

# ***Cluster Beowulf* de baixo custo para o estudo de Sistemas Distribuídos**

**Naiara B. David, Yan A. S. Almeida, Juvenal S. Neto, Dárley D. de Almeida**

Departamento de Computação  
Universidade do Estado de Mato Grosso (UNEMAT) – Alto Araguaia, MT – Brasil

{naiara.borges@unemat.br, yansouzasilva@hotmail.com,  
juvenalneto@unemat.br, darleydalmeida@gmail.com}

**Abstract.** *This paper aims to present some features of the assembly and functional tests of a Linux-based Beowulf Cluster built from recycled computers. The research analyzes the performance of some parallel codes executed in a practical study of parallel and distributed computing. Early tests show that using the Cluster to perform high computationally costly tasks can provide considerable gains in runtime compared to performing the same task on a single machine.*

**Resumo.** *Este trabalho tem como objetivo apresentar algumas características da montagem e testes de funcionamento de um Cluster Beowulf baseado em Linux construído a partir de computadores reciclados. A pesquisa analisa o desempenho de alguns códigos paralelos executados em um estudo prático de computação paralela e distribuída. Os primeiros testes realizados demonstram que o uso do Cluster para realização de tarefas com alto custo computacional pode prover ganhos consideráveis em tempo de execução, em comparação com a execução da mesma tarefa por uma única máquina.*

## **1. Introdução**

De acordo com a Associação Brasileira da Indústria Elétrica e Eletrônica (ABINEE), no Brasil é produzida uma grande quantidade de Resíduos de Equipamentos Eletroeletrônicos (REEE), muitas vezes descartados sem o devido tratamento, provocando acúmulo de resíduos tóxicos como os metais pesados, chumbo, mercúrio e cádmio que oferecem risco para o meio ambiente e para a sociedade (ABINEE, 2019). Os REEE são considerados um novo tipo de resíduo que vem crescendo de maneira rápida ao redor do mundo devido à grande inovação tecnológica e a vida curta dos eletroeletrônicos. A gestão destes resíduos é um desafio tanto em países desenvolvidos quanto em desenvolvimento, devido a rápida troca dos equipamentos em função inclusive da obsolescência programada adotada por grandes empresas (FERREIRA, 2018; STEP, 2018; SUEMITSU; ARAÚJO, 2015).

A proposta deste trabalho originou-se com o Projeto de Pesquisa “Vantagens e Desvantagens na Reutilização e Reciclagem de Equipamentos Eletroeletrônicos em municípios com até 20 mil Habitantes” da Universidade do Estado de Mato Grosso (UNEMAT), Campus de Alto Araguaia (SANTOS; ANDRADE; OBANA, 2017), com o objetivo de reaproveitar computadores e suas peças, que foram descartados pela Universidade como lixo sem a intenção de reutilização. Através do processo de remanufatura simples, construiu-se um *Cluster Beowulf* de alto desempenho e de baixo custo no Laboratório de Hardware da UNEMAT. Um *Cluster* pode ser compreendido

como um sistema distribuído de computadores independentes e interligados, cujo o objetivo é suprir a necessidade de um grande poder computacional com um conjunto de computadores de forma transparente ao usuário (BACELLAR, 2009). O *Cluster* do tipo *Beowulf* foi desenvolvido pela NASA na década de 90, e é utilizado em várias empresas e universidades pelo fato de não precisar de um hardware específico, oferecendo facilidade na montagem e configuração, podendo alcançar um nível de processamento semelhante ao de supercomputadores, máquinas de elevado custo (DATTI; UMAR; GALADANCI, 2015).

O projeto de um *Cluster Beowulf* envolve diversos desafios, iniciando pela montagem e configuração física que exigiu um conhecimento mínimo e suficiente sobre hardware; a escolha da distribuição GNU/Linux Ubuntu; a instalação e configuração dos pacotes necessários; a configuração dos servidores, a elaboração do tutorial de instalação com o cuidado para fazê-lo o mais didático possível e por fim, o desafio a parte que foi descobrir o quanto o *Cluster* poderia ser eficiente nas tarefas de programação paralela.

## 2. Metodologia

O desenvolvimento deste trabalho foi dividido em dois momentos, iniciando pela classificação e montagem de quatro computadores Intel® Pentium® 4 CPU 2.8 GHz, 2 GB de memória principal e disco rígido de 40 GB, sendo um computador com a função de nó “mestre” e os demais como nós “escravos”. Adotou-se o uso do Sistema Operacional *Linux Ubuntu*, e todas as máquinas, por ser uma distribuição mais leve do *Ubuntu*, ser específico para trabalhar com redes e próprio para computadores mais antigos com baixo poder de processamento e pouca memória principal. Nas configurações de rede foram instalados o SSH (*Secure Shell*) responsável por fazer a comunicação entre os computadores por meio de protocolos seguros e o NFS (*Network File System*), com o objetivo de compartilhar arquivos e diretórios entre os computadores da rede. E para realização de processamento em paralelo adotou-se o MPI.

Na segunda etapa realizou-se testes de desempenho. A primeira bateria de testes consistiu na execução de código que exigia pouco poder de processamento sendo a execução de operações aritméticas sobre vetores do tipo *long int* de tamanho menor que mil posições; e, em seguida, o mesmo código com custo computacional mais elevado, pois considerava um vetor de mesmo tipo de dado com no máximo dez milhões de posições. Em todos os casos, o código foi executado primeiramente apenas no nó mestre, e em seguida no *Cluster*, considerando todos os nós do conjunto em operação.

## 3. Considerações Finais

Para a primeira bateria de testes, aqueles com menor custo computacional, observou-se que a execução em apenas um nó foi mais eficiente que a execução em paralelo, ou seja, pelo *Cluster*, o que pode ser justificado pelo fato de os processos de trocas de mensagens demandarem tempo razoável implicando no maior tempo final para realização da mesma tarefa. Porém, como esperado, para os testes com vetores de ordens de grandeza mais elevados, foi observado ganhos de 42% e 58% quando utilizado, respectivamente, dois e três nós escravos, em comparação com a execução do mesmo código apenas no nó mestre. Casos para os quais, o custo das trocas de

mensagens via MPI é pequeno frente ao ganho em tempo de processamento quando da execução em paralelo no *Cluster*.

Como trabalhos futuros pretende-se implantar um quarto nó ao *Cluster*, no intuito de aumentar o desempenho em trabalhos que envolvam um custo computacional mais elevado, utilizando códigos de complexidade maior. Ademais, serão executados outros testes de desempenho a fim de mensurar eficiência, *speedup* e avaliação da Lei de *Amdahl* para diferentes cenários, tanto para avaliação do *Cluster* montado, como para execução de práticas laboratoriais no âmbito da disciplina de Computação de Alto Desempenho.

Por meio do desenvolvimento deste trabalho foi possível constatar que o *Cluster Beowulf* é uma forma muito econômica de sistema de computação paralela, porque pode ser construído a um custo mínimo ou zero, especialmente em instituições como a UNEMAT, onde os computadores estão em uso há muito tempo. Também ajudará na reciclagem de seus computadores desktop usados.

Além disso, o trabalho de construção do *Cluster* de Baixo Custo está possibilitando o estudo de Computação Paralela e Distribuída para alunos de graduação do curso de Bacharelado em Ciência da Computação da UNEMAT de Alto Araguaia. As ações práticas, mais completas e significativas, em que os alunos experimentam fisicamente os conceitos e ideias, desenvolvendo habilidades, os prepararão para carreiras de projeto e gerenciamento de sistemas de computação de alto desempenho.

## Referências

ABINEE. Logística Reversa: governo deve enviar proposta à consulta pública. Disponível em: <<http://www.abinee.org.br/noticias/com40.htm>>. Acesso em: 5 jul. 2019.

BACELLAR, Hilário Viana. *Cluster: Computação de Alto Desempenho*. . UNICAMP Campinas, São Paulo: [s.n.], 2009.

DATTI, Ahmad A.; UMAR, Hadiza A.; GALADANCI, Jamil. A *Beowulf Cluster* for Teaching and Learning. *Procedia Computer Science*, v. 70, p. 62–68, 2015.

FERREIRA, Vivian Fernandes Marinho. Resíduos de Equipamentos Eletroeletrônicos: Um estudo do Campo de Ação Estratégica no Contexto Brasileiro. 2018. 202 f. Universidade de São Paulo, 2018. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/106/106132/tde-13112018-093618/en.php>>.

SANTOS, R.T.; ANDRADE, N.P.P.; OBANA, Fernando Yoiti. Reutilização, Reciclagem e utilização de equipamentos de informática. 2017, Cáceres: UNEMAT, 2017.

STEP. Step e-waste world map. Disponível em: <<http://step-info.org/step-e-waste-world-map.html>>. Acesso em: 5 jul. 2019.

SUEMITSU, Walter Issamu; ARAÚJO, Marcelo Guimarães. Gestão de Resíduos Eletroeletrônicos no Centro de Tecnologia da UFRJ. *Revista de Sustentabilidade e Tecnologia Ambiental*, p. 60–65, mar. 2015.