**What is missing?**

* After populate the array a,b,c with numbers, we have to copy it into the device memory before launching the kernel
* For the constant memory, we have to initialize it from the host using cudaMemcpyToSymbol
* The data size N is not defined anywhere. If it is large, we must be careful when we use the shared memory and const memory since we might go over the limit. Each particular device might have a limit on shared memory per block and the total available const memory. They can not be violated.