

Parallel computing - Exercise 5

Michela Venturini

Spring 2019

1 Implement the Fast-Transpose Algorithm in CUDA C with shared memory

The code implements the fast transpose algorithm in CUDA C and produces a comparison between the performances in both the case by means of the *memory bandwidth*. The implementation is based on the NVIDIA document MatrixTranspose and exploits the benefits gained through the usage of a tile to improve the access memory pattern. The code is executed on Ulysses on a host node provided with a GPU.

2 Results

The image 1 shows the results of the comparison between the Naive algorithm and the improved one. It is straightforward that the performance of the improved algorithm overcomes the base case. We can notice that the peak in terms of bandwidth is the case of 500 threads and the worst those with 1000 threads.

3 References

[1] <http://www.cs.colostate.edu/cs675/MatrixTranspose.pdf>

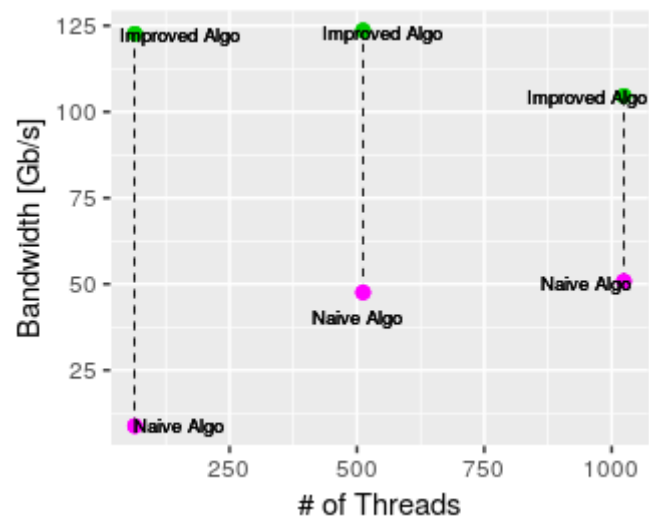


Figure 1: Comparison between Naive and Improved implementations