Resenha da Unidade 10 – Graphics Processing Units (GPU)

Artigo de Referência:

Changxin Li; Hongwei Wu; Shifeng Chen; Xiaochao Li; Donghui Guo; , "**Efficient implementation for MD5-RC4 encryption using GPU with CUDA,** " *Anti-counterfeiting , Security, and Identification in Communication, 2009. ASID 2009. 3rd International Conference on* , vol., no., pp.167-170, 20-22 Aug . 2009 doi: 10.1109/ICASID.2009.5276924. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5276924&isnumber=5276889

Neste artigo, os autores procuraram demonstrar as vantagens da arquitetura GPU na execução de algoritmos de criptografia. Essa arquitetura tem conquistado grande relevância nos últimos anos dado, principalmente, o crescimento explosivo no mercado de games. No entanto, com a evolução das GPUs para uma arquitetura de propósito geral, passou-se a utilizá-las em escala para uma grande variedades de aplicações, desde simulações(

nas mais diversas áreas de conhecimento), até aplicações no mercado de valores. Para desfrutar do que a arquitetura tem a oferecer, as aplicações precisam, basicamente, atender a um requisito: possibilidade de paralelização!

Coincidência ou não, os algoritmos de criptografia escolhidos pelos autores desfrutam da capacidade de paralelização, principalmente o de criptografia simétrica RC4, que consiste em operações XORs independentes, bit a bit, do texto claro com a chave de criptografia (expandida). Neste caso em particular, pode-se perceber facilmente que tal algoritmo também deve ter uma alta performance em uma CPU x86, haja vista que operações XORs
são executadas diretamente como um instrução neste arquitetura, inclusive podendo desfrutar de pipeline. Desta forma, notamos que a característica de ser altamente paralelizável por si só não é suficiente para demonstrar a eficiência da GPU sobre CPU. A escolha de outro algoritmo poderia ter tornado essa diferença ainda mais evidente. Quanto ao MD5, por se tratar de um algoritmo de resumo/
hash
(
não de criptografia), que realiza um conjunto maior de instruções, pode-se perceber melhor as vantagens de GPU sobre uma CPU.

Foi utilizada uma NVIDIA GeForce 9800GTX, contendo 128 GPU Cores e com suporte a plataforma CUDA. A implementação foi relativamente simples, sendo divida nos três estágios a seguir: 1) Os dados e uma chave de entrada são transferidas da CPU
para a GPU.
Ambos são armazenados na memória global da GPU.
2
) O
kernel
Grid é invocado para processar o problema paralelamente. Como resultado, temos os dados criptografados em RC4 e seu respectivo
hash
em MD5.
3
) Os dados são devolvido para a CPU.

Foi utilizada uma configuração de execução para processamento com 32 blocos, cada um executando 32 threads em paralelo(este é aquele parâmetro kernel <<<32,32>>>), sobre blocos de dados de 32bytes cada. Isso nos dá 32x32 linhas de processamento em paralelo. Com essa configuração eles conseguiram, inicialmente, um

throughput
de 71MB/s, o que, posteriormente, foi
otimizado
para 126MB/s). A otimização consistiu em fazer um melhor
uso(
ou seria o uso adequado?) das memórias internas da GPU. Cada thread no Grid consumiu 256bytes da memória compartilhada da GPU. Dado que cada bloco de Threads na NVIDIA utilizada é de 16384bytes, eles poderiam ter incrementado o número de threads por bloco para até 64 threads. No ensaio final, foram utilizados 256 blocos, cada um com 60 threads
cada(
ou seja, 15360 objetos rodando em paralelo), permitido um
throughput
de 217MB/s.

Para comparar os resultados com uma CPU, foi utilizado um AMD Sempron LE-1200, com cloc
k de
2.12
Ghz
(o
clock
da NVIDIA é de 1.83Ghz). A velocidade do processamento deste problema está diretamente relacionado com o tamanho dos blocos a serem
criptografados(
blocos esses que são processados em paralelo). Segundo os autores, o
throughput
da CPU foi de 70MB/s, contra os 217MB/s da GPU. No entanto, não foi mencionado qualquer detalhe da implementação utilizada na CPU, tal como características de paralelização explicita, linguagem de programação/compilador utilizado e etc. Embora, do ponto de vista de
performance
, tenha ficado claro as vantagens da arquitetura GPU sobre uma CPU neste estudo de caso, ainda assim a comparação pecou pela impossibilidade de ser reproduzida, uma vez que não foram apresentados os detalhes do experimento.