

Resolução do Problema de Difusão de Calor usando GPUs

Edmar André Bellorini, Guilherme Galante

¹Colegiado de Informática, UNIOESTE
Rua Universitária 2069, Cascavel-PR

{cabellorini, gui.galante}@gmail.com

1. Introdução

As Unidade de Processamento Gráfico - GPUs têm evoluído muito na última década, se tornando processadores flexíveis com um vasto poder computacional, chegando facilmente a 250 GFLOPs. Essas GPUs proporcionam amplos recursos computacionais tanto para o processamento gráfico, como para o processamento de uso geral voltado a área científica. Isso se deve à especialização em computação intensa e altamente paralelizada, possível pela atual arquitetura empregada nesses dispositivos.

Ferramentas tais como o CUDA (Compute Unified Device Architecture), permitem utilizar as capacidades computacionais das GPUs sem a necessidade de usar as APIs, evitando a adaptação dos dados [NVIDIA 2007]. Sob este escopo, o objetivo desse trabalho é apresentar um estudo sobre o uso das GPUs e da ferramenta CUDA na solução de problemas de natureza científica. Como estudo de caso é mostrado uma solução para o problema da Equação de Difusão de Calor usando a GPU para realizar os cálculos necessários.

2. Compute Unified Device Architecture - CUDA

CUDA é uma nova arquitetura de hardware e software que permite utilizar e gerenciar as capacidades computacionais paralelas das GPUs NVIDIA sem a necessidade de usar um mapeamento das computações para as APIs gráficas. A estrutura de software CUDA consiste em 4 camadas como pode ser observado na Figura 1, a primeira é a aplicação, seguida das bibliotecas de alto nível matemático (CUFFT e CUBLAS), sua API e *runtime* e finalmente o *driver* de hardware [CARAPETO and MAURÍCIO 2007].

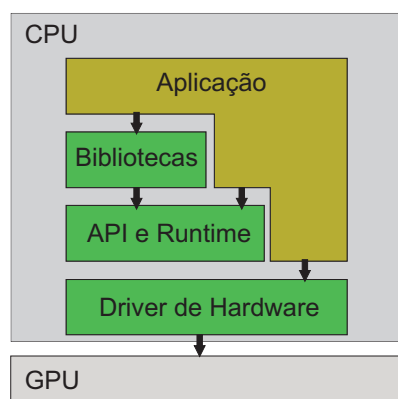


Figura 1. Arquitetura Cuda

Sendo uma extensão a linguagem de programação C, a interface de programação CUDA tem como objetivo fornecer uma forma simples e rápida de desenvolvimento de aplicações gerais para serem executadas pela CPU, diminuindo a curva de aprendizado e permitindo ao programador selecionar quais partes do código fonte sejam executados na CPU ou na GPU.

3. Estudo de Caso

Como estudo de caso para realização de testes comparativos do tempo de processamento entre CPU e GPU, foi escolhido um problema clássico de aplicação de métodos numéricos conhecido como difusão de calor em uma placa homogênea onde o objetivo é determinar a temperatura em qualquer ponto interno da placa em um dado instante de tempo levando em consideração que essa temperatura no instante inicial t_0 seja diferente das temperaturas das bordas.

O processo inicia com a escolha do domínio, onde é determinado o número de nós, isto é, a discretização da placa, quanto maior for o número de nós, maior será a precisão da temperatura em uma área da placa. Como condição inicial, defini-se o tempo total determinado para a simulação, dado em segundos, e também o tempo de cada iteração, juntamente com os valores das temperaturas das bordas e internas da placa. Em seguida, é gerado a matriz de coeficientes e o vetor de termos independentes [BORGES and PADOIN 2006]. Uma síntese do processo pode ser visto na Figura 2.

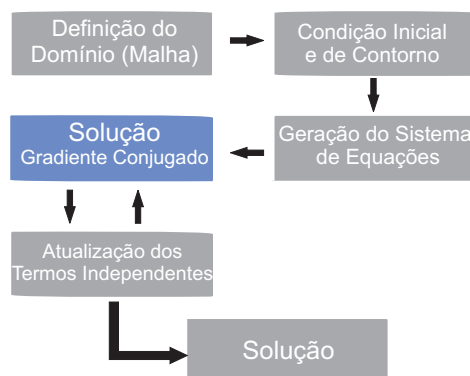


Figura 2. Processo de Solução

Na sequência utiliza-se o método do Gradiente Conjugado para encontrar as aproximações para o vetor resposta (temperaturas internas da placa) na iteração n . Atualiza-se o vetor de termos independentes e verificam-se as condições de parada, que é definido por um número máximo de iterações do gradiente conjugado ou se a diferença da resposta atual com a resposta da iteração de tempo anterior é menor que a precisão desejada. Caso uma dessas condições seja satisfeita, temos a resposta para essa iteração de tempo, caso contrário, uma nova iteração da simulação é realizada.

4. Testes e Resultados

Como ambiente de software foi utilizado o sistema operacional Windows XP SP2, ambiente de desenvolvimento Microsoft Visual Studio 2005 (por questão de compatibilidade

com a ferramenta CUDA 2.0). Utilizou-se uma CPU C2D E8400 3Ghz em um sistema com dois módulos de 1Gb de memória RAM DDR2 operando a 800Mhz em modo Dual Channel (totalizando 1600Mhz) e o chipset NVIDIA 650 SLI com suporte a um slot PCI-E 16x ou a dois slots PCI-E 8x.

Em relação às GPUs, foram utilizados nos testes dois modelos diferentes de dispositivos: NVIDIA 8500GT e NVIDIA 8600GT. Realizaram-se testes comparando os desempenhos entre as GPUs em modo compartilhado com o sistema (uma GPU gerando imagens para o monitor), em modo dedicado (onde são utilizadas duas GPUs sendo uma utilizada para saída de vídeo e outra apenas para os cálculos) e também em modo SLI (*Serial Link Interface*), também utilizando duas GPUs mas trabalhando em conjunto como se fosse apenas uma única GPU para o sistema (para cálculos e saída de vídeo).

Para os testes realizados com a solução da equação da difusão de calor foram utilizados dois códigos: o primeiro código em linguagem C de execução serial, isto é, sem nenhum tipo de paralelização das operações, para ser executado pela CPU e tomado como ponto de referência para avaliar o segundo código, que foi escrito utilizando a ferramenta CUDA em conjunto com a biblioteca CUBLAS para que a GPU ficasse responsável pelos cálculos requeridos na resolução do problema.

Foram utilizadas como condições iniciais da simulação, tempo total da simulação de 10 minutos, tempo de cada iteração de 0,5 segundo, assim resultando em 1.200 iterações de simulação. Para as condições de bordas, utilizaram-se os valores de 25°C para todo o domínio e apenas uma das bordas contendo valores de 45 °C. Optou-se por utilizar três tamanhos de domínio: o primeiro de tamanho 22 x 22 células, o segundo 42 x 42 células e o último como sendo domínio de 82 x 82 células.

As Tabelas 1 e 2 mostram os resultados obtidos da solução da equação de calor entre a CPU e as GPUs nos três domínios com as condições iniciais determinadas anteriormente.

Tabela 1. Tempos de Execução (segundos) - CPU x 8500GT

Domínio	CPU	Compartilhado	Dedicado	SLI
22x22	19,81	3,54	3,60	3,60
42x42	301,09	25,31	26,32	27,08
82x82	4.407,22	335,56	368,62	382,02

Tabela 2. Tempos de Execução (segundos) - CPU x 8600GT

Domínio	CPU	Compartilhado	Dedicado	SLI
22x22	19,81	2,52	2,46	2,54
42x42	301,09	17,84	17,28	18,12
82x82	4.407,22	212,40	210,70	213,18

A Figura 3 mostra os resultados obtidos na simulação (a) no passo inicial, (b) após 5 minutos e (c) após 10 minutos.

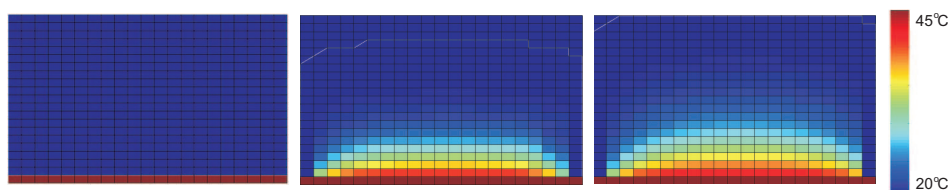


Figura 3. Passos da Solução de Difusão de Calor

5. Conclusão

De modo geral, os resultados obtidos foram extremamente satisfatórios em relação à solução da equação de calor, pois o ganho no tempo de execução em GPU em comparação com a execução em CPU chegou a valores aproximados de 90% no domínio 22x22, 95% no domínio 42x42 e significativos 98% para o domínio 82x82. Considerando que o custo da GPU 8600GT, que obteve os melhores resultados, não passa de 50% do valor da CPU utilizada.

Finalizando, como principal conclusão tem-se que o uso das GPUs (utilizando CUDA) aplicado a execuções de soluções para problemas que requerem grandes níveis de computações é uma abordagem válida em relação ao poder computacional, pois apresenta grande desempenho, certa facilidade na aprendizagem da linguagem, grande comunidade de discussão e custo/benefício atrativo, tornando uma alternativa barata para computação intensa e paralela.

Referências

- BORGES, P. A. P. and PADOIN, E. L. (2006). Exemplos de métodos computacionais aplicados a problemas na modelagem matemática. In *Anais ERAD*.
- CARAPETO, Ricardo R. and FRANCISCO, S. M. and MAURÍCIO, J. G. (2007). Programação em GPU: CUDA, [CTM. Technical report, UNIVERSIDADE TÉCNICA DE LISBOA.
- NVIDIA (2007). *NVIDIA CUDA: Compute Unified Device Architecture*. NVIDIA. Guia de Programação.