

苏州国芯科技股份有限公司

关于公司研发的神经网络处理器 DPNPU 新 IP 产品内部测试成功的自愿性披露公告

本公司董事会及全体董事保证本公告内容不存在任何虚假记载、误导性陈述或者重大遗漏，并对其内容的真实性、准确性和完整性依法承担法律责任。

苏州国芯科技股份有限公司（以下简称“公司”或“国芯科技”）研发的神经网络处理器 DPNPU（Dataflow Parallel NPU）新 IP 产品于近日在公司内部测试中获得成功。现将相关事项公告如下：

一、新产品的基本情况

公司最新研制的神经网络处理器 DPNPU 新 IP 产品面向端侧与边缘计算的高性能 AI 处理器，它针对人工智能应用中复杂多变的计算任务进行了多方面优化，致力于在功耗、性能和灵活性之间取得最佳平衡，为各类智能设备提供高效、可靠且易于扩展的算力支持。

国芯科技 DPNPU 单核支持 0.5~4.8 TOPS 的灵活算力配置，支持算力线性扩展，可以为不同场景提供定制化的 AI 算力解决方案。国芯科技 DPNPU 采用符合 RISC-V 指令集架构（RISC-V ISA）标准的创新开放架构，该架构通过将 RISC-V 核心与高性能神经网络加速单元在架构层面进行深度优化设计，并设计了专用的 TDS（Task Distribution&Synchronization）硬件调度引擎作为核心控制单元，将网络模型中的算子序列转化为高效的节点化任务流，实现了任务管理、数据流控制与 AI 专用计算的统一高效调度。该新 IP 产品采用脉动阵列高效动态融合技术，确保设备在长时间运行中的稳定性和卓越能效表现。DPNPU 内置 90+ 神经网络算子，全面覆盖 CNN、RNN 神经网络架构，并支持 LSTM、GRU 等 RNN 变体。通过 RISC-V 指令和硬件通用性设计可扩展支持更多算子，为适配未来不断涌现的 AI 模型预留充足空间。DPNPU 支持训练后量化（PTQ）技术，提供对称量化、非对称量化、逐层量化和逐通道量化四种方式。同时支持 INT8 和 FP16 混合精度量化，在保持模型精度的前提下，该新 IP 产品大幅减少计算资源和存储空间占用，实现精度与性能的完美平衡。该 DPNPU 集成了自研 PCD

(Parameter Compression&Decompression) 模块，支持参数压缩与硬件自动解压缩技术，有效缓解数据吞吐瓶颈，降低带宽压力，进一步提升整体系统性能。

为了降低 AI 应用开发门槛，国芯科技围绕自研的 DPNPU 构建了完整的软件生态——C*Core NPU Studio，集成完整易用的工具套件、驱动和运行时软件，缩短从开发到部署的周期，提升 AI 应用落地效率。C*Core NPU Studio 工具套件为 DPNPU 提供端到端的模型部署能力，包括模型转换、预处理、量化、编译、仿真等工具软件。C*Core NPU Runtime 提供 DPNPU 运行时推理支撑，涵盖了推理框架软件和各种扩展软算子库；C*Core NPU Driver 适配 RISC-V 等主流 CPU 平台，支持 Linux/RTOS/Bare-metal 等不同应用环境需求。

人工智能技术正加速渗透至工业控制、智能家电、汽车电子、医疗健康等关键领域。与云端 AI 相比，端侧与边缘侧 AI 具备实时响应、数据隐私保护、低网络依赖等显著优势，这对芯片的能效比和算力密度提出了更高要求。国芯科技 DPNPU 的架构可行性、能效表现及软件栈均已完成验证，为后续公司 NPU 技术持续研发以及端侧和边缘侧 AI 芯片发展和应用奠定了基础。

二、对公司的影响

公司对上述 DPNPU 新 IP 产品拥有知识产权。本次神经网络处理器 DPNPU 新 IP 产品研发成功，是公司在人工智能领域长期坚持创新驱动的结果，对公司未来市场拓展和业绩成长性预计都将产生积极的影响。

三、风险提示

本次测试目前是公司内部测试成功，本次公司推出的神经网络处理器 DPNPU 新 IP 产品尚处于市场导入初期，尚未实现客户应用和销售，存在市场推广与客户开拓不及预期、客户验证失败等风险，将对公司收入及盈利带来不确定性，公司将及时根据后续进展履行信息披露义务，敬请广大投资者注意投资风险。

特此公告。

苏州国芯科技股份有限公司

董事会

2026 年 1 月 5 日