



微路智能 VELOXIS AI

加入微路智能

杭州 · NPU / AI 芯片团队 · 正式岗位 + 实习岗位

从 NPU 架构、RTL、验证、编译和 Runtime, 到
SoC、性能建模、后端实现与模型部署。

欢迎正式岗位与实习岗位候选人, 一起把 AI 芯片工程链路真正做通。

NPU 架构与算法

RTL / SoC / 验证

Compiler / Runtime



VELOXIS AI

招聘邮箱: careers@veloxis-tech.com | 杭州微路智能科技有限公司 | 2026年9月

公司概况

以数据流驱动端侧 AI 推理芯片、IP 与系统产品

我们在做什么

微路智能成立于2026年7月，已完成数千万元融资。

公司面向 AI 推理提供计算 IP、互连与设计服务，并开发 VLX 加速产品；我们关注模型能否高效运行、系统能否稳定集成，以及客户设备能否以合理成本获得本地大模型能力。

2026.07

杭州主体成立

数千万元

融资已完成

Qwen-8B

FPGA 端到端验证

IP → 芯片

推进首颗芯片工程化

你将加入的阶段

01 IP 交付与客户集成

技术模块已进入真实客户项目，工程反馈持续回流。

02 首颗芯片工程化

整芯片集成、签核、MPW 与回片验证正在推进。

03 产品化与规模交付

复用 IP、软件和验证资产，走向加速卡与模组产品。

团队覆盖体系结构、低精度计算、片上互连、编译调度与物理实现；研发工作以客户项目和芯片工程节点推进。

端侧大模型推理是一项跨层工程

算力、内存、编译、运行环境和设备集成需要同时解决

0
1

计算不是只看矩阵

解码阶段受权重容量与内存带宽约束；归一化、激活、归约、Softmax
和专家路由还需要向量与特殊函数通路。

0
2

软件不能成为部署负担

从模型格式、量化、算子映射到运行时，任何一个环节不通，客户项目都会停在评估阶段。

0
3

产品必须进入现有设备

工控机、终端和机器人已有稳定的CPU、操作系统和业务软件，不能为增加推理能力重做整套平台。

所以，我们需要的是一支跨层协作的团队

算法与模型 · 编译与算子 · 芯片架构与 RTL · 验证与后端 · 驱动、板卡与系统验证

核心团队与带教团队

从体系结构、芯片工程到编译系统，团队覆盖完整实现链路

CEO / 联合创始人

战略、产品与系统架构

EPFL 博士，研究方向涵盖计算机体系结构与异构系统协同优化；参与 CPU 微架构、AI 加速器系统建模及大模型推理能效研究。

CTO / 联合创始人

芯片架构、编译与系统软件

浙江大学博士生导师。UCSD 博士，专注 AI 芯片架构、近存计算、混合精度计算与片上互连；负责 FusionCore 等核心模块的技术路线与研发。

架构师 / 联合创始人

芯片工程与验证

清华大学集成电路工程硕士，具有 RTL、综合、物理实现、PPA 分析和多工艺节点流片经验，负责芯片架构与工程验证。

首席科学家 / 科学顾问

长期架构、低功耗与量化评审

密歇根大学博士，浙江大学教授，研究方向包括设计自动化、低功耗芯片及设计制造协同，参与长期技术路线与架构评审。

工程与运营骨干覆盖编译器、物理实现、签核与流片、项目管理及客户交付，为候选人的工程成长提供直接带教。

招聘方向与岗位地图

工作地点杭州；正式岗位与实习岗位同步招聘

01

架构 / 算法

NPU 架构及微架构
AI / HW-SW Co-design
性能建模与架构仿真

02

RTL / SoC / Memory

RTL / NPU 设计
NoC / DMA / AXI / CHI
SRAM / LPDDR / HBM

03

验证 / 性能

NPU DV / UVM
Coverage / Assertion
Cycle-level Simulator

04

Compiler / Runtime

Graph / IR / Tiling / CodeGen
Runtime / Driver / Firmware
AI Runtime / 模型部署

05

后端 / DFT / 原型

Physical Design / STA / P&R
DFT / ATPG / MBIST / JTAG
FPGA / Emulation

06

HR / 研发实习

HR / 技术招聘
芯片设计 / 验证实习
Compiler / Runtime / 模型部署实习

正式与实习岗位均以杭州为主；实习优先线下，可接受部分时间远程协作，具体职责和级别结合经历沟通。

NPU 架构、RTL 与 SoC 设计

从工作负载和微架构，到 RTL、SoC、NoC 与 Memory System

你会参与

- 设计 NPU 计算阵列、数据流、片上存储、NoC 与 DMA 等关键架构。
- 分析 Transformer / CNN workload，完成算子映射、性能建模与瓶颈分析。
- 实现计算单元、数据通路、SRAM / Buffer、DMA、NoC 与接口 RTL。
- 完成 Lint、CDC、综合、PPA 优化，并推进 SoC 子系统集成。
- 优化 SRAM、NoC、DRAM / LPDDR / HBM 的带宽、延迟与 QoS。

适配背景

体系结构

熟悉

GEMM、Attention、Convolution、Dataflow、Tiling、Memory Hierarchy 等。

RTL 技能

Verilog / SystemVerilog，熟悉

Pipeline、FSM、FIFO、Memory Interface。

工具与协议

VCS / Verdi / Questa / DC / Genus; AXI / ACE / CHI 至少一种。

加分经历

NPU / GPU / TPU / 加速器、Systolic Array、SoC、流片或时序收敛经验。

同时招聘 RTL、SoC/NoC/Memory、性能建模和 FPGA/Emulation 方向；实习岗位欢迎有课程、科研或开源项目经验的同学。

Compiler、Runtime 与 AI 模型部署

覆盖 Graph/IR、代码生成、驱动运行时和端到端模型执行链路

你会参与的技术链路

PyTorch / ONNX 模型



Graph / IR 优化



Compiler: Tiling / Scheduling / CodeGen



Runtime / Driver / Firmware



NPU 执行与模型部署

模型性能需要在
Compiler、Runtime、Driver
和硬件资源之间联合优化。

工作内容

NPU Compiler Graph / IR、算子融合、Tiling、Scheduling、Memory Planning 与 CodeGen。

编译基础 熟悉 LLVM / MLIR / TVM 至少一种；理解 Layout、Loop Optimization。

Runtime / Driver 调度、Command Queue、Memory、DMA、Interrupt 与 Linux 系统开发。

AI Runtime PyTorch / ONNX 到 NPU、算子库、图执行、异步调度与 Profiling。

加分背景 NPU / GPU Compiler、CUDA / TensorRT / ONNX Runtime、模型推理优化。

正式与实习岗位均招聘 Compiler、Runtime/Driver/Firmware 和 AI Runtime/模型部署方向。

验证、后端实现与原型验证

覆盖 NPU DV、Physical Design、DFT 和 FPGA/Emulation

0
1

Design Verification

UVM、Test
Case、Checker、Coverage、Assertion
与波形调试；支持性能和系统级验证。

熟悉 VCS / Verdi / Xcelium / Questa，掌握
Python / C / C++ 至少一种。

0
2

Physical Design / DFT

综合、STA、P&R、PPA
收敛、Scan / ATPG / MBIST / JTAG、物理实现与
签核。

熟悉
DC / Genus / Innovus / PrimeTime / Tempu
s，具备 Tcl / Python / Shell 自动化能力。

0
3

FPGA / Emulation

RTL
集成、综合、部署、板级调试及软硬件联合验证和
性能测试。

熟悉
Vivado / Quartus、C/C++/Python；PCIe、
DDR/HBM、SerDes 经验优先。

正式岗位和实习岗位均开放：验证、Physical Design、DFT 与 FPGA/Emulation。

岗位类型与候选人期待

正式岗位与实习岗位同步招聘，实习优先线下并支持部分远程协作

你能获得的经历

正式岗位 架构、算法、RTL、SoC、DV、Compiler、Runtime、PD、DFT、FPGA 与 AI Runtime。

实习岗位 芯片设计、验证、架构算法、编译器、Runtime、SoC、性能、PD、DFT、FPGA。

工作地点 杭州；实习优先线下，可接受部分时间远程协作，稳定实习 3 个月以上优先。

申请对象 在校本科生、硕士生、博士生均可申请；正式岗位按经历与团队阶段沟通。

材料建议 课程、科研、开源、竞赛、FPGA、RTL、Compiler 项目均可作为能力证明。

我们看重

- 01 基础扎实，能阅读代码并独立完成小模块开发。
- 02 愿意主动定位问题、Debug，并推进工程闭环。
- 03 对 NPU、AI 芯片、体系结构、编译器或系统软件有真实兴趣。
- 04 能够跨专业协作，不把问题留在接口处。
- 05 欢迎附 GitHub、项目材料或能够说明个人贡献的作品。

正式岗位和实习岗位均可投递；实习岗位不要求完整芯片开发经验，更看重基础、动手能力和学习速度。



微路智能 VELOXIS AI

加入我们

杭州 | 正式岗位 + 实习岗位 | NPU / AI 芯片团队

招聘流程

01

简历与项目材料

02

技术沟通

03

团队互选

04

Offer 与入职安排

简历投递

简历投递

careers@veloxis-tech.com

邮件标题: 姓名 + 应聘方向 + 正式/实习

NPU 架构 / RTL / DV / Compiler / Runtime
SoC / NoC / Memory / 性能建模 / PD / DFT / FPGA
AI Runtime / 模型部署 / HR / 研发实习



正式与实习岗位均可投递; 工作地点杭州, 具体职责、级别与实习安排以沟通为准。