Project 1: An Introduction to Using GPU Systems
Due: March 7, 2012

1. Introduction

The objective of this project is to gain hands-on experience of porting a simple program to GPUs. You will have three weeks to complete this project. During this project, you will also learn how to design and conduct an experimental study. Your goal is to obtain experimental results to answer a few research questions regarding the performance implications of porting code to run on a GPU. The first part of the project involves you taking a sequential C kernel and getting it to run on a GPU using OpenCL or OpenACC. Please just use OpenCL or OpenACC for this project. Once ported, you should measure the running time of the parallel and sequential version of the code and report how parallelization affects the performance of the code. The second part of the project involves taking a large program and compiling and profiling it to identify the most important parts of the code. You should then describe what “hot” portions of the code could potentially be parallelized and how (either data or task parallelism).

2. Project Deliverables

Your first deliverable is to meet with the instructor or Scott Grauer-Gray (sgrauerg@gmail.com) to checkpoint your progress. You should be far enough along that you would have 10 minutes of information to discuss. A checkpoint should not consist of problems logging in and/or using the tools. Of course, if you are having problems it is important that you email the instructor or Scott to discuss this with either of us, but this does not constitute a checkpoint. The checkpoint will count for roughly 20% of your project grade. The second deliverable for this project will be an original and complete project report discussing experiments and results. The report should be two pages in length.

3. Project Report Guidelines

Reports must be well written and formatted correctly. We expect you to clearly describe the project and experimental results. You are required to present experimental results using graphs. You are required to use ACM proceedings format, and we will provide a latex or Word template to help with formatting.

4. Steps in Designing and Conducting the Experiments

Before running experiments, you should come up with a set of research questions you want to answer. Developing and selecting a concise set of research questions (see sample research questions below) will help you focus on the experiments you need to run.
1. Justify your selection of the two programs you chose in terms of conducting an unbiased experiment with results others can reproduce and (hopefully) that can generalize to other programs.

2. Identify the independent and dependent variables of the experiment. What will you measure in order to answer the research questions? Which metrics will you vary and which will you keep constant? Be sure that these are variables that you can indeed measure with the available tools.

3. Develop an experimental procedure.

4. Decide how to present the results. What will be the columns and rows of data tables? What graphs would help illustrate the results to easily draw conclusions to the research questions? What will the axes of the graphs be? What form will the data points be shown in, e.g., bar graph, scatter plot, line graph, box and whiskers graph, etc.?

5. Write a draft of the report based on the above steps, including an experimental methodology and results section that contains place markers for tables and graphs where you will eventually put your results. You should conduct peer editing among the members of your team and create different versions of this document.

6. Given Step 3 revisions, conduct experiments on the two programs and collect data.

7. You can use a variety of different tools to construct graphs. I suggest using Calc (the free open-source spreadsheet in Open Office), Excel, or Gnuplot. You can find lots of helpful information on the Internet on how to use these tools.

8. You should conduct an analysis of the data and make observations of various trends. This analysis and observations will become part of your results discussion. One way of doing this is to first delineate the initial set of observations as a bulleted list and conduct discussion of data and your observations with your team. Often, during this discussion additional experiments that you should have run will be identified. Run an additional set of experiments if necessary and continue iterating this process until you and your team are satisfied.

9. Write the complete experimental study report

Format of Experimental Study Report:

1. **Title, Authors, and Abstract** – This report should look like a conference paper you are going to submit, including properly formatted title and authors. Also, include a brief abstract.

2. **Research Question(s)** – The paper should include a concise statement of the research questions and why these are important to investigate.

3. **Variables and Measures** – You should include a description of the independent and dependent variables and how you measured the dependent variables.

4. **Programs** – Include a table with the two program names, lines of code (LOC), and short phrase describing what each program does. Also, include a paragraph on the programs you chose and why you chose them.

5. **Experimental Methodology** – Your paper should include paragraphs describing what you actually did to generate results. For example, include details such as what compiler you used and what other tools were used, how many times you ran
each program, and how measurements were taken (e.g., what was used to time program execution?). Specify enough details so the instructor could replicate your experiment.

6. **Threats to Validity** – You should describe why results might not generalize to other contexts (e.g., may not hold for another compiler, other programs, programs of other language, etc.) and how you tried to reduce the bias and increase the generality in the experimental design. Any shortcomings of your study go in this section.

7. **Results and Discussion** – This section includes tables, graphs, and an explanation of each table/chart starting with what data it contains and how to read it and any observations you can draw. This is a discussion that pertains to your original research questions you embarked to ask and how this data answers those questions.

8. **Related Work** – You should include paragraphs on related work that is relevant to this study. Summarize briefly (two or three sentences) this work and how it related to your study.

9. **Conclusions** – Finally, your paper should end with a summary of the experiment and results and any conclusions that could possibly be made from the experiment.

**Benchmarks:**

We have prepared a set of sequential kernels that you can use to parallelize. You should work on parallelizing one of these using OpenCL or OpenACC. You can find the codes in the following location:

www.eecis.udel.edu/~cavazos/cisc879/Project1/parallelizableLoops

You will find several directories each with three files, a dummy file, a makefile, and the kernel program you will be parallelizing. Please email the instructor to let him know which kernel you will be parallelizing.

**Larger Applications**

We provide links to several larger codes that can be used for running, compiling, and profiling. These are just suggestions, and if you have a program that you are interested in profiling that is not from this list, please feel free to use that.

MediaBench II Benchmarks: http://euler.slu.edu/~fritts/mediabench
MiBench Benchmarks: http://www.eecs.umich.edu/mibench
NAS Parallel Benchmarks: http://www.nas.nasa.gov/publications/npb.html
San Diego Vision Benchmark Suite: http://parallel.ucsd.edu/vision