

Quote by Devroop Dhar, India CEO & Co-founder, Primus Partners

Published in India Today
July 04, 2026 | 10:34 AM IST

The rise and rise of Nvidia and how it became world's most important company



Authored By Diya Bhati

Read on: <https://www.indiatoday.in/amp/technology/features/story/the-rise-and-rise-of-nvidia-and-how-it-became-worlds-most-important-company-2940373-2026-07-04>

Article Content:

It began with video games, a paintball experiment and a bold bet that few understood. Today, Nvidia has become a company every tech giant depends on to build the future of artificial intelligence.

The large San Jose Convention Center is packed. Thousands of developers, gamers, engineers and technology enthusiasts have gathered for the grand finale of NVIDIA's three-day NVISION 08 conference. They are expecting new graphics technology, faster gaming hardware and a glimpse into what Nvidia has built for the future of computing.

What they don't know is that, in the next few minutes, they are about to witness a demonstration that will give a glimpse of – long before ChatGPT, Gemini or generative AI exist – how Nvidia's technology would one day become the foundation of the AI revolution. Technology that will make it the most valuable company in the world.

On stage, we have television personalities Adam Savage and Jamie Hyneman of MythBusters standing beside two unusual machines. The first machine, nicknamed Leonardo, carries a single paintball gun mounted on a robotic arm. Savage turns to the audience and begins what sounds less like a product launch and more like a science lesson. "The GPU does most of its work in concert with the ultra-important CPU," he says. "We thought we'd show you a comparative demonstration."

The robotic arm slowly moves across a blank canvas, firing one paintball at a time. Each shot represents a single chip calculation, sequentially creating an image similar to how Intel or AMD CPUs function, doing one calculation at a time. "A series of discrete actions performed sequentially, one after the other," Savage explains.

The audience applauds politely and then their attention is drawn towards a machine hidden beneath a giant tarp. This machine is fitted with 1,100 individually addressed paintball barrels. Savage smiles before revealing what it can do. "Each of those paintballs will fly across seven feet of space and, in 80 milliseconds, reach its target," he says.

A countdown begins.

Three.

Two.

One.

With a deafening burst, all 1,100 paintballs fire simultaneously. In less than a tenth of a second, a giant pixelated Mona Lisa appears on the canvas.

"Ladies and gentlemen," Savage declares as the crowd erupts into applause, "science class is now over."

To most people in the audience, it looked like an unforgettable stage spectacle. But Nvidia wasn't demonstrating a paintball machine. It was explaining one of the most important ideas in modern computing – parallel processing.

Unlike a CPU, which works on one task at a time, Nvidia's Graphics Processing Unit (GPU) can work on thousands of tasks simultaneously. It was originally built to make video games look faster and more realistic.

The demo on that afternoon in San Jose was the earliest public glimpse of the computing architecture that would eventually make Nvidia one of the world's most important companies.

Originally, GPU stood for Graphics Processing Unit, a chip built to make games look better by handling thousands of calculations at the same time. But over the years, that graphics processor evolved into something much bigger. It became a general-purpose computing engine capable of solving complex scientific problems, processing massive amounts of data and, eventually, training artificial intelligence models.

But before GPUs become the backbone of AI, let's rewind the clock even further.

The year is 1993

Inside a booth at a Denny's Diner in East San Jose, three engineers: Jensen Huang, Chris Malachowsky, and Curtis Priem were sketching out an idea. Personal computers were beginning to run 3D games, but they lacked the computing power to produce realistic graphics. The trio believed that graphics deserved their own specialised processor rather than relying entirely on the computer's CPU.

That conversation inside the roadside diner became the foundation of Nvidia.

In 1999, the company unveiled the GeForce 256, which it called the world's first Graphics Processing Unit (GPU). Unlike a CPU, which processes one task after another, the GPU could work on thousands of calculations simultaneously. But unlike CPU, which could handle any workload that needed a chip to do calculations, the GPU was fine-tuned for producing only the graphics in a game.

The GPU chip meant richer graphics, smoother gameplay and more immersive virtual worlds for gamers. It was a breakthrough that changed gaming and established Nvidia as one of Silicon Valley's most promising chipmakers. But it wasn't yet the innovation that would matter to anyone outside the gaming world.

That moment came in 2006

That year Nvidia introduced CUDA, or Compute Unified Device Architecture, a software platform that unlocked an entirely new purpose for its GPUs. Instead of using them only to render graphics, developers could now programme these chips to solve complex computing problems that required parallel processing. Not everything worked well on GPU. But some tasks did.

In simple terms, CUDA turned a gaming chip into a tool that scientists, researchers and engineers could use.

Suddenly, Nvidia's GPUs were analysing massive datasets, simulating molecules, forecasting weather, running physics simulations and solving mathematical problems that would have taken traditional processors much longer to complete.

It sounded revolutionary. The reaction, however, was anything but.

"When I launched CUDA, the audience was in complete silence. Nobody wanted it. Nobody asked for it. Nobody understood it," Nvidia CEO Jensen Huang later recalled in an interview with The Closer.

Huang, however, believed he was building for a future the rest of the industry couldn't yet see.

That belief soon paid off.

In 2012, researchers used Nvidia GPUs to train an AI model called AlexNet, which revolutionised image recognition. This proved that GPUs were powerful tools for AI training, not just for gaming, marking a major turning point in the field. Huang realised AI wasn't just another market, it was the future of computing.

Supercomputer to train AI

In 2016, Nvidia doubled down on that belief by launching the DGX-1, a machine purpose-built for training artificial intelligence. It was one of the world's first AI supercomputers, carrying a price tag of around \$129,000.

But once again, only few people understood why anyone would need such an expensive machine dedicated to AI.

Then came a phone call.

It was from Elon Musk, who was helping set up a new non-profit AI research lab called OpenAI. Musk wanted one of Nvidia's DGX-1 systems to accelerate the lab's research. Jensen Huang didn't simply arrange a shipment. He personally loaded one of the first DGX-1 systems into his car and drove it to OpenAI's office in San Francisco.

"Would like to thank @Nvidia and Jensen (Huang) for donating the first DGX-1 AI supercomputer to @OpenAI in support of democratising AI technology," Musk posted on Twitter after receiving the shipment from Huang himself.

Inside that modest office were a handful of researchers, including a young Ilya Sutskever, who would later become one of the key architects behind ChatGPT.

The road which Nvidia built for AI

A decade later, when OpenAI launched ChatGPT, the world suddenly discovered the power of generative AI. But Nvidia wasn't starting from scratch. Its chips were already doing the heavy lifting behind the scenes.

While everyone else was talking about AI, the company had already spent nearly two decades quietly building the hardware, software and developer ecosystem that made the AI boom possible.

Today, almost every major AI model, from OpenAI's ChatGPT and Google's Gemini to Anthropic's Claude and Meta's Llama, has been trained on Nvidia GPUs using CUDA. There are other chip makers, AMD for example. But at least for now Nvidia is ahead of everyone else because it has created the AI ecosystem on which the current AI models have been built.

"The superiority of Nvidia does not only come from offering the best chip in the market," Devroop Dhar, Co-Founder and India CEO of Primus Partners, tells India Today Tech. "During almost two decades, Nvidia has been building a complete AI ecosystem based on its hardware: CUDA, software libraries, developer tools, networking technology and AI framework optimisations. This provided an obvious ecosystem advantage as it is easier to build and cheaper to move to Nvidia's platform compared to anything else."

That ecosystem is Nvidia's biggest competitive advantage.

Atul Arya, Founder and CEO of Blackstraw, concurs. "Nvidia's real advantage isn't just the GPU," he says. "It's the ecosystem it has built over nearly two decades: the software, the developer tools and the

infrastructure that everything else now assumes is there."

In other words, Nvidia didn't just build faster chips. It built the roads that the AI industry now runs on.

The most valuable company on Earth

Today, Nvidia's chips do far more than power gaming PCs.

They train and run large language models like ChatGPT and Gemini, accelerate scientific discoveries, help autonomous vehicles navigate roads, power robots in factories, create digital twins for industries and run some of the world's fastest AI supercomputers.

Or as Sachin Dev Duggal, Founder and CEO of SekondBrain, puts it, "Nvidia is arguably the most important infrastructure company in AI today. Almost every major breakthrough in modern AI has, directly or indirectly, been accelerated by access to Nvidia's hardware, software stack and developer ecosystem."

That infrastructure has become Nvidia's biggest competitive advantage.

"NVIDIA's significance to the AI industry today is difficult to overstate," says Amarjeet Singh Tak, Head of Research and Microscopy Solutions at ZEISS India. "The company has emerged as a key enabler of the AI revolution by providing the computing infrastructure that powers everything from large language models and autonomous systems to scientific discovery and industrial innovation."

But can Nvidia remain king forever?

Probably not.

Industry experts believe Nvidia's dominance won't disappear overnight, but they also agree that no company stays on top forever.

"If Nvidia disappeared tomorrow, innovation would not stop, but it would slow down significantly," says Atul Arya, Founder and CEO of Blackstraw. "Other major players have viable alternatives, but no single player could immediately replace the scale and maturity Nvidia has built."

That competition is already taking shape. AMD and Intel are expanding their AI chip portfolios. Google has developed its in-house Tensor Processing Units (TPUs), Amazon has introduced its Trainium and Inferentia processors, Microsoft is investing in Maia AI chips, while start-ups such as Groq and Cerebras are building specialised AI hardware for inference. Even OpenAI just a few days ago announced that it has built its own specialised AI chip in flat 9 months.

But Duggal suggests that the bigger question isn't who will beat Nvidia. "The question is whether the world should be so dependent on any single supplier for a capability that is becoming foundational to economic growth, national security and scientific progress," he asks.

He believes countries such as India should look beyond building AI models and invest across the entire AI stack to gain an edge in the AI revolution.

As he puts it, "In many ways, the GPU has become the modern equivalent of an oil field. AI is no longer just a software discussion, it is increasingly a geopolitical one."

Ironically, Nvidia never set out to dig for the digital oil. It simply wanted to build better graphics for video games. But even as its popularity surged among gamers over two decades ago, it was beginning to dream of a world beyond video games. And as it worked year after year to make its video cards capable of running more and more cinematic games, it discovered that what can be used to recreate a virtual world can also be used to understand the real world. That was the beginning of CUDA and parallel computing. And rest, as they say, is history.