UTEP-PNNL-DDL meeting

May 10, 2024

Agenda

1. Debrief on Year 1 report
   1. Summary of report
   2. Response from program manager
   3. Project milestones
2. New NVIDIA hardware at UTEP by August
   1. H100 DGX with 8 GPUs
   2. HGX with 4 H100 GPUs
3. NERSC Perlmutter
   1. Should make full use of allocation if possible, next reduction for underuse will be in July
   2. Install group tools in common area
   3. Datasets in common area if possible, or in scratch (but scratch gets purged)
4. MLPerf HPC Benchmarks
   1. How to divide them up among us
   2. Reproduce previous work on Roofline models (done with V100 GPUs)
   3. Communication overheads
5. Collaborations
   1. KTH Sweden -- data movement
   2. UTK PAPI -- SDEs for operation counts, Python interface for PAPI
6. Summer plans
   1. Extend CNN work to GNNs, auto-tuning + compiler -- Briana
   2. Communication measurements using Score-P – Christopher (with Deepak)
      1. <https://www.vi-hps.org/tools/score-p.html>
      2. <https://github.com/score-p/scorep_binding_python>
      3. <https://scorepci.pages.jsc.fz-juelich.de/scorep-pipelines/docs/scorep-4.1/html/quickstart.html>
   3. Communication overhead analysis – William (with Deepak)
   4. PNNL internship -- Alejandro
7. Possible paper venues
   1. PMBS’24 – August 7/September 13 deadlines for full/short papers
   2. Parallel tools workshops (ProTools 24 at SC’24, HLRS Parallel Tools Workshop usually in fall in Germany)
   3. MLSys’25 (deadline usually in October)
8. Updates from students
   1. Briana
   2. Samantha
   3. Irvin
   4. William
9. Meeting time for summer and next fall