“It has been a long road from inventing the GPU to accelerate gaming to reinventing the GPU to be the most diverse and powerful coprocessor we have ever seen.”

THE NEXT PLATFORM

NVIDIA pioneered accelerated computing to tackle challenges ordinary computers cannot. We make computers for the da Vincis and Einsteins of our time so that they can see and create the future.

Accelerated computing requires more than just a powerful chip. We achieve incredible speedups through full-stack invention—from the chip and systems to the algorithms and applications they run.
“NVIDIA now enjoys a robust and self-sustaining ecosystem of software, universities, startups, and partners that have enabled it to become the master of its own newly created universe.”

FORBES

It took more than a decade for the total number of developers in the NVIDIA ecosystem to reach one million. Less than two years later, that number has doubled—and stands today at 2.3 million.
NVIDIA REINVENTS MODERN COMPUTER GRAPHICS

We’ve led the field of visual computing for decades.

Our invention of the GPU in 1999 made real-time programmable shading possible, giving artists an infinite palette for expression.

In 2018, our introduction of NVIDIA RTX™ ray-tracing technology fulfilled another vision of computer scientists, paving the way to new levels of art and realism in real-time graphics.

Now, game developers, design tool makers, and film studios—leaders like Pixar, Sony, Epic, Adobe, Microsoft, and others—have wholeheartedly embraced real-time ray tracing.
NVIDIA Omniverse™ brings professional design applications and real-time photorealistic simulation together into a single collaboration environment.

Architects and designers in different locations can iterate on concepts more effectively and present compelling visualizations of 3D models faster.

It's the ultimate collaboration platform for industries with complex creative and design pipelines—a virtual world rendered in stunning RTX.
NVIDIA RTX™ Studio laptops are thin and light, yet powerful enough to be a next-generation gaming platform or a creative design workstation.

With PC partners Asus, Dell, HP, MSI, Razer, and others, we’ve designed and launched more than 70 new models of RTX laptops based on the NVIDIA Ampere architecture.
Gaming is the world’s largest entertainment industry. With more than 200 million gamers, NVIDIA GeForce® is its largest platform. GeForce RTX GPUs and the GeForce Experience™ application transform everyday PCs into powerful gaming machines.

Powered by the NVIDIA Ampere architecture, our 2nd-gen RTX design—GeForce RTX™ 30 Series GPUs deliver incredibly realistic ray-traced graphics and cutting-edge AI features for gamers. And with Max-Q design, 30 Series laptops can deliver twice the power efficiency of previous generations.

“WITH ITS SERIES OF RTX GRAPHICS CARDS, NVIDIA IS WOWING THE GAMING WORLD”

TECH RADAR
RTX has come to the world’s best-selling video game: Minecraft. Before RTX, game developers painstakingly pre-rendered lighting and shadow effects to make their worlds more realistic. But in user-created virtual worlds like Minecraft, only real-time ray tracing can accomplish these beautiful effects.
“GEFORCE NOW SHOWCASES THE LIMITLESS POTENTIAL OF CLOUD GAMING”

PC WORLD

Billions of gamers don’t have GeForce PCs. With our cloud gaming service GeForce NOW™, players can experience the power of a GeForce PC in the cloud, using nearly any device they own.

After nearly 10 years of development and refinement, GeForce NOW is live and hosts six million members around the world. They can instantly access more than 800 games.
In 2006, the creation of our CUDA® programming model and data center GPU platform brought parallel processing to general-purpose computing. A powerful new approach to high performance computing was born.

Today, the universe of supercomputing has expanded rapidly to include AI, advanced data analytics, and cloud computing. The era of the CPU-centered monolithic supercomputer is coming to a close. The next era has begun.
Accelerated computing is the way forward for the world’s most powerful and efficient computers.

Eight of the world’s top 10 supercomputers now use NVIDIA GPUs, InfiniBand networking, or both. NVIDIA powers 346 of the overall TOP500 systems on the latest list.

Selene, NVIDIA’s own supercomputer, is ranked 5th in the world and is the fastest industrial supercomputer.
NVIDIA accelerates more than 700 applications today, including the top 15 in scientific computing. By addressing the entire computing stack, we can drive continuous speed improvements on these applications even without releasing new GPUs.

Over the last four years, we increased performance of core HPC applications by 400%. With the introduction of the NVIDIA Ampere architecture, that improvement goes to 1,100%.

“NVIDIA’S TIMING IS PERFECT. THE WORLD NEEDS ALL THE ADVANCEMENTS IT CAN GET TO SUPPORT HIGH-PERFORMANCE COMPUTING.”

FORBES
Machine learning turns vast pools of data into applications that recommend products, converse with people, and optimize supply chains.

The Apache Spark platform is essential to the process, turning an entire data center into a single computing engine to tackle huge datasets.

NVIDIA now accelerates Spark 3.0, clearing the way for incredible speedups—just as we reach the limits of CPU computing.

“NVIDIA is inventing the future of machine learning right in front of our eyes”

DATA CENTER KNOWLEDGE
NVIDIA is aiming our technology, our superpower, at COVID-19. When every second counts and much is at stake, scientists gear up with NVIDIA to jump to light speed. They are using NVIDIA-accelerated computing to sequence and image the virus, search for new treatments, and build AI robots to disinfect hospitals.
We used to think of a CPU server as the basic unit of computing. But to meet the demands of today’s machine learning and AI workloads, we must optimize the entire data center, from end to end.

It’s an orchestration of three pillars: the CPU for general-purpose computing, the GPU for accelerated computing, and the DPU, which processes and moves data around the data center.

The new unit of computing is the data center itself.
The NVIDIA Ampere GPU architecture unifies the three primary workloads of machine learning: data processing, training, and inference. The NVIDIA A100 data center GPU, the first of the NVIDIA Ampere generation, is the largest processor the world has ever made.

Compared to the previous generation, V100, the NVIDIA A100 has up to 20 times the performance for processing AI neural networks. NVIDIA A100 is in full production, is shipping now, and has won the support of the world’s top cloud providers and server makers.

"NVIDIA'S CHIPS HAVE POWERED NEARLY EVERY MAJOR AI BREAKTHROUGH"

FORBES
NVIDIA DGX™ is a system designed in-house, from the ground up, for the specific purpose of leading-edge AI and data science.

NVIDIA DGX A100™ consists of eight A100 GPUs and nine Mellanox CX6 InfiniBand DPUs. A single DGX system has the equivalent performance of 150 high-end servers costing well over a million dollars. The DGX A100 sells for $199,000 and is in full production and shipping worldwide.

“NVIDIA’S NEW DGX A100 SYSTEM PACKS 5 PETAFLOPS OF PERFORMANCE IN A SINGLE NODE”

TWEAKTOWN
The NVIDIA BlueField-2® DPU is a data center infrastructure on a chip, delivering a broad range of advanced networking, storage, and security services for complex compute and AI workloads.

The NVIDIA DOCA SDK makes it easy for developers to create high-performance, software-defined, cloud-native applications and services for BlueField-2.
NVIDIA and VMware are delivering a new architecture for the modern data center. By bringing VMware, the data center operating system for 70% of the world’s enterprises, to the BlueField-2 DPU, essential enterprise workloads no longer have to share CPU resources with virtualized networking, storage, and security. And to accelerate enterprise adoption of AI, the partnership integrates NVIDIA AI software with the VMware platform.
"IT’S OFFICIAL: NVIDIA IS NOT JUST A GAMING COMPANY ANYMORE"

DATA CENTER KNOWLEDGE

NVIDIA SuperPOD brings together NVIDIA DGX A100 and the advanced networking capabilities of NVIDIA Mellanox HDR InfiniBand networking to create the world’s first turnkey AI infrastructure. SuperPOD makes it possible for organizations to install incredibly powerful AI supercomputers in just a few weeks. It captured the top spot on the November 2020 Green500 list of the most efficient supercomputers and is the platform for Cambridge-1, the UK’s most powerful supercomputer.
“PUT THEM BOTH TOGETHER, AND YOU’VE GOT THE UNDISPUTED KING OF THE SEMICONDUCTOR INDUSTRY”

In September 2020, NVIDIA announced an agreement to acquire Arm for $40 billion, creating the world’s premier computing company. By uniting NVIDIA’s AI capabilities with Arm’s vast ecosystem, we can advance computing from the cloud, smartphones, PCs, self-driving cars, robotics, and edge IoT, and expand AI computing to every corner of the globe.
AI RENAISSANCE

Since 2012, we’ve pushed the boundaries of AI research and development, harnessing accelerated computing to create new breakthroughs and push others to new levels.

Now these capabilities are converging, opening an exciting future for interactive AI applications that will change our lives. But they will require intense computing performance.

Accelerated computing is making that future possible.
NVIDIA accelerates the entire machine learning pipeline—from data loading and processing, to training, to inference. The suite of acceleration libraries is called NVIDIA AI. On top of this, we also target specific applications.

"WE HAVE NOTED TIME AND AGAIN FOR NVIDIA: ITS LEAD DOES NOT JUST LAY IN HARDWARE"

ZDNET
“NVIDIA MAXINE REINVENTS THE VIDEO CALL WITH AI”

TECH HQ

More than 30 million web meetings are estimated to take place every day. NVIDIA Maxine is a cloud-AI video streaming platform that can breathe new life into the video calls that bring us together for work, study, and personal connection.

It vastly improves streaming quality while giving users AI-powered features like super-res, noise cancellation, face alignment, and more.

Maxine is the latest platform brought to life by NVIDIA AI. It joins NVIDIA Jarvis for conversational AI and NVIDIA Merlin for large-scale recommender systems that predict user preferences like products a shopper would buy, movies to watch, or news of interest.
“MICROSOFT OFFICE JOINS SQUARE, TWITTER, EBAY, GE HEALTHCARE, AND ZOOX, AMONG OTHERS, USING NVIDIA GPUS FOR INFERENCE”

<table>
<thead>
<tr>
<th>AUTO</th>
<th>CONSUMER</th>
<th>CSP</th>
<th>ROBOTICS</th>
<th>MEDICAL</th>
<th>RETAIL &amp; FSI</th>
<th>INDUSTRIAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMW</td>
<td>ByteDance</td>
<td>Alibaba Cloud</td>
<td>bossonova</td>
<td>GE Healthcare</td>
<td>CapitalOne</td>
<td>AUO</td>
</tr>
<tr>
<td>BOSCH</td>
<td>criteo</td>
<td>AWS</td>
<td>braincorp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continental</td>
<td>NAVER</td>
<td>Baidu</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ford</td>
<td>Pinterest</td>
<td>Google Cloud</td>
<td>John Deere</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>POSTMATES</td>
<td></td>
<td>MiR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Square</td>
<td>Komatsu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZOOX</td>
<td></td>
<td>Tencent</td>
<td>Toyota</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The smartphone was the first wave of the Internet of Things revolution. Software-defined, imbued with sensors, and connected to the cloud, the smartphone was fertile ground for millions of developers to create apps that are now central to our daily lives.

The next wave has begun. Now almost everything around us can be smart, cloud-connected, and based on a platform on which imaginative services can be built for almost every industry.
“WITH NVIDIA EGX, HOSPITALS, STORES, AND FACTORIES CAN CARRY OUT REAL-TIME PROCESSING OF MASSIVE AMOUNTS OF DATA FROM TRILLIONS OF SENSORS”

INSIDE HPC

AI is spilling out of the cloud and into the edge of the internet, where oceans of data are generated. Sensors connected to AI computers can make seemingly intelligent decisions to speed checkouts, direct forklifts, manage traffic, and save power.

The NVIDIA EGX™ Edge AI platform is bringing AI to the world’s largest industries. Walmart, the US Postal Service, Procter & Gamble, and Samsung Electronics are among the first customers.
The NVIDIA Clara™ AI platform is advancing the science and practice of healthcare. Clara Discovery supercharges the entire drug development process. Clara Imaging helps radiologists develop AI models to detect early signs of disease. And with Clara Guardian, hospitals can analyze smart sensors throughout buildings to spot fevers and other warning signs.

“NVIDIA LAUNCHES CLARA GUARDIAN TO POWER SMART HOSPITALS”
VENTUREBEAT
NVIDIA Jetson™ AGX Xavier delivers the energy-efficient computational power needed for embedded systems like robots, drones, and smart cities. And the new Jetson Xavier NX is enabling millions more small, low-power AI systems for embedded IoT apps. From the edge to the data center, all of NVIDIA’s AI computers run on the same CUDA-X™ AI software stack.

“NVIDIA JETSON NANO IS RESPONSIBLE FOR THE BIGGEST INDUSTRIAL IOT REVOLUTION THESE DAYS”

UPSWIFT
Eventually, all machines that move will be autonomous. NVIDIA Isaac™ is a platform to accelerate the development and deployment of robotics. BMW is partnering with NVIDIA to build its factories of the future. Running NVIDIA Isaac robotics software on NVIDIA EGX, BMW will deploy robot fleets to automate building “the ultimate driving machines.”
One of the most exciting applications for AI is autonomous vehicles. Essentially data centers on wheels, self-driving cars continuously gather sensor data, process it, and make decisions in fractions of a second. NVIDIA DRIVE™ is an open platform for creating software-defined autonomous vehicles. It’s an end-to-end system that includes training AI models, simulating millions of testing miles safely in virtual reality, and deploying autopilots and in-cockpit assistants in next-generation cars and trucks.

“NVIDIA DRIVE IS THE INDUSTRY’S ONLY END-TO-END PLATFORM FROM CLOUD TO CAR AND BACK TO THE CLOUD”

YAHOO FINANCE
"From automakers to research teams and startups, all depend on NVIDIA for hardware and software solutions for self-driving vehicles.

“The company is partnering with automotive companies to drive innovation in simulation software, automotive sensors, cameras and lidars to facilitate driverless transport.”

ZACK’S
Tomorrow’s cars and trucks will be high-performance, updatable computing devices. Leading the way, Mercedes-Benz is working with NVIDIA to bring software-defined vehicles to its fleet starting in 2024. The next-generation Mercedes vehicles will be powered by a first-of-its-kind computing architecture based on the NVIDIA DRIVE platform and backed by a team of AI engineers and powerful AI supercomputers.
NVIDIA CULTURE

NVIDIA is united by a unique culture—the operating system of our company. We dream big, take risks, and learn from our mistakes together. Speed is the key to our success. Craftsmanship is a passion. There are no org charts—the project is the boss.

These beliefs inform everything we do, from designing amazing products to building one of the world’s great companies—a place where people can do their life’s work.
“Huang has risen to the elite among Silicon Valley’s visionary leaders. Scores of reports show NVIDIA employees love working for him and his addresses are often technical yet accessible. He commands an audience through his passion for the technology his company is creating.

“He’s been at the helm of NVIDIA since co-founding the company at age 30 in 1993 and has led NVIDIA from the maker of computer graphics cards to become the premier platform for artificial intelligence and machine learning. This positions NVIDIA at the forefront as the computing industry contemplates a fundamental shift in processing.

“NVIDIA saw it coming.”

TECHCRUNCH
INNOVATIVE COMPANIES LIKE NVIDIA ARE AN AMERICAN TREASURE

THE STREET

NVIDIA has continuously reinvented itself over more than two decades.

Our 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 computing ignited the era of AI.

NVIDIA is a “learning machine” that constantly evolves by adapting to new opportunities that are hard to solve, that only we can tackle, and that matter to the world.
"NVIDIA'S CULTURE IS THE PRODUCT OF A FOUNDER CEO WHO EMBRACES COMMUNITY"

FORTUNE

NVIDIA employees are dedicated to building technology that moves humanity forward and supporting the communities in which they work and live.

We’ve been recognized as a top company in social responsibility, and our employees are passionate donors to hundreds of charities around the globe. In 2020, employees joined the company in contributing more than $16 million to support COVID-19 response efforts.
“Best Places to Work in 2021”
GLASSDOOR

“100 Best Companies to Work For”
FORTUNE

“World’s Best Performing CEO”
HARVARD BUSINESS REVIEW

“Most Innovative Companies”
FAST COMPANY

“50 Smartest Companies”
MIT TECH REVIEW

“World’s Best CEOs”
BARRON'S

Founded in 1993
Jensen Huang, Founder & CEO
19,000 Employees
$16.7B in FY21