PRODUCT BRIEF

AMD ALVEO™ U50

Adaptable Accelerator Cards for Data Center Workloads

COMPUTE, NETWORKING, AND STORAGE ACCELERATOR FOR CLOUD AND EDGE DATA CENTERS

The AMD Alveo™ U50 Data Center accelerator cards provide optimized acceleration for workloads in financial computing, machine learning, computational storage, and data search and analytics. Built on AMD UltraScale+™ architecture and packaged in an efficient 75-watt, low-profile form factor, the U50 includes 8GB HBM2, 100GbE networking, and PCI Express 4.0 and is designed for deployment in any server.

POWERFUL DATA CENTER ACCELERATION

Built for Performance & Efficiency

- Faster application performance from 8GB of HBM memory (32 AXI channel access) and PCIe Gen4 interconnect
- Low latency network capability through 100G networking with support for
- 4x 10GbE, 4x 25GbE, or 1x 40GbE or 1x 100GbE

Optimized Performance Across Broadest Range of Workloads

- Accelerates compute, network, storage workloads
- Maximized application performance as workloads and algorithms evolve through reconfigurable fabric - unlike fixed-architecture alternatives

Deploy in Any Server – From On-premises to Cloud

- Built for scale out architectures and any server - Low-profile form factor and low 75-watt power envelope

Powerful Developer Platform

- Take advantage of a large and growing library of AMD and partner **applications**
- Develop differentiated solutions leveraging **AMD's SDAccel™ development environment** and **Machine Learning Suite**



SPECIFICATION

Features	AMD Alveo™ U50
Architecture	UltraScale+
Form Factor	Half-Height, Half length single slot Low-Profile
Look Up Tables	872,000
HBM2 Memory	8GB
HBM2 Bandwidth	316GB/s1
Network Interface	1 x QSFP28 (100GbE)2
Clock Precision	IEEE 1588
PCI Express	PCIe Gen3 x 16, dual PCIe Gen4 x 8, CCIX
Thermal Solution	Passive
Power (TDP)	75W

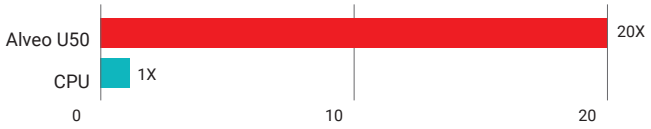
1: For A-U50DD-P00G-ES3-G and A-U50-P00G-PQ-G measured 316 GB/s peak HBM2 bandwidth, 201 GB/s nominal
2: During ES, U50 card will have 2 SFP-DD ports

SUPERCHARGING A BROAD RANGE OF DATA CENTER APPLICATIONS

Computational Storage Acceleration

- Alveo U50 delivers fastest and most flexible compression/ decompression acceleration
- Lower cost – Alveo U50 accelerated compression delivers 33% lower cost. (Based on 10GB/sec throughput and 2:1 compression)

GZIP Compression Throughput (GB/sec)



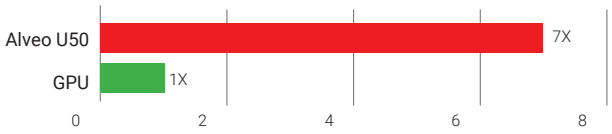
Intel Skylake-SP 6152 @2.10GHz CPU (Ubuntu 16.04)
GB/s compression per CPU core = .0229. Alveo U50=10GB/s (estimate)

Financial Simulation – Grid Computing

- Fastest time to insight
- Reduced operational costs and maximum power efficiency
- Deterministic latency delivers consistent performance

Monte Carlo Simulation

Performance & Efficiency (paths/sec/W)



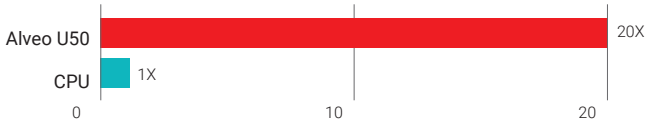
Intel Xeon E5-2697 v4 GCC 5.4.0 Nvidia Tesla V100 16GB PCIe CUDA 10.1 / GCC 5.4.0
Alveo U50 SDAccel 2018.3 (estimate)

Ultra-Low Latency Networking

- 20x lower latency
- Alveo U50 delivers sub-500ns trading time vs CPU latency of 10us
- Deterministic throughput timing

Speedup of Trading Timing

Market data to TCP message (speedup)



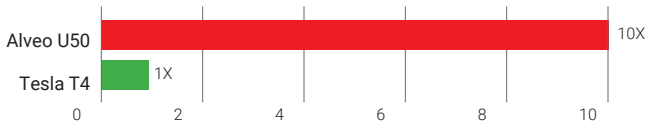
Alveo U50 latency is <0.5us, CPU latency is 10us. Measured from start of packet in on Tick (Market Data) to Start of Packet out on the order to Start of Packet Out on the Order (estimate)

Deep Learning Inference Acceleration

- 10x Higher throughput – translated symbols per second
- 25x lower latency
- ignificantly improved power efficiency per node

Speech Translation Throughput

Transformer NMT (symbols/sec speedup)



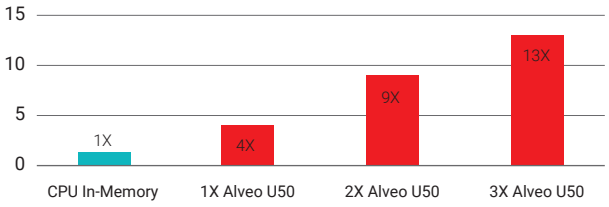
Performance of Alveo U50 - with both Alveo U50 and
Tesla T4 running (B=2, L=8), Tesla T4 (B=8, L=8) (estimate)

Data Analytics Acceleration

- Higher query throughput & response time than CPU
- Higher cost effectiveness per node
- Reduced Operational cost

Database Query Acceleration (TCP-H Query 5)

(Queries / hour speedup)



CPU Query time = 210ms, 34k query/hr. Alveo U50 = 24ms, 150k query/hr
Intel® Xeon® Platinum 8260 Processor (35.75M Cache, 2.40 GHz) 24 core (estimate)

TAKE THE NEXT STEP

Contact your local sales representative or complete the Product Inquiry form at **AMD ALVEO™ U50**

DISCLAIMERS
The information contained herein is for informational purposes only and is subject to change without notice. While every precaution has been taken in the preparation of this document, it may contain technical inaccuracies, omissions and typographical errors, and AMD is under no obligation to update or otherwise correct this information. Advanced Micro Devices, Inc. makes no representations or warranties with respect to the accuracy or completeness of the contents of this document, and assumes no liability of any kind, including the implied warranties of noninfringement, merchantability or fitness for purposes, with respect to the operation or use of AMD hardware, software or other products described herein. No license, including implied or arising by estoppel, to any intellectual property rights is granted by this document. Terms and limitations applicable to the purchase or use of AMD's products are as set forth in a signed agreement between the parties or in AMD's Standard Terms and Conditions of Sale.

© Copyright 2023 Advanced Micro Devices, Inc. All rights reserved. Xilinx, the Xilinx logo, AMD, the AMD Arrow logo, Alveo, Artix, Kintex, Kria, Spartan, Versal, Vitis, Virtex, Vivado, Zynq, and other designated brands included herein are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies. AMBA, AMBA Designer, ARM, ARM1176JZ-S, CoreSight, Cortex, and PrimeCell are trademarks of ARM in the EU and other countries. PCIe, and PCI Express are trademarks of PCI-SIG and used under license.
Printed in the U.S.A. SF11-10-21 PID# 232236903-A