

## Aula 15

• Avalia  o de desempenho de sistemas computacionais

- Defini  es e m  tricas de desempenho
- Desempenho do CPU
- Lei de Amdahl
- M  todos de avalia  o do desempenho
- Os *benchmarks* SPEC
- Desempenho, pot  ncia, efici  ncia energ  tica e cust

Bernardo Cunha, Jos   Luis Azevedo, Arnaldo Oliveira e Silva

## Introdu  o

• O que significa um computador apresentar melhor desempenho que outro?

• Qual a m  trica usada?

• Analogia com modelos de avi  es. Poss  veis m  tricas

• Capacidade, alcance, velocidade ou d  bito?

Modelo do avi��o	Capacidade (passageiros)	Alcance (Km)	Vel. Cruzeiro (Km/h)	D��bito (passageiros x Km/h)
Boeing 747	470	6700	980	460.600
Concorde	132	6500	2170	286.440
Douglas DC-8	146	14000	875	127.750

## O que é o desempenho?

O desempenho de um sistema pode definir-se relativamente a diversos factores:

- Tempo de resposta

- Número de operações executadas num dado intervalo

- Consumo de energia (eficiência energética)

- ...

Nos computadores, o tempo de execução dos programas está, normalmente, associado à definição de desempenho:

- Quanto menor for o tempo de execução, melhor será o desempenho

Para diferentes tipos de sistemas poderão, no entanto, ser usadas diferentes métricas de quantificação de desempenho.

## O que é o desempenho? - exemplos

Num **data center** que aceita pedidos de transações de vários utilizadores

- Um gestor de **data center** está interessado em aumentar o **throughput** de processamento das transações (trabalho total realizado num dado intervalo de tempo)

- O computador mais rápido é aquele que **realiza mais transações por unidade de tempo**

Num sistema embutido (**embedded system**) com restrições de tempo-real, a resposta a certos eventos deve/tem de ocorrer dentro de um intervalo de tempo limitado (pré-definido)

- Se o tempo de resposta estiver garantido, então **designers** poderão aumentar **throughput** ou reduzir o custo

Conclui-se novamente que, dependendo do sistema, podem ser usadas diferentes métricas para a avaliação de desempenho.

## Qual a complexidade na caracterização do desempenho

• Dificuldade da determinação do desempenho resultam complexidade dos sistemas:

- **Hardware** os computadores utilizam, cada vez mais, técnicas complexas (**não determinísticas**) de melhoria de desempenho
- **Software** complexidade (e dimensão) dos programas
- Um computador formado por vários blocos
- Cada um deles afecta o desempenho global de modo distinto
- Cada fabricante tem as suas próprias soluções implementadas
- Dependendo da complexidade do sistema pode ser impossível determinar o desempenho de forma analítica

## Qual a complexidade na caracterização do desempenho

• O desempenho global de um sistema é afectado **por hardware**, quer pelo **software**

• **Hardware:**

- **Microprocessador** é o elemento mais rápido do sistema: a melhoria do desempenho é obtida pelo aumento da frequência de relógio e pela minimização do número médio de ciclos de relógio para a execução das instruções
- **Memória** é mais lenta que o microprocessador: a melhoria do desempenho centra-se na minimização do tempo médio de acesso e escrita
- **Periféricos** podem realizar operações muito distintas: de impacto no desempenho global, sobretudo em programas que os usam de forma intensiva

• **Software:**

- **Compilador** a qualidade do código gerado tem impacto no desempenho global: sequências de instruções usadas, utilização de recursos internos, ...

## Definições de desempenho baseadas no tempo de execução

O tempo é a medida mais fiável de avaliação de desempenho dos computadores

O computador que realizar uma dada quantidade de trabalho em menos tempo é o mais rápido

Tempo de execução

Tempo entre o início e o fim da execução de uma tarefa

Se um mesmo programa for executado em dois computadores distintos (em igualdade de circunstâncias)

O computador que terminar a sua execução em menos tempo é o mais rápido

$$Desempenho_X = \frac{1}{T_{Execução}}$$

$$Desempenho_X > Desempenho_Y$$

$$\frac{1}{T_{Execução_X}} > \frac{1}{T_{Execução_Y}}$$

$$T_{Execução_X} < T_{Execução_Y}$$

$$\frac{Desempenho_X}{Desempenho_Y} = \frac{T_{Execução_Y}}{T_{Execução_X}} = n$$

## Métrica MIPS (Millions of Instructions Per Second)

Medida da velocidade de execução do programa baseado no número de instruções executadas por unidade de tempo

$$MIPS = \frac{\# \text{ Instruções Executadas}}{\text{Tempo de Execução}} \times 10^6$$

Problemas

Não toma em consideração o tipo e a complexidade das instruções

Não permite a comparação de computadores numa única vez que o número total de instruções para a mesma aplicação será diferente

Varia entre programas que executam no mesmo computador

Tem um comportamento anormal em certas situações: MIPS pode variar inversamente com o desempenho ver exemplo 4

Os fabricantes divulgam esta métrica assumindo que o processador executa repetidamente a instrução mais rápida (normalmente a instrução de movimento)

## Parâmetros temporais para a medição do desempenho

### Response time (latency ou execution time ou elapsed time)

O tempo que decorre entre o início e o fim de uma tarefa

Tempo de execução no CPU incluindo acessos à memória e a unidades de Input/Output (I/O), carga do sistema operativo e tudo

Pode ser impossível de avaliar analiticamente

### CPU time (ou CPU execution time)

O tempo que o CPU demora a executar uma dada tarefa em situação ideal. Não inclui:

O tempo em que o programa está suspenso e espera uma unidade de I/O (periférico) esteja pronta para transferir informação

O tempo extra necessário para aceder à memória

O tempo gasto a executar outros programas

CPU time pode ser medido executando o programa

## Desempenho do CPU (baseado no *CPU execution time*)

$$\text{Desempenho}_{CPU} = \frac{1}{\text{Tempo}_{CPU}}$$

$$\text{Tempo}_{CPU} = \text{Ciclos}_{CPU} \cdot \frac{1}{\text{freq}_{ciclo}^{CPU}} = \frac{\text{Ciclos}_{CPU}}{\text{Frequência}_{CPU}}$$

### Exemplo 1

Considere um computador com um CPU a funcionar com uma frequência de 4 GHz. O *CPU Execution Time* medido na execução de um dado programa foi de 10s.

Pretende-se desenvolver um novo CPU que execute o mesmo programa em 6s. O *hardware designer* verificou que é possível um aumento da frequência de trabalho do CPU, mas isso acarreta um aumento do número total de ciclos de relógio de 1,2 vezes ~~relativa~~ ao existente.

Qual a frequência de trabalho que deverá ter o novo CPU?

## Desempenho do CPU (baseado no CPI)

**CPI - Ciclos por Instrução** - número médio de ciclos de relógio que cada instrução de um programa demora a executar. Uma vez que instruções diferentes podem apresentar diferentes tempos de execução, o CPI é uma média de todas as instruções executadas no programa

$$Ciclos_{CPU} = \# Instruções \times CPI \quad \text{Tempo}_{CPU} = \frac{\# Instruções \times CPI}{Frequência_{CPU}} \quad (2)$$

A expressão (2) evidencia os três aspectos que afetam o desempenho. Como obter estes valores?

A frequência do CPU - é conhecida

O número de instruções - pode ser obtido pelo perfil de execução, durante a execução, de estatísticas de execução ou através de um simulador da arquitectura

O CPI - pode ser obtido por simulação ou através de contadores de hardware (quando o CPU já está operacional)

## Desempenho do CPU - Exemplo 2

Considerem-se duas máquinas com implementações diferentes da mesma arquitectura do conjunto de instruções (ISA). Um dado programa,

Máquina A: Clock\_cycle = 350 ps; CPI = 2,0

Máquina B: Clock\_cycle = 400 ps; CPI = 1,5

Qual a máquina mais rápida? Qual a relação entre o desempenho

## Desempenho do CPU – Exemplo 3

O projectista de um compilador para uma dada máquina está a tentar decidir entre duas sequências de código para a execução de uma instrução de alto nível.

Do conhecimento da implementação do processador sabe-se que há 3 classes de instruções A, B e C cuja execução requer 1, 2 e 3 ciclos de relógio, respectivamente.

A primeira sequência de código tem 10 instruções do tipo A, 5 do tipo B e 3 do tipo C.

A segunda sequência de código tem 12 instruções do tipo A, 8 do tipo B e 1 do tipo C.

Q1: Qual das duas sequências será mais rápida? Qual

Q2: Qual o CPI de cada sequência?

## Desempenho do CPU – Exemplo 4

Dois compiladores diferentes estão a ser testados numa máquina com frequência de CPU de 2 GHz. Está disponível um conjunto de instruções A, B e C cuja execução requer 1, 2 e 3 ciclos de relógio, respectivamente.

Os dois compiladores são utilizados para produzir código para um dado programa :

O 1º compilador gera código que contém 5 milhões de instruções da classe A, 1 milhão da classe B e 1 milhão da classe C.

O 2º compilador gera código que contém 10 milhões de instruções da classe A, 1 milhão da classe B e 1 milhão da classe C.

Q1: Qual das duas sequências produzidas será mais rápida de acordo com a máquina MIPS?

Q2: Qual será mais rápida de acordo com a máquina MIPS executando o código?



## Lei de Amdahl

Quantifica a melhoria que se pode esperar no tempo de execução de um programa num dado computador, se se melhorar de  $N$  vezes um dos factores que determina esse tempo

Considerando que o tempo de execução, antes da optimização, corresponde a:

$$T_{Exec} = T_{Exec}^{Afectado} + T_{Exec}^{N\%Afectado}$$

O tempo de execução após a optimização fica:

$$T_{Exec}^{Melhorado} = \frac{T_{Exec}^{Afectado}}{N} + T_{Exec}^{N\%Afectado}$$

## Lei de Amdahl - exemplo

Se o tempo de execução do programa for dado por:

$$T_{EXEC} = T_{CPU} + T_{MEM} + T_{PERIF}$$

Considere-se que:  $T_{CPU} = 1s$ ,  $T_{MEM} = 4s$ ,  $T_{PERIF} = 5s$  ( $T_{EXEC} = 10s$ )

Se se passar a frequência do CPU para o dobro,  $T_{CPU}$  será metade e o tempo de execução melhorado será:

$$T_{EXEC\_MELHORADO} = (T_{CPU} / 2) + T_{MEM} + T_{PERIF} = (1 / 2) + 4 + 5 = 9.5s$$

O aumento para o dobro da frequência de relógio resulta assim, numa melhoria no desempenho de:

$$Melhoria\_desempenho = T_{EXEC} / T_{EXEC\_MELHORADO} = 10 / 9.5 \cong 1,05 \text{ (5\%)}$$

A melhoria global é inferior à melhoria de apenas um factor

O impacto da melhoria é tanto maior quanto maior for o factor objecto de optimização no tempo total de execução:

Se interessa otimizar o(s) caso(s) mais frequente(s)



## Lei de Amdahl Exemplo 5

Considere um programa que executa em 100 s numa dada máquina, sendo que a unidade de multiplicação responde por 80 s desse tempo.

Q1: Qual o factor de melhoria que é necessário introduzir na unidade de multiplicação para que a execução do programa seja 4 vezes mais rápida?

Q2: Se se pretender uma melhoria do desempenho global de 5 vezes, qual o factor de melhoria a introduzir na unidade de multiplicação?

## Métodos usados na avaliação do desempenho de ArqC

O que faz sentido, para o utilizador final, é o conhecimento do desempenho do computador como um todo, isto é, executar programas reais, em cenário de utilização real

A metodologia usada consiste na definição de **benchmarks** para a avaliação normalizada

Cada fabricante pode testar os seus computadores de acordo com as normas do *benchmark* usado, e divulgar os resultados

Estes resultados permitem compará-los entre diferentes computadores

Métodos usados

**Benchmarks** sintéticos

**Workloads**

**Benchmarks** baseados em aplicações reais

## Avaliação do desempenho **Benchmarks sintéticos**

• Whetstone e Dhrystone

• Pequenos programas com sequências de instruções com base na frequência estatística de ocorrência em programas reais

• Muito usadas no passado (e em fases iniciais de desenvolvimento de uma nova arquitectura)

• Problemas:

• Estas sequências de código podem não representar características dos programas que efectivamente são executados

• Encorajam o recurso, por parte dos fabricantes, a optimizações especializadas (não gerais) dos compiladores e das arquitecturas de modo a obter bons resultados com estes *benchmarks* (as sequências de instruções são conhecidas)

## Avaliação do Desempenho

### • **Workload**

• Conjunto representativo dos programas (e frequências relativas) que vão de facto executar no computador

• A avaliação é efectuada com base no tempo de execução do *workload*

### • **Benchmarks**

• Um *benchmark* é composto por um conjunto de programas reais de uma dada classe de aplicações

Computadores de alto desempenho

Gráficos

Servidores, ...

• Um *benchmark* destina-se a comparar o desempenho dos vários sistemas dos computadores nessas classes de aplicações

## Os benchmarks **SPEC** ([www.spec.org](http://www.spec.org))

System Performance Evaluation Corporation

- Fundada em 1989 por fabricantes de computadores
- Disponibiliza benchmarks standard baseados em aplicações reais para avaliação de computadores pessoais e servidores
- Define as regras de execução das aplicações e a interpretação dos resultados

Classes de benchmarks

- CPU (SPEC CPU2006)
- Gráficos, Computação de alto desempenho
- Aplicações Java
- Servidores de e-mail e Web
- Sistemas de ficheiros
- Potência
- etc.

## Desempenho, potência, eficiência energética e custo

O consumo de potência é, cada vez mais, um **limite** chave no desempenho dos processadores, especialmente em sistemas embutidos e equipamentos portáteis

Alimenta as baterias e sistemas de dissipação passiva

**O consumo de potência é um aspecto tão importante como o desempenho ou o custo**

**Técnicas para reduzir o consumo:** redução da frequência (o consumo é proporcional à frequência), redução da tensão de alimentação, suspensão (parcial) do funcionamento

## Desempenho, potência, eficiência energética e custo

O desempenho é avaliado em diferentes modos de funcionamento

Potência máxima → alto desempenho

Potência mínima → maximiza a autonomia

Potência intermédia → reduz o consumo de energia

Em aplicações com limitações de consumo de energia, a eficiência energética é provavelmente a métrica mais importante

$$\text{Eficiência Energética} = \text{Desempenho} / \text{Energia consumida}$$

## Conclusão

Projecto realista de um sistema computacional deve em conta

Desempenho e a funcionalidade requeridas pelas aplicações alvo

O consumo de potência

O custo

Tipos de computadores com diferentes restrições

Computadores de alto desempenho e servidores de gama alta

Computadores de secretoria e servidores de gama baixa

Equipamentos portáteis e sistemas embutidos

O único método fiável de medir e reportar o desempenho é medir o tempo de execução de aplicações reais

Os três factores da expressão em conjunto exprimem o desempenho

$$\text{Tempo}_{CPU} = \frac{\# \text{Instruções}}{\text{Frequência da CPU}}$$