



面向 AI 和 HPC 工作 负载的 IBM Cloud GPU 解决方案

克服复杂的挑战，提升业务价值。

IBM Cloud

IBM

引言

GPU 技术日新月异，企业对人工智能和高性能计算的需求日益高涨，这使得越来越多的行业改变了挖掘数据价值、提升 IT 性能的方式。

现在，医疗卫生机构利用聊天机器人提供更优质的客户服务；石油企业利用油藏模拟器降低环境风险和勘探成本；汽车企业教会无人驾驶汽车遵守交通规则。这些数据密集型工作负载都有一个重要的共同点，那就是一个具有 GPU 计算增强力量的快速、安全且经济高效的云基础架构。

IBM Cloud 知道高性能的基础架构的合适基础会提升您的‘底线’，能够在不增加管理难度的情况下帮助您增加利润，提高产出潜力，为您的 IT 团队提供所需要素。我们最擅长的就是帮助您在特定的裸机和虚拟服务器环境内设计和配置领先的 NVIDIA GPU 架构。

从分析和图形增强，到能源勘探和机器学习，GPU 的增强力量无可争议。使用 IBM Cloud 构建 GPU 解决方案具有简单化和成本节约的优势，您将借此发掘出最大价值。

“借助 IBM Cloud 裸机服务器强大的 GPU，我们知道我们能够大幅缩短优化和验证机器学习算法的时间。”

— Jeffrey Thatcher, Spectral MD 首席科学家

目录

为什么选择 IBM Cloud 解决方案	4
GPU。不仅仅只是图形	5
各个行业所用的 GPU	6
裸机与虚拟机	7
基于 IBM Cloud 的 NVIDIA GPU	8
具体细节	9

为什么为 GPU 解决方案 选择 IBM Cloud?

选择 IBM Cloud 来满足您的 GPU 要求，正是我们的企业基础架构、平台和服务的能力所在，广泛的裸机和虚拟服务器环境，使您能够直接接入业内最灵活的服务器选择流程。同时，成本节约优势、控制力和可视性，会使您的团队克服掉采用 GPU 的障碍。

我们提供的 NVIDIA GPU 经过了精挑细选，能够无缝集成到您的 IBM Cloud 架构、API、应用和工作负载中。我们了解您的数据位置需求，IBM Cloud 能轻松满足您的需求。您可以随时访问分布在全球各地的现代化数据中心网络，提升高可用性和安全性等必备要素，满足当今企业“始终在线”、“始终安全”的需求。

易于使用	广泛分布	先进要素
可定制的裸机和虚拟服务器	全球数据中心	混合云生态系统专业知识
按需配置和快速扩展	高速私有网络选项	网络应用种类
预配置的工作负载选项	尽可能提高控制力度和可视性，满足高可用性要求	强大的 AI 软件和应用数据解决方案
无供应商锁定	快速的数据摄取产品	“白手套”服务和支持

“IBM 支持我们灵活地按需选择合适的 GPU，并可随着需求的变化，选择按月或按小时付款。目前，我们正在研究如何基于 IBM Cloud 弹性扩展 GPU 计算需求。”

—Danny Tomsett, FaceMe 公司创始人兼 CEO

GPU 不仅仅只是图形

GPU 逐渐成为了先进的通用处理器

首先，我们来快速研究一下为什么选择 GPU。通过在云环境中运行，GPU 能够与服务器的 CPU 协同运行，加速提升应用和处理性能。CPU 将应用的计算密集型工作负载转交给 GPU，GPU 能够一次性处理大量数据，而非逐步提升服务器环境的总体性能。GPU 通过同时运行多个小任务，能够面向性能高耗工作负载的要求，协调控制处理能力。

面向 GPU 增强型服务器的行业工作负载已经不再局限于图形设计。随着技术的进步，成本的降低，以及深度学习和 AI 的崛起，曾经价格高昂的附加组件变成了价格低廉的必备组件。借助 GPU 的力量，深度学习算法能够执行复杂的数学计算和统计计算，满足 IT 架构内的各类应用的需求。



各个行业所用的 GPU

GPU 让各行各业受益匪浅

行业	企业	GPU 解决方案带来的机会
金融行业	投资银行、经纪公司、商业银行、零售银行	<ul style="list-style-type: none">• 风险规避应用• 更快速地处理复杂交易
医疗保健行业	医院、研究中心	<ul style="list-style-type: none">• 安全应用问题• 大型医疗成像支持
制造业	汽车公司、航空航天公司、消费电子品公司	<ul style="list-style-type: none">• 加快上市速度• 更快速地交付创新成果
政府	国防部门、情报机构、州政府和地方政府	<ul style="list-style-type: none">• 安全和数据漏洞• 网络威胁• 不断变化的法规
建筑、工程和建设行业 (AEC)	建筑公司、工程和设计公司	<ul style="list-style-type: none">• 提供面向应用的处理能力• 支持在现场工作时快速提供远程访问
一线科研	国内外机构科研业务部门和营销机构	<ul style="list-style-type: none">• 执行复杂的分子模型计算• 挖掘人口统计数据
媒体和娱乐行业	数字工作室和机构	<ul style="list-style-type: none">• 提供基于 AI 的工具集，利用深度学习创建内容• 从媒体内容中挖掘客户洞察力• 视频会议、转码和媒体归档

裸机与虚拟机

裸机服务器的威力与虚拟服务器的扩展和快速配置优势

IBM Cloud 裸机服务器：

- 提供高度定制优势，满足任意工作负载要求
- 提高一流的安全性
- 不产生管理程序支出
- 所有资源专用于一位用户；没有“吵闹的邻居”

IBM Cloud 虚拟服务器：

- 配置时间短；预配置的虚拟服务器
- 按存储类型、内存和 GPU 数量提供灵活的实例选项
- 提供按月付费和按小时付费选项，其高效、可伸缩的特点能够满足性能高耗工作负载的要求

“我们的企业内有太多计算任务。通过在 IBM Cloud 上运行 NVIDIA GPU，过去通常需要 10 天的流程现在只需要几个小时。”

— Jeffrey Thatcher, Spectral MD 首席科学家

基于 IBM Cloud 的 NVIDIA GPU

GPU 让各行各业受益匪浅

选项	NVIDIA Tesla M60	NVIDIA Tesla K80	NVIDIA Tesla P100	NVIDIA Tesla V100
理想的环境	基础的企业性能，用于虚拟化和专业图像	可靠的企业性能，用于入门级 AI 计算	高性能，用于越来越多的先进 AI 和 HPC 功能	最大的性能，用于增强型深度学习工作负载
可用性	按月付费的裸机服务器	按月付费的裸机服务器	按月付费和按小时付费的裸机服务器 按月付费和按小时付费的虚拟服务器	按月付费的裸机服务器 按月付费和按小时付费的虚拟服务器
支持的数据中心	北美洲、欧洲、亚洲、澳大利亚和南美洲的数据中心	北美洲、欧洲、亚洲、澳大利亚和南美洲的数据中心	(裸机) 德克萨斯州达拉斯 加州圣何塞 华盛顿特区 阿姆斯特丹 首尔 东京 (虚拟) 德克萨斯州达拉斯 华盛顿特区 伦敦	德克萨斯州达拉斯 华盛顿特区
规格	1TB SATA-3.8TB SSD, 12-28 个内核, 100 MBps – 10 GBps, 64 GB – 1.5 TB			16 GB 的内存容量, 900 Gb/s 的内存带宽
入门级价格	650 美元/月	500 美元/月	1.95 美元/小时 500 美元/月	3.06 美元/小时 900 美元/月

具体细节

充分了解基础架构的细节

增加 GPU，处理最复杂的工作负载，包括机器学习、深度学习和人工智能 (AI)。根据您的选择，为裸机服务器或虚拟服务器配置使用最新的 NVIDIA Tesla GPU 或 NVIDIA GRID GPU 可实现科学计算、数据分析和大规模计算的高性能加速，而 GRID 可支持专业级虚拟化图表。

- **裸机服务器** | <https://ibm.com/cloud/bare-metal-servers>
高性能的云服务器，提供按月和按小时付费两种选项，可满足任意工作负载需求
- **虚拟服务器** | <https://ibm.com/cloud/virtual-servers>
允许您按需配置和扩展的计算资源，让您在几分钟内满足工作负载要求
- **存储** | <https://www.ibm.com/cloud/storage>
灵活的存储方法，应对由新数据源和不断发展的技术所带来的迅速增长的存储需求
- **网络** | <https://www.ibm.com/cloud/network>
致力于快速、一致地增强全球流量的解决方案
- **Watson Studio 深度学习** | <https://www.ibm.com/cloud/deep-learning>
设计复杂的神经网络，然后开展大规模试验，以部署优化的深度学习模型
- **基于 IBM Cloud 的 Veeam** | <https://ibm.co/veeam>
自动化备份、恢复和复制解决方案，全年全天候可用，旨在确保企业持续运营，并保护您的数据
- **基于 IBM Cloud 的 Zerto** | <https://ibm.co/zerto>
保护、扩展您现有的 VMware vSphere 和其他管理程序工作负载，或将它们迁移到 IBM Cloud 的全球高性能基础架构
- **IBM Cloud Object Storage** | <https://ibm.co/object-storage>
提供高持久性、弹性和安全性的高度可扩展的云存储服务



IBM Cloud

借助基于 IBM Cloud 的 NVIDIA GPU 加速 AI 和 HPC 工作负载。

如欲进一步了解或预订基于 IBM Cloud 解决方案的 NVIDIA GPU，请访问：

<https://www.ibm.biz/nvidia-gpus>

如何入门 →

© Copyright IBM Corporation 2018

IBM Corporation New Orchard Road Armonk, NY 10504

美国印刷 2018 年 10 月

IBM、IBM 徽标及 ibm.com 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 www.ibm.com/legal/copytrade.shtml 上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表。