

# Análise da Eficiência Energética de Operações de E/S em Arquiteturas de Baixa Potência \*

Pablo J. Pavan<sup>1</sup>, Ricardo K. Lorenzoni<sup>1</sup>, Jean L. Bez<sup>2</sup>,  
Edson L. Padoin<sup>1,2</sup>, Francieli Z. Boito<sup>3</sup>, Philippe O. A. Navaux<sup>2</sup>, Jean-François Méhaut<sup>4</sup>

<sup>1</sup> Universidade Reg. do Noroeste do Estado do Rio G. do Sul (UNIJUI) – Ijuí, RS – Brazil

<sup>2</sup> Universidade Federal do Rio Grande do Sul (UFRGS) – Porto Alegre, RS – Brazil

<sup>3</sup> Universidade Federal de Santa Catarina (UFSC) – Florianópolis - SC – Brazil

<sup>4</sup> Universidade de Grenoble – Grenoble – France

{pablo.pavan, ricardo.lorenzoni, padoin}@unijui.edu.br,  
francieli.boito@posgrad.ufsc.br, {jean.bez, navaux}@inf.ufrgs.br, jean-francois.mehaut@imag.fr

## Resumo

*This work presents a performance and energy efficiency analysis of I/O operations when using ARM processors with SSD and HDD storage. We aim to evaluate the viability of using such low-power architectures as file systems servers in HPC environments. To address this question we used the IOzone benchmark to analyze the execution time and total energy consumption. Preliminary results point that ARM can achieve energy savings of up to 136 times when compared to conventional architectures.*

## 1. Introdução

Em decorrência do aumento significativo da demanda de potência dos sistemas computacionais, inúmeras pesquisas vem sendo desenvolvidas pela comunidade de computação de alto desempenho em busca de alternativas que reduzam demanda de potência dos sistemas. Nesta premissa, pesquisadores sugerem no relatório DARPA um limite de 20MW para os futuros sistemas exascale [4].

Para atender este limite pesquisas têm sido realizadas com diferentes alternativas. Uma delas é o uso de processadores *Advanced RISC Machine* (ARM). Estes, apesar de possuírem menos desempenho, propiciam uma

melhor eficiência energética para algumas aplicações científicas [9].

Os processadores representam um percentual significativo na demanda de potência dos sistemas de HPC. Da mesma forma, sistemas de arquivos também apresentam impacto na demanda de potência. Nesses sistemas, as operações de E/S geralmente são feitas em sistemas de arquivos paralelos. Entretanto todo o poder de processamento do sistema geralmente não é utilizado durante operações de E/S. Nesses cenários, o uso de arquiteturas de baixa demanda de potência, como servidores de armazenamento, poderia apresentar uma melhor eficiência energética [10].

A fim de avaliar a viabilidade dessa ideia, o presente artigo apresenta um estudo comparativo de eficiência energética entre um computador com processador tradicional e um *Multi-Processor System-on-Chip* (MPSoC) com um processador ARM. Essa avaliação inclui discos rígidos (HDDs) e dispositivos de estado sólido (SSDs) para o armazenamento.

O restante do trabalho está assim organizado. A Seção 2 discute os trabalhos relacionados Na Seção 3 descreve o método experimental e os ambientes utilizados nos testes. Resultados são discutidos na Seção 4 seguidos das conclusões e trabalhos futuros.

## 2. Trabalhos Relacionados

Diversos trabalhos da literatura focam no consumo energético de operações de E/S. Tais operações representam grande parte do tempo de execução de muitas aplicações. Pesquisas menos recentes exploravam a

\* Trabalho parcialmente apoiado por CNPq, CAPES, FAPERGS. Possui recursos do projeto Europeu EU H2020 Programme e do MCTI/RNP-Brasil no Projeto HPC4E, sob número 689772. É realizada no contexto do Laboratório Internacional Associado entre a UFRGS e Université de Grenoble-LICIA.

utilização de discos de múltiplas velocidades para servidores de armazenamento [1, 3, 13]. Outras, mais atuais, empregam *Dynamic Voltage and Frequency Scaling* (DVFS) para reduzir o consumo do processador durante as operações de E/S.

Neste contexto, Ge *et al.* [2] propõem uma estratégia para arquiteturas de HPC que envolve a realização de DVFS nos nós de processamento durante as suas operações de E/S. Esta abordagem leva em consideração as características dos acessos das aplicações para decidir a frequência mais apropriada. Uma técnica semelhante havia, anteriormente, sido aplicada para aplicações sequenciais [11].

O consumo de energia é um dos principais problemas para o desenvolvimento da próxima geração de supercomputadores. Pesquisas buscam avaliar o consumo de energia de processadores de baixo consumo. McKenney *et al.* [5] atingiu um ganho de eficiência energética de 10% em sistemas Mobile utilizando Cortex-A15.

Ou *et al.* [7] comparou ARM e clusters baseados em Intel Nehalem para web services, e concluiu que ARM proporciona na media uma eficiência energética de 1.3 vezes melhor e é capaz de ter uma performance melhor do que o Intel em alguns cenários.

Por outra lado, Nijim *et al.* [6] combinam dispositivos de armazenamento baseados em flash (SSDs) com discos HDD para prover armazenamento com menor consumo energético. Isso é alcançado utilizando os dispositivos mais rápidos como uma cache para os discos rígidos. Essa estratégia híbrida (SSD+HDD) de armazenamento é explorada em diversos trabalhos para prover alto desempenho para servidores de E/S [12]. Nesses casos, o SSD é utilizado como uma cache por causa do seu alto custo por byte, que inviabilizaria a substituição total dos discos rígidos.

Desta forma, independentemente das técnicas discutidas para redução do consumo, em muitas aplicações, o poder de processamento as vezes subutilizado durante os períodos de E/S. Nestes casos, o uso de processadores ARM poderia ser uma alternativa para redução da demanda de potência nos sistemas de armazenamento.

Apesar da viabilidade do uso dessas arquiteturas para a computação de alto desempenho ter sido objeto de trabalhos, não foram encontradas pesquisas que investiguem o consumo energético de operações de E/S em arquiteturas de baixa potência.

### 3. Método experimental

Dois ambientes foram utilizados para esse trabalho. O primeiro, chamado de PC, é um computador tradicional com um processador Intel Core2Duo modelo E8400. Esse processador é da arquitetura *Wolfdale* e possui pipeline de 14 estágios com execução de até 4 instruções por ciclo.

O equipamento possui 6 GB de RAM com frequência de 800 MHz.

O segundo ambiente é um MPSoC CubieTruck com um SoC A20 fabricado pela AllWinnerTech e uma dual GPU MALI400 MP2, chamado de MPSoC. O processador é um Dual Core ARM Cortex-A7. Esse processador possui uma arquitetura superescalar, duas unidades de execução parcial, *pipeline* com 8 estágios e execução em ordem. Ele é baseado na arquitetura ARMv7-A e permite escalabilidade e controle sobre o consumo de energia, uma vez que possibilita o desligamento de qualquer um dos cores quando estiverem ociosos. O equipamento possui 2 GB de RAM com frequência de 480 MHz.

O sistema operacional instalado em ambos equipamentos é GNU/Linux. No PC utilizou-se Ubuntu com kernel 3.16.0 – 38 e no MPSoC Debian com kernel 3.4.106. O sistema de arquivos utilizado para os experimentos foi o *ext4*. A Tabela 1 apresenta as principais características dos ambientes experimentais.

	PC	MPSoC
<b>Processador</b>	Intel Core2Duo	ARM Cortex A7
<b>Modelo Processador</b>	E8400	AllWinnerTech SoC A20
<b>Técnica de Fabricação (nm)</b>	45	40
<b>Frequência de Clock</b>	3.0GHz	960MHz
<b>Número de processadores</b>	1	1
<b>Cores/Processador (#)</b>	2	2
<b>TDP do processador (W)</b>	65	0,25
<b>Cache L1/Core (KB)</b>	64 x 2	64
<b>Cache L2/Core (KB)</b>	6144	1024
<b>Memória (GB)</b>	6 (DDR2)	2 (LP DDR3)

**Tabela 1:** Configuração dos Ambientes

Também foi utilizado quatro dispositivos de armazenamento, dois SSDs e dois HDDs, a fim de cobrir diferentes características. Eles são apresentados na Tabela 2. Os nomes apresentados na primeira coluna da tabela, serão utilizados no restante do texto para referenciá-los. Em todos os experimentos foi utilizada a interface SATA II, suportada pelas duas arquiteturas e por todos os dispositivos. Foram realizados testes com os quatro dispositivos de armazenamento nos dois equipamentos, totalizando oito configurações.

	Tipo	Fabricante	Capacidade (GB)	RPM	Especificações Fabricante	
					Tensão (VDC)	Corrente (A)
<b>HDD1</b>	HDD	Western Digital	160	5400	5	0,55
<b>HDD2</b>	HDD	Seagate	500	7200	5	0,45
<b>SSD1</b>	SSD	Samsung	240	-	5	0,50
<b>SSD2</b>	SSD	Kingston	120	-	5	1,00

**Tabela 2:** Dispositivos de Armazenamento Utilizados

O benchmark utilizado para os testes foi o IOzone<sup>1</sup>, escolhido por ser amplamente aplicado e por permitir a descrição de diversos padrões de acesso. Os experimentos foram realizados em cada uma das configurações *com* e *sem* o uso da buffer cache. Foram testados os seguintes padrões de acesso: escrita sequencial; escrita randômica; leitura sequencial; e leitura randômica.

Outro parâmetro avaliado nos testes foi o tamanho das requisições. Neste caso, foram utilizados 32 KB ou 4 MB. Portanto, em cada configuração, foram realizados 16 testes, totalizando 128 experimentos, sendo que cada experimento foi repetido 10 vezes para o cálculo da média. Para medição de potência foi utilizado o equipamento Dranetz PP-4300, que mensura tensão e corrente alternada (CA) de todo o equipamento [8].

#### 4. Resultados

Nesta seção, serão apresentados os resultados obtidos durante a execução do benchmark, bem como da medição da potência instantânea consumida para a realização dos testes.

A Tabela 3 demonstra a potência média dos equipamentos durante os testes do benchmark.

Equipamento	Requisição	Cache	HDD1	HDD2	SSD1	SSD2
PC	32KB	off	60,071	56,240	59,247	59,987
PC	32KB	on	59,214	57,205	64,752	69,503
PC	4096KB	off	58,153	57,084	58,306	57,962
PC	4096KB	on	60,537	59,301	66,432	65,500
MPSoC	32KB	off	15,269	12,883	8,473	10,265
MPSoC	32KB	on	16,443	12,979	9,233	11,235
MPSoC	4096KB	off	16,931	13,659	8,661	10,954
MPSoC	4096KB	on	17,405	14,479	9,257	11,268

Tabela 3: Potência Média (W) dos Equipamentos

Estes valores foram obtidos através da média aritmética das medições de potência instantânea providas pelo equipamento de medição durante cada teste. A potência é analisada, uma vez que ela não depende da duração dos testes, ou seja, independe do tempo dos testes realizados pelo benchmark.

Os dados de tempo foram alcançados de modo que o benchmark acessava 2 GB no dispositivo de armazenamento teste, estes acessos aconteciam com vários padrões, já discutidos na Seção 3, assim a Tabela 4 traz uma média aritmética do tempo de execução de todos os padrões. O tempo de execução do benchmark entre os dispositivos difere consideravelmente entre os HDDs em relação aos SSDs. No PC as diferenças de tempo de execução são mais

Equipamento	Requisição	Cache	HDD1	HDD2	SSD1	SSD2
PC	32KB	off	253,768	237,647	221,071	16,694
PC	32KB	on	17,449	18,037	16,500	3,432
PC	4096KB	off	36,756	28,541	8,306	8,743
PC	4096KB	on	13,400	8,752	4,125	4,018
MPSoC	32KB	off	275,091	252,950	44,589	58,651
MPSoC	32KB	on	101,812	104,275	33,622	34,881
MPSoC	4096KB	off	51,316	43,863	34,142	36,666
MPSoC	4096KB	on	37,004	36,043	32,124	32,006

Tabela 4: Tempo Médio (s) de Execução do Benchmark

evidentes, apresentando reduções de tempo entre 2 e 20 vezes. Já no MPSoC, estas diferenças ficam menos evidentes, apresentando tempos de execução muito semelhantes na maioria dos testes.

A Figura 1 apresenta valores obtidos quando os dispositivos estavam ociosos e executando o benchmark. A partir destes dados, notou-se que ao acessar os HDDs e SSDs no PC, a demanda de potência aumenta 3,7% e 6,4% respectivamente. No MPSoC, a potência aumenta 2,6% e 68,4% ao acessar HDDs e SSDs. Comparado ao PC, o MPSoC apresenta demanda de potência 74,2% menor com HDDs e 83,7% com SSDs.

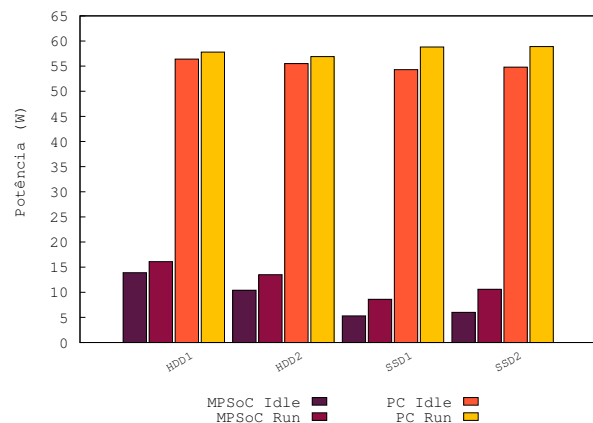


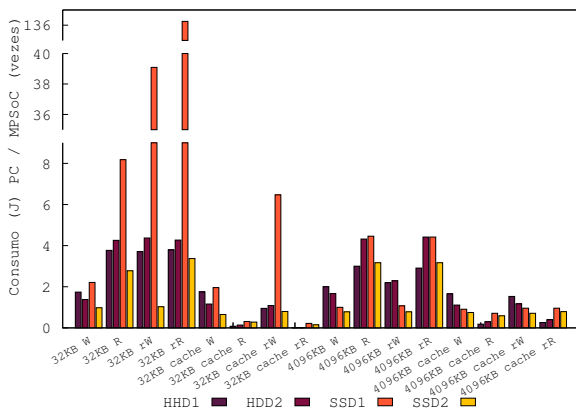
Figura 1: Potência Média em *idle* e Durante os Testes.

A energia consumida (J) de cada experimento foi calculada pela multiplicação da potência média mensurada (W) pelo tempo de execução (s), como mostra a Equação 1.

$$E = P_{avg} * t \quad (1)$$

Percebeu-se que o padrão de acesso não interfere na potência do equipamento, porém o padrão de acesso resulta em tempos de execução diferentes. Desta forma, almejando relacionar a energia consumida dos sistemas, dividiu-se o consumo do PC pelo consumo do MPSoC. Esta relação é apresentada na Figura 2.

<sup>1</sup> Disponível em <http://www.iozone.org/>



**Figura 2:** Relação da Energia Consumida pelos Sistemas

Foi observado que, utilizando discos rígidos para armazenamento, o uso do MPSoc leva a uma eficiência energética até 4,4 vezes maior do que o observado usando o PC. Para leitura utilizando SSDs, o MPSoc leva a uma eficiência energética até 136 vezes maior. Essa diferença acontece porque, apesar do maior tempo de execução medido nesse ambiente, a demanda de potência é até 6,7 vezes menor.

## 5. Conclusões e Trabalhos Futuros

Esse artigo apresentou uma análise de desempenho e eficiência energética de uma arquitetura de baixa demanda de potência - um MPSoc usando um processador ARM - para operações de E/S. Esses resultados foram comparados com um computador tradicional a fim de estudar a viabilidade do uso de arquiteturas de baixa potência como servidores de armazenamento. Esse estudo incluiu diferentes padrões de acesso e dispositivos de armazenamento a fim de cobrir diferentes situações e características.

Os resultados mostraram que a demanda de potência não é afetada pelo padrão de acesso. No entanto, este possui um impacto no tempo de execução, o que afeta o consumo de energia.

Concluiu-se que a substituição de um servidor de configuração PC + HDD por múltiplos servidores de baixa potência com SSDs seria viável e manteria um desempenho semelhante. Dependendo da carga de trabalho esperada, essa substituição diminuiria a demanda de potência, e consequentemente o consumo de energia, em até 85%.

Como trabalho futuro pretende-se expandir a análise apresentada a outros modelos de equipamentos de dispositivos de armazenamento. Além disso, serão avaliados ambientes distribuídos em que múltiplos servidores de baixa potência oferecem armazenamento.

## Referências

- [1] E. V. Carrera, E. Pinheiro, and R. Bianchini. Conserving disk energy in network servers. *Proceedings of the 17th annual international conference on Supercomputing - ICS '03*, page 86, 2003.
- [2] R. Ge, X. Feng, and X. H. Sun. SERA-IO: Integrating energy consciousness into parallel I/O middleware. In *Proceedings - 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, CCGrid 2012*, pages 204–211, 2012.
- [3] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke. DRPM: dynamic speed control for power management in server class disks. *30th Annual International Symposium on Computer Architecture, 2003. Proceedings.*, 2003.
- [4] P. Kogge, K. Bergman, S. Borkar, D. Campbell, W. Carson, W. Dally, M. Denneau, P. Franzon, W. Harrod, K. Hill, et al. Exascale computing study: Technology challenges in achieving exascale systems. pages 1–297, 2008.
- [5] P. E. McKenney, D. Eggeman, and R. Randhawa. Improving energy efficiency on asymmetric multiprocessing systems. Technical report, 2013.
- [6] M. Nijim, A. Manzanares, X. Ruan, and X. Qin. Hybud: An energy-efficient architecture for hybrid parallel disk systems. *Proceedings - International Conference on Computer Communications and Networks, ICCCN*, 0845257(2005), 2009.
- [7] Z. Ou, B. Pang, Y. Deng, J. Nurminen, A. Yla-Jaaski, and P. Hui. Energy- and Cost-Efficiency Analysis of ARM-based Clusters. In *12th IEEE/ACM Cluster, Cloud and Grid Computing (CCGrid)*, Ottawa, Canada, 2012.
- [8] E. L. Padoin, L. L. Pilla, F. Z. Boito, R. V. Kassick, P. Velho, and P. O. A. Navaux. Evaluating application performance and energy consumption on hybrid CPU+GPU architecture. *Cluster Computing*, 16(3):511–525, 2013.
- [9] E. L. Padoin, L. L. Pilla, M. Castro, F. Z. Boito, P. O. A. Navaux, and J.-F. Mehaut. Performance/energy trade-off in scientific computing: The case of ARM big.LITTLE and Intel Sandy Bridge. *IET Computers & Digital Techniques*, 2(3):1–14, 2014.
- [10] P. J. Pavan, R. K. Lorenzoni, J. L. Bez, E. L. Padoin, F. Z. Boito, P. O. A. Navaux, and J.-F. Méhaut. Análise da eficiência energética de operações de e/s com arquiteturas de baixa potência - (artigo submetido). In *XVII Simpósio de Sistemas Computacionais (WSCAD-SSC)*, pages 1–8, Aracaju, SE, 2016.
- [11] P. Shang and J. Wang. A novel power management for CMP systems in data-intensive environment. *Proceedings - 25th IEEE International Parallel and Distributed Processing Symposium, IPDPS 2011*, pages 92–103, 2011.
- [12] B. Welch and G. Noer. Optimizing a hybrid SSD/HDD HPC storage system based on file size distributions. In *2013 IEEE 29th Symposium on Mass Storage Systems and Technologies (MSST)*, pages 1–12, 2013.
- [13] Q. Zhu, Z. Chen, L. Tan, Y. Zhou, K. Keeton, and J. Wilkes. Hibernator: helping disk arrays sleep through the winter. *Proceedings of the twentieth ACM symposium on Operating systems principles - SOSP '05*, 39:177, 2005.