

# Atlas 200 AI加速模块

型号：3000



## 极致性能

- 半张信用卡大小即可提供16 TOPS INT8算力，支持16路高清视频实时分析
- 多级算力配置，支持16/8 TOPS两级算力

## 超低功耗

- 支持毫瓦级休眠、毫秒级唤醒，最大功耗仅9.5 W，使能边缘AI应用

## 应用场景



### 智能摄像机

输电线路检测、人脸识别



### 智能机器人

智能巡查、无人送餐



### 无人机

输电线路监测

Atlas 200 AI加速模块（型号：3000）集成了昇腾310 AI处理器，可以在端侧实现人脸识别、图像分类等，广泛用于智能摄像机、机器人、无人机等端侧AI场景。

## 产品规格

AI芯片	昇腾310
AI算力	16 TOPS INT8 8 TOPS INT8
内存规格	LPDDR4X, 8 GB/4 GB, 总带宽51.2 GB/s
编解码能力	<ul style="list-style-type: none"><li>• 支持H.264硬件解码, 16路1080P 30 FPS (2路3840*2160 60 FPS)</li><li>• 支持H.265硬件解码, 16路1080P 30 FPS (2路3840*2160 60 FPS)</li><li>• 支持H.264硬件编码, 1路1080P 30 FPS</li><li>• 支持H.265硬件编码, 1路1080P 30 FPS</li><li>• JPEG解码能力1080P 256 FPS, 编码能力1080P 64 FPS, 最大分辨率: 8192*4320</li><li>• PNG解码能力1080P 24 FPS, 最大分辨率: 4096*2160</li></ul>
接口	<ul style="list-style-type: none"><li>• ×4 PCIe Gen3.0</li><li>• ×1 USB2.0 / USB3.0</li><li>• ×1 RGMII</li></ul>
串行总线	UART / I2C / SPI
接口规格	144 pin BTB连接器
功耗	最大9.5 W
工作环境温度	-25°C ~ 80°C
结构尺寸	52.6 mm * 38.5 mm * 8.5 mm

# Atlas 200 DK 开发者套件

型号：3000



## 高集成度

- 基于华为昇腾310AI处理器，集成丰富的外设接口和 Mind Studio开发环境，方便开发者简捷接入、快速开发

## 易用的软件环境

- Mind Studio编程界面友好，具备图形化调试能力，可全自动管理离线模型并提供仿真环境

## 应用场景



### 开发者方案验证

模型验证  
方案验证



### 高校教育

AI入门教育  
人才培养



### 科学研究

应用研究  
算法研究

Atlas 200 DK 开发者套件（型号：3000）是一款高性能AI应用开发板，集成了昇腾310 AI处理器，方便用户快速开发、快速验证，可广泛应用于开发者方案验证、高校教育、科学研究等场景。

## 产品规格

AI芯片	昇腾310
AI算力	16 TOPS INT8 8 TOPS INT8
内存规格	LPDDR4X, 8 GB, 总带宽51.2 GB/s
编解码能力	<ul style="list-style-type: none"><li>• 支持H.264 硬件解码，16路1080P 30 FPS (2路3840*2160 60 FPS)</li><li>• 支持H.265 硬件解码，16路1080P 30 FPS (2路3840*2160 60 FPS)</li><li>• 支持H.264 硬件编码，1路1080P 30 FPS</li><li>• 支持H.265 硬件编码，1路1080P 30 FPS</li><li>• JPEG解码能力1080P 256 FPS，编码能力1080P 64 FPS，最大分辨率：8192*4320</li><li>• PNG解码能力1080P 24 FPS，最大分辨率：4096*2160</li></ul>
接口	<ul style="list-style-type: none"><li>• 网络：1个GE RJ45</li><li>• USB：1个USB2.0 / USB3.0</li><li>• 其他：1个40 pin IO连接器</li></ul>
电源	5~28 V DC，默认配置12 V / 3 A适配器
功耗	典型功耗20 W
工作环境温度	0°C ~ 45°C
结构尺寸	137.8 mm * 93.0 mm * 32.9 mm

# Atlas 300 推理卡

型号：3000



## 超强算力

- 单卡提供64 TOPS INT8算力，支持64路高清视频实时分析，为边缘推理提供更强大算力

## 硬件编解码

- 支持JPEG和视频硬件编解码，提升图片和视频类应用性能

## 低时延

- 内存容量大、带宽高，满足特征匹配场景内存需求，降低应用时延

## 应用场景



**智慧城市**  
人脸识别  
车牌识别



**智慧交通**  
信号灯优化  
智能疏导



**智慧社区**  
住户识别  
陌生人识别



**智能制造**  
智能质检  
柔性制造



**无人零售**  
无人货柜  
智慧门店



**智能楼宇**  
身份验证  
融合分析

Atlas 300 推理卡（型号：3000）基于昇腾 310 AI 处理器，提供超强AI推理性能，单卡算力可达64 TOPS INT8，支持64路高清视频实时分析，可广泛应用于智慧城市、智慧交通、智慧金融等场景。

## 产品规格

形态	半高半长PCIe卡
AI芯片	昇腾310
AI算力	64 TOPS INT8
内存规格	LPDDR4X 32 GB，总带宽204.8 GB/s
编解码能力	<ul style="list-style-type: none"><li>• 支持H.264硬件解码，64路1080P 30 FPS（8路 3840*2160 60 FPS）</li><li>• 支持H.265硬件解码，64路1080P 30 FPS（8路 3840*2160 60 FPS）</li><li>• 支持H.264硬件编码，4路1080P 30 FPS</li><li>• 支持H.265硬件编码，4路1080P 30 FPS</li><li>• JPEG解码能力4*1080P 256 FPS，编码能力4*1080P 64 FPS，最大分辨率：8192*4320</li><li>• PNG解码能力4*1080P 48 FPS，最大分辨率：4096*2160</li></ul>
PCIe	x16 PCIe Gen3.0
功耗	最大67 W
工作环境温度	0°C ~ 55°C
结构尺寸	169.5 mm * 68.9 mm

# Atlas 500 智能小站

型号：3000



## 智能边缘

- 业界领先的集成AI处理能力的边缘产品
- 无风扇散热，支持-40℃至70℃室外工作

## 小身材大能量

- 机顶盒大小即支持16 TOPS INT8算力
- 实现16路高清视频处理能力

## 边云协同

- 支持WiFi和LTE无线传输
- 云边协同，模型实时更新
- 可在云端统一进行设备管理和固件升级

## 应用场景



**平安城市**  
人脸识别  
车牌识别



**智慧交通**  
信号灯优化  
智能疏导



**智慧社区**  
住户识别  
陌生人识别



**环境监测**  
污染物检测



**智能制造**  
智能质检  
柔性制造



**智慧营业厅**  
客流分析  
客户精准识别



**无人零售**  
无人货柜  
智慧门店



**智能楼宇**  
身份验证  
融合分析

Atlas 500智能小站 (型号：3000) 是面向边缘应用的产品，具有超强计算性能、体积小、环境适应性强、易于维护和支持云边协同等特点，可以在边缘环境广泛部署，满足在安防、交通、社区、园区、商场、超市等复杂环境区域的应用需求。

## 产品规格

AI芯片	昇腾310
AI算力	16 TOPS INT8
内存规格	LPDDR4X, 8 GB / 4 GB, 最大51.2 GB/s
编解码能力	支持H.264硬件解码, 16路1080P 30 FPS (2路3840*2160 60 FPS) 支持H.265硬件解码, 16路1080P 30 FPS (2路3840*2160 60 FPS) 支持H.264硬件编码, 1路1080P 30 FPS 支持H.265硬件编码, 1路1080P 30 FPS JPEG解码能力1080P 256 FPS, 编码能力1080P 64 FPS, 最大分辨率: 8192*4320 PNG解码能力1080P 24 FPS, 最大分辨率: 4096*2160
接口	网络: 2个GE RJ45 其他I/O: 1个HDMI接口, 1对3.5 mm立体声输入输出接口; 2个外部和1个内部USB2.0接口 (Type-A)
典型功耗	无盘配置: 25 W 有盘配置: 40 W
环境条件	无盘配置: -40℃ ~ 70℃ 有盘配置: -40℃ ~ 60℃
结构尺寸	无盘配置: 45 mm * 235 mm * 220 mm 有盘配置: 45 mm * 355 mm * 220mm

# Atlas 800 推理服务器

型号：3000



## 超强算力

- 支持8张Atlas 300 AI推理加速卡，满足多场景推理需求；整机可提供512路高清视频实时分析
- 搭载2个业界性能最高、具有超强算力的鲲鹏920处理器，高效加速应用

## 超高能效

- 发挥鲲鹏架构多核、低功耗优势，为推理场景构建高效能、低功耗的AI计算平台
- Atlas 300 单卡功耗仅67W，为AI服务器算力加速同时提供更优的能效比

## 应用场景



精准营销



医疗影像分析



视频分析



OCR



智慧零售



智慧医疗



智慧城市



智慧金融

Atlas 800 推理服务器（型号：3000）是基于昇腾310芯片的推理服务器，最大可支持8个Atlas 300 AI推理加速卡，提供强大的实时推理能力，广泛应用于中心侧AI推理场景。

## 产品规格

形态	2U AI服务器
CPU	2 * 鲲鹏920
CPU内存	32个DDR4内存插槽，最高2933 MT/s
AI加速卡	最大支持8个Atlas 300 AI加速卡 满足512路人/车/物视频智能分析
AI算力	512 TOPS INT8
NPU内存	最大256 GB，总带宽最大1638.4 GB/s
本地存储	25*2.5 SAS/SATA 12*3.5 SAS/SATA 8*2.5 SAS/SATA+12x2.5 NVMe
RAID支持	RAID 0/1/10/5/50/6/60等
PCIe	最多支持9个PCIe4.0 PCIe接口，其中1个为RAID扣卡专用的PCIe扩展槽位，另外8个为标准的PCIe扩展槽位
电源	2个热插拔900 W或2000 W交流电源模块，支持1+1冗余备份
风扇	4个热插拔风扇，支持N+1冗余备份
工作环境温度	5°C ~ 40°C
结构尺寸	447 mm * 790 mm * 86.1 mm

# Atlas 800 推理服务器

型号：3010



## 灵活配置，适配多项负载

- 支持SAS/SATA/NVMe/M.2 SSD硬盘多种组合灵活配置
- 支持板载网卡和灵活I/O卡，提供丰富多样的网络接口

## 智能视频分析

- 支持7个AI加速卡，支持各种视频图像分析
- 借助Atlas 300专用解码引擎，对多达448路高清视频流进行实时转码和推理

## 应用场景



精准营销



医疗影像分析



视频分析



OCR



智慧零售



智慧医疗



智慧城市



智慧金融

Atlas 800 推理服务器（型号：3010）是基于Intel处理器的推理服务器，最多可支持7个Atlas 300 推理加速卡，最大支持448路高清视频实时分析，广泛应用于中心侧AI推理场景。

## 产品规格

形态	2U AI服务器
CPU	1/2个Intel® Xeon® SP Skylake 或 Cascade Lake处理器，最高205W
CPU内存	24个DDR4内存插槽，最高2933 MT/s
AI加速卡	最大支持7个Atlas 300 AI推理加速卡，支持448路视频智能分析
AI算力	448 TOPS INT8
NPU内存	最大224 GB，总带宽最大1433.6 GB/s
本地存储	8*2.5 SAS/SATA 12*3.5 SAS/SATA 8* 2.5 SAS/SATA+12*2.5 NVMe 24*2.5 SAS/SATA 24*2.5 NVMe 25*2.5 SAS/SATA
RAID支持	RAID 0/1/5/6/10/1E/50/60等
PCIe	x10 PCIe Gen3.0 (含1个RAID控制卡+1个灵活LOM)
电源	可配置2个冗余热插拔电源，支持1+1冗余备份，选择规格如下： 550 W AC 白金电源、900 W AC 白金/钛金电源、1500 W AC 白金电源 1500 W 380 V 高压直流电源、1200 W -48 V ~ -60 V 直流电源
风扇	4个热插拔风扇，支持N+1冗余备份
工作环境温度	5°C ~ 45°C
结构尺寸	3.5英寸硬盘机箱尺寸： 86.1 mm * 447 mm * 748 mm 2.5英寸硬盘机箱尺寸： 86.1 mm * 447 mm * 708 mm

# Atlas 800 训练服务器

型号：9000



## 最强算力密度

- 4U高度提供 2 PFLOPS FP16超强算力
- 算力密度达到业界2.5倍

## 超高能效

- 单机支持风冷和液冷两种散热方式
- 提供2 PFLOPS/5.5 kW<sup>①</sup>超高能效比，达到业界1.8倍

## 高速网络带宽

- 8\*100G RoCE v2高速接口，宽带达到业界2倍
- 芯片间跨服务器互联时延缩短10~70%

## 应用场景



模型训练



HPC



智慧城市



智慧医疗



天文探索



石油勘探

Atlas 800 训练服务器（型号：9000）是基于华为鲲鹏920+昇腾910处理器的AI训练服务器，具有最强算力密度、超高能效与高速网络带宽等特点。该服务器广泛应用于深度学习模型开发和训练，适用于智慧城市、智慧医疗、天文探索、石油勘探等需要大算力的行业领域。

## 产品规格

形态	4U AI服务器
CPU	4 * 鲲鹏920
CPU内存	<ul style="list-style-type: none"><li>• 最多32个DDR4内存插槽，支持RDIMM</li><li>• 内存速率最高2933 MT/s</li><li>• 单根内存条容量支持32 GB/64 GB</li></ul>
AI处理器	8 * 昇腾910
HBM	32 GB, 1200 GB/s
NPU内存	16 GB, 总带宽38.4 GB/s
AI算力	2 PFLOPS FP16
本地存储	<ul style="list-style-type: none"><li>• 2 * 2.5 SAS/SATA+3 * 2.5 NVMe</li><li>• 2 * 2.5 SATA+3 * 2.5 NVMe</li><li>• 2 * 2.5 SAS/SATA+6 * 2.5 NVMe</li><li>• 2 * 2.5 SATA+6 * 2.5 NVMe</li><li>• 2 * 2.5 SATA+8 * 2.5 SAS/SATA</li></ul>
RAID支持	支持 RAID 0/1/10/5/50/6/60
网络	8 * 100GE+ 4* 25GE/2 * 100GE
PCIe扩展	最多支持2个PCIe 4.0扩展插槽
电源	4个热插拔3 kW交流电源模块，支持2+2冗余
供电	<ul style="list-style-type: none"><li>• 200 ~ 240 V AC</li><li>• 240 V DC</li></ul>
功耗	最大功耗5.5 kW <sup>①</sup>
散热方式	风冷/液冷
风扇	支持8个热插拔风扇模组，支持N+1冗余
温度	<ul style="list-style-type: none"><li>• 工作温度：5°C ~ 35°C</li><li>• 存储温度：-40°C ~ 70°C</li></ul>
结构尺寸	175 mm * 447 mm * 790 mm

① 持续调优中，数值根据调优结果动态更新