

18-645: How to Write Fast Code I

Project 2:- Kernel Design

Instructions. A 3 to 4 page design document is to be submitted via Canvas. **The project updates must be submitted as a pdf document. Documents must be submitted individually.**

The main focus of this submission is to design the kernel(s) of the project, and to benchmark the existing implementation. Teams should feel free to use office hours to talk about their projects as well.

Requirements: The document must address the following:

1. *Kernel Description.* Provide a description of the kernels you have identified. For each of the identified kernels, you need to specify the following:
 - What are the independent operations?
 - What are the dependent instructions that make up the independent operation?
 - What functional unit(s) are used to compute those instructions?
2. *Performance Peak.* Based on your description of the kernel(s), answer the following questions:
 - (a) Describe the machine of your choice (model & architecture)
 - (b) What are the latency and throughput of the instructions identified in the previous question?
 - (c) Do you have the appropriate SIMD instructions on your machine?
 - (d) Are there specialized units (beyond SIMD) that can be used for this kernel?
 - (e) Which is the instruction/instructions that is the bottleneck for the kernel?
 - (f) What is the theoretical peak achievable for the kernel?
 - (g) Explain how you computed the theoretical peak of your kernel(s)
3. *Performance Baseline.* Provide baseline performance plots for the algorithm(s) you have identified. **You should not be writing this code from scratch.** This should be a performance plot based on existing code. The data set use to create these plots should also be the same data set used to test your implementation.
 - (a) Provide performance numbers for different input data types and sizes if applicable.
 - (b) How close to the theoretical peak is the existing baseline?
4. *Design of your implementation.* This is the main focus of this portion of your project where you are required to design kernels based on the information gathered previously.

Provide answers to the following questions:

- (a) What is the size (how many outputs) of the kernel?
- (b) How many registers do you need, and how many registers do you have on the selected machine?
- (c) How would you allocate registers to inputs, outputs and temporary variables?
- (d) How would you implement your kernel?

This description should be sufficiently detailed so that it is possible to implement the kernel based on your description.