UPPSALA
UNIVERSITET

Accelerator-Based Programming

Assignment 3
A TENSORFLOW EXAMPLE

Jinglin Gao

September 27, 2023

1 Compute time of GPU node on UPPMAX

In this task, we use a GPU node on Uppmax to compute the implementation of Conway's Game of Life. First we use 1000 iterations and verify the code's correctness. The result is shown below:

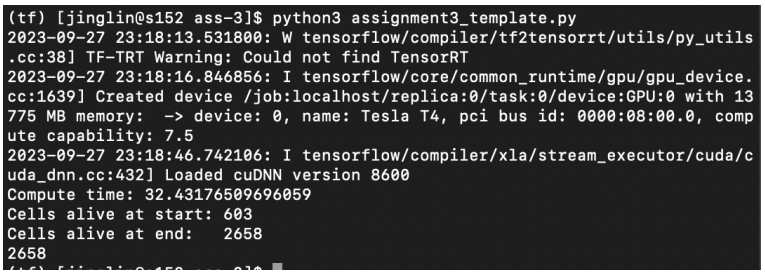

Figure 1: Compute time on GPU node

From the result, we can see that the compute time is 32.43 seconds. The result after 1000 iterations is 2658, which indicates my implementation is correct.

2 Compute time of CPU node on UPPMAX

Next, we disable the GPU by adding two lines to the start of the node and compute the time needed to finish 1000 iterations again. The result is shown below:

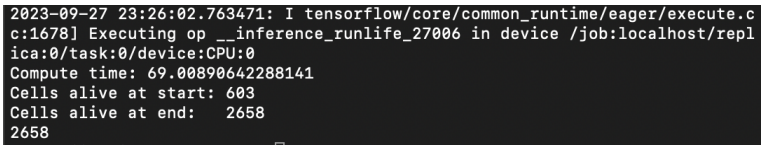

Figure 2: Compute time on CPU node

So when we disable the GPU node we can notice that there is a significant performance drop. Now we need 69.01 seconds to finish 1000 iterations on the CPU node, which is around 2.15× slower than the GPU node.

3 Switch Data Type and Compare Run Time

Run Time		
Data Type	GPU node (sec)	CPU node (sec)
float16	32.43	69.01
float32	34.56	77.89
int32	48.69	85.16
bfloat16	34.42	61.38

Generally, we can expect that GPU will provide better performance than CPU, and our experiments align with this.

GPUs are optimized for certain data types and operations. Here we use Tesla T4, which is optimized for FP16 and FP64, so here in the result, we can see that we get better performance for **float** datatype than **int** datatype on GPU. If we change the iteration times from 1000 to 3000, we will see a more significant difference between the compute time of float16 and float32, which is 101.63 and 108.77, respectively.

In conclusion, GPU has better performance than CPU in general. But for different data types, GPU performs better for some specific data types since it is optimized for some specific computing.