

Nebius brings compute pricing into the open as Europe's AI builders scale on its cloud

- Introduces public live prices for GPU spot capacity to expand transparent and flexible access to AI compute
- Expands AI Builder Program to Europe with NVIDIA to accelerate European AI development
- Across countries and industries, Europe's most ambitious AI companies are building on Nebius

BERLIN — NVIDIA GTC, October 21, 2026 — Nebius (Nasdaq: NBIS), the AI cloud company, today announced public live spot pricing for GPU compute, giving teams greater visibility and control over their compute costs.

AI teams are increasingly experimenting, training and running production services at the same time. Each calls for a different way to access and use compute capacity. Nebius is expanding its portfolio of pricing mechanisms to meet these needs, as part of its vision to promote an open market for compute.

Alongside live preemptible spot pricing, reserve auctions offer defined capacity blocks for short-term compute needs, while pay-as-you-go enables access to compute on demand without a fixed-term reservation. Nebius is also piloting a capacity-rebate program that enables customers to recover value by releasing idle GPUs back into the spot capacity pool.

Arkady Volozh, founder and CEO of Nebius, said:

“Companies should be able to see what compute costs and access it on terms that fit their work. We are combining that flexibility with the infrastructure, software and engineering expertise that AI builders in Europe and around the world need to turn ideas into products that can compete globally.”

Nico Caprez, VP AI Infrastructure Growth at NVIDIA, said:

“NVIDIA AI infrastructure is productive, durable and fungible across models, workloads and customers. Transparent pricing and flexible access help AI companies put that infrastructure to work on terms that fit their needs. Nebius is helping connect available capacity with growing AI demand, keeping AI factories productive and giving Europe's AI builders more ways to scale.”

Nebius is working with the ecosystem to make live pricing more widely available. Public live pricing for Spot PVMs is available now at nebius.com/spot-prices and on [Shadeform](#), [Silicon Data](#) and [GPU Finder](#).

Europe's most ambitious AI leaders build on Nebius

Nebius is bringing AI compute online across a European capacity footprint that currently spans Finland, France, the UK, Spain, Estonia and Iceland, with a pipeline of potential future deployments under evaluation.

Across the continent, digital enterprises and AI natives are building models and products with global ambition on Nebius — from personal finance and e-commerce platforms to understanding disease, developing world models, and deploying industrial robots.

In Germany, Black Forest Labs used Nebius's NVIDIA B200 infrastructure for the research behind Self-Flow, a training method that became foundational to its FLUX 3 models.

Sandeep Mandala, Head of Engineering, Black Forest Labs:

"Training frontier visual intelligence models depends on infrastructure that holds up under sustained load. Nebius has been a dependable part of that foundation for Black Forest Labs."

Nebius is also strengthening its presence in Germany as a compute supplier for the Next Frontier AI Challenge of SPRIND, the Federal Agency for Breakthrough Innovation.

In France, Advanced Machine Intelligence uses Nebius compute and storage for research into world models. Paris-based Photoroom uses Nebius infrastructure and engineering support to develop commercial image-editing models.

Alex LeBrun, co-founder and CEO of Advanced Machine Intelligence:

"We are building AI that can understand the physical world, anticipate the consequences of actions and plan. That ambition requires infrastructure that lets researchers test new approaches at scale. Nebius provides the compute and storage behind our research, helping us develop world models with potential applications across industry, robotics and healthcare."

In the UK, Prima Mente has trained its biological foundation models on Nebius to discover new biomarkers for early detection and new hypotheses for novel drug targets for Alzheimer's and other neurodegenerative diseases.

Hannah Madan, co-founder and COO of Prima Mente:

"The future of brain care we're building is one where our parents have impactful treatments, those in our generation can detect disease earlier and do something about it, and our children never have to worry."

In the Netherlands, Prosus uses Nebius Token Factory for high-volume open-model workloads supporting its lifestyle and e-commerce businesses.

Euro Beinat, Chief AI Officer, Prosus:

“Our AI systems need to be reliable and economical in production. With Nebius Token Factory, we have doubled throughput, while autoscaling helps us handle peaks without holding GPU capacity in reserve. Combined with Nebius AI Cloud for training and fine-tuning, this gives us better economics and room to scale.”

Neobank bunq’s agentic assistant Finn now runs on open weight models. Powered by NVIDIA accelerated computing, deploying on Nebius matches the quality of leading closed models at 60% of the cost per answer.

In Switzerland, Zurich-based Jua trained and refined its Earth Physics Transformer on Nebius to deliver hourly global weather forecasts for energy markets.

European scientific AI and healthcare pioneers running on Nebius include Portugal-founded Sword Health, which worked with Nebius engineers to scale Dawn, its mental health AI conversational assistant, on dedicated NVIDIA Blackwell infrastructure.

Basecamp Research trained EDEN on Nebius with 28 billion parameters and 9.7 trillion nucleotide tokens to design programmable medicines. Inherent Labs used Nebius infrastructure to post-train Faraday, an AI scientist agent that learns to replicate published research as a step towards more autonomous scientific discovery.

European robotics and physical AI companies including Sereact and NormaCore are leveraging Nebius to run the full physical AI development cycle. Nebius today announced the [addition of agentic workflows](#) to its [Physical AI Workbench](#), with customers reporting reductions in engineering time of up to 40%.

Empowering Europe’s AI builders to compete globally

From its headquarters in Amsterdam, Nebius supports the entire European AI ecosystem.

The Nebius AI Builder Program [is expanding across Europe](#) with an additional \$10 million in compute grants, as well as founder funding opportunities and hands-on community programs.

Nebius works with communities including [{Tech: Europe}](#) and [AI Tinkerers](#) to connect builders through hack days, demo days and technical training. For GTC, [Nebius Berlin Build Week](#) culminated in a hackathon with NVIDIA, using [NVIDIA Nemotron](#) open models on Nebius infrastructure.

The company’s education and research arm, [Nebius Academy](#), is collaborating with [NVIDIA Deep Learning Institute](#) to [launch an education collaboration at GTC Berlin](#) to expand hands-on AI training for developers. Nebius also [provides compute grants](#) to support cutting-edge research in AI and AI-driven science.

Building the cloud for production AI with NVIDIA

Nebius AI Cloud is designed and built as one integrated system, combining NVIDIA accelerated computing and networking with Nebius custom servers and software, including ML native tooling and managed inference.

A fast-growing partner ecosystem gives customers more ways to integrate Nebius into their existing stack, deploy and manage production AI applications, and purchase through trusted technology, services, and channel providers.

Nebius and NVIDIA are collaborating to scale the next generation of hyperscale AI cloud across the full AI technology stack — from AI factory design and deployment to co-building the inference and agentic stack and taking customers into production. This collaboration supports early adoption of new computing architectures and optimisation across hardware and software.

Nebius's AI infrastructure leverages successive generations of NVIDIA accelerated computing and reference architectures including the NVIDIA DSX™ AI Factory Platform.

Nebius was among the first AI clouds to bring up and validate NVIDIA Vera Rubin NVL72 rack-scale systems, delivering inference with up to 30x more throughput per watt, and up to 45x lower token cost and training with one-fourth the GPUs.

It is also the first AI cloud to adopt NVIDIA Groq 3 LPX and will be an early deployer of NVIDIA Vera CPUs for agentic workloads. NVIDIA Groq 3 LPX is the inference accelerator designed for the low latency and large context demands of agentic systems. NVIDIA Vera is the CPU for agents, delivering 80% faster agentic CPU performance for longer-thinking agents, faster-learning RL systems, and data-driven AI factories.

The companies' engineers work together to optimize inference across hardware and software, translating new architectures into greater capacity for production AI. This co-engineering delivers performance that is validated by independent evaluations and benchmarks.

Nebius holds NVIDIA Exemplar Cloud recognition for training workloads on NVIDIA HGX B300, NVIDIA GB300 NVL72 and NVIDIA HGX H200, and Platinum status in SemiAnalysis's ClusterMAX rating of AI clouds. Nebius regularly achieves top MLPerf results across training, inference and storage.

Nebius supports an approach to AI safety, outlined in the NVIDIA OpenShell Reference Architecture, which combines software and hardware safeguards with clear boundaries around what agents can access and do, continuous monitoring, and the ability to interrupt, correct or stop their actions when needed.

About Nebius

Nebius, the AI cloud company, is building the full-stack platform for developers and companies to take charge of their AI future — from data and model training to production deployment. Founded on deep in-house technological expertise and operating at scale with a rapidly expanding global footprint, Nebius serves startups and enterprises building AI products, agents, and services worldwide. Nebius is listed on Nasdaq (Nasdaq: NBIS) and headquartered in Amsterdam.

For more information please visit nebius.com

Media kit nebius.com/media-kit

Contacts

Media relations

media@nebius.com

Investor relations

askIR@nebius.com

Forward-looking statements

This press release contains forward-looking statements within the meaning of the Private Securities Litigation Reform Act of 1995, which involve risks and uncertainties. All statements contained in this press release other than statements of historical fact, including, without limitation, statements regarding our future financial and business performance, strategy, expected growth, planned investments and capital expenditures, capacity expansion plans, anticipated future financing transactions and expected financial results, are forward-looking statements. The words "anticipate," "believe," "continue," "estimate," "expect," "guide," "intend," "likely," "may," "will" and similar expressions and their negatives are intended to identify forward-looking statements.

These forward-looking statements are subject to risks, uncertainties and assumptions, some of which are beyond our control. Actual results may differ materially from the results predicted or implied by such statements, and our reported results should not be considered as an indication of future performance. The potential risks and uncertainties that could cause actual results to differ from the results predicted or implied by such statements include, among others: market, macroeconomic and geopolitical conditions; our ability to build, operate and manage our businesses to the desired scale; competitive pressures; technological developments; our ability to secure and retain clients; our ability to secure additional capital to enable the growth of the business; unpredictable sales cycles; and potential pricing pressures; as well as those risks and uncertainties related to our continuing businesses included under the captions "Risk Factors" and "Operating and Financial Review and Prospects" in our Annual Report on Form 20-F for the year ended December 31, 2025, filed with the SEC on April 30, 2026, which is available on our investor relations website at <https://nebius.com/investor-hub> and on the SEC website at www.sec.gov.

All information in this press release is as of the date hereof (unless stated otherwise). Except as required by law, we undertake no obligation to update or revise publicly any forward-looking statements, whether as a result of new information, future events or otherwise, after the date on which the statements are made or to reflect the occurrence of unanticipated events.

In addition, statements that "we believe" and similar statements reflect our beliefs and opinions on the relevant subject. These statements are based upon information available to us as of the date hereof and, while we believe such information forms a reasonable basis for such statements, such information may be limited or incomplete, and our statements should not be read to indicate that we have conducted an exhaustive inquiry into, or review of, all potentially available relevant information. These statements are inherently uncertain, and investors are cautioned not to unduly rely upon these statements.