Supplemental Discussion

MLPerf™ HPC v2.0 Results Discussion

The submitting organizations provided the following descriptions as a supplement to help the public understand the submissions and results. The statements **do not reflect the opinions or views of the MLCommons™ Association.**

Dell

Dell Technologies has long been dedicated to advancing, democratizing, and optimizing HPC to make it accessible to anyone who wants to use it. Together, Dell and NVIDIA have partnered to deliver unprecedented acceleration and flexibility for AI, data analytics and HPC workloads to help enterprises tackle some of the world's toughest computing challenges.

For the MLPerf HPC Training 2.0 testing, Dell submitted model 32x PowerEdge XE8545 servers with 128 NVIDIA A100 SXM GPUs for DeepCAM training model. This submission is from the Rattler supercomputer at the Dell Technologies Edge Innovation Center. The HPC system, stemming from a partnership with NVIDIA, is designed to showcase extreme scalability and was previously recognized on the TOP500 list of the world's fastest supercomputers.

There are always going to be bigger questions and bigger data sets requiring HPC solutions to keep pace with the speed of innovation. Dell has the engineering expertise needed to build large scale GPU solutions to meet these growing demands across industries. Scientific researchers at Oregon State University (OSU) are using Dell servers with NVIDIA GPUs for climate change research, among other areas. For them, innovative HPC technology in tailored configurations is the must-have capability to drive meaningful discoveries. "It used to take about 10 years to fully sequence a seawater sample", says Christopher Sullivan, Assistant Director of Biocomputing at OSU's Center for Genome Research and Biocomputing. "Now it takes about less than a week to analyze and sequence all of the DNA in a sample."

Experience Dell's solutions for HPC for yourself in one of our worldwide Customer Solution Centers. Tap into one of our HPC & AI Centers of Excellence and/or collaborate with our HPC & AI Innovation Lab. When you engage with the Lab, you work directly with experts to design a solution for your unique HPC workloads.

Fujitsu + RIKEN

RIKEN and Fujitsu jointly developed the world's top-level supercomputer—the supercomputer Fugaku—capable of realizing high effective performance for a broad range of application software, and started its official operation on March 9, 2021 [1]. RIKEN and Fujitsu submitted CosmoFlow results to closed division using 512 nodes for strong scaling and 81,536 nodes (=128 nodes×637 model instances) for weak scaling.

For both weak and strong scaling, LLIO (Lightweight Layered IO Accelerator) was used to cache library and program files from FEFS (Fujitsu Exabyte File System) storage. We developed customized TensorFlow and optimized oneAPI Deep Neural Network Library (oneDNN) as the backend [2]. The oneDNN uses JIT assembler Xbyak_aarch64 to exploit the performance of A64FX.

For weak scaling, since the job scheduler cannot launch a large number of instances immediately, inter-instance synchronization across jobs was added to align start times among instances. Moreover, to avoid excessive access to the FEFS from all instances, the dataset is staged to node local memory using a MPI program that only the first instance reads the dataset from FEFS and broadcasts it to the other instances. We actually ran 648 instances (82,944 nodes) but submitted 637 instance results of them. The pruned instances consist of 1 instance that hung during training, 6 instances that used the same seed value as others unintentionally, and 4 instances that took particularly long time.

For strong scaling, we used reformatted uncompressed TFRecord dataset to improve training throughput. The reference dataset is compressed with gzip and needs decompression at each training step. Since the number of nodes increases from weak scaling and the amount of staging data per node decreases, the uncompressed dataset could be used.

In this round, the performance of the Fugaku half-system with more than 80,000 nodes can be evaluated using the weak scaling metric.

[1] https://www.fujitsu.com/global/about/innovation/fugaku/

[2] https://github.com/fujitsu

HelmholtzAI

In Helmholtz AI, Germany's largest research association has teamed up to bring cutting-edge AI methods to researchers from the natural sciences. With this in mind, the Helmholtz AI members from the Steinbuch Centre for Computing (SCC) at Karlsruhe Institute of Technology (KIT) and the Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich have jointly submitted their results for the MLPerf® HPC benchmarking suite. We are proud of our large-scale training runs using NVIDIA A100 GPUs on both the HoreKa supercomputer at SCC and the JUWELS Booster at JSC. On the latter, we used up to 3,072 NVIDIA A100 GPUs during these measurements.

The MLPerf® HPC benchmarking suite is a great opportunity for us to fine-tune both code-based and system-based optimization methods and tools. For CosmoFlow, we were able to improve our submission by over 300% compared to last year! While fine-tuning our IO operations, for example, we discovered ways for our filesystems to more reliably deliver read and write performance.

As the impacts of climate change become more apparent, it is also imperative to be more conscious about our environmental footprint, especially with respect to energy consumption. To that end, the system administrators at HoreKa have enabled the use of the Lenovo XClarity Controller to measure the energy consumption of the compute nodes*. For the submission runs on HoreKa, 1,127.8 kWh were used. This is more than it takes to drive an average electric car from Miami to Vancouver or from Portugal to Finland.

The MLPerf® HPC benchmarking suite is vital to determining the utility of our HPC machines for modern work flows. We look forward to submitting again next year!

*This measurement does not include all parts of the system and is not an official MLCommons methodology, however it provides a minimum measurement for the energy consumed on our system. As each system is different, these results cannot be directly transferred to any other submission.

NVIDIA

The HPC community is amid a second renaissance – one associated with adopting AI methods to augment or replace traditional HPC approaches. Over the last five years, the number of research papers published about AI-accelerated simulation has increased from less than 100 per year to nearly 5,000 in the last year.

MLPerf HPC benchmarks measure training time and throughput for three types of high-performance simulations that have adopted machine learning techniques. Peer-reviewed industry-standard benchmarks are a critical tool for evaluating HPC platforms, and we believe access to reliable performance data will help guide HPC architects of the future in their design decisions.

The MLPerf HPC benchmarks seek to model the types of workloads HPC centers perform:

- Cosmoflow - physical quantity estimation from cosmological image data
- Deepcam - identification of hurricanes and atmospheric rivers in climate simulation data
- Opencatalyst – prediction of molecular configuration energy levels based on graph connectivity

Importantly, MLPerf HPC exercises, and is sensitive to the impact of every key subsystem from memory bandwidth to shared filesystem throughput. Therefore, we believe the MLPerf HPC benchmark represents one of the best tools for HPC and AI centers system bring-up and acceptance testing while also being the best metric to use for system comparison during design and acquisition phases.

NVIDIA continues to improve scores year over year for this submission by bettering the strong scaling scores of Cosmoflow by 2.1X and the best Opencatalyst score by 5.1X compared to last year.

NVIDIA and partner ecosystem submitted using two generations of NVIDIA GPUs (V100 and A100). Supercomputing centers Jülich, the Texas Advanced Computing Center, and NVIDIA partner Dell made submissions.

All software used for NVIDIA submissions is available from the MLPerf repository. NVIDIA is constantly making performance improvements, including those from MLPerf, to our software available on NGC, our software hub for GPU applications.

Chris Porter is a Technical Marketing Manager in NVIDIA's Accelerated Computing Group

Texas Advanced Computing Center

MLCommons HPC workgroup provides an excellent opportunity to evaluate Machine Learning applications on supercomputing platforms. In the v2.0 submission round, Dr Amit Ruhela ran two Machine Learning applications, i.e. Cosmoflow and Deepcam, on the TACC Longhorn system and submitted the performance numbers a third time. These benchmarks allow TACC staff to envisage and plan specifications for their upcoming supercomputing systems.