



CLICK HERE

The AI boom has sparked a new kind of arms race—not just in algorithms and models, but in the raw computing power needed to run them. This week, Microsoft took a bold step to secure its AI future by signing a **\$17.4 billion, multi-year agreement** with Amsterdam-based [Nebius](#), a fast-growing AI infrastructure company.

The deal ensures Microsoft gets guaranteed access to GPU-powered infrastructure, while Nebius cements itself as one of the most important new players in the global cloud ecosystem.

Key Highlights

- Microsoft signs a **five-year, \$17.4 billion deal** with Nebius for dedicated AI cloud infrastructure.
- The agreement could grow to **\$19.4 billion** if Microsoft expands usage through 2031.
- Services will be delivered from Nebius's **new 300 MW Vineland, New Jersey data center** starting later in 2025.
- Nebius shares soared **over 50%** after the announcement, reflecting strong investor confidence.
- The deal positions Nebius as Microsoft's largest infrastructure partner and accelerates its global growth plans.

Who is Nebius?

Nebius may not be as well-known as Amazon Web Services or Google Cloud, but it is quickly building a reputation as a **specialist in AI-native cloud platforms**. Originally spun out from Yandex, the company is now headquartered in Amsterdam and listed on Nasdaq (ticker: NBIS).

Its technology stack is built specifically for **AI workloads**—from training large models to fine-tuning and running applications at scale. Nebius combines proprietary hardware and software to deliver the compute, storage, and managed services AI developers need.

Beyond cloud infrastructure, Nebius has stakes in multiple ventures:

- **Avride**, a self-driving technology developer.
- **TripleTen**, an edtech firm reskilling people for tech careers.
- Investments in **ClickHouse** and **Toloka**.

But the Microsoft deal firmly shifts Nebius's identity to that of a **tier-one AI cloud provider** capable of serving hyperscale demand.

Inside the Deal

Here's what the agreement entails:

- **Value:** \$17.4 billion guaranteed, with an option to increase to **\$19.4 billion**.
- **Duration:** Runs until 2031.
- **Deployment:** Starts later this year, with phased rollout in 2025 and 2026.
- **Infrastructure:** Delivered from Nebius's new **Vineland, New Jersey data center**, with 300 MW capacity.

This contract is transformative for Nebius. In 2023, the company generated **\$117.5 million in revenue**. Under this deal, it expects at least **\$17.4 billion through 2031**—a leap that makes Microsoft its largest customer by a wide margin.

Why Microsoft is Betting on Nebius

Microsoft has become the face of enterprise AI through its investments in **OpenAI** and AI-powered features like **Copilot** across its software suite. But behind the scenes, AI success hinges on one scarce resource: **GPU capacity**.

Nvidia chips power most of today's AI training and inference workloads, and supply remains tight. For Microsoft, partnering with Nebius offers several advantages:

- **Secured supply:** Guaranteed access to GPU clusters at a time of global shortage.
- **Faster scaling:** Outsourcing infrastructure avoids years of construction delays.
- **Capital efficiency:** Microsoft doesn't need to bear the full cost of building new facilities.
- **Flexibility:** If AI demand shifts, Microsoft isn't locked into owning all the hardware.

This model—big tech renting capacity from AI-native cloud specialists—has become increasingly common. Rivals like **CoreWeave** have struck similar deals, showing how hyperscalers are diversifying their supply lines to stay competitive.

Nebius's Growth Story

For Nebius, the Microsoft agreement is more than just revenue—it's a ticket to accelerated global expansion.

CEO Arkady Volozh described it as the first of several expected long-term contracts with major AI players. He noted that the deal would allow Nebius to **"accelerate the growth of our AI cloud business even further in 2026 and beyond."**

The financial markets responded immediately:

- Nebius stock jumped **60% in extended trading** on announcement day.
- Shares continued climbing, ending more than **50% higher** within two days.
- The stock is now valued at **\$95.72, up from \$64.19** before the deal.

The company's revenue growth also tells a story of surging demand:

- **Q1 2025 revenue:** \$55.3 million.
- **Q2 2025 revenue:** \$105.1 million (a 625% year-over-year increase).

To support the Microsoft contract and beyond, Nebius has launched a **\$3 billion capital raise**—including \$2 billion in convertible notes and \$1 billion in new shares. The funds will finance land acquisition and expanded computing power. Backed by Microsoft's credit quality, Nebius secured favorable terms for debt issuance, with maturities in 2030 and 2032.

Global and Market Impact

This deal is significant not only for Microsoft and Nebius but for the broader AI ecosystem:

- **For Microsoft:** It strengthens Azure's position in the global AI race, ensuring that OpenAI partnerships and enterprise AI services won't be bottlenecked by hardware shortages.
- **For Nebius:** It marks the transition from serving startups to competing in the top tier of cloud providers. With Microsoft as its anchor client, Nebius has financial stability and credibility.
- **For the industry:** The partnership highlights a shift in cloud economics—where hyperscale demand is increasingly being met by specialized, agile providers.
- **On the global stage:** The U.S. is emerging as the primary hub for GPU-heavy data centers. Similar deals could soon spread to Europe and Asia as demand intensifies.

Even competitors are benefitting from the AI surge. Oracle recently reported a boost in its cloud business tied to AI workloads, while Nvidia projects AI infrastructure spending could reach **\$3–4 billion annually by 2030**.

Challenges Ahead

Despite the optimism, risks remain:

- **Execution risk:** Nebius must complete its data centers and deliver on schedule.
- **GPU supply:** Nvidia hardware is still scarce, and any disruptions could slow deployments.
- **Energy demands:** Running a 300-MW data center requires enormous electricity and cooling capacity.
- **Competition:** Other AI-native clouds like CoreWeave and Lambda are also chasing hyperscale partnerships.

Nebius's ability to balance aggressive expansion with reliable delivery will be key to sustaining investor and customer confidence.

The Bigger Picture

This deal signals the rise of a new cloud era. Traditional hyperscalers like Microsoft, Google, and Amazon are no longer building all infrastructure in-house. Instead, they're partnering with **AI-first cloud providers** to secure capacity more quickly and efficiently.

For Microsoft, the Nebius deal is about staying ahead in the AI race. For Nebius, it's about proving that smaller, agile companies can play in the big leagues of global infrastructure. And for the industry at large, it shows that the AI revolution will be powered not only by clever algorithms but by the massive computing engines humming inside data centers worldwide.

Conclusion

The \$17.4 billion Microsoft–Nebius partnership is one of the most significant AI infrastructure deals in history. It ensures Microsoft has the computing power to keep driving innovation in AI while catapulting Nebius into the top tier of cloud providers.

As the world's appetite for AI grows, this deal may be remembered as a turning point—where a once little-known spinoff from Yandex became a key pillar of the global AI economy.