
Cloud, HPC & Bio

Laboratorio 6: GPUs con Python

Informe de valoración de prácticas.

En este laboratorio se continuó el trabajo iniciado en prácticas anteriores, centrándose en el uso de aceleradores hardware, concretamente GPUs, para aplicaciones de cálculo científico en Python dentro de un entorno HPC gestionado por SLURM. A lo largo de la práctica se exploró el uso de distintas bibliotecas, como CuPy, Numba y PyTorch, con el objetivo de analizar las mejoras de rendimiento obtenidas al ejecutar código en GPU y comprender el impacto de la transferencia de datos entre CPU y GPU. Asimismo, se reforzó el uso del control de versiones con Git y la ejecución de notebooks de forma automatizada mediante scripts batch.

Entre los aspectos positivos del laboratorio, destaca la facilidad con la que bibliotecas como CuPy permiten adaptar código originalmente escrito con NumPy para su ejecución en GPU, obteniendo mejoras significativas en el tiempo de ejecución. El uso de Numba permitió además comprender el modelo de ejecución en GPU mediante funciones vectorizadas, así como las limitaciones asociadas a operaciones de reducción. Por otro lado, la práctica facilitó una mejor comprensión de la importancia de minimizar las copias de datos entre CPU y GPU, aspecto clave para obtener un rendimiento óptimo en aplicaciones aceleradas por hardware.

Como principal desventaja, la práctica presenta ciertas dificultades operativas propias del entorno HPC. En particular, la ejecución de los trabajos con GPU requiere lanzarlos desde la cola bohr, de lo contrario, los trabajos no se ejecutan correctamente. Este requisito puede no resultar evidente inicialmente y puede generar errores difíciles de diagnosticar. Además, aunque PyTorch ofrece una interfaz muy potente y flexible para el uso de GPUs, su ejecución está limitada por la compatibilidad de la versión instalada con el hardware disponible, lo que obliga a ejecutar los notebooks únicamente en nodos específicos del clúster.