



NVIDIA A16

为具有丰富图形功能的虚拟桌面基础架构 (VDI) 提供前所未有的用户体验和密度

基于 NVIDIA Ampere 架构的 NVIDIA A16 与 NVIDIA vGPU 软件相结合，提高具有丰富图形功能的虚拟桌面基础架构 (VDI) 的用户体验标准。随着越来越多的组织将远程工作作为长期战略，采用 NVIDIA vPC 或 vApp 软件的 A16，让各种产业的知识工作者能够更大限度地提高生产效能，其性能可以媲美本地 PC。与上一代 M10 相比，NVIDIA A16 可提供高达 2 倍的用户密度，从而减少所需的硬件资源数量并降低总拥有成本 (TCO)。

在与 NVIDIA RTX vWS 软件结合使用时，A16 可以作为经济实惠的入门级虚拟工作站，是运行计算机辅助设计 (CAD) 等工作负载的理想工作站。A16 采用独特的四 GPU 主板设计，可布局混合用户配置文件大小，因此 IT 部门可以支持轻型虚拟 PC 工作负载以及具有更大显存和图形需求的用户。同时还支持在主板上混合用户类型，从而能够在单个主板上布局虚拟 PC、虚拟工作站，甚至虚拟化计算。

出色的流媒体性能

NVIDIA A16 具有更多的视频编码器和解码器，单个 A16 主板中整合了 4 个片上硬件编码器 (NVENC) 和 8 个解码器 (NVDEC) 单元。这提供了出色的编码、解码和转码性能，进而最大化每一个 A16 主板的视频流数量，且与其他产品相比，具有更高的性价比。

规格	
GPU 架构	NVIDIA Ampere 架构
GPU 显存	4x 16 GB GDDR6
显存带宽	4x 200 GB / 秒
纠错码 (ECC)	是
基于 NVIDIA Ampere 架构的 CUDA 核心	4x 1280
NVIDIA 第三代 Tensor Core	4x 40
NVIDIA 第二代 RT Core	4x 10
FP32 TF32 TF32 ¹ (TFLOPS)	4x 4.5 4x 9 4x 18
FP16 FP16 ¹ (TFLOPS)	4x 17.9 4x 35.9
INT8 INT8 ¹ (TOPS)	4x 35.9 4x 71.8
系统接口	PCIe Gen 4 (x16)
最大功耗	250 瓦
散热解决方案	被动
外形尺寸	全高、全长 (FHFL) 双插槽
电源接口	8-pin CPU
编码 / 解码引擎	4 个 NVENC / 8 个 NVDEC (包括 AV1 解码)
通过 GPU 硬件信任根进行安全与测试的开机	是
vGPU 软件支持	NVIDIA 虚拟 PC (vPC) NVIDIA 虚拟应用 (vApp) NVIDIA RTX 虚拟工作站 (vWS) NVIDIA AI Enterprise NVIDIA 虚拟计算 服务 (vCS)
图形 API	DirectX 12.07 ² ShaderModel 5.17 ² OpenGL 4.68 ³ Vulkan 1.18 ³
计算 API	CUDA、DirectCompute、OpenCL™、OpenACC™
MIG 支持	否

实现 VDI 现代化并提升流媒体性能



专为加速 VDI 而设计

优化用户密度，并与 NVIDIA vPC 软件结合，可从任何位置存取图形功能丰富的虚拟 PC。



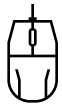
平价的虚拟工作站

为入门级虚拟工作站提供庞大的每个用户的大帧缓存配备 NVIDIA RTX vWS 软件，可运行计算机辅助设计 (CAD) 等工作负载。



灵活支持不同用户类型

独特的四 GPU 主板设计，支持在单个主板上布局混合用户设定文件大小和用户类型，例如虚拟 PC 和虚拟工作站。



卓越的用户体验

与仅使用 CPU 的 VDI 相比提供更高的帧率和更低的最终用户延迟，从而提高应用程序的响应、速度，并媲美原生电脑或工作站的用户体验。



用户密度翻倍

专为图形功能丰富的 VDI 而构建，采用双插槽外形规格，每个主板最多可支持 64 个并发用户。



高分辨率显示器

支持多个高分辨率显示器，在 VDI 环境中实现更高的工作效率和逼真的品质。



超过 2 倍的编码器吞吐量

相较于上一代 M10 编码器吞吐量增加了两倍，可以提供高性能转码以及多媒体和多流视频所需的用户性能。



更高质量的视频影像

支持最新的编解码器，包括 H.265 编码 / 解码、包括 H.265 编码 / 解码、VP9 和 AV1 解码，确保获得高品质的视频影像体验。



NVIDIA Ampere 架构

基于 NVIDIA Ampere 架构的 CUDA 核心，第二代 RT Core，以及第三代 Tensor Core，可以灵活的控制搭载 NVIDIA RTX vWS 软件的虚拟工作站，或利用未使用的 VDI 资源，通过 NVIDIA AI Enterprise 软件来运行计算工作负载。



PCI EXPRESS GEN 4

支持来自 CPU 内存的 PCI Express Gen 4 数据传输速度，以处理资料密集型任务。

特性

- 专为搭载 NVIDIA vPC 的图形丰富型虚拟桌面架构 (VDI) 而构建
- 借助 NVIDIA RTX vWS，为每位虚拟工作站用户提供最低的成本⁴
- 支持所有 NVIDIA vGPU 软件版本：NVIDIA vPC、NVIDIA vApp、NVIDIA RTX vWS、NVIDIA AI Enterprise 和 NVIDIA vCS
- PCI Express Gen 4
- 最新的编解码器支持：H.265 编码 / 解码、VP9 和 AV1 解码

准备好开始了吗

如需深入了解 NVIDIA A16，请访问：

nvidia.cn/data-center/products/a16-gpu/

¹ 启用结构稀疏性

² GPU 支持 DX 12.0 API，硬件功能级别 12+1

³ 产品基于已发布的 Khronos 规格为准，预计在上市时通过 Khronos 符合性测试流程。如需了解当前的符合性状况，请访问 www.khronos.org/conformance

⁴ NVIDIA A16 与 T4、RTX 6000、RTX 8000 和 A40 虚拟工作站的比较。

