Architecture · Technical Paper

The Constraints We Put in Orbit, Part I: The Heat Problem We Keep Calling a Bandwidth Problem

Part I of III. Orbital AI data centers don't have a bandwidth problem, they have a heat problem: a thermodynamic law no one can bargain with, hiding one layer beneath the laser downlinks everyone argues about.

Jennifer McKinney · August 15, 2026 · 21 min read

This is Part I of a three-part series on what orbital AI infrastructure actually escapes, and what it does not. Part II turns to jurisdiction. Part III turns to governance and trust.

Thesis

I want to put the argument up front and then spend the rest of this series deserving it, because the route I took to get here matters as much as the destination.

The pitch for orbital AI data centers is, at root, an escape story. Up there, the reasoning goes, you escape the grid fights and the water rights, you escape the cooling bills, and, though it is said more quietly, you escape the regulators and the jurisdictions and the people asking who gave you permission. Lift the data center off the planet and you lift it out of the planet's constraints.

My claim is that this is precisely backwards. Orbit does not *subtract* constraints. It *relocates* them, to places where they are harder to see, harder to measure, and harder to enforce, and then it calls that relocation progress. This holds across three completely different layers of the problem, and the fact that it holds across all three is what convinced me it is a real pattern and not a clever line. The physics constraint does not disappear; it moves from "cooling bill" to "the law of thermodynamics you cannot bargain with." The legal constraint does not disappear; it moves from "one regulator" to "several incompatible sovereigns stacked on one moving object." The accountability constraint does not disappear; it moves from "institutions that watch the powerful" to "a gap filled by nothing but the discretion of whoever got there first, assessed by experts those same actors pay."

Three movements, then. Heat, jurisdiction, trust. Each is a story about a hard limit the romance of "just put it in space" pretends is not there. I'll walk through them the way I actually got there, because I started this only meaning to sanity-check a claim about lasers, and ended up somewhere much bigger.

Movement I: The Heat Problem We Keep Calling a Bandwidth Problem

Starting with the question I set out to answer

The question that set this off was a good one, and a leading one, the kind that sounds settled the moment you hear it: are lasers and RF inadequate, unreliable, and wasteful for AI data centers in space, dependent on long chains of hardware and software just to power the things and shuttle data to and from Earth?

My honest first reaction was that the question contains a trap, and the trap is instructive enough that I want to start there rather than hide it. Four accusations are bundled into one claim: that the links are *inadequate*, that they're *unreliable*, that they're *wasteful*, and that they're *dependency-chain-heavy*. We say them in one breath as if they rise and fall together. They don't. A communications link can be supremely reliable and still power-hungry. It can be efficient and fragile. It can be adequate in raw capacity and still sit at the end of a Rube Goldberg machine of ground stations and error-correction. If you let the four travel as a pack, you'll end up agreeing or disagreeing with the whole bundle without ever noticing that two of the four are true, one is misdirected, and one is the most important claim in the discussion, just not for the reason it was offered.

So this movement does something a little contrarian with its own founding question. It takes the question seriously enough to take it apart. And in taking it apart, it ends up somewhere I didn't expect when I started: the link layer, the very thing I set out to indict, is mostly fine. The crime is happening one layer down, in a place the popular conversation barely mentions, and it is a thermodynamics crime, not a communications one.

The seductive version, where the lasers work

Here is the story that makes orbital data centers feel inevitable, and I want to tell it straight because it's genuinely strong before we complicate it.

Start with power. Sunlight in orbit arrives at about 1,361 watts per square meter, uninterrupted by atmosphere, weather, or night if you choose your orbit well (Kopp & Lean). The Project

Suncatcher team at Google estimates a solar panel up to roughly eight times more productive in the right orbit than the same panel on the ground (Google Research). Put your compute in a dawn-dusk sun-synchronous orbit and you get near-continuous illumination. Free real estate, free fuel, and what feels like free cooling. For an industry currently fighting municipalities over water rights and grid interconnects, this is a seductive pitch, and the people making it are not cranks. By early 2026 the field includes Suncatcher, with 81-satellite clusters flying in formation on optical inter-satellite links and two prototypes contracted with Planet for early 2027 (Google Research); Starcloud, which actually flew an NVIDIA H100 on orbit in November 2025 (Starcloud); the Thales Alenia EU-funded ASCEND feasibility study (Thales Alenia Space); Jeff Bezos publicly forecasting gigawatt-class space data centers within a decade or two (Reuters); and a SpaceX FCC filing in January 2026 gesturing at up to a million orbital-data-center satellites and, with no apparent embarrassment, "a Kardashev Type II civilization." (FCC Notice DA 26-113)

Now the part my own question flagged as broken: getting the data out. This is where the seductive story is strongest, because the evidence is already in hand. In 2023, NASA and MIT Lincoln Laboratory flew TBIRD, a laser terminal on a shoebox-sized 6U CubeSat, and pulled 200 gigabits per second down to a ground station, moving as much as 4.8 terabytes in a single pass lasting under five minutes, which NASA characterized as more than a thousand times faster than a comparable radio link, from a transmitter drawing roughly 100 watts (NASA / MIT Lincoln Laboratory). The Suncatcher bench demonstrator hit 1.6 terabits per second across a single transceiver pair (Google Research). The SpaceX Starlink V3 targets a terabit per second of downlink per satellite, more than ten times the previous generation (Via Satellite). The Blue Origin optical-terminal venture claims up to six (Blue Origin).

Pause on what that means against that opening word, *inadequate*. The physics here is not subtle: an optical carrier oscillates near 193 terahertz, while space radio sits at or below 40 gigahertz, about four orders of magnitude lower. Shannon-Hartley says capacity scales with bandwidth, and bandwidth scales with carrier frequency, so optical buys you something like ten to a hundred times the throughput for a given engineering effort, in a tighter, lower-power, harder-to-intercept beam (NASA / MIT Lincoln Laboratory). On the specific charge of *inadequate*, the data is already in and the verdict is no. The lasers work. They demonstrably move data-center-scale volumes today.

If you stopped here, and a lot of the breathless coverage does stop here, one would conclude the opening claim is simply wrong and the only thing between us and orbital hyperscale is launch cost. Hold that thought, because it is a trap of its own.

The two accusations that survive

Two of the four charges don't dissolve under scrutiny. They just relocate.

Unreliable is half-true, but not where one would expect. The lasers themselves are reliable; the *ground* is the problem. A narrow optical beam that screams through vacuum runs straight into a wall of water vapor at the atmosphere. Cloud cover doesn't gently attenuate an optical downlink, it can impose attenuation well beyond 100 decibels per kilometer, which is engineering-speak for "the link is gone." (NASA / MIT Lincoln Laboratory) Radio shrugs off weather that optics cannot survive. The honest fix is not a better laser; it's a *network* of geographically scattered optical ground stations so that some subset always sees clear sky, plus radio as the weather-immune fallback for the housekeeping data you cannot afford to lose. Which is exactly the architecture the serious players are filing for (Via Satellite). So *unreliable* is real, but it's a property of the Earth's troposphere and the ground segment, not of the link technology, and it's why the fourth charge is the fairest one in the bundle.

Dependency-chain-heavy is the charge I'll concede most cleanly. To make optical downlink work you need relay constellations or inter-satellite meshes, a diverse ground-station network, modems, automatic-repeat-request error correction to claw back data lost to atmospheric scintillation, and pointing systems that hold a beam steady to within microradians. the TBIRD pointing budget was about 20 microradians, roughly the angle a dime subtends from a kilometer away, maintained from a tumbling CubeSat (NASA / MIT Lincoln Laboratory). That is a long chain, and every link in it is a thing that can fail. That charge is right. But, and this matters for the direction of this argument, notice that this is true of *every* high-performance communications system ever built, including the fiber under the ocean that most of the internet's traffic travels through. Complexity is the cost of capacity. The question is never "is there a dependency chain," it's "is the chain buying you something worth its fragility," and for capacity, it is.

That leaves **wasteful**. And here is where the movement turns, because *wasteful* is the accusation that is pointed at the wrong target.

The floor gives way

The energy cost of the communications link is, in the full system budget, a rounding error. That 200-gigabit TBIRD terminal sipped on the order of 100 watts (NASA / MIT Lincoln Laboratory). In a facility measured in megawatts, the comms layer is not where the energy fight is won or lost. So if you indict the lasers for being wasteful, you've aimed at the one component that is thrifty.

Here is the thing the seductive story quietly skipped. Every watt of sunlight you so cleverly harvested has to go somewhere after your TPUs are done thinking with it. On Earth, "somewhere" is air and water; convection and conduction carry your waste heat away, which is why data centers fight over water rights in the first place. In vacuum, there is no air and no

water. There is no convection. There is no conduction to an ambient that is not there. There is exactly one way for heat to leave a body in space, and it is to radiate as infrared light. That's it. That's the whole menu.

And radiation is governed by the Stefan-Boltzmann law, which is the quiet villain of this entire field: the power a surface radiates scales with its area and with the *fourth power* of its temperature. The only knobs you have are how big the radiator is and how hot the operator is willing to run it. There is no third option. Physics did not leave you a back door.

Run the numbers and the seductive story inverts. the Starcloud white paper concedes that a two-sided radiator at around 20 degrees C sheds roughly 633 watts per square meter, and notes this is on the order of a thousand times slower than water cooling on the ground (Starcloud). IEEE Spectrum, working from an ABI Research analysis in June 2026, put it in hardware terms: a single 700-watt accelerator, the kind already in racks today, needs about 1.4 square meters of radiator at 60 degrees C. Leave it in orbit for five years and ultraviolet exposure plus atomic oxygen degrade the radiator coatings until the system needs closer to 2.0 square meters to do the same job, a roughly 40 percent physics tax that just accrues with time on station (IEEE Spectrum). Scale that: a single dense rack pushing 40 kilowatts needs on the order of 80 square meters of radiator. A 100-megawatt facility, modest by hyperscale standards, needs thousands of them.

For a sense of how heavy this gets, look at the one orbital thermal system we've actually flown at scale. The International Space Station rejects about 70 kilowatts of heat through six deployed radiator wings, running two ammonia loops, with a measured areal mass around 8 kilograms per square meter (NASA). Scale *that* naively to a one-megawatt compute facility and the mass ratio means launching something like 100 tonnes of radiator to cool roughly 10 tonnes of computer. The cooling outweighs the thing being cooled by an order of magnitude, before you've launched a single gram of solar array, structure, or propellant.

So look back at those four opening words with fresh eyes. *Inadequate?* No, the links work. *Unreliable?* Only the ground segment, only because of weather. *Wasteful?* Yes, catastrophically, but the waste is thermal, not communicational. The accusation was aimed at the messenger. The real culprit was sitting one layer down the whole time, radiating into the void at 633 watts a square meter and dragging a hundred tonnes of aluminum behind it. We keep arguing about a bandwidth problem when we have a heat problem.

Are the assumptions "antiquated"? Yes, but be precise about which ones

The deeper version of that question asks whether the materials, thermal, and transport assumptions underneath all this are antiquated. My answer is a qualified yes, and the qualification is the entire point, because there is a sloppy way to be right here that will get you

laughed out of a serious room.

The sloppy version says the *physics* is antiquated. It is not. Stefan-Boltzmann is not going to be repealed. Shannon-Hartley is not waiting for a firmware update. The free-carrier physics that makes silicon modulators sensitive to radiation is not a bug in our understanding that a clever startup will patch. When a pitch deck implies that some breakthrough will get us *around* these, that deck is selling something. The physics is exactly what dooms the naive lift-and-shift design, and respecting it is what separates analysis from marketing.

The defensible version says the *deployed practice* is antiquated, that our flight hardware and our procurement habits are lagging badly behind materials that already exist on lab benches and in the literature. That version is true, and it is where the real opportunity lives.

Consider the radiator. The ISS flies conventional aluminum-and-ammonia at about 8 kilograms per square meter (NASA). The NASA target for advanced radiators is closer to 2 kilograms per square meter, the regime where the cooling stops outweighing the computer, and bare carbon-fiber radiators running at 800 to 1000 kelvin have approached it (NASA Advanced-Radiator Research). Graphite and graphene heat spreaders reach thermal conductivities on the order of 1,700 watts per meter-kelvin, well past copper, and they're commercially available today (Balandin et al., Nano Letters 2008). Liquid-droplet radiators, where you spray coolant droplets into vacuum, let them radiate across an enormous effective surface, and recapture them downstream, have been studied for decades and are finally showing up in startup testbeds. None of this is exotic. It's just not what we fly, because flying it requires someone to absorb the qualification risk, and conservatism in space hardware is rational right up until it becomes the thing holding the field back.

The photovoltaics tell the same story. Space-grade solar still leans heavily on germanium substrates, with a supply chain concentrated in one country, when perovskite and advanced III-V chemistries promise better power-to-mass off that substrate (IEEE Spectrum). Antiquated practice, not antiquated physics. But the place this argument gets genuinely interesting, the part where the analytical originality lies, is one layer below even the radiators, inside the machine itself.

The asymmetry nobody is pricing in

Here is what almost no one in the orbital-data-center conversation is talking about, and it is the most important unpriced risk in the whole enterprise.

Compute scales with *volume*. You stack chips, you stack racks, density climbs, and improvement following the chip-density curve keeps making each cubic meter think harder roughly every year. Heat rejection scales with *area*, that is Stefan-Boltzmann again, surface

area is the only free variable. And the gap between a thing that grows with volume and a thing that grows with surface area is not a detail; it's a structural divergence. The better your compute gets, the *worse* your area-bound radiator falls behind it. Every generation of denser chips makes the thermal asymmetry more lopsided, not less. On Earth you paper over this with more water. In vacuum you cannot.

Now layer on attrition. Your radiator doesn't just fail to keep up, it degrades, losing dozens of percent of its capacity to UV and atomic oxygen over a five-year mission (IEEE Spectrum). And the chips themselves die in a place where you cannot send a technician with a replacement. A dead GPU in a terrestrial data center is a hot-swap on a Tuesday. A dead GPU in orbit is dead capacity for the life of the satellite, unless the constellation has flown enough cheap redundant nodes to vote around failures, sacrificing a chunk of your compute to triple-redundancy, or the business model depends on an in-space servicing economy that does not yet exist at the price the economics would require.

And here's the connective tissue that ties this back to the opening subject, photonics, because the conversation treats the laser downlink and the optics *inside* the machine as unrelated, and they are not. Inside a modern AI accelerator, data increasingly moves on light, not copper, because copper at these speeds can't reach across a meter. Those on-chip optical modulators, the silicon microrings that switch the light, are exquisitely sensitive to temperature. the thermo-optic coefficient of silicon drifts the resonance about 80 picometers per kelvin, which means a temperature swing of just a few degrees pushes a high-Q ring clean off its resonance and forces you to spend continuous power on heaters just to hold it in tune (arXiv: Silicon Ring Thermometers). Now put that part in low Earth orbit, where the platform plunges from sunlight into the shadow of Earth and back every 90 minutes. MISSE experiments conducted by NASA measured external temperature swings from roughly minus 17 to plus 38 degrees C across exactly that day-night cycle (AIAA / MISSE-16 Thermal Data). You are asking a component that hates a few degrees of drift to survive a 70-degree thermal cycle, sixteen times a day, for years.

And radiation does its own damage. In a landmark 2024 study in *Science Advances*, Mao and colleagues flew foundry silicon photonic modulators externally on the ISS for 325 days, roughly 6,700 orbits, and reported a genuinely encouraging headline (the electro-optic tuning efficiency was essentially unchanged) alongside the quieter bad news: modulator extinction ratios fell from around 24 to 14 decibels, intrinsic Q dropped about 30 percent, and some centimeter-scale device arms simply failed where heavy ions struck them (Science Advances). Follow-up gamma-irradiation work by Zhao and colleagues in 2025 watched modulator bandwidth fall from 52 to 31 gigahertz under dose, though much of it recovered with a thermal anneal (ACS Photonics). The materials that *don't* have this problem are known: silicon nitride drifts roughly ten times less with temperature and shrugs off radiation; thin-film lithium niobate modulates through the Pockels effect with no free-carrier damage pathway at all; indium phosphide rings

have shown only single-digit-percent degradation under doses that would wreck silicon (PubMed). They are sitting in the literature. We fly silicon anyway, because silicon is cheap and has a foundry. Antiquated practice. Not antiquated physics.

The asymmetry, taken together, resolves to four interconnected failures: compute grows with volume and improves yearly; heat rejection grows with area and degrades with exposure; the very photonics that move the data are thermally and radiologically fragile in precisely the cycling, irradiated environment the industry proposes to put them in; and none of it can be repaired in place. That is the unpriced risk. That is the thing the pitch decks skip.

Where the physics lands

If the heat problem is the master constraint, then the rational near-term architecture is not the lift-and-shift design that started this inquiry at all. Lifting a terrestrial data center to orbit to serve Earth, training and inference whose results come straight back down, is the *hardest* possible version, because it maximizes both the compute you must cool and the data you must shuttle across the weather-blocked air-vacuum boundary. The version that actually pencils out is the inverse: compute that lives in orbit to serve *space-side* workloads, where the data is already up there. Earth-observation satellites generate hundreds of terabytes a day that congest their radio downlinks; process it on orbit and send down answers, not raw pixels. Collision avoidance in a crowded LEO is a real-time problem, SpaceX reported on the order of 300,000 Starlink avoidance maneuvers in 2025, roughly one every two minutes across the constellation, and the loop is tighter if the thinking happens locally (New Scientist). This is the conclusion IEEE Spectrum, ABI Research, and the orbital-edge-computing literature keep independently arriving at (IEEE Spectrum). Edge compute in orbit is a business. Hyperscale-in-orbit-for-Earth is, for now, a thermodynamics problem wearing a business plan.

I want to be honest about what is genuinely contested. The economics are not reconcilable on current data. Google and Starcloud argue cost parity once launch falls below a couple hundred dollars per kilogram in the 2030s, contingent on a Starship cadence nobody has demonstrated (Google Research, Starcloud). ABI Research argues the per-GPU-year cost stays at least ten times terrestrial *regardless* of launch price, because power and thermal hardware are 65 to 70 percent of satellite mass and that fraction does not fall when launch gets cheaper (IEEE Spectrum). Both can cite real numbers. Pick your assumptions and you pick your conclusion, which is usually a sign that the assumptions are the actual content.

The first constraint, then, did not vanish when we left the planet. It changed costume, from a utility bill you negotiate to a physical law you cannot. Hold that shape in mind, because the next two constraints have exactly the same one.

Continue to Part II: Sovereignty Doesn't Subtract in Orbit, It Stacks.

Sources

- Kopp & Lean
- Google Research
- Starcloud
- Thales Alenia Space
- Reuters
- FCC Notice DA 26-113
- NASA / MIT Lincoln Laboratory
- Via Satellite
- Blue Origin
- IEEE Spectrum
- NASA
- NASA Advanced-Radiator Research
- Balandin et al., Nano Letters 2008
- arXiv: Silicon Ring Thermometers
- AIAA / MISSE-16 Thermal Data
- Science Advances
- ACS Photonics
- PubMed
- New Scientist