

全爱 QA200-RC/PE Atlas 200 开发者套件

QA200RC 推理卡切换至 EP 模式操作手册

文档版本 01

发布日期 2021-12-3



版权所有 全爱科技（上海）有限公司 2021。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



和其他全爱商标均为全爱科技（上海）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、服务或特性等应受全爱科技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，全爱公司对本文档内容不做任何明示或默示的声明或保证。

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。

全爱科技（上海）有限公司

电话：021-64025956

地址：上海市闵行区剑川路 930 号 D 栋 3 层 邮编：200240

网址：www.quanaichina.com

概述

本文档简单介绍全爱 QA200-RC Atlas 200 开发者套件（型号 3000）切换至 QA200-EP 的操作方法。

QA200-RC/EP 可提供 RC、EP 两种应用模式， QA200-RC 默认为 RC 模式，SD 卡存储。

出厂时可根据需求选择方式；出厂后如需要更换应用模式需要进行工程变动，需要专业硬件工程师进行操作。








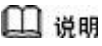
读者对象

本文档主要适用于以下人员：




- 全爱售前工程师
- 渠道伙伴售前工程师
- 企业售前工程师

注意事项

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
	 本手册所有操作事项，均应在防静电保护环境下，由硬件工程师进行操作。
	 操作前，请切断电源并确认板卡电感、电容放电完成。
符号	说明
	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。
	表示如不避免则可能导致死亡或严重伤害的具有中等等级风险的危害。
	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。
	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。

全爱 QA200-RC Atlas 200 开发者套件 QA200RC 推理卡切换至 EP 模式操作手册：

符号	说明
	本手册所有操作事项，均应在防静电保护环境下，由硬件工程师进行操作。
	 危险 操作前，请切断电源并确认板卡电感、电容放电完成。

需要完成 5 个步骤：

1. SW1 如下图拨至左侧；



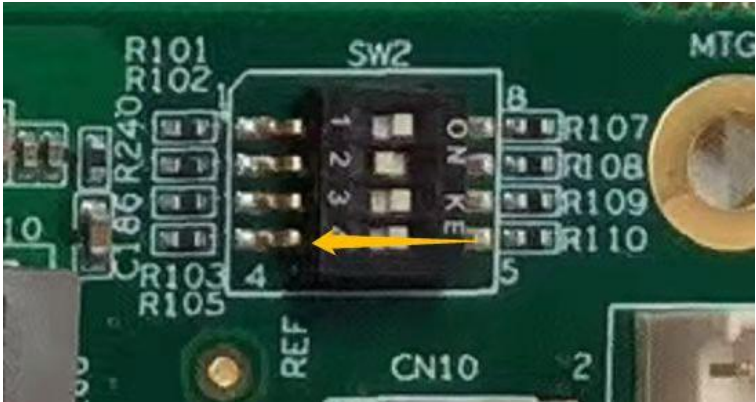
RC：右侧，

EP：左侧

2. SW2 按照下图 ON 是 0;OFF 是 1

SW2 的 1/2/3，三位，往左拨是 1，往右拨是 0

1/2/3 分别对应 BOOT_MD 的[0, 1, 2]，含义如下：



RC：SW2：0101

EP: SW2: 0001

BOOT_MD[2:0]
000 on-chip-rom + SFC(xloader) + PCIe(UEFI, EP mode only)。
001 on-chip-rom + USB加载。
010 on-chip-rom + SFC(SPI Nor)加载 (xloader+UEFI), 缺省。

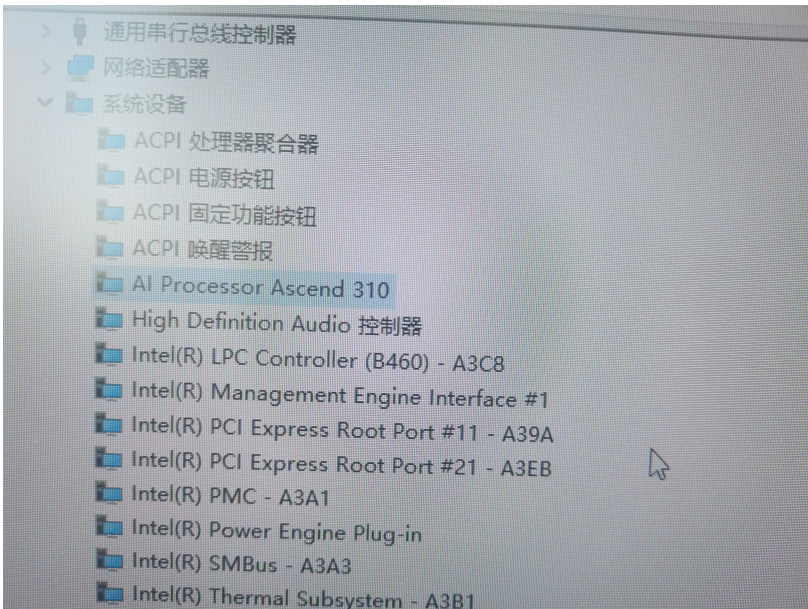
3. 切换 SD 卡或者 EMMC 才需要进行如下参数 设置:

SW3 描述:

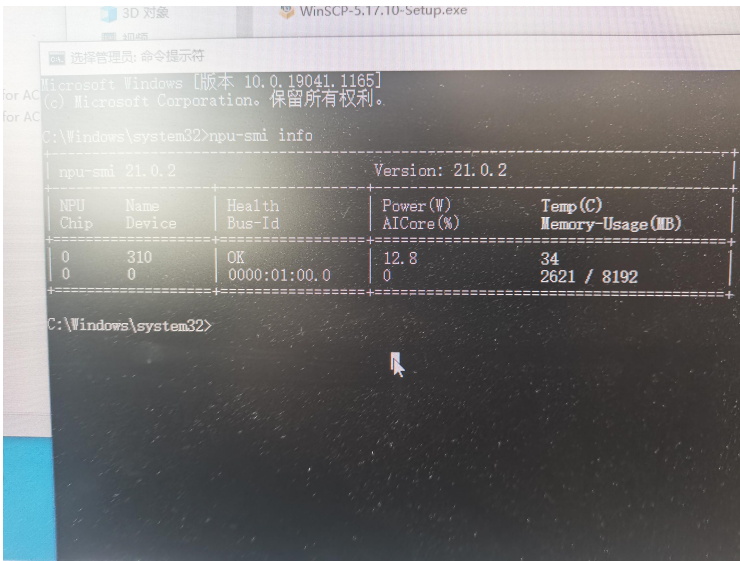
RC 模式下 SW3 的 1 为 1,

EP 模式下 SW3 的 1 为 0,

4. 电脑端固件安装后如下图:



5. 电脑端驱动安装后 ping: npu-smi info
显示如下图，即为正确按照，可以进行 AI 开发。



软件部署请参阅

Atlas200 加速模块 EP 模式配置手册

说明：EP 模式（即 Atlas200 加速模块作为 PCIe 使能从设备模式），相关环境部署配置 host（开发环境）侧即可。大家有任何建议或意见请进行评论，相关信息会及时进行修正。

一、物理环境

客户需确认物理设备无误，即：host 侧设备和 Atlas200 加上模块正常连接可用。

二、运行环境（device）说明

运行环境即为 Atlas200 加速模块，在 host 安装相关固件和驱动后，可正常调用。

三、开发环境（host）部署

a.硬件等环境准备参考：https://support.huaweicloud.com/instg-cli-cann502-alpha001/atlasdeploy_03_0009.html

b.配置开发环境参考：https://support.huaweicloud.com/instg-cli-cann502-alpha001/atlasdeploy_03_0089.html

c.按照流程依次安装固件和驱动：<https://support.huawei.com/enterprise/zh/doc/EDOC1100149865>

d.按照流程依次安装开发套件包等环境：https://support.huaweicloud.com/instg-cli-cann502-alpha001/atlasdeploy_03_0097.html（通过 `atc -h` 可验证是否安装成功，无报错返回即可）

e.样例运行：相关环境安装完毕后，下载权重和模型文件进行测试，按照以下连接进行调试：https://support.huaweicloud.com/aclcppdevg-cann502alpha2infer/atlassample_14_0009.html（此目录下有多个样例：https://support.huaweicloud.com/aclcppdevg-cann502alpha2infer/atlasdevelopment_01_0006.html）

四、常见问题说明

a.若确认各项环境按照无误，在运行样例报错时，需排查当前运行账号是否正确，相关环境变量是否配置完毕。

b.在验证按照流程进行作业，还是报错时，可到 Atlas200 加速模块进行发帖求助：
<https://bbs.huaweicloud.com/forum/forumdisplay-fid-1337-orderby-lastpost-filter-typeid-typeid-1835.html>