Universidade Federal do Rio Grande do Sul Instituto de Informática

Organização de Computadores Aula 9

Avaliação de Desempenho

Avaliação de desempenho

- 1. Introdução
- 2. Métricas básicas

Ciclos de clock

Tempo de CPU

Ciclos por instrução

3. Outras métricas e seus problemas

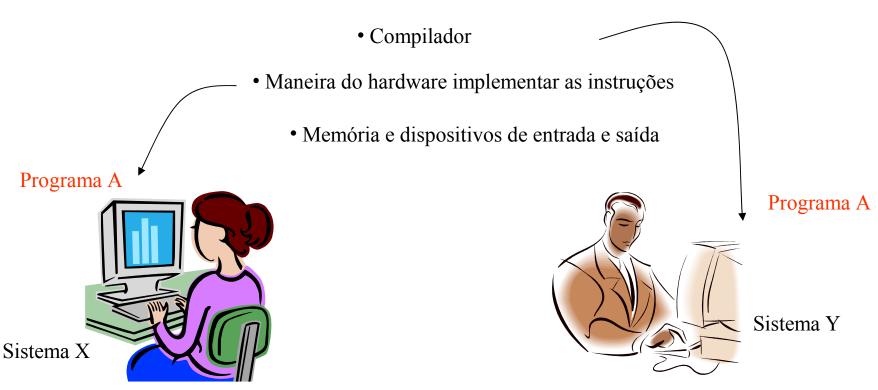
MIPS

MFLOPS

4. Benchmarks

1. Introdução Avaliação de Desempenho

- Desempenho = eficácia do sistema composto por hardware e software.
- Desafio em calcular a performance (desempenho) em sistemas cada vez mais complexos com softwares e hardwares diferentes.



Exemplo

Avião	Capacidade	AutonomiaVel	ocidade Thoug	hput
	(passageiros)	(milhas)	(milhas/hora)	Pass * milh/ hor
Boeing 777	375	4630	610	228.750
Boeing 747	470	4150	610	286.700
Concorde	132	4000	1350	178.200
Douglas	146	8720	544	79.424

Qual avião tem melhor desempenho se definirmos desempenho segundo sua velocidade?

- melhor velocidade de cruzeiro (levar 1 unico passageiro mais rapidamente)
- mais passageiros no menor tempo

Qual avião tem melhor desempenho se definirmos desempenho como a capacidade de levar mais passageiros?

Pensando em um usuário no computador..., ele está interessado mais é no tempo de resposta a um certo programa, ou seja, no tempo decorrido do inicio ao final do seu programa.

Introdução - Parâmetros

PARAMETROS DE MEDIDAS

- Tempo de relógio real necessário para execução de um programa (ou <u>tempo de resposta</u>) inclui ...
 - acessos a disco
 - atividades de I/O
 - overhead do sistema operacional, da rede, ...
- No entanto, seguidamente um sistema executa diversos programas simultaneamente
- Throughput, vazão, é uma medida do número de tarefas executada por unidade de tempo
- O que otimizar?
 - Tempo de resposta de um programa isolado?
 - Throughput do sistema?

Introdução - Tempo de Execução x Throughput

Para um dado sistema de computação:

- Subtituição de um processador por outro mais rápido, diminui o tempo de resposta e quase sempre melhora o throughput.
- Alocação de processadores adicionais a um sistema que utiliza varios processadores para executar programas diferentes, não faz com que as tarefas sejam realizadas mais rápidas pois os processadores são os mesmos, porém aumenta o throughput caso a demanda de processamento anteriormente fosse maior que a possivel.

Primeiramente iremos discutir tempo de resposta:

• Desempenho = 1

Tempo de execução

Introdução

- Distinção entre ...
 - tempo de relógio real gasto em um programa, e
 - tempo do processador CPU efetivamente gasto com tarefas específicas do programa
 - não inclui tempo de execução de outros programas, tempo de espera por I/O, ...
- Tempo do processador (CPU) está dividido em ...
 - tempo de usuário tempo gasto executando instruções do programa do usuário
 - tempo de sistema tempo gasto com tarefas do S.O. necessárias para a execução do programa do usuário

Introdução

• Exemplo no UNIX com o comando time

90.7s 12.9s 2:39m 65% (90.7 + 12.9) / 2:39 = 0.65 tempo do processador tempo do sistema tempo decorrido percentual do processado pelo decorrido Tempo em segundos e minutos.

- * Note que a maioria do tempo é gasto em entrada e saida e/ou rodando outros programas.
- Tendo em vista os objetivos desta disciplina de Organização, desempenho de CPU será medido apenas em função do tempo de CPU do usuário, ou seja, o tempo do processador gasto para a execução das instruções do programa em questão.

Introdução

- Desempenho é consequência de otimizações feitas em 3 dimensões do hardware:
- Arquitetura
 - processadores RISC
 - instruções otimizadas
- Organização
 - paralelismo
 - memória cache
- Tecnologia
 - velocidade de chaveamento das portas lógicas
 - tempo de acesso à memória
- Desempenho também depende no entanto do compilador
 - geração de código otimizada
 - bom aproveitamento dos recursos da organização

2. Métricas básicas – Período do Relógio

- <u>Ciclos de relógio</u> = intervalos básicos de tempo nos quais são executadas as operações elementares de uma instrução
 - transferências de valores entre registradores
 - operações aritméticas na ALU
- <u>Período do relógio</u> T = tempo de duração de um ciclo de clock
 - p.ex. 1 ns
- <u>Frequência de operação</u> f = número de ciclos de clock por unidade de tempo

$$f = 1 / T$$

se T = 1 ns
Então $f = 1 / 1.10^{-9} = 10^9 = 1 \text{ GHz}$

Métricas básicas – Tempo de CPU de um programa

χ período do clock

por exemplo:

- programa gasta 10.10⁶ ciclos
- período do clock é 1 ns



tempo de CPU do programa =

 $10.10^6 \times 1.10^{-9} = 10.10^{-3} = 10 \text{ ms}$

alternativamente:

nº de ciclos de clock do programa

frequência do clock

Formas de Aumento de Desempenho

Formas de aumento do desempenho:

- diminuir o período do clock
- > tecnologia, organização
- diminuir nº de ciclos necessários para execução do programa
 - organização, arquitetura

Objetivos muitas vezes conflitantes.

Métricas básicas – Exercício Frequência

Exercício

- programa roda em 10 s na máquina A, que tem clock de 1 GHz
- queremos rodá-lo em 6 s numa máquina B com nova tecnologia e nova organização
- a máquina B pode ter um aumento substancial de freqüência de clock
- no entanto, a organização da máquina B exigirá 1.2 vezes mais ciclos de clock para executar instruções do que a máquina A
- qual é a frequência de clock necessária para a máquina B?

Métricas básicas – Tempode CPU

Solução do exercício

$$10 s = \frac{n^{\circ} \text{ de ciclos de clock em A}}{10^{\circ} \text{ ciclos / s}}$$

nº de ciclos de clock em A = 10 . 109 ciclos

$$6 \text{ s} = \frac{1.2 \text{ x n}^{\circ} \text{ de ciclos de clock em A}}{\text{freqüência do clock em B}} = \frac{1.2 \text{ x } 10 \cdot 10^{9}}{\text{freqüência do clock em B}}$$

frequência do clock em B = 2 GHz

Métricas básicas: CPI

nº de ciclos de clock do programa = nº instruções de clock por instrução

CPI = nº médio de ciclos de clock por instrução é uma média sobre todas as instruções executadas no programa.

Exercício:

- Duas máquinas implementam o mesmo conjunto de instruções
- Um mesmo programa é rodado em ambas as máquinas
- Máquina A tem período de clock de 1 ns e CPI = 2.0 para este programa
- Máquina B tem período de clock de 2 ns e CPI = 1.2 para o mesmo programa
- Qual máquina é mais rápida ?

Métricas básicas CPI

Solução do exercício

```
n^{\circ} de ciclos de clock do programa = N n^{\circ} (n^{\circ} instruções do programa) X CPI
```

 n° de ciclos de clock de A = 2.0 x N

 n° de ciclos de clock de B = 1.2 x N

tempo de CPU e nº de ciclos de clock de um programa e nº de ciclos de clock do programa x período do clock

tempo de CPU de $A = 2.0 \times N \times 1 \text{ ns} = 2 \text{ N ns}$

tempo de CPU de B = $1.2 \times N \times 2 \text{ ns} = 2.4 \text{ N ns}$

portanto: máquina A é 1.2 vezes mais rápida do que B na execução deste programa

Métricas básicas: Tempo de CPU

tempo de CPU = nº de instruções X CPI X período do clock

- Como obter estes valores?
 - Tempo de CPU é medido executando-se o programa
 - Período (ou frequência) do clock é divulgado pelo fabricante
 - nº de instruções e CPI são mais difíceis de obter
 - nº de instruções depende da arquitetura e do compilador
 - CPI depende da organização e do programa
 - Conhecendo-se nº de instruções ou CPI, pode-se determinar o outro a partir da fórmula

Métricas básicas

Nº de ciclos de clock de um programa pode ser determinado ...

- Pela análise dos diferentes tipos de instruções executados pelo programa
- Pelo nº de ciclos de clock necessários para cada tipo de instrução

$$n^{\circ}$$
 de ciclos de clock do programa = $\sum_{i=1}^{n}$ ($CPI_{i} \times C_{i}$)

onde n = nº de tipos de instruções

CPI, = ciclos de clock por instrução do tipo i

C_i = nº de instruções do tipo i no programa

Métricas básicas

Exercício:

- O projetista de um compilador deseja decidir entre duas possíveis seqüências de código para a resolução de um problema
- Dados os tipos de instruções e o nº de ciclos por instrução de cada tipo, qual seqüência é mais rápida?

tipo de instrução	СРІ
A	1
B	2
C	3

	nº de instruções (x N)				
código	tipo A	tipo B	tipo C		
1	2	1	2		
2	4	1	1		

Métricas básicas

Solução do exercício

O código 1 executa 2 + 1 + 2 = 5 instruções

O código 2 executa 4 + 1 + 1 = 6 instruções

$$\frac{n^{o} \text{ de ciclos de}}{\text{clock do programa}} = \sum_{i=1}^{n} (CPI_{i} X C_{i})$$

 N° ciclos de clock para código 1 = (2 x 1) + (1 x 2) + (2 x 3) = 10 ciclos

 N° ciclos de clock para código 2 = (4 x 1) + (1 x 2) + (1 x 3) = 9 ciclos

$$CPI = \frac{\text{ciclos de clock}}{\text{n° de instruções}}$$

CPI código
$$1 = 10 / 5 = 2.0$$

CPI código
$$2 = 9 / 6 = 1.5$$

Código 2 é mais rápido, mesmo que execute uma instrução a mais, pois tem CPI bem mais baixo.

3. Outras métricas: MIPS

MIPS =
$$\frac{n^{\circ} \text{ instruções}}{\text{tempo de CPU X 10}^{\circ}}$$
 = $\frac{n^{\circ} \text{ instruções}}{\text{ciclos X período X 10}^{\circ}}$ = $\frac{n^{\circ} \text{ instruções X freqüência}}{n^{\circ} \text{ instruções X CPI X 10}^{\circ}}$ = $\frac{\text{freqüência}}{\text{CPI X 10}^{\circ}}$

Problemas com a medida em MIPS:

- Não se pode comparar máquinas com conjuntos de instruções diferentes, pois certamente o nº de instruções será diferente para um mesmo programa
- MIPS varia de um programa para outro na mesma máquina
- MIPS pode variar inversamente ao desempenho

MIPS

Exemplo:

tipo de instrução	СРІ
Α	1
В	2
С	3

	nº de instruções (x N)				
código	tipo A	tipo B	tipo C		
compil. 1	5	1	1		
compil. 2	10	1	1		

freqüência do clock = 1 GHz

Calcular o desempenho do código gerado por cada compilador:

- em MIPS
- em tempo de CPU

MIPS

ciclos de clock =
$$\sum_{i=1}^{n} (CPI_i X C_i)$$

$$CPI = \frac{\text{ciclos de clock}}{\text{n° instruções}}$$





$$CPI_1 = \frac{(5x1 + 1x2 + 1x3) \cdot 10^6}{(5 + 1 + 1) \times 10^6} = \frac{10}{7} = 1.43$$

$$CPI_2 = \frac{(10x1 + 1x2 + 1x3) \cdot 10^6}{(10 + 1 + 1) \times 10^6} = \frac{15}{12} = 1.25$$

$$MIPS_1 = \frac{1 \text{ GHz}}{1.43 \times 10^6} = 699$$

$$MIPS_2 = \frac{1 \text{ GHz}}{1.25 \times 10^6} = 800$$

MIPS

tempo de CPU =
$$\frac{n^{\circ} \text{ instruções X CPI}}{\text{freqüência}}$$

tempo₁ =
$$\frac{(5+1+1).10^6 \times 1.43}{1.10^9}$$
 = 10 ms

tempo₂ =
$$\frac{(10+1+1).10^6 \times 1.25}{1.10^9}$$
 = 15 ms

O código 2 gasta portanto mais tempo.

Apesar de ter MIPS com maior valor, o código 2 gasta bem mais instruções.

Outras métricas: MFLOPS

MFLOPS = $\frac{n^{\circ} \text{ de operações de ponto flutuante no programa}}{\text{tempo de CPU x } 10^{6}}$

Não se contam <u>instruções</u> de PF e sim <u>operações</u> para evitar comparações injustas entre máquinas com instruções diversas

- somas, subtrações, multiplicações, divisões
- precisão simples, precisão dupla
- instruções mais complexas: seno, raíz quadrada

Problemas na comparação:

- máquinas diferentes têm não apenas conjuntos diferentes de instruções, mas também de operações
- certas operações (p.ex. soma em precisão simples) são muito mais rápidas do que outras (p.ex. divisão em precisão dupla)
- como na medida de MIPS, pode-se ter gasto maior em tempo mesmo tendo mais MFLOPS

4. Benchmarks

- Benchmarks são conjuntos de programas representativos da carga de trabalho de uma máquina
 - Utilizados para avaliação de desempenho, segundo métricas já discutidas
 - Geralmente de domínio público
 - Procuram explorar repertório de recursos da arquitetura

Problemas com benchmarks

- Melhor avaliação seria feita com a carga de trabalho efetivamente utilizada em cada máquina (workload)
- Escolha dos benchmarks e aplicação rigorosa da metodologia de avaliação
- Fabricantes podem tentar otimizar arquitetura e/ou organização e/ou compilador para executar de forma mais eficiente apenas os benchmarks

Tipos de benchmarks

- Aplicações reais ou modificadas
 - Quais ?
 - Dados de entrada ?
 - Exemplos:
 - programas do SPEC (empresa de padronização de avaliação)
 - iCOMP (benchmarks da Intel, aplicações sob Windows, PC-workload)
 - transações de banco de dados (TPC-A, TPC-B, etc.)
 - simuladores de equações diferenciais (NAS: benchmarks NASA)
- Kernels
 - Linpack: rotinas de resolução de sistema de equações lineares
 - Livermore Loops, NAS kernels (simulação numérica) ...
- Benchmarks de "brinquedo"
 - Puzzle, Quicksort, Sieve, ...
- Benchmarks sintéticos
 - Dhrystone, Whetstone, ...

Exemplos de benchmarks

Dhrystone

- desenvolvido em 1984 e escrito em ADA, C ou Pascal

Whetstone

- primeiro grande programa de benchmark sintético
- exemplo de uma tabela de resultados:

www.dl.ac.uk/TCSC/disco/Benchmarks/whetstone.html

SPEC

- Standard Performance Evaluation Corporation

www.spec.org/benchmarks.html

- SPEC2000 www.specbench.org/cpu2000
 - CINT2000: 12 programas para cálculos intensivos com inteiros
 - CFP2000: 14 programas para cálculos intensivos de ponto flutuante
 - máquina de referência: SUN Ultra 5_10 com processador SPARC de 300MHz e 256MB de memória
- SPEC2004 www.specbench.org/cpu2004

What is SPEC?

SPEC is the Systems Performance Evaluation Cooperative.

www.spec.org

- The current membership list includes 16 companies: AT&T, Control Data Corp., Data General, Digital Equipment Corp., DuPont, Hewlett-Packard, IBM, Intel, Intergraph, MIPS Computer Systems, Motorola, Multiflow, Solbourne, Stardent, Sun and Unisys.
- The SPEC applications represent a large body of code (over 14 megabytes) which span a range of application arenas.
- The membership to SPEC is open to any interested company.
- SPEC is not devoted to any single architecture nor any particular philosophy of computing systems.
- SPEC has created a framework in which a wide variety of applications can be tested by a very large audience of computer users.

SPEC95 Programs

Artificial intelligence; plays the game of Go

Plasma physics; electromagnetic particle simulation

Description

Integer

Benchmark

go

m88ksim Motorola 88k chip simulator; runs test program The Gnu C compiler generating SPARC code gcc Compresses and decompresses file in memory compress Lisp interpreter Graphic compression and decompression ijpeg Manipulates strings and prime numbers in the special-purpose programming language Perl perl A database program vortex tomcatv A mesh generation program Shallow water model with 513 x 513 grid swim Quantum physics; Monte Carlo simulation su2cor Astrophysics: Hydrodynamic Naiver Stokes equations hydro2d Multigrid solver in 3-D potential field mgrid Parabolic/elliptic partial differential equations applu trub3d Simulates isotropic, homogeneous turbulence in a cube Solves problems regarding temperature, wind velocity, and distribution of pollutant apsi Quantum chemistry fpppp

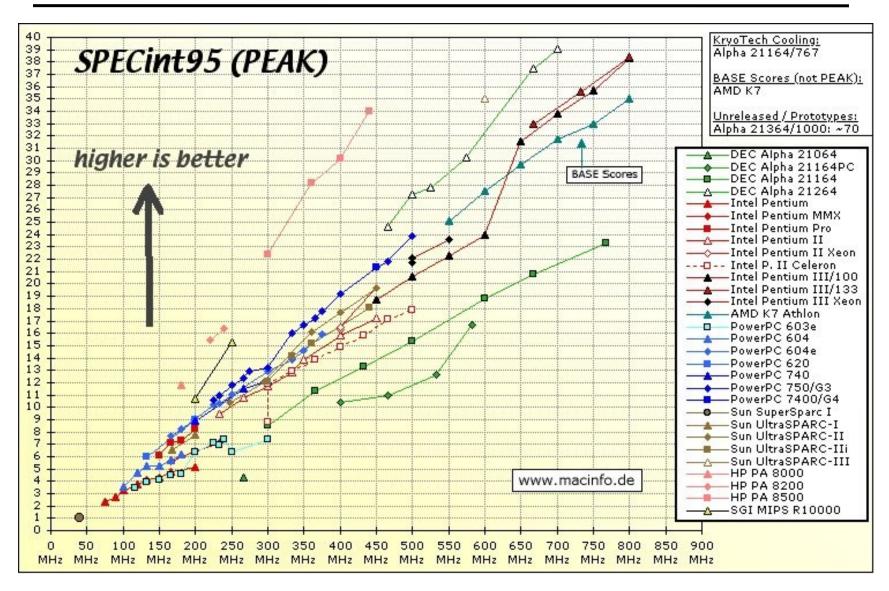
Floating Point

wave5

Medidas de Tempo de Execução no SPEC

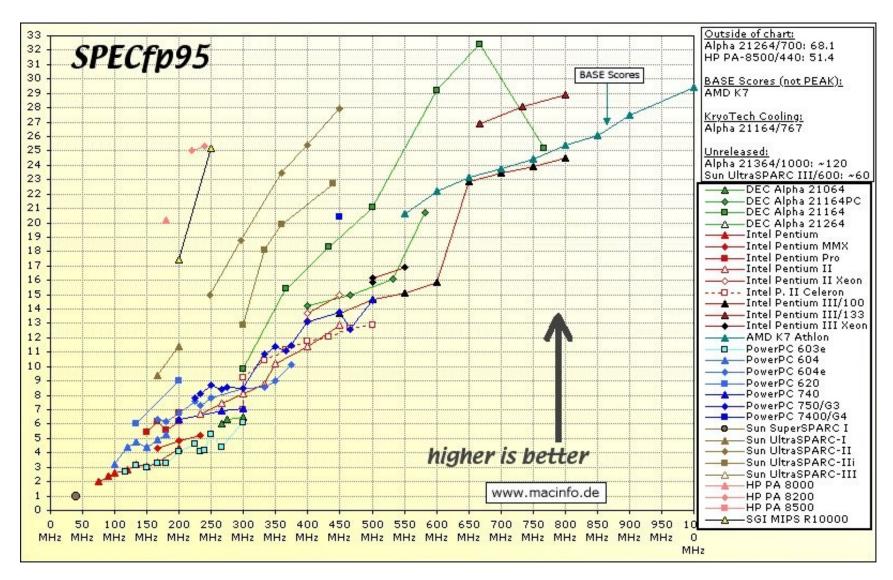
- As medidas de tempo de execução são normalizadas por meio da divisão pelo tempo gasto para executar o benchmark na estação de trabalho SUN SPARC station 10/40, pelo tempo de execução onde a medição esta sendo realizada. = razão SPEC.
- A medida final de performance é medida pela média geométrica das razões SPEC para cada programa do conjunto, por exemplo SPECint95 e SPECfp95.

Sample SPECint95 Results



Source URL: http://www.macinfo.de/bench/specmark.html

Sample SPECfp95 Results



Source URL: http://www.macinfo.de/bench/specmark.html

SPEC CPU2000

• CINT 2000

Benchmark	Função	Fonte
gzip	compressão de arquivos	С
vpr	posicionamento e roteamento de circuitos FPGA	С
gcc	compilador C	С
mcf	otimização combinacional	С
crafty	jogo: xadrez	С
parser	parser	С
eon	computação gráfica	C++
perlmbk	linguagem de programa PERL	С
gap	linguagem e biblioteca para computação em grupo	С
vortex	banco de dados orientado a objetos	С
bzip2	compressão de arquivos	С
twolf	simulador de posição e rota	С

SPEC CPU2000

• CFP 2000

Benchmark	Função	Fonte
wupwise	computação da propagação de quarks	Fortran77
swim	modelo da superfície d'água com 513x513 pontos	Fortran77
mgrid	soluções multigrid em campos 3D	Fortran77
apply	equações parciais diferenciais	Fortran77
mesa	biblioteca gráfica 3-D	С
galgel	dinâmica de fluídos	Fortran90
art	reconhecimento de imagens / redes neurais	С
equake	simulação de propagação de ondas sísmicas	С
facerec	processamento de imagens: reconhecimento de faces	Fortran90
ammp	modelo de um grande sistema de moléculas	С
lucas	teoria de números / teste de números primos	Fortran90
fma3d	simulação de ruídos elétricos em elementos finitos	Fortran90
sixtrack	energia nuclear: acelerador de partículas	Fortran77
apsi	problemas com temperatura, vento e poluição	Fortran77

Top 20 SPEC CPU2000 Results (As of March 2002)

		Top 20 SPECi	nt2000			Top 20 SPECfp	2000	
#	MHz	Processor	int peak	int base	MHz	Processor	fp peak	fp base
1	1300	POWER4	814	790	1300	POWER4	1169	1098
2	2200	Pentium 4	811	790	1000	Alpha 21264C	960	776
3	2200	Pentium 4 Xeon	810	788	1050	UltraSPARC-III Cu	827	701
4	1667	Athlon XP	724	697	2200	Pentium 4 Xeon	802	779
5	1000	Alpha 21264C	679	621	2200	Pentium 4	801	779
6	1400	Pentium III	664	648	833	Alpha 21264B	784	643
7	1050	UltraSPARC-III Cu	610	537	800	Itanium	701	701
8	1533	Athlon MP	609	587	833	Alpha 21264A	644	571
9	750	PA-RISC 8700	604	568	1667	Athlon XP	642	596
10	833	Alpha 21264B	571	497	750	PA-RISC 8700	581	526
11	1400	Athlon	554	495	1533	Athlon MP	547	504
12	833	Alpha 21264A	533	511	600	MIPS R14000	529	499
13	600	MIPS R14000 500	483	675		SPARC64 GP	509	371
14	675	SPARC64 GP 478	449	900		UltraSPARC-III	482	427
15	900	UltraSPARC-III	467	438	1400	Athlon	458	426
16	552	PA-RISC 8600	441	417	1400	Pentium III	456	437
17	750	POWER RS64-IV	439	409	500	PA-RISC 8600	440	397
18	700	Pentium III Xeon	438	431	450	POWER3-II	433	426
19	800	Itanium	365	358	500	Alpha 21264	422	383
20	400	MIPS R12000 353	328	400		MIPS R12000	407	382
								36

INF01113 - Organização de Computadores

Source: http://www.aceshardware.com/SPECmine/topujspx 2010

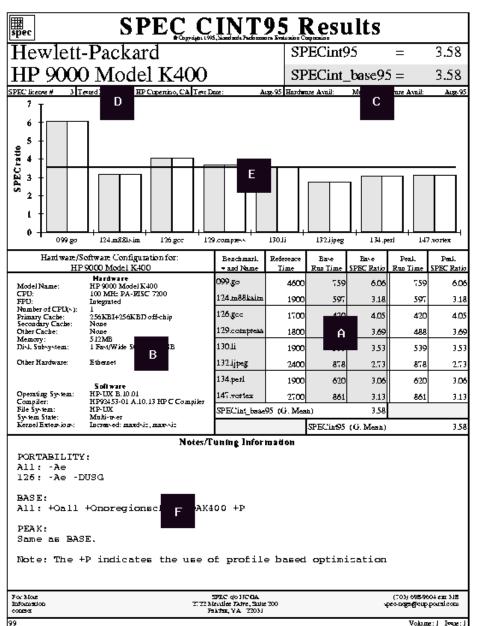
Resultados do benchmark Whetstone-97

www.dl.ac.uk/TCSC/disco/Benchmarks/whetstone.html

Rank	Máquina	TotalCPU (s)	MWIPS
1	Pentium 4/2666 (ifc)	10.4	3532
2	HP RX5670 Madison/1500 (+)	10.6	3532
3	IBM pSeries 690Turbo/1.7	10.8	3472
4	Compaq Alpha ES45/1250	10.9	3441
5	IBM Regatta-HPC/1300	11.5	3281
6	IBM pSeries 690Turbo/1.3	11.7	3260
7	Compaq Marvel EV7/1000	13.0	2894
8	Compaq Alpha ES45/1000	13.6	2778
9	IBM pSeries 630/1000	15.4	2461
10	HP RX2600 Itanium2/1000	15.9	2369
11	AMD Opteron244/1800 (pgi)	15.9	2358
12	SGI Onyx 300/R14k-600	16.4	2280

interações do Benchmark / s

Exemplo de relatório de resultados do SPEC



A: seção de resultados

B: configuração do sistema

C: disponibilidade

D: patrocinador (sponsor)

E: gráfico

F: seção de notas

Resumo:

- Qual a métrica mais importante? MIPS? MFLOPS?
- Como escolher uma CPU?
 - Spec int?
 - Spec FP?
 - Média dos dois?
 - Whetstone?
- Que outros fatores impactam a escolha de uma CPU?

END

Exercício 1

Para um dado programa:

Classe	CPI em M1	CPI em M2	Codigo por C1	Codigo por C2	Codigo por terceiros
A	4	2	30%	30%	50%
В	6	4	50%	20%	30%
C	8	3	20%	50%	20%

A máquina M1 tem um clock de 400MHz e a M2 tem um clock de 200MHz.

Usando o compilador C1 na maquina M1 e o compilador C2 na máquina M2, qual é a mais rapida e de quanto comparado com a outra?

Se tiveres que comprar a máquina M2, qual compilador deverias usar para conseguir uma melhor performance para o programa em questão.

Exercício 2

 Considere duas implementações diferentes M1 e M2 para o mesmo conjunto de instruções. Existem 4 classes de instruções.

Classe	CPI para M1	CPI para M2
A	1	2
В	2	2
C	3	4
D	4	4

- M1 tem um clock de 500MHz e M2 tem um clock de 750MHz.
- Suponha que o desempenho de pico é definida como a taxa mais elevada que uma máquina possa executar uma seqüência de instruções, escolhida para maximizar essa taxa. Determine o desempenho de pico para M1 e M2 dada em instruções executadas por segundo.
- Se o numero de instruções de um programa for dividido igualmente entre as classes, qual máquina executa o programa mais rapidamente?

Exercício 3

• A tabela a seguir mostra o numero de operações em ponto flutuante executadas por dois programas diferentes além do tempo de execução em 3 máquinas distintas.

Programa Operaçõ	ses em ponto flutuante	Tempo de execução (s)			
-	-	Computador A	Computador B	Computador C	
Programa 1	10.000.000	1	10	20	
Programa 2	100.000.000	1.000	100	20	

Qual das máquinas é mais rápida considerando o tempo de execução. E quanto mais rápida ela é comparada as outras duas?