

## INVESTOR RELATIONS NEWSLETTER



### NVIDIA Turing launch redefines gaming, addresses two new markets

In August NVIDIA unveiled the Turing GPU architecture, the greatest leap since the invention of the CUDA GPU in 2006. Turing – the result of more than 10,000 engineering-years of effort - makes real-time ray tracing possible for the first time and delivers a 40X performance speed-up vs. CPUs for deep learning inference.

Unlike any prior architecture launch, Turing opens opportunities for NVIDIA in 3 markets, 2 of which are new to us:

1. **Redefines computer graphics:** 2X performance increase vs. the prior Pascal generation, with cinematic visuals enabled by ray tracing.
2. **Opens the \$20B inference market:** 40X greater inference performance than CPUs, with a new hyperscale-ready software stack.
3. **Opens the \$4B rendering market:** 30X speed-up for render farms, currently served by an installed base of 1.5M CPUs/\$15B in infrastructure.

We expect the Turing architecture to be a core foundational technology across all our market platforms, including gaming (GeForce RTX), datacenter (T4 universal deep learning accelerator and NVIDIA RTX server), professional visualization (Quadro RTX), and automotive (DRIVE Pegasus autonomous vehicle computer).



### GeForce RTX - the biggest ever generational leap in gaming GPUs

The GeForce RTX series, **announced** in August, includes the first gaming GPUs based on the new NVIDIA Turing™ architecture and the NVIDIA RTX™ platform for real-time ray tracing and all-new AI capabilities. Turing delivers up to 2X the performance of its predecessor, Pascal, and 6X more for ray-traced graphics. These are the biggest generational jumps we have ever delivered to gaming GPUs.

The first two GeForce RTX gaming cards are on shelves now: the GeForce 2080 Ti (starting at \$999) and the GeForce 2080 (\$699). They can deliver 4K HDR gaming at 60 frames per second on even the most advanced AAA titles - a feat even our prior generation flagship GeForce

GTX 1080 Ti GPU couldn't manage. This is quickly becoming the new performance baseline, as 4K displays are now reaching affordable price points at <\$300. In addition, the GeForce 2070 (\$499) is due out in mid-October.

GeForce RTX is coming to market at the perfect time. New games are the No. 1 reason most gamers upgrade, and this coming holiday season is full of highly anticipated blockbuster launches. Two of the biggest upcoming titles – *Battlefield V* and *Shadow of the Tomb Raider* – will ship with NVIDIA RTX technology, enabling lifelike reflections and shadows on the new GeForce RTX GPUs. All in, more than 30 games so far have announced upcoming support for NVIDIA RTX features, which can be enjoyed on the new NVIDIA RTX GPUs.

Real-time ray tracing, or the physical computation of the paths traced by rays of light, has long been considered the “Holy Grail” of computer graphics. It produces lifelike shadows, reflections and refractions that make special effects look real. While offline ray tracing has long been used for special effects in movies, the computational power needed to run it real-time is enormous, so industry watchers didn't expect it to arrive in videogames for another decade.



## Turing T4 GPU with new software stack opens \$20B+ inference opportunity

At GTC Japan in September we [announced](#) the NVIDIA TensorRT Hyperscale Platform featuring the Tesla T4 GPU based on the Turing architecture and a comprehensive set of new inference software. With this launch, NVIDIA is poised to take on the data center inference market, which we estimate will exceed \$20B annually in 5 years.

The data center inference market is currently served by CPUs, with an estimated installed base of over 20 million server nodes across hyperscale datacenters globally. Our inference platform delivers 40X faster performance than CPUs, and with the new TensorRT software stack, it is ideally suited for hyperscale datacenters. As Moore's Law has now ended, we expect that one day, every server node in hyperscale datacenters will need to be accelerated in order to keep up with demand.

There are two critically important aspects of the new TensorRT software stack that can help enable deployment of the NVIDIA inference platform at scale in hyperscale datacenters. First, it enables multiple models and multiple frameworks to run on the same GPU at the same time. This can drive higher datacenter utilization, directly translating to significant savings. Second, it integrates

with Kubernetes, the leading orchestration layer for hyperscale datacenters. Completing our inference platform, the new Tesla T4 GPU delivers 12X the peak inference performance of its P4 predecessor, or 4X in measured applications.

Support for the Tesla T4 GPU and GPU accelerated inference is coming from major server OEMs, hyperscalers and cloud providers worldwide, including Microsoft and Google. Following its launch of NVIDIA's prior generation Tesla P4 processor in August, Google Cloud plans to soon add support for the Tesla T4. T4 will ship in volume in the fourth quarter.



## Turing ray tracing opens the \$4B rendering compute market to GPUs

Also in August, we **announced** three Turing architecture-based GPUs for the professional visualization market: the NVIDIA Quadro RTX 8000, Quadro RTX 6000 and Quadro RTX 5000. In adding ray-tracing and AI capabilities, these GPUs can revolutionize the work of 50 million designers and artists by enabling them to render photorealistic scenes in real time and leverage AI in their workflows. The Quadro RTX series will be initially available in the fourth quarter.

In addition to driving an upgrade cycle in the installed base of workstations and PCs, Turing opens a new market to GPUs: photoreal rendering, or the creation of computer-generated images that look real. Rendering is instrumental to large industries such as media & entertainment (e.g. for visual effects in movies and TV shows), advertising (for commercials and print ads), retail (catalogs), product design, manufacturing, architecture, engineering and construction.

Prior to the launch of Turing and its ray tracing capabilities, GPUs were not able to address this workload, so most rendering today is done on CPU server “render farms.” Because of its tremendous computational intensity, rendering is big business: we estimate an installed base of 1.5M CPUs, or \$15B in compute infrastructure, translating to a \$4B annual TAM.

This market can now be addressed by GPUs for the first time; in fact, our Turing-based GPUs not only support ray tracing, but deliver a 30X speed-up compared to CPU server nodes, slashing rendering times from hours to minutes. With the render farm opportunity in mind, the new NVIDIA RTX Server reference architecture includes 8 Quadro RTX 8000 GPUs. An RTX-accelerated render farm is 1/4 the cost, 1/10 the space, and 1/11 the power of an equivalent performance CPU server render farm. NVIDIA's

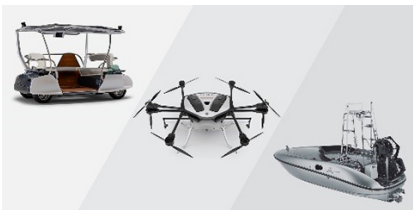
RTX platform has **garnered** major industry support, including from key developers such as Adobe, ANSYS, Autodesk, Dassault, EA, Epic Games, Siemens, Weta Digital and many others.



## Daimler and Bosch select DRIVE Pegasus as AI brain for robotaxi fleets

In July we **announced** that Daimler and Bosch will offer customers a shuttle service with autonomous vehicles based on the NVIDIA DRIVE Pegasus platform in Silicon Valley in the second half of 2019. Pegasus is our AI supercomputer designed specifically for autonomous vehicles, delivering 320 trillions of operations per second. This automotive-grade, functionally safe production solution uses two NVIDIA Xavier SoCs and two Turing GPUs designed for AI and vision processing.

Our collaboration with Daimler and Bosch will unite each company's strengths. NVIDIA brings leadership in AI and self-driving platforms. Bosch, the world's largest tier 1 automotive supplier, brings its hardware and system expertise. Daimler brings total vehicle expertise and a global brand that's synonymous with safety and quality.



## NVIDIA Jetson AGX Xavier autonomous machine platform opens \$5B robotics TAM and \$2B medical instruments TAM

At GTC Japan in September, we also **announced** that our Jetson AGX Xavier platform for autonomous machines will be adopted by leading Japanese companies, including FANUC, Komatsu, Yamaha Motor, Musashi Seimitsu, and Kawada Technologies. For example, Yamaha Motor **plans** to standardize on Jetson AGX Xavier as the brain for a broad portfolio of intelligent mobile machines, including unmanned agriculture vehicles, last-mile vehicles and marine products. We estimate the robotics compute TAM at \$5B by 2022.

The Jetson AGX Xavier platform was built to meet the needs of intelligent devices with large computational requirements, such as robots, medical instruments or drones. It is the first computer of its kind in the world, designed from the ground up. At its core is the Xavier SoC, whose 32 trillion operations per second offer the processing capability of a powerful workstation in the palm of your hand, with operating modes at 10W, 15W and 30W. Jetson AGX Xavier also supports a full software stack to enable companies to quickly develop the next wave of intelligent machines.

One of the most promising applications of Jetson AGX Xavier is medical imaging. Ultrasounds, CT, MR, endoscopy, sequencing and pathology instruments can leverage the platform to deliver supercomputing capabilities for high-speed sensor processing, reconstruction, AI and visualization, upgrading legacy FPGA and CPU-based solutions. We estimate this TAM at \$2B by 2022.

As we expand our business into autonomous machines, we are also expanding our go-to-market reach. We **announced** a partnership with Arrow Electronics, Inc. to bring NVIDIA Jetson AGX Xavier to companies worldwide. The collaboration combines NVIDIA's world-leading AI capabilities with Arrow's global roster of industrial customers and its broad support network of engineers and designers. This opens the door to the development and deployment of AI solutions for manufacturing, logistics, smart cities, healthcare and more.

---

#### Notable Links

[Investor presentation](#)

[NVIDIA #1 on JUST Capital's Top Ten Companies for Workers](#)

[NVIDIA Turing Architecture white paper](#)

[NVIDIA Inference Platform white paper](#)

[NVIDIA Blog: DLSS: What Does It Mean for Game Developers?](#)

[The Street: NVIDIA is Taking Over the Autonomous Driving Industry](#)

[Forbes: Robotics Heavyweights Embrace NVIDIA's Jetson AGX Xavier For AI Edge Intelligence](#)

#### Legal Notices

---

#### FOLLOW US



[Manage Preferences](#) | [Unsubscribe](#) | [Contact Us](#)

© 2018 NVIDIA Corporation. All rights reserved.

NVIDIA Corporation, 2788 San Tomas Expressway, Santa Clara, CA 95051.