

DRAM

Kelly Sandim, Lucas Queiroz, Vitória Freitas

¹Faculdade de Computação – Universidade Federal de Mato Grosso do Sul
(FACOM–UFMS) Campo Grande, Brasil.

1. Introdução

A memória RAM - ou memória principal - é um tipo de tecnologia que permite o acesso aos arquivos armazenados no computador. Diferentemente da memória do HD, a RAM não armazena conteúdos permanentemente. É responsável, no entanto, pela leitura dos conteúdos quando requeridos. Ou seja, de forma não-sequencial, por isso, a nomenclatura em inglês de *Random Access Memory* (Memória de Acesso Aleatório). [Monteiro]

A memória principal oferece tempos de acesso brutalmente mais baixos que o HD e trabalha com taxas de transferência muito mais altas. [Morimoto b] Dito isso, a quantidade de RAM disponível tem um grande efeito sobre o desempenho, já que sem uma quantidade suficiente dela o sistema passa a usar memória virtual, que é lenta. [Morimoto a]

O avanço da tecnologia nesta área é surpreendente. A pouco tempo foram descobertos novos materiais para produzir o núcleo de armazenamento das memórias. É muito provável que logo estejam no mercado memórias RAM com 10 GB, ou mais. Isto é necessário, já que os programas e aplicativos para computador não param de evoluir também. A cada dia são lançados programas e jogos que usam mais da capacidade do computador. [Pacievitch]

O presente trabalho foi realizado como parte da disciplina de Tópicos em Arquitetura de Computadores. Ele se propõe a avaliar as configurações de memórias principais, comparando diferentes parâmetros, como latência, largura de banda e potência média. Mais especificamente, o trabalho dedica-se a analisar tais dados e, baseando-se neles, escolher um modelo de memória RAM e, a partir dele, avaliar o comportamento da potência, da largura de banda e da latência conforme quantidade de Ranks e de Bancos, que são variados em diferentes configurações.

Este trabalho está organizado conforme segue: A Seção 2 descreve a metodologias adotada nos experimentos e as configurações básicas da máquina utilizada nos testes; Na Seção 3 são mostrados os experimentos, seus resultados e análise; As conclusões finais são apresentadas na Seção 4.

2. Metodologia

Nesse experimento foram analisados três modelos de memórias principais com diferentes características referentes a canais, bancos e ranks. Esses experimentos foram executados com a ferramenta DRAMSim2, um simulador de memórias (DRAM, SDRAM, DDR) que está acoplado ao MPSoCBench. Os experimentos foram divididos em duas partes.

Os três modelos de memórias definidos neste trabalho estão presentes em arquivos de configuração disponíveis em (MPSoCBench/ip/tlm memory/ini), são eles: *DDR2_micron_16M_8b_x8_sg3E.ini*, *DDR3_micron_16M_8B_x8_sg15.ini* e

DDR3_micron_32M_8B_x8_sg15.ini. O nome de cada arquivo indica o padrão de DDR, capacidade, número de quantidade de bancos, dentre outros.

Para a realização dos primeiros testes, foram feitas modificações nos três arquivos dos módulos de memória do MPSoCBench, e também no arquivo System.ini, presente na pasta tlm_memory do MPSoCBench, alterando-se a quantidade de bancos. Nesta primeira parte dos testes, foram avaliados cada um dos modelos de memória, com a quantidade de oito bancos, todos em um ambiente single core, em um conjunto de 4 aplicações, retornando valores relativos a Potência, Largura de Banda e Latência.

Após a obtenção dos valores, iniciou-se segunda parte dos testes, onde foram variados os valores de ranks e bancos em (2, 4 e 8) ranks e (2, 4 e 8) bancos, ambos como parte de um canal. Estes testes foram realizados apenas no modelo *DDR2_micron_16M_8b_x8_sg3E.ini*, onde o objetivo foi avaliar os impactos positivos e negativos destas alterações no projeto de memória.

Os experimentos foram feitos para as aplicações BasicMath, Lu, SHA e Water, pertencentes ao benchmark MPSoCBench, com apenas um processador utilizado na execução. A máquina na qual os experimentos foram realizados possui as seguintes configurações básicas:

- **8 GB de memória RAM DDR3**, 798.1MHz ;
- **1 TB de HDD**;
- Sistema Operacional: **Ubuntu 16.04 LST**;
- Arquitetura: **AMD64**;
- **Processador Intel Core i5 - 5200U** com tecnologia *Turbo Boost 2.0*, que acelera o desempenho do processador, fazendo com que os núcleos do processador trabalhem mais rapidamente do que a frequência operacional nominal quando estiverem operando abaixo dos limites especificados para energia, corrente e temperatura [Intel b], fazendo com que a frequência varie de 2.2 GHz até 2.7 GHz. Este processador possui cache de 3 MB, 2 cores e 4 *threads* [Intel a].

3. Resultados

3.1. Análise dos modelos de memória RAM

O primeiro experimento realizado comparou os modelos *DDR2_micron_16M_8b_x8_sg3E*, *DDR3_micron_16M_8B_x8_sg15* e *DDR3_micron_32M_8B_x8_sg15*, (nas aplicações BasicMath, Lu, SHA e Water), quanto à potência média, à largura de banda e à latência usadas na execução.

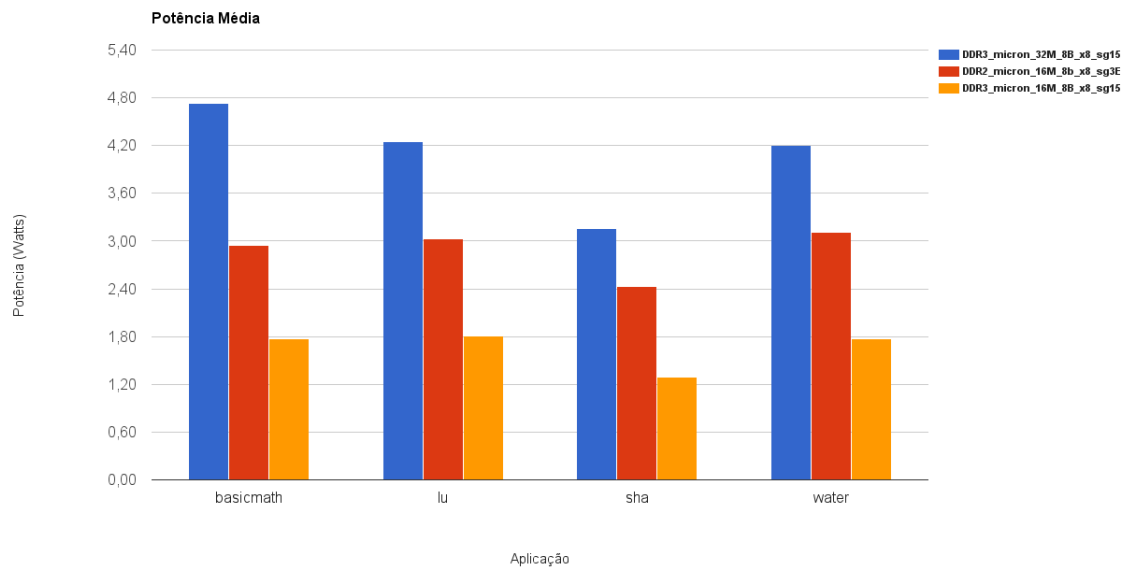


Figura 1. Potência média obtida para a RAM de cada modelo

Para a potência, foi observado, como apresentado no gráfico 1, o mesmo comportamento entre os modelos para as quatro aplicações: DDR3_micron_32M_8B_x8_sg15 necessita de maior potência média, DDR3_micron_16M_8B_x8_sg15 é o que precisa da menor quantidade de potência, e DDR2_micron_16M_8b_x8_sg3E foi intermediário entre os dois. Provavelmente, a grande potência de DDR3_micron_32M_8B_x8_sg15 se deveu à capacidade de memória que esse modelo descreve, 32M. Já entre os outros dois, que descrevem a mesma capacidade, é sabido que o padrão DDR3 é mais novo e projetado para ser mais econômico que o DDR2 [Jordão], por isso sua latência mais baixa é também plausível.

O simulador MPSoCBench fornece valores de largura de banda e latência para cada Banco. Os gráficos a seguir apresentam os valores obtidos para um Rank inteiro (calculados através da média aritmética entre todos os seus Bancos).

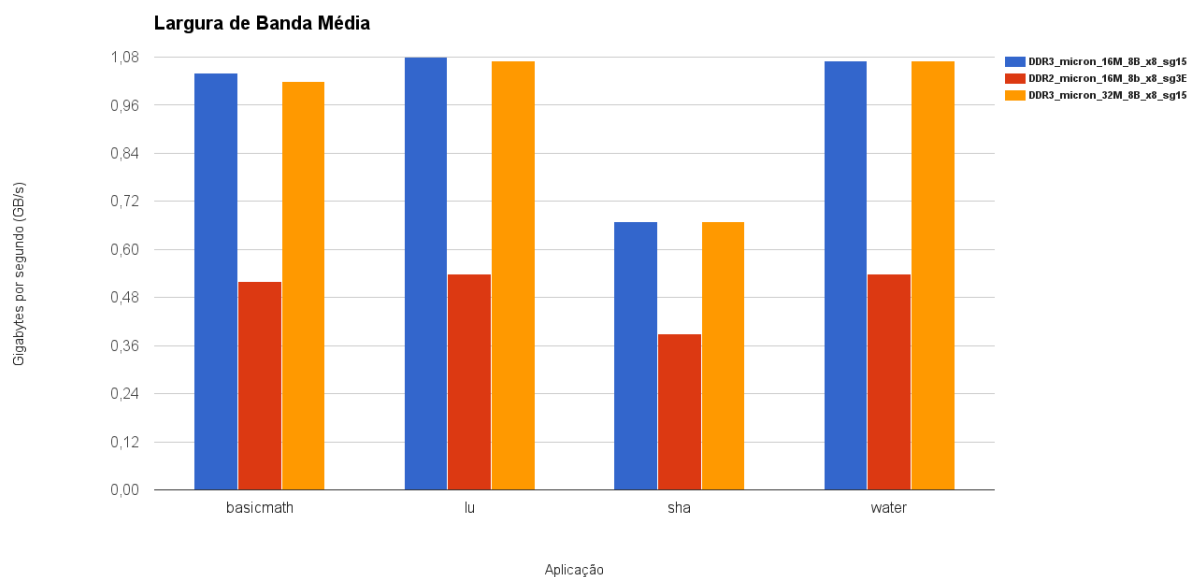


Figura 2. Largura de banda média obtida para a RAM de cada modelo

Quanto à largura de banda média, os modelos DDR3 obtiveram praticamente os mesmos resultados, os quais foram visivelmente mais altos que os do DDR2. O gráfico 2 indica, portanto, que o padrão DDR3 exige quase o dobro de largura de banda que o DDR2.

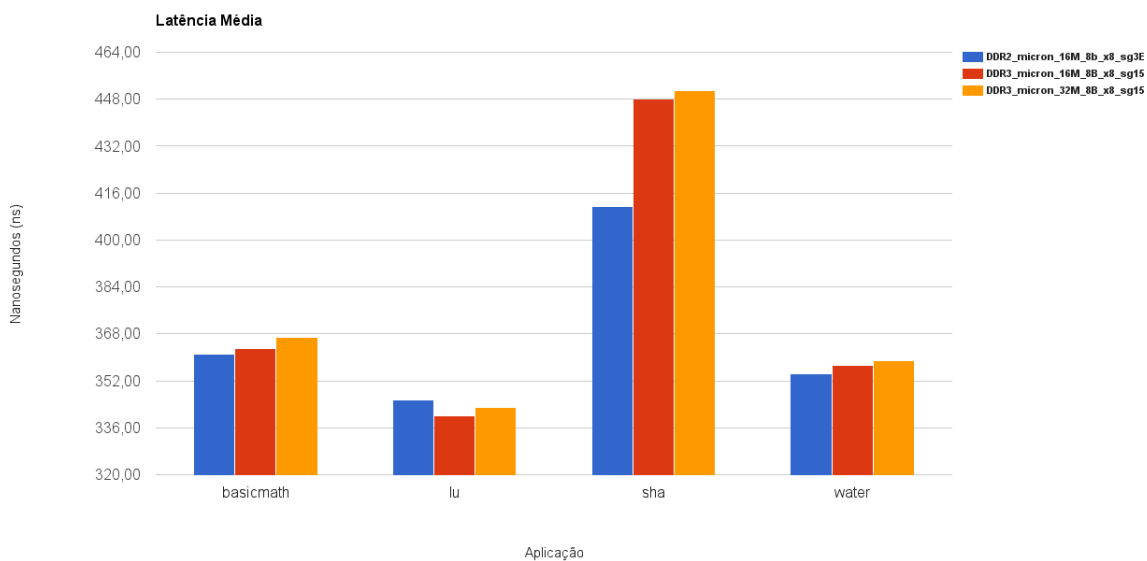


Figura 3. Latência média obtida para a RAM de cada modelo

Enquanto, no gráfico 2 quase não houve, para cada modelo, variação dos resultados entre as aplicações, o gráfico 3 mostrou que a latência é muito mais sensível à aplicação que ao modelo. Para cada aplicação, a variação de latência média entre os modelos foi pouca, ainda que se note que os DDR3 tenham, em geral, maior latência que o DDR2, comportamento apenas alterado na aplicação Lu. Comparando as aplicações, percebe-se que a Lu possui a menor latência média, enquanto BasicMath e Water têm latências muito próximas, além de um pouco maiores, e SHA atinge até 100ns a mais que estas, destacando-se no gráfico.

Para visualização da largura de banda e da potência discriminada por Banco, foi escolhida a aplicação LU:

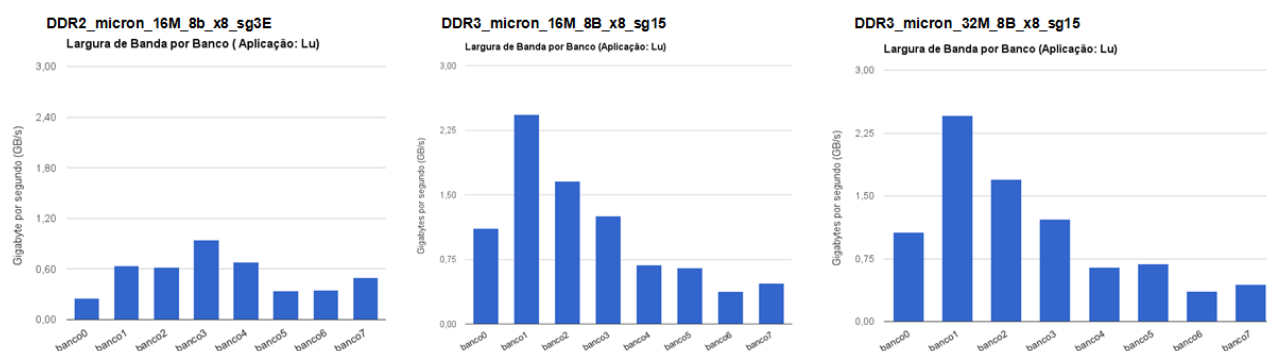


Figura 4. Largura de banda em cada Banco na aplicação LU, para os modelos DDR2_micron.16M.8b.x8.sg3E (esquerda), DDR3_micron.16M.8B.x8.sg15 (centro) e DDR3_micron.32M.8B.x8.sg15 (direita)

O gráfico 4 reitera a diferença entre o DDR2 e os DDR3 para a aplicação Lu, o primeiro tem menores larguras de banda que os seguintes. Da mesma forma, é também reiterado que os DDR3 exigem largura de banda maior e quase idêntica entre si, como pode ser visto neste gráfico, onde cada bloco nos dois modelos tem praticamente a mesma largura de banda.

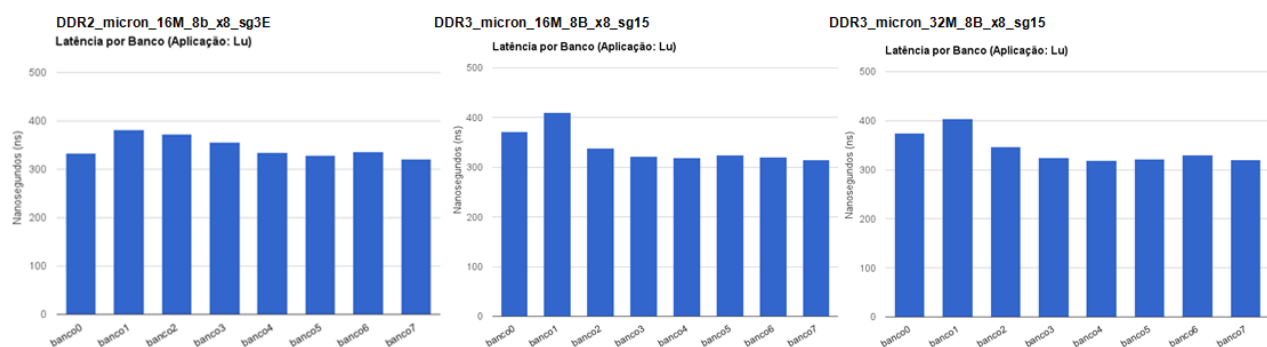


Figura 5. Latência na aplicação LU, para os modelos DDR2_micron.16M.8b.x8.sg3E (esquerda), DDR3_micron.16M.8B.x8.sg15 (centro) e DDR3_micron.32M.8B.x8.sg15 (direita)

Como já foi previamente afirmado, a latência depende mais da aplicação que do modelo. Portanto, já era esperado que, ao fixar a aplicação (no caso, Lu) a diferença de

latência, para cada banco da memória RAM, entre os modelos não seria grande, como é mostrado no gráfico 5.

3.2. Testes com o modelo DDR2_micron_16M_8b_x8_sg3E

A segunda fase de experimentos consistiu em realizar uma série de testes sobre o modelo de memória DDR2_micron_16M_8b_x8_sg3E, variando, em um Canal, a quantidade de Ranks e de Bancos por Rank. O simulador MPSoCBench fornece os valores de potência média por Rank, porém, no presente trabalho, foi decidido analisar a potência média por canal (obtida através da média aritmética entre as potências dos Ranks do canal). O gráfico a seguir apresenta, para as 6 configurações testadas, as potências médias resultantes:

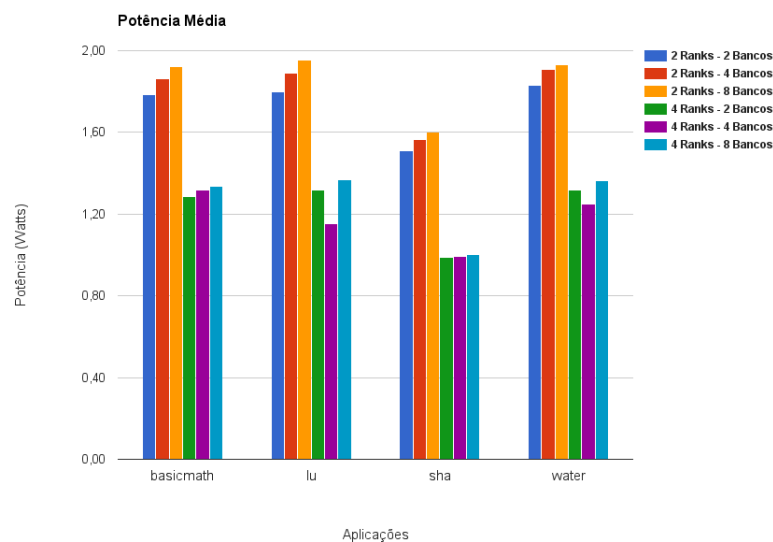


Figura 6. Potência média para o modelo DDR2_micron_16M_8b_x8_sg3E variando a quantidade de Ranks e Bancos

No gráfico 6, pode-se observar uma queda brusca da potência na passagem de 2 para 4 Ranks. Como os dados do gráfico resultam de contas de média aritmética entre os Ranks, essa observação significa que, para 4 Ranks, a Potência fica melhor distribuída entre os Ranks do que para 2. Fixando a quantidade de Ranks em 2, o aumento do número de Bancos causou ligeiro aumento na potência para todas as aplicações. Fixando em 4 Ranks, porém, o comportamento diante da variação de Bancos é menos uniforme. Na BasicMath e na SHA, a potência aumentou com o aumento do tamanho da RAM, mas na Lu e na Water a potência média diminuiu na configuração com 4 Bancos. Dessa diferença de comportamento pode ser concluído que o fator determinante na quantidade de potência média é a quantidade de Ranks.

Para a análise de largura de banda e latência, novamente foi escolhido representar nos gráficos, para cada configuração, o resultado da média aritmética por canal, ou seja, média calculada entre Ranks e Bancos.

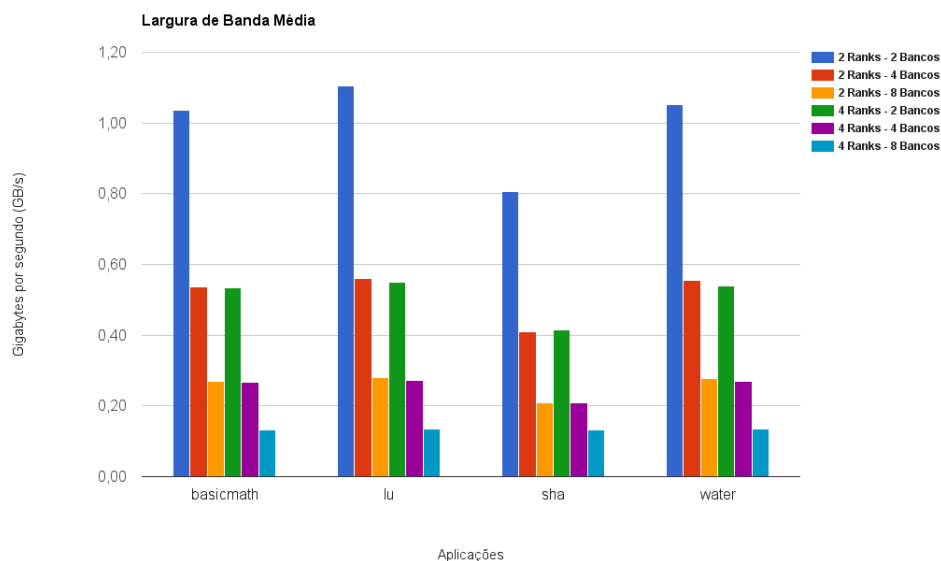


Figura 7. Largura de banda para o modelo DDR2_micron_16M_8b_x8_sg3E variando a quantidade de Ranks e Bancos

Em 7, percebe-se que o aumento da RAM causou diminuição na largura de banda média. Esse efeito é também consequência da média aritmética entre Bancos e entre Ranks, novamente indicando que, quanto maior a quantidade desses elementos, melhor é a distribuição da largura de banda entre eles. É notável também que as colunas de 2 Ranks com 4 Bancos e 4 Ranks com 2 Bancos (em vermelho e verde no gráfico), assim como as de 2 Ranks com 8 Bancos e 4 Ranks com 4 Bancos (em laranja roxo no gráfico), ficaram praticamente iguais entre si. Esse era um comportamento esperado, já que tais configurações de RAM possuem, entre si, o mesmo tamanho total.

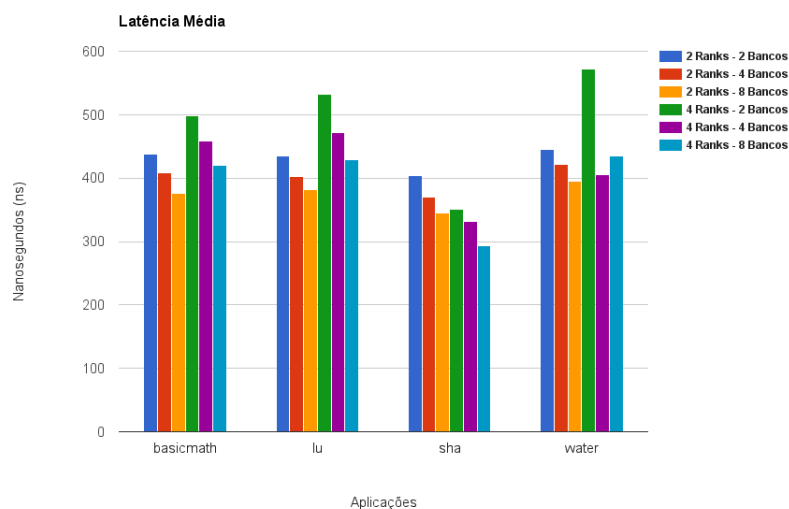


Figura 8. Latência para o modelo DDR2_micron_16M_8b_x8_sg3E variando a quantidade de Ranks e Bancos

Na latência, o comportamento das configurações entre as aplicações é menos uniforme, reforçando a ideia anteriormente constatada de que, em questão de latência, a aplicação é o fator preponderante. Assim como na análise sobre a largura de banda no gráfico anterior, fixando-se 2 Ranks e aumentando a quantidade de Bancos, a latência média diminuiu. Ao testar para 4 ranks, porém, nas aplicações BasicMath, Lu e Water, a latência aumentou significativamente, e diminuiu novamente ao aumentar o número de Bancos, com excessão da Water, cuja potência, para 8 Bancos, voltou a aumentar. Na SHA, por outro lado, o aumento progressivo do tamanho da memória RAM causou a diminuição da sua latência. Certamente essas diferenças são consequência das particularidades de cada aplicação.

4. Conclusão

A análise e comparação dos modelos DDR2_micron_16M_8b_x8_sg3E, DDR3_micron_16M_8B_x8_sg15 e DDR3_micron_32M_8B_x8_sg15 revelou que, a princípio, o modelo DDR3 é mais econômico, assim utilizando menos potência, que o DDR2. Porém, uma maior capacidade da RAM também exige mais potência. Portanto, em questão de potência média, padrão e capacidade devem ser considerados. Também pôde ser constatado que modelos DDR3 exigem praticamente o dobro da largura de banda do DDR2, já a quantidade de latência depende da aplicação.

Fixando-se o modelo percebeu-se que, quanto mais elementos na memória RAM (Ranks e Bancos), melhor é a distribuição de potência e largura de banda entre eles, especialmente para a potência ao aumentar a quantidade de Ranks. Entretanto, essas constatações não indicam que a quantidade absoluta de potência média e largura de banda é menor conforme o tamanho da memória aumenta. Quanto à latência, mesmo aumentando a quantidade de Ranks e Bancos, o fator determinante em seu tamanho continua sendo o perfil da aplicação executada.

Dessa forma, o presente trabalho levou às conclusões de que, numa memória RAM, sua potência e largura de banda são definidas por seu modelo, ambas podendo ser melhor distribuídas com a inclusão de Ranks e Bancos. A latência, por outro lado, independe de modelo ou capacidade da memória, sendo determinada de acordo com as características das aplicações a serem rodadas em tal memória.

Referências

- Intel. Intel® core™ i5-5200u processor (3m cache, up to 2.70 ghz). Disponível em: <http://ark.intel.com/pt-br/products/85212/Intel-Core-i5-5200U-Processor-3M-Cache-up-to-270-GHz> . Acesso em: 08 Jul. 2016.
- Intel. Tecnologia intel® turbo boost 2.0. Disponível em: <http://www.intel.com.br/content/www/br/pt/architecture-and-technology/turbo-boost/turbo-boost-technology.html> . Acesso em: 08 Jul. 2016.
- Jordão, F. O que é memória ddr? Disponível em: <http://www.tecmundo.com.br/hardware/1775-o-que-e-ddr-.htm> .Acesso em: 28 Jul. 2016.
- Monteiro, J. O que é memória ram e qual é sua função? Disponível em: <http://www.techtudo.com.br/artigos/noticia/2012/02/o-que-e-memoria-ram-e-qual-sua-funcao.html> .Acesso em: 28 Jul. 2016.
- Morimoto, C. E. Memória - hardware, o guia definitivo. Disponível em: <http://www.hardware.com.br/livros/hardware/memoria.html> .Acesso em: 28 Jul. 2016.
- Morimoto, C. E. Memória ram - definição de memória ram. Disponível em: <http://www.hardware.com.br/termos/memoria-ram> .Acesso em: 28 Jul. 2016.
- Pacievitch, Y. Memória ram. Disponível em: <http://www.infoescola.com/informatica/memoria-ram/> .Acesso em: 28 Jul. 2016.