



加速计算是可持续计算

在 AI 时代提高能效并降低成本



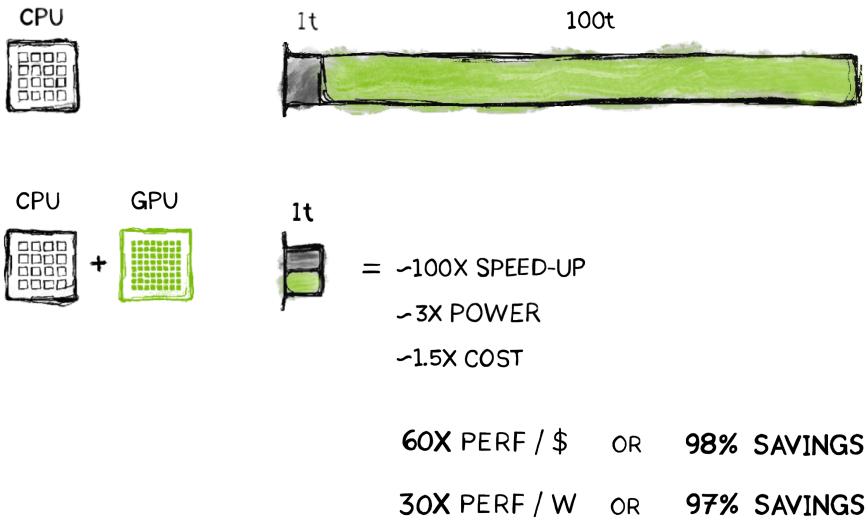
过去 60 年来，从 PC 的推出到移动技术和云技术的出现，计算行业经历了几次重大转变。整个行业现在正处于一个关键时刻——广泛使用的处理引擎 CPU 不断发展，但成本高昂。当今的 CPU 性能代际提升逐渐减弱，需要的功率也越来越大。与此同时，随着**企业持续创新**，人们对更大计算能力的需求也在飞速增长。

不断增长的数据量与 CPU 性能滞后之间的差距日益扩大，导致计算膨胀。数据量呈指数级增长但 CPU 性能无法持续跟进。随着数据中心为了满足需求而不断扩展，能耗越来越高。在这种日益扩大的差距下，需要转向更高效的计算解决方案，以更少的能耗处理计算密集型工作负载。

每单位面积的晶体管数量大约每两年翻一番（被称为“摩尔定律”），这项存在已久的预测已不再现实。

提升性能和能效

NVIDIA 开发了一种可将 GPU 与 CPU 配合使用的架构，可以显著提升处理速度。将这两种处理器并行运行，并且每个处理器自主独立运行，可以大幅提升性能和能效。以前需要 100 时间单位的任务现在只需 1 时间单位即可完成。这样可以显著节省每项任务所需的时间、精力和成本，为科学计算、医疗健康、气候技术等领域带来了突破。



在 PC 行业，这种做法非常普遍。例如，将价值 500 美元的 NVIDIA® GeForce™ GPU 集成到价值 1000 美元的 PC 中，可以显著提升 PC 的性能。同样，从纯 CPU 系统向 GPU 加速系统过渡，高性能计算和 AI 工作负载**每年可以节省超过 40 太瓦时的电力**，相当于近 500 万美国家庭的电力需求。

面向未来的基础设施

为了适应生成式 AI 等全新工作负载，同时实现环境、社会和治理 (ESG) 目标，加速计算投资至关重要。**生成式 AI** 正在重塑各个行业，为创新和增长创造了新的机遇。它意味着主要基于检索的传统计算从根本上转变为计算机生成上下文相关内容的模型。这种转变需要全新硬件和软件解决方案，还需要重新评估现有计算策略。加速计算提供全面的硬件、软件和服务套件，不仅可以为生成式 AI 等未来工作负载提供支持，还可以助力工业数字化发展，实现AI 的行业级部署。

展望未来，加速计算的影响不仅仅局限于技术方面的增强，也预示着人们将转向可持续性更强、更加经济高效的计算实践。许多公司在数据处理方面投入了数亿美元。将加速计算应用于该流程，可以大幅节省财务成本或大幅提升计算能力，同时为全新的创新工作负载奠定基础。

长期以来，计算通胀使通用计算不堪重负。选择加速各项流程，企业就可以弥补大量固有成本损失并持续创新，而不会受到硬件的物理限制。

准备好开始了吗？

详细了解 NVIDIA 和 **加速计算**。

© 2024 NVIDIA Corporation 及其关联公司。保留所有权利。NVIDIA 和 NVIDIA 徽标均为 NVIDIA Corporation 及其关联公司在美国和其他国家/地区的商标和/或注册商标。其他公司名称和产品名称可能是各个相应所有者的商标。3356796。OCT24

