2025
NVIDIA Corporation
Annual Review

The Future Runs on Accelerated Computing

We’re entering a new industrial revolution. This time, intelligence is the product. Generative and agentic AI generate, reason, and act. Like electricity, intelligence is becoming a ubiquitous utility. Nations are building AI infrastructure, and NVIDIA powers this transformation. We’re not just advancing AI. We’re building the infrastructure of intelligence, turning computing into a generative force for every industry.

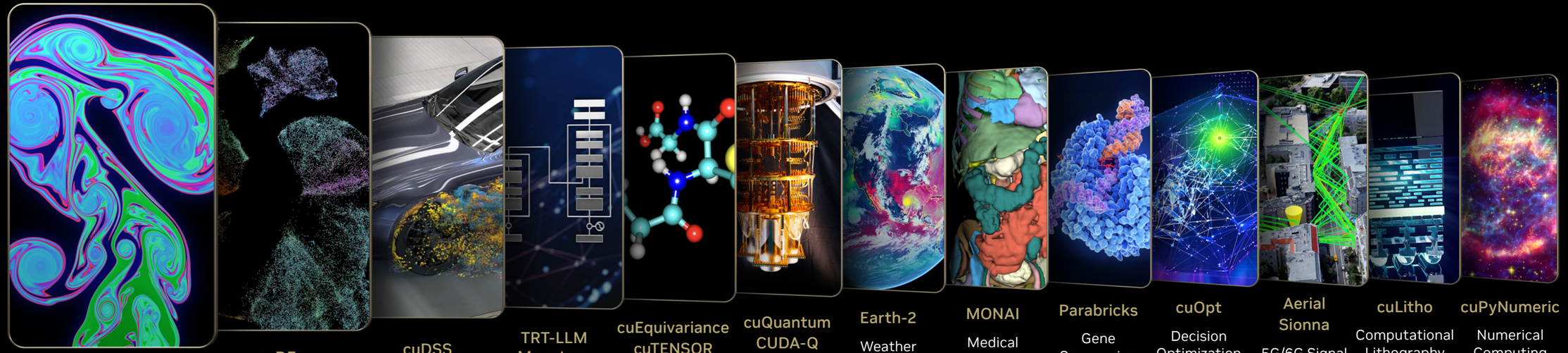
Photograph by: Katie Thompson

The Intelligence Supply Chain Has Arrived

AI factories are a new class of infrastructure, transforming data into tokens of intelligence. Built on the NVIDIA full-stack platform with NVIDIA Blackwell GPUs and NVLink, they deliver unmatched performance and efficiency, powering trillion-dollar industries. Intelligence is no longer stored and retrieved. It's continuously generated by AI factories. NVIDIA is building the global supply chain for this new essential resource.

CUDA-X: Turning Acceleration Into Impact

NVIDIA CUDA-X is the language of accelerated computing. It's a layered software stack of domain-specific libraries, frameworks, and tools built on NVIDIA CUDA. It abstracts complexity and delivers performance across industries for applications spanning materials science, weather modeling, and industrial design. CUDA-X is how accelerated computing becomes accessible, powerful, and practical—transforming raw GPU capability into real-world breakthroughs.



Warp
Physics

cuDF
cuML
Data Science and Processing

cuDSS
cuSPARSE
cuFFT
AmgX
Computer-Aided Engineering

TRT-LLM
Megatron
NCCL
cuDNN
CUTLASS
cuBLAS
Deep Learning

cuEquivariance
cuTENSOR
Quantum Chemistry

cuQuantum
CUDA-Q
Quantum Computing

Earth-2
Weather Analytics

MONAI
Medical Imaging

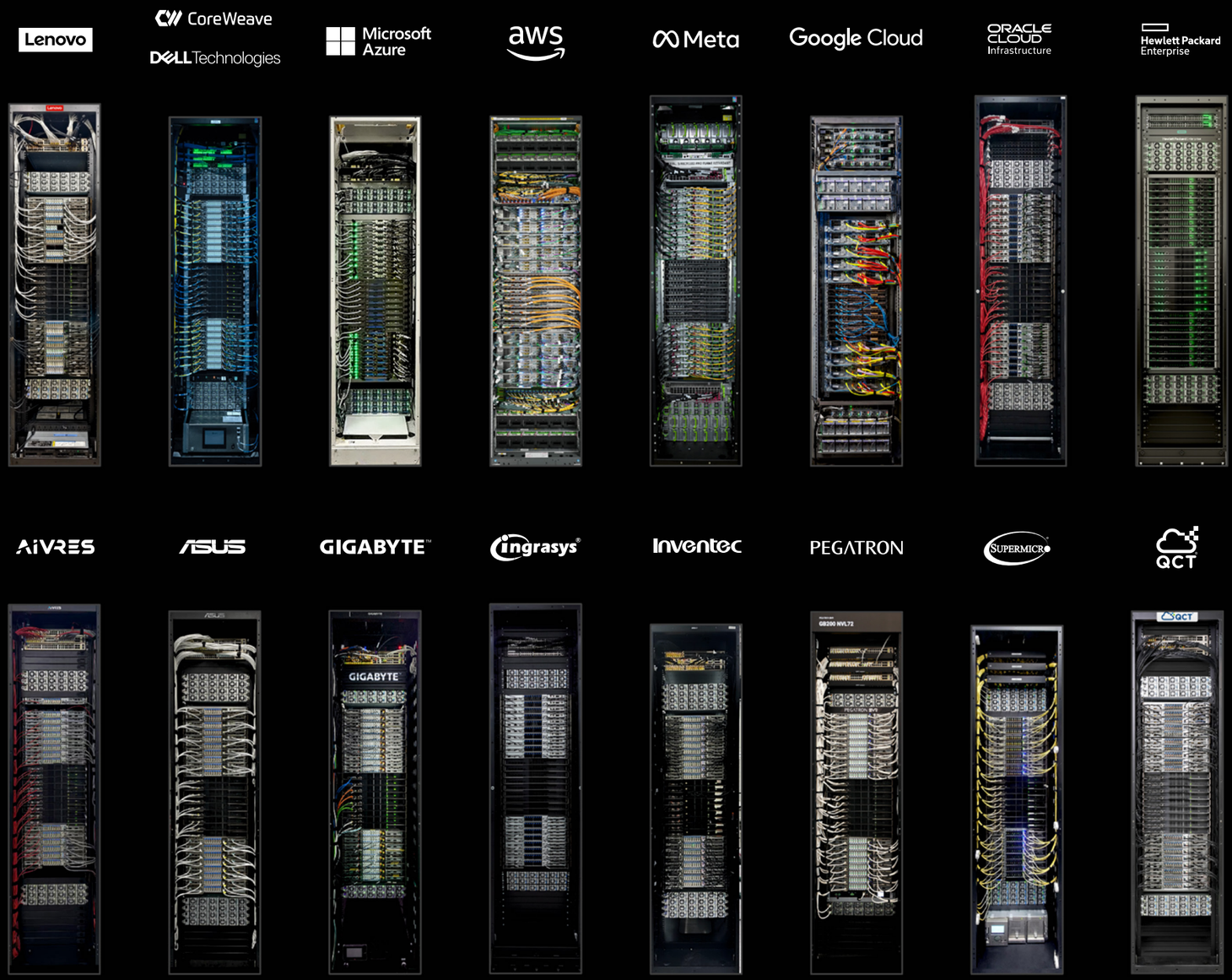
Parabricks
Gene Sequencing

cuOpt
Decision Optimization

Aerial
Sionna
5G/6G Signal Processing

cuLitho
Computational Lithography

cuPyNumeric
Numerical Computing



The AI Superchip Powering the Next Industrial Revolution

Blackwell is NVIDIA's fastest-ramping platform ever—powering the era of AI reasoning. It delivers 30x faster inference with 25x lower cost of ownership, redefining AI economics. NVIDIA Blackwell Ultra boosts AI factory output by 50x over NVIDIA Hopper, accelerating digital biology, robotics, and autonomous systems. With dramatically reduced energy per token, Blackwell makes AI production scalable, efficient, and foundational to the new industrial revolution.

A Blueprint for Tomorrow's Digital Workforce

Agentic AI is the next frontier—intelligent systems that perceive, reason, and act. With NVIDIA NeMo, NIM microservices, and AI Blueprints, businesses can deploy custom AI agents at scale. From ServiceNow to SAP, industry leaders are using agentic AI to transform workflows, preserve knowledge, and drive innovation. These agents reach beyond software into robotics and automation, fueling a trillion-dollar economy.

Transforming $100 Trillion of Industry With Digital Twins and AI

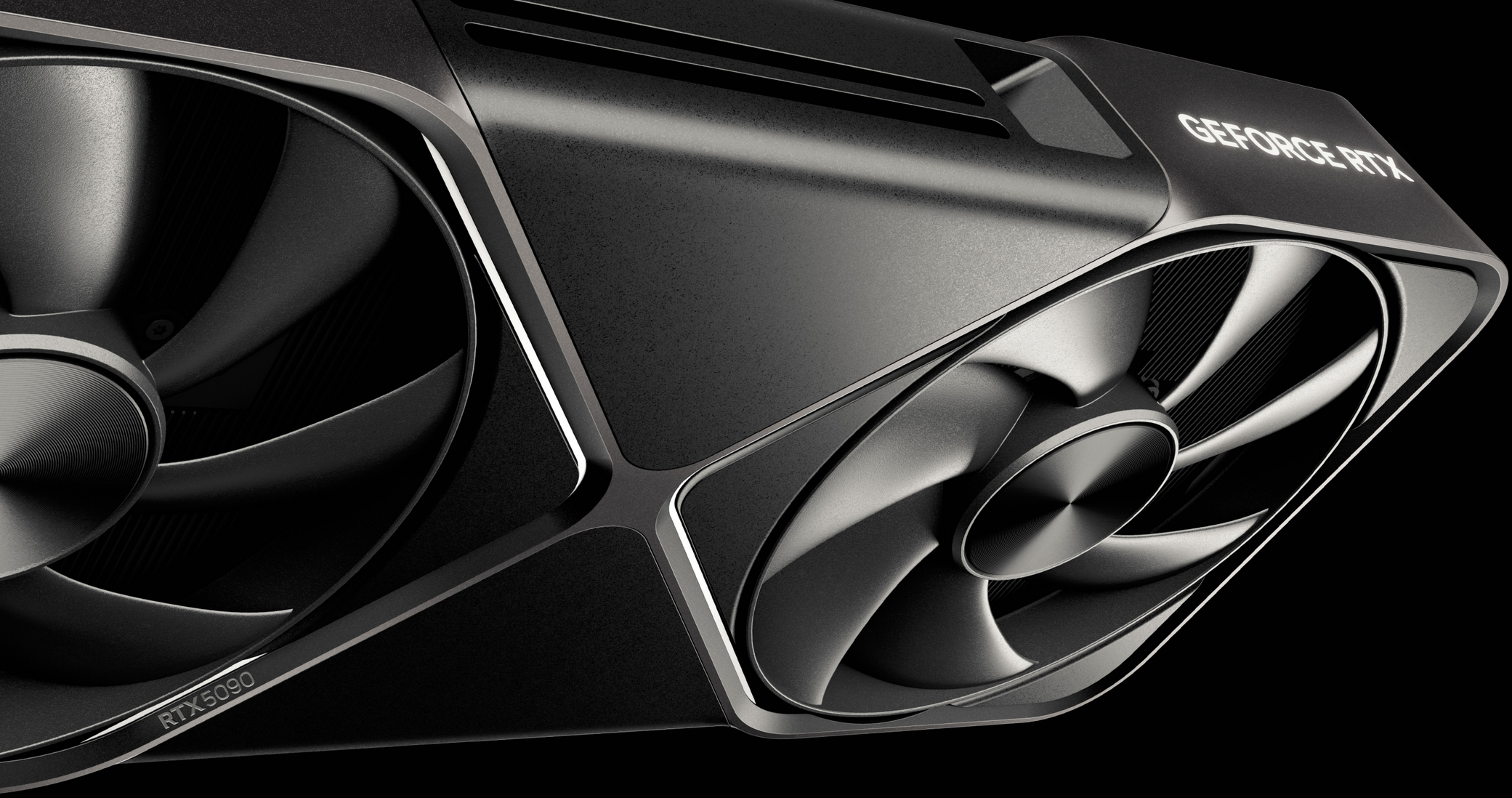
NVIDIA Omniverse is the operating system for industrial digitalization, enabling enterprises to build and optimize digital twins of factories, supply chains, and autonomous machines. Paired with NVIDIA Cosmos—world foundation models trained on real-world physics—teams can simulate, test, and refine intelligent systems before deployment. From Foxconn to Accenture, industry leaders are using Omniverse and Cosmos to accelerate the future of physical AI.

The Robotics Era Has Arrived

Robotics represents a multitrillion-dollar opportunity driven by a global worker shortage and breakthroughs in physical AI and world simulation at scale. NVIDIA powers the world's robot developers with a three computer solution:

1. NVIDIA DGX for training AI models
2. NVIDIA Omniverse and Cosmos for simulation and synthetic data generation
3. NVIDIA Jetson AGX for deployment—and development stacks like NVIDIA Isaac GR00T for humanoid robots, NVIDIA Metropolis for vision AI, and NVIDIA DRIVE for autonomous vehicles

From Figure AI to Universal Robots, and Uber to Mercedes-Benz, robotics leaders are building on NVIDIA technologies.



GeForce RTX 50 Series: AI-Powered Gaming and Creativity

The NVIDIA GeForce RTX 50 Series, built on Blackwell, introduces a new era of AI-powered gaming and creation. With 92 billion transistors and NVIDIA DLSS 4 neural rendering, it delivers up to 8x frame boosts. GeForce RTX 50 Series laptops pack GeForce RTX 4090-class power into ultra-thin designs with 40% better battery life. RTX AI PCs now enable generative AI for gamers, creators, and developers everywhere.

Photograph by: John Shearer

Dear NVIDIANs and Stakeholders,

A new industrial revolution is underway—transforming energy into artificial intelligence.

The factory of this industrial revolution, an AI factory, is powered by computers created by NVIDIA.

For decades, computing advanced through hand-coded software executed on CPUs. That model has reached its limits. Today, computers learn to understand, reason, and solve problems using machine learning algorithms running on GPUs.

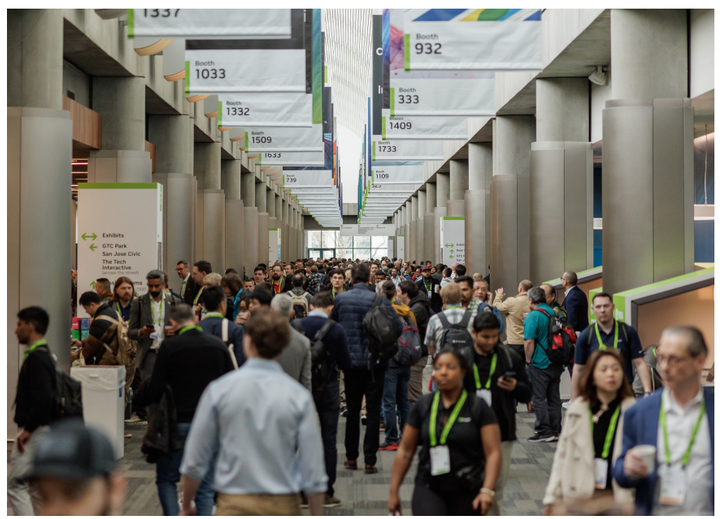
AI is the new model of computing. This transformation is global, sweeping across every market and institution—from digital customer service and drug discovery to financial trading, software development, robotics, and self-driving cars.

Just as bits are the fundamental units of classical computing, tokens are the atomic units of digital intelligence—the symbols through which AI expresses thought. Each token is a fragment of meaning: a word, a piece of code, a strand of DNA. In sequence, they form thoughts, instructions, stories, and reasoning.

NVIDIA computers are the thinking machines that generate these tokens, the currency of the AI age.

A new type of infrastructure has emerged—AI factories. Just as earlier factories forged steel or assembled cars, AI factories generate tokens, transforming data into intelligence. And just as power plants fueled the industrial economy, AI factories will power the intelligence economy.

At NVIDIA, we are building this future—designing the engines of the AI age and creating the factories that will power the next wave of human progress.



GTC 2025

Our annual GTC conference—hailed as the "Woodstock of AI"—is the launchpad of the AI industry. This year, nearly 27,000 attendees gathered in San Jose, with hundreds of thousands more joining online. They came from some 100 countries, more than 10,000 companies, and more than 400 of the world's most promising startups.

GTC 2025 showcased the accelerating frontier of AI: reasoning and inference-time scaling that will drive a 100x increase computing demand; our computing roadmap from NVIDIA Blackwell to Rubin to Feynman; and the rise of the next waves—agentic AI and physical AI—beyond perception and generative AI.

We introduced new platforms and partnerships to bring AI to every enterprise and industry, and unveiled transformative technologies for robotics and autonomous machines.

GTC 2025 made it unmistakenly clear: NVIDIA has transformed from a chip company to a full-stack computing platform, to data center-scale AI systems, to the AI infrastructure company.

Today, companies and countries rely on NVIDIA to build the essential infrastructure of the AI age.

In the decade ahead, trillions of dollars of AI infrastructure will be built, reshaping the $100 trillion global economy.

Financials and Strategic Positioning

This year marked a structural shift in the computing economy. Fiscal 2025 was not a transition, it was a break from the past. AI models moved beyond answering questions to reasoning through complex tasks. Inference workloads surpassed training. And AI infrastructure expanded beyond centralized data centers into factories, vehicles, hospitals, and homes around the world.

This metamorphosis drove exceptional performance.

- Revenue more than doubled to $130.5 billion, up 114% year-over-year
- Operating income rose 147% to $81.5 billion
- Earnings per share grew 147% to $2.94

The fourth quarter capped a defining year:

- $39.3 billion in revenue, up 78% year-over-year and 12% sequentially
- $22.1 billion in net income
- $16.6 billion in operating cash flow

From Training to Reasoning

Data center revenue reached $115.2 billion, up 142% year-over-year.

Q4 alone delivered $35.6 billion, including $11 billion in Blackwell revenue, the fastest product ramp in our company's history.

What began with training is now shifting toward real-time reasoning. Inference has become the dominant workload, and agentic systems are entering production across

sovereign AI initiatives, Fortune 500 enterprises, and edge deployments.

Full Stack, End-to-End AI Infrastructure

The Blackwell platform integrates GPU, CPU, NVLink, Spectrum-X, and NVIDIA software into a unified system, optimized for every stage of AI, from pre-training and post-training to real-time reasoning and inference.

As AI factories become multi-billion-dollar investments, the need for assured performance, reliability, and long-term support is critical. AI continues to advance and change at an incredible pace. With NVIDIA's leadership across every phase of AI, rich ecosystem, programmability, and generations of experience, customers can have complete confidence to scale to the largest infrastructures with us.

We made tremendous progress this year. NVIDIA is now a full-stack AI company—spanning chips, systems, software, and AI models delivered through NVIDIA Inference Microservices (NIMs), a fully operational AI model packaged in a container. To accelerate deployment, we introduced NIMs, AI Blueprints, and the Nemotron model family, enabling enterprises across finance to healthcare to adopt AI faster.

NVIDIA is an end-to-end AI infrastructure platform, from GPUs (Blackwell) and CPUs (Grace) to network processors (CX and Bluefield), scale-up NVLink switches, and scale-out networking with Quantum (InfiniBand) and Spectrum (Ethernet). Each is world-class, highly differentiated, and deeply integrated with NVIDIA software.

We've fundamentally redefined what we build. Instead of assembling AI systems from separately designed components, we now architect and design full-stack, end-to-end AI factories. Blackwell, Rubin, and Feynman aren't just chips, they're AI factories: purpose-built to scale, deeply optimized, and supported across their operational lifetimes.

Though we architect AI factories as a unit, we open the architectures to integrate seamlessly into the diverse environments of our customers, delivered by a global network of system makers across the supply chain. This enables a scale and velocity in deploying these one-and-a-half-ton AI supercomputers the world has never seen before.

Operational Scale, Disciplined Investment

We're investing at pace—with discipline, efficiency, and purpose. Operating expenses rose 45% year-over-year to $16.4 billion, reflecting sustained investment in engineering, compute infrastructure, our global supply chain, developers, and go-to-market partners.

AI demands massive infrastructure scale—more than silicon. It requires a full-stack ecosystem: software, systems, and a global supply chain capable of delivering end-to-end AI factories at incredible speed and scale. We ended the fiscal year with:

- $43.2 billion in cash, cash equivalents, and marketable securities
- $64.1 billion in operating cash flow
- $34.5 billion of cash returned to shareholders, including $33.7 billion in buybacks and $834 million in dividends

Engineering for Continuity

We're operating in a dynamic environment—marked by shifting policies, evolving demand, and rapid technological adoption. Our model is designed for resilience:

- Geographically diversified manufacturing across foundries, substrates, and systems makers
- $30.8 billion in supply commitments to support customer continuity
- Deep partnerships with Amazon, Cisco, Microsoft, Oracle, Google, CoreWeave, and others to scale across markets
- Sovereign AI clusters in Japan, France, India, and Canada
- Enterprise AI partners including Dell, HPE, Accenture, Deloitte, ServiceNow, SAP, and more
- Industrial AI partners including Cadence, Synopsys, Siemens, Foxconn, and more

In the coming years, every company will operate AI factories they built or rent, and every country will build and operate AI factories as part of their critical infrastructure. NVIDIA is their partner to build these foundational deployments, at scale.

This year, for the first time, NVIDIA began manufacturing AI chips and supercomputers in the United States. In partnership with TSMC, Foxconn, Wistron, Amkor, and SPIL, we secured over a million square feet of advanced manufacturing space across Arizona and Texas. These facilities will produce Blackwell chips and assemble full AI systems—bringing AI infrastructure closer to the industries and institutions that depend on it.

Within the next four years, we expect up to half a trillion dollars of AI infrastructure to be built on U.S. soil—strengthening supply chains, creating new jobs, and accelerating our nation's leadership in advanced compute.

From Perception to Physical Intelligence

A decade ago, AI was in its infancy. In 2012, AlexNet ignited the deep learning revolution, demonstrating that GPUs could train neural networks with superhuman accuracy. What began with pattern recognition quickly evolved into machines that could see, hear, and understand.

Perception AI enabled computers to understand the world—powering breakthroughs in speech recognition, medical imaging, and personalized recommendations.

Then came generative AI: AI that creates new content. From text and images to code and proteins, generative AI is a new computing model, shifting from processing data and

retrieving information to generating original content.

Now we're entering the era of agentic AI; AI that can analyze, reason, plan, and use tools to solve problems. These intelligent systems act as digital teammates, automating workflows, accelerating development, and assisting in everything from customer service to healthcare.

Our Nemotron models, purpose-built for reasoning and deployed as NIMs, along with our tools for building custom agents, make it easy for any company to create, train, and deploy agentic AI.

Next is physical AI—intelligent systems that operate in the real world. From autonomous machines in warehouses to robotic assistants in hospitals and self-driving cars, physical AI is built to work safely alongside us.

From One to Three AI Scaling Laws

AI is now governed by three scaling laws—each demonstrating more compute leads to smarter AI:

1. Pre-Training Scaling: The original law: more data, larger models, and more compute produce more capable systems. This remains foundational.

2. Post-Training Scaling: Models continue to improve after initial training. Fine-tuning, reinforcement learning, human feedback, and synthetic data generation drive exponential refinement. At scale, post-training can be even more compute-intensive than pre-training.

3. Inference (Test-Time) Scaling: Inference is no longer a one-shot process, it's real-time reasoning. Models dynamically allocate compute to think through complex tasks, explore alternatives, and plan actions. Agentic AI depends on this capability. Reasoning or "thinking" can require 100x more compute than one-shot inference.

With hundreds of millions of users engaging AI, and agentic AI systems operating continuously in the background, inference has become the most compute-intensive phase of AI.

Every interaction with a computer will trigger AI inference. And as agentic AI begins preempting tasks and responding to complex requests, inference workloads will scale dramatically—driving massive deployments of AI factories.

Enter Blackwell

Blackwell isn't just faster. It's the first architecture purpose-built for every stage of AI, from training to reasoning at scale. With FP4 precision, long-context processing, and dynamic compute allocation, it makes test-time scaling real time as well as energy and cost efficient. Now in full production, Blackwell is NVIDIA's fastest product ramp ever—powering AI infrastructure across hyperscalers, enterprises, and sovereign clouds.

At its core is the NVIDIA GB200 NVL72: 72 GPUs, 36 Grace CPUs, unified by NVLink, Spectrum-X Ethernet, and liquid or air cooling. It delivers up to 30x faster inference for trillion-parameter models like GPT-4 or Grok-3—at the lowest total cost per watt, per rack, and per dollar invested.

A Full-Stack Platform for Scaling Intelligence

Blackwell is more than silicon. Its supporting technologies are foundational to the next wave of AI factories:

- NVLink 5 and NVSwitch connect 72 Blackwell GPUs into a massive compute resource enabling exascale performance for AI factories.
- Spectrum-X extends networking capabilities to support scale-out AI factories with hundreds of thousands of GPUs.
- Silicon photonics unlocks efficient multi-rack GPU scaling—laying the foundation for ultra-large-scale AI factories with millions of GPUs.

These advances make Blackwell the compute engine behind AI's next industrial phase.

The Operating System of AI Factories

Announced at GTC, NVIDIA Dynamo is the inference operating system of the AI factory. Dynamo smartly schedules incoming queries and orchestrates compute resources to minimize response time while maximizing factory throughput. In DeepSeek-R1, Dynamo boosted inference throughput by 30x, cutting the cost of long-context reasoning dramatically.

The world's top cloud providers—Amazon, Microsoft, Google, and Oracle—have increased their investment in Blackwell, with early orders already exceeding Hopper's peak volumes.

What's Next: Blackwell Ultra and Rubin

Later this year, Blackwell Ultra arrives—boosting memory bandwidth, energy efficiency, and long-context support for agentic and multimodal models. It's purpose-built for real-time planning, decision-making, and AI-native applications. Compared to Hopper, it delivers up to 50x more AI factory value.

Rubin, our next-generation platform, is built for a world defined by trillion-parameter agentic AI models. Blackwell today. Blackwell Ultra tomorrow. Rubin on the horizon.

And Feynman beyond. Each generation pushes technology to the limits of physics, increasing AI factory throughput while reducing cost per token generated. And every generation is architecturally compatible, preserving the massive software investments made by the global ecosystem of AI researchers and software developers building on the NVIDIA platform.

NVIDIA is the engine powering the world's AI infrastructure.

By providing a multi-year view, we enable the world's vast supply chain—upstream and downstream—to prepare, align, and innovate alongside NVIDIA as we push technology to the limits of physics.

CUDA-X Accelerates Industries

The AI revolution is powered as much by software as by hardware. CUDA-X, our suite of over 400 acceleration libraries, bridges NVIDIA GPUs with real-world applications, delivering exponential performance gains across AI, supercomputing, and data analytics. Optimized for highly efficient parallel processing, these libraries unlock up to 200x acceleration, enabling massive-scale training, real-time inference, and high-throughput data processing.

Key libraries include:

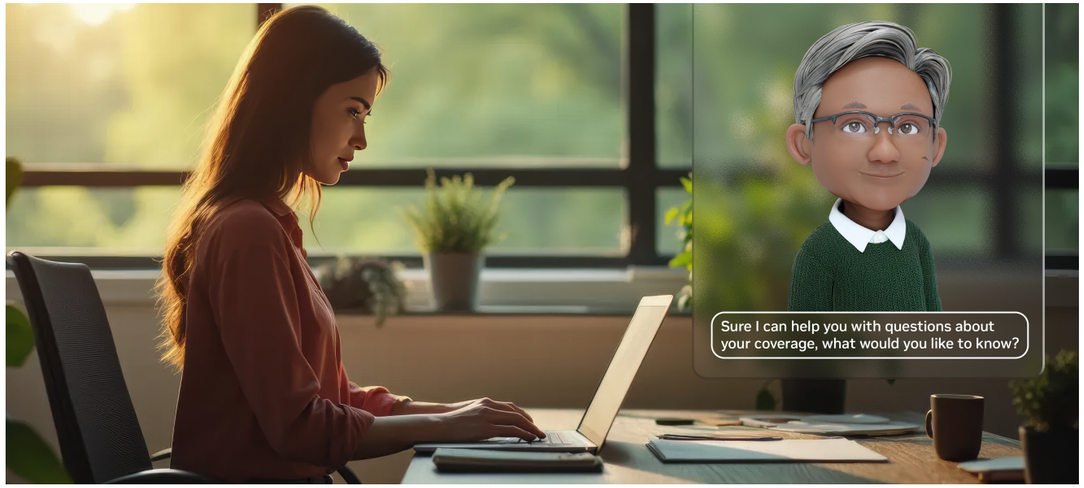
- **NVIDIA cuDNN** is the backbone of deep learning training and inference
- **TensorRT** and **TensorRT-LLM** deliver accelerated inference for LLMs
- **RAPIDS** accelerates data analytics and ML pipelines by up to 50x
- **NVIDIA Parabricks** reduces genome sequencing from days to minutes
- **NVIDIA cuQuantum** accelerates quantum circuit simulation and lays the foundation for hybrid computing, where every data center will one day have a QPU alongside a GPU

Across sectors—from healthcare to finance to robotics—CUDA-X is the software foundation of modern AI computing. With over 6 million developers and 40,000 companies building on our platforms, the CUDA ecosystem has become one of the largest technical communities in the world, driving a flywheel of innovation, platform adoption, and long-term value creation. CUDA doesn't just accelerate compute—it accelerates progress.

A Digital Workforce, Working Alongside Humans

For decades, computing meant people using software to do work. We are now entering a new era where software does the work itself.

Agentic AI goes far beyond chatbots. These systems understand the task, interface directly with business systems, read and understand retrieved information, and use tools and automate complex work—operating like digital employees, but with a speed, scale, and consistency no human workforce can match.



Enterprises can hire or build an AI workforce. Most will do both.

AI agents are being deployed across customer service, engineering, finance, legal, marketing, and operations. Some are generalist business operations AI agents. Others are highly specialized in software coding, supply chain optimization, or chip design. They don't just assist. They learn, codify, and preserve institutional knowledge, ensuring continuity across generations.

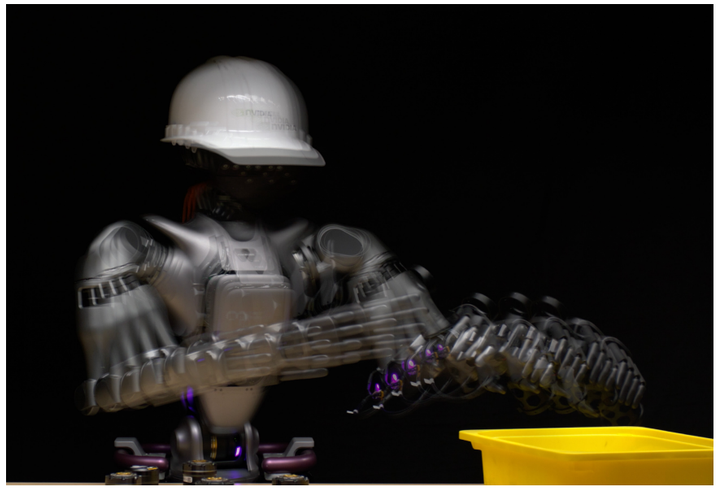
Enterprise IT is becoming the HR function for AI agents—onboarding, training, and managing the digital workforce. Agents are given roles, evaluated on performance, and continuously improved.

Supporting this evolution requires a new kind of infrastructure. Agentic AI isn't powered by a single model—it runs on systems of models working together, in real time. NVIDIA provides the full-stack foundation:

- **NeMo** to train and fine-tune agents on proprietary enterprise data
- **NeMo Guardrails** to align behavior, enforce policy, and ensure safety
- **NIMs**—AI models in a box—deployable as microservices across cloud, on-prem, and edge
- **Agent Blueprints** to launch ready-made, domain-specific agents at speed

This global rollout is supported by consulting and systems integration leaders like Accenture, Deloitte, and Infosys, with thousands of trained professionals helping enterprises build agentic AI on the NVIDIA platform.

Soon, billions of digital agents will work alongside humans, reshaping how businesses operate. This is more than a technical shift—it's an economic transformation.



AI Enters the Physical World

The next wave of AI is here: physical AI—robots and autonomous machines that can perceive, reason, and act in the real world.

Physical AI is transforming warehouses, factories, hospitals, and mobility itself.

Everything that moves—cars, forklifts, surgical arms, warehouse bots—is becoming autonomous. Industry leaders are already deploying these systems:

- GM, Toyota, and Mercedes-Benz are building autonomous vehicles with NVIDIA DRIVE
- Amazon is automating logistics with NVIDIA-powered robotics
- Johnson & Johnson is building NVIDIA AI-powered systems to assist surgery

Humanoid robots that can operate and work in environments made for humans are making rapid advances. At GTC, we announced NVIDIA Isaac GR00T N1—the world's first open humanoid foundation model, trained in Omniverse by observing human demonstrations. Companies like Boston Dynamics, Figure AI, and Unitree Robotics are already building on it.

Train. Simulate. Deploy.

Bringing AI into the physical world requires more than intelligent machines. It demands an integrated pipeline supported by three computing platforms:

- **Train** foundation models on NVIDIA AI supercomputers with massive, real-world datasets
- **Simulate** skills in Omniverse, our physically accurate digital twin and simulation platform
- **Deploy** AI models on Jetson Thor, an AI computer purpose-built for robotics

This closed loop shortens development, improves safety, and accelerates real-world deployment.

Understanding the Physical World

To operate in the physical world, AI must understand it. That's why we built Cosmos, a World Foundation Model for robotics and physical AI.

Like language models learn grammar, Cosmos learns physics— gravity, friction, motion, object permanence, and cause and effect.

Companies like Uber, Agile Robots, Neura Robotics, and many others are already using Cosmos to build the next generation of intelligent machines.

AI in Motion

Robots are no longer science fiction. Autonomous machines are no longer prototypes.

Physical AI represents a $50 trillion opportunity. It will redefine how work gets done—from assembly lines to city streets.

Physical AI is here. And NVIDIA is powering it.

AI Comes Home: GeForce and Desktop Computing

The AI revolution began with a GeForce GPU in a PC. It's expanded to the world's largest data centers. Now, it's coming full circle—back to the PC.

When researchers first turned to GPUs to train neural networks, it ignited a wave that redefined computing. With the launch of the GeForce RTX 50 Series, that wave has returned to where it started—bringing AI and graphics together on a single platform.

Powered by the Blackwell architecture, RTX 50 delivers up to 2x the performance of the previous generation. But more importantly, it changes what a PC can do.

AI is no longer enhancing graphics—it *is* graphics.

- With DLSS 4 and neural rendering, RTX GPUs don't just draw pixels—they predict and generate entire frames
- The result: up to 8x faster performance, real-time cinematic lighting, and dynamic realism
- NVIDIA Max-Q optimizations boost battery life by up to 40%—enabling portable, AI-native workflows for creators, developers, and researchers

Lightspeed Ahead

This year, NVIDIA went through a fundamental transformation from a chip maker to a builder of infrastructure—building AI factories that make up the infrastructure of the AI industry. The transformation is clear, and we're still at the beginning.

Just as traditional computing became essential to every company, industry, and country, so too will AI. But AI demands more than yesterday's data centers. It requires a new kind—AI factories—that produce intelligence at scale. And just as intelligence is foundational to every human endeavor, AI factories will become essential infrastructure.

In the decade ahead, AI will evolve—from tools to teammates, from assistants to the digital workforce of companies, from software to physically intelligent robots that move through factories, hospitals, farms, and cities—alongside us, helping us, filling tens of millions of unfilled jobs and offsetting labor shortages around the world. Economies will grow. Productivity will increase.

A new era of computing has begun. A new industry is born, perhaps the largest of all.

At NVIDIA, we're not merely anticipating this future. We're building it.

To our employees around the world, thank you. You are building a once-in-a-generation company, driven by a once-in-a-lifetime opportunity to reinvent an industry we've devoted our life's work to.

We carry this responsibility with deep humility and unwavering commitment to the journey ahead.

Jensen Huang
CEO and Founder, NVIDIA
May 2025

NVIDIA Corporation

Notice of 2025 Annual Meeting
Proxy Statement and Form 10-K

Forward-Looking Statements

Certain statements in this document including, but not limited to, statements as to: the impact, benefits, abilities, features, performance, and availability of our products, services, and technologies; third parties using or adopting NVIDIA products, technologies and platforms, and the benefits and impacts thereof; our partnerships and collaborations with third parties, and the benefits and impacts thereof; AI factories powering the intelligence economy; NVIDIA designing the engines of the AI age and creating the factories that will power the next wave of human progress; in the decade ahead, trillions of dollars of AI infrastructure being built, reshaping the 100 trillion dollar global industries and economy; agentic systems entering production across sovereign AI initiatives, Fortune 500 enterprises, and edge deployments; AI continuing to advance and change at an incredible pace; with NVIDIA's leadership across every phase of AI, rich ecosystem, programmability, generations of experience, customers being able to have complete confidence to scale to the largest infrastructures with us; we enabling a scale and velocity in deploying these one-and-a-half ton AI supercomputers the world has never seen before; in the coming years, every company operating AI factories they built or rent, and every country building and operating AI factories as part of their critical infrastructure; facilities we secured producing Blackwell chips and assembling full AI systems-bringing AI infrastructure closer to the industries and institutions that depend on it; within the next four years, we expecting up to half a trillion dollars of AI infrastructure to be built on U.S. soil-strengthening supply chains, creating new jobs, and accelerating our nation's leadership in advanced compute; reasoning or "thinking" requiring 100x more compute than one-shot inference; every interaction with a computer triggering AI inference; as agentic AI begins preempting tasks and responding to complex requests, inference workloads scaling dramatically – driving massive deployments of AI factories; Blackwell powering AI infrastructure across hyperscalers, enterprises, and sovereign clouds; we enabling the world's vast supply chain-upstream and downstream-to prepare, align, and innovate alongside NVIDIA as we push technology at the limits of physics; enterprises being able to hire or build an AI workforce; AI agents being deployed across customer service, engineering, finance, legal, marketing, and operations; billions of digital agents working alongside humans, reshaping how businesses operate; everything that moves-cars, forklifts, surgical arms, warehouse bots-becoming autonomous; humanoid robots that can operate and work in environments made for humans making rapid advances; physical AI redefining how work gets done-from assembly lines to city streets; NVIDIA powering physical AI; GeForce GPU coming full circle-back to the PC; AI factories becoming essential infrastructure; in the decade ahead, AI evolving-from tools to teammates, from assistants to the digital workforce of companies, from software to physically intelligent robots that move through factories, hospitals, farms, and cities-alongside us, helping us, making up for the tens of millions of unfilled jobs and offsetting labor shortages around the world; economies growing; productivity increasing; our employees building a once-in-a-generation company, driven by a once-in-a-lifetime opportunity to reinvent an industry we've devoted our life's work are forward-looking statements within the meaning of Section 27A of the Securities Act of 1933, as amended, and Section 21E of the Securities Exchange Act of 1934, as amended, which are subject to the "safe harbor" created by those sections 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.

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, California 95051 | www.nvidia.com

© 2025 NVIDIA Corporation. All rights reserved.