



NVIDIA DGX-1 AI 研究的基石

建立 AI 平台所面临的挑战

数据科学家利用深度学习和分析的强大能力，并依靠计算性能来更快地获得见解并进行创新。GPU 技术提供通往 AI 的更快途径，但是建立平台远不止部署服务器和 GPU。

AI 和深度学习要求在软件工程方面投入大量精力。这项投资可能会导致您的项目延误几个月，因为您需要集成复杂的组件及软件堆栈，其中包括框架、库和驱动程序。完成部署后，在您等待不断升级的开源软件变稳定的过程中，也需要不断投入额外的时间和资源。您还需要等待优化基础架构以提高性能，而管理成本会随着系统规模的扩大而增加。

通往深度学习的更快路径

为满足 AI 与数据科学的需求，NVIDIA® DGX-1™ 通过开箱即用的解决方案来加快实施您的 AI 计划，如此一来，您可以在几小时而非数月内获得见解。借助 DGX-1，再加上集成式 NVIDIA 深度学习软件堆栈和 DGX-1 云管理服务，您只需插入并开启电源，即可开始工作。现在，您只需短短一天时间即可开始深度学习训练，而不必花费数月来集成和配置软硬件以及进行错误排查。

轻松取得工作成果

NVIDIA DGX-1 可消除不断优化深度学习软件的负担，而且提供经优化的即用型软件堆栈，可让您节省数十万美元。它可使用当今最热门的深度学习框架、NVIDIA DIGITS™ 深度学习训练应用程序、第三方加速解决方案、NVIDIA 深度学习 SDK（例如，cuDNN、cuBLAS）、CUDA® 工具包、快速多 GPU 群集 NCCL、NVIDIA Docker 以及 NVIDIA 驱动程序。



系统规格

GPUs	8 块 Tesla V100	8 块 Tesla P100
TFLOPS (GPU FP16)	960	170
GPU 显存	整个系统 128 GB	
CPU	双路 20 核 Intel Xeon E5-2698 v4 2.2 GHz	
NVIDIA CUDA® 核心数量	40,960	28,672
NVIDIA Tensor 核心数量 (基于 V100)	5,120	无
最大功率要求	3,200 W	
系统内存	512 GB 2,133 MHz DDR4 LRDIMM	
存储空间	4 块 1.92 TB SSD RAID 0	
网络	双 10 GbE，高达 4 IB EDR	
软件	Ubuntu Linux 操作系统 如需详细信息，请参阅软件堆栈	
系统重量	60 千克	
系统尺寸	866 深 x 444 宽 x 131 高 (毫米)	
包装尺寸	1,180 深 x 730 宽 x 284 高 (毫米)	
操作温度范围	10–35 °C	

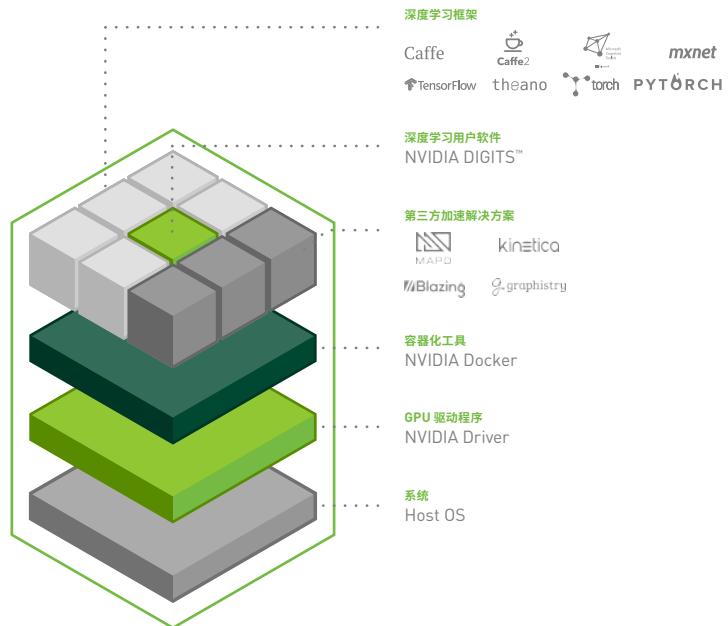
革命性的 AI 性能

尽管许多解决方案都提供 GPU 加速性能,但只有 NVIDIA DGX-1 发挥出了最新 NVIDIA GPU(例如 NVIDIA Tesla® V100)的全部潜力,包括下一代 NVLink™ 和全新 Tensor 核心架构等创新科技。相比其他基于 GPU 的系统,DGX-1 凭借其注重性能的深度学习软件堆栈,将训练速度提升高达三倍。在 NVIDIA DGX-1 的支持下,单一系统的计算能力可媲美 25 个机架的传统服务器;该系统将最新 NVIDIA GPU 技术与当今最先进的深度学习软件堆栈集成,为您提供革命性的性能,让您以前所未有的速度获得见解。

投资保护

您的 AI 计划对于贵公司的成功至关重要,而且依赖于频繁优化的软件堆栈和集成式硬件基础架构。当今,开源软件快速升级,而库、驱动程序和硬件也十分复杂。令人深感欣慰的是每个 DGX-1 背后都有 NVIDIA 企业级支持和软件工程专业知识提供保障。此软件堆栈以多年的研究、创新和深度学习专业知识为基础,并通过每月优化的框架版本进行维护。而且,NVIDIA 支持服务还包括软件升级和优先解决紧急问题;您大可放心,您的环境经过调优,可实现最佳的性能及正常运行时间。

NVIDIA DGX-1 软件堆栈



NVIDIA DGX-1 可将训练速度提高 96 倍



工作负载: ResNet50, 90次数据训练 | CPU 服务器: 双路 Xeon E5-2699 v4, 2.6GHz

如需了解更多信息,请访问 www.nvidia.cn/dgx-1