

INFORME LABORATORIO 5 – BELÉN SALAR

EXPLOTANDO EL PARALELISMO CON PYTHON EN CPUs

Con este laboratorio he podido profundizar todavía más en Python, centrado en el paralelismo en CPUs usando los paquetes multiprocessing y Numba. Además, la práctica me ha resultado muy interesante, ya que ayuda a comprender mejor cómo aprovechar los recursos disponibles en un servidor para reducir los tiempos de ejecución de programas complejos.

He aprendido aspectos relevantes como diferenciar de manera práctica el rendimiento entre un código secuencial y versiones de ese código optimizadas y paralelizadas.

Durante la realización de la práctica me surgieron diversas dudas como comprender por qué el uso de múltiples núcleos no siempre conduce a una reducción del tiempo de ejecución. Por ejemplo, cuando al combinar Numba con multiprocessing observaba que el tiempo aumentaba respecto a la versión secuencial optimizada con Numba. Además, también tuve problemas con la ejecución en el clúster con SLURM a la hora de gestionar los núcleos asignados al job y su uso dentro del script. Por otro lado, me resultó un tanto costoso saber en la salida de SLURM, que líneas del fichero correspondían con cada parte del código del notebook. Ya que cuando se ejecutaba con un solo núcleo, había repeticiones de ejecución y me condujo a confusión en comparación con los otros casos donde se usaban más núcleos.

Como aspecto positivo del laboratorio es que me resulta muy valioso el aprender tantas cosas del lenguaje de programación Python y no centrarnos únicamente en el código original. Me ha permitido adquirir una visión más realista de cómo optimizar programas científicos en Python y valorar la importancia de escribir código eficiente cuando se trabaja con grandes volúmenes de datos y recursos compartidos.

