**平行程式Lab7 109062101 許佳綺**

**Strong Scalability & Observation**

可以看到雙GPU的train sample per second會比單GPU還要高 ，GPU數量跟train sample per second成正相關，但是因為是Distributed Data Parallel，所以推測多GPU時會有不同device間communication的overhead，有才讓scalability沒有到成倍數正比的成長。

**Experiment Process**

我使用hades跑單GPU跟雙GPU的script，分別使用run\_DDP\_1GPU.sh與run\_DDP\_2GPU.sh做實驗。下完sbatch run\_DDP\_{n}GPU.sh後會出現.err跟.out檔案在下指令的該資料夾下。而在1GPU.err裡也可以看到total train batch size是1，2GPU.err裡是total train batch size=2，也可以看到確實是DDP，也就是把data切割成兩個batch給不同的device運算。而最後出現在1GPU.out跟2GPU.out都有出現train metrics，也用該數據來繪製scalability的圖並分析。