

a bibliografia dos primeiros relatórios conclui este livro interessante.

Public Broadcasting System [1992]. *A máquina que mudou o mundo*, fitas de vídeo.

Esses cinco programas de 1 hora incluem filmagens raras e entrevistas com pioneiros da indústria de computadores.

Slater, R. [1987]. *Portraits in Silicon*, MIT Press, Cambridge, MA.

Biografias curtas de 31 pioneiros da computação.

Stern, N. [1980]. "Quem inventou o primeiro computador digital eletrônico?" *Annals of the History of Computing* 2: 4 (outubro), 375–76 *A perspectiva de um historiador sobre Atanasoff versus Eckert e Mauchly.*

Weicker, RP [1984]. "Dhrystone: um benchmark de programação de sistemas sintéticos", *Communications of the ACM* 27 (10): 1013–1030.

Descrição de um programa de benchmarking sintético para código de sistemas.

Wilkes, MV [1985]. *Memoirs of a Computer Pioneer*, MIT Press, Cambridge, MA.

Uma visão pessoal da computação por um dos pioneiros.

1.13 Exercícios

As classificações de tempo relativas dos exercícios são mostradas entre colchetes após cada número de exercício. Em média, um exercício classificado como [10] levará o dobro da duração de um exercício classificado [5]. As seções do texto que devem ser lidas antes de tentar um exercício serão fornecidas entre colchetes; por exemplo, <§1.4> significa que você deve ter lido a [Seção 1.4](#), Sob as Capas, para ajudá-lo a resolver este exercício.

1.1 [2] <§1.1> Além dos telefones celulares inteligentes usados por um bilhão de pessoas, liste e descreva quatro outros tipos de computadores.

1.2 [5] <§1.2> As oito grandes idéias em arquitetura de computador são semelhantes às idéias de outros campos. Combine as oito ideias da arquitetura de computador, "Design for Moore's Law," "Use Abstraction to Simplify Design," "Make the Common Case Fast",

"Performance via Parallelism," "Performance via Pipelining,"

"Performance via Prediction," "Hierarquia de memórias" e "Confiabilidade via redundância" para as seguintes ideias de outros campos:

- a. Linhas de montagem na fabricação de automóveis
- b. Cabos da ponte suspensa
- c. Sistemas de navegação aérea e marítima que incorporam informações sobre o vento
- d. Elevadores expressos em edifícios
- e. Mesa de reserva da biblioteca
- f. Aumentar a área da porta em um transistor CMOS para diminuir seu tempo de chaveamento

- g. Adicionar catapultas eletromagnéticas de aeronaves (que são movidas eletricamente em oposição aos modelos atuais movidos a vapor), permitido pelo aumento da geração de energia oferecida pela nova tecnologia de reator
- h. Construir carros autônomos cujos sistemas de controle dependem parcialmente de sistemas de sensores existentes já instalados no veículo de base, como sistemas de saída de faixa e sistemas de controle de cruzeiro inteligentes

1.3 [2] <§1.3> Descreva as etapas que transformam um programa escrito em uma linguagem de alto nível, como C, em uma representação que é executada diretamente por um processador de computador.

1.4 [2] <§1.4> Suponha uma exibição em cores usando 8 bits para cada uma das cores primárias (vermelho, verde, azul) por pixel e um tamanho de quadro de 1280 × 1024.

- uma. Qual é o tamanho mínimo em bytes do buffer de quadros para armazenar um quadro?
- b. Quanto tempo levaria, no mínimo, para o quadro ser enviado por uma rede de 100 Mbit / s?

1.5 [4] <§1.6> Considere três processadores diferentes P1, P2 e P3 executando o mesmo conjunto de instruções. P1 tem uma taxa de clock de 3 GHz e um CPI de 1,5. P2 tem uma taxa de clock de 2,5 GHz e um CPI de 1,0. P3 tem uma taxa de clock de 4,0 GHz e um CPI de 2,2.

- uma. Qual processador tem o maior desempenho expresso em instruções por segundo?
- b. Se os processadores cada um executar um programa em 10 segundos, encontre o número de ciclos e o número de instruções.
- c. Estamos tentando reduzir o tempo de execução em 30%, mas isso leva a um aumento de 20% no IPC. Que taxa de clock devemos ter para obter essa redução de tempo?

1.6 [20] <§1.6> Considere duas implementações diferentes da mesma arquitetura de conjunto de instruções. As instruções podem ser divididas em quatro classes de acordo com seu CPI (classes A, B, C e D). P1 com uma taxa de clock de 2,5 GHz e CPIs de 1, 2, 3 e 3, e P2 com uma taxa de clock de 3 GHz e CPIs de 2, 2, 2 e 2.

Dado um programa com uma contagem de instrução dinâmica de Instruções 1.0E6 divididas em classes da seguinte forma: 10% classe A, 20% classe B, 50% classe C e 20% classe D, o que é mais rápido: P1 ou P2?

- uma. Qual é o CPI global para cada implementação?
- b. Encontre os ciclos de clock necessários em ambos os casos.

1.7 [15] <§1.6> Compiladores podem ter um impacto profundo no desempenho de um aplicativo. Suponha que, para um programa, o compilador A resulte em uma

contagem de instrução dinâmica de $1,0E9$ e tenha um tempo de execução de $1,1$ s, enquanto o compilador B resulte em uma contagem de instrução dinâmica de $1,2E9$ e um tempo de execução de $1,5$ s.

uma. Encontre o CPI médio para cada programa, dado que o processador tem um tempo de ciclo de clock de 1 ns.

b. Suponha que os programas compilados sejam executados em dois processadores diferentes. Se os tempos de execução nos dois processadores são iguais, quão mais rápido é o clock do processador que executa o código do compilador A em comparação com o clock do processador que executa o código do compilador B?

c. Foi desenvolvido um novo compilador que usa apenas $6.0E8$ instruções e tem um CPI médio de $1,1$. Qual é a aceleração do uso desse novo compilador em comparação com o uso do compilador A ou B no processador original?

1.8 O processador Pentium 4 Prescott, lançado em 2004, tinha clock de $3,6$ GHz e tensão de $1,25$ V. Suponha que, em média, ele consumisse 10 W de potência estática e 90 W de potência dinâmica. O Core i5 Ivy Bridge, lançado em 2012, tem clock de $3,4$ GHz e tensão de $0,9$ V. Suponha que, em média, ele consumiu 30 W de potência estática e 40 W de potência dinâmica.

1.8.1 [5] <§1.7> Para cada processador, encontre as cargas capacitivas médias.

1.8.2 [5] <§1.7> Encontre a porcentagem da potência dissipada total composta pela potência estática e a razão entre a potência estática e a potência dinâmica para cada tecnologia.

1.8.3 [15] <§1.7> Se a potência dissipada total deve ser reduzida em 10% , quanto a tensão deve ser reduzida para manter a mesma corrente de fuga? Nota: a potência é definida como o produto da tensão e da corrente.

1.9 Suponha que para as instruções aritméticas, carregar / armazenar e ramificar, um processador tem CPIs de 1 , 12 e 5 , respectivamente. Suponha também que em um único processador um programa requer a execução de instruções aritméticas $2.56E9$, instruções carregar / armazenar $1.28E9$ e 256 milhões de instruções de desvio. Suponha que cada processador tenha uma frequência de clock de 2 GHz.

Suponha que, como o programa é paralelizado para rodar em vários núcleos, o número de instruções aritméticas e carregar / armazenar por processador é dividido por $0,7 \times p$ (onde p é o número de processadores), mas o número de instruções de ramificação por processador permanece o mesmo.

1.9.1 [5] <§1.7> Encontre o tempo total de execução para este programa em 1 , 2 , 4 e 8 processadores e mostre a aceleração relativa do resultado de 2 , 4 e 8 processadores em relação ao resultado de processador único.

1.9.2 [10] <§§1.6, 1.8> Se o CPI das instruções aritméticas fosse duplicado, qual seria o impacto no tempo de execução do programa em 1 , 2 , 4 ou 8 processadores?

1.9.3 [10] <§§1.6, 1.8> A que o CPI das instruções de carregamento / armazenamento deve ser reduzido para que um único processador corresponda ao desempenho de quatro processadores usando os valores originais do CPI?

1.10 Suponha que um wafer de 15 cm de diâmetro tenha um custo de 12, contenha 84 matrizes e 0,020 defeitos / cm². Suponha que um wafer de 20 cm de diâmetro tenha um custo de 15, contenha 100 matrizes e tenha 0,031 defeitos / cm².

1,10.1 [10] <§1.5> Encontre o rendimento para ambas as bolachas.

1,10.2 [5] <§1.5> Encontre o custo por dado para ambas as bolachas.

1.10.3 [5] <§1.5> Se o número de matrizes por wafer for aumentado em 10% e os defeitos por unidade de área aumentarem em 15%, encontre a área da matriz e o rendimento.

1,10.4 [5] <§1.5> Suponha que um processo de fabricação melhora o rendimento de 0,92 para 0,95. Encontre os defeitos por unidade de área para cada versão da tecnologia dada uma área de matriz de 200 mm².

1.11 Os resultados do benchmark SPEC CPU2006 bzip2 rodando em um AMD Barcelona tem uma contagem de instruções de 2.389E12, um tempo de execução de 750 se um tempo de referência de 9650 s.

1.11.1 [5] <§§1.6, 1.9> Encontre o IPC se o tempo de ciclo do relógio for 0,333 ns.

1.11.2 [5] <§1.9> Encontre a SPECratio.

1.11.3 [5] <§§1.6, 1.9> Encontre o aumento no tempo de CPU se o número de instruções do benchmark for aumentado em 10% sem afetar o CPI.

1.11.4 [5] <§§1.6, 1.9> Encontre o aumento no tempo de CPU se o número de instruções do benchmark aumentar em 10% e o CPI em 5%.

1,11.5 [5] <§§1.6, 1.9> Encontre a alteração na SPECratio para esta alteração.

1.11.6 [10] <§1.6> Suponha que estejamos desenvolvendo uma nova versão do processador AMD Barcelona com uma taxa de clock de 4 GHz. Adicionamos algumas instruções adicionais ao conjunto de instruções de forma que o número de instruções foi reduzido em 15%. O tempo de execução é reduzido para 700 se a nova SPECratio é 13.7. Encontre o novo CPI.

1,11.7 [10] <§1.6> Este valor de CPI é maior do que o obtido em 1.11.1, pois a taxa de clock foi aumentada de 3 GHz para 4 GHz. Determine se o aumento no IPC é semelhante ao da taxa de clock. Se eles são diferentes, por quê?

1,11.8 [5] <§1.6> Em quanto o tempo de CPU foi reduzido?

1,11,9 [10] <§1.6> Para um segundo benchmark, libquantum, assumo um tempo de execução de 960 ns, CPI de 1,61 e taxa de clock de 3

GHz. Se o tempo de execução for reduzido em 10% adicionais sem afetar o CPI e com uma taxa de clock de 4 GHz, determine o número de instruções.

- 1.11.10 [10] <§1.6> Determine a taxa de clock necessária para dar uma redução adicional de 10% no tempo de CPU enquanto mantém o número de instruções e com o CPI inalterado.
- 1.11.11 [10] <§1.6> Determine a taxa de clock se o CPI for reduzido em 15% e o tempo de CPU em 20% enquanto o número de instruções permanecer inalterado.

1.12 A [seção 1.10](#) cita como uma armadilha a utilização de um subconjunto da equação de desempenho como uma métrica de desempenho. Para ilustrar isso, considere os dois processadores a seguir. P1 tem uma taxa de clock de 4 GHz, CPI médio de 0,9 e requer a execução de instruções $5.0E9$. P2 tem uma taxa de clock de 3 GHz, um CPI médio de 0,75 e requer a execução de instruções $1.0E9$.

1.12.1 [5] <§§1.6, 1.10> Uma falácia comum é considerar o computador com a maior taxa de clock como tendo o melhor desempenho. Verifique se isso é verdade para P1 e P2.

1.12.2 [10] <§§1.6, 1.10> Outra falácia é considerar que o processador que executa o maior número de instruções precisará de um tempo de CPU maior. Considerando que o processador P1 está executando uma sequência de instruções $1.0E9$ e que o CPI dos processadores P1 e P2 não muda, determine o número de instruções que P2 pode executar no mesmo tempo que P1 precisa para executar as instruções $1.0E9$.

1.12.3 [10] <§§1.6, 1.10> Uma falácia comum é usar MIPS (*milhões de instruções por segundo*) para comparar o desempenho de dois processadores diferentes e considerar que o processador com o maior MIPS tem o maior desempenho. Verifique se isso é verdade para P1 e P2.

1.12.4 [10] <§1.10> Outro número de desempenho comum é MFLOPS (milhões de operações de ponto flutuante por segundo), definido como

mas este número tem os mesmos problemas que MIPS. Suponha que 40% das instruções executadas em P1 e P2 sejam instruções de ponto flutuante. Encontre os números MFLOPS para os processadores.

1.13 Outra armadilha citada na [Seção 1.10](#) é esperar melhorar o desempenho geral de um computador melhorando apenas um aspecto do computador. Considere um computador executando um programa que requer 250 s, com 70 s para executar instruções FP, 85 s para executar instruções L / S e 40 s para executar instruções de desvio.

1.13.1 [5] <§1.10> Em quanto o tempo total é reduzido se o tempo para operações de PF for reduzido em 20%?

1.13.2 [5] <§1.10> Em quanto o tempo para operações INT é reduzido se o tempo total for reduzido em 20%?

1.13.3 [5] <§1.10> O tempo total pode ser reduzido em 20% reduzindo

apenas o tempo para instruções de ramal?

1.14 Suponha que um programa requer a execução de 50×10^6 instruções FP, 110×10^6 instruções INT, 80×10^6 instruções L / S e 16×10^6 instruções de desvio. O CPI para cada tipo de instrução é 1, 1, 4 e 2, respectivamente. Suponha que o processador tenha uma taxa de clock de 2 GHz.

1.14.1 [10] <§1.10> Em quanto devemos melhorar as instruções CPI de FP se quisermos que o programa rode duas vezes mais rápido?

1.14.2 [10] <§1.10> Em quanto devemos melhorar o IPC das instruções L / S se quisermos que o programa rode duas vezes mais rápido?

1.14.3 [5] <§1.10> Em quanto é o tempo de execução do programa melhorado se o IPC das instruções INT e FP for reduzido em 40% e o IPC de L / S e Branch for reduzido em 30%?

1.15 [5] <§1.8> Quando um programa é adaptado para rodar em múltiplos processadores em um sistema multiprocessador, o tempo de execução em cada processador é composto de tempo de computação e o tempo de overhead necessário para seções críticas bloqueadas e / ou para enviar dados de um processador para outro.

Suponha que um programa exija $t = 100$ s de tempo de execução em um processador. Ao executar p processadores, cada processador requer t / p s, bem como 4 s adicionais de overhead, independentemente do número de processadores. Calcule o tempo de execução por processador para 2, 4, 8, 16, 32, 64 e 128 processadores. Para cada caso, liste a aceleração correspondente em relação a um único processador e a razão entre a aceleração real e a aceleração ideal (aceleração se não houver sobrecarga).

Respostas para verificar você mesmo

§1.1, página 10: Questões para discussão: muitas respostas são aceitáveis.

§1.4, página 24: Memória DRAM: volátil, tempo de acesso curto de 50 a 70 nanossegundos e custo por GB é de \$ 5 a \$ 10. Memória de disco: não volátil, os tempos de acesso são 100.000 a 400.000 vezes mais lentos que a DRAM e o custo por GB é 100 vezes mais barato que a DRAM. Memória Flash: não volátil, os tempos de acesso são 100 a 1000 vezes mais lentos que a DRAM e o custo por GB é 7 a 10 vezes mais barato que a DRAM.

§1.5, página 28: 1, 3 e 4 são razões válidas. A resposta 5 pode ser geralmente verdadeira porque o alto volume pode tornar o investimento extra para reduzir o tamanho da matriz em, digamos, 10% uma boa decisão econômica, mas não precisa ser verdade.

§1.6, página 33: 1. a: ambos, b: latência, c: nenhum. 7 segundos.

§1.6, página 40: b.

§1.10, página 51: a. O computador A tem a classificação MIPS mais alta. b. O computador B é mais rápido.

