

“Aggressive demand” from Musk’s AI startup for Nvidia chips led to supply restrictions

The letter states that although Nvidia meets the needs of all its clients, the growing demand from Elon Musk’s AI startup xAI affects the speed of deliveries, necessitating some restrictions. This is related to Musk’s announcement in September that his startup xAI launched the Colossus training cluster in just 122 days, setting a record. Musk called Colossus “the most powerful AI training system in the world,” operating on 100,000 Nvidia H100 training chips.

Musk also stated that the cluster “will double in size in a few months” and will consist of 200,000 chips, including 50,000 more powerful Nvidia H200 chips. Competitors of xAI in the AI field, including OpenAI and Meta*, also have hundreds of thousands of Nvidia chips, but Colossus, as a separate computing cluster, has the largest number of processors in the world.

It is noted that the xAI supercomputer will be used to train the next generation of the large language model (LLM) xAI Grok, which is expected to compete with OpenAI’s GPT-4.

□ The organization is recognized as extremist, and its activities are banned in Russia.