GIGABYTE™

DNN Training Appliance

MYELINTEK

GIGABYTE's DNN Training Appliance is a well-integrated software and hardware package that combines powerful computing performance together with a user-friendly GUI. In turn, it provides DNN developers with an easy to use environment to conduct dataset management, training jobs management, real time system environment monitoring, and model analysis. The appliance includes powerful hardware and software optimizations that can improve the performance while reducing the time required for DNN training.

Provides Developers and Data Scientists

Reduces the complexity of DNN training environment setup and management

To generate a production grade DNN model, a developer will need to go through many difficult and time consuming steps, including dataset collection, dataset cleansing, dataset labeling, dataset augmentation, dataset format conversion, models selection, model design, hyperparameters tuning, model training, model evaluation, and model format conversion. Each step requires different tools and configurations that require time and effort for preparation, and switching between these tools often requires additional time writing code to convert different formats to use with different tools.

GIGABYTE's DNN Training Appliance aims to reduce this complexity by providing a complete training and management platform, while incorporating all these processes into an easy to use web-browser based GUI. Users can import, convert and manage their dataset; design, train and evaluate different DNN models; and test inferencing of their trained models. Based on GIGABYTE's G481-HA1 server, the platform is fully optimized to use the bare metal resources available to deliver improved training performance on cost-efficient hardware.

Reduces the time and improves the accuracy for each DNN training job

DNN models need to be trained on a large dataset to achieve an acceptable level of accuracy. Depending on the dataset size, this training could take days or even weeks. And in order to adapt to the latest business circumstances or situations (such as new products, new regulations, etc.), the DNN model needs to be periodically retrained through the latest datasets. If running a DNN training job takes too long, it will have a serious impact on an organization's operations, resource management, and competiveness.

GIGABYTE's DNN Training Appliance helps to reduce training time by incorporating many different optimization features:GPU memory optimization to accommodate a large amount of training input or to fit a large model into GPU memory, automatic hyperparameter tuning (during a training job) to achieve higher accuracy, and dataset cleaning features to reduce the training time generated by mislabeled or duplicated training data.

Summary of DNN Training Appliance Benefits

Saves Time
All in one package eliminates the painful process of setting up your own hardware & software environment.

Team Collaboration
Multiple teams and their members can share the hardware resources with unified scheduling and secured by a container sandbox.

Version Control
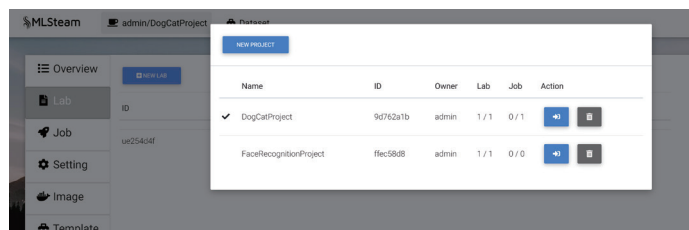Versioning all your hyper-parameters, environ-ments, and datasets to ensure reproducibility of DNN models.

Share Knowledge
DNN models can be shared as templates. There is no need to reinvent the wheel.

Effortless Training Management and Monitoring

Project-based training management

Gigabyte's DNN Training sytem uses a group training process following a project concept. You can import new datasets, and create new training jobs in a project easily via the web portal UI. In one project you can keep track of your model training histories, including hyperparameter modification, training job results, and training job models.

Project table view

Training jobs view

Cloud IDE & utilities interface

Users can easily create a Cloud IDE (based on Jupyterlab) for DNN model development or data preprocessing by attaching their dataset. The Cloud IDE also provides utilities, such as hyperparameter passing, 3rd-party IDE integration (VSCode and PyCharm), tensorboard and GPU monitoring to simplify the training process.
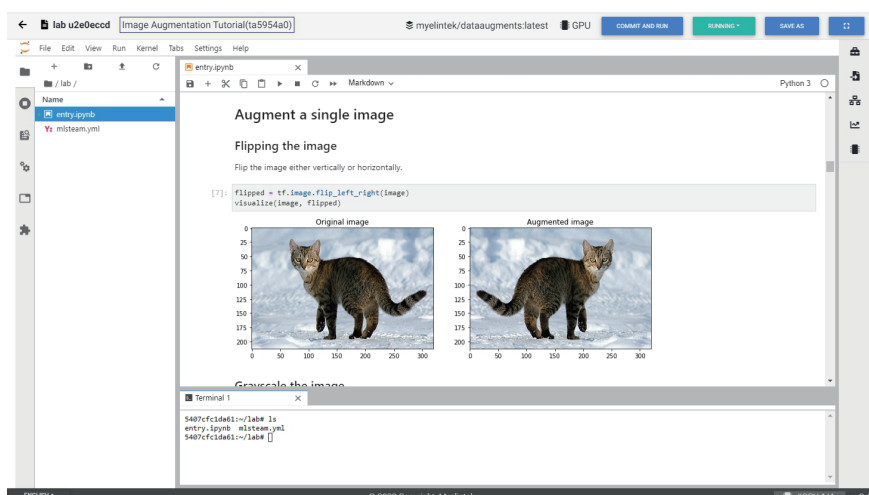
Cloud IDE

Model templates and optimization tutorial

Gigabyte's DNN Training sytem has built-in templates, guiding the user on how to train different types of models (for image classification, object detection, etc.) with various optimization techniques, such as GPU memory optimization and mixed precision training. These templates allow the user to easily choose the dataset, DNN models, and hyperparameter settings needed based on the DNN application type. Thus, the user can easily leverage templates for collaboration.

Real-time monitoring and quick result verification.

Once training starts, it is possible to keep track of the progress in real-time via the training monitoring chart. After each training job is completed, you can quickly verify your DNN model with the Cloud IDE workspace.

Effective Dataset Management Tools

User Friendly File Browser

The platform provides a file browser style management interface. The user can preview image files, delete files, and download files by selecting target files. To upload files, simply drag and drop files from your PC to the dataset.

Dataset Annotation Visualization

The platform supports multiple dataset annotated formats so that the user can preview the annotated dataset on the dataset page. Ex. Bounding box, segmentation images, etc.
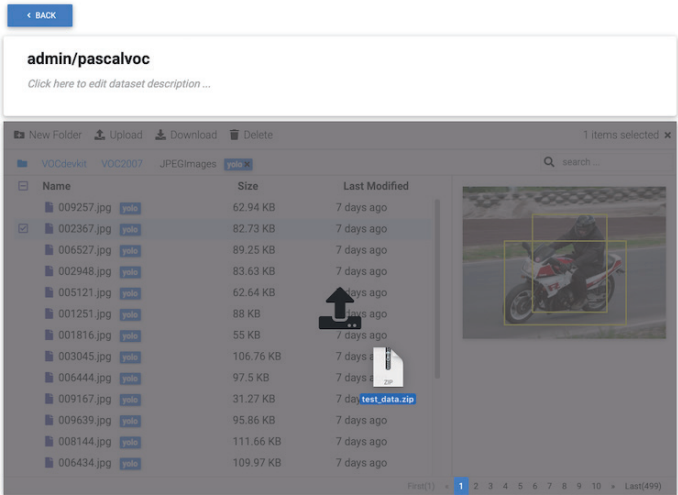
Upload files by drag and drop
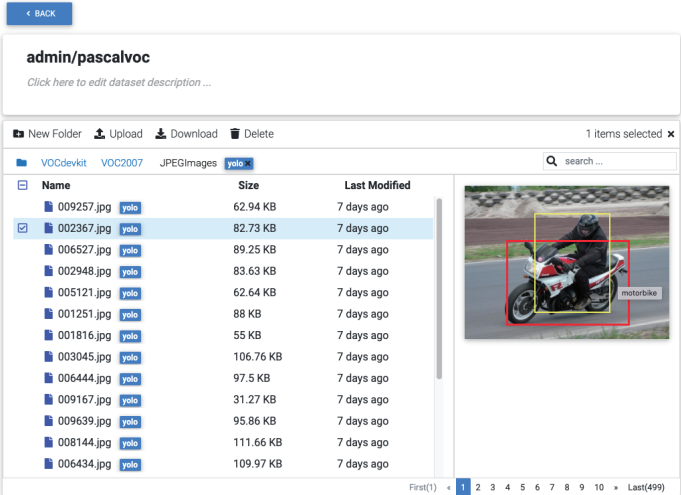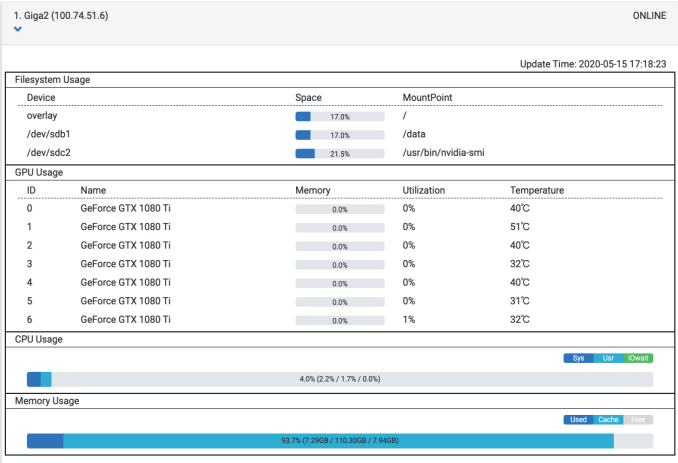

Bounding box annotation visualization

System Monitoring and Administration

System resource monitoring

GIGABYTE's DNN Training Appliance features real-time GPU (including GPU utilization, GPU memory usage, and temperature), CPU, Disk, and memory usage monitoring.

Administration Dashboard

GIGABYTE's DNN Training Appliance features a dashboard for administration, including an audit log, training tasks overview, dataset overview, and user account management.

Real-time system resource monitoring

Audit log of system operations

Optimized Hardware Platform

Single-Root GPU Server

GIGABYTE's DNN Training Appliance is built in the G481-HA1, which is a server optimized for a single cluster training appliance by employing a single root GPU system architecture (one CPU controls all of the GPUs). Since a DNN training job requires frequent communication between each GPU in the system, utilizing a single-root architecture (all GPUs can communicate via the same CPU root) helps reduce GPU to GPU latency and decrease DNN training job time.
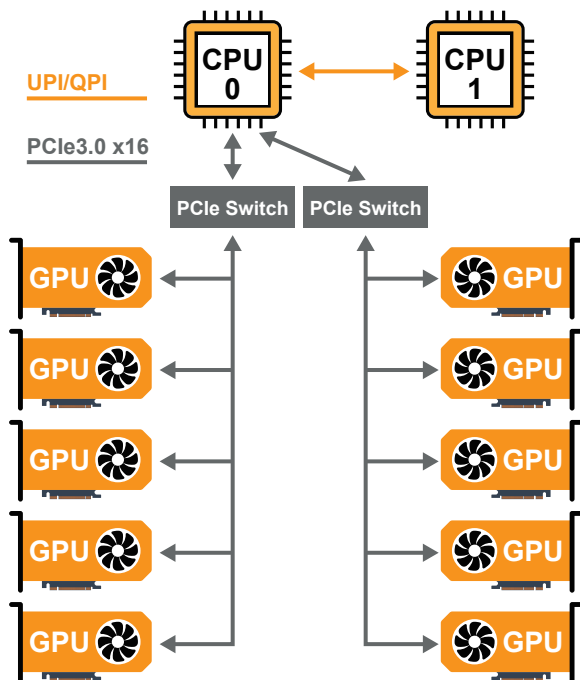
Single-Root Configuration
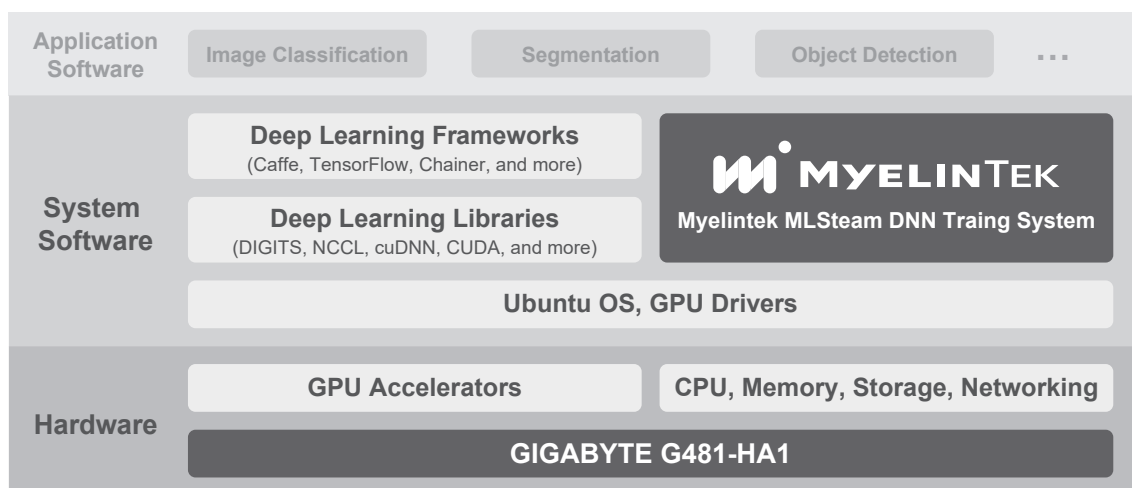Minimize GPU to GPU latency for small scale deployments
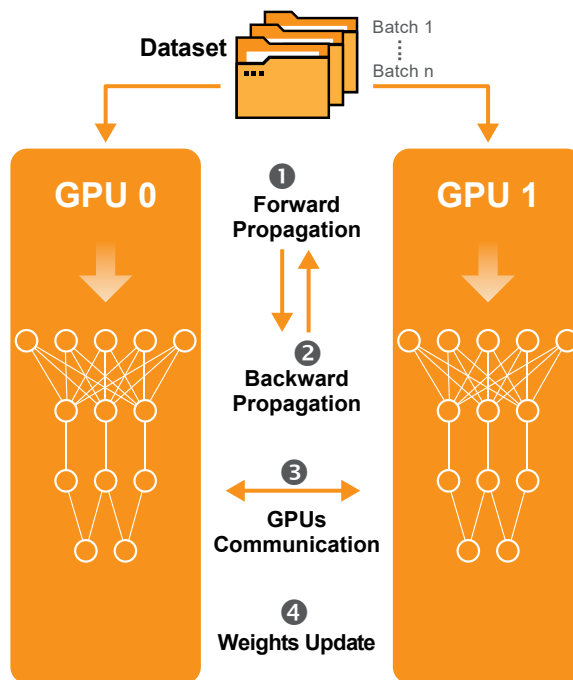
DNN Model Training Method
GPU to GPU synchronization is an important feature of DNN model training

G481-HA1 Specifications:

CPU	Dual 2nd Generation Intel Xeon Scalable Processors, TDP up to 205W
Memory	6-Channel DDR4 memory, 24 x DIMMS Intel Optane DC Persistent Memory Ready
Networking	2 x 10GbE BASE-T LAN ports 2 x 1GbE BASE-T LAN ports (Optional: 4 x Omni-Path QSFP28 LAN ports)
Storage	8 x 2.5" NVMe + 2 x 2.5" SATA / SAS hot-swap SSD 12 x 3.5" SATA / SAS hot-swap HDD
Expansion Slots	10 x PCIe x16 (3.0 x16) for GPUs 1 x PCIe 3.0 x16, LPHL 1 x PCIe 3.0 x16, LPHL (occupied by RAID card)
Power	3 x 2200W 80 PLUS Platinum redundant PSUs
Management	1 x Dedicated management port Aspeed AST2500 management controller GIGABYTE Server Management remote management platform

GIGABYTE DNN Training Appliance

Application Software	Image Classification	Segmentation	Object Detection	...

System Software

Deep Learning Frameworks (Caffe, TensorFlow, Chainer, and more)	**MYELINTEK** Myelintek MLSteam DNN Traing System
Deep Learning Libraries (DIGITS, NCCL, cuDNN, CUDA, and more)	

Ubuntu OS, GPU Drivers

Hardware

GPU Accelerators	CPU, Memory, Storage, Networking

GIGABYTE G481-HA1

Dataset management

- Multi-format dataset support: image classification datasets, object detection datasets, etc.
- Automatic conversion of image classification / object detection dataset into a format required by a deep learning framework
- Use GUI to quickly browse dataset content, edit dataset tags, and delete, copy, retag, etc.

Training management

- Guided training workflow for image classification and object detection
- Use GUI for system management, DNN model editing and visualization, hyperparameters setting, and real-time model training performance monitoring
- Support for multiple mainstream deep learning frameworks
- Multiple training job scheduling support with containers
- Web console interface: providing an environment for developers to prepare training materials for pre-processing in a browser
- CLI (command-line interface): for managing resources and running training jobs remotely
- Support model version control, efficient systemized management training, deep learning model

Training optimization

- One-click enabled optimization features, include mixed precision training, automatic hyperparameters tuning, and GPU memory usage optimization
- Confusion matrix feature for advanced analysis of image classification training results
- Simultaneously run model training with different hyperparameters setting combinations
- Multiple GPU training support
- Flexibility to scale GPU resources according to GPU resource usage rate, and perform training job scheduling
- Single-root GPU architecture for reduced latency GPU to GPU synchronization during training job

Environment monitoring

- System monitoring interface, displaying information such as usage rate of system resources required for model training
- GPU thermal aware feature prevents the GPU from being damaged due to excessive temperature