Tomorrow's Affairs
Analysis of today
Assessment of tomorrow

By: Tomorrow's Affairs Staff

Can space solve AI's energy problem?

Nvidia, the world's largest producer of artificial intelligence processors, and Cisco, a major supplier of networking equipment, launched a project on 21 August that aims to lift data centres into low-Earth orbit.

Starcloud has raised an additional $250 million, bringing the US start-up's total funding to $450 million and valuing the company at $2.3 billion.

It has one processor demonstration satellite that has been tested in space, a production facility under construction, and plans for a constellation of up to 88,000 spacecraft, but has not yet launched commercial service.

The entry of two large technology companies is changing the status of an idea that, until recently, belonged to a small circle of space engineers.

The spread of artificial intelligence on Earth increasingly depends on power plants, transmission lines, water, land and permits, so the industry is looking for new locations for computing capacity.

If part of data processing moves into orbit, rockets, communication frequencies and laser links will become part of the digital infrastructure on which companies and countries depend.

Starcloud's value rests on the expectation that demand for AI capabilities will grow faster than power grids can expand, and that new rockets will reduce the cost of transporting heavy equipment.

The company must therefore prove that a computer in orbit can become a reliable service that customers pay for regularly, because this is where the line lies between an interesting experiment and a new industry.

The first satellite proved that AI can work in space

In November 2025, Starcloud launched Starcloud-1 with Nvidia's H100 processor, a chip designed for demanding AI workloads in large terrestrial data centres.

During the mission, Starcloud, according to its own data, ran AI inference – that is, it used an already trained model to process new data – and then trained a small experimental language model, nanoGPT, in orbit.

The mission showed that a powerful commercial processor can continue to operate after launch and in the orbital environment, although Starcloud has not released a complete, independently verified analysis of its behaviour under radiation and temperature.

> The much larger Starcloud-3 is expected to use Starship, a heavy rocket being developed by US space company SpaceX

The scope of the test is limited, however, because Starcloud-1 is a demonstration platform, whereas a real data centre must provide stable, uninterrupted service to many users.

Training a small model validates the operation of hardware and software in orbit, but reveals little about training large models, which require thousands of interconnected processors, massive data exchanges and reliable power over many months.

The first commercial step should be Starcloud-2, a system of two eight-kilowatt satellites planned for 2027. The satellites would process data for customers that the company says include US government agencies, although the original 2026 deadline has already been postponed.

The much larger Starcloud-3 is expected to use Starship, a heavy rocket being developed by US space company SpaceX, while the new capital will fund its development, production and future launches.

Why the AI industry is looking to orbit

New AI complexes on Earth consume hundreds of megawatts of electricity and require connections to heavily loaded grids, large areas of land, permits, and heat removal systems.

Their construction often depends on new power plants and transmission lines, so the schedules of energy companies increasingly determine the pace at which the technology industry can expand.

In orbit, the sunlight is more intense and, on certain trajectories, available for most of the day, while there are no costs for land, grid connections, or cooling water.

However, solar panels, batteries, cables and protective electronics are all very heavy, and every kilogram has to be tested and launched from Earth, which is why free sunlight does not necessarily mean cheap energy.

Cooling can be even more demanding, since there is no air around the satellite to absorb heat from the processor, nor water to carry it away, as in many terrestrial data centres. The heat therefore must reach the radiators – large flat surfaces that emit it into space in the form of infrared radiation – via pipes or liquid inside the spacecraft.

Despite their name, they do not heat the environment but act as an external cooling system, without which computers would quickly overheat; more powerful processors require larger surfaces, a heavier satellite, and a more complex construction.

> An Earth observation satellite could recognise images of fires, floods, ships, or military movements while still in orbit and send back the processed results

In a paper, "Orbital Data Centers: Spacecraft Constraints and Economic Viability," published as a preprint in late April, Slava G. Turyshev of Caltech's Jet Propulsion Laboratory calculated that an orbital system providing one megawatt of electrical power for computing equipment would require approximately 5,640 square metres of solar panels and 2,500 square metres of heat-dissipation area.

His model shows that today's costs of launching and building space equipment do not yet allow such a system to compete with ground-based data centres for general use. As a more realistic initial market, Turyshev identifies the processing of data already generated in space and computing services for satellite communications.

An Earth observation satellite could recognise images of fires, floods, ships, or military movements while still in orbit and send back the processed results instead of transmitting a huge volume of raw images to ground stations.

Faster delivery of usable data has immediate value for meteorology, disaster management, agriculture, insurance, intelligence services, and the armed forces, so orbital computing will initially target these users as its primary market.

Everything depends on the rocket, which is not ready yet

Starcloud's long-term projections depend on Starship, a rocket system that SpaceX is developing to carry very large payloads and be reusable.

Starship is expected to carry computing equipment at a cost significantly lower than today's launches, but it does not yet have regular commercial flights, proven rapid reusability, or the market price on which Starcloud's projections rest.

Company CEO Philip Johnston has admitted that not being able to reserve SpaceX's capacity for 2029 would be a serious problem.

Starcloud is therefore considering buying an entire flight of Falcon 9s, SpaceX's smaller rocket that already regularly carries satellites, as well as signing contracts with other operators. However, competing heavy rockets do not yet have the pace needed to mass-build a computer network in space.

The price of energy becomes crucial only once the cost of delivering panels, processors, radiators and other equipment has fallen sufficiently.

> Nvidia may benefit long before Starcloud starts making serious money

Dependence on SpaceX also extends to communications between satellites, as in May Starcloud contracted more than 50 laser terminals from Starlink, SpaceX's satellite internet network, for at least 25 of its spacecraft.

Each satellite should receive two devices that use a directed beam of light to transmit data at speeds of up to 25 gigabits per second over distances of up to 4,000 kilometres, according to the supplier. SpaceX thus becomes simultaneously a carrier, a communications provider and a key partner of a company that wants its own infrastructure in orbit.

Nvidia may benefit long before Starcloud starts making serious money, because in March it unveiled the Space-1 Vera Rubin Module, a package of processors and related equipment adapted to work on a satellite, which Starcloud plans to send on its first flight in late 2028.

The data collected by Starcloud-1 from the H100 processor helps in choosing the operating temperature, radiation protection and size of the cooling surfaces, so Nvidia is using Starcloud to test future equipment in real conditions.

The regulator faces a request for 88,000 satellites

Starcloud has asked the Federal Communications Commission, the US regulator responsible for radio frequencies and some satellite licences, to approve a constellation of up to 88,000 satellites, and has announced a long-term computing power requirement of 20 gigawatts.

The commission has accepted the request for consideration, but the constellation has not yet been approved, and a single demonstration satellite provides little evidence of the safety of a network that would comprise more than five times the number of all operational satellites currently in Earth orbit.

The regulator is simultaneously considering SpaceX’s proposal for a million orbital centres and a request from Blue Origin, the space company founded by Jeff Bezos, for 51,000 satellites.

> Full approval on the basis of a single demonstration therefore seems unlikely

All such systems use radio frequencies to track spacecraft, transmit data on their status and send commands, even when the main content is transmitted by lasers, while tens of thousands of new objects increase the risk of collisions and raise issues of interference, the removal of worn-out satellites and the consequences of their burning up in the atmosphere.

Full approval on the basis of a single demonstration therefore seems unlikely, whereas a gradual approach with an initially limited number of satellites, mandatory reporting, and further expansion only after proven reliability is more realistic.

Such an outcome would slow the most ambitious version of the business plan, but it would give the company a more sustainable growth path and reduce the risk that a single incident could bring the entire project to a halt.

The first real test comes in 2027

In 2027, Starcloud-2 will provide the first serious answer to whether there is a market for orbital data processing beyond technology demonstrations. A successful mission, with work ordered and paid for by users, would show that the company can deliver a reliable service.

By the end of the decade, orbital centres are likely to complement ground-based AI complexes, mainly through a smaller, more expensive and strategically important network for processing satellite, government and security data

The next test will come with Nvidia's space module and the larger Starcloud-3 system, possibly in 2028 or later. Until then, the company's estimated value will depend primarily on customer contracts, booked launches and progress in securing permits.

By the end of the decade, orbital centres are likely to complement ground-based AI complexes, mainly through a smaller, more expensive and strategically important network for processing satellite, government and security data.

Wider everyday use of already trained AI models will depend on the reliability of Starship, while training the largest models will require even lower launch costs, more powerful laser links and the ability to transport the heat generated by thousands of processors safely into space.

Nvidia and SpaceX can make money with a much smaller network, while Starcloud must build its own value through reliable service and customers willing to pay for it.

By 2029, the company will need to secure regular contracts, sufficient launch dates and permission to expand gradually, while the key cost of the whole endeavour – getting the equipment into orbit – will still depend on other companies.

It is this dependence, far more than the value of the latest investment round, that will determine whether Starcloud has opened up a new market or merely created the most expensive laboratory in orbit above Earth.