

HG9680服务器

以智能计算驱动产业数智化变革



HG9680服务器采用8模组高效推理架构，具备强劲的算力、内存带宽和超大互联能力。支持2.2 PFLOPS FP16算力，能够高效处理生成式大模型推理任务，如智能客服、文案生成和知识沉淀等内容生成应用。其高速互联带宽确保模组间数据传输低延迟，提升整体推理效率和响应速度。凭借卓越的计算性能和通信能力，HG9680为企业和科研机构提供坚实的平台，推动AI在各行业的深度应用，特别适合需要大规模推理和高吞吐量的智能场景，助力实现高效智能化转型升级。

更高算力密度

4U机箱设计内集成8颗昇腾AI模组，提供高达2.2 PFLOPS FP16算力，有效提升计算能力，满足多样化AI推理需求

强劲算力支持

采用高性能昇腾AI处理器，提升应用加速效率，助力生成式模型和其他AI任务实现流畅快速的推理处理

高效能耗表现

系统能效比达到0.4 PFLOPS/kW水平，帮助降低能耗和运营成本，实现经济且持续的计算资源利用

高速互联能力

配备8个200GE RoCE v2接口，结合HCCS Full Mesh互联技术，保障模组间高速数据传输和稳定通信体验

产品规格

名称	规格参数
NPU	8 × 昇腾 910
CPU	4 × 鲲鹏 920
AI 算力	整机提供最大 2.2 PFLOPS@FP16
片上内存	整机支持 256 GB、512 GB 两种规格
内存	32 个 DDR4 内存插槽，最高 3200 MT/s 单根内存条容量支持 16 / 32 / 64 GB
本地存储	8 × 2.5 SATA+2 × 2.5 NVMe 4 × 2.5 SATA+6 × 2.5 NVMe
RAID 支持	支持 RAID 0 / 1 / 10 / 5 / 50 / 6 / 60
网络	8 × 200GE QSFP 接口直出，RoCE 协议
PCIe 接口	最多支持 3 个 PCIe 4.0 扩展插槽
电源	4 个热插拔 2.6 kW 电源模块，支持 2+2 冗余
散热方式	风冷
风扇	支持 8 个热插拔风扇模组，支持 N+1 冗余
工作环境温度	5°C~35°C (41°F~95°F)
结构尺寸	175 mm (高) × 447 mm (宽) × 790 mm (深)



👤 | 了解更多信息，欢迎咨询
🌐 | www.linkupai.cn
☎ | 400-869-9865