



卓越性能 灵活扩展

KunLun G5580 AI 推理服务器





KunLun G5580

KunLun G5580是一款4U2路AI推理服务器，搭载NPU/GPU加速卡和鲲鹏920处理器，具有计算性能卓越、配置灵活均衡、高效部署运维等特点，适合为加速计算、科学计算、视频分析和数据库等应用加速，针对百亿级/千亿级大模型推理、训练等业务类型进行了优化，支持多种类型的异构处理器，支持企业私有化部署和数据中心部署。



领先架构，卓越性能

- 支持8张全高全单宽NPU或双宽GPU加速卡，2个鲲鹏920处理器
- 支持12个3.5英寸SAS/SATA盘，其中可支持4个NVMe盘



灵活扩展，稳定可靠

- 11个PCIe槽位和2个Flex IO扩展槽位，其中8个NPU/GPU加速卡专用槽位
- 最多支持4个3000W高效钛金级电源，支持2+2/3+1冗余



智能管理，开放集成

- 版本自动推送，无接触升级，升级效率提升20倍
- 75%工具化，流水线式部署，部署效率提升10倍
- 资产位置自动识别，实时轨迹追踪，盘点准确率100%

形态	4U AI训练推理服务器
加速卡	支持8张全高全长单宽 NPU 或双宽 GPU 加速卡
CPU	支持2颗鲲鹏 920 系列处理器 (48/64 核) , CPU 主频 2.6GHz
DDR4内存	最大支持32条 DDR4 RDIMM 内存, 最大容量可支持 4TB
本地存储	支持多种不同的硬盘配置, 硬盘支持热插拔 <ul style="list-style-type: none"> •支持12个3.5英寸 SAS/SATA 硬盘, 其中支持4个 NVMe 盘 •支持2个 M.2 SSD
RAID支持	支持RAID0、1、10、5、50、6、60等, 支持 Cache 超级电容保护, 提供 RAID 缓存, 支持超级电容掉电数据保护。
PCIe拓展	最多11个PCIe 槽位, 其中10个 AI 加速卡专用槽位
IO	多种网络扩展能力, 支持2张半高半长网卡和2个 Flex IO 扩展槽位, 按需选配
风扇	提供6/8个热插拔风扇模块, 支持 N+1 冗余, 提升系统整体可用性。
管理	BMC芯片集成1个专用管理GE网口, 提供全面的故障诊断、自动化运维、硬件安全加固等管理特性 <ul style="list-style-type: none"> •BMC支持Redfish、SNMP、IPMI2.0等标准接口; 提供基于HTML5/VNC KVM的远程管理界面; 支持监控、诊断、配置、Agentless及远程控制等带外管理功能, 简化管理复杂度 •可选配管理软件, 提供无状态计算、OS批量部署、固件自动升级等高级管理特性, 实现全生命周期智能化、自动化管理
电源	支持4个2000W、2600W、3000W AC 白金/钛金电源模块, 支持 2+2/3+1 冗余, 支持热插拔
操作系统	BC - Linux、Ubuntu、CentOS、openEuler、统信、银河麒麟等操作系统
工作温度	5°C ~ 35°C (41°F~95°F) (符合 ASHRAE CLASS A1/A2)
尺寸(高x宽x深)	175mm × 447mm × 898mm

河南昆仑技术有限公司

技术咨询电话: 400-080-9000 **技术支持邮箱:** support@kunlunit.com


地址: 河南省郑州市郑东新区龙子湖智慧岛中道东路时埂街北创智天地大厦10层

网址: www.kunlunit.com

版权所有 © 河南昆仑技术有限公司 2025。保留一切权利。

非经河南昆仑技术有限公司书面同意，任何单位和个人不得擅自摘抄、复制本手册内容的部分或全部，并不得以任何形式传播。

商标声明

 KunLunKunLun是河南昆仑技术有限公司的商标或者注册商标。在本手册中以及本手册描述的产品中，出现的其他商标、产品名称、服务名称以及公司名称，由其各自的所有人拥有。

免责声明

您购买的产品、服务或特性等应受河南昆仑技术有限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，河南昆仑技术有限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。