高效能深度學習訓練系統
High-performance DNN Training Appliance (ITRI DNN Training System, IDTS)

Eric Ming-Chiang Chen 陳明江 **PhD**
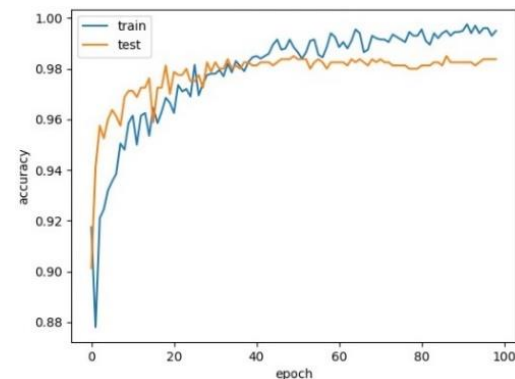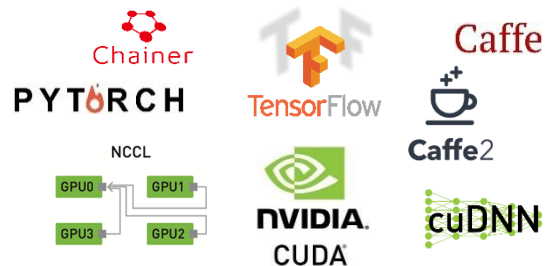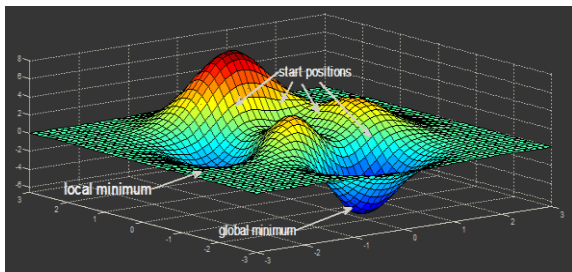
ericmcc@itri.org.tw

資訊與通訊研究所

Information and Communications Research Laboratories, ICL

ITRI DNN Training System

Total Solution of DNN Training Environment for DNN Developers

- **Software**
  - Full DNN training system software stack
  - Intuitive training management interface, Resource management
  - Distributed training, Performance monitoring, Sample DNNs, etc.
- **Hardware**
  - Commodity X86 machines + GPUs
- **Optimization**
  - Faster DNN training
  - Better accuracy
  - Complicated DNN accommodation
  - Large dataset training
  - Hyperparameters optimization
  - More DNN training visibility
  - Flexibility & scalability of DNN training resources
  - **DNN related research and development technologies**

  - Without NVLink
  - Low-end GPU with less GPU memory and lower GPU computing power

Features

- **Q2, 2018**
  - Faster training without scarifying accuracy
    Reduce training time to achieve the accuracy

- **Q3, 2018**
  - Automatic hyperparameters optimization
    Automatically configuring hyperparameters for higher accuracy

- **Q4, 2018**
  - Complicated DNN accommodation
    Make a huge network fit in GPU memory

- **2019**
  - Distributed DNN training
  - Other optimization technologies

Faster DNN Training Without Sacrificing Accuracy

Image Per Second, Per $USD

	0.65X	**1X**	**0.37X**
	X86	**IDTS (ITRI DNN Training System)**	**DGX-1**
Hardware	A X86 server + 8 NVIDIA GeForce GTX 1080 Ti GPUs	A X86 server + 8 NVIDIA GeForce GTX 1080 Ti GPUs	8 NVIDIA Tesla V100 SXM2 (NVLink) GPUs
Software	Open Source Community	IDTS	DGX-1 Software (DIGITS)
Price	under USD$34,000 for hardware	under USD$34,000 for hardware	more than USD$149,000
ResNet-50, ILSVRC2012 (1,281,167 images), 30 epochs, NVIDIA Caffe			
Training Time	11h 19m	7h 20m	4h 25m
Accuracy	68.48%	68.86%	67.89%
Images Per Second	946	1,466	2,417

ITRI
Industrial Technology
Research Institute

Faster DNN Training Without Sacrificing Accuracy

ResNet-50, ILSVRC2012 (1,281,167 images), 30 epochs, NVIDIA Caffe		
	X86	**IDTS** **(ITRI DNN Training System)**
Training Time	**11h 19m**	**7h 20m**
Multi-GPU Synchronization	**21.27%**	**1.64%**
Convolution	46.27%	55.02%
Batch Normalization	15.06%	19.92%
ReLU	9.24%	12.67%
Others	8.15%	10.75%

IDTS speeds up the DNN training by **reducing the multi-GPU synchronization**, where there is **NO NVLink**, and also **keeps the accuracy**.

ITRI
Industrial Technology
Research Institute

Faster DNN Training Without Sacrificing Accuracy

80% Training Time

120% Images Per Second

	X86	IDTS (ITRI DNN Training System)
Hardware	A X86 server + 8 NVIDIA GeForce GTX 1080 Ti GPUs	A X86 server + 8 NVIDIA GeForce GTX 1080 Ti GPUs
Software	Open Source Community	IDTS
Price	under USD$34,000 for hardware	under USD$34,000 for hardware
ResNet-50, ILSVRC2012 (1,281,167 images), 90 epochs, TensorFlow		
Training Time	1d 3h 56m	22h 26m
Accuracy	75.63%	75.14%
Images Per Second	1,183	1,428

ITRI
Industrial Technology
Research Institute

Production Use

- **Production in ITRI DNN Training Farm**
  - Centralized GPU datacenter for ITRI DNN developers, including 資通、生醫、材化、機械、綠能、電光、巨資、智機、服科 labs
  - 4x NVIDIA DGX-1, 6x X86 GPU servers
  - NVIDIA Tesla V100 GPUs, NVIDIA GeForce GTX 1080 Ti GPUs
  - GPU thermal-aware management
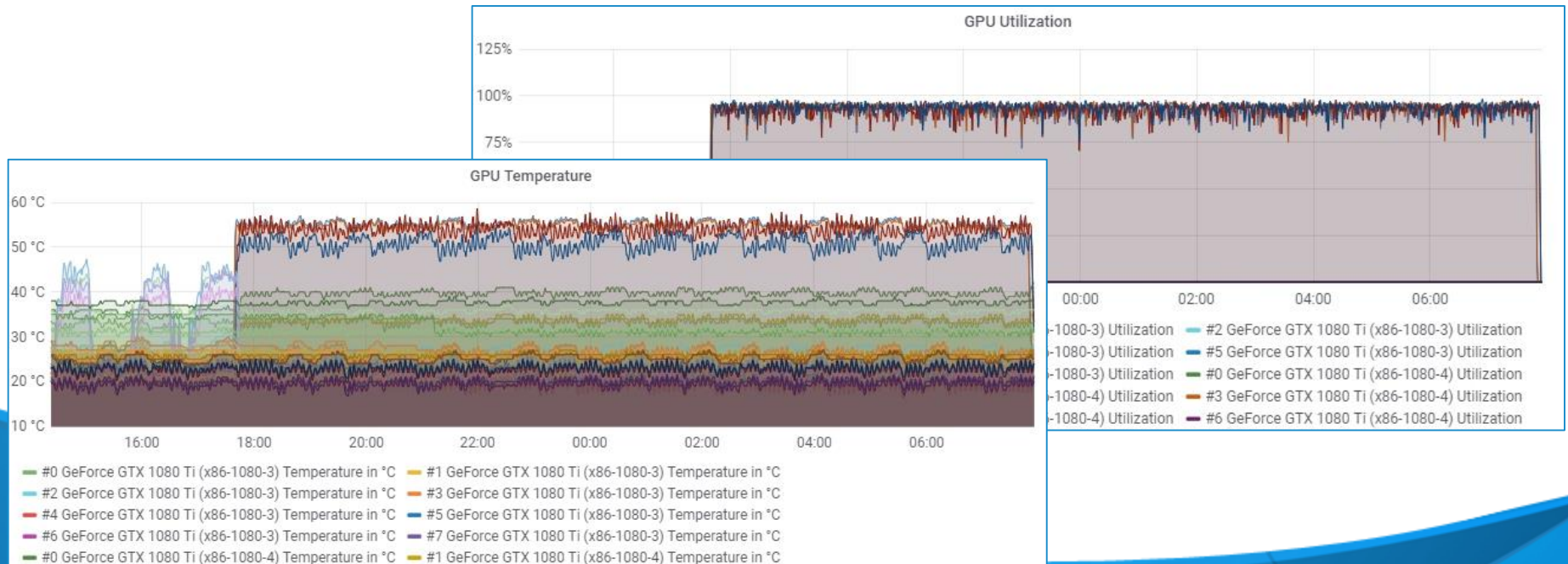
ITRI DNN Training System

Total Solution of DNN Training Environment for DNN Developers
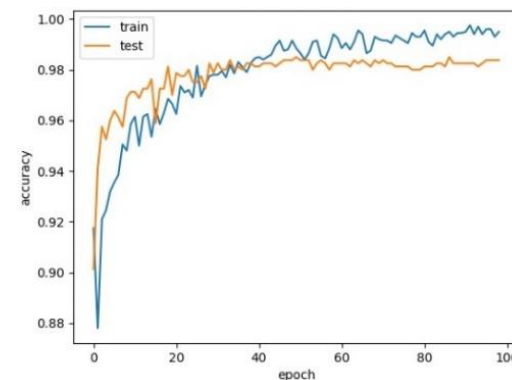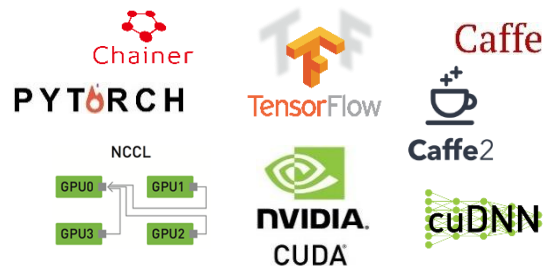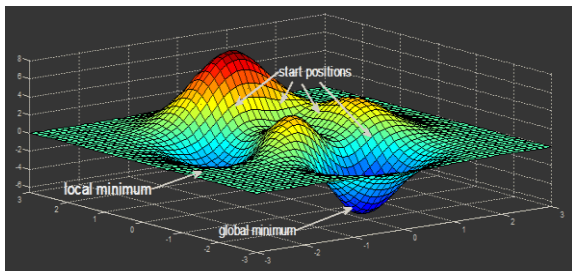
ITRI
Industrial Technology
Research Institute

Thank you!

Eric Ming-Chiang Chen 陳明江 **PhD**
ericmcc@itri.org.tw
+886-920-603351
Skype: emc.chen
ICL 資訊與通訊研究所