

# MLU100智能处理卡

MLU100智能处理卡是寒武纪推出的第一款智能处理板卡产品，搭载了MLU100芯片，为云端推理提供强大的运算能力支撑。与传统架构处理器相比，MLU100在处理人工智能任务时可获得巨大的性能及能效提升，是真正适合人工智能的处理器。

MLU100智能处理卡的等效理论峰值运算能力128TOPS (INT8)，可通过PCIe接口集成于现有的各类型数据中心的服务器中，半高半长版本功耗不超过75W，全高全长版本功耗不超过110W。MLU100智能处理卡支持最高16GB的DDR4内存容量，并具备ECC数据校验功能。



MLU100-C系列智能处理卡



MLU100-D系列智能处理卡

## 核心特性

### 通用性好

通用智能处理器，支持各类深度学习技术，支持多模态智能处理（视觉、语音和自然语言处理），应用领域广泛。

### 高效能 高可靠性

相比GPU处理器，采用了针对深度学习和人工智能应用特点定制的指令集和处理器架构，具有更优的能效比。

### 支持稀疏化

将稀疏化技术用于智能处理器，通过稀疏化处理，可以达到128TOPS的INT8运算能力。

### 完善的软件开发环境

寒武纪为MLU100智能处理卡提供了一整套成熟的开发环境Cambricon NeuWare，具体包括应用开发、功能调试、性能调优等在内的一系列工具。

## 规格

|                  |  |           |
|------------------|--|-----------|
| 产品型号             | MLU100-C   | MLU100-D  |
| 核心频率             | 1 GHz  |           |
| 半精度浮点运算速度 (FP16) | 16 TFLOPS (关闭稀疏模式时的峰值)<br>64 TFLOPS (打开稀疏模式时的峰值) |           |
| 整数运算速度 (INT8)    | 32 TOPS (关闭稀疏模式时的峰值)<br>128 TOPS (打开稀疏模式时的峰值)    |           |
| 内存容量             | 8GB/16GB   |           |
| 内存位宽             | 256-bit  |           |
| 内存带宽             | 102.4 GB/s                                       |           |
| 系统接口             | PCI Express 3.0 x16                              |           |
| 外形               | 全高全长, 单槽位  | 半高半长, 单槽位 |
| 是否支持解码           | 支持解码   | 不支持解码     |
| TDP功耗            | 110W   | 75W       |
| ECC 保护           | 是  |           |
| 散热方式             | 被动散热   |           |

## 应用领域

### 机器视觉

#### 典型应用

人脸身份认证, 智能交通, 智慧城市, 机器人, 无人车, 图像搜索引擎, 图像/视频的理解和美化



### 语音识别

#### 典型应用

语音输入, 语音控制, 智能助手, 机器翻译, 机器人听觉



### 自然语言

#### 典型应用

聊天机器人, 智能助手, 智能客服, 视频理解, 机器翻译



## 应用场景

### 视觉智能处理

内置硬件编解码引擎, 强大的INT8算力, 超低的计算延迟, 可以支持最高32路高清视频或480fps高清图像的解码和智能处理。

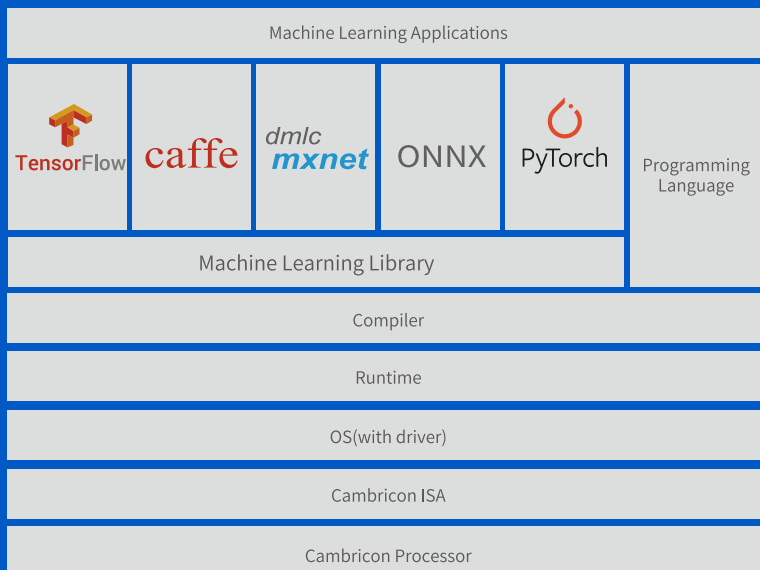
### 大数据处理

相比普通CPU/GPU服务器, 具有更高的吞吐量和更低的延迟。

## 寒武纪提供软件开发环境 Cambricon NeuWare

NeuWare全面支持各类主流编程框架(如TensorFlow、Caffe、MXNet和ONNX等)。用户可面向上述编程框架, 便捷地在寒武纪MLU100上开发和部署深度学习应用。同时, NeuWare提供了完整的运行时系统和驱动软件, 方便系统快速集成。

NeuWare还提供了包括应用开发、功能调试、性能调优等在内的一系列工具。其中应用开发工具包括机器学习库、运行时库、编译器、模型重训练工具和特定领域(如视频分析领域) SDK等; 功能调试工具可以满足编程框架、函数库等不同层次的调试需求; 性能调优工具包括性能剖析工具和系统监控工具等。



如需详细了解, 请访问 [www.cambricon.com](http://www.cambricon.com)

电话: 86-10-83030003 邮箱: [business@cambricon.com](mailto:business@cambricon.com) 地址: 北京市海淀区知春路7号致真大厦D座11层