Qiusi Shen

CS 35L

004749315

Instructor Isha Verma

# Analysis on advanced GPU

In the 2017 GPU Technology Conference in the San Jose, the new NVIDIA Tesla V100 which is also called Volta was introduced by the CEO Jen Hun Huang. By now it is the most powerful Graphic processer unit accelerator that has ever been built. This new GPU Architecture will bring AI more convenient to every industry. AI exists in all the fields from healthcare, big data analysis and the supercomputing. This is the fastest growing industry and this newest and most powerful graphic processor unit will satisfy all the requirement and create a perfect platform for all the highest technology performance.

It has more features. New streaming Multiprocessor Architecture Optimized for deep learning. The new generation of the design gains more energy efficiency and it also enabling major boost in FP32 and FP64. The new parallel threading will increase the calculation speed. The new L1 Data cache improves the performance. HBM2 Memory make the memory subsystem deliver speed increased and the it provides a 1.5 times delivered memory bandwidth. Also, it improves the Enhanced Unified Memory and Address Translation Service

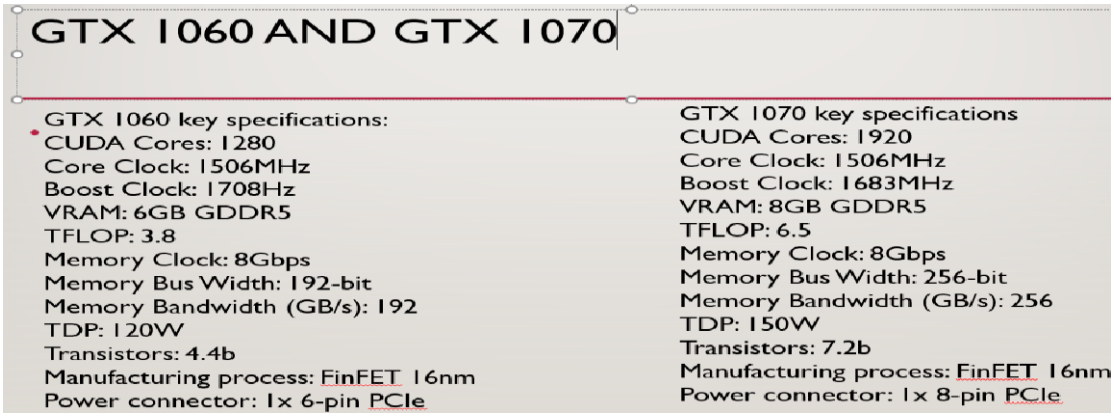
and the Cooperative Groups and new cooperative Launch APIs. The new Volta also allow the GPU to access the CPU page table directly.

The Graphic Processor Unit is not strange to most of us. All the computers have one Graphic Processor Unit to display on the monitor. Graphic Processor Unit which is also shorted for GPU is an electronic circuit designed to rapidly manipulate and alter memory to accelerate the creation of images in a frame buffer intended for output to display device by the definition. The first Graphic Processor Unit was built in 1970. It is an arcade system boards that used to accelerate the drawing sprite graphic. During that time, the random-access memory buffer was extremely expensive so the chips were to help display on the monitor.

Modern Graphic Processor Unit Manufactory are dominated by Nvidia, AMD, and Intel. The Nvidia has the market share of 27.8% and the AMD is 20.6% while the Intel has the most market share of 49.4%.

In the 2016 the Nvidia which is the most popular Graphic Processor Unit company release the newest 10 series card. The most popular ones are GTX 1060, GTX 1070, and GTX 1080. The GTX is means to have higher performance than GT which means the standard Version. The first two digits means the series number which is 10 and if it is in the nine series it could be GTX 980m. The number after the series number is for the performance purpose. The larger number it has the better performance it has. For example, the GTX 1070 has a better performance

that GTX 1060. And the letter m stands for the mobile and it is for the laptop to be exactly. If it has no m then it is for desktop. Usually the desktop version has better performance than the mobile one. But since the 10 series the performance difference has shrink.




From this form, the GTX1070 has more cores and higher Boost clock, larger Virtual RAM and higher Memory Bus Width. As a result, the GTX 1070 has nearly 40% better performance than that of GTX 1060 but it also has more TDP power value which means it produce more heat and noise which is bad for the laptop performance.

However, in 2017 the Nvidia Company came up with a new MAXQ technology. Which sacrifices small part of the performance to reduce heat and noise. This modern technology will make the laptop thinner. It will produce the thinnest, fastest, quietest gaming laptop.

Work Cited

https://www.nvidia.com/en-us/geforce/products/10series/laptops/max-q/

Inside Volta: The World's Most Advanced Data Center GPU

https://devblogs.nvidia.com/parallelforall/inside-volta/

GEFORCE® GTX WITH MAX-Q DESIGN TRANSFORMING GAMING LAPTOPS