

**IMPLANTAÇÃO E AVALIAÇÃO DE UM *CLUSTER BEOWULF* PARA FINS
EDUCACIONAIS E DE PESQUISA: UM ESTUDO DE CASO APLICADO AO
IFSULDEMINAS-CÂMPUS MUZAMBINHO**

Marcelo FASSBINDER¹; Augusto JÚNIOR²

RESUMO

A área educacional, de forma geral, tem se beneficiado da evolução da internet e dos computadores. De forma específica, considerando o aspecto acadêmico e de pesquisa, o número de dados coletados, armazenados e capturados em projetos de pesquisa é cada vez maior. Entretanto, certas atividades exigem um processamento de alto desempenho, em especial, aquelas realizadas durante as disciplinas de Sistemas Distribuídos e Programação Paralela, bem como pesquisas científicas. Nesse contexto, considerando que super computadores são, normalmente, muito caros, e que instituições de ensino necessitam prover a alunos e docentes estratégias de baixo custo, este artigo teve como objetivo estudar os fundamentos de redes de computadores, programação paralela, a fim de implantar e avaliar um Cluster Beowulf tendo o IFSULDEMINAS – Câmpus Muzambinho como estudo de caso. Como resultados alcançados têm-se os passos principais utilizados nesta implantação, bem como as estratégias utilizadas para testar o cluster construído. Espera-se que tais resultados possam ser úteis à comunidade acadêmica do Câmpus Muzambinho.

Palavras-chave: Cluster Beowulf, Programação Paralela, Alto Desempenho.

¹ Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Câmpus Muzambinho. Muzambinho/MG - E-mail: marcello.fassbinder@gmail.com

² Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Câmpus Muzambinho. Muzambinho/MG - E-mail: augustomsj@gmail.com

INTRODUÇÃO

A evolução dos computadores e da internet vem beneficiando diversas áreas, entre elas, a pesquisa científica, que faz uso das tecnologias da informação e comunicação (TICs) para facilmente coletar e armazenar dados importantes, a fim de resolver problemas.

Essa facilidade faz com que o conjunto de dados relacionados com tais pesquisas seja cada vez maior. Para computar esses dados, extrair informações relevantes ou realizar simulações, de forma rápida e eficiente, os computadores pessoais não são adequados, pois sua capacidade de processamento é limitada. Segundo Fischbom (2006), necessita-se, portanto, de computadores potentes ou super computadores, também denominados Plataformas de Computação de Alto Desempenho.

Ainda de acordo com Fischbom(2006), essa definição envolve não apenas computadores, mas também a tecnologia de rede, os algoritmos e as ferramentas necessárias para permitir a utilização dessas plataformas, que podem variar desde *clusters* de computadores pessoais até grandes supercomputadores vetoriais.

Os supercomputadores vetoriais, no entanto, possuem um custo extremamente elevado e demandam profissionais capacitados para mantê-los funcionando corretamente. Um exemplo é o supercomputador adquirido pela USP³, em 2012, para criação do novo laboratório de astroinformática, estimado em torno US\$ 1 milhão, além do valor para instalação.

A outra opção, conforme já citado anteriormente, é utilizar os *clusters* de computadores pessoais. Segundo Conti (2008), trata-se de um tipo de sistema para processamento paralelo e distribuído, ou seja, é uma coleção de computadores interconectados que trabalham juntos, mas como se fossem um único recurso computacional integrado. Com esse aglomerado de computadores é possível realizar processamentos que até então somente computadores de alto desempenho seriam capazes de realizar.

Dentre os vários tipos de *cluster* existentes, o denominado *Cluster Beowulf* utiliza *hardware* de baixo custo, sendo implementado com computadores pessoais, inclusive com configurações distintas. Além disso, todos os *softwares* utilizados são

³<http://agencia.fapesp.br/15091>

baseados na licença *Open Source*, caracterizando-se, desta forma, uma plataforma de computação de alto desempenho, mas com foco principal no baixo custo de construção e implantação.

O *Cluster Beowulf* vem sendo utilizado em diversas áreas, como alternativa para super computadores. Segundo Brown (2006), seu campo de aplicações é imenso, podendo ser aplicado para processar dados relacionados à meteorologia, astrofísica, exploração de petróleo, otimização, pesquisa operacional, biotecnologia, inteligência artificial, servidores de internet, pesquisas científicas, jogos e até renderização de efeitos visuais para o cinema.

Na área de educação, as instituições de ensino superior (IES), que são um ambiente propício para desenvolvimento de pesquisas e estudos, estão cada vez mais investindo na construção de *clusters* de baixo custo, conforme observado em Junior (2013), pois vislumbram neles a oportunidade de oferecer aos seus estudantes e pesquisadores meios para que eles possam realizar estudos e pesquisas que envolvam cálculos complexos alcançados apenas com máquinas de alto poder de processamento. Dessa forma, para Prado (2010), a utilização de *Cluster Beowulf* garante à instituição uma ferramenta viável e com a possibilidade de expansão de seus recursos sem a necessidade de grandes investimentos.

MATERIAL E MÉTODOS

Essa sessão, apresenta uma síntese das principais atividades realizadas para que os objetivos pretendidos com o desenvolvimento deste artigo, e estão categorizadas em i) análise e coleta das máquinas subutilizadas do campus Muzambinho para a construção do cluster proposto, ii) construção e configuração do Cluster Beowulf e iii) avaliação do cluster construído.

i) Análise e coleta das máquinas subutilizadas do Campus Muzambinho para construção do cluster proposto:

Esta etapa consistiu de uma reunião, com os responsáveis pelo Núcleo de Tecnologia da Informação (NTI) do câmpus, a fim de discutir e verificar quais computadores poderiam ser utilizados no desenvolvimento do Cluster. Após tal reunião, ficou definido que as máquinas a serem utilizadas seriam, inicialmente, aquelas disponíveis no Laboratório de Redes e, após a substituição de alguns computadores, os quais não suportam mais as tarefas executadas nos setores do câmpus em questão, esses também seriam disponibilizados para a construção do

cluster. Sendo assim, cinco computadores foram selecionados e suas configurações são descritas na Tabela 1.

Tabela 1: Configuração dos Computadores utilizados na construção do cluster

Processador	Memória Ram	HD
Intel core 2 Duo 2.80 GHz	2 GB	160 GB
Intel 2140 1.60 GHz	2 GB	80 GB
Intel core 2 Duo 2.53 GHz	2 GB	120 GB
Intel core 2 Duo 2.53 GHz	2 GB	80 GB
Intel 2140 1.60 GHz	2 GB	80 GB

ii) Construção e Configuração do Cluster Beowulf para o Campus Muzambinho:

Nessa etapa foi realizado a configuração do cluster Beowulf com a instalação do sistema Operacional (SO), configuração de rede, instalação de pacotes e edição de arquivos do linux. A configuração iniciou com a instalação do SO Debian na versão 6.0.10. Após a instalação do SO em todas as maquinas, as mesmas foram interligadas a uma rede local, utilizado cabo par trançado, categoria 5, conectados a um Switch fast Ethernet de 5 portas.

Após a configuração de rede, foi realizado a instalação de todos os pacotes necessarios para a configuração do cluster, os seguintes pacotes foram instalados: build-essential, lam-mpidoc, lam-runtime, lam4-dev, mpich-bin libmpich1.0-dev, libmpich1.0gf, rsh-client, rsh-server, nfs-kernel-server, nfs-common, libgmp3-dev, libgmp3-doc, libgmp3c2.

Após a instalação dos pacotes em todos os nós, foi realizado a edição de alguns arquivos do Linux e dessa forma terminada a configuração do cluster Beowulf.

iii) Avaliação do Cluster Construído:

Os testes de desempenho realizados no cluster consistiram na execução de algoritmos desenvolvidos na disciplina de programação paralela na turma do 7º período do curso de Ciência da Computação do Câmpus Muzambinho. Os alunos deveriam escolher um problema, e desenvolver um algoritmo sequencial e outro em paralelo, que resolvessem o problema escolhido. A avaliação foi realizada com 12 alunos e um total de 13 códigos em paralelo. O teste de desempenho se deu executando os algoritmos e verificando o tempo de execução em diferentes números de nós do Cluster Beowulf.

Outra forma utilizada para avaliar o cluster consistiu na utilização da ferramenta de renderização de imagem Persistence of Vision Ray Tracer (PovRay). Essa ferramenta permite o processamento de imagens vetoriais criadas a partir de um código fonte. Os testes realizados com a ferramenta consistiram em renderizar uma imagem em apenas um nó do Cluster Beowulf e repetir o teste com a mesma imagem, mais dessa vez aumentando o número de nós. Após os testes, foi comparado o tempo, e verificado o ganho de desempenho.

RESULTADOS E DISCUSSÃO

Considerando as principais atividades realizadas durante este trabalho e descritas anteriormente, de forma geral têm-se como resultados o Cluster Beowulf implantado no laboratório de redes do câmpus Muzambinho.



Figura 1: Cluster Beowulf Implementado no IFSULDEMINAS Câmpus Muzambinho

E à análise dos dados coletados com a execução dos algoritmos implementados pelos alunos da disciplina de Programação Paralela, onde percebeu-se um grande desempenho dos algoritmos desenvolvidos de forma paralela em comparação aos sequenciais. Dos treze algoritmos testados, dez demonstraram um ganho de desempenho a medida que se foi aumentando o número de nós durante a execução. Três algoritmos apresentaram um melhor desempenho com dois nós, e a medida que foi aumentando o número de nós o tempo de execução do algoritmo foi aumentando, isso ocorre devido ao tipo de problema e ao modo de como o problema foi paralelizado pelo aluno.

Sequencial X Paralelo

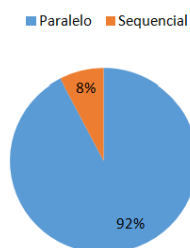


Figura 2: Comparação de desempenho entre os algoritmos desenvolvidos de forma sequencial e em paralelo na disciplina de Programação Paralela.

Fonte: Dados de pesquisa (2015)

Outra forma utilizada para testar o cluster Beowulf, foi com a ferramenta PovRay, para isso foi feito a renderização da imagem skyvase.pov, a imagem foi renderizada com uma resolução de 640x480 e o resultado dos intervalos de tempo se encontram na tabela 2.

Tabela 2: Resultado dos tempos obtidos na renderização da imagem skyvase.pov no Cluster Beowulf

Número de nós	Tempo
1	46.24 seg
2	42.12 seg
3	35.54 seg
4	21.23 seg
5	13.07 seg

Os resultados da análise demonstram o potencial do Cluster Beowulf no processamento de imagem, podendo ser utilizado para auxiliar na renderização de imagens que necessitem um alto poder de processamento.

CONCLUSÕES

Este artigo teve como objetivo geral implantar e testar um Cluster do tipo Beowulf para ser utilizado por alunos e docentes do câmpus Muzambinho. De forma mais específica, espera-se que os alunos da Disciplina de Programação paralela possam utilizá-lo como um recurso a fim de executar, testar e validar algoritmos desenvolvidos durante a referida disciplina.

Adicionalmente, também espera-se que projetos de pesquisa que envolvam computação de alto desempenho possam obter rapidez no processamento de dados, tendo este cluster como apoio.

Como trabalhos futuros pretende-se aumentar o número de nós do cluster, disponibilizar ao nó mestre o acesso a internet e fornecer uma forma de acesso remoto ao cluster, de modo que, de qualquer laboratório do campus, os alunos, professores, pesquisadores e demais interessados possam executar suas aplicações, sem a necessidade de se locomoverem até o laboratório de redes.

REFERÊNCIAS

- BROWN, R. G. (2006) What makes a Cluster Beowulf?. **Duke University Physics Department**. Disponível em <www.beowulf.org/overview/howto.html>. Acesso em: 31 Ago. 2014.
- CONTI, F. (2008) Grades computacionais para processamento de alto desempenho. **Universidade de Santa Maria**. Disponível em <<https://svn.inf.ufsc.br/luis.custodio/TCCDantas/Artigos/artigo3.pdf>>. Acessado em 30Ago. 2014.
- FISCHBOM, M. (2006) Computação de alto desempenho aplicada à análise de dispositivos eletromagnéticos. **Universidade de Santa Catarina**. Disponível em <<https://repositorio.ufsc.br/handle/123456789/88235>>. Acesso em 30 ago. 2014.
- JUNIOR, E. E. (2013) Estudo de Tecnologias para Computação Paralela e Distribuída: Implementação de um ClusterBeowulf. **Dissertação de Mestrado, UFF**. Disponível em <http://www.uff.br/mcct/downloads/Dissertacao_8.pdf>. Acesso em 30 ago. 2014.
- PRADO, C. L., & da Silva, J. M. A. (2010) Aplicação de ClusterBeowulf em Instituições de Ensino. Monografia. **Faculdade Tecnológica de Guaratinguetá**. Disponível em <http://portallivre.org/blog_fatecti/tccs/tcc_joao_claudio_final.pdf>. Acessado em 23 Ago. 2014.