

Esercitazione 5: Carmine D'Angelo, Emanule Vitale, Francesco Aurilio

GPU: Tesla K80

Compute capability: 3.7

Massimo numero di thread per blocco per SM: 2048

Numero massimo di blocchi residenti per SM: 16

Massimo numero di registri a 32 bit per multiprocessor/thread: 131072

Massima dimensione della memoria condivisa: per thread: 48K

Strategia 1

N	Tempo CPU	Tempo GPU	Sp
100096	1	0,611648	1,236424
200192	1,681408	1,040896	1,615347
400384	3,424576	1,973312	1,735446
800768	6,94304	3,504288	1,981298
1601536	13,620352	6,80832	2,000545
3203072	28,305729	14,99392	1,887814

128 thread : $2048/128 = 16$ blocchi per occupare tutto un SM.

Con 16 blocchi: $128 \times 16 = 2048$ thread per SM. Piena occupazione dello SM!

Uso dei registri

Eseguendo l'istruzione “!nvcc -Xptxas -v” ottengo che il numero di registri utilizzato da ogni thread è .

Dunque, moltiplicando il numero di registri, per il numero di thread e per il numero di blocchi ottengo:

$$8 * 128 * 16 = 16384 < 131K$$

Strategia 2

N	Tempo CPU	Tempo GPU	Sp
100096	0,875552	0,249504	3,50917
200192	1,756992	0,280448	6,264948
400384	3,176096	0,295232	10,75797
800768	6,610272	0,47936	13,78979
1601536	13,251552	0,76224	17,38501
3203072	27	1,385985	19,70778

128 thread : $2048/128 = 16$ blocchi per occupare tutto un SM.

Con 16 blocchi: $128 \times 16 = 2048$ thread per SM. Piena occupazione dello SM!

Uso dei registri

Eseguendo l'istruzione “!nvcc -Xptxas -v” ottengo che il numero di registri utilizzato da ogni thread è 17.

Dunque, moltiplicando il numero di registri, per il numero di thread e per il numero di blocchi ottengo:

$$17 * 128 * 16 = 34816 < 131K$$

Info sulla shared memory

Dimensione della shared memory con allocazione dinamica: 64

Dimensione float: 4 byte

Dimensione della shared memory di ogni blocco * dimensione in byte dei float * il numero di blocchi:

$$128 * 4 * 16 = 8192 < 48K$$

Strategia 3

N	Tempo CPU	Tempo GPU	Sp
100096	0,75856	0,232736	3,259315
200192	1,77968	0,258752	6,877937
400384	3,344192	0,30336	11,02384
800768	7,1072	0,355616	19,9856
1601536	13,706016	0,603392	22,71494
3203072	28,863071	1,017856	28,35673

128 thread : $2048/128 = 16$ blocchi per occupare tutto un SM.

Con 16 blocchi: $128 \times 16 = 2048$ thread per SM. Piena occupazione dello SM!

Uso dei registri

Eseguendo l'istruzione "`!nvcc -Xptxas -v`" ottengo che il numero di registri utilizzato da ogni thread è 10.

Dunque, moltiplicando il numero di registri, per il numero di thread e per il numero di blocchi ottengo:

$10 \times 128 \times 16 = 20480 < 131K$.

Info sulla shared memory

Dimensione della shared memory con allocazione dinamica: 64

Dimensione float: 4 byte

Dimensione della shared memory di ogni blocco * dimensione in byte dei float * il numero di blocchi:

cublas

N	Tempo CPU	Tempo GPU	Sp
100096	0,791968	0,094176	8,409446
200192	1,728992	0,081952	21,09762
400384	3,527488	0,075008	47,02816
800768	7,069312	0,105216	67,18856
1601536	14,365536	0,162016	88,66739
3203072	28,634785	0,250976	114,0937

