



Adaptable Accelerator Cards for Data Center Workloads

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

The Xilinx® 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 Xilinx UltraScale+™ architecture and packaged in an efficient 75-watt, low-profile form factor, the U50 includes HBM2 with 460 GB/s bandwidth, 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 460GB/s 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 Xilinx and partner [applications](#)
- Develop differentiated solutions leveraging Xilinx's [SDAccel™ development environment](#) and [Machine Learning Suite](#)

Alveo U50



SPECIFICATIONS

FEATURES	ALVEO U50
Architecture	UltraScale+
Form Factor	Half-Height, Half length single slot Low-Profile
Look Up Tables	872,000
HBM2 Memory	8GB
HBM2 Bandwidth	460GB/s
Network Interface	1 x QSFP28 (100GbE)*
Clock Precision	IEEE 1588
PCI Express	PCIe Gen3 x 16, dual PCIe Gen4 x 8, CCIX
Thermal Solution	Passive
Power (TDP)	75W

* During ES, U50 card will have 2 SFP-DD ports

Product Brief

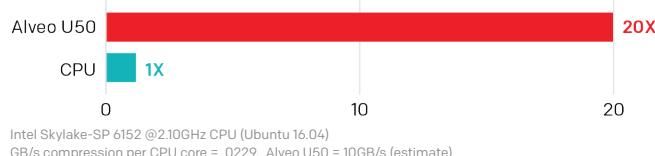
Xilinx Alveo U50

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)

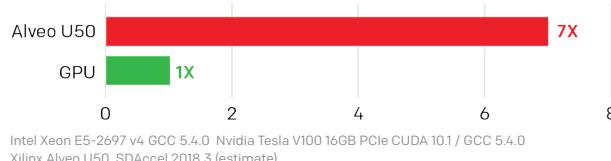


Financial Simulation – Grid Computing

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

Monte Carlo Simulation

Performance & Efficiency (paths/sec/W)

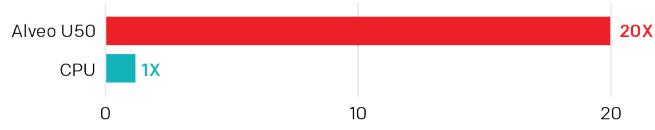


Ultra-Low Latency Networking

- 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)



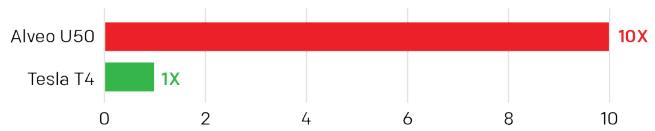
Alveo U50 latency is <0.5us, CPU latency is 10us. Measured from start of packet in on Tick (Market Data) to Start of Packet out on the order to Start of Packet Out on the Order (estimate)

Deep Learning Inference Acceleration

- 10x Higher throughput – translated symbols per second
- 25x lower latency
- Significantly 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 (TPC-H Query 5)

(Queries / hour speedup)



TAKE THE NEXT STEP

Learn more about Xilinx Alveo U50, www.xilinx.com/U50