

Xilinx MLPerf v1.0 Inference Demo on Alveo VCK5000

INTRODUCTION

MLPerf.org issued its third round of inferencing result (MLPerf Inference v1.0). In this benchmark, Xilinx has achieved a result of ResNet-50 inference throughput at 5,921 FPS (available category, server model, single card) using the latest Versal ACAP PCI-e card VCK5000 for Data Center. Outperformed T4 5,423FPS performance under the same mode.

As the inventor of the FPGA and adaptive SoCs, Xilinx adaptive computing devices enable Domain-Specific Architectures (DSAs) optimized for AI inference and associated pre/post-processing workloads. They support not only fast AI inferencing but also end-to-end application acceleration.

With the powerful software AI development platform Vitis AI, developers could take models trained using TensorFlow, Caffe or PyTorch, and compile, quantize and optimize for Xilinx's software-programmable adaptive computing platforms without hardware development knowhow







Key Features

- > Optimized DPU architecture
- > Support batch computing
- > Standard workflow with Vitis AI
- > Optimized compiler strategy
- > Optimized device power



DEMO OVERVIEW

In this demo, you will get the containerized MLPerf v1.0 benchmark demo on VCK5000, including the preset AI acceleration engine DPUv4E for Versal, compiled and deployed ResNet-50 model, and the software dependencies.

Demo performance:

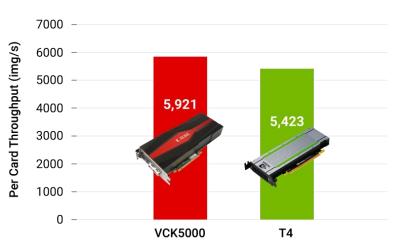
1.2GHz AIE: Server 5269 FPS / Offline 5732 FPS

Note: we recommend to use the default clock frequency at 1.2GHz to run this demo.

MLPERF V1.0 SUBMISSION RESULT

AMD Server + DPUv4E @VCK5000 end-to-end performance:

1.33GHz AIE: Server 5921 FPS / Offline 6257 FPS



MLCommons v1.0 Data Center Closed Division Server ResNet-50

TAKE THE NEXT STEP

- Download the demo and get the accelerated performance on your VCK5000 card.
- Contact with Xilinx for more details bingging@xilinx.com

Corporate Headquarters

San Jose, CA 95124

Tel: 408-559-7778

www.xilinx.com

Xilinx, Inc. 2100 Logic Drive

USA

Europe

Xilinx Europe Bianconi Avenue Citywest Business Campus Saggart, County Dublin Ireland Tel: +353-1-464-0311 www.xilinx.com

Xilinx K.K.

Japan

Art Village Osaki Central Tower 4F 1-2-2 Osaki, Shinagawa-ku Tokyo 141-0032 Japan Tel: +81-3-6744-7777 japan.xilinx.com

Asia Pacific Pte. Ltd.

Xilinx, Asia Pacific 5 Changi Business Park Singapore 486040 Tel: +65-6407-3000 www.xilinx.com

India

Xilinx India Technology Services Pvt. Ltd. Block A, B & C, 8th, 12th and 13th Floors, Meenakshi Tech Park, Survey No. 39, Gachibowli (V), Serilingampally (M), Hyderabad - 500 084 India Tel: +91-40-6721-4747 www.xilinx.com



© Copyright 2019 Xilinx, Inc. Xilinx, the Xilinx logo, Alveo, Artix, Kintex, Spartan, Versal, Virtex, Vivado, Zynq, and other designated brands included herein are trademarks of Xilinx in the United States and other countries. All other trademarks are the property of their respective owners. Printed in the U.S.A. Initials&Date