

Laboratorio 3

PYTHON

El tercer laboratorio se centra en el uso de Python para cálculos científicos de alto rendimiento, especialmente en aplicaciones bioinformáticas. Se destaca el uso de Jupyter notebook, Numpy, Scipy, Numba y CuPy para mejorar la eficiencia del código. Aunque Python ha sido considerado tradicionalmente como un lenguaje "lento", el uso de paquetes específicos y técnicas de paralelización ha cambiado esta percepción.

Los objetivos del laboratorio incluyen el uso de Jupyter notebook para programar en el cluster de GACOP, evaluar el rendimiento del código, escribir código más rápido con NumPy, manejar multiprocessing, optimizar con Numba y utilizar la GPU con CuPy.

En la primera parte, se trabaja con el notebook "primes-exercise.ipynb", realizando ejercicios que incluyen el uso de funciones mágicas de IPython, Numba para aceleración, y la implementación de paralelismo. Además, se crea un script para ejecutar el notebook en colas específicas del cluster y se evalúa el rendimiento.

La segunda parte se enfoca en el cálculo del número π mediante el algoritmo de Monte-Carlo. Se proporciona un notebook llamado "pi.ipynb" con implementaciones en Python, Numpy y Numba. Se realizan mejoras en el código y se evalúa el rendimiento para diferentes tamaños de problema. Se crea un script para medir el tiempo de ejecución con SLURM.

Como actividad extra, se sugiere paralelizar la aplicación del cálculo de π de tres formas diferentes: con Numba para usar varios núcleos de la CPU, con multiprocessing para aprovechar varios núcleos de la CPU, y con CuPy para utilizar una GPU.

Desde mi perspectiva personal, encuentro que este laboratorio es esencial para el desarrollo profesional debido a que proporciona herramientas excepcionales que mejoran significativamente la eficiencia temporal, un factor crucial para potenciar el rendimiento en cálculos que tienden a ser lentos.