INTELLIGENCE UPDATE

Data centers in space: sky high costs, astronomical challenges?

Andy Lawrence   24 Dec 2025

OPINION

One evening somewhere around 2015, the CTO of a global data center infrastructure company was talking excitedly over cocktails. "Within a decade," he told analysts and customers, "we'll be building the next generation of data centers in space." His chief executive was listening (as were we) and quietly called him aside. "Keep it real," the CTO was told. "We're a serious company. It gives a bad impression."

The idea, however, has continued to rumble on, driven in part by the cost and limited availability of power on Earth. In 2025, the topic crashed out of its orbit and became part of the latest AI hype cloud. Almost as if flying in formation, the chief executives of Tesla/SpaceX (Elon Musk), Amazon (Jeff Bezos) and Google (Sundar Pichai) all stated in October and November that they expected large data centers to be built in space within 10 years.

The Uptime Intelligence team has a tendency toward skepticism — or at least about the speed at which new technologies are adopted — so when we were asked about the idea on stage at the GITEX technology conference in Dubai in October, we were dismissive. There may be abundant solar power in space, but terrestrial concerns over costs, resiliency and carbon were unlikely to be solved by repeatedly blasting servers into orbit. Other experts on the panel, however, were much more favorable.

Are we wrong? Is there serious momentum building? Google, in particular, has set up a formal project (known as Suncatcher) to explore the concept, with the first satellite launches scheduled for 2027. The ultimate goal is a huge, kilometer-wide data center constellation comprising 81 satellites. And Google is not alone. Starcloud, a startup formerly known as Lumens, has raised $20 million in seed funding to date and has already had a light satellite, carrying a small server equipped with an Nvidia GPU, inserted into Low Earth Orbit (LEO) for testing. This will be followed by a second, larger mission into space in 2026. The company's stated long-term aim is to build gigawatt-scale data centers in orbit. The startup already has a neocloud partner in US company Crusoe.

The underlying driver — and assumption — behind these efforts is that generative AI will become a

foundational, ubiquitous technology that will need vast amounts of compute and power. Proponents argue that it may ultimately be more efficient to power the training work in space and transmit the resulting bits to Earth, rather than move the electrical energy to Earth after solar conversion or to use terrestrial power sources for the training process.

The prospect of building data centers in space has some appealing benefits. Solar panels there can provide free energy run for most of the day, every day, at more than 30% efficiency, compared with less than 20% on Earth. Cooling could, in theory (see below), be completely free; real estate is endless; and laser-based optical communications in a vacuum can outperform even physical fiber, opening up many new configurative possibilities. Launch costs, meanwhile, are falling year by year.

Google appears to believe that the idea is worth some investment, but it is not yet clear whether it is as committed as some media reports suggest. Discussing the Suncatcher project, chief executive Sundar Pichai said that "like any moonshot, it's going to require us to solve a lot of complex engineering challenges." The key words in this sentence are "a lot" and "complex."

Alongside its announcement, Google published a technical paper examining the issues (see *Towards a future space-based, highly scalable AI infrastructure system design*). The paper states that it is "working backward from an eventual future in which the majority of AI computation happens in space." Using projections and comparisons with terrestrial data centers, it suggests that such an approach could work in theory and might even be economically competitive — but only if many factors go in its favor.

One immediate issue is that, as on Earth, renewable energy — even from solar panels in orbit — lacks energy density, so a lot of real estate is needed. Powering tens of thousands of GPUs and other IT would require many thousands of square meters of solar panels. The result is a vast, large-scale operation, with significant weight to lift into space. Rather than build one very large modular satellite, Google and others propose deploying many smaller ones orbiting in a close formation (Google's concept envisions 81).

These smaller satellites would be positioned just a few hundred meters apart and communicate using optical links, which can offer lower latency and higher bandwidth than current optical fiber. In theory, this would allow a GPU cluster to be distributed across many satellites while still communicating with Earth at acceptable latencies for most applications. However, such networks use active laser links that, like the servers themselves, require continuous power.

This raises another issue. To maximize solar exposure, the satellites would need to operate in a dawn-to-dusk Sun-synchronous orbit (SSO), passing above a certain point on the ground exactly the same time, every day. This would maintain near-continuous sunlight for most of the year — but not all the time due to equinox eclipses. If the data center is to operate without regular downtime, it would need batteries to ride through these interruptions, roughly two days a year, usually in short periods of up to 5 minutes each. In other words, orbital data centers are going to need batteries — possibly in similar configurations to terrestrial facilities. Unlike on Earth, however, these solar outages occur every 90 minutes during equinoxes, so a reserve would be needed to

allow for recharging, further adding complexity and weight.

An additional facility resiliency challenge is the risk of cascading damage: space debris could strike one satellite to generate further debris that hits adjacent satellites, setting off a chain of collisions known as the Kessler Syndrome. Data center operators may also be concerned that large data centers are more exposed to hostile action and electronic interference than their terrestrial sites.

Cooling presents a major challenge as well. The liquid cooling systems, needed for high-performance compute, would require major reengineering because the movement of fluids and bubbles is different in microgravity. Thermal management systems researched by NASA for space electronics — such as two-phase loops and capillary-driven flows — would need to be adapted, if possible, for dense data centers. Even then, removing heat from electronics is only half of the problem. Heat rejection, in the absence of an atmosphere, would need vast surface areas for radiators (high emissivity in the infrared spectrum) pointing toward deep space — the exact opposite direction of where the solar panels must face.

A further problem is radiation exposure, which significantly shortens component lifespans of microelectronics, from embedded controllers to IT silicon. Mitigating this requires extensive shielding, further adding weight and costs. This also raises the question of how operators would handle equipment failures or maintenance of cooling and power systems. While failed processors can be left in situ, failures in cooling or power are more problematic. More than outright hardware failures, space radiation also multiplies the probability of data corruption and computational errors going undetected. To date, proponents have not clearly addressed this issue.

For all these issues — and many others — some companies still intend to push ahead. Yet even if the many technical hurdles are overcome, the economics remain tough. Google has estimated that launch prices would need to fall to around $200 per kilogram before space-based data center costs are in a similar range to those of terrestrial facilities. Today, launch costs are more than $3,500 per kilogram. They would have to fall by nearly 95% — and Google agrees this might only happen if the industry is working at a large scale.

At that scale, however, carbon emissions — including soot from rocket launches — from these launches become such a serious issue that building data center constellations in space would likely be viewed as a backward environmental strategy; one that also contributes to space junk and non-recyclable waste. In 2024, Microsoft reported that 90% of its server components for the cloud were recycled. In space, this figure might drop to zero.

For all the celestial engineering and economic challenges, however, the biggest hurdle to the development of commercial-scale space data centers will likely lie on Earth. At least three technologies could start making a big difference to power availability in the 2030s and 2040s — small modular reactors, carbon capture at scale (new metal organic frameworks are promising), enabling continued large-scale use of fossil fuels and, longer term, fusion energy. If any of these succeed at scale — alongside the more prosaic use of renewable energy and battery storage — then power is likely to become less of a constraint. Some projects are moonshots, but some are

simply more moonshot than others.

ABOUT THE AUTHOR

Andy Lawrence

30 Dec 2025

Andy is a founding member and the Executive Director of Research for Uptime Institute Intelligence, which analyzes and explains trends shaping the critical infrastructure industry. He has extensive experience analyzing developments in IT, emerging technologies, data centers and infrastructure, and advising companies on technical and business strategies.

alawrence@uptimeinstitute.com

About Uptime Institute

Uptime Institute is the Global Digital Infrastructure Authority. Its Tier Standard is the IT industry's most trusted and adopted global standard for the proper design, construction, and operation of data centers – the backbone of the digital economy. For over 25 years, the company has served as the standard for data center reliability, sustainability, and efficiency, providing customers assurance that their digital infrastructure can perform at a level that is consistent with their business needs across a wide array of operating conditions.

With its data center Tier Standard & Certifications, Management & Operations reviews, broad range of related risk and performance assessments, and accredited educational curriculum completed by over 10,000 data center professionals, Uptime Institute has helped thousands of companies, in over 100 countries to optimize critical IT assets while managing costs, resources, and efficiency.