

**UFPR - DInf - Bacharelado em Ciência da Computação
CI212 - Organização e Arquitetura de Computadores**

Primeira Prova de 2002-2 (25nov02)

- 1) A instrução JM (jump-memory) é similar a uma instrução LW, mas o valor copiado da memória é gravado no PC e não no registrador rt. Mostre no diagrama em anexo como implementar esta instrução. Acrescente todas as ligações e componentes que julgar necessário. Desenhe o ramo da máquina de estados correspondente a esta instrução e indique os eventos relevantes em cada estado. Mostre a codificação da instrução JM. [10p]
- 2) Um número de ponto flutuante de precisão simples (float) está armazenado no endereço E. (a) Mostre como este número pode ser multiplicado por 2 SEM_USAR instruções de ponto flutuante. (b) Escreva um programa em assembly do MIPS para efetuar esta multiplicação. [10 pontos]
- 3) Defina e descreva os cinco modos de endereçamento do MIPS. [10 pontos]

Segunda Prova de 2002-2 (07fev03)

- 1) Considere uma hierarquia de memória com dois níveis de cache (L1 e L2) e memória principal. Suponha que em 1000 referências à memória ocorrem 40 faltas na L1 e 20 faltas na L2. [a] Qual a taxa de faltas na cache primária? [b] Quais as DUAS taxas de faltas na cache secundária? [5pts]
- 2) Projete uma cache com capacidade para 16 palavras que, simultaneamente [1] produza a maior taxa de acertos para a seguinte seqüência de referências (endereços de palavras), e [2] o melhor desempenho global. Considere que o barramento tem largura de duas palavras. Referências marcadas com * são escritas. [10 pontos]

12 13* 20 15 16* 14* 13 22 24 26 12* 14 8 4 2 7 5 13* 15* 2* 3 27
- 3) Os dois trechos de programa mostrados abaixo adicionam uma constante aos elementos de um vetor (com número par de elementos). O loop2 foi "desenrolado" para permitir execução mais rápida no processador SEM adiantamento. Reordene as instruções do loop2 para obter o melhor desempenho possível, e compare o tempo de execução (número de ciclos) das duas versões. [10 pontos]

```

loop1:                                loop2:
    lw   t0, 0(s1)                    lw   t0, 0(s1)
    addu t0, t0, s2                    addu t0, t0, s2
    sw   t0, 0(s1)                    sw   t0, 0(s1)
    addi s1, s1, -4                    lw   t1, -4(s1)
    bne  s1, $0, loop1                addu t1, t1, s2
                                        sw   t1, -4(s1)
                                        addi s1, s1, -8
                                        bne  s1, $0, loop1

```

- 4) Re-calcule os tempos de execução dos dois programas acima, considerando processador COM adiantamento. Indique CLARAMENTE quais são os circuitos de adiantamento que você considerou. [5 pontos]

Terceira Prova de 2002-2 (17mar03)

- 1) Tabelas de páginas (TP) ocupam grandes áreas de memória, mesmo que a maioria de seus elementos sejam inválidos. Uma solução para reduzir o espaço ocupado pela TP é usar uma hierarquia de TPs. O número da página virtual pode ser quebrado em dois campos, um "número da tabela de páginas" e um "deslocamento na tabela de páginas". O número da tabela de páginas pode ser usado para indexar a TP de primeiro nível (TP-1n), que contém os endereços físicos para TPs de segundo nível (TP-2n), supondo que estas (TP-2n) residem em memória (se não estiver em memória, ocorrerá uma falta de página e a TP-2n deverá ser copiada do disco para a memória). O deslocamento na tabela de páginas é usado para indexar a TP-2n para então obter o número da página física. Um esquema óbvio para isso é fazer a com que as TP-2n ocupem exatamente uma página. Supondo endereços virtuais de 32 bits e páginas de 4Kbytes, e 4 bytes em cada elemento da TP, quantos bytes cada programa necessita para armazenar a TP-1n (que deve estar sempre em memória)? Desenhe um diagrama detalhado que demonstre sua compreensão desta idéia. [20 pontos]

- 2) Considere um barramento com as seguintes características:
- barramento e memória suportam transferências em blocos com tamanho de 4 a 16 palavras de 32 bits;
 - o barramento é síncrono, com 64 bits de largura e com clock de 200MHz (5ns). Cada transferência de 64 bits leva um ciclo, e um ciclo é necessário para transferir um endereço;
 - São necessários dois ciclos entre cada transação (arbitragem);
 - o tempo de acesso à memória para as quatro primeiras palavras é de 200ns; cada conjunto adicional de quatro palavras pode ser acessado em 20ns. Considere que o barramento consegue sobrepor a transferência das palavras lidas mais recentemente com a leitura das quatro próximas palavras.

Calcule a vazão sustentada e a latência para a leitura de 256 palavras para transferências que usam blocos de 4 palavras e para blocos de 16 palavras.

Calcule também o número efetivo (sustentado) de transações por segundo em cada um dos dois casos acima. Uma "transação de barramento" consiste da emissão de um endereço seguida da transmissão dos dados. [15 pontos]

- 3) Para cada um dos parâmetros de projeto listados abaixo, descreva sua influência no desempenho de um computador. [5 pontos]
- associatividade em cache ou TLB;
 - arquitetura Harvard (instruções e dados em circuitos diferentes);
 - segmentação da interface de memória (pipelineing);
 - super-escalaridade;
 - transações multi-fase no barramento de memória (split transactions).

Exame Final de 2002-2 (31mar03)

- 1) Projete uma cache com associatividade maior que 1, com capacidade para 16 palavras que, simultaneamente [1] produza a maior taxa de acertos para a seguinte seqüência de referências (endereços de palavras), e [2] o melhor desempenho global. Considere que o barramento tem largura de duas palavras. Referências marcadas com * são escritas. Explique o porquê de sua escolha. [30 pontos]

12 31* 20 15 16* 14* 13 22 24 26 21* 14 8 4 2 7 5 31* 15* 29* 13 28

- 2) Considere o processador SEM PIPELINE (multiciclo) do capítulo 5. Um programa tem a seguinte mistura de instruções:
alu 40%, lw 25%, sw 15%, beq/bne 20%
- Suponha que existem três implementações do processador:
- o processador visto em aula (abaixo), com clock de 500MHz (2ns);
 - um processador similar ao do item (a), exceto que a atualização dos registradores ocorre no mesmo tic de relógio em que, uma leitura de memória, ou uma operação de ULA completam. Esta versão tem clock de 400MHz (2,5ns) porque a atualização dos registradores alonga o ciclo do relógio.
 - um processador similar ao do item (b), exceto que o cálculo do endereço efetivo ocorre no mesmo ciclo em que o acesso à memória. Esta versão possui clock de 250MHz (4ns) por causa do longo intervalo necessário para o cálculo do endereço e o acesso à memória.
- Qual das três é a mais rápida? Existem outras misturas de instruções que fariam outra máquina mais rápida, se sim, qual (ou quais)? [40 pontos]
- 3) Considere três máquinas com diferentes organizações de caches:
- Cache 1 - mapeamento direto com blocos de uma palavra;
 - cache 2 - mapeamento direto com blocos de quatro palavras;
 - cache 3 - associatividade binária, com blocos de quatro palavras.
- As seguintes medidas de taxas de faltas foram observadas:
- Cache 1 - taxa de faltas de instruções de 4% e de dados de 8%;
 - cache 2 - taxa de faltas de instruções de 2% e de dados de 5%;
 - cache 3 - taxa de faltas de instruções de 2% e de dados de 4%;
- Para estas máquinas, metade das instruções contém uma referência a dados. Suponha que a penalidade por falta é de $6 + |\text{tamanho do bloco}|$. O CPI para estes programas foi medido na máquina com a cache-1 e é de 2,0. Determine qual das máquinas dispense mais ciclos em faltas. [30pts]