



New NVIDIA RTX GPUs Power Next Generation of Workstations and PCs for Millions of Artists, Designers, Engineers and Virtual Desktop Users

NVIDIA Ampere Architecture GPUs Provide Computing Power for Professionals to Work from Anywhere

GTC -- NVIDIA today announced a range of eight new NVIDIA Ampere architecture GPUs for next-generation laptops, desktops and servers that make it possible for professionals to work from wherever they choose, without sacrificing quality or time.

Using the new NVIDIA RTX™ GPUs, artists can create dazzling 3D scenes, designers can iterate on stunning building architecture in real time, and engineers can create groundbreaking products on any system.

“Hybrid work is the new normal,” said Bob Pette, vice president of Professional Visualization at NVIDIA. “RTX GPUs, based on the NVIDIA Ampere architecture, provide the performance for demanding workloads from any device so people can be productive from wherever they need to work.”

For desktops, the new [NVIDIA RTX A5000](#) and [NVIDIA RTX A4000](#) GPUs feature new RT Cores, Tensor Cores and CUDA® cores to speed AI, graphics and real-time rendering up to 2x faster than previous generations.

For professionals on the go needing thin and light devices, the new NVIDIA RTX A2000, NVIDIA RTX A3000, RTX A4000 and RTX A5000 [laptop GPUs](#) deliver accelerated performance without compromising mobility. They include the latest generations of Max-Q and RTX technologies and are backed by the [NVIDIA Studio™](#) ecosystem, which includes exclusive driver technology that enhances creative apps for optimal levels of performance and reliability.

For the data center, there are the new [NVIDIA A10 GPU and A16 GPU](#). The A10 provides up to 2.5x the virtual workstation performance of the previous generation for designers and engineers, while the A16 GPU provides up to 2x user density with lower total cost of ownership and an enhanced virtual desktop infrastructure experience over the previous generation.

Combined with [NVIDIA RTX Virtual Workstation \(vWS\)](#) and [NVIDIA Virtual PC \(vPC\) software](#), the A10 and A16 GPUs deliver more power, memory and speed to boost any workflow, from graphics and AI to VDI, so users can focus more time on working efficiently and productively.

Customer Adoption

Among the first to tap into the RTX A5000 is Woods Bagot, a global architectural and consulting practice.

“This new NVIDIA Ampere GPU microarchitecture is a huge leap forward. I was shocked by the performance gain I saw with the NVIDIA RTX A5000 versus the equivalent model previous generation RTX when running applications such as Chaos Vantage and NVIDIA Omniverse,” said Robert Cervellione, studio leader of Design Technology at Woods Bagot. “It’s clear the NVIDIA RTX A5000 professional GPU was designed for this intensive multi-tasking compared to game-level cards.”

Polaris, a manufacturer of motorcycles, snowmobiles, ATV and electric vehicles, is also using the RTX A5000.

“Speed is everything when we need to evaluate new concepts for the most adventurous vehicles, and the NVIDIA RTX A5000 really delivers what we need,” said EM Green, 3D/CGI lead at Polaris. “The basic viewport rendering is incredibly fast in Octane Render — 5x faster — and unlocks things we couldn’t have even tried before.”

“At Important Looking Pirates, we’re constantly pushing creative boundaries, and breakthrough technology is the great enabler,” said Ola Björling, producer of Virtual Production & Real-Time Graphics at Important Looking Pirates. “With the RTX A5000 powering NVIDIA Omniverse, we’re seeing high-end cinematic results in real time that used to take hours per frame.”

Next-Generation RTX Technology

The new NVIDIA RTX GPUs feature the latest generation of NVIDIA RTX technology to accelerate graphics, ray tracing, AI and compute. The new RTX laptop GPUs also include the latest generation of NVIDIA Max-Q technology, enabling a new generation of powerful thin and light laptops designed for professionals. Additionally, RTX technology powers [NVIDIA Omniverse](#) — the only platform that delivers high-performance, physically accurate simulation for complex 3D worlds and true real-time ray- and path-traced rendering with ease.

All the NVIDIA RTX GPUs feature the latest technologies in the NVIDIA Ampere architecture:

- **Second-Generation RT Cores:** Up to 2x the throughput of the previous generation, with the ability to run concurrent

ray tracing, shading and denoising tasks.

- **Third-generation Tensor Cores:** Up to 2x the throughput of the previous generation, up to 10x with sparsity, with support for new TF32 and BFloat16 data formats.
- **CUDA Cores:** Up to 2.5x the FP32 throughput of the previous generation for significant increases in graphics and compute workloads.

Desktop GPU features and technology include:

- **Up to 24GB of GPU memory:** Double the memory of the previous generation, the RTX A4000 with 16GB GDDR6 memory and the RTX A5000 with 24GB of GDDR6 memory both support ECC memory for error-free computing. The RTX A5000 is expandable up to 48GB of memory using NVIDIA NVLink[®] to connect two GPUs.
- **Virtualization:** The RTX A5000 supports NVIDIA RTX vWS software for multiple high-performance virtual workstation instances that enable remote users to share resources to drive high-end design, AI and compute workloads.
- **PCIe Gen 4:** Doubles the bandwidth of the previous generation and speeds up data transfers for data-intensive tasks such as AI, data science and creating 3D models.

Laptop GPU features and technology include:

- **Third-Gen Max-Q technology:** For thin and light laptops that perform quieter and more efficiently with Dynamic Boost 2.0, WhisperMode 2.0, Resizable BAR and NVIDIA DLSS technology.
- **Up to 16GB of GPU memory:** For the largest models, scenes, assemblies and advanced multi-application workflows.

NVIDIA also introduced the NVIDIA T1200 and NVIDIA T600 laptop GPUs, based on its previous-generation Turing architecture. Designed for multi-application professional workflows, these GPUs are a significant upgrade in performance and capabilities from integrated graphics.

Availability

The new NVIDIA RTX desktop GPUs and NVIDIA data center GPUs will be available from global distribution partners and OEMs starting later this month.

The new NVIDIA RTX laptop GPUs will be available in mobile workstations anticipated in Q2 this year from global OEMs.

About NVIDIA

[NVIDIA](https://nvidianews.nvidia.com/)'s (NASDAQ: NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market and has redefined modern computer graphics, high performance computing and artificial intelligence. The company's pioneering work in accelerated computing and AI is reshaping trillion-dollar industries, such as transportation, healthcare and manufacturing, and fueling the growth of many others. More information at <https://nvidianews.nvidia.com/>.

Certain statements in this press release including, but not limited to, statements as to: the features, performance, benefits, impact and availability of NVIDIA RTX GPUs, the NVIDIA Ampere architecture, and the NVIDIA T1200 and NVIDIA T600 laptop GPUs; and hybrid work as the new normal are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the most recent reports NVIDIA files with the Securities and Exchange Commission, or SEC, including, but not limited to, its annual report on Form 10-K and quarterly reports on Form 10-Q. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

© 2021 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, CUDA, NVIDIA RTX, NVIDIA Studio and NVLink are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. MAXQ is the registered trademark of Maxim Integrated Products, Inc. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.

Kasia Johnston
+1-415-813-8859
kasiaj@nvidia.com