МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ

(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Институт №8 «Компьютерные науки и прикладная математика»

Кафедра 806 «Вычислительная математика и программирование»

**Лабораторная работа №0**

**по курсу «**Параллельные и распределенные вычисления**»**

**Изучение технологии CUDA**

Выполнил: *Буловятов А.М.*

Группа: *М8О-310*

Преподаватель:  А.Ю. Морозов

Москва, 2023

**Условие**

Цель работы: Ознакомление и установка программного обеспечения для работы с программно-аппаратной архитектурой параллельных вычислений (CUDA). Реализация одной из примитивных операций над векторами.

В качестве вещественного типа данных необходимо использовать тип данных double. Все результаты выводить с относительной точностью , ограничение n < 255.

Вариант 6: Поэлементное возведение в квадрат вектора.

**Программное и аппаратное обеспечение:**

GPU:

Графическая карта: NVIDIA Tesla T4

Compute capability : 7.5

Глобальная память : 15835398144

Разделяемая память : 49152

Регистров в блоке : 65536

Макс. размер блока : (1024, 1024, 64)

Макс. размер сетки : (2147483647, 65535, 65535)

Константная память : 65536

Количество мультипроцессоров : 40

CPU: Intel(R) Xeon(R) CPU @ 2.30GHz

**Метод решения** заключается в вычислении возведения в квадрат параллельно на GPU. Алгоритм состоит из 5 частей 1.

* Ввод данных из stdin и сохранение их в RAM
* Копирование вектора из RAM на память GPU
* Поэлементное умножение
* Копирование из памяти GPU в память RAM
* Освобождение памяти и вывод результата

**Описание программы**: Программа реализована в одном файле main.cu, в ядровой функции мы получаем элементы для последующего возведения в квадрат. Для определения элемента использовали только одну размерность, так как задача у нас также в размерности 1.

**Результаты**

Тесты работы с разными конфигурациями на 1 000 000 000 записей.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| <1, 32> | <32, 32> | <64, 64> | <128, 128> | <515, 515> | <1024, 1024> |
| 11.26 s | 4.59 s | 1.19s | 0.73s | 0.93 s | 0.68 s |

Тесты работы с разными конфигурациями на 100 000 записей.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| <1, 32> | <32, 32> | <64, 64> | <128, 128> | <515, 515> | <1024, 1024> |
| 2.08 s | 1.42 s | 0.72 s | 0.36 s | 0.11 s | 0.11 s |

Тесты работ на CPU c 1 000 000 000 записей.

Время работы: 13.19 s.

**Выводы**

Использование технологии CUDA позволяет значительно ускорить выполнение операций поэлементного возведения в квадрат вектора. В данной лабораторной работе производительность на GPU оказалась выше, чем на CPU. Однако, для достижения максимальной производительности на GPU необходимо правильно выбирать параметры запуска.