



U200 | U250

Adaptable Accelerator Cards for Data Center Workloads

OVERVIEW

Xilinx® Alveo™ Data Center accelerator cards are designed to meet the constantly changing needs of the modern Data Center, providing up to 90X performance increase over CPUs for common workloads, including machine learning inference, video transcoding, and database search and analytics.

With complex algorithms evolving faster than silicon design cycles, it's impossible for fixed function GPU and ASIC devices to keep pace. Built on Xilinx 16nm UltraScale™ architecture, Alveo U200 and U250 accelerator cards provide reconfigurable acceleration that can adapt to continual algorithm optimizations, supporting any workload type while reducing overall cost of ownership.

Enabling Alveo accelerator cards is a growing ecosystem of Xilinx and partner applications for common Data Center workloads. For custom solutions, Xilinx's Application Developer tool ([SDAccel™ tool](#)) and [Machine Learning Suite](#) provide the tools for developers to bring differentiated [applications](#) to market.

HIGHLIGHTS

Fast – Highest Performance

- > Up to 90X higher performance than CPUs¹ on key workloads at 1/3 the cost²
- > Over 4X higher inference throughput³ and 3X latency advantage over GPU-based solutions⁴

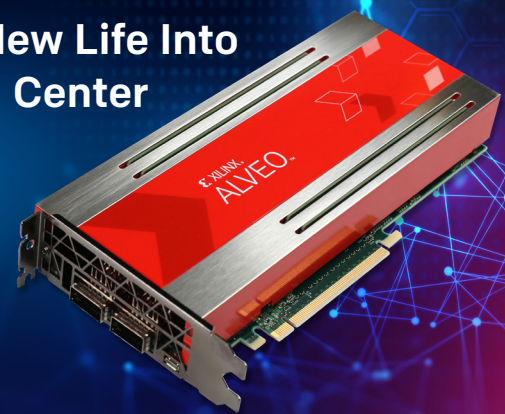
Adaptable – Accelerate Any Workload

- > Machine learning inference to video processing to any workload using the same accelerator card
- > As workloads algorithms evolve, use reconfigurable hardware to adapt faster than fixed-function accelerator card product cycles

Accessible – Cloud ↔ On-Premises Mobility

- > Deploy solutions in the cloud or on-premises interchangeably, scalable to application requirements
- > [Applications available](#) for common workloads, or build your own with the [application developer tool](#)

Breathe New Life Into Your Data Center



ADAPTABLE TO ANY WORKLOAD

- > Database Search & Analytics
- > Financial Computing
- > Machine Learning
- > Storage Compression
- > Video Processing/Transcoding
- > Genomics

1: BlackLynx Elasticsearch on Alveo versus EC2 c4.8xlarge

2: Based on CapEx & OpEx savings for DNN inference on Alveo vs dual-socket Intel Xeon Platinum servers

3: [Accelerating DNNs with Alveo Accelerator Cards White Paper](#)

4: Measured on CNN+BLSTM Speech-to-Text ML inference against Nvidia P4



Adaptable. Intelligent.

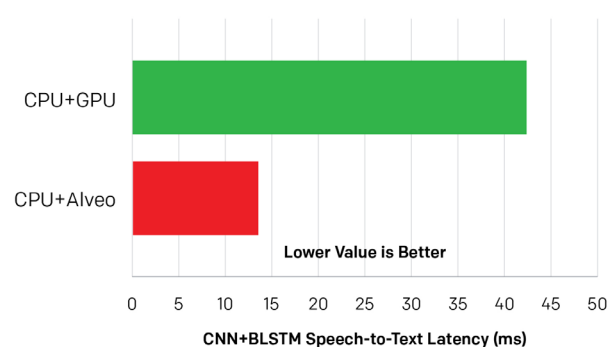
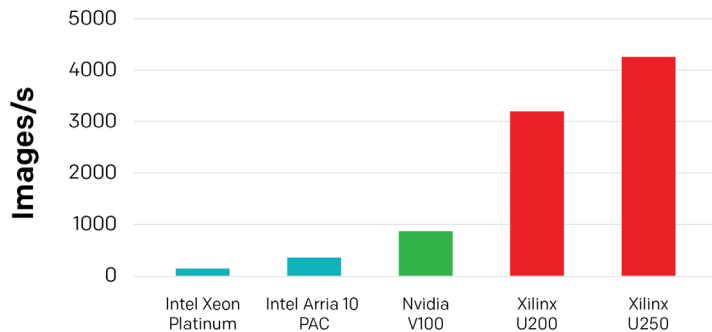
BENCHMARKS

Adapt and Accelerate Any Workload

AREA	PARTNER WORKLOAD	ALVEO ACCELERATION VS CPU
Database Search and Analytics	BlackLynx Unstructured Data Elasticsearch	90X
Financial Computing	Maxeler Value-at-Risk (VAR) Calculation	89X
Machine Learning	Xilinx Real-Time Machine Learning Inference	20X
Video Processing / Transcoding	NGCodec HEVC Video Encoding	12X
Genomics	Falcon Computing Genome Sequencing	10X

CPU Comparisons: Xeon c4.8xlarge AWS | Xeon E5-2643 v4 3.4GHz | Xeon Platinum c5.18xlarge AWS | Dual Socket E5-2680 v3 2.5GHz | Xeon f1.16xlarge

Increase Real-Time Machine Learning* Throughput by 20X Reduce ML Inference Latency by 3X



*GoogLeNet V1: Accelerating DNNs with Xilinx Alveo Accelerator Cards White Paper

CPU+GPU: Nvidia P4 + Xeon CPU E5-2690 v4 @2.60GHz (56 Cores)
 CPU+Alveo: Alveo U200 or U250 + Xeon CPU E5-2686 v4 @2.3GHz (8 Cores)

FEATURES	ALVEO U200 ACCELERATOR CARDS	ALVEO U250 ACCELERATOR CARDS
Peak INT8 TOPs	18.6	33.3
DDR Memory Bandwidth	77GB/s	77GB/s
Internal SRAM Bandwidth	31TB/s	38TB/s
Look-up Tables (LUTs)	892,000	1,341,000
Thermal Options	Passive or Active	Passive or Active

TAKE THE NEXT STEP

Visit www.xilinx.com/alveo for more information. To get started now, visit [Nimblex cloud services](#) to test drive Data Center workloads or purchase Alveo [U200](#) or [U250](#) Data Center accelerator cards for on-premises deployment.

Corporate Headquarters

Xilinx, Inc.
 2100 Logic Drive
 San Jose, CA 95124
 USA
 Tel: 408-559-7778
 www.xilinx.com

Xilinx Europe

One Logic Drive
 Citywest Business Campus
 Saggart, County Dublin
 Ireland
 Tel: +353-1-464-0311
 www.xilinx.com

Japan

Xilinx K.K.
 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

Meenakshi Tech Park
 Block A, B, C, 8th & 13th floors,
 Meenakshi Tech Park, Survey No. 39
 Gachibowli(V), Seri Lingampally (M),
 Hyderabad -500 084
 Tel: +91-40-6721-4747
 www.xilinx.com

