



## Computação Heterogênea de Alto Desempenho

Mestrado em Engenharia Eletrotécnica e de Computadores

LAB 1

– Conceitos genéricos sobre HPC e exercícios sobre Lei de Amdahl –

1. Defina ou expanda cada um dos seguintes termos ou acrónimos:

- Flops, gigaflops, teraflops, petaflops, exaflops, zettaflops
- Throughput
- Benchmark
- Parallel processing
- Moore's law
- Strong scaling
- Starvation
- Latency
- ALU
- von Neumann architecture
- DRAM e as diferenças principais relativamente a tecnologia SRAM
- SIMD
- SIMT
- SPMD
- VLSI

2. Qual é considerado o requisito principal que diferencia HPC de outros computadores? Que outros requisitos são importantes?

3. Indique seis técnicas para melhorar *throughput performance* num sistema .

4. Descreva o que entende por “*end of Moore's law*”.

5. Qual era o computador mais rápido no ano em que nasceu? Que tecnologias eram usadas nesse computador? Quão mais rápido é hoje o computador mais veloz do mundo?

6. Assuma que o cálculo da raiz quadrada é responsável por 20% do tempo gasto por um benchmark gráfico. Considere as seguintes alternativas:

- i. tornar o cálculo da raiz quadrada 10 vezes mais rápido;
- ii. tornar todas as instruções de vírgula flutuante 60% mais rápidas.

As instruções de vírgula flutuante perfazem 50% do tempo total de execução. Compare as duas alternativas e explique qual é mais vantajosa.

7. Considere que tem de acelerar um programa que se encontra a 30% de conseguir executar em tempo-real, numa CPU dual-core, demorando 1.3 seg a processar 60 *frames* (o objetivo consiste em atingir 60 *frames* por segundo). Após análise do *profiling* temporal do programa sequencial, descobriu-se que, entre outras funcionalidades, executa 6 FFTs que representam 40% do tempo total de processamento.

a) No sentido de acelerar o programa, admita que vamos usar 2 GPUs para processar as 6 FFTs. Sabendo que o potencial de aceleração de cada GPU é 22x, qual é o *speedup* obtido (ignore tempos de transferência de dados entre *host* e *device*)? Consegue atingir o objetivo inicial?

b) Admita agora que os tempos de transferência *host*  $\leftrightarrow$  *device* representam 35ms em cada sentido (*host*  $\Rightarrow$  *device* e *device*  $\Rightarrow$  *host*), para cada GPU. O objetivo é atingido?

8. Suponha que certo modo de execução de um programa é melhorado 10 vezes. O novo modo de funcionamento é usado 50% do tempo do programa original.

a) Determine o ganho de rapidez (*speedup*) obtido pela utilização do modo melhorado de funcionamento.

b) Que percentagem de tempo de execução original foi convertido para usar o novo modo de funcionamento?