

Impacto da Organização dos Dados em Operações com Matrizes Esparsas na GPU

Paula Prata^{1,2}, Gilberto Melfe², Ricardo Pesqueira², João Muranho^{1,3}

¹ Instituto de Telecomunicações,

² Departamento de Informática,

Universidade da Beira Interior, 6201-001 Covilhã, Portugal

³ IMAR – Instituto do Mar, Departamento de Zoologia, FCTUC,

Universidade de Coimbra, 3004-517 Coimbra

e-mail: pprata@di.ubi.pt, a23049@ubi.pt, ricardopesqueira705@hotmail.com, muranho@mail.telepac.pt.

Resumo. A utilização de placas gráficas (GPU) para operações sobre matrizes, nomeadamente sobre matrizes esparsas, tem sido alvo de intensa investigação. Pretende-se obter o máximo desempenho de uma arquitectura com centenas de cores e um modelo de paralelismo de dados com execução simultânea de milhares de *threads*. Numerosas aplicações científicas e de engenharia manipulam matrizes esparsas, sendo o produto matriz-vector a operação base para vários algoritmos iterativos de resolução de sistemas de equações esparsas. Neste trabalho avaliamos o desempenho da operação matriz-vector para dois dos formatos mais importantes de armazenamento de matrizes esparsas: CSR e ELL. Mostramos como a ordenação das linhas pode aumentar significativamente o desempenho da operação estudada, no caso do formato ELL, e estudamos o comportamento da solução proposta na resolução de sistemas de equações esparsas correspondente a um problema real.

Palavras-Chave: placa gráfica (GPU), CUDA, paralelismo de dados, matrizes esparsas, sistemas de equações lineares, formatos de armazenamento para matrizes esparsas.