



FP5295G2 服务器

面向企业 AI 与 HPC 应用的全球最高性能计算平台节点



FP5295G2 是基于 OpenPOWER9 处理器的下一代产品系统，专为深度学习 (DL) 和人工智能 (AI)，高性能数据分析 (HPDA) 和高性能计算 (HPC) 而设计。企业可充满信心地部署深度学习框架和加速数据库等数据密集型工作负载。

产品特性

全新 AI 基础架构： CPU-GPU NVLink 互联，共享内存

全新 AI 基础架构，为企业数据密集型工作负载提供高效支撑。POWER9 CPU 与 NVIDIA Tesla V100 GPU 通过 NVLink 实现互联，唯一能在 CPU 和 GPU 之间提供高达 5.6 倍性能加速。NVlink 加速支持 CPU-GPU 共享内存，简化编程过程，使客户尽情享受 GPU 的高性能体验。

提供更丰富更高效的 IO 支持，PCIe Gen4、CAPI 2.0、OpenCAPI 等诸多特性，可提供的带宽总量是 x86 服务器中 PCIe Gen3 的 2-5.6 倍，并且延迟更低，响应更快。

为 AI 时代设计的最卓越 CPU 与 最佳 GPU 联合实践

专为 AI 设计的 POWER9 CPU，2 倍于 x86 线程数量，

更大 L3 缓存，NVLink 原生直连，超快的运行速度，把周围的一切性能充分调动起来。

- 搭载 4 块支持 NVLink 技术的 NVIDIA® Tesla® V100 GPU，支持 32G 显存特性，单 GPU 提供双精度 15.7 TFLOPS 超高计算能力。

FP5295G2 单节点提供超过每秒 500 万亿次级的 Tensor 性能，您甚至可以将他们组成每秒 3 百万亿亿次 AI 运算即服务的庞大集群。

企业级 AI 平台就绪

企业级 PowerAI DL 框架，简化深度学习部署于性能，为 AI 用户提供更强悍、更简单的端到端工具链。您可以从一个节点开始，高效地扩展至堆栈或数千个节点，且它们的性能几乎是线性增长的。

FP5295G2 能够帮助您轻松、高效地部署 AI，是您实现 AI 愿景的理想之选。

产品配置信息

	FP5295G2—标准版	FP5295G2—增强版
处理器	2颗POWER9 Monza处理器 16核 with NVLink	
处理器到内存带宽	每插槽170GB/s,每系统340GB/s	
L2到L3缓存带宽	7TB/s片上带宽	
内存	DDR4 Registered RDIMM	
	16个,高达2TB容量	
GPU	4 x 16GB NVIDIA Tesla V100 with NVLink	4 x 32GB NVIDIA Tesla V100 with NVLink
PCIe扩展插槽	4个 PCIe Gen4 插槽	
I/O接口	2 x USB 3.0, 2 x 1 GB网卡, VGA接口	
硬盘	2块2.5寸SATA HDD/SSD硬盘, 可选PCIe NVMe SSD	
RAID支持	可选支持RAID 0/1/5/6/10/50/60, 支持Cache超级电容保护	
光驱	支持外插USB光驱	
操作系统	Radhat Enterprise Liunx, Ubuntu server, 中标麒麟, 红旗	
电源	1+1冗余白金电源	
物理尺寸	441.5mm x 86mm x 822mm (宽*高*深)	
重量	30kg	
工作环境温度	5°C - 35°C	