

Axe Compute Secures \$260 million, Three-Year Enterprise Contract for 2,304-GPU NVIDIA B300 Deployment

April 22, 2026

Redefining enterprise AI infrastructure: enterprises no longer adapt to cloud constraints — they specify what they need, and Axe Compute delivers it

PITTSBURGH, April 22, 2026 (GLOBE NEWSWIRE) -- Axe Compute Inc. (NASDAQ: AGPU), a neocloud AI infrastructure platform delivering dedicated enterprise GPU compute capacity at global scale, today announced the signing of a 36-month enterprise infrastructure contract with aggregate contract value of approximately **\$260 million** to deliver a dedicated cluster of **2,304 NVIDIA B300 GPUs** and AI-focused high-speed storage for massive data processing and training, deployed in a Tier 3 data center in the United States. The contract represents the **largest enterprise engagement in Axe Compute's history**.

Under the 36-month agreement, which has options to renew for additional years, Axe Compute will deliver dedicated GPU compute and AI-focused high-speed storage infrastructure from a single U.S. Tier 3 data center facility. The cluster is purpose-built to support large-scale AI model training, fine-tuning, and high-throughput inference workloads, powered by current-generation NVIDIA B300 GPUs.

"This agreement is a signal. Enterprise AI customers are no longer willing to adapt their infrastructure roadmaps to the capacity constraints of legacy hyperscalers. A 2,304-GPU B300 deployment, contracted, dedicated, U.S.-based, and priced to compete, is what purpose-built AI infrastructure looks like. We intend to replicate this commercial structure at scale."

— Christopher Miglino, Chief Executive Officer, Axe Compute Inc.

Contract Highlights

Aggregate Contract Value: Approximately \$260 million over 36 months, subject to the terms of the definitive agreement, across both GPU compute and high-speed storage.

Infrastructure: 2,304 NVIDIA B300 GPUs and large AI-focused high-speed storage for massive data processing and training, purpose-built for large-scale AI model training, fine-tuning, and high-throughput inference. All dedicated and committed, while maintaining NVIDIA reference architecture.

Deployment Geography: Single U.S. Tier 3 data center facility.

Power Infrastructure: 4.8 megawatts of dedicated power capacity, delivered on an N+1 redundant basis, providing the fault-tolerant power foundation required for uninterrupted large-scale AI workloads.

Targeted Deployment Start: Q3 2026.

Contract Structure: Secured with a deposit, prepayment, and monthly in advance payments against contracted pricing on a take-or-pay basis. Supported by enterprise-grade service levels, with the ability to add ancillary value-added services like dedicated local loops. Terms architected and provided by Axe Compute to align with the enterprise, not dictated by provider inventory and requirements.

Strategic Significance

This contract illustrates the commercial architecture Axe Compute is scaling toward: multi-year, dedicated GPU deployments with contracted pricing, service levels, and location specified by the customer. At \$260 million over 36 months, it establishes a new benchmark for enterprise AI infrastructure engagements and provides the Company with meaningful long-dated income visibility.

Two structural capabilities of the Axe Compute platform directly enable engagements of this size and structure. First, the platform's geographic reach means customers can match compute capacity to the regions their workloads actually require — a structural flexibility that incumbent providers, constrained to the facilities they have built, cannot always offer. Second, Axe Compute is able to offer dedicated clusters backed by delivery guarantees, ensuring customers receive the needed GPU compute when they need it and when they want it, to support scaling their businesses and serving their end clients. This combined with Axe Compute's predictability means customers know what they will pay each month, with no hidden fees, aligning to their monetization model with no surprises. This deployment is backed by dedicated, N+1 redundant power infrastructure, totaling 4.8 megawatts committed to this cluster alone, fully dedicated, fully supported by 24/7 on-site resources, which is what Axe Compute feels enterprises deserve.

Axe Compute believes this transaction is representative of a broader, structural shift in how enterprise AI infrastructure is procured: customers specify what their AI workloads require and contract accordingly, rather than adapting their AI roadmaps to the constraints of legacy cloud capacity. This agreement is representative of the engagement profile Axe Compute is built to deliver - providing choice, flexibility, dependability, and scalability to a market that is desperate for an alternative model.

Workload Use Cases

The 2,304-GPU B300 cluster delivered under this agreement is purpose-built to support the most demanding AI workloads at enterprise scale. Representative workloads include:

Foundation Model Training: Pre-training large language models and multimodal foundation models requires sustained, high-throughput GPU compute across thousands of accelerators operating in tight coordination. The B300's memory bandwidth and single-spine interconnect performance make it particularly well-suited for training runs at this scale, where GPU utilization and inter-node communication efficiency directly determine time-to-completion and cost.

Fine-Tuning and Domain Adaptation: Enterprises adapting foundation models to proprietary datasets, whether for legal, financial, biomedical, or

customer-specific applications, require dedicated compute that eliminates the multi-tenancy risks and unpredictable availability that characterize shared cloud environments. Dedicated infrastructure ensures data remains within a controlled facility boundary and compute capacity is available on the enterprise's schedule, not the provider's.

High-Throughput Inference: Production AI deployments serving real-time or near-real-time inference at scale, including recommendation engines, content generation pipelines, fraud detection systems, and autonomous decision-making platforms, all require low-latency, high-availability GPU infrastructure with predictable performance. Dedicated clusters eliminate the noisy-neighbor latency spikes that plague shared cloud environments, delivering consistent, predictable performance at scale.

AI-Intensive Data Processing: The integration of high-speed AI-focused storage (e.g. Vast) with the GPU cluster enables workloads that demand rapid ingestion, transformation, and processing of massive datasets at training time, including multimodal data pipelines processing image, video, audio, and text at scale. Storage throughput and proximity to compute are critical bottlenecks at this data volume; the co-located architecture directly addresses both.

About Axe Compute Inc.

Axe Compute Inc. (NASDAQ: AGPU) is a neocloud AI infrastructure platform built on a fundamental premise: AI innovation should not be constrained by infrastructure supply and performance limits. Axe Compute gives enterprises and AI innovators choice across hardware, geography, and deployment speed. Axe Compute also operates a Strategic Compute Reserve that translates to enterprise GPU access, converting reserve holdings into deployable AI infrastructure capacity. Axe Compute is among the first publicly traded companies delivering this model at scale. Learn more at axecompute.com.

Forward-Looking Statements

This press release contains "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. Forward-looking statements include, but are not limited to, statements regarding the anticipated timing, scope, value, and performance of the contract described herein; the expected deployment schedule; the availability of hardware and facility capacity; the customer relationship and its future progression; the Company's ability to secure additional engagements of similar scale; and Axe Compute's broader business strategy and market positioning. These statements are based on the Company's current expectations and assumptions and are subject to known and unknown risks and uncertainties that could cause actual results to differ materially, including risks related to the execution and enforceability of the definitive agreement, hardware supply chain constraints, facility readiness, customer performance, macroeconomic conditions, competition, regulatory matters, and other risk factors described in the Company's filings with the U.S. Securities and Exchange Commission. Axe Compute undertakes no obligation to update any forward-looking statement, except as required by applicable law.

Investor & Media Contacts

Investor Relations

Erin McMahon

erin@axecompute.com