

Computational GPU Unveiled



SANTA CLARA, Calif. — NVIDIA Corp. today introduced its next generation CUDA GPU architecture, codenamed “Fermi”. An entirely new ground-up design, the “Fermi” architecture is the foundation for the world’s first computational graphics processing units (GPUs), delivering breakthroughs in both graphics and GPU computing.

“NVIDIA and the Fermi team have taken a giant step towards making GPUs attractive for a broader class of programs,” said Dave Patterson, director Parallel Computing Research Laboratory, U.C. Berkeley and co-author of *Computer Architecture: A Quantitative Approach*. “I believe history will record Fermi as a significant milestone.”

Presented at the company’s inaugural GPU Technology Conference, in San Jose, California, “Fermi” delivers a feature set that accelerates performance on a wider array of computational applications than ever before. Joining NVIDIA’s press conference was Oak Ridge National Laboratory who announced plans for a new supercomputer that will use NVIDIA GPUs based on the “Fermi” architecture. “Fermi” also garnered the support of leading organizations including Bloomberg, Cray, Dell, HP, IBM and Microsoft.

“It is completely clear that GPUs are now general purpose parallel computing processors with amazing graphics, and not just graphics chips anymore,” said Jen-Hsun Huang, co-founder and CEO of NVIDIA. “The Fermi architecture, the integrated tools, libraries and engines are the direct results of the insights we have gained from working with thousands of CUDA developers around the world. We will look back in the coming years and see that Fermi started the new GPU industry.”

As the foundation for NVIDIA’s family of next generation GPUs namely GeForce, Quadro and Tesla - “Fermi” features a host of new technologies that are “must-have” features for the computing space, including:

Computational GPU Unveiled

Published on Electronic Component News (<http://www.ecnmag.com>)

- C++, complementing existing support for C, Fortran, Java, Python, OpenCL and DirectCompute.
ECC, a critical requirement for datacenters and supercomputing centers deploying GPUs on a large scale
- 512 CUDA Cores featuring the new IEEE 754-2008 floating-point standard, surpassing even the most advanced CPUs
- 8x the peak double precision arithmetic performance over NVIDIA's last generation GPU. Double precision is critical for high-performance computing (HPC) applications such as linear algebra, numerical simulation, and quantum chemistry
- Support for concurrent-kernel execution, where different kernels of the same application context can execute on the GPU at the same time (eg: PhysX fluid and rigid body solvers).
- NVIDIA Parallel DataCache - the world's first true cache hierarchy in a GPU that speeds up algorithms such as physics solvers, raytracing, and sparse matrix multiplication where data addresses are not known beforehand
- Nexus - the world's first fully integrated heterogeneous computing application development environment within Microsoft Visual Studio

More on "Fermi" can all be found at: www.nvidia.com/fermi [1]

Source URL (retrieved on 08/04/2015 - 10:01am):

<http://www.ecnmag.com/product-releases/2009/10/computational-gpu-unveiled>

Links:

[1] <http://www.nvidia.com/fermi>