Fall 2022: CS 691

Homework – 2

The problem was to implement an optimized version of the Game of Life program. The objective is to improve performance of the program with the Use of GPU's. The GPU used for the purpose of this experiment is Nvidia Volta. A single Volta GPU has a capacity of 84 Streaming Multiprocessors (SM's). The Shared memory of the SM is fast accessible and will save compute time for the process.

To measure the performance of the program, the program was tested both on CPU node(local PC) and GPU nodes(Alabama Super Computer).

It also involves implementation of MPI version of the code and uses MPI communication from the CPU (i.e., copies data from the GPU to the CPU and then uses MPI primitives to communicate between the GPUs) b. uses CUDA-aware MPI to communicate between the GPUs

To compile the program the program GCC version -9.1.0 was used on the CPU execution and nvcc on the GPU.

Algorithm Versions	Execution Time
Naïve CPU Version	1 min 6 sec
Multi-CPU MPI Version	194.155 mseconds
Single GPU version	5.017184 ms
Multi-GPU version (a)	6.254800 ms
Multi-GPU version (b)	8.256400 ms

Fig: Table 1

The GPU version used 4 cores on a volta GPU utilizing 25.77% of the memory and The GPU version is significantly faster than the cpu version almost a million times faster compared to the sequential version. The MPI version which ran on the cpu nodes of the DMC is faster than the sequential version.

All the tests were conducted using a 1024 x 1024 size matrix for 5000 iterations on the dmc cluster