

# Lista de Exercícios

## Revisão de Arquiteturas Paralelas

Segundo Semestre de 2020

1. Do ponto de vista de desempenho de programas, por que razão transístores adicionais são utilizados em microprocessadores atuais para novos *cores* ao invés de serem utilizados para aumentar o número de unidades funcionais num superescalar?
2. Quais as vantagens de uma arquitetura SIMD sobre uma MIMD? Quais as vantagens da MIMD sobre a SIMD?
3. Qual a diferença entre um sistema UMA e um NUMA?
4. Por que razão os sistemas NUMA são em geral preferidos no caso de um número grande de processadores?
5. Qual a função das redes de interconexão num sistema paralelo?
6. Por que barramentos não costumam ser usados em sistemas com grande número de processadores?
7. Por que chaves *crossbar* não costumam ser usadas em sistemas com grande número de processadores?
8. Qual a maior distância entre dois nós em uma estrutura de malha bidimensional  $8 \times 8$ ? E qual a maior distância se ao invés de malha tivermos um toro das mesmas dimensões? Qual a maior distância numa malha tridimensional  $4 \times 4 \times 4$ ? E num toro dessas dimensões? Qual o número total de elos em todos esses casos? [Note que em todos os casos temos 64 nós.]
9. Qual a maior distância entre dois nós numa estrutura de hipercubo de dimensão 6? Qual o número total de elos? [Note que esta estrutura também tem 64 nós.]
10. Por que precisamos de coerência de cache em um sistema de espaço de endereçamento compartilhado?
11. No que consiste o *false sharing*? Por que ele é prejudicial ao desempenho de programas paralelos?
12. Num protocolo de coerência de cache por invalidação, quais eventos provocam a emissão de sinais de coerência no barramento? Quais eventos provocam a necessidade de transmissão de dados entre a memória e o cache ou vice-versa?
13. Por que protocolos de invalidação são usualmente implementados por *snooping*?
14. Indique um problema com *snooping* que é resolvido pelo uso de diretórios, e como isso é conseguido.
15. Por que implementações de diretório devem normalmente ser feitas de forma distribuída?
16. O sistema paralelo  $A$  tem parâmetros de comunicação  $t_s^{[A]} = 10ns$ ,  $t_w^{[A]} = 10ns$ , enquanto o sistema  $B$  tem parâmetros  $t_s^{[B]} = 900ns$ ,  $t_w^{[B]} = 1ns$ . Para esses dois sistemas, encontre o tempo total de comunicação para:

- (a) 100 mensagens de 10 palavras cada.
- (b) 10 mensagens de 100 palavras cada.
- (c) 1 mensagem de 1 000 palavras.