

Ambiente ShareD-GM: Uma Proposta de Integração de Sistemas DSM ao Ambiente D-GM

Gustavo M. Zechlinski¹, Renata Hax Sander Reiser¹, Adenauer Corrêa Yamin¹

¹Programa de Pós Graduação em Informática – Universidade Católica de Pelotas
Rua Félix da Cunha, 412 – 96.010-000 – Pelotas – RS – Brazil

{reiser,adenauer}@ucpel.tche.br, gustavo.mata@ufpel.edu.br

O recente avanço das tecnologias de computadores impulsionaram o uso de vários computadores de pequeno porte para execução de aplicações que exijam um grande esforço computacional, tornando esta prática uma forte tendência atual. Acompanhando esta tendência, o Ambiente D-GM (*Distributed-Geometric Machine*) é uma ferramenta, composta de módulos de software, cujo objetivo é usar *clusters* de computadores para execução paralela e/ou distribuída de aplicações da computação científica que necessitem de alto poder de computação, tais como *Picalc* (cálculo do π pelo método de Monte Carlo) e *Passbreak* (quebra de senhas). O principal objetivo do trabalho é a formalização e definição de uma memória distribuída para o Ambiente D-GM através da concepção e modelagem da integração entre o Ambiente D-GM e um sistema de memória compartilhada distribuída (DSM), obtendo consequentemente, uma melhor dinâmica de execução e possivelmente um aumento de desempenho. Pretende-se alcançar esta integração através da adaptação de um sistema DSM chamado Terracotta [Terracotta 2008] ao Ambiente D-GM, denominando-se então Ambiente ShareD-GM. O Ambiente ShareD-GM, por ser baseado no Ambiente D-GM, usa como núcleo a máquina geométrica (*Geometric Machine-GM*) [Reiser 2002], um modelo de máquina abstrato para computações paralelas e/ou concorrentes cujas definições abrangem os paralelismos existentes para execução de processos. Os módulos de software já desenvolvidos também passam a integrar este novo ambiente, são eles: (i) o módulo de desenvolvimento visual VPE-GM (*Visual Programming Environment for Geometric Machine*), a partir do qual são desenvolvidos os algoritmos da computação científica; e (ii) o módulo de execução VirD-GM (*Virtual Distributed Geometric Machine*) [Fonseca 2008], responsável pelo gerenciamento a execução dos algoritmos construídos no VPE-GM. A Figura 1 mostra a integração entre os módulos de software pré-existentes e o sistema Terracotta.

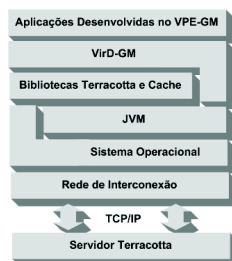


Figura 1. Arquitetura ShareD-GM

Atualmente, os sistemas DSM [Eskicioglu 2004] ressurgiram pela possibilidade de simplificação da programação de aplicações destinadas à executar de forma distribuída em um *cluster*. Dessa forma, as seguintes vantagens podem ser atribuídas aos sistemas DSM: (i) Otimização da comunicação inter-processos, gerenciada de forma automática pelo sistema DSM; (ii) Maior abstração na programação, não necessitando explicitar no código a comunicação dos processos; (iii) Possibilidade de aumento de desempenho, devido à troca de mensagens otimizada entre os processos; e (iv) Capacidade de adquirir funções de *middleware*, gerenciando a execução das tarefas da aplicação nos nodos e realizando o balanceamento automático de carga.

A conclusão dos estudos iniciais sobre DSM permitiu a definição do Terracotta, dentre outras implementações estudadas como JESSICA e JavaParty, como melhor opção de sistema DSM à ser usado nesta proposta. Verificou-se que: (i) É um dos sistemas mais utilizados por aplicações que necessitam de um maior poder de processamento; (ii) É um sistema atual e em constante atualização, desenvolvido na linguagem JAVA, herdando o seu modelo de memória e consistência; (iii) A coerência de memória é mantida através da atualização dos objetos replicados; (iv) Não modifica o código já construído da aplicação; e (v) Se propõe a reduzir o *overhead* de comunicação entre os objetos.

Este menor *overhead* gerado pelo sistema Terracotta deve ser alcançado pela não utilização de RMI para comunicação entre os objetos, evitando o *overhead* gerado pelas serializações e deserializações (*marshalling*) dos objetos quando transmitidos via rede usando este protocolo de comunicação de objetos. O Terracotta gerencia as JVMs do cluster a partir de um servidor e os objetos compartilhados, quando modificados, tem as suas modificações propagadas para as outras JVMs do cluster somente quando estes objetos forem acessados. Alguns testes realizados indicam que é possível obter uma diminuição de até 80% na quantidade de dados transferidos na rede ao usar o Terracotta ao invés do RMI.

No desenvolvimento do trabalho, faz-se necessário o estudo de compatibilidade da integração entre o Terracotta e o módulo de *software* VirD-GM. A integração quando realizada na prática irá viabilizar a validação do Ambiente ShareD-GM a partir da execução de aplicações da computação científica. Isto possibilitará a avaliação de medições de desempenho entre o sistema atual em funcionamento (Ambiente D-GM) e o sistema pretendido (Ambiente ShareD-GM).

Referências

- Eskicioglu, M. R. (2004). *Software distributed shared memory: issues and a case study*. PhD thesis, Edmonton, Alta., Canada.
- Fonseca, V. S. d. (2008). VirD-gm: Uma contribuição para o modelo de distribuição e paralelismo do projeto d-gm. Dissertação de mestrado em ciência da computação, UCPel, Pelotas, RS.
- Reiser, R. (2002). *A Máquina Geométrica - Um Modelo Computacional para Concorrência e Não-determinismo Usando como Estrutura os Espaços Coerentes*. Tese, II/UFRGS, POA, RS.
- Terracotta (2008). *The Definitive Guide to Terracotta: Cluster the JVM for Spring, Hibernate and POJO Scalability*. Apress.