# Nscale Secures $14bn Microsoft Deal to Supply 200,000 AI Chips



British AI infrastructure firm Nscale has signed a landmark deal with Microsoft to supply around 200,000 Nvidia AI chips, in one of the largest AI infrastructure agreements of 2025. Valued at up to $14 billion, the multi-year contract marks a major step in scaling global AI compute capacity amid soaring demand for machine learning and generative AI services.

Under the agreement, Microsoft will receive 104,000 Nvidia GB300 GPUs at Nscale’s Texas data centre and 12,600 at the Start Campus in Portugal, with deliveries beginning next year. Dell Technologies will assist in deploying the infrastructure, building on an earlier $6.2 billion contract with Nscale to supply 52,000 GPUs from its Norway-based AI campus, developed in partnership with Aker.

Nscale, which emerged from an Australian Bitcoin mining spin-out just over a year ago, has rapidly positioned itself as a leading AI infrastructure provider. It raised $1.1 billion last September from investors including Aker, Nokia, Nvidia, Dell and Fidelity, reaching a valuation of $3 billion. The company’s ambition to become a national AI infrastructure champion for the UK is now underpinned by plans to invest up to £2.5 billion in domestic data centre development.

A key part of that expansion is Nscale’s new greenfield site in Loughton, Essex. The facility is set to deliver 50 megawatts of high-performance computing capacity, scalable to 90 megawatts. Construction is expected to complete by the end of 2026, with tens of thousands of GPUs to be installed. Chief executive Josh Payne has outlined plans for a potential public listing by late 2026, subject to progress on construction and chip delivery milestones.

Yet challenges remain. Hyperscale data centre development must navigate complex regulatory environments, environmental standards, and power constraints. AI workloads are energy intensive, and Nscale’s ability to secure stable, green power sources will be closely scrutinised. Logistics and chip supply also pose risks, with global shipping delays or Nvidia capacity constraints potentially impacting delivery schedules.

The deal reflects how technology giants such as Microsoft are scaling their physical infrastructure to meet AI demand. Microsoft’s recent launch of a supercomputer-scale GB300 NVL72 cluster using 4,608 Nvidia GPUs on its Azure platform shows the scale of investment underway. The system delivers 92.1 exaFLOPS of inference performance—significantly reducing AI training times.

Alongside internal build-outs, Microsoft is leasing AI compute from multiple partners—including Nscale, CoreWeave, Nebius and Lambda—bringing its total access to more than 100,000 GB300 GPUs as part of a broader $33 billion cloud infrastructure push. While this helps meet near-term demand, it has raised concerns over energy usage, grid pressure and regulatory oversight, particularly regarding the company’s ties with Nvidia and OpenAI.

Nscale’s rapid evolution from crypto mining to AI infrastructure highlights the strategic role compute capacity plays in the next wave of innovation. With strong investor backing and a UK expansion grounded in sustainability, the firm is well placed to support the UK’s ambitions as a global leader in responsible AI development.

Created by [Amplify](https://www.hbmadvisory.com/amplify): AI-augmented, human-curated content.

## Bibliography

1. <https://blockonomi.com/nscale-signs-14-billion-microsoft-deal-for-200000-nvidia-ai-chips/> - Please view link - unable to able to access data
2. <https://www.reuters.com/world/europe/uks-nscale-signs-deal-with-microsoft-supply-200000-nvidia-ai-chips-2025-10-15/> - British AI firm Nscale has expanded its agreement with Microsoft to supply approximately 200,000 Nvidia AI chips for data centres in Europe and the U.S. The deal, supported by Dell Technologies, could generate up to $14 billion in revenue for Nscale. Deliveries will commence next year from data centres in Texas and Portugal, with plans to further expand operations in Texas. This agreement builds on a previous multi-year deal between Nscale and Microsoft's joint venture with Norway's Aker, involving 52,000 Nvidia chips from a hyperscale AI campus in Narvik, Norway. To support its growing infrastructure, Nscale raised $1.1 billion in September from investors including Aker and Nokia. ([reuters.com](https://www.reuters.com/world/europe/uks-nscale-signs-deal-with-microsoft-supply-200000-nvidia-ai-chips-2025-10-15/?utm_source=openai))
3. <https://www.reuters.com/legal/transactional/blackrock-nvidia-buy-aligned-data-centers-40-billion-deal-2025-10-15/> - An investor consortium including BlackRock, Microsoft, Nvidia, Abu Dhabi’s MGX, Elon Musk’s xAI, the Kuwait Investment Authority, and Singapore's Temasek is acquiring Aligned Data Centers for $40 billion. The purchase of the U.S.-based company from Australian Macquarie Asset Management marks the first major deal for the AI Infrastructure Partnership, established to secure the infrastructure necessary to meet surging demand for artificial intelligence applications. Aligned Data Centers operates over 5 gigawatts of capacity across nearly 80 centres in the U.S. and Latin America. ([reuters.com](https://www.reuters.com/legal/transactional/blackrock-nvidia-buy-aligned-data-centers-40-billion-deal-2025-10-15/?utm_source=openai))
4. <https://www.tomshardware.com/openai-broadcom-to-co-develop-10gw-of-custom-ai-chips> - OpenAI has entered a multi-year agreement with Broadcom to co-develop 10 gigawatts (GW) of custom AI accelerators and rack systems, aiming to start deployment in the second half of 2026 and complete rollout by the end of 2029. Under the deal, OpenAI will design the accelerators and systems, while Broadcom will manage development and deployment. This marks a strategic shift for OpenAI, moving from reliance on Nvidia GPUs toward custom, tightly integrated silicon built for its unique training and inference workloads. The systems will use Ethernet-based networking for better scalability and interoperability. The partnership deepens an 18-month-long existing collaboration between the companies and positions OpenAI to benefit from Broadcom's mature supply chain and ASIC design expertise. OpenAI’s broader hardware strategy now includes around 26 GW of capacity across Nvidia, AMD, and Broadcom technologies. While financial specifics remain undisclosed, the deal is reportedly worth "multiple billions of dollars." This development mirrors a broader industry trend, with other tech giants like Amazon, Google, Meta, and Microsoft also investing in custom accelerators, challenging Nvidia's dominance. Final decisions around foundry, memory, and packaging are pending and will influence production timelines. ([tomshardware.com](https://www.tomshardware.com/openai-broadcom-to-co-develop-10gw-of-custom-ai-chips?utm_source=openai))
5. <https://www.tomshardware.com/tech-industry/artificial-intelligence/microsoft-deploys-worlds-first-supercomputer-scale-gb300-nvl72-azure-cluster-4-608-gb300-gpus-linked-together-to-form-a-single-unified-accelerator-capable-of-92-1-exaflops-of-inference> - Microsoft has launched the world’s first supercomputer-scale GB300 NVL72 cluster on its Azure platform, powered by 4,608 Nvidia Blackwell Ultra GB300 GPUs. This cutting-edge deployment, interconnected with NVLink 5 and Nvidia’s Quantum-X800 InfiniBand fabric, forms a unified accelerator capable of delivering a massive 92.1 exaFLOPS of FP4 inference performance. Each NVL72 rack includes 72 GPUs and 36 Grace CPUs, collectively providing up to 1,440 petaflops and 37TB of fast memory per rack, with 130 TB/s total memory bandwidth. Designed specifically for advanced AI workloads and OpenAI training models, this cluster dramatically reduces training time from months to weeks. It also emphasizes sustainability, using liquid cooling with minimal water consumption. Microsoft plans this as the first of many GB300 clusters, aimed at global hyperscale deployment. The initiative is part of a broader strategic partnership between Nvidia and OpenAI, which includes a $100 billion investment and future deployments involving 10 GW of AI infrastructure. This cluster is a significant milestone in advancing next-gen AI computing capabilities. ([tomshardware.com](https://www.tomshardware.com/tech-industry/artificial-intelligence/microsoft-deploys-worlds-first-supercomputer-scale-gb300-nvl72-azure-cluster-4-608-gb300-gpus-linked-together-to-form-a-single-unified-accelerator-capable-of-1-44-pflops-of-inference?utm_source=openai))
6. <https://www.techradar.com/pro/microsoft-just-got-its-hands-on-100000-nvidia-gb300-chips-and-all-it-took-was-investing-usd33-billion-in-these-startups> - Microsoft has secured access to over 100,000 of Nvidia's latest GB300 GPUs through a $19.4 billion deal with neocloud provider Nebius, part of a broader $33 billion investment in emerging cloud infrastructure companies. This move enables Microsoft to meet booming AI computing demands by leasing power rather than waiting for its own data centres to go live. The deal includes high-end Nvidia GB300 NVL72 server racks, each valued at around $3 million, offering a significant shortcut to scalable AI resources. The strategy reflects Microsoft's shift to rent AI compute from partners like Nebius, CoreWeave, Nscale, and Lambda, while reserving internal infrastructure for clients. Simultaneously, Microsoft is expanding its physical presence with massive data centre projects, including a 315-acre site in Wisconsin designed to host hundreds of thousands of GPUs with independent power and extensive fibre networks. However, this rapid digital expansion is straining energy systems, driving up power prices, and raising environmental and regulatory concerns. Microsoft's deep ties with Nvidia and OpenAI are now under scrutiny amid growing fears of market dominance and antitrust issues. ([techradar.com](https://www.techradar.com/pro/microsoft-just-got-its-hands-on-100-000-nvidia-gb300-chips-and-all-it-took-was-investing-usd33-billion-in-these-startups?utm_source=openai))