

## Integrantes:

- Sebastian Cuevas
- Andres Seguel
- Cristian Ortiz

## Hardware Utilizado:

Para generar de los resultados y gráficos se utilizó la siguiente configuración:

- CPU: AMD Ryzen 7 5700X
- GPU: NVIDIA GeForce RTX 3080
- SO: Linux

## Interpretación de Resultados:

De acuerdo con los gráficos, se observan las siguientes tendencias:

- Escalabilidad de la CPU vs GPU: La CPU presenta un crecimiento de tiempo casi exponencial para  $N > 1024$ . Debido a que el tiempo de cómputo para  $N = 8192$  en CPU es demasiado extenso, se limitaron las pruebas de CPU hasta  $N = 4096$ . Esto causa que la línea azul en el gráfico de tiempos se interrumpa.
- Eficiencia de la Memoria Compartida: La versión GPU<sub>sm</sub> es consistentemente más rápida que la versión GPU básica. Esto se debe a la técnica de tiling, que minimiza los accesos a la memoria global que es mas lenta y aprovecha la memoria compartida dentro de cada bloque de hilos.
- Saturación del Speedup: El speedup sobre la CPU crece rápidamente hasta  $N = 2048$ , alcanzando un factor de aceleración superior a 1750x con GPU<sub>sm</sub> para matrices de 4096. A partir de ese punto, la pendiente del speedup tiende a estabilizarse, lo que indica que la GPU está alcanzando su máxima capacidad de paralelismo para este hardware.
- GFLOPs: En los tamaños más grandes, la GPU logra mantener una tasa de cómputo estable, demostrando que el rendimiento se mantiene óptimo incluso con cargas de trabajo masivas.