

Projeto MultiCluster

Prof. Dr. Philippe Olivier Alexandre Navaux
Denise Stringhini, Elgio Schlemer, Fábio A. D. de Oliveira, Marcos E. Barreto,
Rafael B. Ávila, and Ricardo Cassali

Instituto de Informática — UFRGS
Cx. Postal 15064 — 91501-970 Porto Alegre
Fone: (51) 316-6165 Fax: (51) 319-1576
<http://www-gppd.inf.ufrgs.br/projects/mcluster>
E-mail: mcluster@inf.ufrgs.br

Este projeto¹ objetiva um ambiente de desenvolvimento de aplicações voltado à utilização de agregados de multiprocessadores (*clusters*). Um *cluster* é geralmente formado por um conjunto homogêneo de máquinas PC, cada uma com dois processadores, conectadas por uma rede de comunicação de alto desempenho como Myrinet [3] ou SCI (*Scalable Coherent Interface*) [5]. Além disso, os nodos de um *cluster* são também normalmente conectados por uma rede de comunicação, usada para o gerenciamento da arquitetura, disparo de aplicações, etc., baseada em tecnologia Ethernet (10 ou 100Mbps/s). Duas dessas máquinas paralelas estão disponíveis no Instituto de Informática, uma utilizando tecnologia Myrinet e outra com tecnologia SCI, ambas com quatro nodos.

A idéia principal do projeto [1] é permitir a interconexão de agregados baseados em tecnologias de comunicação diferentes (donde o nome *MultiCluster*), o que, atualmente, não é uma tarefa trivial, porém altamente desejável, dada a imensa popularização desse tipo de máquina paralela. Para tanto, desenvolve-se uma biblioteca de programação, DECK (*Distributed Execution and Communication Kernel*) [2], a qual provê funcionalidades essenciais para programação paralela e distribuída (*threads*, semáforos, caixas postais, comunicação coletiva, etc.). Nesse sentido, DECK pode ser comparado a bibliotecas como PVM [4] e MPI [7], porém oferecendo, além da comunicação, o recurso da multiprogramação. Adicionalmente, DECK é estruturado de forma a suportar a troca de dados através de diferentes tecnologias de comunicação, sendo portanto esta a base para a efetivação do modelo MultiCluster. Atualmente, encontram-se disponíveis implementações de DECK para tecnologias Myrinet e SCI.

Em um nível de abstração mais elevado, também é disponibilizada a linguagem DPC++ (*Distributed Processing in C++*) [9], uma extensão de C++ que facilita a programação de aplicações paralelas e distribuídas através do uso da orientação a objetos. Em DPC++, um objeto pode ser instanciado em um nodo remoto, fazendo com que vários objetos possam executar concorrentemente e, portanto, aumentar o desempenho de uma aplicação. Além disso, os métodos de um mesmo objeto podem ser executados de forma concorrente entre si, o que permite explorar o potencial das máquinas multiprocessadas. A grande vantagem de DPC++ sobre uma biblioteca paralela tradicional é que todos os mecanismos relaciona-

¹ Projeto parcialmente financiado com recursos da FAPERGS, FINEP, CAPES e CNPq.

dos à distribuição e ao paralelismo interno dos objetos são automaticamente introduzidos no código-fonte, por meio de um pré-processador, baseado nas extensões sintáticas utilizadas. Isso permite que mesmo usuários não familiarizados com os detalhes intrínsecos da programação concorrente possam tirar proveito de seus benefícios.

Adicionalmente, o projeto MultiCluster envolve pesquisas na área de depuração visual de aplicações paralelas, desenvolvendo o software PADI — *Parallel Debugger Interface* [10]. O trabalho consiste no desenvolvimento de uma interface gráfica completa para um ambiente de depuração paralela simbólica *on-line*. O principal objeto de pesquisa é o nível de coordenação dos processos distribuídos de uma aplicação paralela; a este nível, estão sendo desenvolvidos mecanismos de seleção de processos e visualização, ainda pouco explorados nas ferramentas existentes na atualidade. PADI já conta com um primeiro protótipo, desenvolvido com o objetivo de validar as idéias principais da ferramenta, que são a seleção e a visualização de processos.

É importante também destacar o envolvimento do projeto MultiCluster com grupos de pesquisa internacionais. Através de programas de cooperação internacional mantidos pelo CNPq (com o INRIA) e pela CAPES (com o COFECUB e o DAAD), o grupo tem trocado experiência com professores e pesquisadores da França, da Alemanha e de Portugal, o que tem contribuído decisivamente para a relevância do trabalho em desenvolvimento.

Atualmente, o grupo trabalha no refinamento do ambiente DECK e no desenvolvimento de aplicações, tendo sido implementadas versões paralelas do gerador de fractais de Mandelbrot [6] e da simulação de dissipação de calor através da equação de Laplace [8]. Essas aplicações vêm sendo usadas também para depuração da biblioteca DECK e avaliação do mecanismo de escalonamento de *DPC++*.

Referências

1. Marcos Barreto, Rafael Ávila, and Philippe Navaux. The MultiCluster model to the integrated use of multiple workstation clusters. In José Rolim et al., editors, *Proc. of the 3rd Workshop on Personal Computer based Networks of Workstations*, volume 1800 of *Lecture Notes in Computer Science*, pages 71–80, Cancun, 2000.
2. Marcos Ennes Barreto. DECK: Um ambiente para programação paralela em agregados de multiprocessadores. Master's thesis, PPGC/UFRGS, Porto Alegre, 2000.
3. N. Boden et al. Myrinet: A gigabit-per-second local-area network. *IEEE Micro*, 15(1):29–36, February 1995.
4. Al Geist et al. *PVM: Parallel Virtual Machine*. MIT Press, Cambridge, MA, 1994.
5. Hermann Hellwagner and Alexander Reinefeld, editors. *SCI: Scalable Coherent Interface: Architecture and Software for High-Performance Compute Clusters*, volume 1734 of *Lecture Notes in Computer Science*. Springer, Berlin, 1999.
6. B. B. Mandelbrot. *The Fractal Geometry of Nature*. W. E. Freeman and Company, New York, NY, 1982.
7. MPI Forum. The MPI message passing interface standard. Technical report, University of Tennessee, Knoxville, April 1994.
8. W. H. Press et al. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University, Melbourne, second edition, 1994.
9. André Silveira, Rafael Ávila, Marcos Barreto, and Philippe Navaux. *DPC++: Object-oriented programming applied to cluster computing*. In H. R. Arabnia, editor, *Proc. of the 2000 International Conference on Parallel and Distributed Processing Techniques and Applications*, volume 5, pages 2515–2521, Las Vegas, 2000. CSREA Press.
10. Denise Stringhini, Philippe Olivier Alexandre Navaux, and Jacques Chassin de Kergommeaux. A selection mechanism to group processes in a parallel debugger. In H. R. Arabnia, editor, *Proc. of the 2000 International Conference on Parallel and Distributed Processing Techniques and Applications*, Las Vegas, 2000. CSREA Press.