Supplemental Discussion

MLPerf™ Training v1.1 Results Discussion

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

Baidu

Baidu started to develop deep learning applications as early as 2012. In 2013, we began developing a deep learning framework, which led to the release of PaddlePaddle in 2016. This year, Baidu released core framework v2.2, which has already begun to be widely deployed in the industry for applications including speech, vision, and NLP.

PaddlePaddle is an Industrial Grade Deep Learning Platform, supporting both declarative programming and imperative programming, while providing a high degree of development flexibility and high runtime performance. Designed to be easy-to-use in both scientific research and industrial applications, PaddlePaddle has been applied by a wide range of companies.

With MLPerf Training 1.1, we have made remarkable optimizations on the PaddlePaddle framework, including CUDA Graph, fully asynchronous GPU executor, convolution-batch normalization fusion and optimizer kernel merging. We have submitted the ResNet50 benchmark results using both the PaddlePaddle and the NGC MXNet 21.05 framework, showing that the ResNet50 model on the PaddlePaddle framework reaches the same performance as that of the NGC MXNet 21.05 framework, with PaddlePaddle ranking among the fastest frameworks as tested on A100 GPUs.

We are grateful to the MLCommons for providing this excellent platform for communication. We look forward to sharing our further performance improvements for the PaddlePaddle framework along with more results in the future.

Dell Technologies

At Dell Technologies, we continue to push technology, so you can go further.

To provide the data you need to compare and select the best options, Dell Technologies submitted 51 results across 12 system configurations on all eight of the MLPerf training models.

- · **Select the best**. See how different CPU, GPU and memory configurations perform for specific AI training workloads.
- · **Speed multi-node results.** As AI models continue to grow with a need for speed, the Dell Technologies Innovation Lab team submitted training results on multiple nodes to show scalable performance.
- · **Save with PCIe.** With an eye toward performance per watt and per dollar, the team submitted benchmarking results for PCIe-connected and NVLINK GPUs.

Come see for yourself in one of our worldwide Customer Solution Centers. Collaborate with our HPC & AI Innovation Lab and/or tap into one of our HPC & AI Centers of Excellence.

Fujitsu

Fujitsu is a leading information and communications technology company that supports business through delivering robust and reliable IT systems by a group of computing engineers. We participated in this round, MLCommons training v1.1 and improved resnet50 and ssd benchmark results. We also add the following results from this round: unet3d, bert and rnnt, which are reproducible with machine specific configurations.

Our system, PRIMERGY GX2460 M1, is a middle range computing node. It consumes less power and smaller area in 2U rack mount size, and can be used for various ways, not only for training but for inference. We also participated in MLCommons previous inference round with this system. The result can be confirmed at MLCommons website.

The system has two AMD EPYC processors and four NVIDIA A100 GPUs as accelerators, which are connected with PCI express and have their own 40GB memory in HBM. Its storage is 1.95TiB NVMe SSD connected via PCIe.

GIGABYTE

GIGABYTE Technology, an industry leader in high-performance servers, partook in MLCommons Training v1.1. This round, we chose dual 3rd Gen Intel Xeon Scalable 8362 for our GIGABYTE G492-ID0 with the NVIDIA HGX A100 80GB 8-GPU solution, a powerful end-to-end AI and HPC platform for data centers. It allows researchers to rapidly deliver real-world results and deploy solutions into production at scale.

We completed the frameworks:
- MXNet NVIDIA v.21.09
- Merlin HugeCTR w/ NVIDIA Framework
- PyTorch NVIDIA v21.09
- TensorFlow NVIDIA v.21.09

Overall, optimization and performance could be improved. Showed strong performance in PyTorch 21.09 and Merlin HugeCTR.

GIGABYTE will continue optimization of product performance, to provide products with high expansion capability, strong computational ability and applicable to various applications at data center scale. GIGABYTE solutions are ready to help customers upgrade their infrastructure.

Google

Throughout the course of this year, the demand for training billion and trillion-parameter scale machine learning models has grown significantly, both from within Google, and from our Cloud customers. This has been driven by findings across the ML industry that model accuracy and generalizability increase with model size. These models are orders of magnitude larger than the MLPerf reference models, and present unique scaling challenges to our infrastructure. Following Google's record-breaking performance results from MLPerf 1.0, we have taken this opportunity to showcase performance for model sizes at the cutting edge of research.

In MLPerf Training 1.1, Google has chosen to make 2 large model submissions to the Open Division of the benchmarking competition. The first is a 480 billion parameter BERT model using Lingvo on TensorFlow, that we trained using 2048 TPU v4 chips. Lingvo is Google's high level framework for building sequence models. The second is a 200 billion parameter BERT model using Lingvo on JAX, that we trained using 1024 TPU v4 chips. For both these models, we were able to achieve record-breaking efficiency, with a TPU FLOPs utilization rate of 63%.

While our fourth-generation TPU chip provides considerable compute power, the exceptional networking within the TPU Pod, as well as the advanced performance optimizations within the frameworks and compiler ensure that these chips are kept busy, even as work is split across thousands of chips. Such high efficiency at scale is critical to ensuring that these models are able to train as quickly as possible.

Our largest scale submission at 480 billion parameters was made using our recently launched Cloud TPU v4 Pods. This means that all of Google's industry-leading ML infrastructure, from frameworks such as Lingvo, TensorFlow and Jax, to the XLA Compiler and our latest generation HW are now accessible to the public.

Graphcore

Graphcore continues its participation with MLPerf Training with the introduction of two new systems, IPU-POD$_{128}$ and IPU-POD$_{256}$ for machine intelligence scale-out, which have been launched since our first MLPerf v1.0 Training submission. These systems are designed both for large scale distributed training and commercial AI inference applications and both are already shipping to customers in production and available in the cloud. As a result, we have submitted them directly into MLPerf's available category.

These new systems are powered by Graphcore's second generation Intelligence Processing Unit (IPU) and linked together by Graphcore's IPU-Fabric to deliver impressive training performance and highly efficient scaling.

We have demonstrated significant performance improvements since MLPerf v1.0 as a result of new functionality and ongoing optimisation of our standard Poplar SDK. Our Resnet submissions show a 24% improvement on IPU-POD$_{16}$, and 41% improvement on IPU-POD$_{64}$. Our BERT submissions show a 5% improvement on IPU-POD$_{16}$, and 12% improvement on IPU-POD$_{64}$.

We are also providing highly performant results for our IPU-POD$_{128}$ and IPU-POD$_{256}$ further demonstrating the efficiency with which IPUs can be scaled out for large, distributed training jobs. Scaling efficiency is strong and will continue to improve with our regular software releases.

The disaggregation of AI compute and servers means that CPU to IPU ratio in IPU-PODs can be optimized for different AI workloads, reducing the total cost of ownership (TCO), which is extremely important for customers in production. For example, for the NLP-based BERT workloads, the IPU-POD$_{128}$ uses just two dual-CPU servers, while a more data-intensive task such as computer vision (like ResNet) may benefit from an eight server (dual-CPU) setup.

As with all Graphcore hardware, the IPU-POD$_{128}$ and IPU-POD$_{256}$ are co-designed with our Poplar software stack, which provides support for high-level frameworks such as PyTorch and Tensorflow. Poplar manages communication and synchronization between IPUs enabling straightforward scale out for our IPU-POD systems.

All software used for our submissions are available from the MLPerf repository, to allow anyone to reproduce our results. The Graphcore Github repository also covers many other new and emerging models where the IPU's unique architecture can enable innovators to create the next breakthroughs in machine intelligence.

HPE

When data is universally accessible, AI teams can focus on development and deployment, and IT infrastructure is flexible and unbounded. HPE makes AI that is data-driven, production-oriented and cloud-enabled, available anytime, anywhere and at any scale.

We understand that successfully deploying AI workloads requires much more than hardware. That's why we deliver a full complement of offerings that enable customers to embark on their AI journey with confidence. Award-winning HPE AI Transformation Services make some of the brightest data scientists in the industry available to assist with everything from planning, building and optimizing to implementation. Built upon the widely popular open source **Determined Training Platform**, HPE Cray AI Development Environment helps developers and scientists focus on innovation by removing the complexity and cost associated with machine learning model development. Our platform accelerates time-to-production by removing the need to write infrastructure code, and makes it easy to set-up, manage, secure, and share Artificial Intelligence (AI) compute clusters. With HPE Cray AI Development Environment, customers are able to train models faster, build more accurate models, manage GPU costs and track and reproduce experiments.

Today we are publishing our inaugural MLPerf Training results based on the HPE Apollo 6500. Dual AMD EPYC processors and eight NVIDIA HGX A100 GPUs delivered leading results across multiple categories, including image detection/classification and speech recognition. As a founding member of MLCommons, HPE is committed to delivering benchmark results that provide our customers with guidance on the platforms best suited to support a variety of workloads.

Inspur

Inspur Electronic Information Industry Co., LTD is a leading provider of data center infrastructure, cloud computing, and AI solutions, ranking among the world's top 3 server manufacturers. Through engineering and innovation, Inspur delivers cutting-edge computing hardware design and extensive product offerings to address important technology arenas like open computing, cloud data center, AI, and deep learning.

In MLCommons TrainingV1.1, Inspur made submissions on two systems: NF5488A5 and NF5688M6. NF5488A5 is Inspur's flagship server with extreme design for large-scale HPC and AI computing. It contains 8 A100-500W GPUs with liquid cooling. NF5488A5 system is capable of high temperature tolerance with operating temperature up to 40℃. It can be deployed in a wide range of data centers with 4U design, greatly helps to lower cost and increase operation efficiency. NF5688M6 based on 3rd Gen Intel® Xeon® scalable processors increases performance by 46% from Previous Generation, and can support 8 A100 500W GPUs with air cooling. It accommodates more than 10 PCIe Gen4 devices, and brings about a 1:1:1 balanced ratio of GPUs, NVMe storage and NVIDIA Mellanox InfiniBand network.

In closed division, the single node performance of Bert, SSD and DLRM are improved by 16.03%, 4.0% and 10.99% compared with the best performance Inspur achieved in Training v1.0. In addition, Inspur submit the results of Mask R-CNN, Minigo, RNN-T and 3D UNET for the first time, and good performance is achieved on these workloads.

Intel-HabanaLabs

We're pleased to deliver the second results for the Habana® Gaudi® deep learning training processor, a purpose-built AI processor in Intel's AI XPU portfolio. This time at scale!

Intel-HabanaLabs submitted results for language (BERT) and vision (ResNet-50) benchmarks on Gaudi based clusters and demonstrated near-linear scalability of the Gaudi processors. The ongoing efforts to optimize the Habana software stack (SynapseAI® 1.1), which include data packing, sharded optimizers and checkpoint-saving, resulted in more than a 2x improvement in BERT time-to-train using the same Gaudi processors compared to our last round results

Topology	# of Gaudi1s	V1.0 Submission [min]	V1.1 Submission [min]
ResNet50	8	62.55	56.13
	16		33.25
	32		17.38
	64		9.48
	128		5.41
	256		3.39
BERT-Large	8	164.37	80.47
	16		42.29
	32		21.28
	64		11.87

This highlights the usability and scalability of Gaudi and demonstrates the capabilities of our SynapseAI® software platform, which includes Habana's graph compiler and runtime, communication libraries, TPC kernel library, firmware, and drivers. SynapseAI is integrated with TensorFlow and PyTorch frameworks and is performance-optimized for Gaudi.

We are looking forward to the next submission!

Lenovo

Lenovo is an industry trailblazer and global provider of data center infrastructure and solutions. We believe in smarter technology for all and specifically, ***smarter uses AI to rethink the possibilities***. From implementing computer vision for retail loss prevention to social distance monitoring for COVID-19 safety measures, we believe AI is an essential component of all we do, and we must empower organizations to realize the potential of what AI can do for them.

In MLPerf Training 1.1, we increased our number of benchmarked servers from two to three as well as increased the number of benchmarks executed from two to five all while using the fastest GPUs in the market.

Server	**GPU**	**Watts**
Lenovo ThinkSystem SR670 V2	8x 80GB NVIDIA A100 PCIe	300
Lenovo ThinkSystem SR670 V2	4x 80GB NVIDIA A100 SXM4	500
Lenovo ThinkSystem SD650N-V2	4x 80GB NVIDIA A100 SXM4	500

It is worth noting that Lenovo Neptune^{TM} liquid and hybrid cooling enables our servers with 500W cards. With liquid cooling, we can have these cards in 1U chassis while with hybrid we can do a 3U chassis for 4x500W.

Implementing AI can be a complex and seemingly daunting task. Organizations can rely on Lenovo's expertise to simplify and show the real business value of AI deployments. We believe MLPerf Training 1.1 results will bring clarity to those AI infrastructure conversations to allow customers to make informed decisions today to reduce risks associated with AI deployments tomorrow. Start your PoC or discover all Lenovo has to offer including software and services solutions to accelerate your AI initiatives through our Lenovo AI Center of Excellence.

Microsoft

Azure is pleased to share results from our first ever large-scale MLCommons training submission. [AJ4] [JS5] For this submission we used the NDm A100 v4[KR6] [JS7] series virtual machines (VMs) [RP8] [JS9] powered by 8 NVIDIA A100 GPUs (80 GB), 8 NVIDIA 200 Gb/s HDR InfiniBand cards, 96 AMD Rome cores, 1.9 TB of RAM, and 8 * 1TB NVMe disks. This high-end AI training platform allows our customers to scale from 1 – 256+ VMs (8 – 2048+ GPUs) as required by their AI training needs[AJ10] .

Some of the highlights from our MLCommons benchmark results are

1. Ability to train an entire Bert (Natural Language Processing Model) in nearly 25 seconds at 2048 GPUs.
2. Processed as high as 3.8M images/sec using Resnet (image classification) at 2048 GPUs.
3. Completed the Minigo (reinforcement learning) benchmark in under 17.5 minutes using 1792 GPUs.

These benchmark results demonstrate how Azure has

1. raised the bar in terms of scale and performance for AI training in the cloud.
2. is in-line with on-premises performance
3. is committed to democratizing AI at scale in the cloud

To generate these results, we used Azure CycleCloud to orchestrate the cluster environment of 256 VMs. We used the Slurm scheduler configured with NVIDIA Pyxis and Enroot to schedule the NVIDIA NGC MLCommons containers***. This enabled us to set up our environment in a timely manner and perform the benchmarks with strong performance and scalability. For more information on how to deploy this setup please see cc-slurm-ngc.

The NDm A100 v4 series VMs are what we and our Azure customers turn to when large-scale AI and ML training is required. We are excited to see what new breakthroughs our customers will make using these VMs

*** Special thanks to the NVIDIA team for all their support during this benchmarking effort

NVIDIA

In MLPerf v1.1, the NVIDIA AI ecosystem set records on every single benchmark from at-scale performance with the fastest time to solution, to normalized per-chip performance on NVIDIA A100 Tensor Core GPUs. All of these benchmarks were run both on-prem, and in the cloud. Our performance increased over five-fold in just a single year since MLPerf v0.7, on the broadly available NVIDIA A100. Continuous innovation has enabled this leadership performance, and NVIDIA AI is the only platform to submit on every benchmark encompassing diverse use cases, demonstrating both the highest performance and the versatility of the platform.

Direct submissions were made by our partners and accounted for over 90% of closed submissions. Microsoft Azure established itself as the world's fastest cloud for AI powered by NVIDIA A100 and HDR InfiniBand networking, setting records on every benchmark for cloud instances. Baidu, Dell, Fujitsu, Gigabyte, HPE, Inspur, Lenovo and Supermicro submitted on-prem. Dell, Inspur and Supermicro set multiple records on a per-chip basis.

In the last three years since the first MLPerf training benchmark launched, NVIDIA performance has increased over twenty fold. In just five months, NVIDIA's performance on the A100 GPU has increased up to 2.2x between MLPerf v1.0 and v1.1 powered by multiple software improvements including the following:

- Concat/split operations on Unet-3D are 2.5x faster versus MLPerf v1.0.
- Fine-grained overlap computation and communication improved performance, especially at scale up to 27% on DLRM
- CUDA graphs were expanded to encompass the entire iteration, improving performance by 6% on ResNet-50
- Added buffer registration to NCCL, which uses pointers rather than copying weights between GPUs, as well as fusing scaling operations to speedup BERT by 5%

NVIDIA AI continues to provide consistent performance improvements, offering a single leadership platform from cloud to data center to cloud to edge.

All software used for NVIDIA submissions is available freely from the MLPerf repository and these cutting-edge MLPerf improvements are added to containers available on NGC, our software hub for GPU applications.

Samsung

Samsung is delighted to share its first ever set of MLPerf Training result, after submitting to RDI (Research, Development and Internal) category on our debut round. We delivered an extremely strong performance on BERT training, 25.06 seconds on 1024 Nvidia A100 GPUs.

The system used for BERT training consists of 128 nodes, which have two AMD EPYC 7543 processors and eight NVIDIA Tesla A100s as accelerators, which are connected with NVLinks and have their own 80GB memory in HBM.

Based on PyTorch NVidia Release 21.08, we have focused on the large batch training and overlap between computation and communication for performance boost.

For BERT open division, we show x2.37 improvement TTT (Total Time on Test) over our internal baseline based on Nvidia's implementation which was published in Training v1.0.

Our key optimizations are:

- Fully utilize Pytorch DDP and ADAM optimizer for large batch training with communication/computation overlap

- Bucket-wise local gradient clipping which takes the best of both clip-norm-before-reduce and clip-norm-after-reduce

- Efficient input data load balancing for increasing GPU utilization

In addition to AI acceleration in mobile device, Samsung is actively researching on the scalable and sustainable AI computing. We will work to solve the scaling challenge between computing capability and memory bandwidth through innovation in memory and storage products such as HBM-PIM and AX-DIMM.

Supermicro

Supermicro has its long history of providing a broad portfolio of AI-enabled products for different use cases. In MLPerf Training v1.1, we have submitted results based on two high performance systems to address multiple compute intensive use cases, including medical image segmentation, general object detection, recommendation systems, and natural language processing.

Supermicro's DNA is to provide the most optimal hardware solution for your workloads and services. For example, we provide four different systems for NVIDIA's HGX A100 8 GPU platform and HGX A100 4 GPU respectively. Customers can configure the CPU and GPU baseboards based on their needs. Furthermore, we provide upgraded power supply versions to give you choices on using our cost-effective power solutions or genuine N+N redundancy to maximize your TCO. Supermicro also offers liquid cooling for HGX based-systems to help you deploy higher TDP GPU baseboards without thermal throttling. If customers are looking for rack scale design to cluster systems for large machine learning training problems, we can offer rack integration in air cooled solution, RDHx and DLC liquid cooling solution to suit your plug and play need.

Supermicro's SYS-420GP-TNAR, AS-4124GO-NART, AS-2124GQ-NART and upcoming SYS-220GQ-TNAR with NVIDIA's HGX A100 GPUs can pass data directly from GPU to GPU, to avoid the pass-through overhead from processors and system memory. By shortening the data path to the accelerator, it shortens the training time for applications such as computer vision and recommendation system.

With multiple configurations of processors, accelerators, system form factors, cooling solutions, and scale out options, Supermicro would like to provide our customers the most comprehensive and convenient solutions to solve the AI problems. We are happy to see all the results we ran on MLPerf using our portfolio of systems, and we will keep optimizing the solutions for customer's different requirements to help achieve the best TCO.