

Relatório Trabalho Prático - Benchmarks

Erian Alves¹, Guilherme Sérgio², Pedro Cardoso³

¹Instituto de Ciências Exatas e Tecnológicas,
Universidade Federal de Viçosa, Florestal, MG, Brasil

1. Introdução

Medir o desempenho de um computador pode ser algo útil para várias situações, seja para usuários compararem seus dispositivos com modelos mais novos, ou para projetar um hardware, desta forma se introduz os benchmarks. Os benchmarks consistem na realização de um conjunto de tarefas no computador com o propósito de quantificar o desempenho da máquina, por meio de uma pontuação gerada através de uma comparação com um computador referência.

Visto isso, o objetivo deste trabalho é analisar o desempenho dos computadores em posse dos estudantes, utilizando benchmarks. Para isso, serão escolhidos dois computadores chamados de CompA e CompB para serem analisados juntamente com um computador referência (CompRef) disponibilizado.

Para começar o trabalho, primeiramente, foram escolhidos os dois computadores aos quais seriam feitos os testes. O desenvolvimento do trabalho será mostrado a seguir.

2. Especificação dos computadores

Os computadores escolhidos, como dito anteriormente, serão chamados de CompA e CompB. As especificações do CompRef também serão apresentadas. Importante ressaltar que, durante o processo de obtenção das informações, por meio dos programas que serão descritos na próxima seção, a taxa de clock do computador B foi fornecida considerando sua capacidade máxima. Seguem abaixo suas respectivas especificações.

CompA:

Processador: Intel(R) Pentium(R) CPU G4560 @ 3,5GHz

Sistema Operacional: Microsoft Windows 10 pro (64-bit)

Chipset: Gigabyte Core7

Placa-mãe: GigaByte H110M-S2V-CF

Memória Primária (RAM): 1 x 4GB DDR4 2.4GHz SDRAM

Memória Cache de Instrução L1: 32kB, 8-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache de Dados L1: 32kB, 8-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache L2: 256kB, ECC, 4-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache L3: 3MB, ECC, 12-Way, Fully Inclusive, 64bytes Line Size, 16 Thread(s)

CompB:

Processador: AMD Ryzen 5 1600 Six-Core Processor @ 3.7GH

Sistema Operacional: Microsoft Windows 10 Home (64-bit)

Chipset: AMD F17

Placa-mãe: ASUSTeK COMPUTER INC. PRIME A320M-K/BR

Memória Primária (RAM): 2 x 4GB 2.4GHz DDR4 SDRAM

Memória Cache de Instrução L1: 6x 32kB, 8-Way, Exclusive, 64bytes Line Size, 2 Thread(s)

Memória Cache de Dados L1: 6x 32kB, 8-Way, Exclusive, 64bytes Line Size, 2 Thread(s)

Memória Cache L2: 6x 512kB, 8-Way, Fully Inclusive, 64bytes Line Size, 2 Thread(s)

Memória Cache L3: 2x 8MB, 16-Way, Exclusive, 64bytes Line Size, 8 Thread(s)

CompRef:

Processador: Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz

Sistema Operacional: Microsoft Windows 10 (64-bit)

Chipset: Acer Core7H

Placa-mãe: Acer Nitro AN515-51 Acer Nitro 5

Memória Primária (RAM): 2 x 8GB 2.4GHz DDR4 SDRAM

Memória Cache de Instrução L1 : 4x 32kB, 8-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache de Dados L1 : 4x 32kB, 8-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache L2: 4x 256kB, ECC, 4-Way, 64bytes Line Size, 2 Thread(s)

Memória Cache L3: 6MB, ECC, 12-Way, Fully Inclusive, 64bytes Line Size, 16 Thread(s)

3. Medidas de Desempenho

Para a coleta das medidas de desempenho foram utilizados dois programas, o GeekBench e o Sisoftware Sandra. O GeekBench retorna uma pontuação a partir da realização de um teste que inclui tarefas como compressão de dados, processamento de imagens, aprendizado de máquina e simulações de físicas (como colisões e fricções). É importante citar também que o GeekBench realiza o teste duas vezes, uma vez utilizando apenas um único Core e outra utilizando Multi-Cores, gerando assim duas pontuações. Portanto, se tratando desta métrica em questão, quanto maior a pontuação, maior o desempenho.

Já o software Sandra é mais completo, realizando além de testes similares aos do GeekBench, testes mais específicos, como de latência e largura de banda das memórias cache e da memória principal. A seguir são apresentadas as métricas obtidas para cada computador, incluindo o computador referência:

CompA:

Pontuação GeekBench: 890 Single-Core e 1875 Multi-Core

CPU Aritmético:

Integer Score: 70660 MIPS

Floating Point Score: 47430 MFLOPS

Cache L1: Latência - $4/\text{Frequência} = 4/3.5\text{GHz} = 1,143\text{ns}$

Cache L2: Latência - $12.5/\text{Frequência} = 12,5/3.5\text{GHz} = 3,572\text{ns}$

Memória Primária:

Memory Latency: 35.1ns

Largura de banda:

Write int: 11GB/s

Write float: 11.23GB/s

CompB:

Pontuação GeekBench: 929 Single-Core e 5004 Multi-Core

CPU Aritmético:

Integer Score: 216220 MIPS

Floating Point Score: 136000 MFLOPS

Cache L1: Latência - $4/\text{Frequência} = 4/3.7\text{GHz} = 1,08\text{ns}$

Cache L2: Latência - $12.3/\text{Frequência} = 12,3/3.7\text{GHz} = 3,32\text{ns}$

Memória Primária:

Memory Latency: 75.6ns

Largura de banda:

Write int: 27.63GB/s

Write float: 29GB/s

CompRef:

Pontuação GeekBench: 1000 Single-Core

CPU Aritmético:

Integer Score: 102000 MIPS

Floating Point Score: 63230 FLOPS

Cache L1: Latência – 5 ciclos = $(5 \times 1/\text{Frequência}) = 5/2,8\text{ GHz} = 1,78\text{ ns}$

Cache L2: Latência – 11,5 ciclos = $(11,5 \times 1/\text{Frequência}) = 11,5/2,8\text{ GHz} = 4,1\text{ ns}$

Memória Primária:

Latência: 40.8 ns

Largura de Banda:

Write float: 21,92 GB/s

Write int: 21,65 GB/s

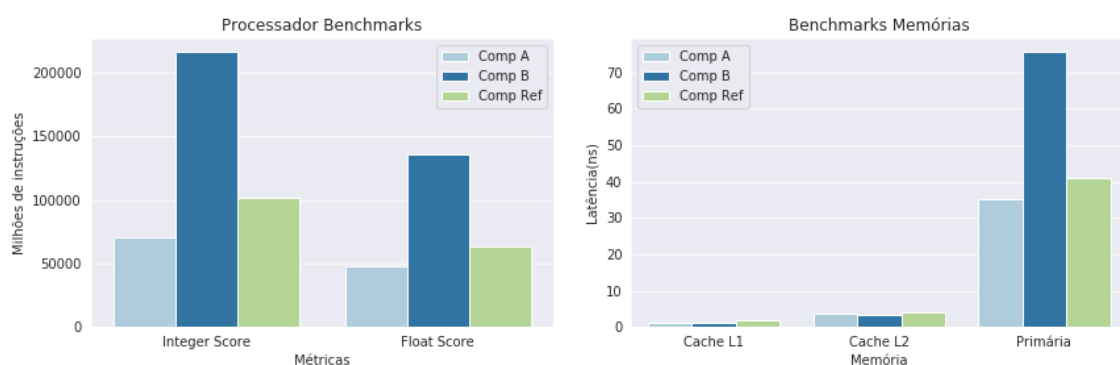


Figura 1. Gráficos da quantidade de instruções e das latências das memórias.

Através dos gráficos é possível notar uma coisa muito interessante, apesar do computador B poder realizar muito mais instruções por segundo, tanto de inteiros quanto de floats, a latência de suas memórias não é a menor entre os três computadores.

4. Solução de Problemas Hipotéticos

O primeiro problema solicitava o cálculo do tempo de execução do programa BABE nos três computadores descritos anteriormente. Para realizar o cômputo foram utilizadas as seguintes fórmulas fornecidas na descrição do trabalho:

- Tempo = número de instruções / MIPS - (TULA)
- Tempo = número de instruções / MFLOPS - (TUPF)
- Tempo = número de acessos x Latência - (TCL1, TCL2, TRAM)

Temos que, a primeira fórmula diz respeito ao Tempo da Unidade de Lógica Aritmética, a segunda o Tempo de Ponto Flutuante, e a terceira fala sobre o tempo de acesso às memórias.

Por fim os valores seriam somados para obter o tempo total de execução do programa.

BABE:

Número instruções INT = $2 \times 10^{12} \times 0.2 = 0.4 \times 10^{12} = 400000$ MI

Número instruções Ponto Flutuante = $2 \times 10^{12} \times 0.8 = 1.6 \times 10^{12} = 1600000$ MI

Número de acessos L1 = $1 \times 10^7 \times 0.72 = 0.72 \times 10^7 = 7200000$

Número de acessos L2 = $1 \times 10^7 \times 0.15 = 0.15 \times 10^7 = 1500000$

Número de acessos RAM = $1 \times 10^7 \times 0.13 = 0.13 \times 10^7 = 1300000$

Após a aplicação dos valores nas fórmulas foram obtidos os seguintes resultados:

CompA:

Tempo total = TULA + TUPF + TCL1 + TCL2 + TRAM

Tempo Total = 5.661s + 33.73s + 0.008573s + 0.005355s + 0.004641s

Tempo Total = 39.41s

CompB:

TempoTotal = TULA + TUPF + TCL1 + TCL2 + TRAM

TempoTotal = 1.84s + 11.76s + 0.00777s + 0.00498s + 0.0982s

TempoTotal = 13.71s

CompRef:

TempoTotal = TULA + TUPF + TCL1 + TCL2 + TRAM

Tempo Total = 3.92s + 25.30s + 0.00128s + 0.00615s + 0.053s

Tempo Total = 29.28s

Com o propósito de promover uma melhor visualização das informações obtidas, foi produzido o gráfico a seguir, o qual compara os tempos de execução.

Na sequência, foi avaliado o desempenho das máquinas selecionadas em relação ao computador referência. Por meio da razão, Tempo de execução CompRef / Tempo de execução máquina selecionada, é possível analisar qual computador foi mais rápido ou lento ao rodar o programa proposto e qual era essa diferença entre os aparelhos. Os

resultados obtidos foram os seguintes:

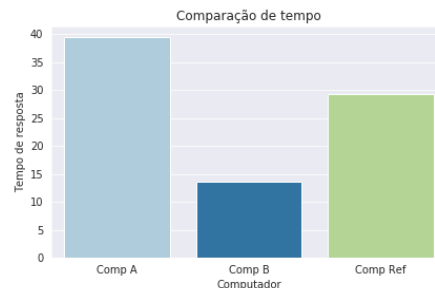


Figura 2. Gráfico tempo de execução de cada máquina para o programa BABE

Aceleração:

$$\text{CompRef} / \text{CompA} = 29.28\text{s} / 39.41\text{s} = 0.742$$

$$\text{CompRef} / \text{CompB} = 29.28\text{s} / 13.71\text{s} = 2.135$$

A partir das informações obtidas, identificou-se o computador B como aquele que apresentou melhor performance e, devido a isso, ele foi a máquina selecionada para os cenários posteriores.

A questão seguinte propunha uma situação na qual a memória cache L2 não seria contabilizada no cálculo do tempo e todos os seus acessos iriam para a memória RAM. Ao realizar um novo cômputo, é possível notar que o tempo de execução aumentou e, portanto, ficou 1,008 mais lento que no cenário determinado inicialmente.

$$\text{TempoTotal} = \text{TULA} + \text{TUPF} + \text{TCL1} + \text{TCL2} + \text{TRAM}$$

$$\text{TempoTotal} = 1.84\text{s} + 11.76\text{s} + 0.00777\text{s} + 0.21196\text{s}$$

$$\text{TempoTotal} = 13.8198\text{s}$$

$$\text{Speedup} = 13.8198 / 13.71 = 1.008$$

Este aumento é explicado ao observar a fórmula de cálculo do tempo de acesso às memórias, logo que o atributo de latência de cada uma impacta diretamente no aumento ou redução. Ao retirar os acessos da L2, a qual possui a menor latência, e transferir para a RAM, além dessa última possuir um valor maior para o fator citado, o incremento no número de acessos também afeta no acréscimo do tempo.

Na sequência, uma diferente circunstância é apresentada, na qual a memória RAM é responsável por todos os acessos.

$$\text{TempoTotal} = \text{TULA} + \text{TUPF} + \text{TRAM}$$

$$\text{TempoTotal} = 1.84\text{s} + 11.76\text{s} + 0.757\text{s}$$

$$\text{TempoTotal} = 14.357\text{s}$$

$$\text{Speedup} = 14.357 / 13.71 = 1.04$$

Após os cálculos, nota-se que o computador BEST ficou 1,04 mais lento em relação a sua performance inicial e isso ocorre pela lógica já explicada no cenário an-

terior.

O problema seguinte aborda, novamente, a relação entre o número de acessos em determinada memória e a variação de desempenho. Nesta nova perspectiva, a memória cache L2 é a única responsável por este trabalho. Ao realizar o cômputo do tempo, é possível perceber que a máquina BEST torna-se 1,0056 mais rápida.

$$\begin{aligned}\text{TempoTotal} &= \text{TULA} + \text{TUPF} + \text{TCL2} \\ \text{TempoTotal} &= 1.84\text{s} + 11.76\text{s} + 0.0332\text{s} \\ \text{TempoTotal} &= 13.6332\text{s} \\ \text{Speedup} &= 13.71 / 13.6332 = 1.0056\end{aligned}$$

Ao analisar um cenário no qual a memória cache L2 fosse do tamanho da memória principal, a partir dos dados a seguir, conclui-se que o custo adicional em reais é um valor muito elevado para o aumento de desempenho encontrado anteriormente.

$$\begin{aligned}\text{Tamanho atual L2: } &6 \times 512\text{kB} = 3072\text{kB} = 3\text{M} \\ \text{Relação tamanho e preço proposto: } &1\text{M} - \text{R\$ } 50 \\ \text{Ajustando as unidades: } &1024 \text{ M} = 1\text{GB} - \text{R\$ } 51.200 \\ \text{Memória principal: } &8\text{GB} \\ \text{Preço L2 tamanho da memória principal: } &8 \times 51.200 = \text{R\$ } 409.600 \\ \text{Custo adicional: } &\text{R\$ } 409.600 - \text{R\$ } 150 = \text{R\$ } 409.450\end{aligned}$$

Na sequência é solicitado o cálculo do tempo de execução do programa BABE, mas, agora, num contexto no qual o processador tivesse diferentes quantidades de unidade de lógica aritmética e unidade de ponto flutuante. Os tempos podem ser observados no seguinte desenvolvimento e as acelerações são expostas no gráfico na sequência:

CompB:

$$\begin{aligned}\text{TempoTotal} &= \text{TULA} + \text{TUPF} + \text{TCL1} + \text{TCL2} + \text{TRAM} \\ \text{TempoTotal} &= 1.84\text{s} + 11.76\text{s} + 0.00777\text{s} + 0.00498\text{s} + 0.0982 \\ \text{TempoTotal} &= 13.71\text{s}\end{aligned}$$

2 ULAs e UPFs

$$\begin{aligned}\text{TULA} &= 1.84/2 = 0.92\text{s} \\ \text{TUPF} &= 11.76/2 = 5.88\text{s} \\ \text{Tempo Total} &= 0.92\text{s} + 5.88\text{s} + 0.00777\text{s} + 0.00498\text{s} + 0.0982\text{s} \\ \text{Tempo Total} &= 6.91\text{s} \\ \text{Speedup} &= 13.71/6.91 = 1.9840\end{aligned}$$

4 ULAs e UPFs

$$\begin{aligned}\text{TULA} &= 1.84/4 = 0.46\text{s} \\ \text{TUPF} &= 11.76/4 = 2.94\text{s} \\ \text{Tempo Total} &= 0.46\text{s} + 2.94\text{s} + 0.00777\text{s} + 0.00498\text{s} + 0.0982\text{s} \\ \text{Tempo Total} &= 3.51\text{s} \\ \text{Speedup} &= 13.71/3.51 = 3.906\end{aligned}$$

8 ULAs e UPFs

$TULA = 1.84/8 = 0.23s$
 $TUPF = 11.76/8 = 1.47s$
 $Tempo\ Total = 0.23s + 1.47s + 0.00777s + 0.00498s + 0.0982s$
 $Tempo\ Total = 1.811s$
 $Speedup = 13.71/1.811 = 7.5704$

16 ULAs e UPFs
 $TULA = 1.84/16 = 0.115s$
 $TUPF = 11.76/16 = 0.735s$
 $Tempo\ Total = 0.115s + 0.735s + 0.00777s + 0.00498s + 0.0982s$
 $Tempo\ Total = 0.961s$
 $Speedup = 13.71/0.961 = 14.2663$

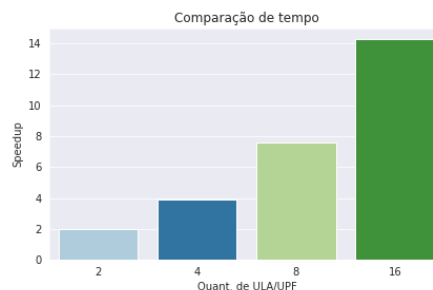


Figura 3. Tempo de execução do programa BABE com diferentes quantidades de ULA e UPF

O impacto do aumento do número desses componentes na performance deve-se ao fato de que, inicialmente, havia apenas um de cada para lidar com as respectivas instruções, a partir do momento em que há o acréscimo, o número de instruções pode ser dividido, reduzindo a carga de um único ULA e UPF, ele terá um tempo de execução menor.

A tarefa seguinte estabelecia algumas alterações nas características do programa BABE com o objetivo de obter algum ganho no tempo de execução. O tempo de execução das instruções de ponto flutuante seria melhorado em 50%, em outras palavras, reduzido pela metade, e a latência de acesso a cache L1 diminuído em 30%. Após computar os novos valores nas fórmulas, podemos observar por meio da aceleração de cada computador o ganho geral no tempo de execução. Ademais, o gráfico a seguir mostra a nova relação entre os tempos de execução.

CompA:

Cache L1: $1.143 \times 10^{-9}s \times 0.7 = 0.8 \times 10^{-9}s$
 $TUPF: 1600000\ MI / 47430\ MFLOPS = 33.73s \times 0.5 = 16,86s$
 $Tempo\ total = TULA + TUPF + TCL1 + TCL2 + TRAM$
 $Tempo\ Total = 5.661 + 16.86 + 0.006 + 0.005355 + 0.004641$
 $Tempo\ Total = 22.53s$
 $Tempo\ Total\ (Antes\ das\ modificações) = 39.41s$
 $Ganho\ de\ desempenho = 39.41 / 22.53 = 1.75$

CompB:

Cache L1: $1.08 \times 10^{-9} \text{ s} \times 0.7 = 0.756 \times 10^{-9} \text{ s}$

TUPF = $1600000 \text{ MI} / 136000 \text{ MFLOPS} = 11,76 \text{ s} \times 0.5 = 5,88 \text{ s}$

TempoTotal = TULA + TUPF + TCL1 + TCL2 + TRAM

TempoTotal = $1.84 \text{ s} + 5.88 \text{ s} + 0.00544 \text{ s} + 0.00498 \text{ s} + 0.0982$

TempoTotal = 7.82 s

TempoTotal (Antes das modificações) = $13,71 \text{ s}$

Ganho de desempenho = $13.71 / 7.82 = 1.75$

CompRef:

Cache L1: $1.780 \times 10^{-9} \text{ s} \times 0.7 = 1.24 \times 10^{-9} \text{ s}$

TUPF = $1600000 \text{ MI} / 63230 \text{ MFLOPS} = 25,30 \text{ s} \times 0.5 = 12,65 \text{ s}$

TempoTotal = TULA + TUPF + TCL1 + TCL2 + TRAM

Tempo Total = $3.92 \text{ s} + 12.65 \text{ s} + 0.00892 \text{ s} + 0.00615 \text{ s} + 0.053$

Tempo Total = 16.63 s

Tempo Total (Antes das modificações) = $29,28 \text{ s}$

Ganho de desempenho = $29.28 / 16.63 = 1.76$

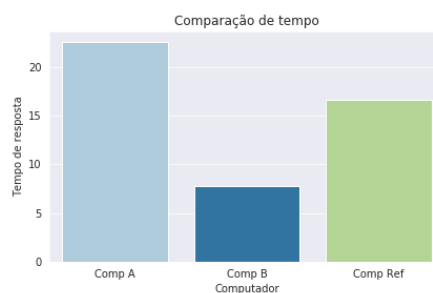


Figura 4. Medidas de tempo obtidas após alterações no programa BABE

Por fim, é possível notar que o computador B continua apresentando o melhor desempenho, isto é explicado pelo fato das mudanças terem sido aplicadas a todos os computadores, portanto, como o CompB já tinha o melhor tempo, ele continuou a demonstrar uma performance superior.

5. Conclusão

A partir das medidas e análises realizadas neste trabalho, é possível concluir que o computador B apresentou um desempenho superior em relação aos demais, no quesito de tempo de execução do programa BABE, e isso ocorre devido às características de seu hardware como um conjunto. Entretanto, ao considerar um outro grupo de aplicações, como processamento gráfico e compressão de dados, o computador referência se destaca entre todos. Essa conclusão é baseada na pontuação do benchmark fornecido pelo software Geekbench 5, logo que, para o cômputo desses valores, ele leva em conta essas aplicações e muitas outras.

No âmbito dos programas de benchmark utilizados, o Geekbench 5 tem uma grande vantagem em ter uma interface mais simples e convidativa, contudo, o conjunto de

informações que ele fornece é limitado. Outra desvantagem desta ferramenta, é que ela não está disponível para todos os sistemas operacionais, como o Windows 7, por exemplo.

Já a outra ferramenta, o Sisoftware Sandra, oferece uma coleção mais completa de dados, desde testes de benchmark, à informações sobre o hardware e software da máquina. Todavia, ela também é restrita a alguns sistemas e, além disso, possui um menu muito complexo, fator que dificulta no momento de sua utilização. Em relação a cada uma das medidas de benchmark dos programas, o Geekbench dispõe de uma pontuação com uma unidade própria, criada a partir de uma seleção de testes. Em contrapartida, dentre todas as possibilidades presentes no Sisoftware, é possível encontrar os dados em medidas mais comuns, como o MIPS, por exemplo.

Após o processo de pesquisa e produção deste trabalho, torna-se evidente que não é tarefa trivial determinar quais são as melhores métricas para medir o desempenho de um computador, logo que, dependendo da aplicação planejada, algumas métricas fornecem informações mais precisas que outras.

6. Referências

Interpreting Geekbench 5 Scores. Disponível em:
<http://support.primatelabs.com/kb/geekbench/interpreting-geekbench-5-scores>. Acesso em: 14 de set. de 2020

- Explicação da pontuação do GeekBench

<https://www.geekbench.com/GeekBench>. Acesso em: 12 de set. de 2020

- GeekBench - Benchmark a ser utilizado para descobrir o score

[https://www.sisoftware.co.uk/SiSoft Sandra](https://www.sisoftware.co.uk/SiSoft%20Sandra). Acesso em: 12 de set. de 2020

- SiSoft Sandra - Benchmark a ser utilizado para descobrir o MIPS, MFLOPS e latência da cache