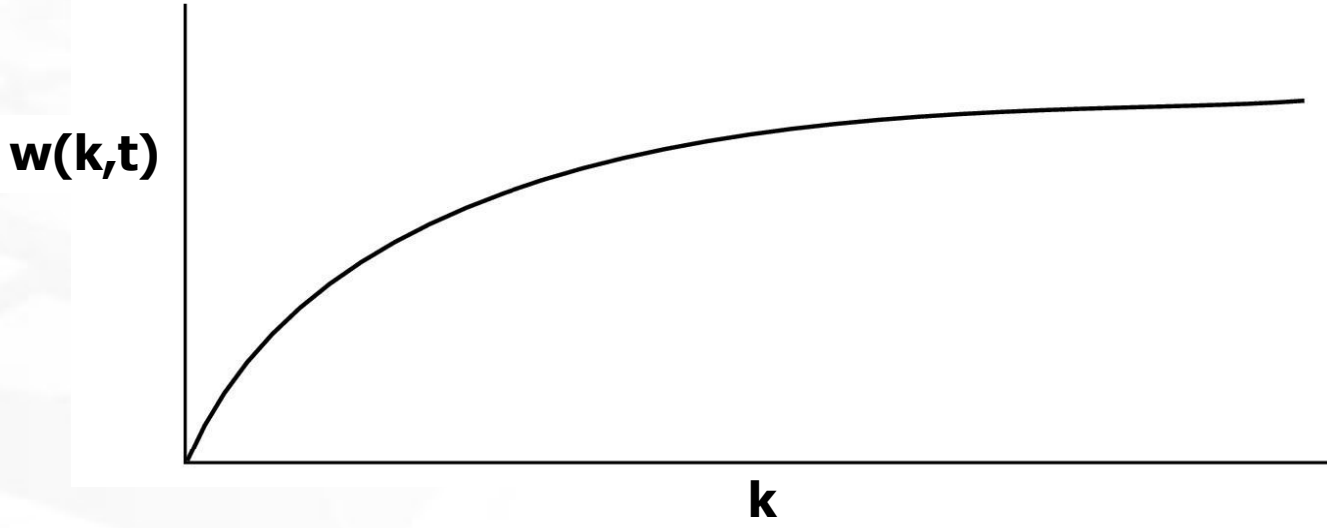
Working Set (3)

- Como prevenir o *Trashing*?

Quanto mais frames alocamos para um processo, a tendência é que ocorram menos Page Faults neste processo….

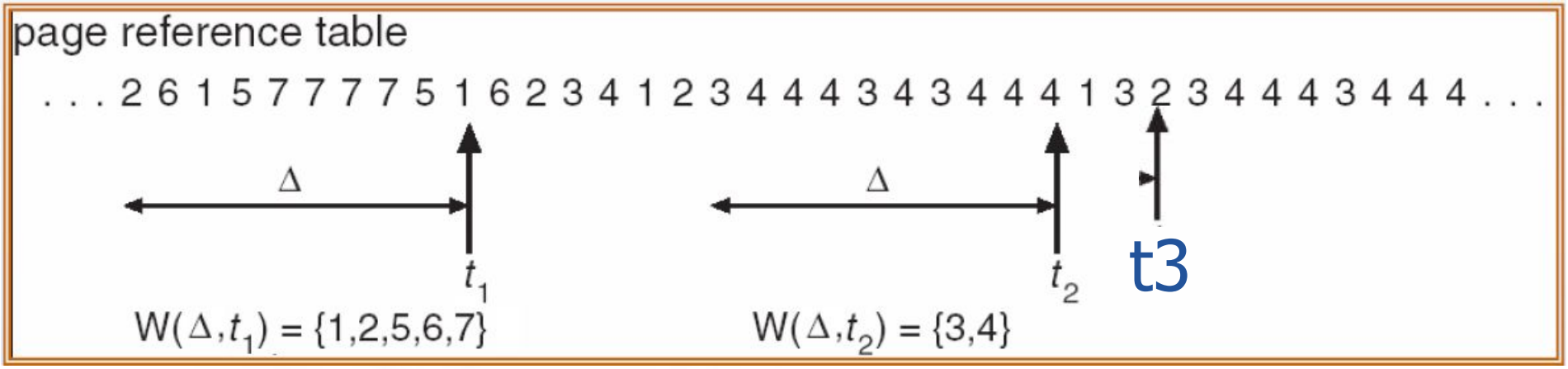


Definindo o Working Set formalmente... (1)



- O *Working Set* = as páginas usadas (referenciadas) pelas **k** referências mais recentes à memória
  - Ou aquelas usadas nos últimos **τ** segundos.
- A função **w(k,t)** [ou **w(τ,t)**] retorna a quantidade de páginas do Working Set no instante *t.*

Definindo o Working Set formalmente... (2)



0 4 3 8 50 61

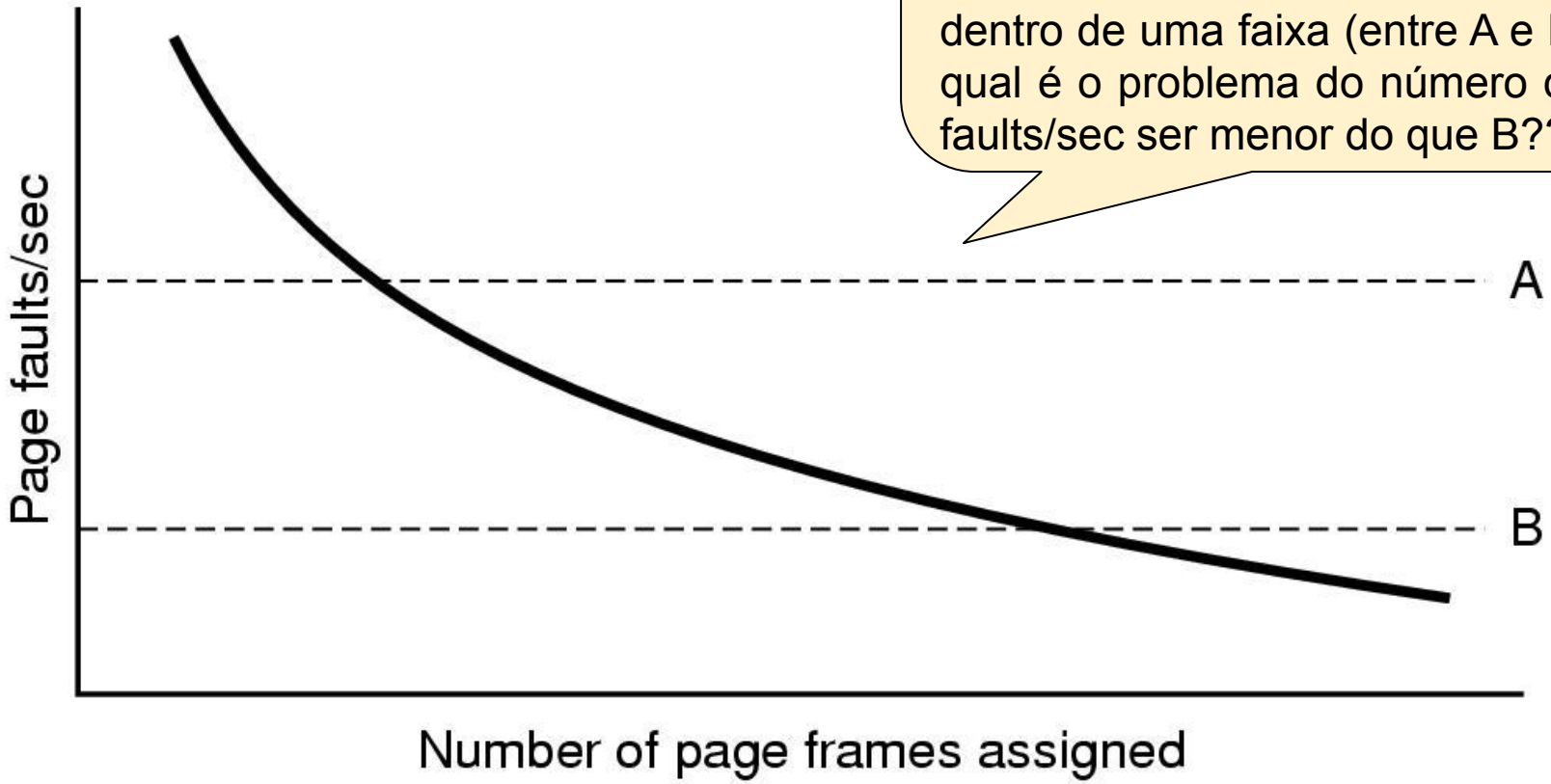
Estratégias de alocação de frames

- Alocação **fixa** de quantidade de frames:
  - cada processo recebe um número fixo de frames
  - em caso de falta de páginas, uma das residentes é trocada
- **Alocação variável:** número de páginas varia durante a execução do processo
  - Utilização de valores máximo e mínimo de dimensão do espaço de trabalho para controlar a paginação
  - Estes valores devem-se adaptar dinamicamente a cada aplicação

Considerando o Working S usar?

-

Quando alocamos mais ou menos frames para um processo, podemos observar uma diminuição ou aumento no número de page faults / sec... o interessante é manter esse número dentro de uma faixa (entre A e B). Mas qual é o problema do número de page faults/sec ser menor do que B??



Agora que você sabe o que é o Working Set...

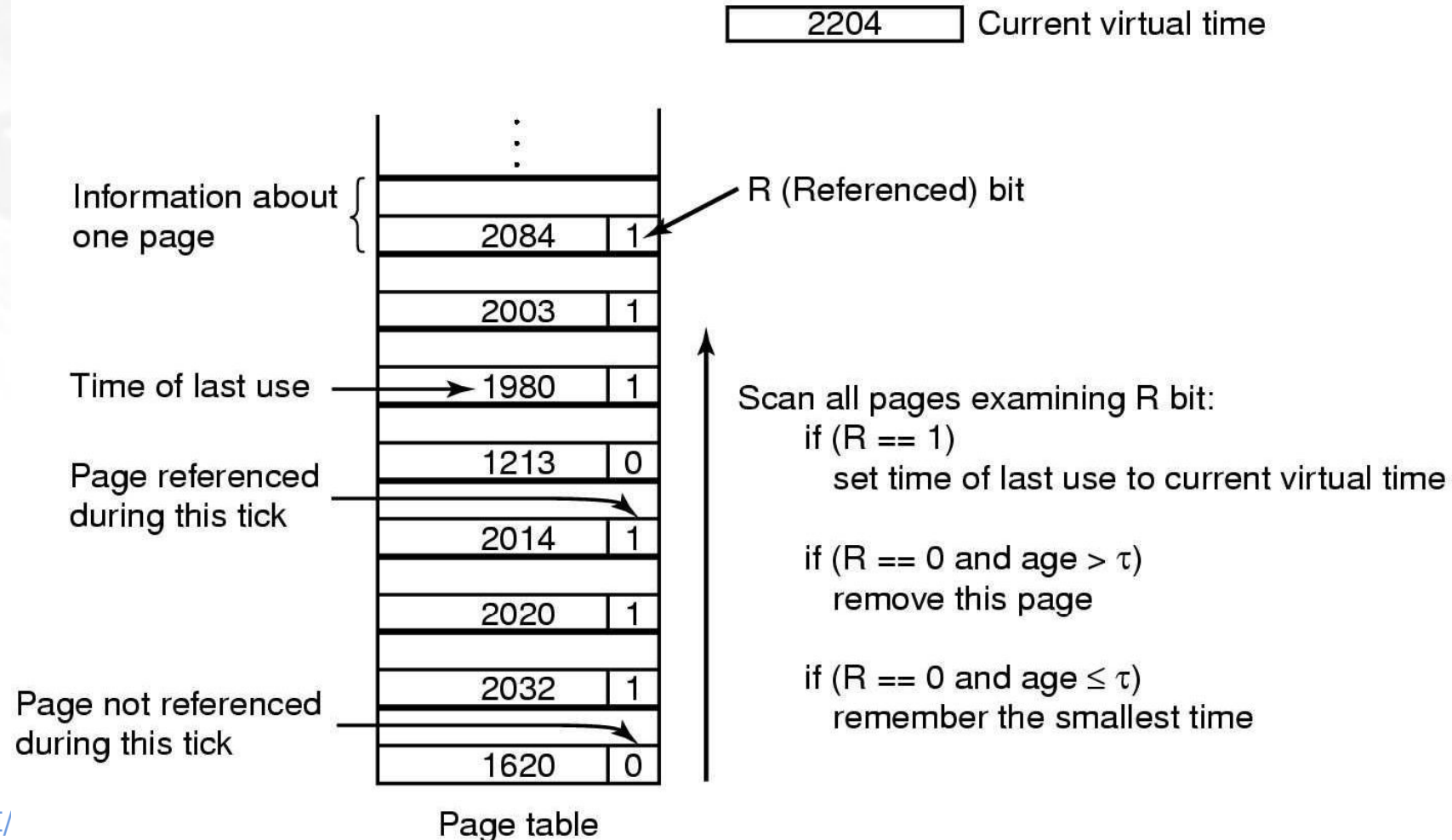
"Políticas de Busca de Páginas de um Processo" (voltando ao slide 2)

- Nos sistemas time-sharing processos estão constantemente bloqueados, podendo ser "jogados para disco"
- *Swapping*
  - Técnica para resolver problema de processos que aguardam por espaço livre adequado;
  - Processos não ficam mais na memória o tempo todo (são então suspensos).
  - Um processo residente na memória é levado para o disco (Swapped-Out), dando lugar a outro;
  - O processo Swapped-Out retorna à memória (Swapped-In), sem "perceber" o que ocorreu.
- Se paginação **por demanda**, 20, 50, 100... Page faults cada vez que o processo é re-carregado na MP
  - Processo fica lento, perda de tempo de CPU

Working Set (7)

- **Pré-paginação**
  - Carregar em memória as páginas do Working set do processo antes que o mesmo possa continuar sua execução
  - Garantimos que ocorrerá menos Page faults quando o processo for executado
- Como monitorar o Working set do processo de modo que ele esteja sempre atualizado?
  - Se a página não foi referenciada nos *n clock ticks* consecutivos , sai do Working set

Um algoritmo de substituição de páginas baseado no Working Set ... uma ideia de implementação:



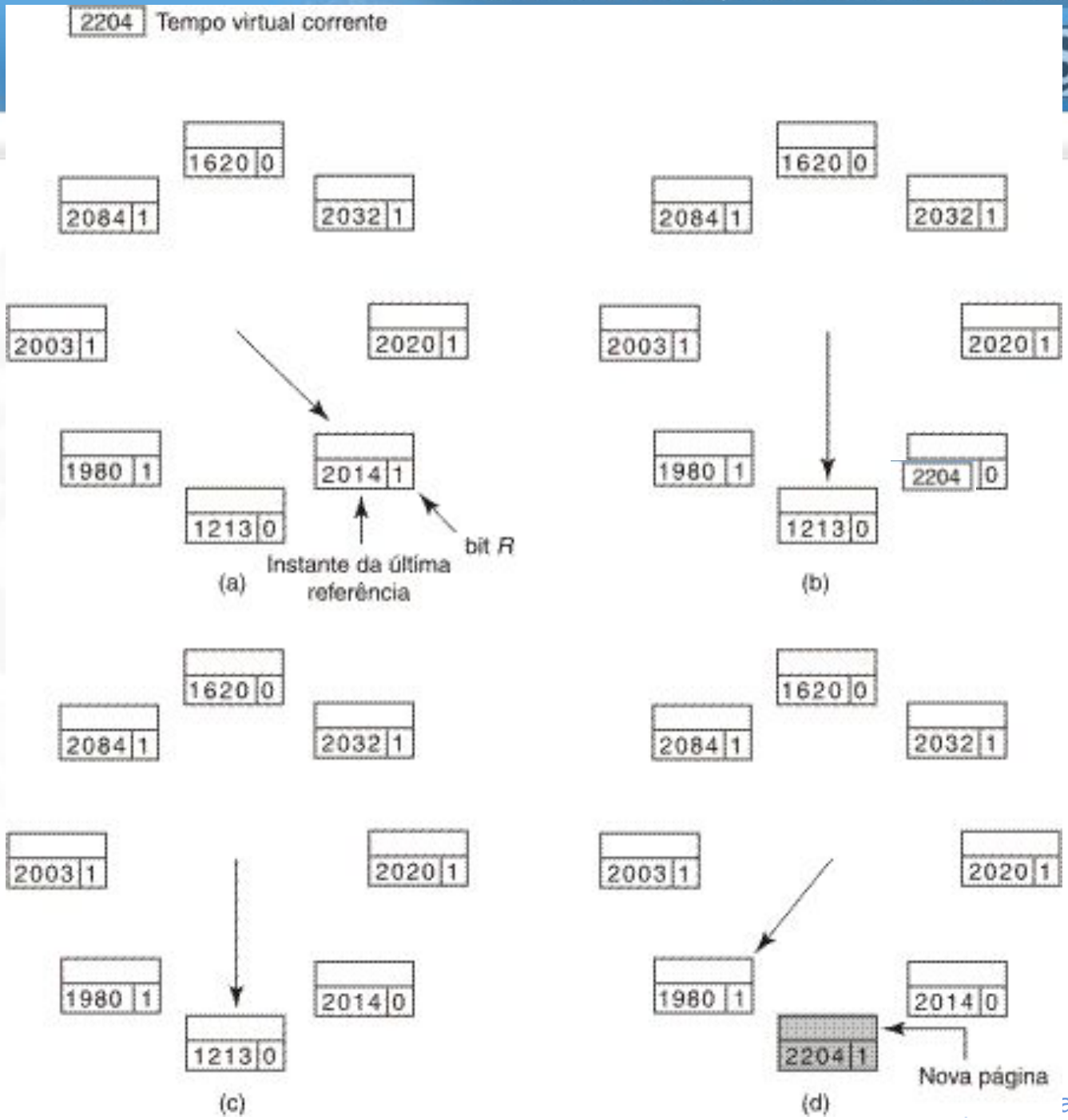
Algoritmo *WSClock* ... uma implementação mais eficiente...

- No *WSClock (Working Set Clock*), na troca de páginas só são avaliadas as páginas presentes em uma lista circular
- Cada entrada dessa lista possui os bits R e M, além de um *timestamp* (tempo da última referência)
- À medida que as páginas são carregadas em memória, elas são inseridas na lista circular
- O algoritmo é executado quando precisa-se "liberar" molduras
  - Isto pode ocorrer quando ocorrem page faults no processo em questão ou mesmo em outros processos (política Global de alocação de molduras)
  - Troca-se a primeira página (eventualmente, libera-se mais páginas) a partir da posição do ponteiro na lista que tenha R=0 e cuja idade supera T
    - Na verdade, verifica-se se a pag. está limpa (i.e. se ela ñ foi modificada). Caso ela tenha M=1, é escalonada uma escrita dessa pag. no disco e pula-se p/ a próxima página da lista circular.

Algoritmo *WSClock*

2204

2204



Resumo dos Algoritmos

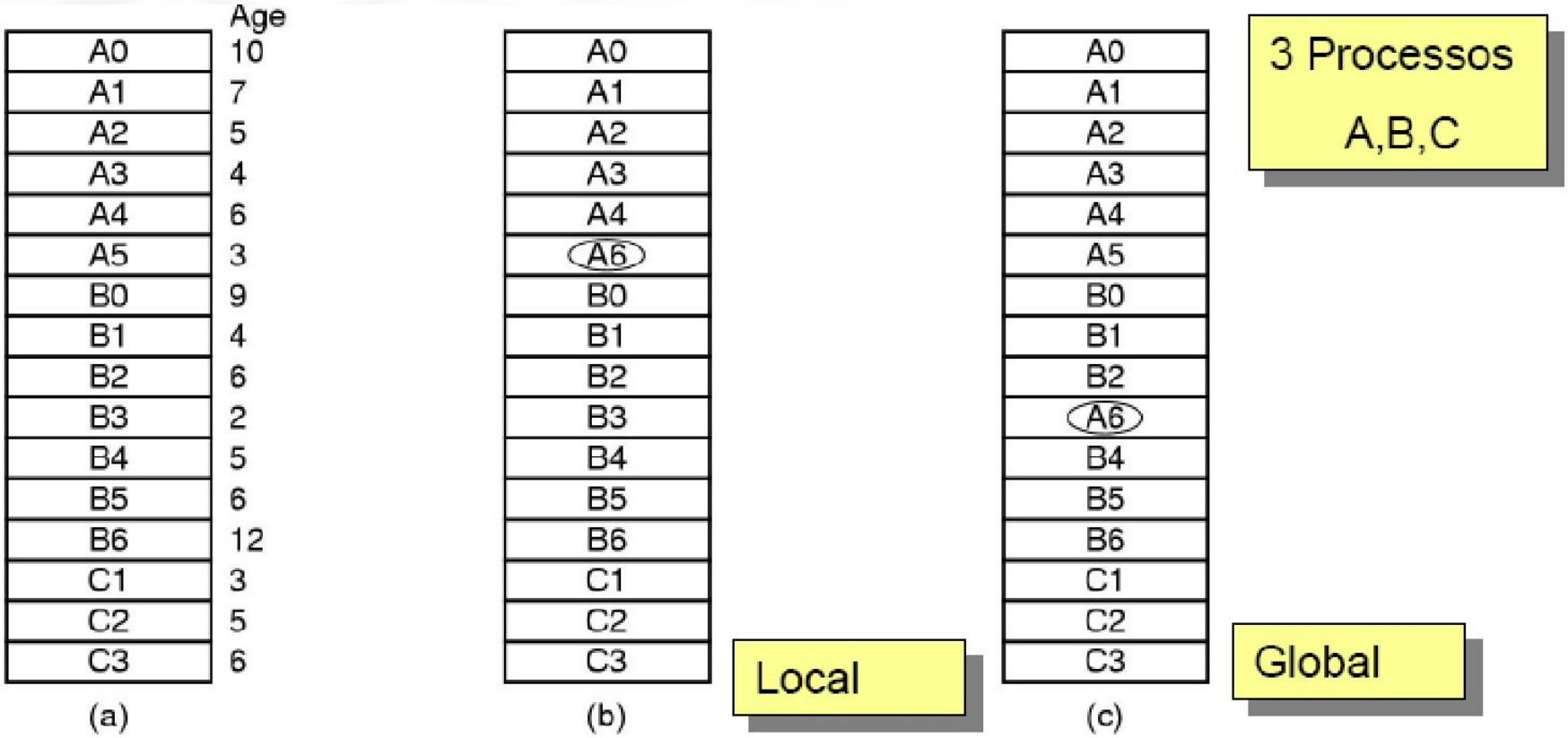
Ótimo	Não é possivel(referência)
NRU	Fácil de implementar; Pouco eficiente
FIFO	Pode retirar páginas importantes
Segunda Chance	Melhorias ao FIFO
Relógio	Implementação eficiente do SC; Realista
LRU	Excelente, dificil de implementar (HW)
NFU	Fraca aproximação do LRU
Aging	Eficiente que se aproxima do LRU
Working Set	Dificil de implementar
WSClock	Boa eficiência

Mais Considerações no Projeto de Sistemas de Paginação

- Política de alocação: Local x Global
- Anomalia de Belady
- Controle de Carga
- Tamanho da página
- Espaços de Instruções e Dados Separados
- Páginas compartilhadas

Política de alocação: Local x Global (1)

- O LRU deve considerar as páginas apenas do processo que gerou o Page Fault, ou de todos os processos?



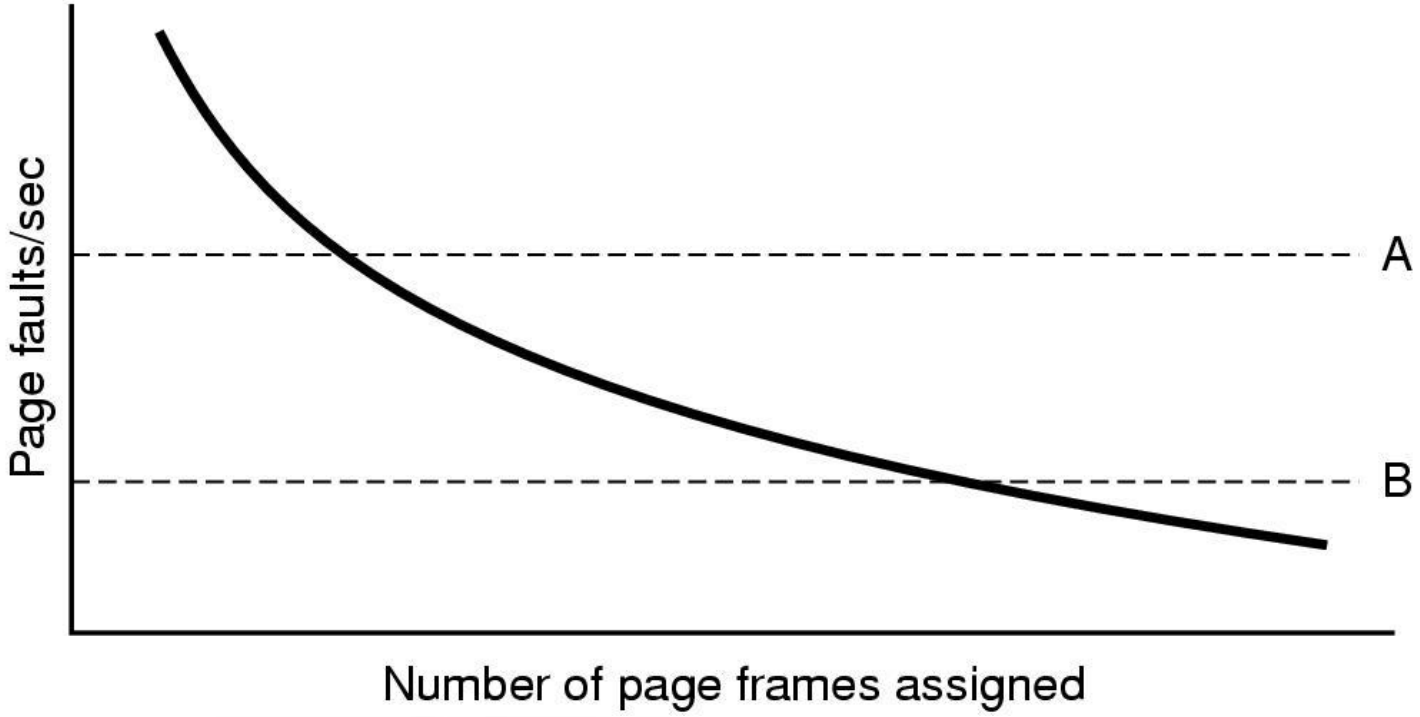
Política de alocação: Local x Global (2)

- Política LOCAL (temos uma "Alocação **fixa** de qde. de fames")
  - Alocam uma fração **fixa** de memória para cada processo
- Política GLOBAL (temos uma "Alocação **variável** de qde. de fames")
  - Alocam molduras de páginas entre os processos em execução
    - O número de molduras alocadas para cada processo varia no tempo
- Working set varia durante a execução de um processo
  - Quando a política é local
    - Há trashing quando o tamanho do WS aumenta
    - Há desperdício quando o tamanho do WS diminui
  - Algoritmos com política global são mais adequados
    - Usa-se os bits de "tempo da ultima referencia" para monitorar o Working Set
    - Não necessariamente evita o trashing -> o Working set pode variar de tamanho em questão de microssegundos (os bits de aging são alterados a cada interrupção de relógio)

Política de alocação: Local x Global (3)

- Outra abordagem determinar periodicamente o número de processos e dividir as molduras entre os mesmo
  - 12.416 molduras ; 10 processos => 1.241 molduras / processo
  - É justo? E se processos têm tamanho diferentes?
- Solução:
  - Alocar para cada processo um número mínimo de páginas proporcional ao tamanho do processo
  - Atualizar a alocação dinamicamente
- Algoritmo de alocação ***Page Fault Frequency*** **(PFF)**
  - Informa quando aumentar ou diminuir a alocação de molduras para um processo
  - Tenta manter a taxa de Page Fault dentro de um intervalo aceitável
  - Usa-se em combinação com algum algoritmo de substituição de página

Política de alocação: Local x Global (4)



- Se maior do que A, taxa muito alta
  - Deve-se alocar mais molduras
- Se menor do que B, taxa muito baixa
  - Algumas molduras podem ser eliminadas

Anomalia de Belady (1)

- Intuitivamente, quanto maior o número de molduras, menor será o número de *Page Faults*
  - Nem sempre isso será verdadeiro!
- Belady et al. descobriram um contra-exemplo para o algoritmo FIFO
  - Suponha que as páginas sejam referenciadas nesta ordem:

    0 1 2 3 0 1 4 0 1 2 3 4
  - Qual será o número de Page Faults em um FIFO alocando 3 molduras para o processo? E 4 molduras?
- **Belady et al definiram um modelo ("Modelo de Pilha") para mostrar se algoritmos apresentam a anomalia**

Anomalia de Belady - Rodando o FIFO

	0	1	2	3	0	1	4	0	1	2	3	4
Página mais nova	0	1	2	3	0	1	4	4	4	2	3	3
		0	1	2	3	0	1	1	1	4	2	2
Página mais velha			0	1	2	3	0	0	0	1	4	4
	P	P	P	P	P	P	P			P	P	

9 Page Faults

	0	1	2	3	0	1	4	0	1	2	3	4
Página mais nova	0	1	2	3	3	3	4	0	1	2	3	4
		0	1	2	2	2	3	4	0	1	2	3
			0	1	1	1	2	3	4	0	1	2
Página mais velha				0	0	0	1	2	3	4	0	1
	P	P	P	P			P	P	P	P	P	P

10 Page Faults

Controle de Carga

- Mesmo com paginação, swaping é ainda necessário
- Determina o número de processos residentes em MP (escalonador de médio prazo)
  - Poucos processos, possibilidade de processador vazio;
  - Muitos processos, possibilidade de *trashing*
- Regra dos 50% de utilização do dispositivo de paginação (acionado por *Page fault*)
- Swapping é usado para reduzir demanda potencial por memória, em vez de reivindicar blocos para uso imediato

Tamanho de Páginas (1)

- Página de pequeno tamanho
  - tempo curto para transferência de página entre disco e memória
  - muitas páginas de diferentes programas podem estar residentes em memória
  - menor fragmentação interna
  - exige tabelas de páginas muito grandes, que ocupam espaço em memória
  - mais adequada para instruções
- Página de grande tamanho
  - Tabelas de páginas pequenas
  - Transferência de 64 páginas de 512 B pode ser mais lenta do que a transferência de 4 páginas de 8KB
  - Tempo longo para transferência de uma página entre disco e memória
  - Mais adequada para dados (gráficos exigem páginas muito grandes)

Tamanho de Páginas (2)

- Custo adicional devido à paginação
  - s : tamanho médio dos processos
  - p: tamanho da página em bytes
  - e: tamanho de cada entrada da tabela de páginas

$\frac{s}{p} . e$ -> tamanho aproximado da tabela de páginas

$\frac{p}{2}$ -> memória desperdiçada na última página do processo

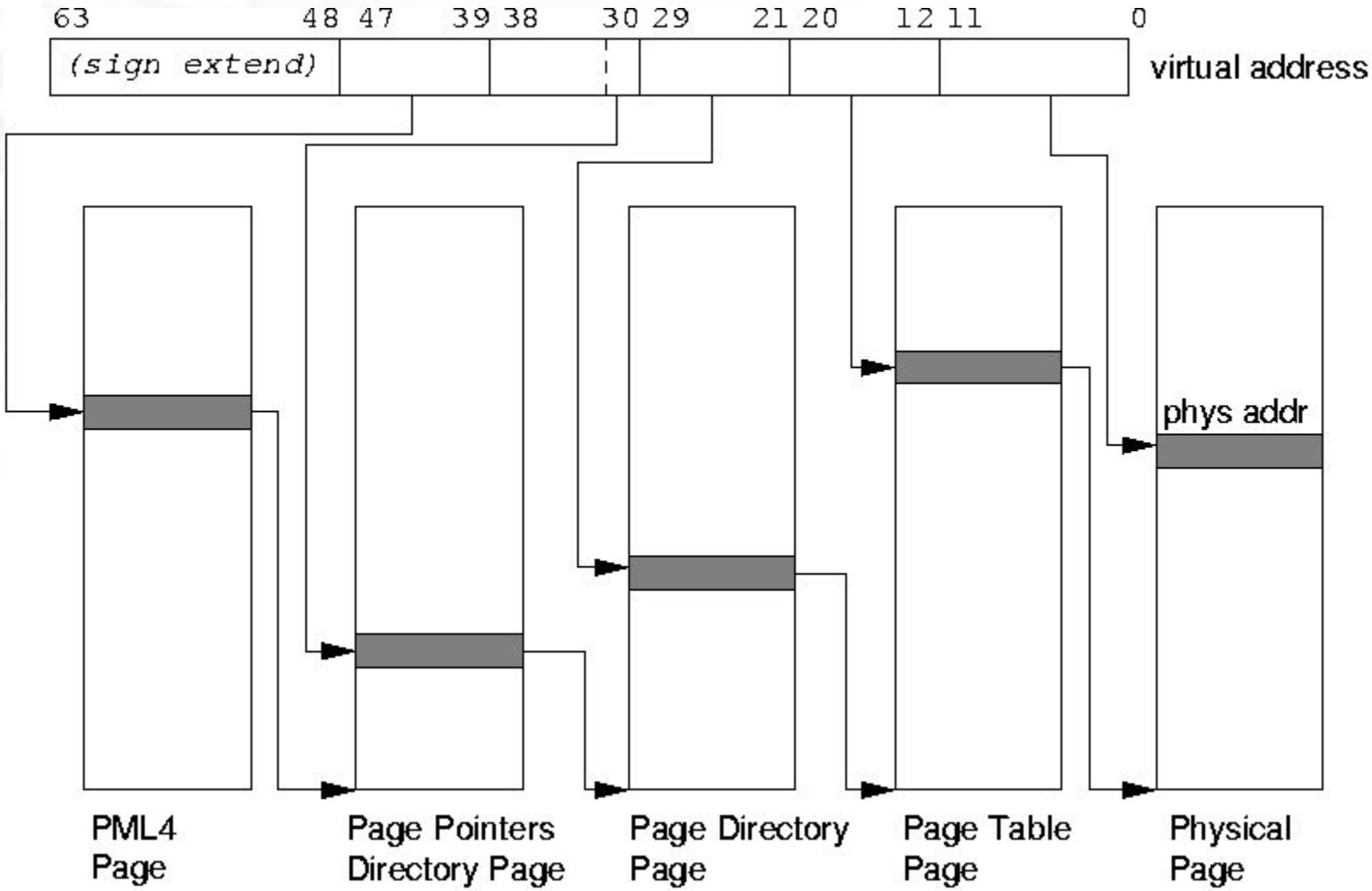
$$\text{custo adicional} = \frac{s.e}{p} + \frac{p}{2}$$

Derivando em relação a **p**: o tamanho ótimo será: $p=\sqrt{2.s.e}$

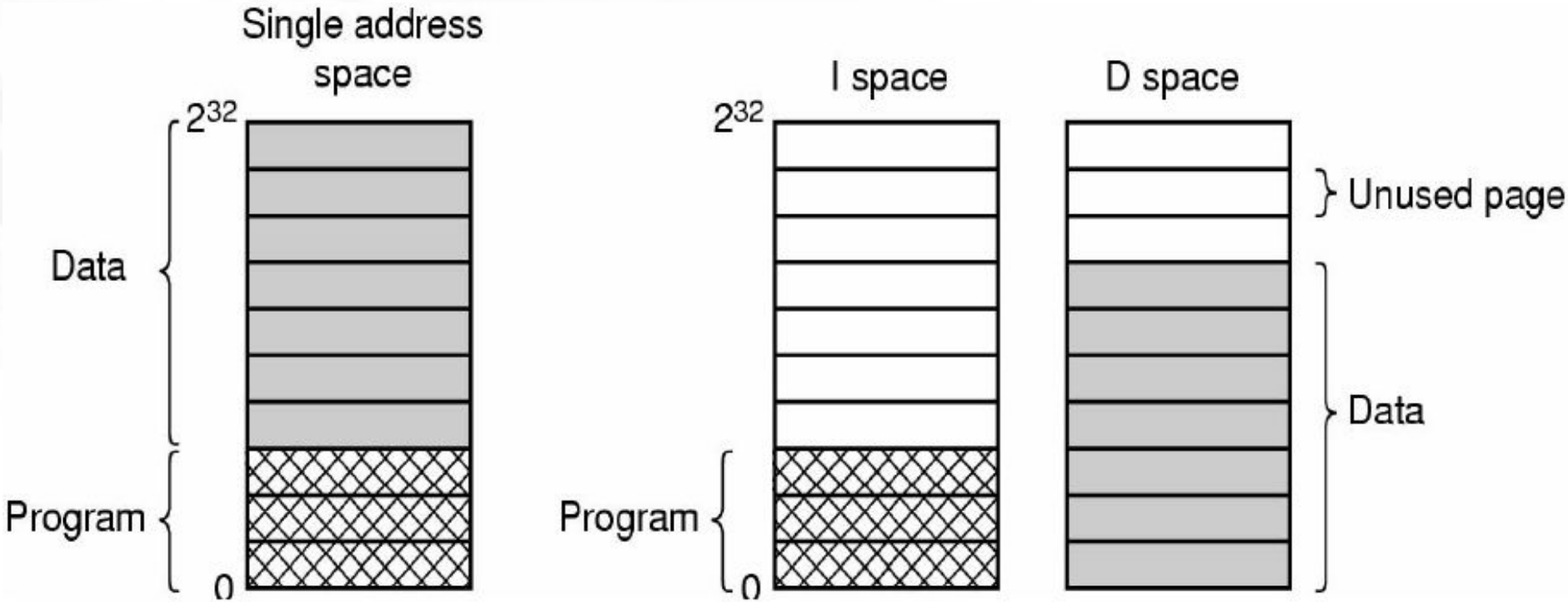
Tamanho de Páginas (3)

- Solução de compromisso: permitir páginas de tamanhos diversos para código e dados
- Tamanhos de páginas variam muito, de 64 bytes a 4 Mbytes
  - Pentium (... x86) permite selecionar página de 4 K ou 4 Mbytes
  - Motorola MC88200
    - páginas de 4 Kbytes para programas de usuário
    - páginas de 512 Kbytes para programas do sistema, que devem residir sempre em memória
- **Máquinas de 64 bits: páginas de 4kb, 2Mb and 1Gb**
- **ARM: 4kb, 64kb, e 1Mb**

Tamanho de Páginas (4)



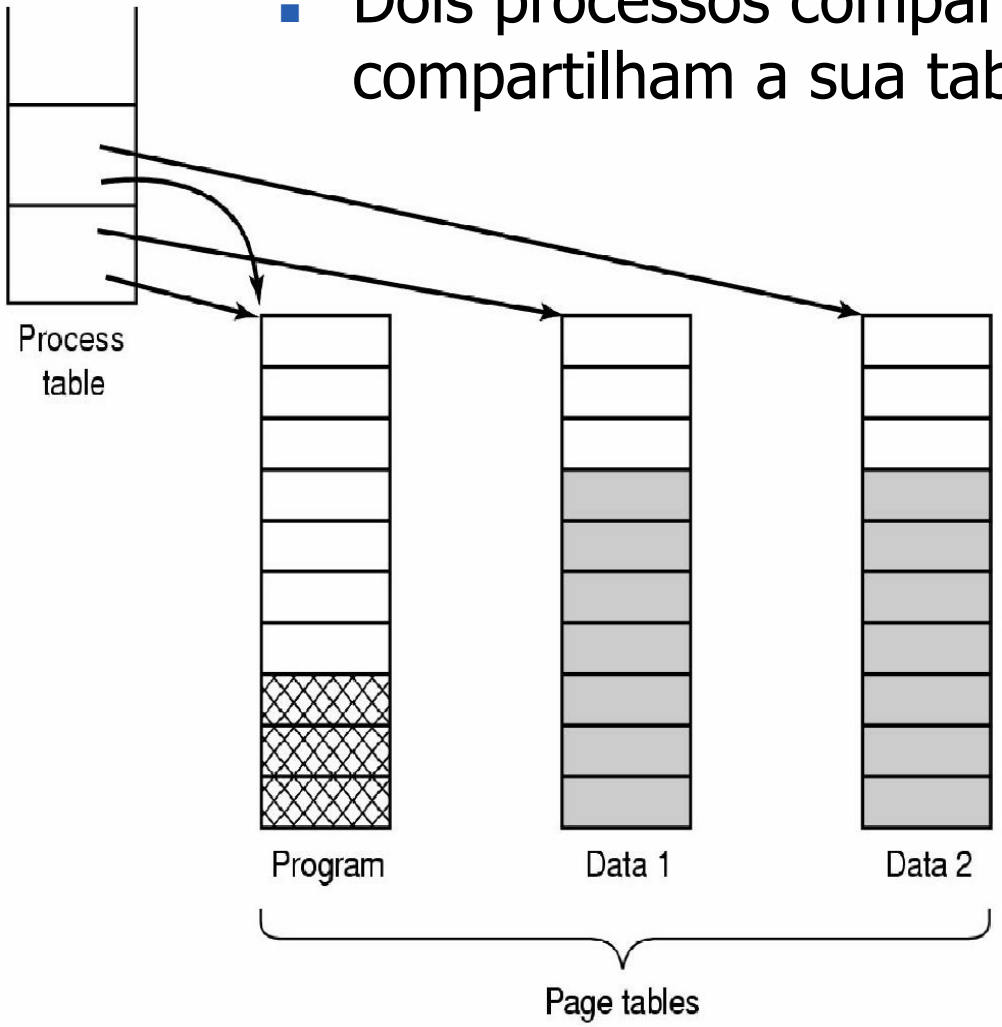
Espaços de Instruções e Dados Separados



- Duplica o espaço de endereçamento disponível
- Uma tabela de páginas para cada espaço de endereçamento

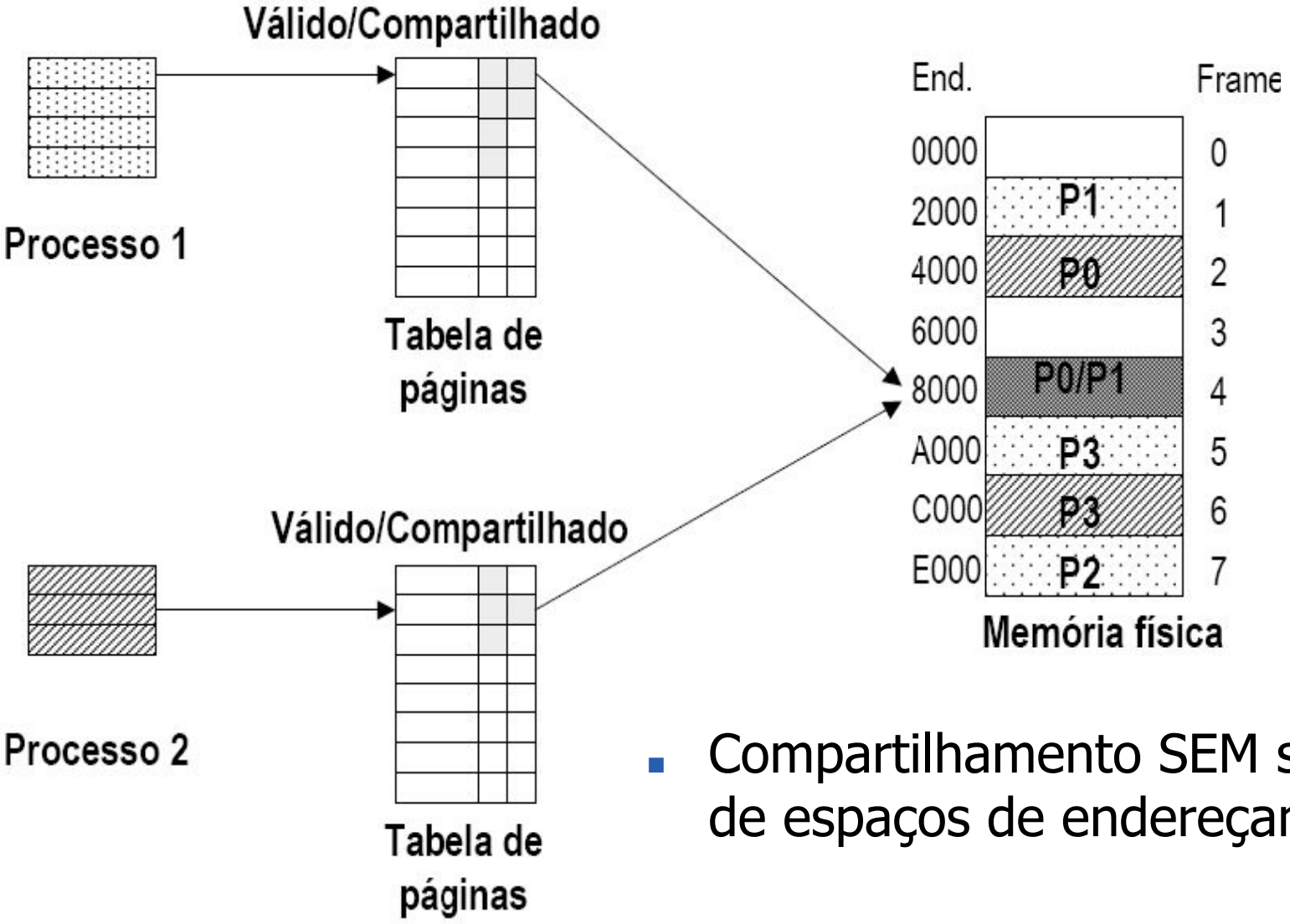
Páginas Compartilhadas (1)

- Dois processos compartilhando o mesmo programa, compartilham a sua tabela de páginas



- Mesmo não havendo dois espaços de endereçamento (Código e Dado) é possível compartilhar páginas, mas o mecanismo não é tão direto
- Usando tabelas invertidas o mecanismo é mais complicado ainda...

Páginas Compartilhadas (2)



- Compartilhamento SEM separação de espaços de endereçamento

Páginas Compartilhadas (3)

- **Código Reentrante**
  - Código que não modifica a si próprio, ou seja, ele nunca é modificado durante a execução
  - Dois ou mais processos podem executar o mesmo código "simultaneamente"
  - Exemplo:
    - Editor de texto com código reentrante de **150 K** e área de dados de **50 K**
    - **40** usuários utilizando o editor em um ambiente de tempo compartilhado, seriam necessários **200 K x 40 = 8000 K**
    - Se o código executável for compartilhado, serão consumidos apenas **(50 K x 40) + 150 K = 2150 K**