Cardiff School of Computer Science and Informatics

Coursework Assessment Pro-forma

Module Code: CM3103
Module Title: High Performance Computing
Lecturer: Professor David W. Walker
Assessment Title: Programming with CUDA
Assessment Number: 2B
Date Set: 13 November 2019
Submission Date and Time: 13 December 2019 at 9:30am
Return Date: 31 December 2019

This assignment is worth 15% of the total marks available for this module. If coursework is submitted late (and where there are no extenuating circumstances):

1. If the assessment is submitted no later than 24 hours after the deadline, the mark for the assessment will be capped at the minimum pass mark;
2. If the assessment is submitted more than 24 hours after the deadline, a mark of 0 will be given for the assessment.

Your submission must include the official Coursework Submission Cover sheet, which can be found here:

https://docs.cs.cf.ac.uk/downloads/coursework/Coversheet.pdf

Submission Instructions

All submissions should be via Learning Central unless agreed in advance with the Director of Teaching. Students should submit all their files as a single zip (*.zip) file.

Description		Type	Name
Cover sheet	Compulsory	One PDF (*.pdf) file	[student number].pdf
Report	Compulsory	One PDF (*.pdf) file, including any relevant plots.	report_[student number].pdf
Code	Compulsory	One or more C source files	No restriction
Input image file	Compulsory	One Postscript (*.ps) file	David.ps
Output image file	Compulsory	One Postscript (*.ps) file	DavidBlur.ps
Output data file	Optional	One Excel (*.xlsx or *.xls) file	Spreadsheet of your timing data

Any code submitted will be run on a computer in the Linux Lab (C/2.08) and must be submitted as stipulated in the instructions above.

Any deviation from the submission instructions above (including the number and types of files submitted) will result in a mark of zero for the assessment or question part.

Staff reserve the right to invite students to a meeting to discuss coursework submissions

---

Assignment

Note: You should either do this coursework (2B), or Coursework 2A. You should not do both.

In this coursework you will modify an existing sequential program in C to produce a parallel version of the code that uses CUDA. You will then run your CUDA code on a machine with an NVidia GPU installed, such as the machines in the Linux lab. You will time different sections of your code and discuss any opportunities for optimising it. In addition to the lecture notes, you may find NVidia's "CUDA Zone" web site: https://developer.nvidia.com/category/zone/cuda-zone.

Here's what you need to do:

1. Download the sequential code, blur.c, and the input file, David.ps, from Learning Central.
2. Log into your account on one of the machines in the Linux lab (or some other machine with a CUDA-capable GPU – see https://developer.nvidia.com/cuda-gpus), and copy blur.c and David.ps to your file space (for example, using the scp command).
3. Edit blur.c to correctly refer to the input and output files in your file system.
4. Compile and run blur.c several times (say, between 6 and 10 times), noting the time for the execution of the main computational work. Evaluate the average and standard deviation of these times.
5. Produce a parallel CUDA version of blur.c named blurCUDA.cu. Edit the code so that the time spent on the following is measured:
    a. Reading in the input file.
    b. Allocation of device memory.
    c. Transferring data between host memory and device memory.
    d. Doing the blurring.
    e. Outputting the blurred image.

    Run the code several times and record the average time, and the standard deviation, for each phase. Vary the value of nblurs (try 10, 20, 40, 80, and 160) and check if the time per blur is a constant.

    [8 marks for correct CUDA code]
6. Write a report (2 or 3 pages of text, plus figures) that presents and interprets the results of your timing experiments. The report should include:
    a. A description of the hardware and software environment.

[1 mark]

b. A description of the timing experiments carried out.

[1 mark]

c. Appropriate graphs of your timing experiments.

[2 marks]

d. A discussion of the results that accounts for any unusual features, and suggests ways to improve the performance of the CUDA program.

[2 marks]

7. A short section presenting any overall conclusions, and giving a reflection on what you have learned from this coursework.

[1 mark]

Learning Outcomes Assessed

1. The ability to recognize the potential for exploiting parallelism to solve a specific problem.
2. The ability to write simple, regular parallel applications using CUDA.

Criteria for assessment

Credit will be awarded as given in the Assignment section above.

Feedback and suggestion for future learning

Feedback on your coursework will address the above criteria. Feedback and marks will be returned by 31 December 2019 via email.

Feedback from this assignment will be useful for you.