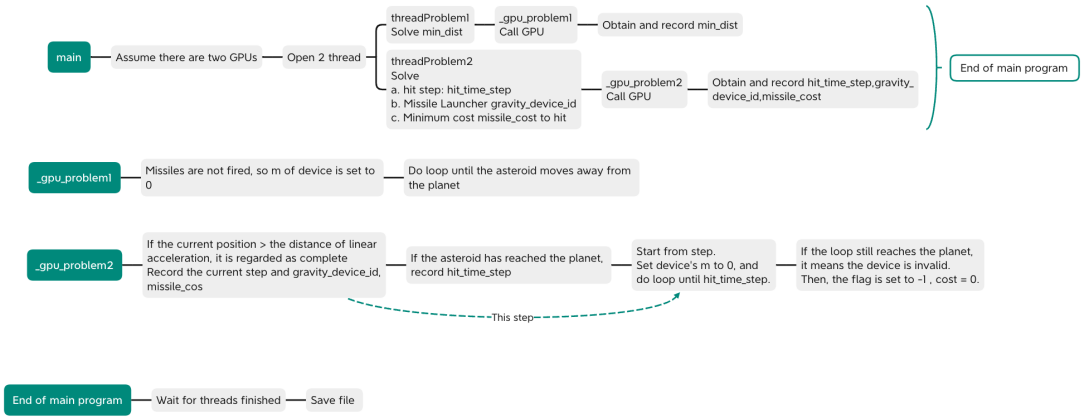
HW5: N-body Simulation

110065504 葉之晴

1. What is your parallelism strategy?



Open two threads, one of which is used to deal with the minimum distance (min_dist), first call the GPU through _gpu_problem1 to obtain and record the min_dist. The other one is used to solve the hit step (hit_time_step), the missile launch (gravity_device_id), and the calculation of the minimum cost of the hit (missile_cost). First, call the GPU through _gpu_problem2 as the same above, and then obtain and record the answers of the above three.

The methods of speeding up are as follows:

a. Since the time cost of sqrt is found to be too high, we changed to use rsqrt instead in _gpu_problem2, while in _gpu_problem1, we do not do sqrt, and record the dx, dy, and dz first, and the square value is directly used to judge, and then do the calculation sqrt after completion.
b. During the test, I found that pow would generate a lot of time cost. The specific reason is uncertain, but I switched to manually calculating the power of 1.5.
c. Originally, to speed up the operation, the compiler will convert some division operations into multiplication operations. In order to reduce the time of this conversion, we directly manually convert the division in the program to multiplication.

d. Use __shared__ instead to set q , v and m first, and 0-6 threads can do one for each __shared__ array. But 1024 Failed to pass the build when we did this change. So I think maybe we need to use global variables at this time, but there is no time to change.
e. When the number of thread is used to 1024, the program also exploded. I think that maybe we can plan to use the loop method and the maximum thread number is 512. Also, when it exceeds, the thread will be counted to the next one. However, there is no time to do the change, and maybe I can try it in the future.

2. If there are 4 GPUs instead of 2, what would you do to maximize the performance?

   With the current architecture of HW5, I can't figure out how to disperse the processing. If this is for HW2, HW3, and lab2 can be used, it may be my architecture problem this time.

3. If there are 5 gravity devices, is it necessary to simulate 5 n-body simulations independently for this problem?

   I guess it is not necessary, because the role should be similar. But to be honest, I don't know the answer clearly. In other words, I know how to sort and set the works, but the overall model is still in a state of confusion, which makes it impossible to optimize the entire process, and can only improve the specific part by parallel computing methods. This time, I appreciate TA help very much. The detailed explanation gave me a lot of hints. Thank you very much!