

宝德AI超节点服务器PR420KI G3



宝德PR420KI G3服务器是基于华为自研鲲鹏920处理器、昇腾910 AI处理器、总线交换芯片和CDR芯片的AI计算节点，面向互联网、运营商、金融等行业的大模型推理等重要场景，提供高性能、高可靠、易部署的AI推理服务器，满足通用风冷机房部署，以算力/互联带宽/可靠性极致竞争力为目标，结合工程优势持续构筑AI风冷推理服务器硬件竞争力。

超高性能

- 计算xPU采用昇腾HiAM模组
- AI系统采用多种处理器，整机支持4*鲲鹏920处理器+8*昇腾910 AI模组

超高密度

- 整机为10U机框，支持标准19英寸机柜安装。整机结构设计采用正交盲插架构，集成 CPU抽屉、NPU抽屉、灵衢总线板、IO框等关键部件，集成密度高、易维护

运维管理智能化

- 支持CPU和内存智能故障诊断和隔离，提升系统可靠性。算网融合，柜内支持通用交换机，简化机房部署

应用场景



互联网



金融



人工智能计算中心



微调



强化学习



预训练

产品规格

形态	● 10U风冷AI训推服务器
CPU处理器	● 支持4个鲲鹏920处理器，处理器支持≥64核 ● 每个core集成64KB L1 Cache，≥512KB L2 Cache，≥64MB L3 Cache。
内存	● 提供64个DDR5内存插槽，每个CPU基础板均提供32个内存插槽 ● 支持RDIMM ● 最大内存传输速率为5200MT/s ● 单根内存条容量支持64GB
本地存储	● 前置2*2.5英寸SATA+8*2.5英寸NVMe硬盘位
PCIe接口	● 最多支持5个PCIe5.0标准接口插槽，2个PCIe x16和3个PCIe5.0 x8插槽
HiAM AI模组	● 支持8个HiAM AI模组，单AI模组支持7路灵衢互联1.0+1路RoCE，用于AI处理器之间互连 ● 单NPU AI算力 560 TFlops (FP16)，总AI算力 4.48 PFlops，单NPU显存128G HBM
L1交换网络	● 7个交换芯片用于连接灵衢互联1.0网络，1:1无收敛 ● HiAM模组直出CDR芯片用于RoCE网络，1:1无收敛
能耗	● 最大输入功耗：14.6kW
工作环境温度	● 工作温度：5°C~35°C(41°F ~ 95°F) (符合ASHRAE Class A2) ● 工作海拔高度：≤3050m
结构尺寸	● 442mm(H)×447mm(W)×920mm(D)

* 具体配置请以实际配置为准