

目录

目录	1
产品概述	2
产品功能	2
快速创建	2
Web化管理	2
VPC支持	2
监控统计	2
产品优势	2
高性能	2
大集群	2
轻资产	2
高安全	2
使用场景	2
深度学习离线训练	2
深度学习在线推理	2
图形图像渲染设计	2

产品概述

金山云GPU裸金属服务器（GPU Elastic Physical Compute，简称GEPC）是云金山云提供的搭载 GPU 的物理服务器，将物理服务器的高性能与 GPU 的高算力融合，以物理机形式交付客户，提供高性能、可独享的物理机集群服务，拥有超强的并行计算及浮点计算能力。GPU裸金属服务器支持对物理服务器全周期的云上管理，为客户提供与云物理服务器一致的云上管理体验。产品服务于AI、深度学习、图像渲染、云游戏、AR/VR等多种应用场景，为用户提供稳定、快速、弹性可扩展的计算服务，助力用户业务创新及高速发展。

产品功能

快速创建

一键式创建，分钟级部署。

Web化管理

通过Web控制台可实现对GPU裸金属服务器实例的创建，查看，续费和开关机等生命周期管理操作。

VPC支持

GPU裸金属服务器实例原生支持VPC专有网络，提供灵活的网络规划选择，便捷用户使用VPC内的各种资源。

监控统计

提供实时详细的实例监控，性能高低一目了然。具体监控指标可参考[云监控](#)。

产品优势

高性能

GPU裸金属服务器实例总计可以为用户提供8颗NVIDIA Tesla V100加速器，配备512GB内存，总计提供40960CUDA cores、5120Tensor cores和最高112TFLOPS的单精度浮点计算能力，同时采用最新的Volta架构，为深度学习和高性能计算应用提供了卓越的性能。

大集群

GPU裸金属服务器可以支持最大256个节点的100Gbps无损无收敛RDMA网络集群组网，满足用户对超算集群高带宽、低延迟的需求。

轻资产

低服务器投资风险，低服务器运维成本，始终用到最新GPU加速卡硬件。

高安全

用户私有网络（VPC）隔离，全面应用金山云安全产品。

使用场景

深度学习离线训练

利用GPU裸金属服务器的强大计算能力，将GPU裸金属服务器作为深度学习的训练平台，结合对象存储KS3提供的云存储服务、云数据库RDS提供的在线数据库服务、金山云大数据平台KMR提供的海量分布式处理服务，您可以搭建一个功能完备的深度学习离线训练系统，帮助您高效、安全的进行各种深度学习的模型训练。

深度学习在线推理

利用GPU裸金属服务器的强大计算能力，将GPU裸金属服务器作为深度学习的推理平台，结合对象存储KS3提供的云存储服务、云数据库RDS提供的在线数据库服务、金山云缓存服务KCS提供的在线缓存服务，再加上金山云VPC网络和云安全服务以及金山云内容分发网络CDN，您可以构建一套功能强大的深度学习在线推理系统，满足您的各种在线服务需求。

图形图像渲染设计

利用GPU裸金属服务器的图像处理能力，将GPU裸金属服务器作为图形图像的渲染设计平台，结合对象存储KS3提供的云存储服务、云数据库RDS提供的在线数据库服务、金山云缓存服务KCS提供的在线缓存服务，再加上金山云VPC网络和云安全服务以及金山云内容分发网络CDN，您可以构建一个强大的图形图像处理系统，完成3D设计、图像渲染、云端游戏等各种业务。