

Cambricon MLU100

MLU100, designed by Cambricon Technologies, is a dedicated processor for deep learning, delivering higher computing performance at lower power consumption, than general purpose processors (CPUs and GPUs).

Designed for a wide range of deep learning systems and applications, Cambricon MLU100 can accelerate numerous AI tasks, including:

- High-accuracy speech, image, and text recognition
- Intelligent video analytics
- Real-time language translation
- Financial modeling, and data analytics
- Online search optimizations, and personalized user recommendations
- Robotics and factory automation
- Autonomous vehicle
- Scientific research, and weather forecasting
- Drug discovery, and disease diagnosis



MLU100-C Series



MLU100-D Series

Cambricon MLU100 can deliver 128 TFlops of deep learning peak performance and support up to 32 channels of 1080P video decoding.

Paired by 16x PCIe gen3 and 8GB/16GB DDR4 memory, Cambricon MLU100 can be deployed among various cloud servers and workstation systems. A detailed specification of Cambricon MLU100 is showed in Table 1.

Specifications

Type	MLU100-C3/C4	MLU100-D3/D4
Core architecture	Cambricon MLU100	
Core frequency	1GHz	
Float performance (FP16)	16 TFLOPS (Dense performance) 64 TFLOPS (Sparse performance)	
Integer performance (INT8)	32 TOPS (Dense performance) 128 TOPS (Sparse performance)	
Video decoding	Yes	No
Memory size	8GB(C3/D3) / 16GB(C4/D4)	
Memory bit width	256-bit	
Memory bandwidth	102.4 GB/s	
Host interface	PCI Express 3.0 x16	
TDP	110W	75W
ECC Support	Yes	

MLU100 Core Strength

Versatility

Support all mainstream deep learning frameworks, accelerate wide ranges of AI applications

Power Efficiency

Dedicated architecture and ISA for deep learning, offering higher performance at lower power over GPU

Sparse Mode

Utilize sparse techniques to boost 128 TFTPDS (INT8) deep learning performance

Comprehensive Software Suite

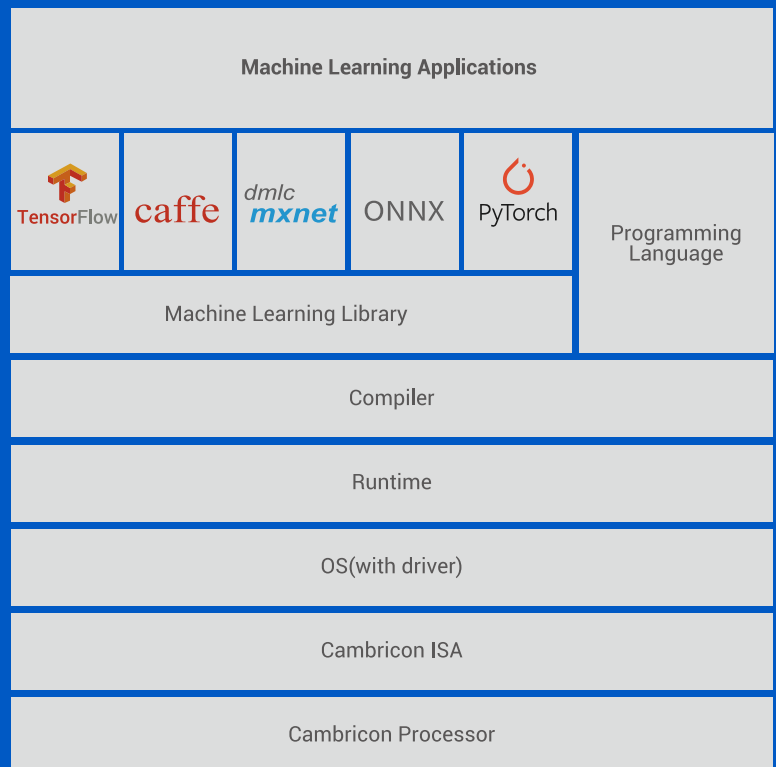
A verified software developer suite, Cambricon NeuWare, offers toolkits from application development, debugging, to performance tuning

Cambricon NeuWare

Cambricon NeuWare, a software developer suite, supports all mainstream deep learning frameworks. (TensorFlow, Caffe, MXNet and ONNX). Leveraging this suite, developers can with minimum efforts, quickly deploy deep learning application on MLU100 processor, and start harvesting its computing performance.

NeuWare provides an extensive toolkits, including:

- Application development tools: machine learning library, runtime library, compiler, model retraining tool and domain specific SDKS (video analytics)
- Debugging tools on different levels: frameworks and libraries
- Performance tuning tools: analysis and monitoring



Contact us: Website: www.cambricon.com  Tel:86-10-83030003 Email:business@cambricon.com

Address:11/F, Block D, Zhizhen Tower, No.7 Zhichun Road, Haidian District Beijing,China