

Construção de um Cluster Computacional baseado no modelo Beowulf a partir de Máquinas obsoletas

Leonardo S. Mendes¹, Carla Castanho¹, Eduardo Ferreira¹, Victor Machado¹,
Eduardo Spies¹

¹Departamento de Engenharias e Ciência da Computação – Universidade Regional Integrada do Alto Uruguai e das Missões (URI) – Câmpus Santiago
Caixa Postal 97700-000 – Santiago – RS – Brasil

lsm.ccomp@gmail.com, carla.castanho, eduardo.ferreira, victor.alves,
eduardo.spies {@urisantiago.br}

Abstract. *This article aims to explain what is a Computational Cluster and to show its different types, besides presenting the present project that is realized in one of the laboratories of the University Uri Santiago, as well as it's progress and the configurations used to implement the cluster. Bringing as focus of mirroring in the Beowulf cluster.*

Resumo. *Este artigo visa explicar o que é um Cluster Computacional e mostrar os seus diferentes tipos, além de apresentar o presente projeto que se realiza em um dos laboratórios da universidade Uri Santiago, bem como seu andamento e as configurações utilizadas para implementação do cluster. Trazendo como foco de espelhamento o cluster Beowulf.*

1. Introdução

Com o intuito de ampliar o incentivo a pesquisa e análise de dados, uma iniciativa que vem crescendo em universidades é o projeto de *clusters* computacionais. Estes modelos computacionais favorecem tanto áreas que necessitam de poder computacional, como também o meio ambiente, por diminuir o lixo eletrônico e apresentarem uma alternativa de grande baixo custo ao reutilizarem os hardwares disponíveis.

Uma outra alternativa para áreas que requerem grande poder computacional são os supercomputadores, *mainframes* desenvolvidos com o propósito de resolver problemas complexos de forma eficiente e que contam com grande poder de processamento. Contudo, o alto custo destes supercomputadores muitas vezes se torna um empecilho para pequenas instituições, logo, o *cluster* acaba por se tornar a melhor saída. Construído a partir de computadores pessoais que seriam descartados, tem a finalidade do reaproveitamento de máquinas gerando um baixo custo de montagem.

Diante disso, a URI-Santiago, em seu laboratório, conta com um projeto de implementação de um *cluster* computacional utilizando o hardware disponível em sua universidade. As máquinas usadas no presente projeto foram retiradas de departamentos da universidade e adquiridas através de coleta de lixo eletrônico realizado na praça da cidade, em pareceria com a prefeitura de Santiago.

O projeto tem como objetivo a implementação de um *cluster* computacional baseado no modelo Beowulf, que é um modelo quem vem sendo amplamente utilizado, formado por máquinas de baixo poder computacional.

2. Clusters de Computadores

Cluster é o nome dado a um sistema distribuído que relaciona dois ou mais computadores para que estes trabalhem de maneira conjunta no intuito de processar uma determinada tarefa (ALECRIM, 2013). O sistema age de forma transparente, operando como se houvesse apenas um processador, seus nós são interligados por uma rede Ethernet capaz de transferir informações e controlar o transporte de dados.

No seu desenvolvimento há requisitos a serem cumpridos, ele precisa ter mais de dois nós (*back-end*), sendo que precisa ter ao menos um nó mestre (*front-end*) (BACELLAR, 2010).

A ideia inicial dos *clusters* de computadores foi desenvolvida na década de 60 pela IBM como uma alternativa para ligar *mainframes* e oferecer uma forma efetiva de paralelismo (BUYA, 1999). Mas foi em 1994 que este sistema distribuído foi desenvolvido a partir de computadores convencionais priorizando combater o alto custo dos supercomputadores. Esse modelo de *cluster* foi chamado de Beowulf.

O *cluster* Beowulf foi idealizado pelos pesquisadores da NASA, Thomas Sterling e Donald J. Becker, para ser voltado à computação paralela com a finalidade de processar as informações espaciais que a entidade recolhia (BACELLAR, 2010), e chegou a ocupar o ranking 5 dos supercomputadores mais poderosos do mundo (GROPP, 2003, tradução nossa).

3. Tipos de Clusters

A escalabilidade em um projeto de *cluster* é um fator importante a ser ressaltado, ela garante que a remoção de qualquer nó, não deve parar com o funcionamento do *cluster*, assim como a inserção de um novo nó. Aplicações que necessitam de um pleno funcionamento constante e exigem computação de alto desempenho, como ferramentas de mapeamento genético, simuladores geotérmicos, entre outros, podem ter o *cluster* como uma solução viável, desde que o tipo mais adequado seja escolhido (ALECRIM, 2013).

Para isso, se apresentam os seguintes tipos:

- Cluster de Alto Desempenho: Destinam-se a aplicações de maior exigência no que diz respeito ao processamento, são usados para analisar grandes quantidade de dados rapidamente, como também para realizar cálculos bastante complexos. Seu foco é fornecer resultados precisos em tempo mais hábil possível (PEREIRA, 2016 apud INFOWESTER, 2013).
- Cluster de Alta Disponibilidade: O principal foco aqui é manter a aplicação sempre em pleno funcionamento, é inaceitável que a aplicação pare de funcionar, mas caso isso ocorra, ela deve ser o menor possível, e para atender a esse propósito, existem recursos como: ferramentas de monitoramento que identificam nós defeituosos, uso de geradores para casos de queda elétrica, entre outros (PEREIRA, 2016 apud INFOWESTER, 2013).
- Cluster para Balanceamento de Carga: Como o próprio nome já sugere, o objetivo deste *cluster* é manter o equilíbrio entre os nós, e não necessariamente que ele divida uma tarefa com outras máquinas, é necessário que ele distribua as requisições de forma que seja executado garantindo o equilíbrio da aplicação. É

possível monitorar os nós constantemente para verificar se alguma máquina está lidando com uma menor quantidade de tarefas, e direcionar novas requisições para esta (PEREIRA, 2016 apud INFOWESTER, 2013).

É válido ressaltar que uma solução de *cluster* não precisa estar presa a somente um tipo, eles podem combinar características para atender as necessidades de uma aplicação da melhor forma possível (ALECRIM, 2013).

4. Metodologia

Partindo do princípio de levantamento bibliográfico, foram analisados quais as bibliotecas disponíveis – PVM, MPI, OpenMPI –, e o sistema operacional mais adequado – OpenSource –. Desta etapa em diante foram testados os computadores que apresentavam melhores condições e separados de acordo com sua disponibilidade de utilização, além de terem sido classificados de acordo com seu desempenho para uma melhor performance como parte do *cluster*.

Atualmente, o projeto de implementação do *cluster* computacional na universidade conta com cinco computadores funcionantes, um *switch* e mais três máquinas que aguardam restauração, sendo estes cinco computadores com processadores Intel Pentium 4 e Celeron, com sistema operacional Debian Stretch GNU/Linux 9.5 com kernel 4.9. A arquitetura do *cluster* conta com quatro nós escravos e um nó controlador sendo monitorados pela ferramenta Ganglia, que é um software livre de controle de performance para ambientes computacionais, além disso, ele trabalha de forma escalável e distribuída em *clusters* de alto desempenho.

Para acompanhar a temperatura dos núcleos do processador durante sua utilização nos diversos testes será utilizada a ferramenta Psensor, que é um monitor gráfico de temperatura de hardware para Linux. Para comunicação entre as máquinas, foi usado a biblioteca MPI (*Message Passing Interface*) que é um padrão para comunicação paralela de dados.

5. Resultados Parciais

O *cluster* mostrado na Figura 1 apresenta o atual modelo do *cluster* Beowulf sendo desenvolvido na universidade. Com cinco nós escravos, um nó controlador, e um *switch* fazendo a conexão entre as máquinas, o progresso do projeto para a construção de um *cluster* de alto desempenho com balanceamento de carga tem se mostrado satisfatório, seguindo como base a metodologia aplicada no desenvolvimento do trabalho.

Com isso, já visando um *benchmark* a ser instalado, o *cluster* possibilitará uma abertura para projetos interdisciplinares, como na resolução de cálculos complexos desempenhados pelas áreas de engenharias e matemática, e uma possível implementação de exemplos para aulas práticas, como uma forma de aliar a teoria com a prática. Embora, contudo, seguindo os critérios do *benchmark* a ser selecionado, para se obter resultados mais precisos, livres de erros, esta etapa pode se estender um pouco mais



Figura 1. Cluster Computacional

6. Conclusão

Clusters computacionais tem se tornado cada vez mais utilizados por instituições devido ao seu baixo custo de montagem e pelo fácil acesso aos recursos de montagem. Este artigo visa realçar isso usando como exemplo a universidade mencionada e indo direto para o modelo de *cluster* a estar sendo desenvolvido em um de seus laboratórios.

O modelo Beowulf como sendo um dos mais comumente usados, destaca-se pela sua fácil incrementação, e fácil configuração, além de ser um marco para a evolução dos supercomputadores no quesito custo x benefício.

O projeto, ainda que de forma sucinta, além de apresentar resultados já parciais, tem como o reaproveitamento das máquinas ainda sendo feito de forma gradual, o que tem levado a uma diminuição do lixo eletrônico gerado pelos computadores que antes se acumulavam no laboratório por não terem um destino.

Referências

- Alecrim, E. (2013) “Cluster: conceito e características”, <https://www.infowester.com/cluster.php>, setembro.
- Bacellar, H. Viana. (2010) “Cluster: Computação de Alto Desempenho”, <http://www.ic.unicamp.br/~ducatte/mo401/1s2010/T2/107077-t2.pdf>.
- Barros M. Almeida., Coelho P. Marques. and Teixeira, H. Gonçalves. (2000) “Clusters Beowulf”, http://www.cesarkallas.net/arquivos/apostilas/redes/beowulf_clusters.pdf.
- Buyya, R. (1999) “High Performance cluster computing: architectures and systems”, 1ª Edição, Upper Saddle River, NJ, USA: Prentice Hall. v. 1, 849p.
- Ferreira, Rubem E. (2008) “Linux Guia do Administrador do Sistema”, 2ª Edição.
- Gropp, W., Lusk E. and Sterling T. (2003) “Beowulf Cluster Computing with Linux, Second Edition”.
- Pereira, V. Pacheco. (2016) “Instalação e Configuração de um Cluster Beowulf”, <http://hdl.handle.net/10314/3697>.