

OSCAR: Um Gerenciador de Agregado para Ambiente Operacional Linux

C. Rista, A.R. Pinto, M.A.R. Dantas

Universidade Federal de Santa Catarina – UFSC
Centro Tecnológico – CTC
Departamento de Informática e Estatística – INE
{rista, arpinto, mario}@inf.ufsc.br

Resumo

O cenário que surge com a maior utilização de agregado de computadores, como uma opção para a computação de alto desempenho, tem requerido uma demanda de ferramentas capazes de proporcionar uma melhor configuração e um gerenciamento mais eficiente desses ambientes. Neste artigo apresentamos as principais características do OSCAR (*Open Source Cluster Application Resource*). Este ambiente é uma ferramenta de software aberto e cuja utilização já atinge cerca de vinte e três por cento [LIG 03] das configurações de agregados da comunidade técnico-científica.

Introdução

Os recentes avanços nas tecnologias de microprocessadores e redes locais de alto desempenho têm criado a possibilidade de utilização de agregado de computadores como eficientes ambientes paralelos para execução de um grande número de aplicações. Essa abordagem baseia-se na idéia de utilização de diversos computadores comerciais facilmente encontrados no mercado (COTS – *Components Off The Shelf*) interligados por uma tecnologia de rede local, visando um aumento no desempenho das aplicações, através da exploração da distribuição ou paralelismo. Neste artigo apresentamos um estudo da ferramenta OSCAR (*Open Source Cluster Application Resource*) [OSC 03] para o gerenciamento de agregado de computadores visando um maior desempenho da distribuição/paralelismo.

Nas próximas seções apresentamos uma breve descrição sobre agregado de computadores e o gerenciador OSCAR.

Agregado de Computadores

No contexto tecnológico atual o paradigma de agregado de computadores representa uma forma popular de configuração para execução de aplicações que requeiram um computador paralelo. O termo agregado de computadores pode ser considerado como um conjunto de computadores interligados de tal forma, que os seus usuários tenham a convicção de estar usando um recurso computacional único. Uma boa parte dessa abstração é obtida através das camadas de software, dentro ou fora do

sistema operacional, que permitem essa transparência. Segundo [DAN 02], o paradigma de agregado visa um aumento do desempenho das aplicações, através de uma maior taxa de execução dos aplicativos e um aumento no número de dados a serem considerados na execução.

Em um ambiente de agregado de computadores, pode-se imaginar que para atingir o objetivo de melhoria de desempenho, primeiramente devemos considerar alguns esforços, tais como: aumento na velocidade do processador, uso de algoritmos mais otimizados e adoção de um ambiente de computação concorrente (ou paralela) [DAN 02].

Ambiente Experimental

Para a instalação do gerenciador OSCAR foi utilizado o sistema operacional Linux (Red Hat Linux 9). A escolha pela distribuição ocorreu após uma criteriosa seleção entre as demais distribuições suportadas pelo gerenciador OSCAR [OSC 03], [PER 04]. Os critérios que adotamos estão relacionados essencialmente com desempenho, robustez e confiabilidade.

O ambiente de execução foi o agregado de computadores do LabWeb (Laboratório de Web Software) localizado no INE (Departamento de Informática e Estatística) da UFSC. Este ambiente é caracterizado por cinco máquinas Intel Pentium 4 de 1,8 GHz, com 256 MB de memória principal, disco rígido com 40 GB e como rede de interconexão utilizamos um switch Fast Ethernet.

O Gerenciador OSCAR

Segundo [OSC 03], o OSCAR é um pacote de software que permite simplificar a complexa tarefa de utilização e gerenciamento de um agregado. O OSCAR é essencialmente utilizado para a computação de alto desempenho, podendo perfeitamente ser utilizado por qualquer aplicação, que necessite das funcionalidades de um agregado, para obter um aumento em seu desempenho através da exploração do paralelismo. Cabe ressaltar alguns pacotes padrões instalados nativamente, como implementações de MPI (Message Passing Interface) [MPI 94], de PVM (Parallel Virtual Machine) [PVM 02] e de PBS (Portable Batch System) [PBS 03].

Uma prática comum em agregados é o uso da biblioteca PBS, responsável por controlar a submissão e execução das tarefas no nodo servidor, combinada com alguma implementação MPI ou PVM para o ambiente paralelo. As bibliotecas de troca de mensagens como MPI e PVM fornecem a capacidade necessária ao ambiente OSCAR para a criação de programas paralelos em um ambiente de computação distribuída, como apresentado na figura 1.

Inicialmente, foram realizadas a instalação e configuração do sistema operacional no nodo servidor. Logo após foi iniciada a instalação do gerenciador de agregado OSCAR versão 2.3, como é apresentado na figura 2. Como não havia conhecimento da necessidade de instalação de pacotes extras, e por se tratarem de etapas opcionais, os três primeiros passos não foram explorados, mantendo desta forma as opções padrões do OSCAR.

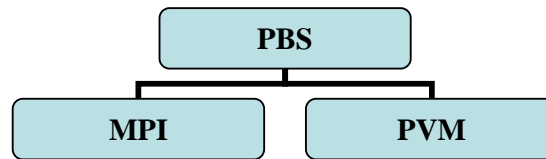


Figura 1 – Hierarquia das bibliotecas no OSCAR

Em seguida, é iniciada a instalação de vários pacotes RPMs adicionais requeridos, além de uma configuração auxiliar no nodo servidor. A fase seguinte consiste na criação da imagem para os nodos escravos. Neste momento, um determinado tempo é necessário para a construção da imagem, sendo que uma barra de progresso indica o status de instalação do processo.

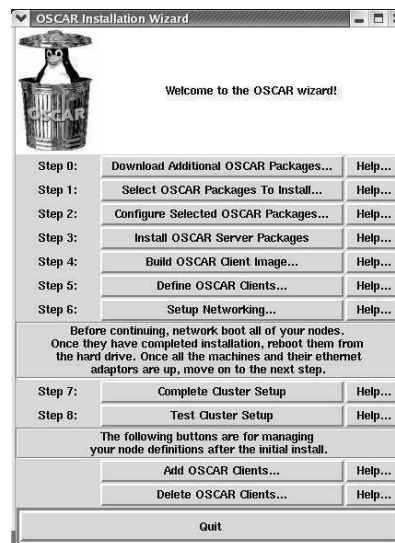


Figura 2 – Instalação do OSCAR

O próximo passo tem a finalidade de definir informações apropriadas para os nodos escravos do gerenciador OSCAR. Nesta fase, são configuradas informações como: número de hosts, atribuição de IP aos nodos escravos, máscara da subrede e gateway, por exemplo.

A seguir é realizada a detecção dos nodos escravos através do endereço MAC, identificado de forma única em uma rede, para posterior atribuição do endereço IP ao nodo escravo. Dois métodos possíveis podem ser utilizados para a inicialização pela rede. O primeiro utiliza uma entrada PXE (*Preboot eXecution Environment*), opção de inicialização da BIOS do nodo escravo, se disponível. Se esta opção não for possível, pode-se criar um disco de inicialização. Cabe lembrar ainda, que no final deste passo, antes de ser iniciado o passo seguinte, é necessário esperar algum tempo para que seja completada a instalação dos nodos escravos. O tempo necessário para a instalação, depende da capacidade do nodo servidor, dos nodos escravos, da rede local, e o número de instalações simultâneas.

As duas etapas que seguem completam a instalação do gerenciador OSCAR. Na fase seguinte, são executados apenas scripts de configurações complementares. E finalmente o ultimo passo completa de fato a instalação do OSCAR, através de testes e demais verificações no ambiente instalado, certificando-se desta maneira do seu perfeito funcionamento.

Conclusões

O estudo do gerenciador OSCAR apresentado neste artigo tem como objetivo prover facilidades para uma melhor configuração e um gerenciamento mais eficiente em ambientes de agregado de computadores. Foram apresentados conceitos básicos utilizados no OSCAR, além das principais funções para troca de mensagens e controle de execução de tarefas, nativas do OSCAR. É interessante comentar algumas dificuldades encontradas durante as fases de instalação e configuração, ocasionadas inicialmente devido a uma determinada complexidade imposta pela configuração do ambiente apresentado, pois não havia conhecimento prévio das reais exigências do gerenciador de agregado OSCAR. A partir do momento em que estes pré-requisitos foram definidos e superados, a instalação do pacote de software OSCAR ocorreu de forma satisfatória e transparente.

Referências

- [DAN 02] DANTAS, M.A.R. **Tecnologias de Redes de Comunicação e Computadores**. Rio de Janeiro: Editora Axcel Books do Brasil, ISBN 85-7323-169-6, 2002.
- [LIG 03] LIGNERIS, B. et al. **Open Source Cluster Application Resources (OSCAR): Design, Implementation and Interest for The Computer Scientific Community**. In. International Symposium on High Performance Computing Systems - HPCS, Canada, May 17, 2003.
- [MPI 94] MPI-FORUM, **MPI : A Message Passing Interface Standard**, Int. J. Supercomputing Applicatons, MPI FORUM 8(1994), 3-4.
- [OSC 03] OSCAR – **Open Source Cluster Application Resources**. Disponível em <http://oscar.sourceforge.net>, August 19, 2003.
- [PBS 03] PBS – **Portable Batch System (OpenPBS)**. Disponível em <http://www.openpbs.org>, 2003.
- [PER 04] PERNAS, A. M.; COLVERO, Taís Appel; DANTAS, M.A.R. **Facilidades para Gerenciamento de uma Configuração de Agregado**, Artigo submetido para o ERAD 2004, Pelotas (RS), 2004.
- [PVM 02] PVM – **Parallel Virtual Machine (PVM)**. Disponível em <http://www.csm.ornl.gov/pvm>, November 24, 2002.