Cluster Unifra: Melhorias e Expansão

Iuri Fiedoruk, Cristiano Berni, Gabriel Pedebos, Rogério C. Turchetti, Ana Paula Canal

Centro Universitário Franciscano
Rua dos Andradas, 1614, (55) 3220-1220
{iuri.unifra, cberni, gpedebos}@gmail.com; {turchetti, apc}@unifra.br

Introdução

O *cluster* do Centro Universitário Franciscano, Unifra, é um sistema composto por três máquinas *Pentium* utilizado para desenvolvimento de programas paralelos, através das bibliotecas *Message Parsing Interface - MPI* [MPI 06] e a Biblioteca de Programação Paralela - BPP [CIO 05], em aulas, projetos de pesquisa e trabalhos de conclusão de curso, não apenas para Computação, como em qualquer área que precise de alto poder de processamento.

Sendo do tipo *Beowulf*, este sistema *cluster* possui uma máquina que é responsável pela coordenação das outras e pelo arquivamento dos dados em disco, tornando possível, de maneira até mesmo simples, adicionar mais nós, como são chamados os componentes à exceção do coordenador, chamado mestre, para aumentar o poder de processamento ou mesmo para realização de testes em que seja útil variar o número de componentes do sistema para averiguação de velocidade ou existência de erros. Desta forma, buscou-se uma alternativa para poder incorporar mais nós ao *cluster* Unifra, sem que seja necessário tornar esta máquina exclusivamente de uso do sistema, a fim de permitir a utilização das estações de trabalho localizadas no mesmo laboratório em que o *cluster* Unifra encontra-se. No entanto, como a utilização das máquinas do laboratório é prioritária para os estudantes, o sistema deve ser feito de tal forma que uma máquina possa ser utilizada a qualquer hora seja como estação de trabalho, seja como nó do *cluster*; para tanto é preciso que a máquina possua em seu sistema de inicialização e desligamento um *script* que avise à máquina mestre que ela está entrando ou deixando o conjunto de nós.

Arquitetura Utilizada

Para o sistema operacional foi utilizado o *Ubuntu Linux*, pois além de ser gratuito, flexível, poder ser instalado em máquinas que já possuam diversos sistemas sem ter de alterá-los ou removê-los, é bastante simples de ser operado e pequeno, sendo preciso apenas um *CD-ROM* para a instalação. Uma particularidade do sistema, é que apenas as três máquinas que atualmente compõe o *cluster* estão ligadas ao mesmo *switch*, e este por sua vez está ligado a outro *switch* da sala, onde também estão ligadas as outras máquinas da sala. Como o Linux permite que uma mesma interface de rede tenha dois endereços IP, foi preciso dar aos nós do *cluster* também endereços da rede do *switch* da sala, para que as máquinas do laboratório possam ter acesso a todos os nós do sistema. Para fazer com que a máquina, ao ser ligada no sistema instalado, integre-se ao *cluster*, foi criado um sistema bastante simples, inspirado no *script lamboot* [LAM 06], parte das ferramentas LAM/MPI, uma das implementações existentes da Biblioteca MPI, e que faz o registro de um nó no computador mestre durante a inicialização através da troca de mensagens. Este

sistema faz com que, ao ser *inicializada*, uma máquina que irá integrar o *cluster*, envie à máquina mestre um comando pedindo para ser registrada, e ao ser desligada, envie um comando dizendo que deve ser removida. A máquina mestre é então encarregada de gerenciar um arquivo contendo a lista de nós, anteriormente estático, inserindo ou removendo o nome e o endereço IP da máquina.

O *cluster* Unifra está configurado para aceitar comandos, sem senha, através de *Remote Secure Shell, rsh*, de máquinas que estejam em sua sub-rede. Abrir a possibilidade de comandos *rsh* para todo um laboratório não é seguro, de forma que se decidiu por executar o comando via *Secure Shell, ssh,* com um usuário criado especialmente para esta tarefa e que tenha restrições de acesso direto e não possua um diretório *home*, de forma que acessos não autorizados através deste usuário não permitam alterações nos arquivos do mestre. Como autenticação utiliza-se um par de chaves no lugar de senha [SOF 06].

A chamada ao comando remoto que registra ou remove o nó no *cluser* é realizada pela criação de um arquivo de *script* chamado *cluster* no diretório */etc/init.d*, de forma que ele execute como se fosse um serviço (*daemon*) do sistema. Para habilitar o serviço no *Ubuntu Linux* é preciso executar o comando *update-rc.d cluster defaults.*

Conclusão e Trabalhos Futuros

O gerenciamento dinâmico de nós do *cluster* pode permitir, além da expansão da capacidade de processamento no número de nós no laboratório de Alta Performance, o desenvolvimento de um *Live-CD*, que possua um sistema já configurado como nós do *cluster*, assim não seria necessária nenhuma alteração nos dados dos discos das máquinas e a instalação do sistema Linux nas mesmas. Outra possibilidade interessante a ser explorada é permitir, em conjunto com um detector de falhas, descobrir quando um dos nós foi desligado de maneira errônea, se travou, está desconectado da rede ou simplesmente não está respondendo, de forma a retirar o nós da lista de componentes do sistema. Além do desenvolvimento de nós dinâmicos, está sendo introduzida tolerância a falhas na biblioteca *BPP* com a utilização de um programa detector de defeitos em cada nó do *cluster*.

Atualmente o *cluster* é utilizado para trabalhos de de final de curso, pesquisa e aulas de Programação Paralela e Distribuída, e uma mudança fácil do número de nós auxilia testes dos programas desenvolvidos nestes casos.

Referências

[MPI 06] **The Message Passing Interface (MPI) standard -**

http://www.mcs.anl.gov/mpi/ acessado em 17/10/2006.

[CIO 05] CIOCARI, Juliano Francisco Cagnini Ciocari (2005) D**esenvolvimento de uma Biblioteca para Programação Paralela - BPP**. ERAD - Escola Regional de Alto Desempenho 2006.

[LAM 06] **LAM/MPI Parallel Computing** - http://www.lam-mpi.org/ acessado em 01/10/2006

[SOF 06] **SourceForge.net: SSH Key Generation and Usage –** http://sourceforge.net/docman/display_doc.php?docid=761&group_id=1, acessado em 08/10/06