



# International Journal of Allied Practice, Research and Review

Website: [www.ijaprr.com](http://www.ijaprr.com) (ISSN 2350-1294)

## Nvidia Tesla The Personal Supercomputer

Sameer Ahmad<sup>1</sup>, Umer Amin<sup>2</sup>, Mr. Zubair M Paul<sup>3</sup>

<sup>1</sup>Student, Department of CSE& SSM College OF Engg.& Tech., India

<sup>2</sup>Student, Department of CSE& SSM College of Engg.& Tech., India

<sup>3</sup>Guide, Department of CSE& SSM College of Engg.& Tech., India

**Abstract:**-The NVIDIA® Tesla™ Personal Supercomputer is based on the new and important NVIDIA CUDA™ parallel computing architecture and powered by up to 900-960 parallel processing cores. The Tesla Personal Supercomputer is a desktop computer that is supported by NVIDIA Corporation. This Tesla Personal Computer is built by Dell, Lenovo and other companies. It has the ability to perform computations 250 times faster than a multi-CPU core PC or workstation. The CUDA™ parallel computing architecture enables developers to utilize C programming with NVIDIA GPUs to run the most complex computationally-intensive applications. CUDA is easy to learn and has become widely adopted by thousands of application developers worldwide to accelerate the most performance demanding applications. In November 2006 NVIDIA'S tesla was introduced in the NVIDIA GeForce 8800 GPU. This helps in parallel computing applications written in C language and thereby enabling high performance. The Tesla Graphics and computing architecture is available in two different series which are GeForce 8-series GPU's and Quadro GPU's for laptops, desktops, workstations, and servers. It also provides the processing architecture for the Tesla GPU computing platforms introduced in 2007 for high-performance computing

**Keywords:** NVIDIA; GPU; CUDA; Supercomputer; Tesla Graphics.

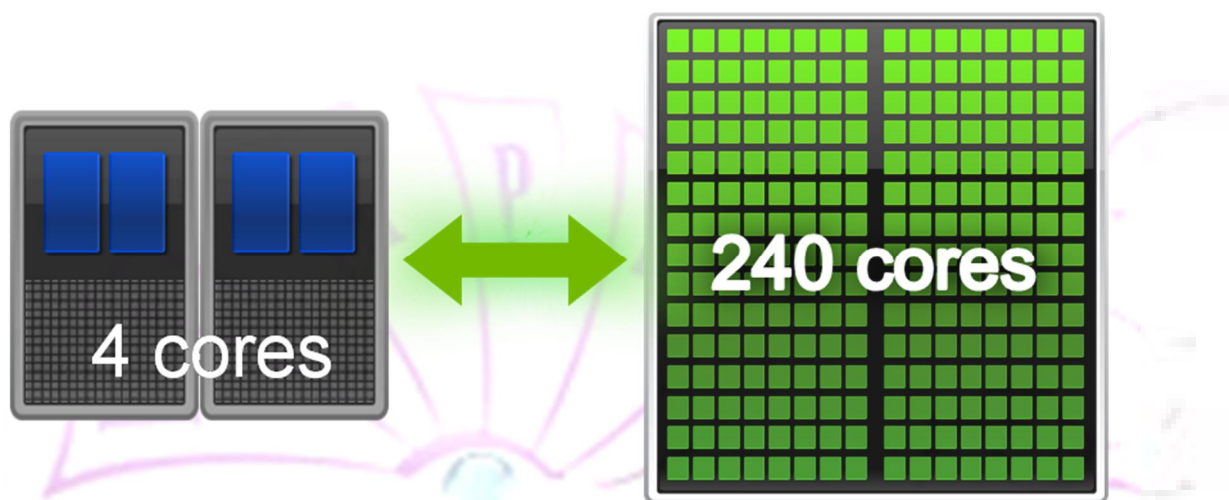
### I. INTRODUCTION

The Tesla Personal Supercomputer is a desktop computer that is backed by Nvidia and was built by Dell, Lenovo and other big companies. The main aim of Tesla Personal Supercomputer to explore the capabilities of Nvidia's Tesla GPGPU brand; it uses NVIDIA's CUDA parallel computing architecture. It is powered by 960 parallel processing cores which gives it enormous computing capability and allows it to achieve a performance up to 250 times faster than standard PCs, as per Nvidia. There are upto 3-4 Nvidia Tesla C1A60 computing processors in the heart of Tesla Personal Supercomputer, which appear similar to other high-performance Nvidia graphics cards, but without any video output ports. Each Tesla C1060

has 240 streaming processor cores running at 1.296 GHz, 4 GB of 800 MHz 512-bit GDDR3 memory and a PCI Express x16 system interface. While typically using only 160-watts of power, each card is capable of 933 GFlops of single precision floating point performance or 78 GFlops of double precision floating point performance.

## II. GPU COMPUTING

GPU computing is the use of a GPU (GPU Stands for graphics processing unit) to do general purpose scientific and engineering computing. GPU computing is to use a CPU and GPU together in a heterogeneous computing model. The sequential part of the application runs on the CPU and the computationally-intensive part runs on the GPU. From the users point of view, the application just runs faster because it is using the High-performance of the GPU which boots its performance.



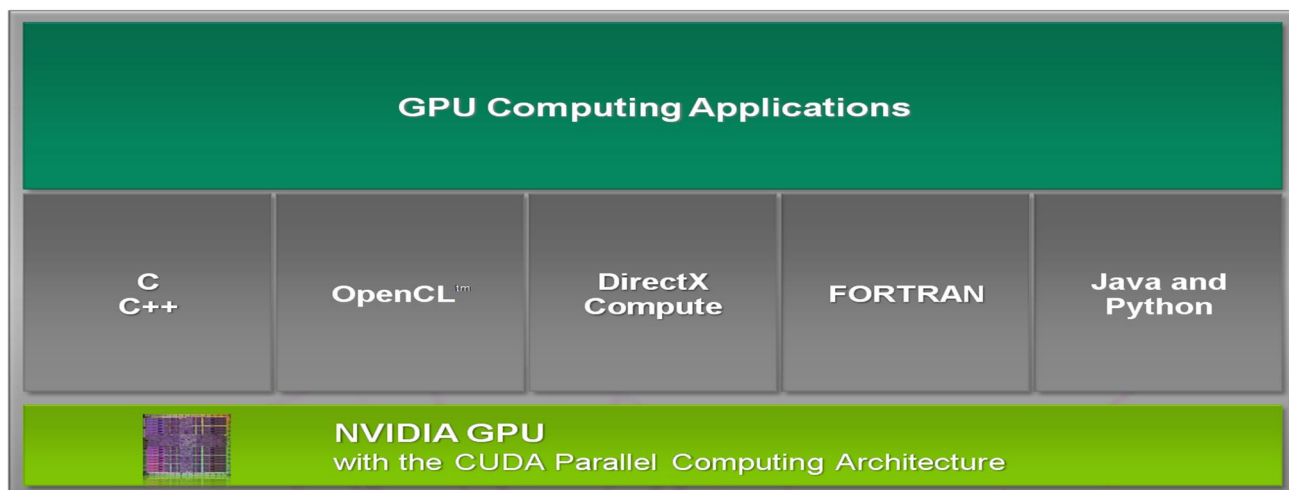
**Fig.1 CPU & GPU Cores**

The developer has to modify their application to take the compute-intensive kernels and map them to the GPU. The rest of the application remains on the CPU. Mapping a function to the GPU involves rewriting the function to expose the parallelism in the function and adding keywords to move data to and from the GPU. GPU computing is enabled by the massively parallel architecture of NVIDIA™s GPUs called the CUDA architecture. The CUDA architecture consists of 100s of processor cores that operate together to crunch through the data set in the application. The Tesla 10-series GPU is the second generation CUDA Architecture

With features optimized for scientific applications such as IEEE standard double precision floating point hardware support, local data caches in the form of shared memory dispersed throughout the GPU, coalesced memory accesses and so on.

### III. CUDA PARALLEL ARCHITECTURE AND PROGRAMMING MODEL

The CUDA (Compute Unified Device Architecture) parallel hardware architecture is accompanied by the CUDA parallel programming model developed by NVIDIA that provides a set of abstractions that enable expressing fine-grained and coarse-grain data and task parallelism. The programmer uses parallelism in high-level languages such as C, C++, Fortran or driver APIs such as OpenCL and DirectX 11 Compute.



**Fig.2 CUDA Parallel Architecture and Programming Model**

The first supported language for parallel programming by NVIDIA was the C language. A set of C for CUDA software development tools enable the GPU to be programmed using C with a minimal set of keywords or extensions. Support for Fortran, OpenCL, et cetera will follow soon.

The CUDA parallel programming model guides programmers to partition the problem into coarse sub-problems that can be solved independently in parallel. Fine grain parallelism in the sub-problems is then expressed such that each sub-problem can be solved cooperatively in parallel.

The CUDA GPU architecture and the corresponding CUDA parallel computing model are now widely deployed with 100s of applications and nearly a 1000 published research papers. CUDA Zone lists many of these applications and papers.

### IV. PROCESSOR OF NVIDIA TESLA

#### Tesla C1060 Computing Processor

At the heart of the new Tesla personal supercomputer are three or four NvidiaTesla C1060 computing processors, which appear similar to a high-performance Nvidia graphics card, but without any video output ports. Each Tesla C1060 has 240 streaming processor cores running at 1.296 GHz, 4 GB of 800 MHz 512-bit GDDR3 memory and a PCI Express x16 system interface. While typically

using only 160-watts of power, each card is capable of 933 GFlops of single precision floating point performance or 78 GFlops of double precision floating point performance.

## **KEY FEATURES**

### **GPU**

- Number of processor cores: 240
- Processor core clock: 1.296 GHz
- Voltage: 1.1875 V
- Package size: 45.0 mm × 45.0 mm 2236-pin flip-chip ball grid array (FCBGA)

### **Memory**

- 800 MHz
- 512-bit memory interface
- 4 GB: Thirty-two pieces 32M × 32 GDDR3 136-pin BGA, SDRAM

### **External Connectors**

- None

### **Internal Connectors and Headers**

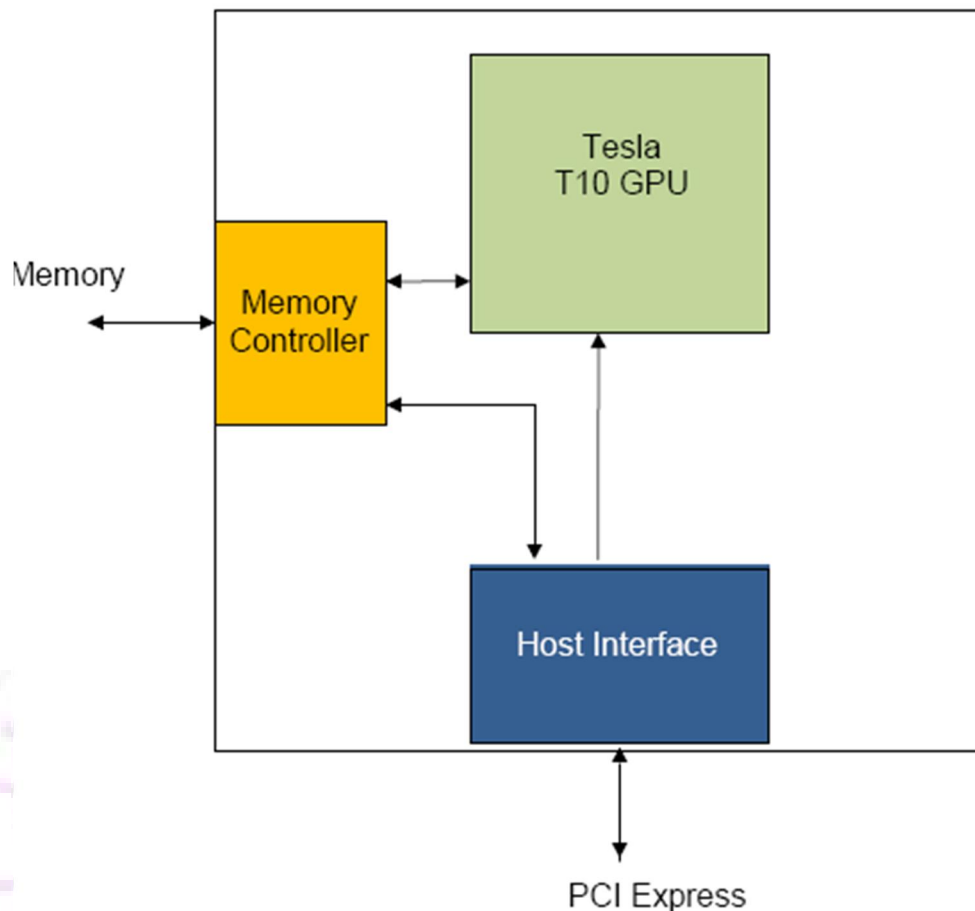
- One 6-pin PCI Express power connector
- One 8-pin PCI Express power connector
- 4-pin fan connector

## **V. COMPUTING PROCESSOR DESCRIPTION**

NVIDIA Tesla C1060 a work station into high-performance computing machine.

NVIDIA Tesla gives a dedicated computing resource to technical professionals at their desk-side Which has high efficiency and is much faster than a shared cluster in data center which is achieved with the help of parallel computing architecture of Nvidia Tesla C1060

## Tesla C1060 Computing Processor



**Fig.3 Tesla C1060 Block Diagram**

## VI. TECHNICAL SPECIFICATIONS

### A. Tesla Architecture

- ✓ 240 scalar processor cores per GPU
- ✓ performance
- ✓ Ultra-fast GPU memory access with 102 GB/s peak bandwidth per GPU
- ✓ IEEE 754 single-precision and double-precision floating point
- ✓ Integer, single-precision and double-precision floating point operations
- ✓ Hardware Thread Execution Manager enables thousands of concurrent threads per GPU

- ✓ Parallel shared memory enables processor cores to collaborate on shared information at local cache
- ✓ Massively-parallel many-core architecture
- ✓ Each Tesla C1060 GPU delivers 933 GFlops Single Precision and 78 GFlops Double Precision performance

### ***B. Supported Platforms***

- ✓ Microsoft® Windows® XP 64-bit and 32-bit (64-bit recommended)
- ✓ Linux® 64-bit and 32-bit (64-bit recommended)
- ✓ Red Hat Enterprise Linux 4 and 5
- ✓ SUSE 10.1, 10.2 and 10.3

### ***C. Software Development Tools***

- ✓ C language compiler, debugger, profiler, and emulation mode for debugging
- ✓ Standard numerical libraries for FFT (Fast Fourier Transform), BLAS (Basic Linear Algebra Subroutines), and CuDPP (CUDA Data Parallel Primitives).

## **VI. CONCLUSION**

Our paper on Tesla supercomputer shows the basic information about GPU computing and parallel Processing. The NVIDIA's Tesla computer could prove invaluable to medical researchers and accelerate the discovery of cancer treatments.

Scientists believe the new systems could help find cures for diseases. The high-performance processor could become invaluable to universities and medical institutions. The Tesla Personal Supercomputer doesn't make supercomputing clusters obsolete but it's a major breakthrough for millions of researchers who can take advantage of the huge heterogeneous computing power of this system. NVIDIA Tesla Personal Supercomputers are available today from the following firms: AMAX (US), Armari (UK), Asus (WW), AzkenMuga (ES), Boxx (US), Comptronic (DE), Concordia (IT), Connoisseur (IN), Dell (WW), Dospara (JP), E-Quattro (IT), JRTI (US), Lenovo (WW), Littlebit (CH), Meijin (RU).

## **VII. ACKNOWLEDGEMENT**

It gives us great pleasure to present our paper on "Nvidia Tesla (personal supercomputer)". No work, however big or small, has ever been done without the contribution of others.

It would be a great pleasure to write few words, which would although not suffice as the acknowledgement of this long cherished effort, but in the absence of which this report would necessarily be

incomplete. So these words of acknowledgement come as a small gesture of gratitude towards all those people, without whom the successful completion of this Paper wouldn't have been possible.

Last but not least I would like to thank our Head of department Mrs. Yasmeen and the Department of Computer Science Engineering for providing us with the facilities and the staff members of such a caliber.

## **VII. Book References**

- [1]. Digit
- [2]. Chip
- [3]. Foley . J. Vandam ,Feiners-Computer Graphics & Principles , Addison Wesley
- [4]. Akenine-Moller .T , Hanes . E , Real Time Rendering – 2<sup>nd</sup> Edition

## **VIII. Web References**

- [1]. [www.nvidia.com](http://www.nvidia.com)
- [2]. [www.wikipedia.org](http://www.wikipedia.org)
- [3]. [www.tomshardware.com](http://www.tomshardware.com)
- [4]. [www.dvhardware.net](http://www.dvhardware.net)