HPC Programming

Requirements

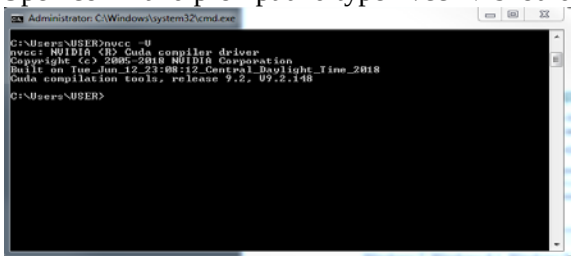
1). NVIDIA QUADRO Graphic Card , Graphic card used in this document is Quadro K5000
2). Download the Cuda toolkit from : https://developer.nvidia.com/cuda-toolkit-70
3). Visual Studio IDE

Introduction

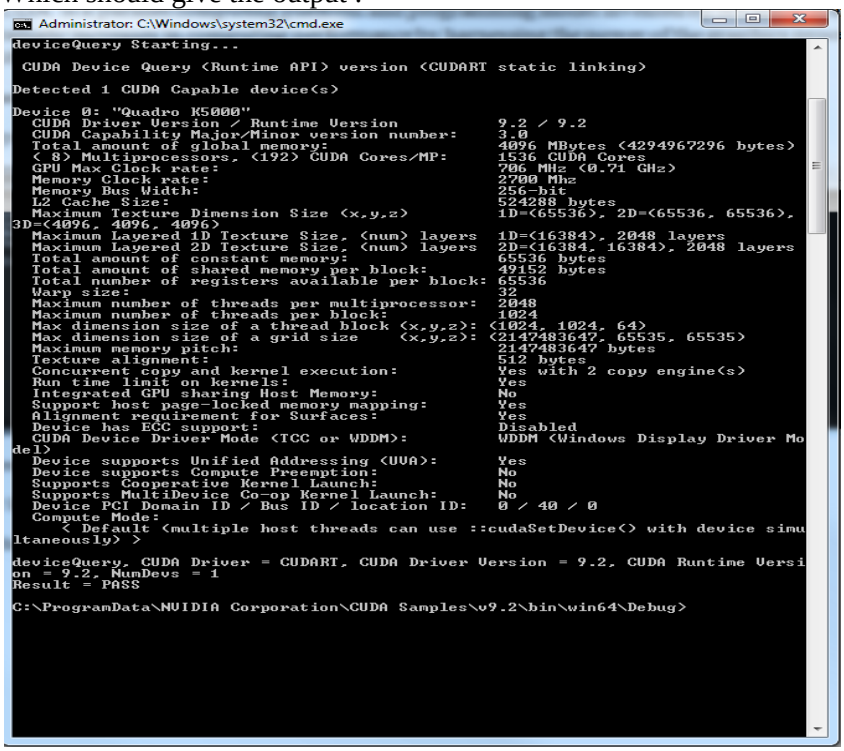
CUDA is a parallel computing platform and programming model invented by NVIDIA. It enables dramatic increases in computing performance by harnessing the power of the graphics processing unit (GPU).

Test installation

Open command prompt and type nvcc -V should give the version that has been installed.

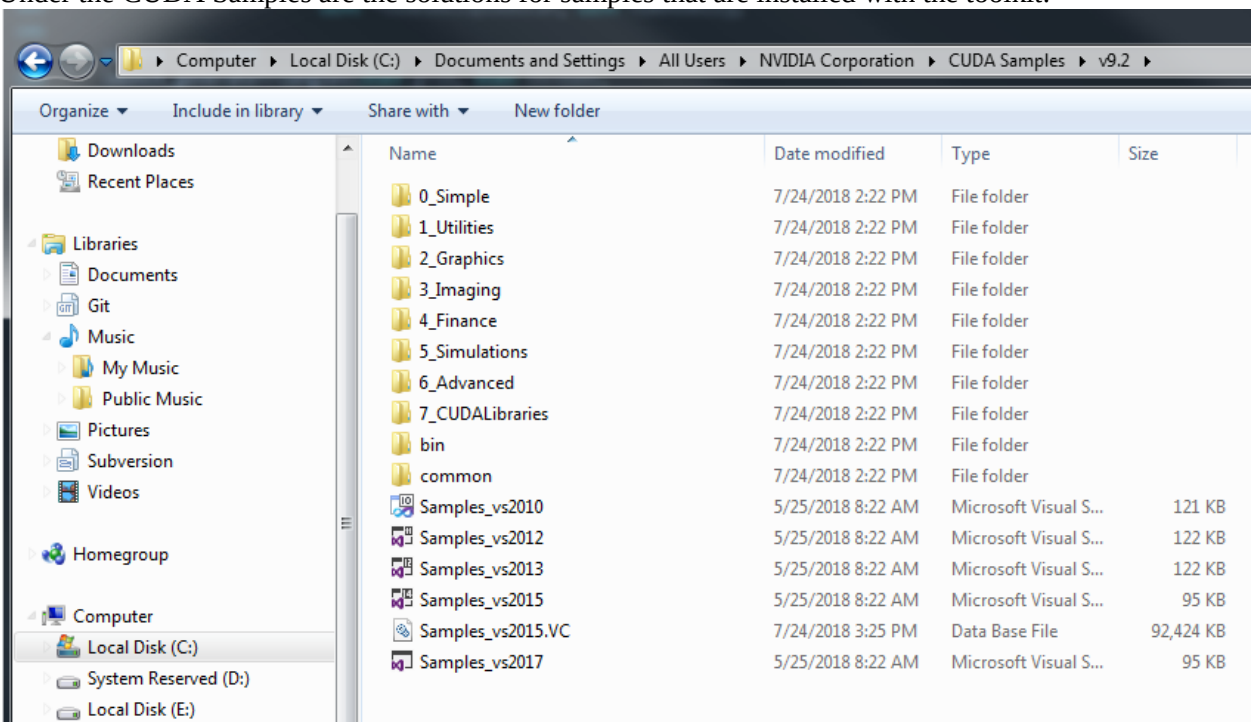


Test correct installation by navigating to :
C:\ProgramData\NVIDIA Corporation\CUDA Samples\v9.2\bin\win64\Debug
Execute deviceQuery
Which should give the output :

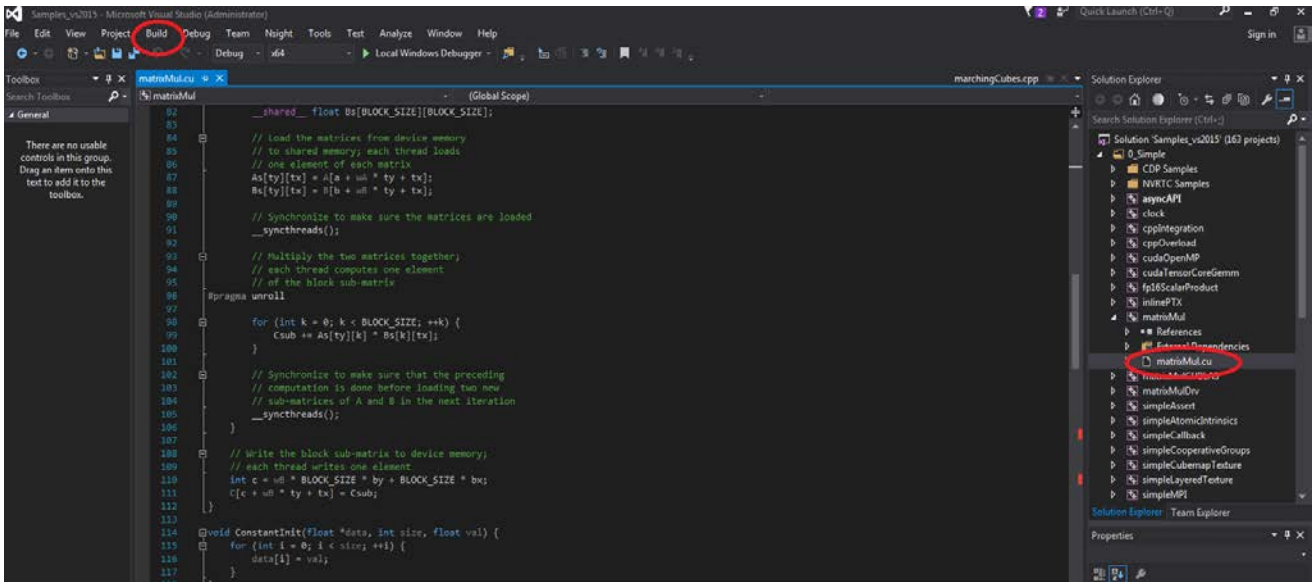


Developing using CUDA toolkit

1). Open the Visual studio IDE (I am using Visual Studio 2015)
Under the CUDA Samples are the solutions for samples that are installed with the toolkit:



2). Open the solution as per the Visual Studio IDE installed on your workstation (Samples_vs2015.VC)
3). Use the build to compile the solution as shown.



4). After compiling

Go to the debug directory of the solution and run i.e

C:\Documents and Settings\All Users\NVIDIA Corporation\CUDA Samples\v9.2\0_Simple\matrixMul\x64\Debug

(I am trying the sample on matrix multiplication)

5). Use the command prompt (Startmenu > CMD)to cd to file matrixMul.cu on the Debug directory

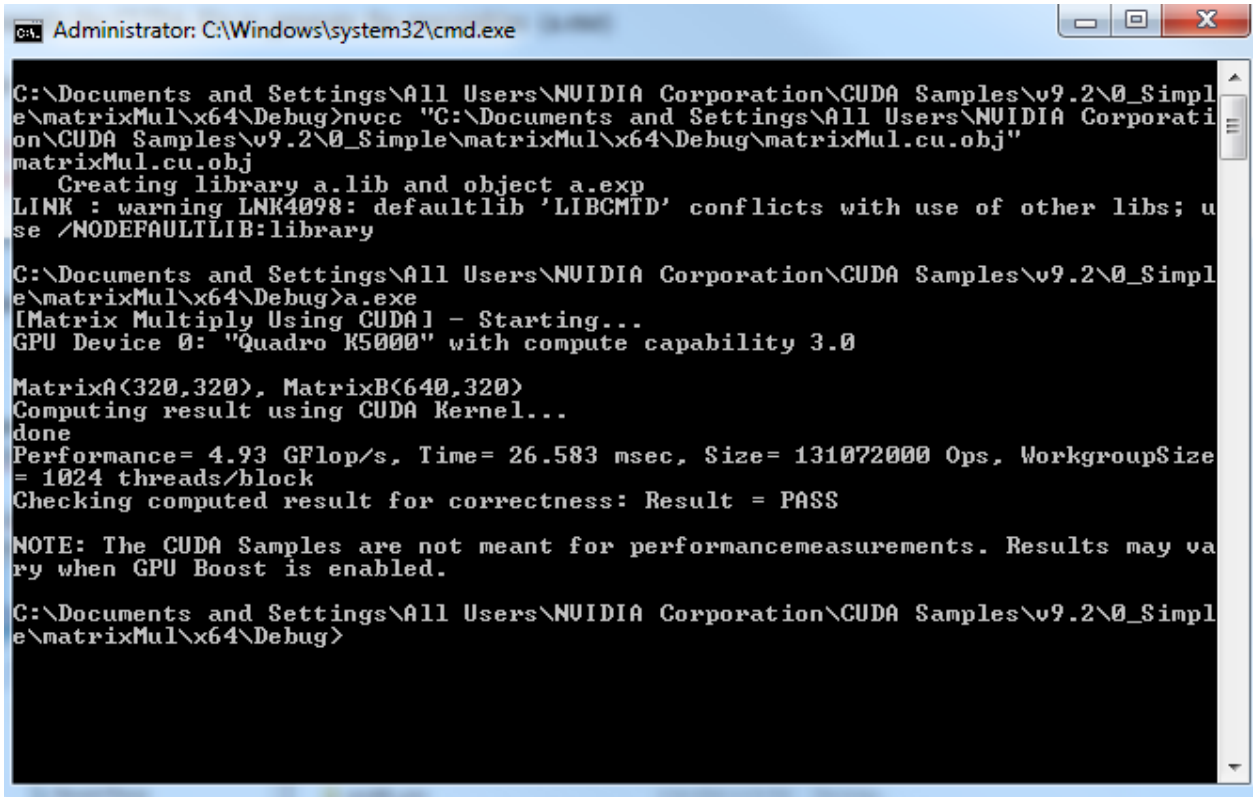
6). Compile the CUDA file to generate the executables (a.exe)

The syntax to compile is :

nvcc " C:\Documents and Settings\All Users\NVIDIA Corporation\CUDA Samples\v9.2\0_Simple\matrixMul\x64\Debug\ matrixMul.cu"

7). It will create the filed a.exe

Run a.exe to give you the output as shown:



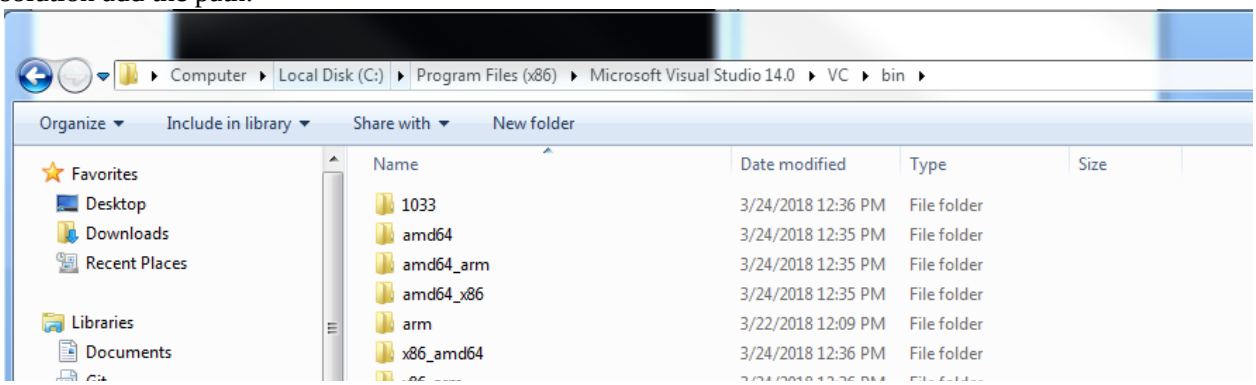
8.) In this case it is showing my Graphic card **Quadro K5000** , and the computation done by the GPU.

Task One - Create a CUDA code to read computation power:

Steps
1). Open visual studio > New Project > Cuda Project > Enter Name of the project
2). Type the code on the .cu file as below:

Errors encountered:
> `nvcc -o a.out matrixMul.cu`

1). nvcc fatal : Cannot find compiler 'cl.exe' in PATH
solution add the path:



to windows path environment.

cl is the c/c++ compiler as located in your linux or windows workstation

Programming using CUDA toolkit

1).Start Visual basic 2015.
2).New project
3). Select CUDA 9.2
4).This will create a file " .cu " where you write GPU specific code

Quick start on a simple hello world to execute on GPU

```
// includes, system
#include <stdio.h>
// includes CUDA Runtime
#include <cuda_runtime.h>
__global__
void kernel(void) {
}
int main(void) {
	kernel << <1, 1 >> >();
	printf("Hello, World!\n");
	getchar();
}
```

CUDA C keyword __global__ indicates that a function.
nvcc splits source file into host and device components — NVIDIA's compiler handles device functions like kernel() — Standard host compiler handles host functions like main()

Memory management on CUDA programs

Host and device memory are distinct entities, where
- Device pointers point to GPU memory which may be passed to and from host code and may not be dereferenced from host code.
- Host pointers point to CPU memory , may be passed to and from device code and may not be dereferenced from device code.

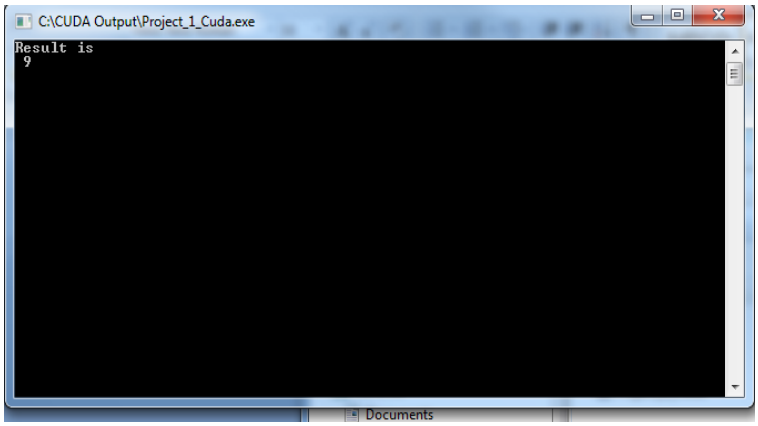
CUDA API for dealing with device memory :
cudaMalloc(), cudaFree(), cudaMemcpy()

I will illustrate with the simple addition code below:

```
// includes, system
#include <stdio.h>
// includes CUDA Runtime
#include <cuda_runtime.h>
__global__
void add(int *a, int *b, int *c) {
	*c = *a + *b;
}
int main(void) {
	//host copies of a, b, c
```

```
        int a, b, c;
    //device copies of a, b, c
        int *dev_a, *dev_b, *dev_c;
    //we need space for an integer
        int size = sizeof(int);
    //allocate device copies of a, b, c
        cudaMalloc((void**)&dev_a, size);
        cudaMalloc((void**)&dev_b, size);
        cudaMalloc((void**)&dev_c, size);
        a = 2;
        b = 7;
        //copy inputs to device
        cudaMemcpy(dev_a, &a, size, cudaMemcpyHostToDevice);
        cudaMemcpy(dev_b, &b, size, cudaMemcpyHostToDevice);
        //launch add() kernel on GPU, passing parameters
        add <<< 1, 1 >>>(dev_a, dev_b, dev_c);
        //copy device result back to host copy of c
        cudaMemcpy(&c, dev_c, size, cudaMemcpyDeviceToHost);
        cudaFree(dev_a);
        cudaFree(dev_b);
        cudaFree(dev_c);
        printf("Result is \n %d " , c);
        getchar();
        return 0;
}
```

Output above will be as below :



Parallel programming on CUDA programs

To be done

3d Max Rendering

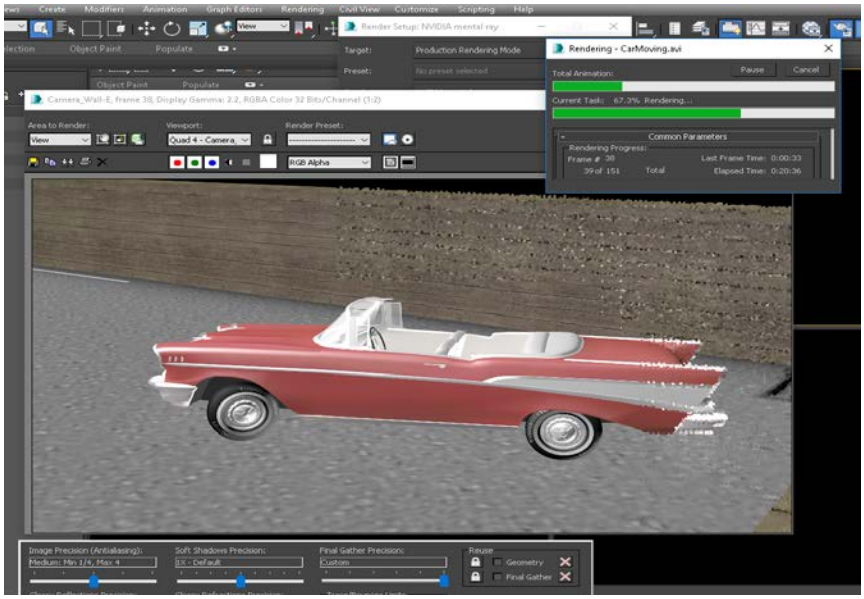
Introduction

This is aimed to make use of the GPU in rendering animation. NVIDIA Tesla K20Xm will be used.

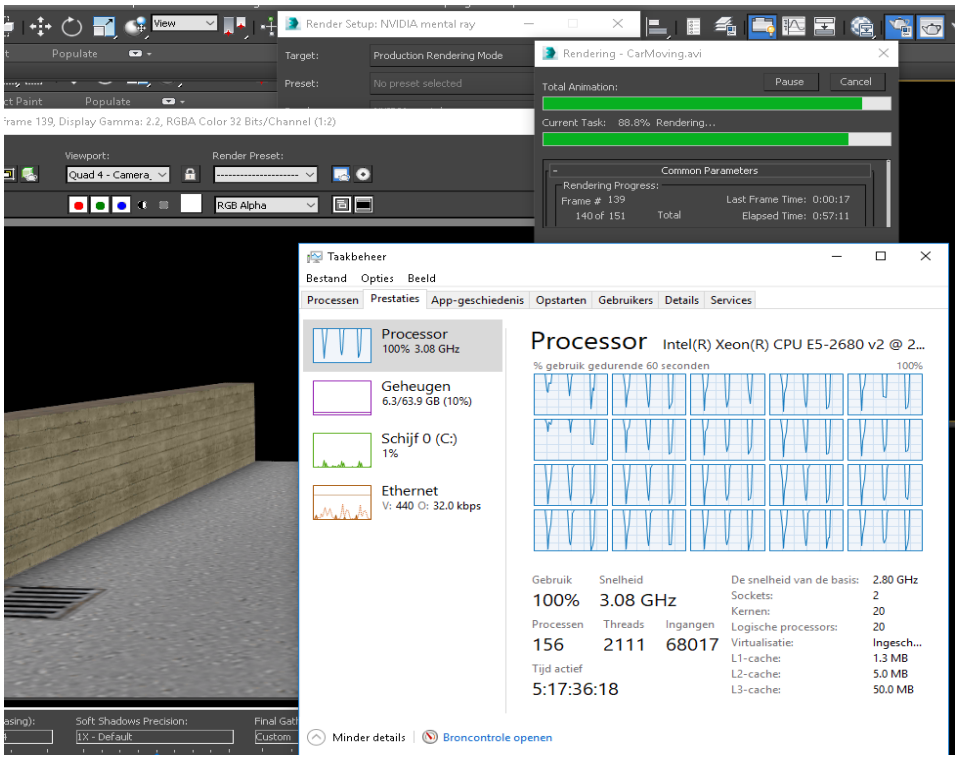
Test 1.

Rendering a car animation of 151 frames on dx360 M4 Server (20CPUs,65536MB RAM)on NVIDIA Tesla K20Xm.

a). Normal 3d max render time is as shown below :



The processor statistics of one of the dx360 M4 Server node as shown below (rendering via 3dmax):

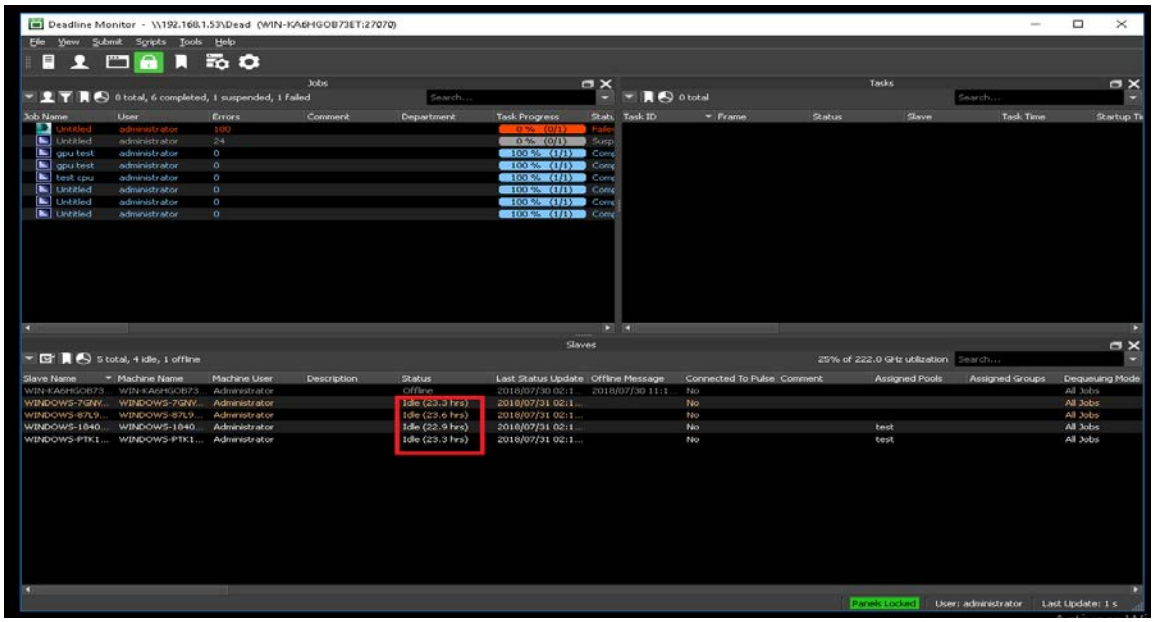


The time taken to render above is :

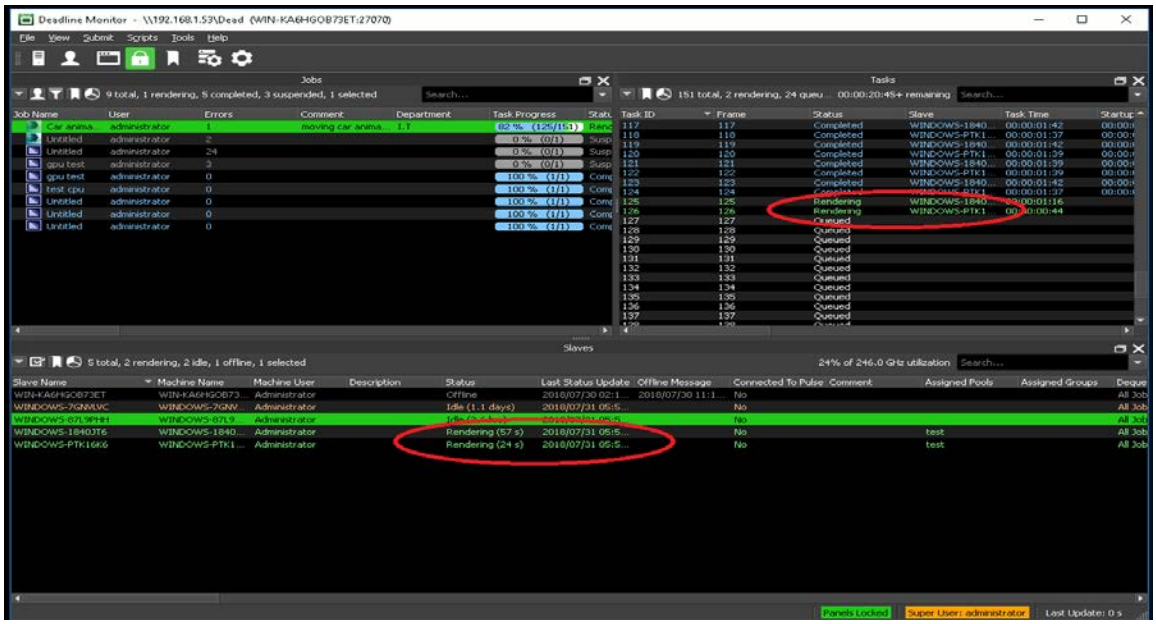
58minutes , the resulting file a **17.3 MB** video file.

b). Deadline rendering of same animation.

The same animation will now be submitted to the render farm with the four slaves.
All the slaves are as shown on the picture below:



Submitted 3dmax job is as shown:

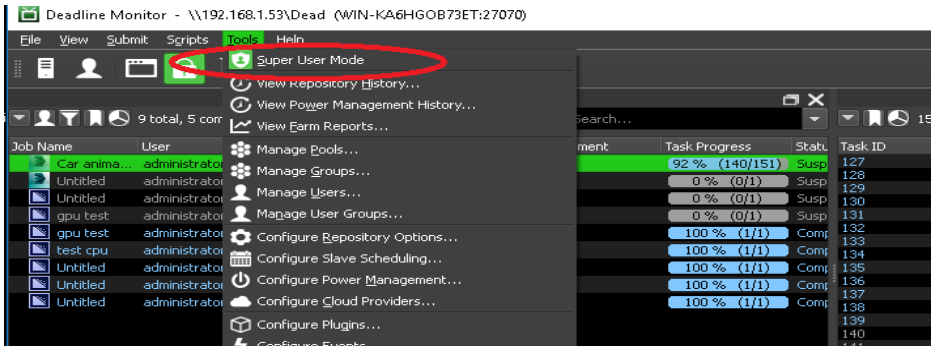


NB: Two nodes are the only being scheduled due to license

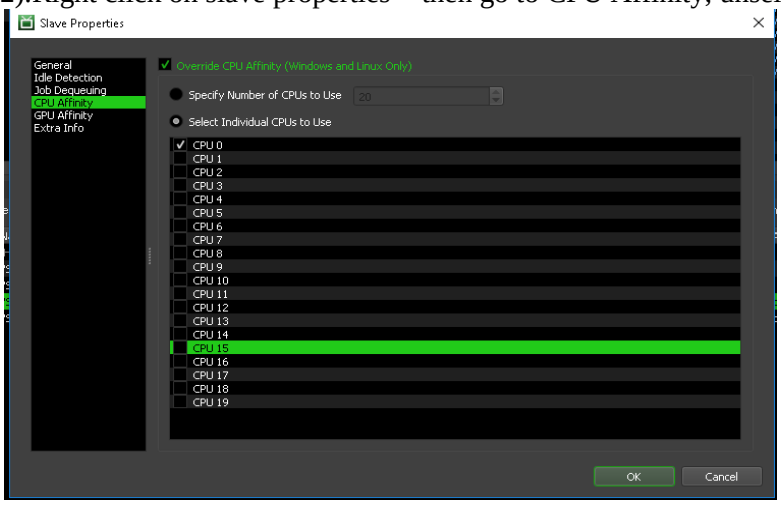
Configuring render farm to use the NVIDIA Tesla K20Xm

Below are the screenshots of the configuration:

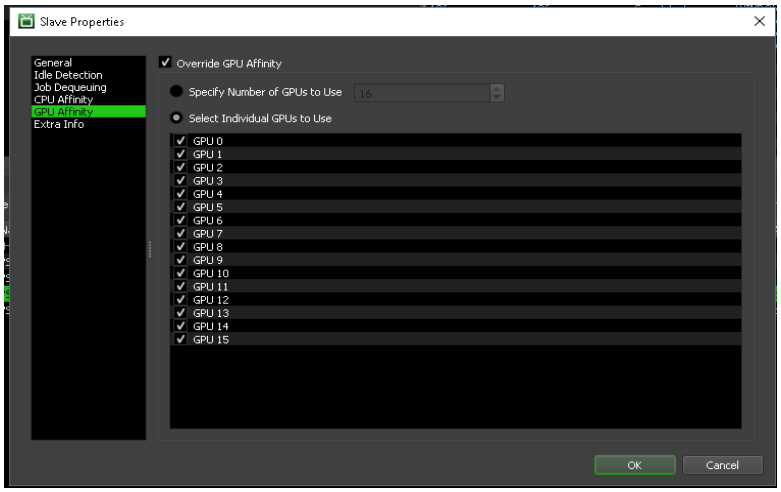
1). Ensure you are on super user mode.



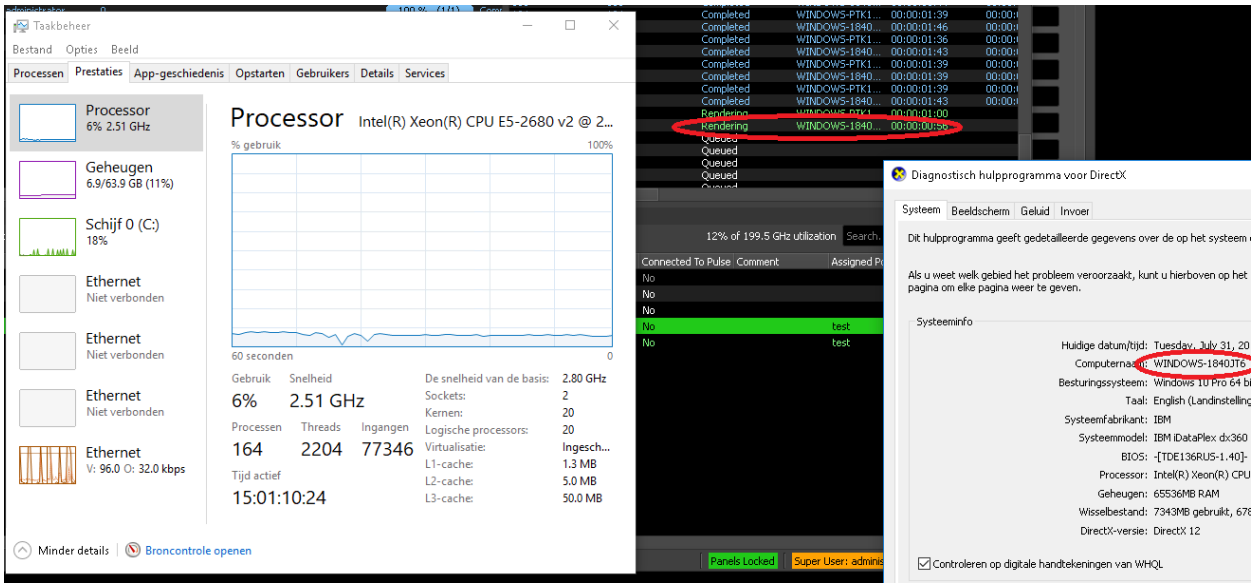
2).Right click on slave properties > then go to CPU Affinity, unselect all cpus



3).Go to GPU Affinity and select the GPU's of that slave that you want to use on the slave render as shown below:



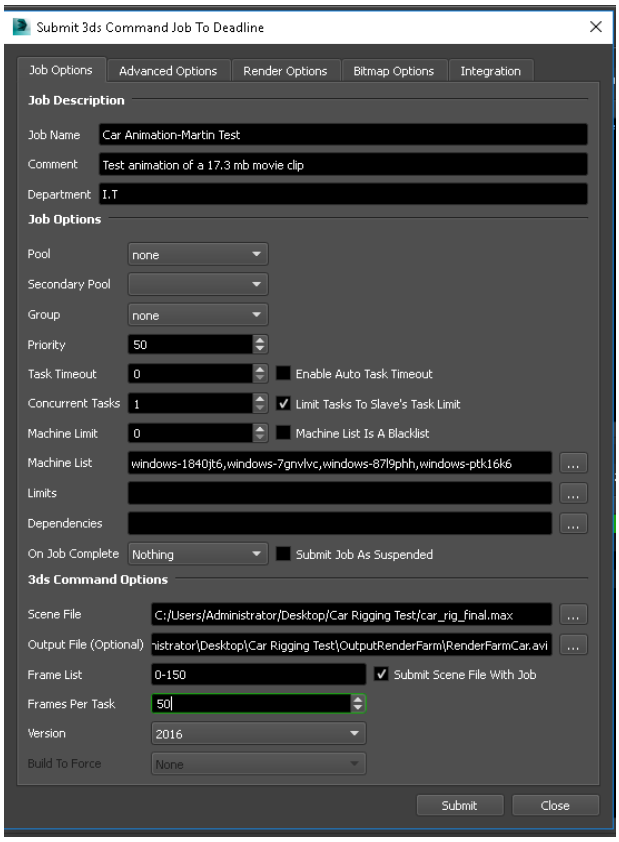
Comparing with the normal CPU render , below is the screenshot of the processor, on one of the m4server nodes.



Time taken to render the script is as shown below:

To be done.

Settings to submit the 3dmax file as shown below:



c). Application of CUDA Programming on render farm

On progress

Reference:

https://docs.nvidia.com/cuda/cuda-installation-guide-microsoft-windows/index.html