

Oct 24, 18 21:10

gpu_vector_add.c

Page 1/2

```

/*
 * Simple CPU program to add two long vectors
 *
 * Author: Inanc Senocak
 * Used by: Dustin (Ting-Hsuan) Ma
 * compile using : nvcc -O2 gpu_vector_add.cu -o exec -gencode arch=compute_61,code=sm_61
 */

#include "timer_nv.h"
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include <sys/resource.h>

__global__ void vector_add_gpu(const int n, const float *a, const float *b, float *c)
{
    int tid = blockIdx.x * blockDim.x + threadIdx.x;
    if (tid < n) c[tid] = a[tid] + b[tid];
}

void vector_add_cpu(const int n, const float *a, const float *b, float *c)
{
    for (int i = 0; i < n; i++)
        c[i] = a[i] + b[i];
}

int main(int argc, char *argv[])
{
    if (argc < 2) {
        perror("Command-line usage: executableName <vector size>");
        exit(1);
    }

    int n = atof(argv[1]);
    float *x, *y, *z;

    cudaMallocManaged(&x, n * sizeof(*x));
    cudaMallocManaged(&y, n * sizeof(*y));
    cudaMallocManaged(&z, n * sizeof(*z));

    for (int i = 0; i < n; i++) {
        x[i] = 3.5;
        y[i] = 1.5;
    }

    StartTimer();

    vector_add_cpu(n, x, y, z);
    printf("vector_add on the CPU. z[100] = %4.2f\n", z[100]);

    double cpu_elapsedTime = GetTimer(); // elapsed time is in seconds

    for (int i = 0; i < n; i++) {
        z[i] = 0.0;
    }

    cudaEvent_t timeStart, timeStop; // WARNING!!! use events only to time the device
    cudaEventCreate(&timeStart);
    cudaEventCreate(&timeStop);
    float gpu_elapsedTime; // make sure it is of type float, precision is milliseconds (ms) !!!

    int blockSize = 256;
    int nBlocks = (n + blockSize - 1) / blockSize; // round up if n is not a multiple of block
size

    cudaEventRecord(timeStart, 0); // don't worry for the 2nd argument zero, it is about cuda
// streams

    vector_add_gpu<<<nBlocks, blockSize>>>(n, x, y, z);
    cudaDeviceSynchronize();

    printf("vector_add on the GPU. z[100] = %4.2f\n", z[100]);

    cudaEventRecord(timeStop, 0);
    cudaEventSynchronize(timeStop);

    // WARNING!!! do not simply print (timeStop-timeStart)!!

    cudaEventElapsedTime(&gpu_elapsedTime, timeStart, timeStop);

    printf("elapsed wall time (CPU) = %5.4f ms\n", cpu_elapsedTime * 1000.);
    printf("elapsed wall time (GPU) = %5.4f ms\n", gpu_elapsedTime);

    cudaEventDestroy(timeStart);

```

Oct 24, 18 21:10

gpu_vector_add.c

Page 2/2

```

    cudaEventDestroy(timeStop);

    cudaFree(x);
    cudaFree(y);
    cudaFree(z);

    return EXIT_SUCCESS;
}

```