

Esercitazione 6 MNI

Studente: Giuseppe Cardaropoli

Matricola: 052251310

GPU: Tesla T4 (Colab)

Compute Capability 7.5
Maximum x -dim of grid of thread blocks = $2^{31}-1$
Maximum number of resident threads per multiprocessor = 1024
Maximum number of resident blocks per multiprocessor = 16
Maximum number of threads per block = 1024
Number of 32.bit registers per thread block = $64K$
Maximum amount of shared memory per thread block = $64KB$
Maximum amount of shared memory per multiprocessor = $64KB$

Tempo di Esecuzione Sequenziale

N	Tempo CPU(s)
1.000.000	0,003100
2.000.000	0,006073
4.000.000	0,012266
8.000.000	0,024248
16.000.000	0,048471
32.000.000	0,099446

Configurazione Ottimale Esecuzione n° 6

GPU: NVIDIA Tesla T4

Compute Capability: 7.5

max numero di blocchi: $2^{31}-1$

max num thread VSM: 1024

max num blocchi VSM: 16

max num thread Vblocco: 1024

num registri VSM: 64 K

shared memory massima VSM: 64 KB

shared memory massima Vblocco: 64 KB

blockDim = $1024/16=64$ $64 < 1024 \Rightarrow$ vincolo sulla dimensione del blocco rispettato

Strategia n°1

num registri usati da ogni thread = 10

registri totali usati VSM: $1024 \times 10 = 10240 < 64K \Rightarrow$ vincolo sui registri rispettato

Strategia n°2

num registri usati da ogni thread = 15

registri totali usati VSM: $1024 \times 15 = 15360 < 64K \Rightarrow$ vincolo sui registri rispettato

shared memory usata Vblocco: $64 \times 4 \text{ byte} = 256 \text{ bytes}$
shared memory usata VSM: $16 \times 256 \text{ byte} = 4096 \text{ bytes}$ } vincolo sulla shared memory usata rispettato

Strategia n°3

num registri usati da ogni thread = 9

registri totali usati VSM: $1024 \times 9 = 9216 < 64K \Rightarrow$ vincolo sui registri rispettato

shared memory usata Vblocco: $64 \times 4 \text{ byte} = 256 \text{ bytes}$
shared memory usata VSM: $16 \times 256 \text{ byte} = 4096 \text{ bytes}$ } vincolo sulla shared memory usata rispettato

In ogni strategia andremo ad utilizzare 16 blocchi e 1024 thread per ogni SM

Strategia n°1 (Global Memory)

N	Tempo GPU(s)	Speedup
1.000.000	0,008071	0,3804
2.000.000	0,016302	0,3725
4.000.000	0,032380	0,3788
8.000.000	0,064869	0,3737
16.000.000	0,130778	0,3706
32.000.000	0,008071	0,3812

Strategia n°2 (Shared Memory)

N	Tempo GPU(s)	Speedup
1.000.000	0,000264	11,47
2.000.000	0,000507	11,97
4.000.000	0,000946	12,96
8.000.000	0,001805	13,43
16.000.000	0,003625	13,37
32.000.000	0,007368	13,49

Strategia n°3 (Shared Memory)

N	Tempo GPU(s)	Speedup
1.000.000	0.000253	12,25
2.000.000	0.000442	13,73
4.000.000	0.000838	14,63
8.000.000	0.001688	14,36
16.000.000	0.003269	14,82
32.000.000	0.006567	15,14

