



NVIDIA DGX-1 人工智能超级计算机

全球首款单机箱人工智能超级计算机

NVIDIA® DGX-1™ 出色的性能，可以让用户在深度学习领域实现更快的训练、处理更大的模型以及得出更精确的结果，这也是全球首款针对深度学习，以及通过人工智能进行加速分析的专用计算机，其性能堪比 250 台常规服务器。该计算机充分集成了硬件、深度学习软件、开发工具以及加速分析应用程序，从而让用户在使用过程中可明显缩短数据处理时间、对更多数据进行可视化处理、加速深度学习框架以及设计更复杂的神经网络。

更快地迭代和创新

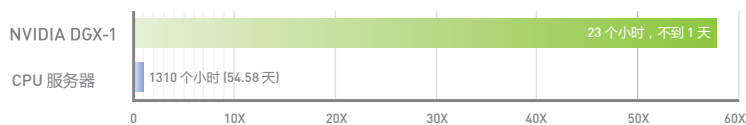
高性能训练可带来工作效率的大幅提升，从而让用户获得更精确的见解并以更短的时间将产品推向市场。



系统规格参数

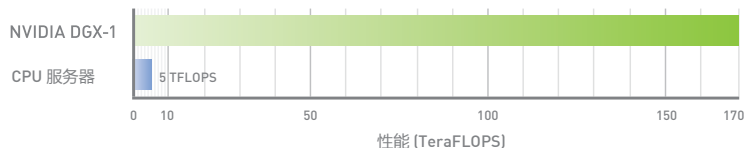
GPUs	8 颗 Tesla P100
TFLOPS (GPU FP16 / CPU FP32)	170/3
GPU 内存	每颗 GPU 16 GB
CPU	双路 20 核英特尔至强 E5-2698 v4 2.2 GHz
NVIDIA CUDA® 核心数量	28672
系统内存	512 GB 2133 MHz DDR4
存储空间	4 块 1.92 TB 固态硬盘 RAID 0
网络	双万兆以太网、4 路 IB EDR
软件	Ubuntu Server Linux 操作系统 DGX-1 专用的 GPU 驱动程序
系统重量	60KG
系统尺寸	深 866 x 宽 444 x 高 131 (mm)
包装尺寸	深 1180 x 宽 730 x 高 284 (mm)
电源最高要求	3200W
工作温度范围	10 - 30°C

NVIDIA DGX-1 可带来 58 倍的训练速度提升



相对性能 (根据训练时间计算得出)
包含 VGG-D 网络的 Caffe 基准测试，在 70 个周期内训练 128 万幅图像 | CPU 服务器采用 2 颗至强 E5-2699v4 CPU

NVIDIA DGX-1 可带来 34 倍以上的性能提升



CPU 为双路英特尔至强 E5-2699v4。170TF 为半精度 (FP16 精度) 下的性能

全新计算带来无限机遇

NVIDIA DGX-1 是第一款采用了 NVIDIA Tesla® P100 加速器 (基于 NVIDIA Pascal™ 架构) 的计算机, 其搭载的 NVIDIA NVLink™ 技术可使 GPU 内存容量大幅提升, 从而让计算机能够学习、观察以及模拟我们的世界。

分析、可视化、人工智能加速

NVIDIA DGX-1 软件堆栈包含各大深度学习框架、NVIDIA DIGITS™ GPU 训练系统、NVIDIA 深度学习软件开发包 (例如 CuDNN 和 NCCL)、NVIDIA Docker、GPU 驱动程序以及 NVIDIA CUDA®, 可用于快速设计深度神经网络 (DNN)。该软件堆栈非常适合用来加速当前主流的分析 and 可视化软件。

这一强大的计算机可访问用于容器创建与部署的云管理服务、可实现系统更新以及访问应用程序资源库。这些软件结合到一起, 借助 Pascal 架构 Tesla GPU 的性能, 可令应用程序的运行速度比使用以往的 GPU 加速型解决方案快 12 倍。

将数据更快地转化为知识

带有创新技术的 NVIDIA DGX-1 计算机让用户能够在数小时甚至数分钟内揭开大型数据集的奥秘, 从而学到新的知识, 获取新的见解。

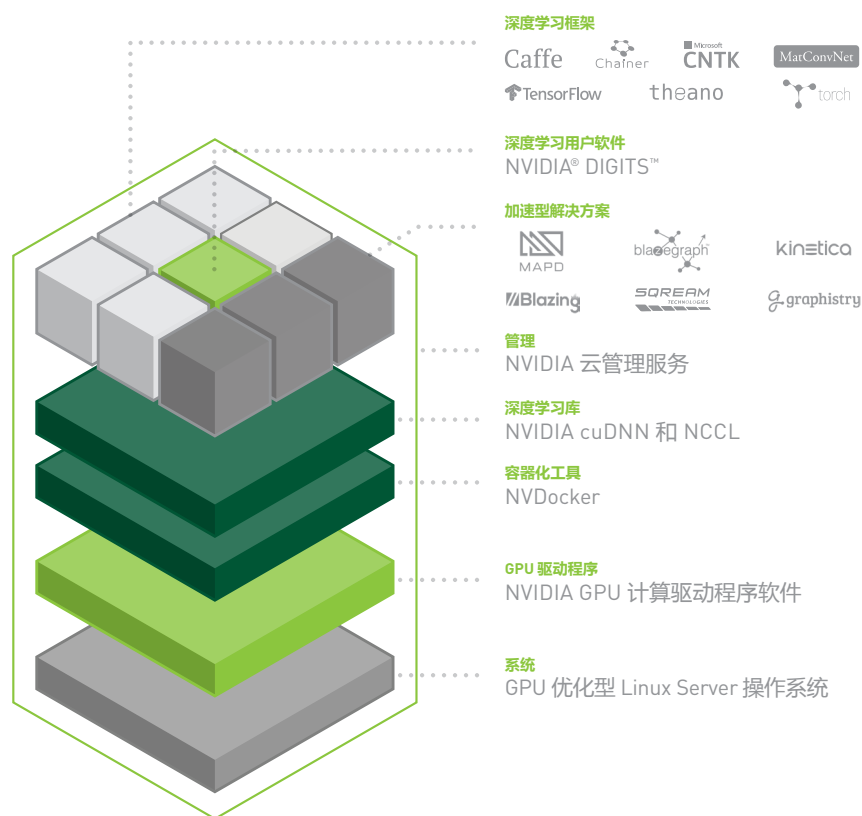
保持竞争优势

NVIDIA DGX-1 采用开创性的技术开发而成, 可为深度学习训练, 以及通过人工智能进行的加速分析提供更快的解决方案。

投资效益最大化

硬件与软件支持服务不但可以让用户畅享 NVIDIA 专业的深度学习技术, 还可以享受云管理服务、软件升级与更新以及优先解决重要问题等服务。

NVIDIA DGX-1 软件堆栈



如需了解有关 NVIDIA DGX-1 的更多信息, 敬请访问 www.nvidia.cn/dgx1

© 2016 NVIDIA 公司版权所有, 保留所有权利。NVIDIA、NVIDIA 徽标以及 IRAY 均为 NVIDIA 公司的商标和/或注册商标。所有公司和产品名称均为相应所有者的商标或注册商标。特性、定价、上市情况以及规格参数可能会随时更改, 恕不另行通知。2016 年 12 月

