**On the cost of a general GPU framework**

The stange case of CUDA 4.0 vs. CUDA 5.0

Mattew Wezowicz and Michela Taufer

University of Delaware

Newark, Delaware, 19716

El autor menciona que CUDA fue lanzado en el 2007, y desde entonces nuevas versiones fueron surgiendo. Dice que alcazó su máximo rendimiento en la versión 4.0. Desde esa versión, los diseñadores de NVIDIA an buscado re-diseñar el framework desde el punto de vista de software engineering. Mattew comenta que esto tiene un impacto significativo en los costos de manteniemiento y en la portabildiad.

La pregunta de investigación del autor es: ¿En la comnunidad HPC que se ha beneficiado de CUDA y programación GPU está dispuesta a pagar el costo?

Los autores utilizan FENZI para hacer el análisis comparativo. FENZI genera simulaciones en el GPU . Dichas simulaciones son ejecutadas en un solo GPU. En la prueba se calcula el campo de fuerza CHARMM y la interacción de la distancia electrostática con el método de suma de Ewald.

Las optimizaciones incluyen el remplazo de ramas con máscaras y el remplazo de operaciones en caracteres e INTs. Consideraron tres distintas picapas de grasa (DMPC) con las mismas estructuras pero diferentes tamaños.

En los experimentos compararon el rendimiento del código FENZI en términos de ns/día para los tres DMPC compilados con cuatro variantes de CUDA y ejecutados en GPUs GTX480 y C2050. Se repitió 20 veces el experimento y se presenta el promedio.

Los autores encontraron que la versión de CUDA que presenta un mejor rendimiento es la 4.0 siempre. Contra la versión CUDA 4.1, 4.2 y 5.0. Después optimizaron el código de FENZI para seguir los lineamientos de NVIDIA y repitieron el análisis. En este caso, la versión 5.0 fue mejor que todas; no obstante, no fue mejor que la versión 4.0 para el código sin obtimizar.

El autor concluye que se observa una afectación en el rendimiento de las aplicaciones actuales debido a los cambios de versión en CUDA. Menciona que la pérdida de rendimiento no puede ser recuperada sin optimizaciones que demandan tiempo.

**Bibliografía**

Wezowicz, M., & Taufer, M. (2012, November). On the Cost of a General GPU Framework: The Strange Case of CUDA 4.0 vs. CUDA 5.0. In *High Performance Computing, Networking, Storage and Analysis (SCC), 2012 SC Companion:* (pp. 1535-1536). IEEE.