



POLITECNICO
MILANO 1863

GPUs and Heterogeneous Systems
(programming models and architectures)

5 Questions...

Question 1

Explain in few sentences which are the main challenges in accelerating histogram computation onto GPU and how they are addressed.

Question 2

In the following snippet of code, how many times will `foo()` and `bar()` be executed? Motivate the answer.

```
#pragma acc parallel num_gangs(64)
{
    foo();
    #pragma acc loop gang
    for (int i=0; i<n; i++) {
        bar(i);
    }
}
```

Question 3

Describe the benefits of the unified memory introduced in the Pascal architecture.

Question 4

Explain which are the benefits of the introduction of a unified shader in the Tesla architecture.

Question 5

Which is the efficiency of global load and store operations of the following CUDA kernel? Assume to have a block of 32 threads and to run the code on a Maxwell architecture where L1 cache has 32-byte access width. Motivate the answer.

```
__global__ void vsumKernel(char* a, char* b) {  
    int i = blockIdx.x * blockDim.x + threadIdx.x;  
    b[i] = a[(i+1)%blockDim.x];  
}
```