

Desenvolvimento do DECK-GM

Clarissa Cassales Marquezan¹

Rafael Bohrer Ávila²

Philippe Olivier Alexandre Navaux³

358763

Resumo

Esse artigo apresenta o projeto da implementação do ambiente DECK utilizando a biblioteca GM para Myrinet. Serão vistos resumidamente os principais pontos da implementação.

1 Introdução

Uma das principais tecnologias de interconexão para clusters de alto desempenho é a Myrinet (BODEN, 1995). Com este panorama, torna-se necessário o suporte do DECK (BARRETO, 2000) para essa tecnologia. Esse trabalho apresenta o projeto do DECK (desenvolvido no Instituto de Informática, UFRGS) para Myrinet, utilizando o GM (MYRICOM INC., 2003).

Para implementação do DECK-GM, como será designada, foi necessária uma fase de estudo e ambientação com as duas bibliotecas. Para entender algumas decisões tomadas nesse trabalho, é preciso que algumas de suas características sejam apresentadas.

O DECK é uma biblioteca de comunicação paralela e distribuída que possibilita troca de mensagens e utilização de memória compartilhada. É baseada em *mail boxes* e dependendo do tamanho das mensagens a comunicação pode ser síncrona ou assíncrona.

O GM é uma biblioteca de troca de mensagens, desenvolvida pela Myricom. É baseada em eventos e em reserva de tokens, além disso é necessária a utilização de DMA no envio e recepção das mensagens.

Como cada biblioteca possui características específicas o maior desafio foi a criação de um modelo de comunicação que conciliasse a semântica do DECK e as propriedades do GM.

2 Implementação

Uma dificuldade na implementação o DECK-GM foi adaptá-lo ao paradigma de eventos. A solução encontrada foi o uso de uma *thread* que centraliza todos os eventos (envio e

¹clarissa@inf.ufrgs.br Bolsista PIBIC/CNPq

²avila@inf.ufrgs.br

³navaux@inf.ufrgs.br

recebimento), mantendo assim a semântica das primitivas de comunicação do DECK.

Outra dificuldade está no tratamento das mensagens. No GM elas são divididas em pequenas e grandes, entretanto devido a restrições de alocação de DMA no DECK-GM foi necessário dividi-las em pequenas, médias e grandes. Mensagens pequenas e médias são assíncronas enquanto grandes são síncronas, mantendo a semântica do ambiente DECK. É importante notar que para o recebimento de uma mensagem é estritamente necessário que se tenha um *token* de recebimento, um buffer alocado com DMA, prioridade e *size* igual aos da mensagem que deverá ser recebida. Essas características do GM implicam na existência de vários buffers alocados de acordo com as possíveis mensagens que chegarão na *mail box*, dificultando o gerenciamento das mensagens quando chegar na *thread* centralizadora.

Com os resultados obtidos com (Figura 1) foi possível a verificação de alguns pontos que podem ser melhorados na implementação, como por exemplo a otimização da utilização dos buffers internos do DECK que propiciará melhor utilização dos recursos.

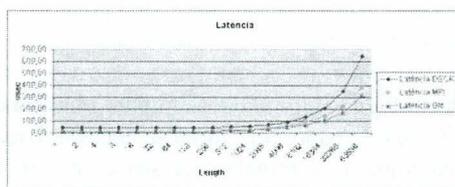


Figura 1: Comparação de Latências

Como conclusões do trabalho a implementação do DECK-GM possibilitou a utilização do ambiente DECK com mais uma tecnologia de interconexão, tornando-o mais portátil. Como trabalhos futuros estão a otimização dos pontos falhos detectados, a implementação da transmissão de mensagens grandes, assim como sua segmentação, a implementação de um mecanismo de retransmissão em caso de erro, a definição do limiar entre mensagens médias e grandes propiciando, no final, um melhor desempenho do ambiente.

Referências

BARRETO, M. E. *DECK : um ambiente para programação paralela em agregados de multiprocessadores*. Dissertação (Mestrado), 2000.

BODEN, N. J. et al. Myrinet – a gigabit-per-second local area network. *IEEE-Micro*, v. 15, n. 1, p. 29–36, Fevereiro 1995. Disponível em <http://www.myri.com/research/publications/Hot.ps>.

MYRICOM INC., M. I. *GM Reference Manual*. 2003. White Paper. Disponível em <http://www.myri.com/scs/GM/doc/refman.pdf>. 2003.