UTEP-PNNL-DDL meeting

September 20, 2024

Agenda and Notes

1. Perlmutter office hour next Friday 10-12 PDT/11-1 MDT

2. Set up overleaf for papers

* 1. Tools evaluation paper

Criteria – functionality, overheads, portability, ease of use

* 1. Extend NUG poster on roofline modeling of MLPerf HPC benchmarks

3. Updates from students

a. Briana -- got OpenCatalyst running with Steven Farrell’s help, problem was old data; will update instructions on project github; will put data in CFS; will start with cyPAPI on ResNet; Josh would welcome help on cyPAPI; Briana working on poster

b. Irvin -- couldn’t login to Perlmutter but supposed to be fixed now, working on outline of thesis, helped Aaron with roofline modeling

c. William -- Tried Score-P but doesn’t work with PyTorch multiprocessing unless using MPI as backend; switched to using Nsight Systems; Nsight results who a lot of time spent in synchronization; putting results on project github

d. Alejandro -- working with Spark data, have to use Pandas -> Numpy; Josh -- high replication of RDDs can cause running out of memory; Spark designed back when we were trying to do big data with key-value stores; Spark parallelizes the key space. Deepak -- can specify amount of resources for docker container.

4. Josh will visit October 25

5. Josh -- NeurIPS MLforSys workshop paper -- off-policy learning,

6. Next year’s NERSC ERCAP proposal due October 7, will need to report on how we used this year’s allocation