Atlas 300I Pro 推理卡

用户指南

文档版本	06
发布日期	2022-09-01

华为技术有限公司

商标声明


和其他华为商标均为华为技术有限公司的商标。
本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： https://e.huawei.com

前言

概述

本文档详细的描述了华为Atlas 300I Pro 推理卡的外观外形、产品规格、所需工具的安装和使用方法，以及日常管理等内容。

读者对象

本文档主要适用于以下工程师：

- 企业管理员
- 企业终端用户

符号约定

在本文中可能出现下列标志，它们所代表的含义如下。

符号	说明
危险	表示如不避免则将会导致死亡或严重伤害的具有高等级风险的危害。
警告	表示如不避免则可能导致死亡或严重伤害的具有中等级风险的危害。
注意	表示如不避免则可能导致轻微或中度伤害的具有低等级风险的危害。
须知	用于传递设备或环境安全警示信息。如不避免则可能会导致设备损坏、数据丢失、设备性能降低或其它不可预知的结果。 “须知”不涉及人身伤害。
说明	对正文中重点信息的补充说明。 “说明”不是安全警示信息，不涉及人身、设备及环境伤害信息。

命令行格式约定

格式	意义
粗体	命令行关键字（命令中保持不变、必须照输的部分）采用**加粗**字体表示。
斜体	命令行参数（命令中必须由实际值进行替代的部分）采用*斜体*表示。
[ ]	表示用“[ ]”括起来的部分在命令配置时是可选的。
{ x | y | ... }	表示从两个或多个选项中选取一个。
[x | y | ...]	表示从两个或多个选项中选取一个或者不选。
{ x | y | ... }*	表示从两个或多个选项中选取多个，最少选取一个，最多选取所有选项。
[x | y | ...]*	表示从两个或多个选项中选取多个或者不选。
&<1-n>	表示符号&前面的参数可以重复1～n次。
#	由“#”开始的行表示为注释行。

修改记录

文档版本	发布日期	修改说明
06	2022-09-01	第六次正式发布。 修改**4.1 基本规格**。
05	2022-07-28	第五次正式发布。 修改AI处理器的名称。
04	2022-03-07	第四次正式发布。 **7 安装驱动和固件**修改文档链接。
03	2022-02-14	第三次正式发布。 刷新**7 安装驱动和固件**。
02	2021-11-05	第二次正式发布。 调整大纲。将“安装与维护”章节拆分为**6 安装硬件**和**7 安装驱动和固件**。
01	2021-08-18	第一次正式发布。

目 录

1 安全

1.1 通用安全注意事项

在安装、操作、维护华为公司制造的设备时，本文档介绍的所应遵守的部分安全注意事项可指导选择测量设备和测试设备。

所有安全注意事项

为保障人身和设备安全，在安装、操作和维护设备时，请遵循设备上标识及手册中说明的所有安全注意事项。

手册中的“注意”、“警告”和“危险”事项，并不代表所应遵守的所有安全事项，只作为所有安全注意事项的补充。

当地法规和规范

操作设备时，应遵守当地法规和规范。手册中的安全注意事项仅作为当地安全规范的补充。

基本安装要求

负责安装维护华为设备的人员，必须先经严格培训，了解各种安全注意事项，掌握正确的操作方法之后，方可安装、操作和维护设备。

- 只允许有资格和培训过的人员安装、操作和维护设备。
- 只允许有资格的专业人员拆除安全设施和检修设备。
- 替换和变更设备或部件（包括软件）必须由华为认证或授权的人员完成。
- 操作人员应及时向负责人汇报可能导致安全问题的故障或错误。

接地要求

以下要求只针对需要接地的设备：

- 安装设备时，必须先接地；拆除设备时，最后再拆地线。
- 禁止破坏接地导体。
- 禁止在未安装接地导体时操作设备。
- 设备应永久性地接到保护地。操作设备前，应检查设备的电气连接，确保设备已可靠接地。

人身安全

- 禁止在雷雨天气下操作设备和电缆。
- 雷雨天气时，应拔掉交流电源连接器、禁止使用固定终端、禁止触摸终端和天线连接器。

说明

上述两则要求适用于无线固定台终端。

- 为避免电击危险，禁止将安全特低电压（SELV）电路端子连接到通讯网络电压(TNV)电路端子上。
- 禁止裸眼直视光纤出口，以防止激光束灼伤眼睛。
- 操作设备前，应穿防静电工作服，佩戴防静电手套或腕带，并去除首饰和手表等易导电物体，以免被电击或灼伤。
- 如果发生火灾，应撤离建筑物或设备区域并按下火警警铃，或者拨打火警电话。任何情况下，严禁再次进入燃烧的建筑物。

设备安全

- 操作前，应先将设备可靠地固定在地板或其他稳固的物体上，如墙体或安装架。
- 系统运行时，请勿堵塞通风口。
- 安装面板时，如果螺钉需要拧紧，必须使用工具操作。
- 安装完设备，请清除设备区域的空包装材料。

1.2 设备上的标志

介绍设备上的标志，有警告标志、接地标志和防静电标志。

表 1-1 安全标志

图示	名称	说明
	警告标志	该标志表示误操作可能会导致设备损坏或人身伤害。
	外部接地标志	该标志是设备外部的接地标识。接地电缆的两端分别接在不同设备上，表示设备必须通过接地点接地，保证设备能够正常运行，同时保证操作人员的人身安全。

图示	名称	说明
	内部接地标志	该标志是设备内部的接地标识。接地电缆的两端都接在同一个设备上的不同组件上，表示设备必须通过接地点接地，保证设备能够正常运行，同时保证操作人员的人身安全。
	防静电标志	该标志表示为静电敏感区，请勿徒手触摸设备。在该区域操作时，请采取严格的防静电措施，例如佩戴防静电腕带或者防静电手套。

1.3 电气安全

介绍高压、雷雨、大漏电流、电源线、保险丝、静电放电的安全注意事项。

高压

危险

- 高压电源为设备的运行提供电力，直接接触或通过潮湿物体间接接触高压电源，会带来致命危险。
- 不规范、不正确的高压操作，会引起火灾或电击等意外事故。

雷雨天气

此要求仅适用于无线基站或带有天馈线的设备。

危险

禁止在雷雨天气下进行高压、交流电操作及铁塔、桅杆作业，否则会有生命危险。

大漏电流

危险

- 在接通电源之前设备必须先接地，否则会危及人身及设备安全。
- 如果设备电源端子附近粘贴有“大漏电流”标志，在连接交流输入电源之前，必须先将设备机壳的保护接地端子接地，以防止设备的漏电流对人体产生电击。

电源线

危险

- 禁止带电安装或拆除电源线。电源线芯在接触导体的瞬间，会产生电弧或电火花，可导致火灾或眼睛受伤。
- 安装、拆除电源线之前，必须先关闭电源开关。
- 连接电源线之前，必须先确认电源线标签标识正确再进行连接。

保险丝

须知

为保证设备运行安全，当设备上的保险丝熔断后，应使用相同型号和规格的保险丝替换。

静电放电

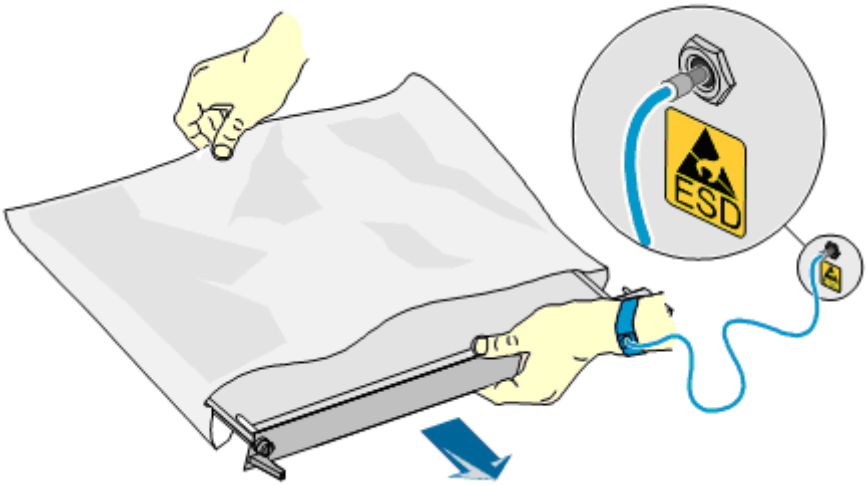
须知

人体产生的静电会损坏单板上的静电敏感元器件，如大规模集成电路（LSI）等。

- 在人体移动、衣服摩擦、鞋与地板的摩擦或手拿普通塑料制品等情况下，人体会产生静电电磁场，在放电前不易消失。
- 在接触设备，手拿单板或专用集成电路（ASIC）芯片等之前，为防止人体静电损坏敏感元器件，必须佩戴防静电腕带，并将防静电腕带的另一端良好接地。

  防静电腕带佩戴如图1-1所示。

图 1-1 佩戴防静电腕带示意图



2 产品简介

2.1 概述

Atlas 300I Pro 推理卡是基于Ascend 310P处理器的新一代高性能推理卡，融合“通用处理器、AI Core、编解码” 于一体，提供超强AI推理、目标检索等功能，具有超强算力、超高能效、高性能特征检索、安全启动等优势，可广泛应用于OCR识别、语音识别、搜索推荐、内容审核等诸多AI应用场景。

2.2 外观

Atlas 300I Pro 推理卡外观如**图2-1**和**图2-2**所示。

图 2-1 Atlas 300I Pro 推理卡半高拉手条外观图

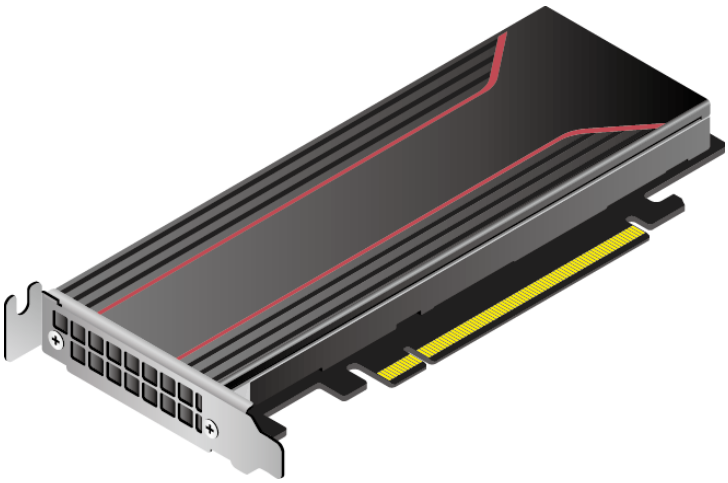
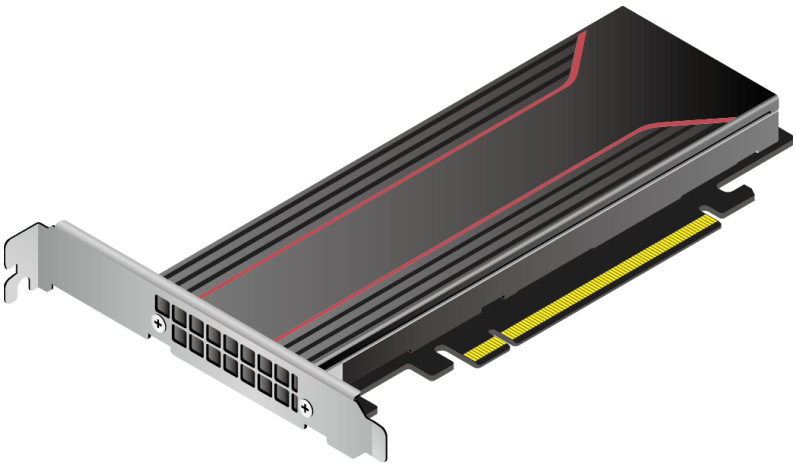


图 2-2 Atlas 300I Pro 推理卡全高拉手条外观图



3 产品特点

3.1 性能特点

采用1个高性能低功耗的Ascend 310P处理器，最高可提供140TOPS INT8的计算能力。

3.2 可维护性特点

- 支持带内的在线升级，方便客户进行日常维护。
- 支持带内及带外获取温度、电压、功耗等设备状态信息，图形界面让管理更简单。
- 完备的命令行管理功能，用户可以通过各种命令进行日常的设备管理。
- 支持带内及带外资产管理功能，提供生产日期、序列号等信息，方便资产管理。

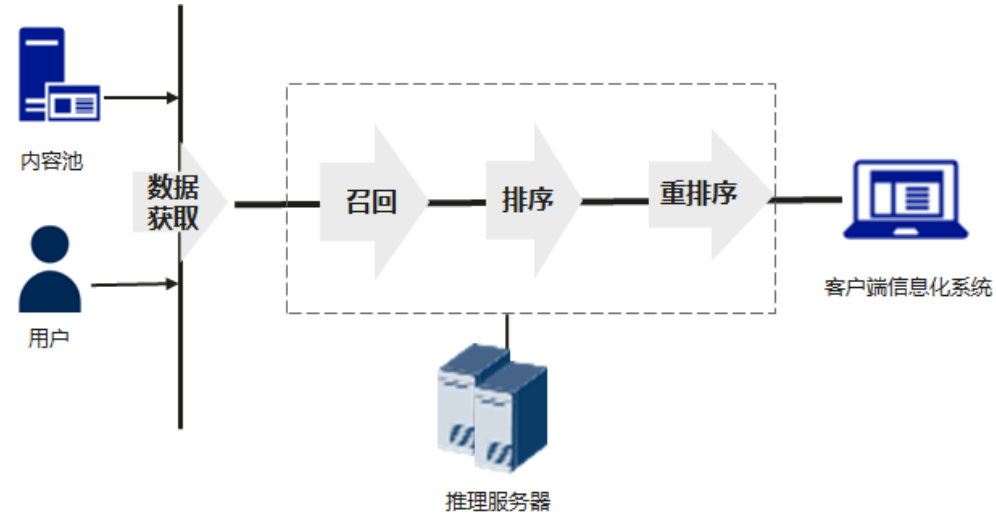
3.3 典型应用场景

Atlas 300I Pro 推理卡典型应用场景为搜索推荐、内容审核和OCR系统。

搜索推荐系统主要根据用户输入（用户画像、搜索词等），通过召回和排序算法，在内容池中筛选出最终推荐的素材（视频、文本等）。主要应用在互联网等领域。

搜索推荐系统如**图3-1**所示，主要部件有推理服务器、客户端信息化系统软件组成，Atlas 300I Pro 推理卡部署在推理服务器中，主要实现用户数据类别召回、排序、重排序等推理功能。

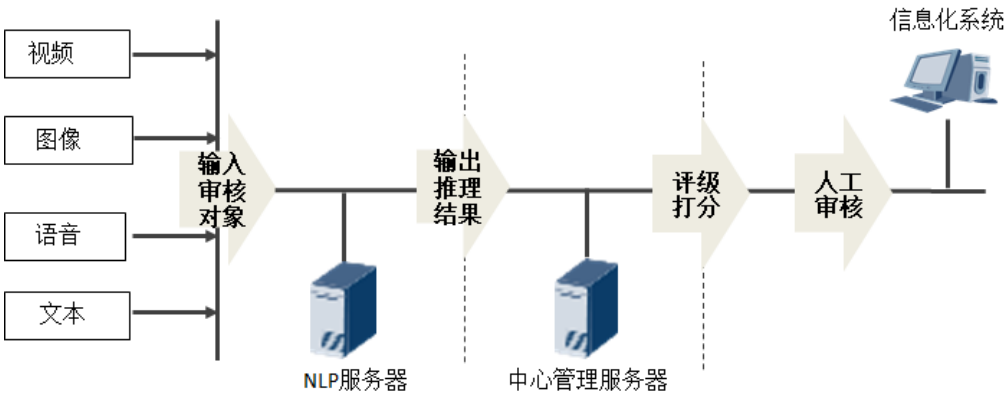
图 3-1 搜索推荐系统架构图



内容审核系统主要采用了数据评级打分算法，实现了视频，图像，语音，文本等审核功能。主要应用在互联网等领域。

内容审核系统如图3-2所示，主要部件有NLP服务器、中心管理服务器、信息化系统软件组成。Atlas 300I Pro 推理卡部署在NLP服务器中，主要实现视频、图像、语音、文本的审核校验等推理功能。

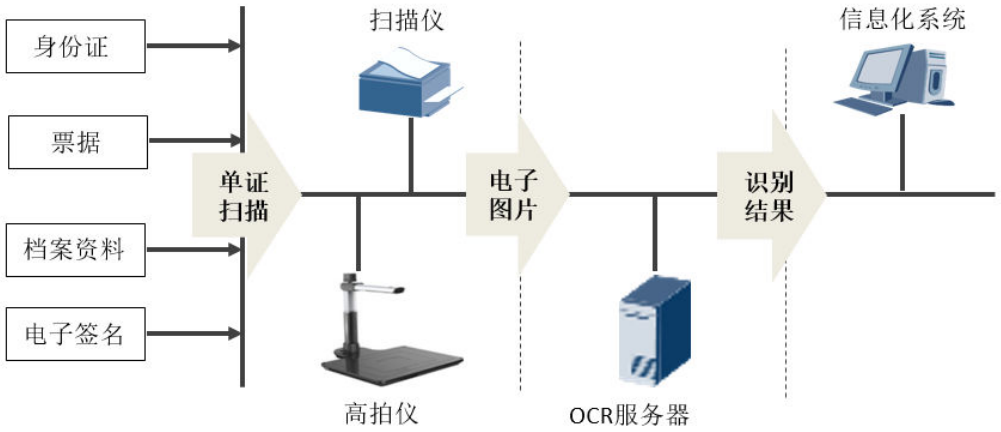
图 3-2 内容审核系统架构图



OCR系统主要采用了文本检测和文本识别算法，实现了身份证实名认证、票据识别、档案资料和信息录入、电子签名识别等功能。主要应用在智慧政务、智慧金融等领域。

OCR系统架构如图3-3所示，主要部件有扫描仪、OCR服务器、中心管理服务器、信息化系统等软件组成。Atlas 300I Pro 推理卡部署在OCR服务器中，主要实现图片预处理、文本检测、文本识别、后处理校验等推理功能。

图 3-3 典型的 OCR 系统架构图



4 产品规格

4.1 基本规格

表 4-1 Atlas 300I Pro 推理卡规格

特征	规格
形态	HHHL Low Profile标卡，支持全高和半高两种拉手条
AI处理器	1 x Ascend 310P 处理器 • 8个DaVinci AI Core • 8个自研CPU核
内存规格	• LPDDR4X • 容量：24GB • 位宽：384bit • 速率：4266Mbps • 总带宽：204.8GByte/s • 支持ECC
AI算力	• 半精度（FP16）：70 TFLOPS • 整数精度（INT8）：140 TOPS
CPU算力	8 core * 1.9 GHz

特征	规格
编解码能力	• H.264、H.265视频编解码 • JPEG图片编解码
PCIe接口	• x16 Lanes，兼容x8/x4/x2 • PCIe Gen4.0，兼容3.0/2.0/1.0
PCI IDs	Vendor ID: 0x19E5 Device ID: 0xD500 Subsystem Vendor ID: 0x0200 Subsystem Device ID: 0x0100
单板功耗	72W
尺寸（长×高×宽）	169.5mm x 68.9mm x 18.45mm
重量（g）	280g
操作系统	详细信息请参见**计算产品兼容性查询助手**。

4.2 环境条件

Atlas 300I Pro 推理卡硬件应用环境条件如**表4-2**所示。

表 4-2 Atlas 300I Pro 推理卡硬件应用环境条件

环境指标	规格
工作温度	0℃～55℃（32℉～131℉）
存储温度	-40℃～+75℃（-40℉～+167℉）
工作湿度	5%RH～90%RH（非冷凝）
存储湿度	5%RH～95%RH（非冷凝）
海拔高度	小于3050m。高于900m使用时，海拔每升高300m最高温度规格降低1℃。

4.3 时钟要求

Atlas 300I Pro 推理卡遵从标准PCIe标卡协议（PCI Express® Card Electromechanical Specification Revision 4.0），整卡只需要提供标准PCIe 4.0（可向下兼容 3.0、2.0及1.0）差分时钟，信号质量满足PCIe规范。

4.4 热插拔

Atlas 300I Pro 推理卡仅支持通知式热插拔，不支持暴力热插拔。

4.5 电源管理

Atlas 300I Pro 推理卡遵从标准PCIe标卡协议（PCI Express® Card Electromechanical Specification Revision 4.0/3.0），单板功耗72W，要求对应Atlas 300I Pro 推理卡槽位可提供5.5A@12V及2A@3.3V标准供电能力。

4.6 散热规格

4.6.1 散热要求

Atlas 300I Pro 推理卡用于带风扇的主动散热环境，支持双向进风出风，风量必须满足散热要求。

表 4-3 Atlas 300I Pro 推理卡散热要求

入风口平均温度/℃	需求最低风量/CFM	压降/Inch H_2O
55	7.0	0.52
50	5.3	0.30
45	4.3	0.22
40	3.9	0.19
35	3.3	0.15
30	2.9	0.12
任何场景	2.9	0.12

说明

- 需求的最低风量为通过Atlas 300I Pro 推理卡散热器风量。
- 散热器入口环境温度为进风口的平均温度。
- 需求的风量是建议值，不同系统提供给Atlas 300I Pro 推理卡的风量和温度可能存在差异，需要根据实际系统进行实测确定。
- Atlas 300I Pro 推理卡上电状态，需要有风量进行散热，需求的最低风量为2.9CFM。

4.6.2 散热规格

Atlas 300I Pro 推理卡支持入口温度为0℃～55℃，内部有温度监控点，带内及带外均可对Ascend 310P、存储芯片进行实时监控，确保该卡在工作过程中，系统的散热情况需保证该卡的温度值低于规格值。

表 4-4 关键器件温度规格

规格	Ascend 310P温度 ℃	存储芯片温度 ℃
下电温度	106	100
降频温度	103	90
长期工作温度	102	85

4.6.3 过温保护

Atlas 300I Pro 推理卡支持带外及带内通道检测Ascend 310P及存储芯片等关键器件的结温，同时也支持检测整板温度。

Atlas 300I Pro 推理卡的主要器件Ascend 310P及其存储芯片最高支持入风口温度为55℃，为保证可靠工作，外界需提供散热需求的风量，并设计了如下温控策略，Atlas 300I Pro 推理卡使用了2级预警机制：

- 第一级为严重告警，Ascend 310P芯片的严重告警阈值为103℃，存储芯片的严重告警阈值为90℃。当芯片结温或环境温度达到该值，固件就会限制设备的性能。
- 第二级为致命告警，Ascend 310P芯片的致命告警阈值为106℃，存储芯片的致命告警阈值为100℃。当芯片结温或环境温度达到该值，Atlas 300I Pro 推理卡会启动自身下电。

5 信号管脚

5.1 管脚定义

Atlas 300I Pro 推理卡遵从标准PCIe标卡协议（PCI Express® Card Electromechanical Specification Revision 4.0），对外提供PCIe X16金手指物理接口，主要提供的数据信号为：一组PCIe 4.0 x16，最大速率16Gbps/lane，用于业务数据的交互传递，一组SMBUS，最大速率100Kbps，用于带外管理单元监控Atlas 300I Pro 推理卡，Atlas 300I Pro 推理卡详细信号分配如**表5-1** ~ **表5-5**所示。

表 5-1 Atlas 300I Pro 推理卡金手指管脚定义（Mechanical key）

序号		管脚名	描述	处理方式
Side B	1	+12V	12V电源	12V电源
	2	+12V	12V电源	
	3	+12V	12V电源	
	4	GND	地	接地
	5	SMCLK	SMBus时钟	接入
	6	SMDAT	SMBus数据	接入
	7	GND	地	接地
	8	+3.3V	3.3V电源	3.3V电源
	9	JTAG1	JTAG接口TRST信号	悬空
	10	3.3Vaux	3.3V auxiliary电源	
	11	WAKE#	链接重新激活的信号	
Side A	1	PRSNT1#	热插拔存在检测1#	在位检测

序号		管脚名	描述	处理方式
	2	+12V	12V电源	12V电源
	3	+12V	12V电源	
	4	GND	地	接地
	5	JTAG2	JTAG接口TCK信号	悬空
	6	JTAG3	JTAG接口TDI信号	
	7	JTAG4	JTAG接口TDO信号	
	8	JTAG5	JTAG接口TMS信号	
	9	+3.3V	3.3V电源	3.3V电源
	10	+3.3V	3.3V电源	
	11	PERST#	基本复位	整板复位

表 5-2 Atlas 300I Pro 推理卡金手指管脚定义（End of the x1 connector）

序号		管脚名	描述	处理方式
Side B	12	RSVD	预留	悬空
	13	GND	地	接地
	14	PETp0	发送差分对 PCIe_TX_0	接PCIe卡
	15	PETn0		
	16	GND	地	接地
	17	PRSNT2#	热插拔存在检测2#	悬空
	18	GND	地	接地
Side A	12	GND	地	接地
	13	REFCLK+	差分时钟	底板提供100M PCIe差分时钟，支持3.0/2.0/1.0，支持SSC
	14	REFCLK-		
	15	GND	地	接地
	16	PERp0	接收差分对 PCIe_RX_0	接PCIe卡
	17	PERn0		
	18	GND	地	接地

表 5-3 Atlas 300I Pro 推理卡金手指管脚定义（End of the x4 connector）

序号		管脚名	描述	处理方式
Side B	19	PETp1	发送差分对 PCIe_TX_1	接PCIe卡
	20	PETn1		
	21	GND	地	接地
	22	GND	地	接地
	23	PETp2	发送差分对 PCIe_TX_2	接PCIe卡
	24	PETn2		
	25	GND	地	接地
	26	GND	地	接地
	27	PETp3	发送差分对 PCIe_TX_3	接PCIe卡
	28	PETn3		
	29	GND	地	接地
	30	RSVD	预留	悬空
	31	PRSNT2#	热插拔存在检测2#	悬空
	32	GND	地	接地
Side A	19	RSVD	悬空	
	20	GND	地	接地
	21	PERp1	接收差分对 PCIe_RX_1	接PCIe卡
	22	PERn1		
	23	GND	地	接地
	24	GND	地	接地
	25	PERp2	接收差分对 PCIe_RX_2	接PCIe卡
	26	PERn2		
	27	GND	地	接地
	28	GND	地	接地
	29	PERp3	接收差分对 PCIe_RX_3	接PCIe卡
	30	PERn3		
	31	GND	地	接地
	32	RSVD	预留	悬空

表 5-4 Atlas 300I Pro 推理卡金手指管脚定义（End of the x8 connector）

序号		管脚名	描述	处理方式
Side B	33	PETp4	发送差分对 PCIe_TX_4	接PCIe卡
	34	PETn4		
	35	GND	地	接地
	36	GND	地	接地
	37	PETp5	发送差分对 PCIe_TX_5	接PCIe卡
	38	PETn5		
	39	GND	地	接地
	40	GND	地	接地
	41	PETp6	发送差分对 PCIe_TX_6	接PCIe卡
	42	PETn6		
	43	GND	地	接地
	44	GND	地	接地
	45	PETp7	发送差分对 PCIe_TX_7	接PCIe卡
	46	PETn7		
	47	GND	地	接地
	48	PRSNT2#	热插拔存在检测2#	悬空
	49	GND	地	接地
Side A	33	RSVD	预留	悬空
	34	GND	地	接地
	35	PERp4	接收差分对 PCIe_RX_4	接PCIe卡
	36	PERn4		
	37	GND	地	接地
	38	GND	地	接地
	39	PERp5	接收差分对 PCIe_RX_5	接PCIe卡
	40	PERn5		
	41	GND	地	接地
	42	GND	地	接地
	43	PERp6	接收差分对 PCIe_RX_6	接PCIe卡
	44	PERn6		

序号		管脚名	描述	处理方式
	45	GND	地	接地
	46	GND	地	接地
	47	PERp7	接收差分对 PCIe_RX_7	接PCIe卡
	48	PERn7		
	49	GND	地	接地

表 5-5 Atlas 300I Pro 推理卡金手指管脚定义（End of the x16 connector）

序号		管脚名	描述	处理方式
Side B	50	PETp8	发送差分对 PCIe_TX_8	接PCIe卡
	51	PETn8		
	52	GND	地	接地
	53	GND	地	接地
	54	PETp9	发送差分对 PCIe_TX_9	接PCIe卡
	55	PETn9		
	56	GND	地	接地
	57	GND	地	接地
	58	PETp10	发送差分对 PCIe_TX_10	接PCIe卡
	59	PETn10		
	60	GND	地	接地
	61	GND	地	接地
	62	PETp11	发送差分对 PCIe_TX_11	接PCIe卡
	63	PETn11		
	64	GND	地	接地
	65	GND	地	接地
	66	PETp12	发送差分对 PCIe_TX_12	接PCIe卡
	67	PETn12		
	68	GND	地	接地
	69	GND	地	接地
	70	PETp13	发送差分对 PCIe_TX_13	接PCIe卡

序号		管脚名	描述	处理方式
	71	PETn13		
	72	GND	地	接地
	73	GND	地	接地
	74	PETp14	发送差分对 PCIe_TX_14	接PCIe卡
	75	PETn14		
	76	GND	地	接地
	77	GND	地	接地
	78	PETp15	发送差分对 PCIe_TX_15	接PCIe卡
	79	PETn15		
	80	GND	地	接地
	81	PRSNT2#	热插拔存在检测2#	PCIe卡上连接
	82	RSVD	预留	悬空
Side A	50	RSVD	预留	悬空
	51	GND	地	接地
	52	PERp8	接收差分对 PCIe_RX_8	接PCIe卡
	53	PERn8		
	54	GND	地	接地
	55	GND	地	接地
	56	PERp9	接收差分对 PCIe_RX_9	接PCIe卡
	57	PERn9		
	58	GND	地	接地
	59	GND	地	接地
	60	PERp10	接收差分对 PCIe_RX_10	接PCIe卡
	61	PERn10		
	62	GND	地	接地
	63	GND	地	接地
	64	PERp11	接收差分对 PCIe_RX_11	接PCIe卡
	65	PERn11		
	66	GND	地	接地
	67	GND	地	接地

序号		管脚名	描述	处理方式
	68	PERp12	接收差分对 PCIe_RX_12	接PCIe卡
	69	PERn12		
	70	GND	地	接地
	71	GND	地	接地
	72	PERp13	接收差分对 PCIe_RX_13	接PCIe卡
	73	PERn13		
	74	GND	地	接地
	75	GND	地	接地
	76	PERp14	接收差分对 PCIe_RX_14	接PCIe卡
	77	PERn14		
	78	GND	地	接地
	79	GND	地	接地
	80	PERp15	接收差分对 PCIe_RX_15	接PCIe卡
	81	PERn15		
	82	GND	地	接地

6 安装硬件

Atlas 300I Pro 推理卡的安装方法可参见各服务器用户指南。

7 安装驱动和固件

Atlas 300I Pro 推理卡驱动和固件的安装与维护请参见《**Atlas 300I Pro 推理卡 NPU 驱动和固件安装指南**》。

8 维护管理

Atlas 300I Pro 推理卡提供了丰富的维护管理功能，包括运行在OS中的带内管理命令集和通过SMBUS提供的带外管理功能。

说明

如果AI芯片没有加载驱动，则带外管理无法准确识别AI芯片是否真正发生故障，因此带外对AI芯片失效场景不做告警提示。带内管理只提供AI芯片健康状况的查询，如果上层业务需要对AI芯片失效场景做实时告警，需要上层业务调用DCMI API中相关接口，并做相关处理。

8.1 带内管理

带内管理的功能有：

- 在线升级功能，升级Firmware，方便用户的设备维护。
- 资产管理功能，提供序列号等信息，方便用户进行资产管理。
  具体资产管理操作请参见《Atlas 300I Pro 推理卡 npu-smi 命令参考》。

8.2 带外管理

Atlas 300I Pro 推理卡提供SMBUS接口，支持服务器的带外管理功能。

A 附录

A.1 术语

A

AI 研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。

B

BIOS 存于计算机主板上的一种固件。包括基本输入输出控制程序、上电自检程序、系统启动自举程序、系统设置信息，为计算机提供底层的硬件设置和控制功能。

C

CPU 中央处理器是计算机的主要设备之一，其功能是解释计算机指令以及处理计算机软件中的数据，与内部存储器、输入及输出设备成为现代电脑的三大部件。

D

DDK DDK是Mind解决方案提供的开发者套件包，Mind Studio通过安装DDK后获得Mind开发必需的API、库、工具链等开发组件。

DKMS 动态内核模块支持是用来生成Linux内核模块的一个框架，其源代码一般不在Linux内核源代码树。当新的内核安装时，DKMS支持的内核设备驱动程序会自动重建。

H

HDC 用于Host和Device之间通信模块，在Host和Device里面均有部署。

I

IDE 集成开发环境。

N

NPU 采用“数据驱动并行计算”的架构，特别擅长处理视频、图像类的海量多媒体业数据，专门用于处理人工智能应用中的大量计算任务。

P

PCIe	PCIe属于高速串行点对点双通道高带宽传输，所连接的设备分配独享通道带宽，不共享总线带宽，主要支持主动电源管理，错误报告，端对端的可靠性传输，热插拔以及服务质量(QOS)等功能。
R	
Runtime	Runtime运行于APP进程空间，为APP提供了设备的Memory管理、Device管理、Stream管理、Event管理、Kernel执行等功能。
T	
TE	用于开发自定义算子。
TEE	在ARM Trustzone的硬件隔离环境基础上，结合硬件可信根设计，实现安全启动、安全存储、安全升级、安全运行等功能，为系统提供可信的基础运行环境。

A.2 缩略语

A		
AI	Artificial Intelligence	人工智能
B		
BMC	Baseboard Management Controller	主板管理控制单元
C		
CFM	Cubic Feet Per Minute	立方英尺每分钟
E		
ECC	Error Checking and Correction	误差核对与改正
F		
FLOPS	Floating-point Operations Per Second	每秒浮点运算次数
FPS	Frames Per Second	每秒传输帧数
H		
HHHL	Half-Height Half-Length	半高半长
I		
I^2C	Inter-integrated Circuit	集成电路总线
J		
JPEG	Joint Photographic Experts Group	联合图像专家组
L		

LPDDR	Low-power Double Data Rate	低功耗双倍速
N		
NLP	Natural Language Processor	自然语言处理器
O		
OCR	Optical Character Recognition	光学字符识别
OS	Operating System	操作系统
P		
PCIe	Peripheral Component Interconnect Express	快捷外围部件互连标准
S		
SMBus	System Management Bus	系统管理总线
T		
TFLOPS	teraFLOPS	每秒万亿次的浮点运算

A.3 免责声明

- 本文档可能包含第三方信息、产品、服务、软件、组件、数据或内容（统称“第三方内容”）。华为不控制且不对第三方内容承担任何责任，包括但不限于准确性、兼容性、可靠性、可用性、合法性、适当性、性能、不侵权、更新状态等，除非本文档另有明确说明。在本文档中提及或引用任何第三方内容不代表华为对第三方内容的认可或保证。
- 用户若需要第三方许可，须通过合法途径获取第三方许可，除非本文档另有明确说明。

A.4 如何获取帮助

日常维护或故障处理过程中遇到难以解决或者重大问题时，请寻求华为技术有限公司的技术支持。

A.4.1 收集必要的故障信息

在进行故障处理前，需要收集必要的故障信息。

收集的信息主要包括：

- 客户的详细名称、地址
- 联系人姓名、电话号码
- 故障发生的具体时间

- 故障现象的详细描述
- 设备类型及软件版本
- 故障后已采取的措施和结果
- 问题的级别及希望解决的时间

A.4.2 做好必要的调试准备

在寻求华为技术支持时，华为技术支持工程师可能会协助您做一些操作，以进一步收集故障信息或者直接排除故障。

在寻求技术支持前请准备好单板和端口模块的备件、螺丝刀、螺丝、串口线、网线等可能使用到的物品。

A.4.3 如何使用文档

华为技术有限公司提供全面的随设备发货的指导文档。指导文档能解决您在日常维护或故障处理过程中遇到的常见问题。

为了更好的解决故障，在寻求华为技术支持前，建议充分使用指导文档。

A.4.4 获取技术支持

华为技术有限公司通过办事处、公司二级技术支持体系、电话技术指导、远程支持及现场技术支持等方式向用户提供及时有效的技术支持。

技术支持网址

查阅技术资料合集：**https://e.huawei.com/cn/** > 技术支持 > 产品和解决方案支持 > 服务器-智能计算 > 昇腾计算

查阅技术资料的使用流程：**https://www.hiascend.com** > 文档

自助平台与论坛

如果您想进一步学习和交流：

- 访问**华为服务器信息服务平台**，获取相关服务器产品资料。
- 访问**华为企业业务智能问答系统**，快速查询产品问题。
- 访问**华为企业互动社区（服务器）**，进行硬件产品学习交流。
- 访问**开发者论坛**，进行AI应用开发学习交流。

公告

有关产品生命周期、预警和整改公告请访问**技术支持 > 公告 > 产品公告**。

案例库

参阅已有案例进行学习：**计算产品案例查询助手**。

说明

计算产品案例查询助手目前仅面向华为合作伙伴及华为工程师开放。

获取华为技术支持

如果在设备维护或故障处理过程中，遇到难以确定或难以解决的问题，通过文档的指导仍然不能解决，请通过如下方式获取技术支持：

- 联系华为技术有限公司客户服务中心。

  中国区企业用户请通过以下方式联系我们：

  - 客户服务电话：400-822-9999
  - 客户服务邮箱：**support_e@huawei.com**

    企业网全球各地区客户服务热线可以通过以下网站查找：**企业用户全球服务热线**

  中国区运营商用户请通过以下方式联系我们：

  - 客户服务电话：400-830-2118
  - 客户服务邮箱：**support@huawei.com**

    运营商全球各地区客户服务热线可以通过以下网站查找：**运营商用户全球服务热线**
- 联系华为技术有限公司驻当地办事处的技术支持人员。