

CUDA-C

di Petraglia Mariangela

Per il corso di Metodi Numerici per l'Informatica

A.A. 2018/2019

Caratteristiche macchina di test

I test sono stati eseguiti sulla macchina multiGPU del dipartimento di Matematica con le seguenti caratteristiche:

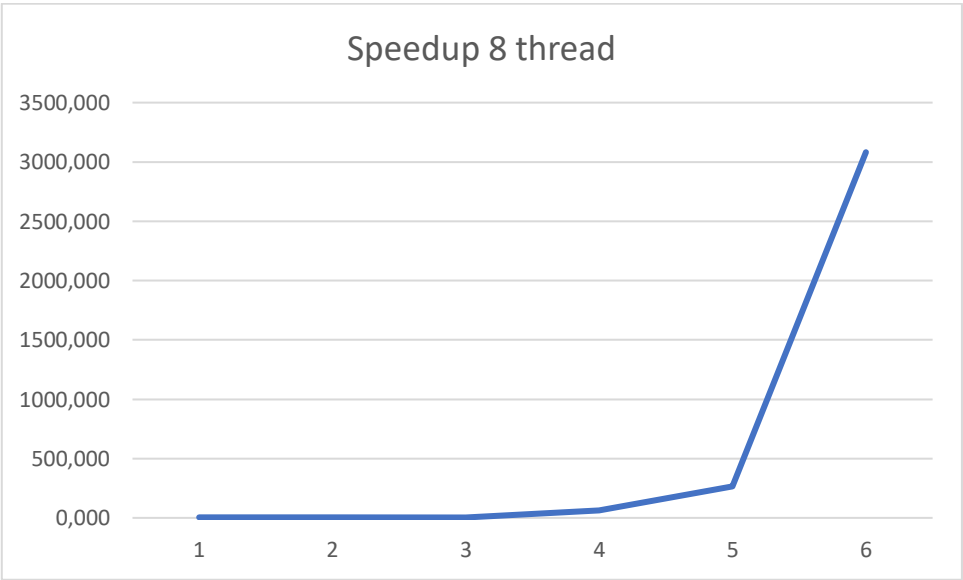
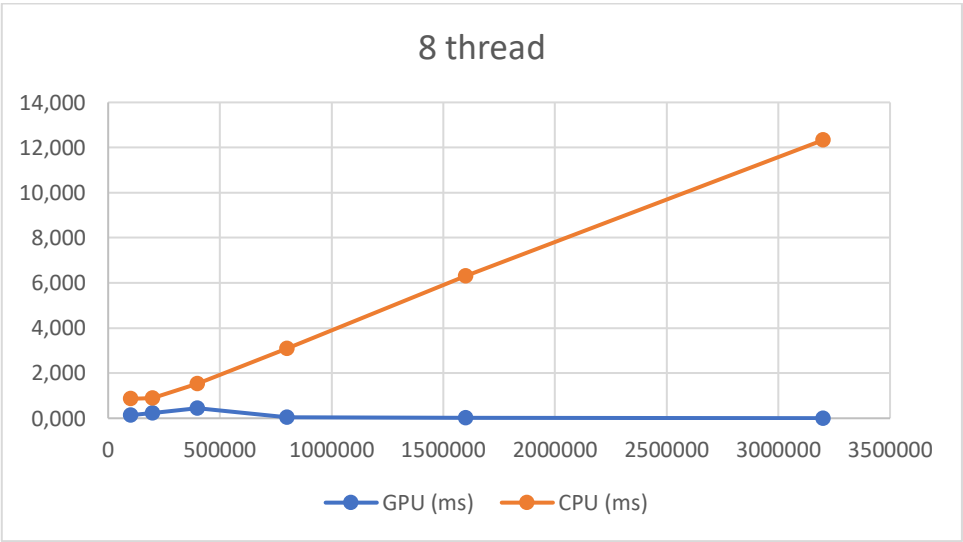
Il sistema utilizzato per la compilazione ed esecuzione del programma è un sistema formato da un cluster multi-GPU che si divide in due computing nodes E7095 costituiti da:

- CPU: Intel Xeon 6 core X5690 3,46GHz;
- RAM: 96GB DDR3 1333MHz;
- Hard Disk: 2 x 600GB SAS 15000 RPM;
- Network: 2 x 1GbE
 - eth0 connessa alla rete privata, tale interfaccia sarà usata anche per management IPMI;
- 4 processori di GPU computing NVIDIA Tesla C2050 (architettura Fermi):
 - Form Factor: 9.75''PCIe x 16;
 - CUDA cores: 448 (= 32 cores x 14 SM);
 - Frequenza CUDA core: 1.15 GHz;
 - Memoria dedicata totale: 3 GB GDDR5;
 - Velocità memoria: 1.5 GHz;
- ed un front-end E7226 costituito da:
 - CPU: INTEL Xeon 4 core E5602 2,13GHz;
 - RAM: 12GB DDR3 1333MHz;
 - Hard Disk: 3 x 500GB SATA 72000 RPM, 5x 2TB SATA 72000 RPM;
 - Network: 2 x 1GbE
 - eth0 connessa alla rete pubblica, tale interfaccia sarà usata anche per management IPMI;
 - eth0 connessa alla rete privata;

Esercizio 1 – tre diverse configurazioni del kernel (*tempi.cu*)

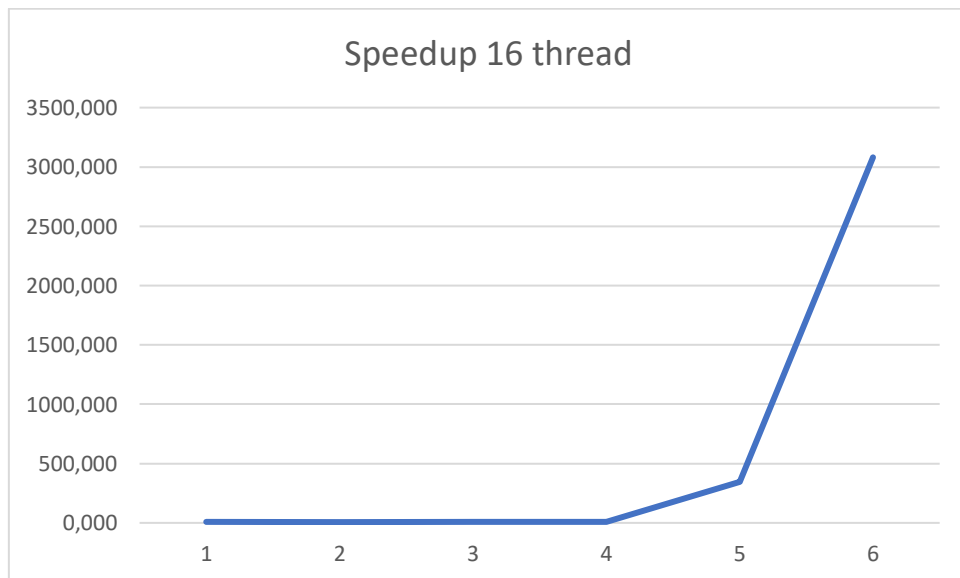
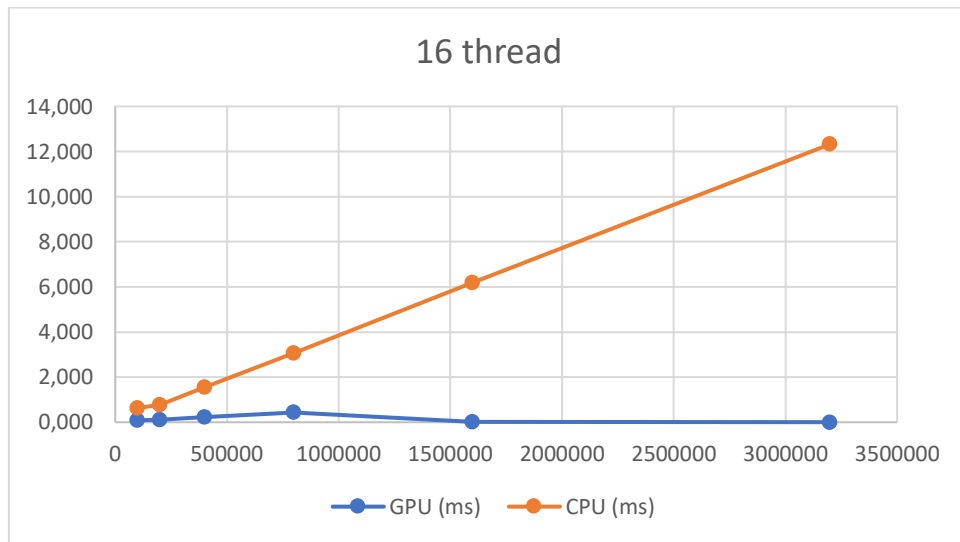
8 thread

N	GPU (ms)	CPU (ms)	Speedup
100000	0,132	0,861	6,523
200000	0,244	0,891	3,652
400000	0,446	1,526	3,422
800000	0,048	3,100	64,583
1600000	0,024	6,318	263,250
3200000	0,004	12,328	3082,000



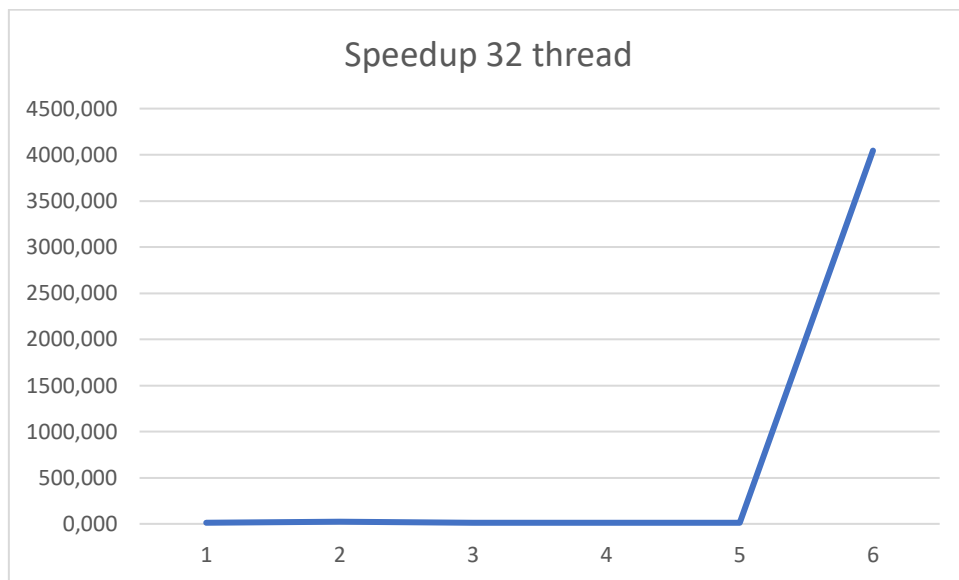
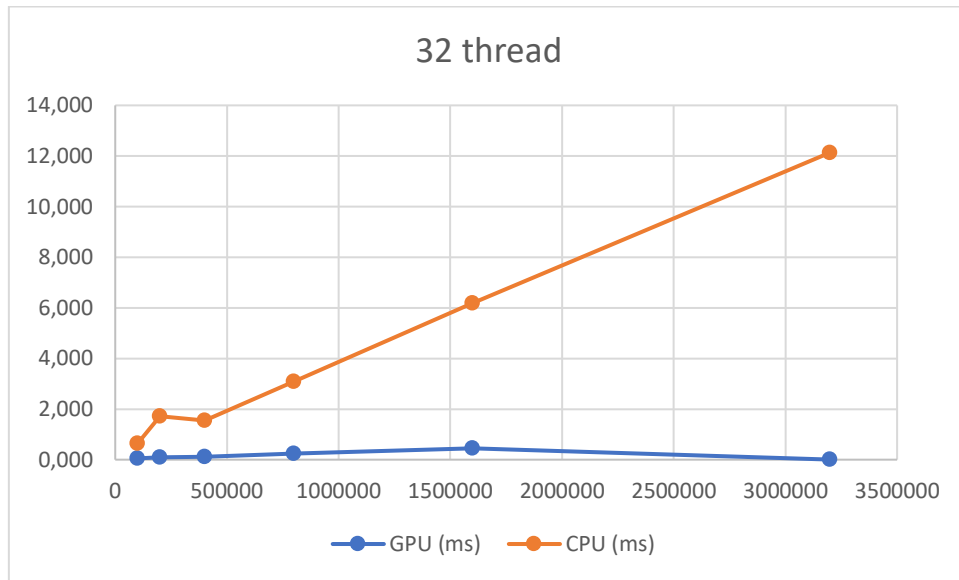
16 thread

N	GPU (ms)	CPU (ms)	Speedup
100000	0,078	0,630	8,077
200000	0,117	0,767	6,556
400000	0,234	1,542	6,590
800000	0,432	3,073	7,113
1600000	0,018	6,190	343,889
3200000	0,004	12,323	3080,750



32 thread

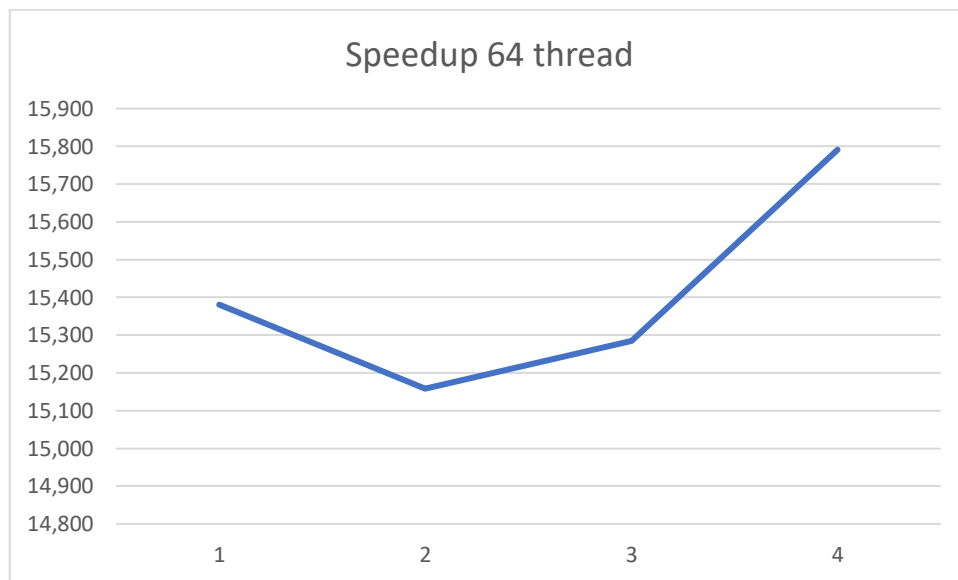
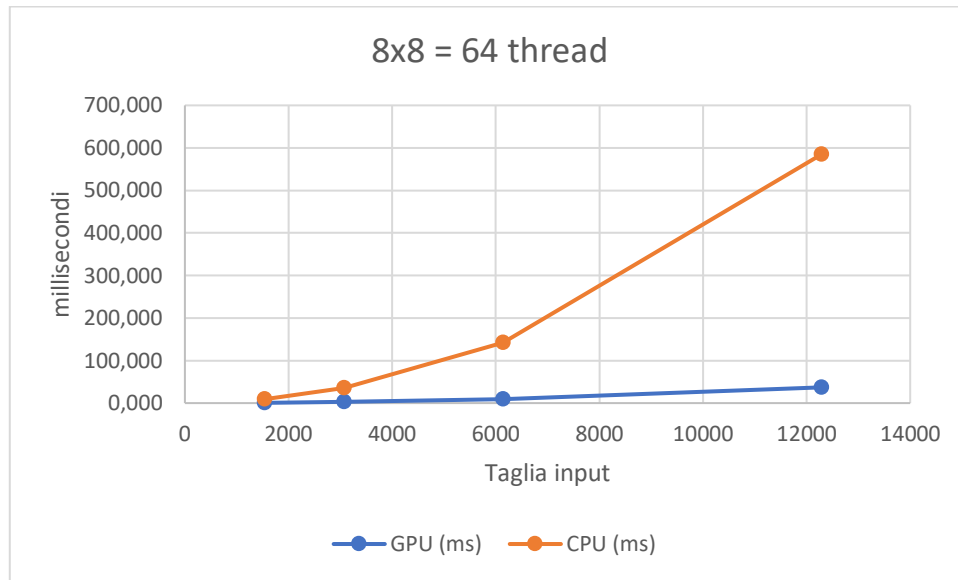
N	GPU (ms)	CPU (ms)	Speedup
100000	0,052	0,638	12,269
200000	0,083	1,72	20,723
400000	0,115	1,537	13,365
800000	0,242	3,092	12,777
1600000	0,448	6,193	13,824
3200000	0,003	12,135	4045,000



Esercizio 2 – somma matrici quadrate

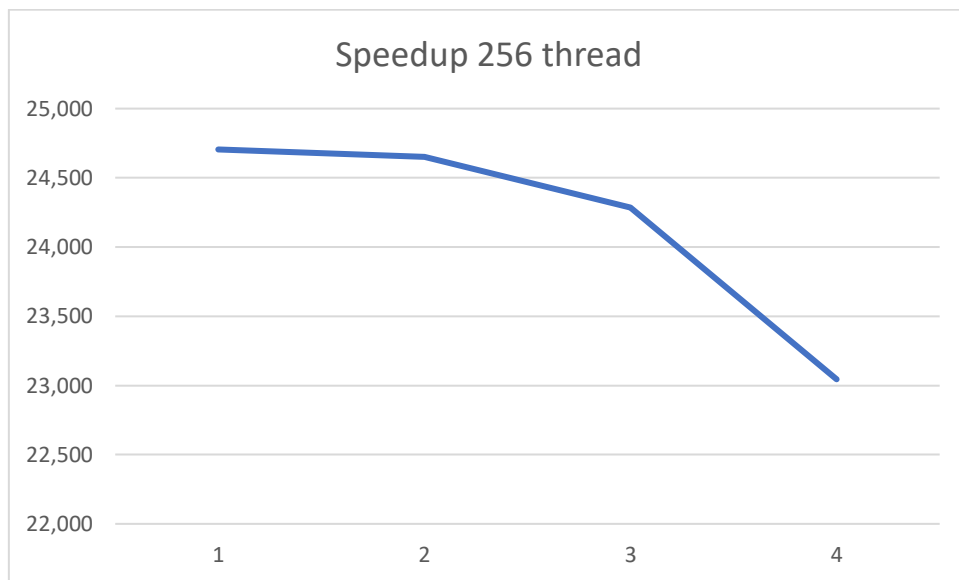
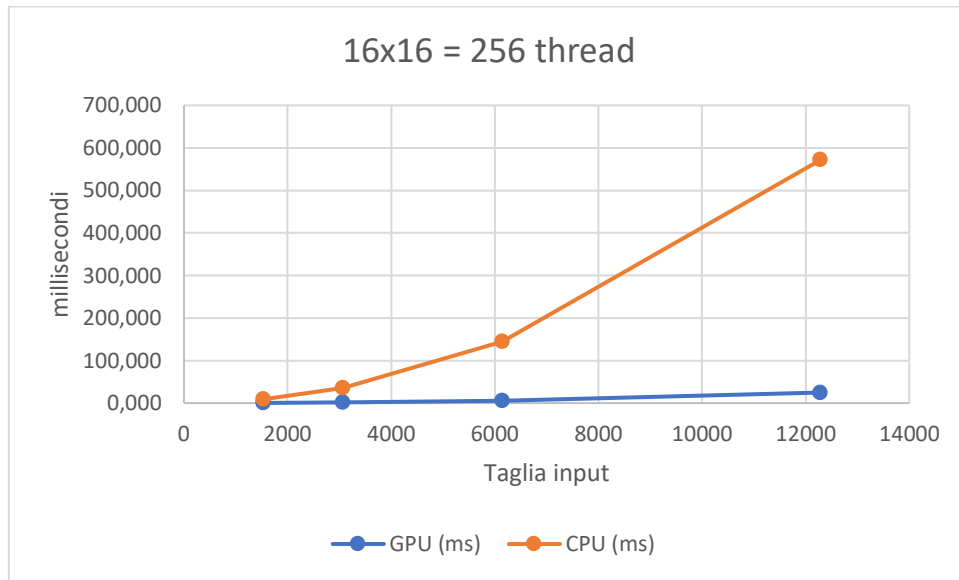
8*8 = 64 thread

N	GPU (ms)	CPU (ms)	Speedup
1536	0,590	9,075	15,381
3072	2,378	36,046	15,158
6144	9,311	142,311	15,284
12288	37,037	584,851	15,791



16*16 = 256 thread

N	GPU (ms)	CPU (ms)	Speedup
1536	0,366	9,042	24,705
3072	1,465	36,112	24,650
6144	5,959	144,714	24,285
12288	24,815	571,851	23,045



32*32 = 1024 thread

N	GPU (ms)	CPU (ms)	Speedup
1536	0,379	9,065	23,918
3072	1,553	36,058	23,218
6144	5,786	141,282	24,418
12288	27,367	566,068	20,684

