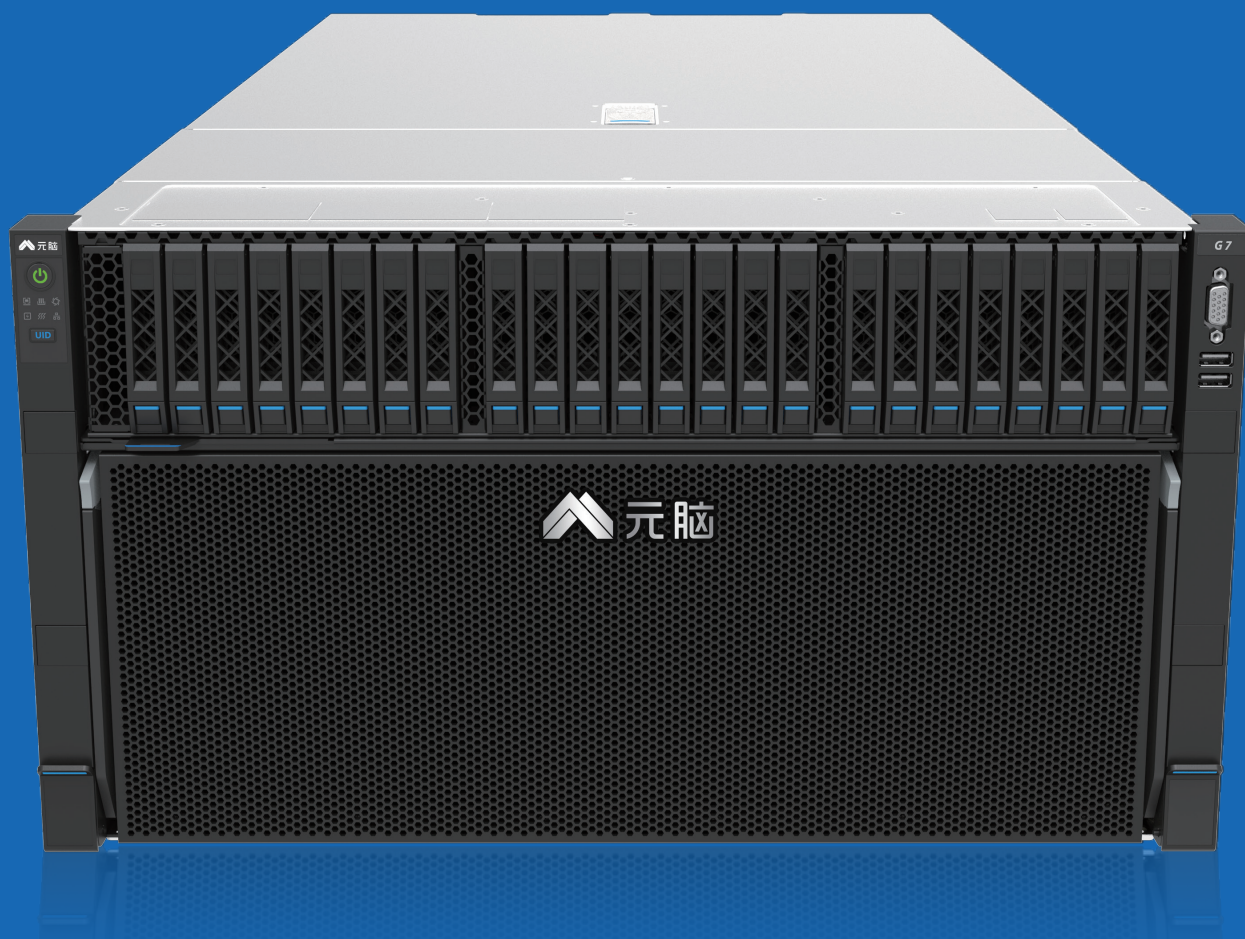


NF5688G7

强劲高扩展AI服务器



产品简介

NF5688G7是基于浪潮信息全新一代AI超融合架构平台，面向超大规模数据中心的强劲性能、极致扩展人工智能服务器，最强算力密度6U空间内搭载1块NVIDIA Hopper架构HGX-8GPU模组，系统支持4.0Tbps网络带宽，满足万亿级参数超大模型并行训练需求。搭载2颗Intel®第四代至强®可扩展处理器(TDP 350W)或2颗AMD第四代EPYC处理器(Max cTDP 400W)，最高可达4TB系统内存，128TB NVMe高速存储，支持高达12个PCIe Gen5 x16扩展槽位，可灵活支持OCP 3.0、CX7、多种智能网卡等，构建面向超大模型训练、元宇宙、自然语言理解、推荐、AIGC等场景的最强AI算力平台。

产品特性

强劲性能	极致能效
8颗NVIDIA最新Hopper架构GPU 2颗Intel®第四代至强®可扩展处理器或2颗AMD第四代EPYCTM 处理器 - 集成Transformer引擎，大幅加速GPT大模型训练速度	- 散热性能极致优化，风道解耦设计提升20%系统能效比 12V和54V N+N冗余电源分离供电设计，减少电源转换损耗 - 智能分层调节散热，降低功耗与噪音
领先架构	多元兼容
- 节点内全PCIe 5.0高速链路，CPU至GPU带宽提升4倍 - 节点间高速互联扩展，4.0Tbps无阻塞带宽IB/RoCE组网 - 集群级优化架构设计，GPU: 计算IB: 存储IB=8:8:2	- 全模块化设计，一机多芯兼容，灵活配置支持本地和云端部署 - 支持大规模GPT-3/LLaMA/ChatGLM模型训练和推理 - 领先支持多样化的SuperPod解决方案，适用AIGC, AI4Science及元宇宙等丰富场景

产品规格

型号	NF5688-M7-A0-R0-00		NF5688-A7-A0-R0-00
高度	6U		
GPU	1块NVIDIA HGX-Hopper-8GPU模组		
处理器	2颗Intel®第四代至强®可扩展处理器, TDP 350W		2颗AMD第四代EPYCTM处理器, 最大cTDP 400W
内存	32条DDR5 DIMMs内存, 速率最高支持4800MT/s		24条DDR5 DIMMs内存, 速率最高支持4800MT/s
存储	最多支持24块2.5英寸SSD硬盘，其中最大支持16块NVMe		
M.2	2块内置M.2 NVMe/SATA（可选）		2块内置M.2 NVMe（可选）
PCIe插槽	支持10个PCIe 5.0 x16插槽（其中1个PCIe 5.0 x16插槽可替换为2个PCIe 5.0 x8速率的x16槽位） 可选支持Bluefield-3、CX7以及多种类型智能网卡		
RAID支持	可选支持RAID 0/1/10/5/50/6/60等, 支持Cache超级电容保护		
前置I/O	1个USB 3.0端口, 1个USB 2.0端口, 1个VGA端口		
后置I/O	2个USB 3.0端口, 1个MicroUSB端口, 1个VGA端口, 1个RJ45管理口		
OCP	可选支持1张OCP 3.0网卡, 支持NCSI功能		
管理	DC-SCM BMC管理模组		
TPM模块	TPM 2.0		
风扇模组	GPU区域：6个54V N+1冗余热插拔风扇 CPU区域：6个12V N+1冗余热插拔风扇		
电源	2块12V 3200W及6块54V 2700W钛金级CRPS电源，支持N+N冗余		
尺寸	宽447mm，高263mm，深860mm		
重量	净重92kg（毛重：107kg）		
环境参数	工作温度：10°C~35°C；贮存温度：-40°C~70°C		
	工作湿度：10%~90% R.H.；贮存湿度：10%~93% R.H.		

产品简介

NF5688G7是基于浪潮信息全新一代AI超融合架构平台，面向超大规模数据中心的强劲性能、极致扩展人工智能服务器，最强算力密度6U空间内搭载1块NVIDIA Hopper架构HGX-8GPU模组，系统支持4.0Tbps网络带宽，满足万亿级参数超大模型并行训练需求。搭载2颗Intel®第四代至强®可扩展处理器(TDP 350W)或2颗AMD第四代EPYC处理器(Max cTDP 400W)，最高可达4TB系统内存，128TB NVMe高速存储，支持高达12个PCIe Gen5 x16扩展槽位，可灵活支持OCP 3.0、CX7、多种智能网卡等，构建面向超大模型训练、元宇宙、自然语言理解、推荐、AIGC等场景的最强AI算力平台。

产品特性

强劲性能

- 8颗NVIDIA最新Hopper架构GPU
- 2颗Intel®第四代至强®可扩展处理器或2颗AMD第四代EPYCTM 处理器
- 集成Transformer引擎，大幅加速GPT大模型训练速度

极致能效

- 12V和54V N+N冗余电源分离供电设计，减少电源转换损耗
- 可选支持系统全液冷设计，液冷覆盖率高于80%，PUE<1.15
- 支持最高45℃进水温度，一次侧支持自然冷却

领先架构

- 节点内全PCIe 5.0高速链路，CPU至GPU带宽提升4倍
- 节点间高速互联扩展，4.0Tbps无阻塞带宽IB/RoCE组网
- 集群级优化架构设计，GPU: 计算IB: 存储IB=8:8:2

多元兼容

- 全模块化设计，一机多芯兼容，灵活配置支持本地和云端部署
- 支持大规模GPT-3/LLaMA/ChatGLM模型训练和推理
- 领先支持多样化的SuperPod解决方案，适用AIGC, AI4Science及元宇宙等丰富场景

产品规格

型号	NF5688-M7-C0-R0-00		NF5688-A7-C0-R0-00
高度	6U		
GPU	1块NVIDIA HGX-Hopper-8GPU模组		
处理器	2颗Intel®第四代至强®可扩展处理器, TDP 350W		2颗AMD第四代EPYCTM处理器, 最大cTDP 400W
内存	32条DDR5 DIMMs内存, 速率最高支持4800MT/s		24条DDR5 DIMMs内存, 速率最高支持4800MT/s
存储	最多支持24块2.5英寸SSD硬盘，其中最大支持16块NVMe		
M.2	2块内置M.2 NVMe/SATA （可选）		2块内置M.2 NVMe （可选）
PCIe插槽	支持10个PCIe 5.0 x16插槽（其中1个PCIe 5.0 x16插槽可替换为2个PCIe 5.0 x8速率的x16槽位） 可选支持Bluefield-3、CX7以及多种类型智能网卡		
RAID支持	可选支持RAID 0/1/10/5/50/6/60等, 支持Cache超级电容保护		
前置I/O	1个USB 3.0端口, 1个USB 2.0端口, 1个VGA端口		
后置I/O	2个USB 3.0端口, 1个MicroUSB端口, 1个VGA端口, 1个RJ45管理口		
OCP	可选支持1张OCP 3.0网卡, 支持NCSI功能		
管理	DC-SCM BMC管理模组		
TPM模块	TPM 2.0		
风扇模组	支持GPU、NVSwitch、CPU等冷板全覆盖，支持最高45℃高温进水 GPU区域：5个54V N+1冗余热插拔风扇 CPU区域：6个12V N+1冗余热插拔风扇		
电源	2块12V 3200W及6块54V 2700W钛金级CRPS电源，支持N+N冗余		
尺寸	宽447mm，高263mm，深860mm		
重量	净重90kg（毛重：105kg）		
环境参数	工作温度：10℃~35℃；贮存温度：-40℃~70℃		
	工作湿度：10%~90% R.H.；贮存湿度：10%~93% R.H.		