NVIDIA Reveals the Titan of Turing: TITAN RTX

Turing-Powered TITAN Delivers 130 Teraflops of Deep Learning Horsepower, 11 GigaRays of Ray-Tracing Performance to World's Most Demanding Users

NeurIPS--NVIDIA today introduced NVIDIA® TITAN RTX™, the world's most powerful desktop GPU, providing massive performance for AI research, data science and creative applications.

Driven by the new NVIDIA Turing™ architecture, TITAN RTX -- dubbed T-Rex -- delivers 130 teraflops of deep learning performance and 11 GigaRays of ray-tracing performance.

"Turing is NVIDIA's biggest advance in a decade - fusing shaders, ray tracing, and deep learning to reinvent the GPU," said Jensen Huang, founder and CEO of NVIDIA. "The introduction of T-Rex puts Turing within reach of millions of the most demanding PC users -- developers, scientists and content creators."

Ultimate PC GPU
NVIDIA's greatest leap since the invention of the CUDA® GPU in 2006 and the result of more than 10,000 engineering-years of effort, Turing features new RT Cores to accelerate ray tracing, plus new multi-precision Tensor Cores for AI training and inferencing. These two engines -- along with more powerful compute and enhanced rasterization -- enable capabilities that will transform the work of millions of developers, designers and artists across multiple industries.

Designed for a variety of computationally demanding applications, TITAN RTX provides an unbeatable combination of AI, real-time ray-traced graphics, next-gen virtual reality and high performance computing. It delivers:

• 576 multi-precision Turing Tensor Cores, providing up to 130 teraflops of deep learning performance.
• 72 Turing RT Cores, delivering up to 11 GigaRays per second of real-time ray-tracing performance.
• 24GB of high-speed GDDR6 memory with 672GB/s of bandwidth -- 2x the memory of previous-generation TITAN GPUs -- to fit larger models and datasets.
• 100GB/s NVIDIA NVLink® can pair two TITAN RTX GPUs to scale memory and compute.
• Incredible performance and memory bandwidth for real-time 8K video editing.
• VirtualLink™ port provides the performance and connectivity required by next-gen VR headsets.

Built for AI Researchers and Deep Learning Developers
TITAN RTX transforms the PC into a supercomputer for AI researchers and developers. TITAN RTX provides multi-precision Turing Tensor Cores for breakthrough performance from FP32, FP16, INT8 and INT4, allowing faster training and inference of neural networks. It offers twice the memory capacity of previous generation TITAN GPUs, along with NVLink to allow researchers to experiment with larger neural networks and data sets.

Perfect for Data Scientists
A powerful tool for data scientists, TITAN RTX accelerates data analytics with RAPIDS. RAPIDS open-source libraries integrate seamlessly with the world's most popular data science workflows to speed up machine learning.

Content Creators Create Their Best Work
TITAN RTX brings the power of real-time ray tracing and AI to creative applications, so 5 million PC-based creators can iterate faster. It also delivers the computational horsepower and memory bandwidth needed for real-time 8K video editing.

Available This Month
TITAN RTX will be available later this month in the U.S. and Europe for $2,499.

About NVIDIA
NVIDIA's (NASDAQ: NVDA) invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI — the next era of computing — with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. More information at http://nvidianews.nvidia.com/.

Certain statements in this press release including, but not limited to, statements as to: the performance, benefits, abilities and availability of the NVIDIA TITAN RTX, including its deep learning and ray-tracing performance, memory, compute, video editing and VirtualLink port which provides the performance and connectivity required by next-gen head-mounted displays; NVIDIA TITAN RTX being the world's most powerful desktop GPU and providing massive performance; Turing representing NVIDIA's most revolutionary advance in more than a decade, and it delivering a new era in graphics and AI; NVIDIA TITAN RTX putting Turing in within reach of millions of the most demanding PC users; Turing being NVIDIA's greatest leap since the invention of CUDA; the number of engineering-years of effort to develop Turing; Turing's features, including RT Cores and Tensor Cores; RT Cores and Tensor Cores enabling capabilities that will transform the work of millions of developers, designers and artists across industries; TITAN RTX providing an unbeatable combination of AI, graphics, virtual reality and high performance computing; TITAN RTX transforming a PC into a supercomputer and its features allowing for faster training and inference of neural networks; TITAN RTX accelerating data analytics with RAPIDS which integrate into data science workflows and speeding up machine learning; and TITAN RTX's ability to bring real-time ray tracing to creative applications so PC-based creators can iterate faster and delivering computational horsepower and memory bandwidth for video editing 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.

Media Contacts

Ken Brown
+1 408 486 2626
kebrown@nvidia.com