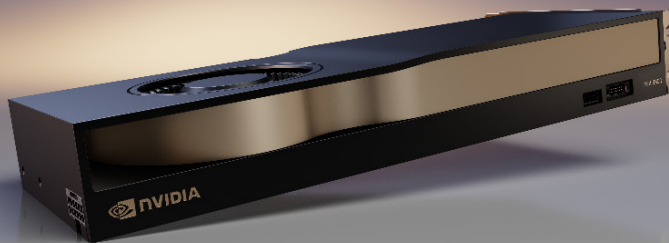




NVIDIA RTX 5000 Ada 架构 GPU

强大性能, 成就无限可能



为下一个创新时代赋能

各行各业都在拥抱加速计算和 AI, 以应对强大的动态变化并解锁变革带来的可能性。生成式 AI 正在重塑从设计与工程、娱乐和医疗健康等各个领域的专业人员的创作和创新方式。NVIDIA RTX™ 5000 Ada 架构 GPU 采用第三代 RTX 技术, 为生成式 AI 带来了实现突破性进展, 彻底改变了工作效率, 并提供了前所未有的创意空间。

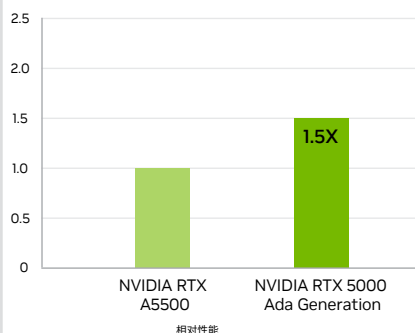
NVIDIA RTX 5000 Ada GPU 基于 NVIDIA Ada Lovelace 架构, 专为当今的专业工作流程而打造。结合使用了 100 个第三代 RT Core、400 个第四代 Tensor Core、12800 个 CUDA® 核心以及 32 GB 的显存, 可提供新一代 AI 图形和 petaFLOPS 推理性能, 在加速渲染、AI、图形和计算工作负载性能方面表现卓越。RTX 5000 赋能的工作站使您能够在当今严苛的商业环境中取得成功。

NVIDIA RTX 专业级显卡经过认证, 适用于由领先的独立软件供应商 (ISV) 和工作站制造商的测试, 并由全球专家团队提供支持。借助适用于任务关键型企业的顶级可视化计算解决方案, 您可安心专注于重要工作。

主要特性

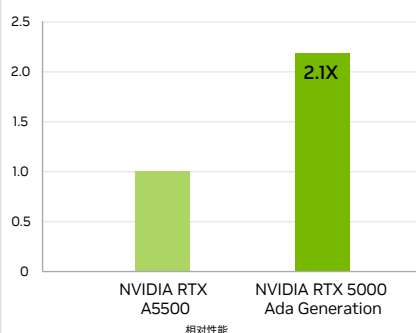
- > PCIe Gen4
- > 四个 DisplayPort 1.4a 接口
- > 支持 AV1 编码和解码
- > 带音频的 DisplayPort
- > 通过立体连接器支持 3D 立体
- > NVIDIA® GPUDirect® 视频支持
- > 支持 NVIDIA GPUDirect 远程直接内存访问 (RDMA)
- > NVIDIA Quadro® Sync II¹ 兼容性
- > NVIDIA RTX Experience
- > NVIDIA RTX 桌面管理软件
- > NVIDIA RTX IO 支持
- > 支持 HDCP 2.2
- > NVIDIA Mosaic² 技术

训练



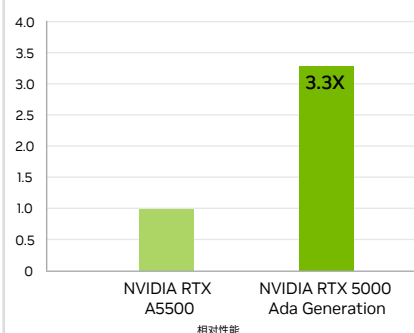
运行测试基于 Intel Core i9-12900K 处理器 @ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, PyTorch v2.1.0, NVIDIA 驱动 528.86。JASPER 训练阶段的相对加速, 精度 = 混合, batch size = 64。性能基于预发布的版本, 可能会发生变化。

渲染



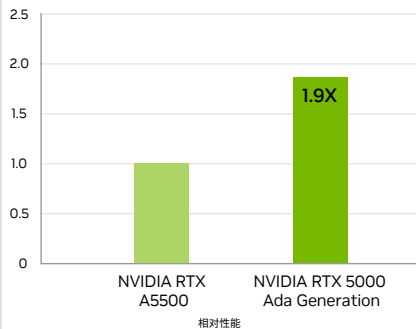
运行测试基于 Intel Core i9-12900K 处理器 @ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, Chaos V-Ray v5.0, NVIDIA 驱动 536.15。1920x1080 分辨率下的相对加速, 场景 12 管线。子测试渲染时间 (秒)。性能基于预发布的版本, 可能会发生变化。

Omniverse



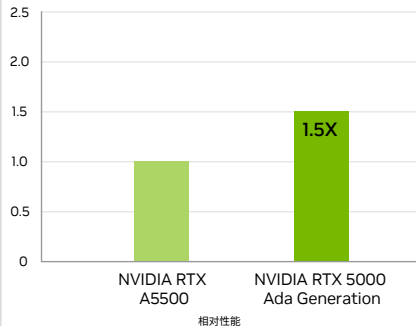
运行测试基于 Intel Core i9-12900K 处理器 @ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, NVIDIA 驱动 528.49。CAD 应用程序性能基于 NVIDIA Omniverse Create 使用不同大小和渲染复杂性的多个模型进行的内部测试。性能以每秒渲染的帧数衡量。NVIDIA DLSS 3 适用于 NVIDIA RTX 5000 Ada 架构的 GPU, DLSS2 适用于非 Ada 架构的 GPU。性能基于预发布的版本, 可能会发生变化。

图形



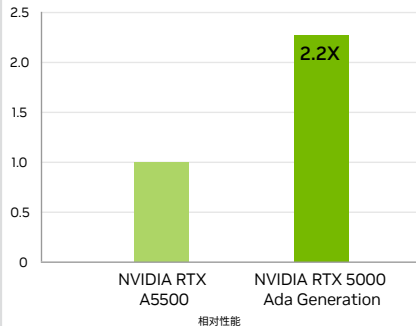
运行测试基于 Intel Core i9-12900K 处理器@ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x 64, SPECviewperf 2020, NVIDIA 驱动 528.49, 4K Siemens NX 综合评分的相对加速。性能基于预发布的版本, 可能会发生变化。

生成式 AI



运行测试基于 Intel Core i9-12900K 处理器@ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x 64, Stable Diffusion WebUI v1.3.1, NVIDIA 驱动 536.15。512 x 512 图像生成。性能基于预发布的版本, 可能会发生变化。

高性能计算



行测试基于 Intel Core i9-12900K 处理器@ 3.2 GHz (5.2 GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, CUDA 11.8 (cuBLAS 性能), NVIDIA 驱动 525.85。GFLOPS 的相对加速, 精度= INT8, input= zero。性能基于预发布的版本, 可能会发生变化。

规格

GPU 显存	32GB GDDR6
显存位宽	256 位
显存带宽	576GB/s
纠错码 (ECC)	是
基于NVIDIA Ada Lovelace 架构的 CUDA 核心	12,800
NVIDIA 第四代 Tensor Core	400
NVIDIA 第三代 RT Core	100
单精度浮点运算性能	65.3 TFLOPS ³
RT Core 性能	151.0 TFLOPS ³
Tensor 性能	1044.4 TFLOPS ⁴
系统接口	PCIe 4.0 x16
功耗	主板总功率: 250 W
散热方式	主动
外形规格	11.2 厘米 (高) x 26.7 厘米 (长), 双插槽
显示器接口	4x DisplayPort 1.4a ⁵
同时显示的最大数量	4x 4096 x 2160 @ 120Hz 4x 5120 x 2880 @ 60Hz 2x 7680 x 4320 @ 60Hz
编码/ 解码引擎	2 个编码, 2 个解码 (+ AV1 编码) 和解码
VR Ready	是
vGPU 软件支持 ⁶	> NVIDIA vPC/vApps > NVIDIA RTX 虚拟工作站
支持的 vGPU 配置文件	查看 vGPU 许可指南
图形 API	DirectX 12, Shader Model 6.7, OpenGL 4.6 ⁷ , Vulkan 1.3 ⁷
计算机 API	CUDA 12.2, OpenCL 3.0, DirectCompute
NVIDIA NVLink [®]	否

准备好开始了吗？

如需详细了解 NVIDIA RTX 5000，请访问：
www.nvidia.cn/design-visualization/rtx-5000/

1 Quadro Sync II 显卡单独销售。2 Windows 10 和 Linux。3 峰值速率基于 GPU 加速频率。4 使用稀疏技术时的有效 FP8teraFLOPS (TFLOPS)。5 默认情况下, RTX 5000 的显示端口处于开启状态。使用 vGPU 软件时, 显示端口处于不开启状态。6 即将发布的 NVIDIA vGPU 版本 (预计于 2023 年第三季度发布) 将提供对 RTX 5000 Ada Generation GPU 的虚拟化支持。7 产品基于已发布的 Khronos 规范, 预计将通过 Khronos 一致性测试流程。当前的一致性状态可在此查看 www.khronos.org/conformance

© 2023 NVIDIA Corporation. 保留所有权利。NVIDIA、NVIDIA 徽标、CUDA、GPUDirect、NVLink、Quadro 和 RTX 均为 NVIDIA Corporation 在美国和其他国家/地区的商标和/或注册商标。其他公司和产品名称可能是其各自关联公司的商标。其他所有商标均为其各自所有者的资产。2911950 2023 年 8 月。

