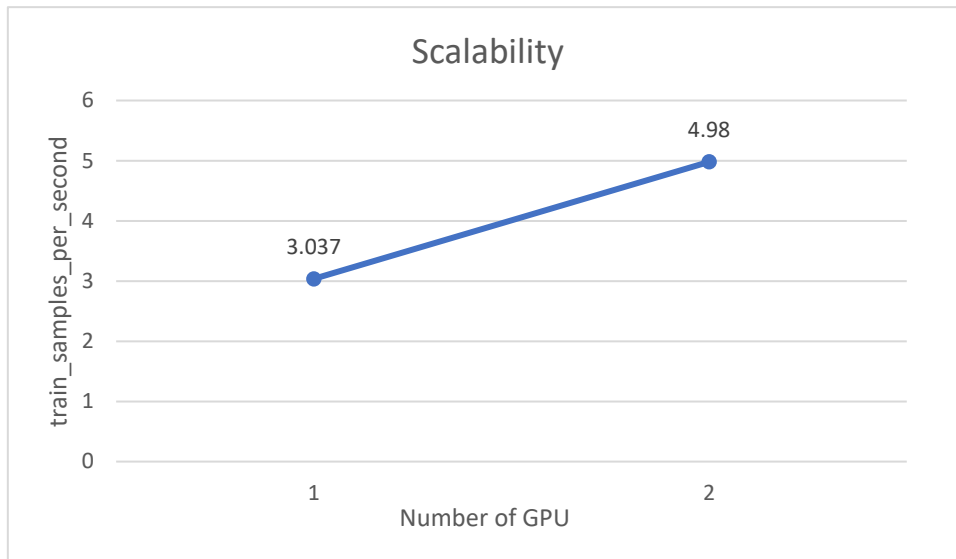


平行程式 Lab7 109062101 許佳綺

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。