

Master's 1st year GPGPU's Dev

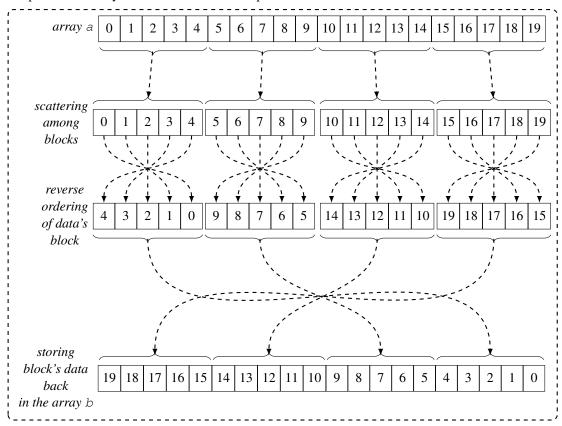
Examination — november 2020

Duration: 1h30 — Permitted access to documents

- 1 We want to write a **CUDA program** putting the elements of an array in **reverse order**:
- **10pts** \triangleright the array a contains *N* elements;

 - ▷ each block:
 - gets 256 associated elements of the array a;
 - performs the reverse ordering of these elements;
 - puts the reverse ordered elements back in the second array b.

Example with an array of 20 slots and 5 threads per block:



Questions:

- a. Are there some **constraints** on the number N of elements of the array a by using blocks of 256 threads, (1pt) with each thread associated to a slot of the array a?
 - Give the **definition of the grid** according to blocks of 256 threads and an array of N elements.
- b. Write a **CUDA program** performing the given algorithm. (4pts)
 - The array a resides in the host memory, its values are defined out of the scope of your program. the arry b resides in the host memory too.
- c. Now, we want store the **reverse ordered** data back in the array a: (1pt) Is this possible?
 - What precautions must be taken?
- d. Write a **CUDA program** that performs the reverse ordering of the data et puts back the results in the (4pts) array a.



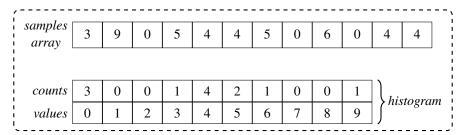


2- We have an array of samples of 10bits, *i.e.* between 0 and 1023 ($2^{10} = 1024$).

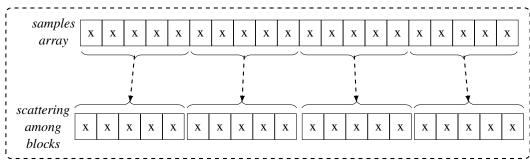
10pts

We want to compute the **histogram**, *i.e.* count the number of times each value is present:

- > traverse each slot of the array;



We want to scatter the array of samples among different blocks:



Questions:

- a. If each block manages a piece of the array of samples, how will we obtain the **histogram** for all the *(1pt)* samples of the array?
- b. What is the **work of each thread**? (1pt)

How much threads per block will be needed?

What is the **size** of the array histogram?

- c. Can the whole computation be performed **only on the GPU**? (1pt) Why?
- d. Give the **grid configuration** for an array of samples with a size of 65536 samples. (1pt)

(5pts)

- e. Write a **CUDA program** computing the histogram.
 - \diamond the array samples has a size of 65536, it resides in the host memory;
 - the arry histogram is on the host too and will be populated by the GPU.
- f. If the samples are now on 12bits, *i.e.* with a value between 0 and 4095 ($2^{12} = 4096$), is there **any** (1pt) **problems** with your provided solution? Why?