

北京电子控股有限责任公司关于光电混合计算加速服务器研制“揭榜挂帅”项目申报指南

一、需求目标

面向支持高性能、高算力的国产化算力中心需求，研制光电混合计算加速服务器，实现对数字经济、元宇宙、智能制造、智慧城市等相关领域提供自主可控的算力支撑。通过项目实施，在 AI 及高性能计算加速方面，实现自主可控高性能的光电混合计算加速服务器规模化推广应用，在保障供应链安全的同时，降低产品成本。

二、产品指标

本项目产品为光电混合计算加速服务器（以下简称“服务器”），服务器采用基于国产 GPU 芯片与光计算模块组成的光电混合计算加速卡（以下简称“加速卡”）实现计算加速，单台服务器包含不少于 8 块加速卡。

（一）服务器功能指标

1. 支持 TensorFlow、Pytorch、paddlepaddle、ONNX 等主流框架的网络模型。
2. 支持 1280 路以上 1080P@30FPS 视频解码。
3. 支持 96 路以上 1080P@30FPS 视频编码。

4. 支持 H.264、H.265 等格式硬件编解码，支持 8K 分辨率。
5. 支持主流 AI 任务与 HPC 任务，可实现对金融分析、元宇宙、智能制造、智慧城市、智慧政务等方面进行硬件加速。

（二）加速卡性能指标

1. 加速卡基于国产 GPU 芯片与光计算模块组成，光计算模块采用可插拔形式，支持后续不断迭代升级。

2. GPU 部分算力不低于 30 TFLOPS@FP32；120 TFLOPS@TF32；240 TFLOPS@FP16；240 TFLOPS@BF16；480 TOPS@INT8。

3. 对特定算法，加速卡等效算力性能不低于 512TOPS；光计算芯片算力不低于 100Tops。

4. 光计算模块通过金手指接口与板卡连接，双向带宽不小于 800Gbps。

5. 单卡内存不低于 64GB

6. 具备光电融合的统一工具链，能够充分发挥光电混合硬件的性能优势。

7. 支持多卡板间互连及 GPU 多实例化。

8. 支持 Ubuntu 等国际主流操作系统，以及 Kylin 等主流国产操作系统。

9. 支持与主流 AI 框架的兼容适配，认证不低于 2 级。

10. 兼容 CUDA。

（三）成本指标

服务器价格不高于 100 万元/台套。

三、交付物与考核指标

1. 提交满足考核指标的 1 台套服务器，并提供相应的软件工具链。

2. 完成主流 AI 框架的兼容适配认证，提供不低于 2 级的兼容适配认证证据。

3. 完成 GPU 芯片算力性能测试，提供具有检测资质的第三方测试报告。

四、项目周期：1 年。

五、榜单金额：不超过 200 万元。