Rationalizing the Existence of Space-Based Data Centers using Physics

Logan Gruss
Brandeis University
B.S. Physics 2027
April 16, 2026

Abstract

…In this article, I am analyzing whether data centers placed in space could be physically better than data centers on Earth. I focus on the main advantages and limitations by looking at thermodynamics, cooling, solar power, radiation, latency, and launch cost. First, I use Landauer's principle to show that lower temperature lowers the minimum energy required for irreversible computation. Next, I use the Stefan-Boltzmann law to show that a space-based system could reject heat through passive radiation rather than large active cooling systems. After that, I compare solar power in space versus on Earth and show that orbit provides much larger continuous solar input per unit area. However, I then show that radiation in space creates serious problems for normal electronics, and that the speed of light places a strict limit on latency that cannot be engineered away. Finally, I examine launch economics and argue that while ordinary orbital data centers are still heavily constrained, a more specialized system at the Sun-Earth L2 point using RSFQ superconducting logic may represent a physically distinct and much more efficient computing platform. Overall, the physics suggests that space-based computing is not a good replacement for normal terrestrial cloud infrastructure, but it may become a strong option for large, energy-intensive, latency-tolerant workloads.

1. Introduction

…As demand for computation keeps increasing, especially from artificial intelligence workloads, the strain on Earth-based data centers is also increasing. Global data center electricity use has already reached about 415 terawatt-hours per year, which is roughly enough to power about 40 million U.S. homes annually (International Energy Agency, 2025; Gorski and Comstock, 2025). Because every unit of electrical energy going into a data center must eventually be rejected as heat, a 100 MW hyperscale facility ultimately releases roughly 100 MW of waste heat, comparable to about 100,000 space heaters running at once (Advanced Research Projects Agency–Energy, 2022). Because of that, I wanted to analyze whether placing computation in space could actually be better from a physics standpoint.

…To answer that, I approach the problem from first principles. Rather than just listing advantages and disadvantages, I look at the physical laws that govern the system. This includes the thermodynamics of computation, radiative heat transfer, solar energy collection, radiation damage to electronics, signal delay from the finite speed of light, and the economics of getting hardware into orbit. Therefore, the question is not just whether space data centers sound futuristic. The real question is whether the environment of space gives real physical advantages for computing; and what disadvantages exist, as well.

…I have restricted this discussion to near and medium term locations, from low Earth orbit, or LEO, at about 550 km, out to the Sun-Earth L2 point at about 1.5 million km. Interplanetary data centers are interesting, but they are too premature for a serious engineering discussion here. The locations I focus on are far enough to show real differences in physics, but close enough that the discussion still connects to present-day technology.

2. The Thermodynamic Foundation of the Problem

2.1 Landauer's Principle and the Role of Temperature

…Any discussion of data center efficiency should begin with *Landauer's principle*. This gives the minimum amount of energy that must be dissipated when a computer performs a logically irreversible operation, such as erasing one bit of information (Landauer, 1961). The lower bound is:

$$\mathrm{E_{min}} = \mathrm{k_B T} \ln(2)$$

Here, k_B is Boltzmann's constant and T is the absolute temperature of the environment. In simple terms, E_{min} is the smallest possible amount of energy a computer must give off as heat when it permanently erases one bit of information. You can think of it as the absolute minimum energy cost of deleting a single bit. Real computers use much more energy than this in practice, but they still cannot go below this physical limit for an irreversible operation. At about room temperature, $T = 293\ K$, this gives $E_{min} = 2.8\ x\ 10^{-21}\ J$ per erased bit.

…The important point is that this minimum energy scales linearly with temperature. Therefore, if I lower the operating temperature, I also lower the minimum possible energy cost per irreversible operation. Modern CMOS processors still operate far above this limit, so *Landauer's* bound is not yet the main design limit for ordinary silicon. However, superconducting systems such as RSFQ logic operate much closer to this floor, so the temperature dependence becomes much more important.

…In simple terms, every computer must produce at least some heat when it irreversibly processes information, and that minimum gets smaller when the system is colder. This is one of the first reasons space is attractive for data centers. It gives access to a much colder thermal environment than Earth does of course.

2.2 Power Usage Effectiveness as the Industry Benchmark

…A standard industry metric for data center efficiency is Power Usage Effectiveness, or PUE. It is defined mathematically as:

$$\mathrm{PUE} = \frac{\mathrm{P_{total}}}{\mathrm{P_{IT}}}$$

…A PUE of 1 would mean that all power goes directly to computation, with no overhead. In reality, some power must be spent on cooling systems, pumps, fans, and support infrastructure. Modern terrestrial facilities can approach about 1.1 to 1.2 in the best cases, but that still means they must spend substantial energy just to remove heat. Therefore, even a very efficient Earth based data center still carries the cooling burden.

…In space, passive radiative cooling could reduce that overhead significantly. So, the advantage is not just lower electricity cost. The deeper point is that space may remove part of the structural inefficiency that comes from operating dense electronics inside a warm atmosphere.

3. The Thermal Advantage: Radiative Cooling Without Infrastructure

3.1 The Stefan-Boltzmann Law in Orbit

…Any object above absolute zero emits thermal radiation. The total radiated power is given by the Stefan-Boltzmann law:

$$\mathrm{P_{rad}} = \varepsilon\sigma \mathrm{A} \mathrm{T}^4$$

Here, ε is emissivity, $\sigma = 5.67\ x\ 10^{-8}\ Wm^{-2}K^{-4}$ is the Stefan-Boltzmann constant, A is radiator area, and T is the radiator temperature. In space, away from atmospheric effects, radiation is the dominant way to remove waste heat. That means thermal management becomes a geometry problem more than a refrigeration problem. In simpler terms, the main issue is no longer "How do I run enough cooling equipment to remove the heat?" but instead "How large and how well placed do my radiating surfaces need to be so the heat can escape on its own?" In space, you are mostly solving for the size, shape, and placement of the radiators rather than relying on fans, chillers, or liquid-cooling plants.

…Thermal equilibrium occurs when the radiated power equals the waste heat being generated:

$$\mathrm{Q} = \varepsilon\sigma \mathrm{A} \mathrm{T}_{\mathrm{eq}}^4$$

…Solving for equilibrium temperature gives:

$$\mathrm{T_{eq}} = \left(\frac{\mathrm{Q}}{\varepsilon\sigma \mathrm{A}}\right)^{\frac{1}{4}}$$

…The fourth-power dependence is useful for intuition here. If I double the waste heat, the equilibrium temperature does not double; it only increases by the fourth root. For example, if a facility rejects $Q = 10\ MW$ through a radiator array of $A = 5{,}000\ m^2$ with $\varepsilon = 0.9$, the equilibrium temperature is about 358 K, or about 85°C. Here I take $\varepsilon = 0.9$ as a representative value for a high-emissivity radiator surface. This means the radiator emits thermal radiation at about 90% of the rate of an ideal blackbody, which is a reasonable approximation for a well-designed space thermal panel. That is warmer than ideal for silicon (the main component in computer chips), but it gives a baseline. If I double the radiator area to $10{,}000\ m^2$, the temperature drops to about 301 K, which is essentially room temperature, and this happens passively. No chillers, no coolant loops, and no electrical power spent directly on cooling.

3.2 Comparison with Terrestrial Active Cooling

…Terrestrial data centers rely on CRAC units, chilled water loops, cooling towers, and increasingly direct liquid cooling. These systems can consume a large fraction of total facility power. So, on Earth, a noticeable part of the electrical input is spent only to remove heat, not to process data.

…In space, that cooling overhead can approach zero in principle. This makes the thermal case for orbital facilities much stronger than it first sounds. The issue is not just that space is cold. The issue is that the vacuum of space lets a system reject heat without carrying the normal energy penalty of an active cooling plant.

4. The Solar Power Advantage

4.1 Solar Irradiance Above the Atmosphere

…At Earth's orbital distance, the solar irradiance perpendicular to the Sun's rays is approximately:

$$S_0 \approx 1{,}361 \text{ W/m}^2$$

…This is the solar constant measured above the atmosphere (Kopp and Lean, 2011). A spacecraft in orbit can access this weakened flux directly. A ground-based solar panel cannot. On Earth, the effective daily average irradiance is reduced by atmospheric absorption, cloud cover, the changing angle of incidence, and of course the day-night cycle.
A rough expression for the terrestrial daily average is:

$$S_{earth} = S_0 \eta_{atm} (< \cos\theta >)\left(1 - f_{night}\right)$$

Using $\eta_{atm} \approx 0.70, < \cos\theta > \approx 0.50$, and $f_{night} \approx 0.50$ gives:

$$S_{earth} \approx 238 \text{ W/m}^2$$

…The values I chose are rough average factors meant to estimate terrestrial solar input. I used $\eta_{atm} \approx 0.70$ to represent atmospheric losses, $< \cos\theta > \approx 0.50$ to represent the average reduction from non-perpendicular sunlight given it is a vector, and $f_{night} \approx 0.50$ to account for the fact that Earth is in darkness for about half of each rotation. Together, these reduce the above-atmosphere solar constant to an average ground-level value of about $238\ W/m^2$.
…By contrast, a tracked space-based panel in high orbit can remain near $S_{space} \approx 1{,}361 \text{ W/m}^2$. That is about 5.7 times higher incident power per unit area than the terrestrial average. Even after accounting for long-term degradation in the radiation environment, orbit still offers a much larger continuous solar energy supply per square meter.

4.2 Power Generation at Scale

…Modern space-grade multi-junction solar cells can reach efficiencies of about 30 to 40 percent (Bett et al., 2009). For a 1 km^2 array with $\eta_{panel} \approx 0.35$, the generated power is:

$$P = \eta_{panel} S_0 A = (0.35)(1{,}361)(10^6) \approx 476 \text{ MW}$$

…That is on the order of a mid-scale terrestrial power plant, but with no fuel cost and continuous solar input. Of course, the array mass still matters, and that mass must be launched many miles at escape velocity. However, the basic physical point remains that space offers more continuous solar power per unit area, and that directly helps any computing platform that can tolerate being off the ground.

5. The Radiation Environment: The Dominant Obstacle

5.1 Sources and Characterization

…The strongest argument against ordinary space-based data centers is radiation. The space environment is much harsher for electronics than the terrestrial environment (Baumann, 2005). There are three main radiation sources we have to consider.

…First, galactic cosmic rays, or GCRs, are high-energy protons and heavier ions that arrive from outside the solar system. A representative flux at 1 AU is about 4 particles $cm^{-2}s^{-1}sr^{-1}$. Integrated over all directions, a 1 cm^2 exposed device sees approximately $\Phi_{GCR} \approx 4\pi \times 4 \approx 50\, particles\ cm^{-2}\ s^{-1}$. My derivation for the flux is as follows, for those uninterested in the derivation, skipping these calculations does not limit the reader's ability to understand my conclusion:

…Let the galactic cosmic ray intensity be:

$$I_{GCR} \approx 4\, particles\ cm^{-2}\ s^{-1}\ sr^{-1}$$

…The sr^{-1} means "per steradian," so to convert this into the total flux from all directions, I want to integrate over a gaussian sphere with respect to r, θ, ϕ. However, that is not particularly efficient, so instead I will convert to spherical coordinates, starting with converting *x, y, and z* (this portion is relevant to bolstering my decision to using a gaussian surface to sum the total I_{GCR} in a specific region of space; it is not relevant to those uninterested in the derivation or in agreeance from prior understanding):

$$x = rsin(\theta)\, cos(\phi)\,, \qquad y = rsin(\theta)cos(\theta), \qquad z = rcos(\theta)$$

Here, *r* is the radial distance outward from the center of the sphere, θ is the angle measured down from the positive z-axis, creating the vertical slice down the center of the sphere, and ϕ is the angle in the xy-plane, creating a horizontal slice.

With that being said, we can now look at a small change in our coordinates:

$$A\ small\ \ change\ in\ r\ gives\ length\ dr$$

$$A\ small\ change\ in\ \theta\ at\ a\ fixed\ r\ sweeps\ out\ an\ arc\ of\ length\ rd\theta$$

$$A\ small\ change\ in\ \phi\ at\ a\ fixed\ r\ and\ \theta\ sweeps\ out\ an\ arc\ on\ a\ circle\ of\ radius\ rsin(\theta)$$

…Therefore, the small differential volume *(dV)* in spherical coordinates is as follows:

$$dV = (dr)(rd\theta)(rsin\theta d\phi) = r^2 sin\theta dr d\theta d\phi$$

This gives us the uninviting full three-dimensional volume integral in spherical coordinates:

$$\iiint f(r,\theta,\phi)dV = \iiint f(r,\theta,\phi)r^2 sin\theta dr d\theta d\phi$$

…However, for the gaussian surface I am not actually interested in the volume, but rather the surface area. The intensity of the intergalactic cosmic rays is defined as a directional flux per unit solid angle, so the relevant aspect is not the radial thickness of the sphere but the angular portion. In order to isolate the angular part, I can restrict the geometry to a sphere of constant radius. This would mean the change in r is now 0 and therefore *dr*=0. So, the differential surface now becomes:

$$dA = (rd\theta)(rsin\theta d\phi) = r^2 sin\theta d\theta d\phi$$

By definition, solid angle is surface area on a sphere divided by the square of the radius:

$$d\Omega = \frac{dA}{r^2}$$

Now substituting dA from above:

$$d\Omega = \frac{(r^2 sin\theta d\theta d\phi)}{r^2} = sin\theta d\theta d\phi$$

Therefore, the full spherical-coordinate measure $r^2 sin\theta d\theta d\phi$ separates the dV into a radial part and angular part:

$$dV = r^2 dr d\Omega$$

…Finally, I can now write the flux as:

$$\Phi_{GCR} = \int I_{GCR}\; d\Omega = \int_0^{2\pi} \int_0^{\pi} I_{(GCR)}\,(\theta, \phi) sin(\theta) d\theta d\phi$$

…Given I am assuming the I_{GCR}

$$\Phi_{GCR} = I_{GCR} \int d\Omega = I_{(GCR)} \int_0^{2\pi} \int_0^{\pi} sin(\theta) d\theta d\phi$$

$$\int_0^{\pi} sin\theta d\theta = 2, \quad \int_0^{2\pi} d\phi = 2\pi$$

Therefore,

$$I_{GCR} \int_{\Omega} d\Omega = \mathrm{I_{GCR}}(4\pi)$$

$$\boxed{\Phi_{GCR} = I_{GCR}(4\pi)}$$

…Now inserting the known value for intensity of galactic cosmic rays in our chosen region,

$$\Phi_{GCR} = I_{GCR}(4\pi)$$

$$\approx (4\, particles\; cm^{-2}\, s^{-1}\, sr^{-1})(4\pi\; sr)$$

$$= 16\pi\; particles\; cm^{-2}\, s^{-1}$$

$$\approx 50\, particles\ cm^{-2}\ s^{-1}.$$

…In conclusion my final value for the flux of intensity of galactic cosmic rays is as follows:

$$\boxed{\Phi_{GCR} \approx 4\pi \times 4 \approx\ 50\, particles\ cm^{-2}\ s^{-1}}$$

…Now reminding ourselves that this was just one instance of particle induced damage, we can now look briefly at two others. The second being solar energetic particles, or SEPs, which are produced during solar flares and coronal mass ejections. These events can increase the local radiation environment by 2 to 5 orders of magnitude for periods of hours to days, and they are not fully predictable (Shea and Smart, 1990).

… Third, the Van Allen belts trap energetic charged particles inside Earth's magnetic field (Van Allen, 1959). Modern NASA descriptions identify two main belts, an inner belt extending from about 3,200 to 16,000 km altitude and an outer belt extending from about 13,000 to 38,000 km, with the outer belt sometimes reaching about 57,000 km depending on solar conditions. This is not just a qualitative hazard. In NASA radiation modeling with 5mm of aluminum shielding, an ISS-like low Earth orbit accumulates about $430\,\frac{\text{rad}}{\text{yr}}$, while geostationary orbit accumulates about $5930\,\frac{\text{rad}}{\text{yr}}$. So, even with millimeter-scale shielding, the annual ionizing dose in GEO is about fourteen times larger than in a moderate LEO environment, which shows mathematically that trapped-belt radiation is a persistent design constraint rather than a small background effect (Van Allen, 1959; Wrbanek and Wrbanek, 2020; Alena, 2022).

5.2 Single Event Upsets: Quantitative Analysis

…For digital electronics, one of the biggest concerns is the Single Event Upset, or SEU. This happens when a charged particle deposits enough ionization energy in a memory cell to flip its stored value from 0 to 1 or from 1 to 0. In more common terms, the SEU rate is set by both the radiation environment and the vulnerability of the electronics. A larger particle flux means more chances for strikes, while a larger upset cross-section means each strike is more likely to flip a bit. Therefore, harsher radiation or more sensitive memory both increase the rate of random errors. The basic relation from a mathematical perspective is the following:

$$SEU\ rate = \Phi\sigma_{SEU}$$

Here, Φ is particle flux and σ_{SEU} is the effective upset cross-section per bit. For modern DRAM (the main short-term working memory used by computers to store data while they are operating), σ_{SEU} is roughly $10^{-13}\ to\ 10^{-12}\,\frac{cm^2}{bit}$. For a 1 GB memory array, this gives a representative upset rate on the order of tens of upsets per day:

$$SEU\ rate \approx (50 \times 10^{-13})(10^{9})\ \approx\ 50\, upsets/day$$

…At modern server-scale capacities, the problem becomes much worse. A server with hundreds of gigabytes of DRAM could see thousands of bit flips per day if left unmitigated. ECC memory helps with single-bit errors, but it does not completely solve multi-bit upsets caused by one energetic track. Therefore, for conventional computing hardware, radiation is not a minor nuisance. It is a major design constraint.

5.3 Total Ionizing Dose Degradation

…Radiation does not just cause temporary bit flips. Over time, it also causes cumulative damage in semiconductors. This is described by the total ionizing dose, or TID. At geostationary orbit, a representative dose rate behind 2 mm of aluminum shielding is:

$$\frac{dD}{dt} \approx 10^4 \rightarrow 10^5 \frac{rad}{year}$$

…Commercial CMOS often tolerates only about $10^2 \rightarrow 10^3$ rad before major parameter drift appears, meaning the electrical behavior of the transistors begins to shift away from its intended values under radiation exposure. That means unshielded commercial hardware could fail very quickly in higher-radiation orbits. Radiation-hardened components can survive much more dose, but they are slower, more expensive, and often lower density than commercial chips. So, radiation creates a direct performance-cost penalty that terrestrial data centers largely do not face.

6. Latency: The Inescapable Speed-of-Light Constraint

6.1 The Round-Trip Delay

…No design improvement can avoid the finite speed of light. If a satellite is at altitude h, then the minimum round-trip latency between ground and orbit is:

$$\tau_{min} = \frac{2h}{c},$$

where c is the speed of light. This is only the ideal physical lower bound. Real systems are always slower than this because they also include routing delay, protocol delay, switching time, and other losses.

…The main point of the table I have below is to show how quickly latency increases as orbital distance increases. In other words, the farther the compute system is from Earth, the longer every back-and-forth exchange must take, even in the best possible case. That makes low orbits much better for interactive applications, while higher orbits become progressively less suitable for tasks that depend on fast response times.

Orbit	Altitude (h)	Minimum round-trip latency	Usability for real-time applications
LEO (Starlink-like shell)	550 km	3.7 ms	Excellent; comparable to terrestrial CDN-scale latency
MEO (GPS altitude)	20,200 km	135 ms	Marginal; noticeable in voice and video

GEO	35,786 km	238 ms	Poor for highly interactive applications

…Human reaction time is on the order of 150 to 250 ms (Kosinski, 2008). So, a geostationary round-trip delay already approaches that scale. That means applications such as competitive gaming, real-time financial trading, live audio processing, and immersive interactive systems become poor candidates for GEO-based compute, not because of weak engineering, but because relativity sets the timing limit.

6.2 Implications for Architecture

…This creates a clear architectural split. LEO can support more latency-sensitive workloads, but it requires large constellations and frequent handoffs, and it does not provide the same uninterrupted geometry for power and thermal design that deeper-space locations can provide. GEO and L2 are much better for steady infrastructure and sunlight access, but their latency rules out interactive services. Therefore, any realistic space-based compute system has to be matched to workloads that do not care about seconds of delay.

7. Launch and Operations Economics

7.1 Current and Projected Launch Costs

…The economics of any space-based infrastructure are controlled first by launch cost. A modern 1U server chassis has a mass of about 20 kg once I include the chassis, power distribution, and cabling fractions. A standard 42U rack therefore has a mass of about 840 kg.

…At a representative reusable Falcon 9 price of about \$1,500 per kilogram, the launch cost per rack is:

$$C_{launch}\ (per\ rack) = 840\ kg \times \$1{,}500\ kg^{-1} = \$1{,}260{,}000$$

…For a hyperscale facility with about 10^5 servers, or about 2,380 racks, the launch cost becomes:

$$C_{total} \approx 2{,}380 \times \$1.26M \approx \$3.0\ billion$$

…That is before adding structural enclosures, shielding, radiator arrays, power systems, and thermal hardware. Once those are included, the real launch bill could easily rise into the \$6 to \$10 billion range. At current prices, that makes a normal orbital hyperscale data center economically very difficult to justify.

7.2 The Starship Inflection Point

…Projected Starship pricing changes the picture. If launch cost drops to around \$200 per kilogram, then the same rack would cost:

$$C_{launch\,projection}\ \ (per\ rack) = 840 \times \$200 = \$168{,}000$$

…The total launch cost for the same 2,380-rack facility would be:

$$C_{total\ projection} \approx 2{,}380 \times \$168{,}000 \approx \$400\ million$$

…That number is still large, but it is now within the range of terrestrial data center construction projects. So, if heavy-lift reusable launch really reaches that price range, the question stops being purely science fiction and becomes a legitimate engineering tradeoff.

7.3 Hardware Replacement and the Upgrade Problem

…Even if launch becomes cheaper, space hardware still faces an upgrade problem. Earth-based data centers can refresh servers every 3 to 5 years. Space-based systems cannot be upgraded that easily. They must be designed for a much longer life, often 10 to 15 years. This matters because computing hardware improves quickly, and a facility that cannot refresh may become obsolete before the end of its physical lifetime.

8. An Original Conjecture: The Thermal-Computational Goldilocks Orbit

8.1 The Limitation of the Wrong Question

…Up to this point, my analysis gives a mixed conclusion. Space offers strong thermal and power advantages, but it also suffers from radiation, latency, launch cost, and hardware replacement problems. So, if I ask whether a space data center should directly compete with a normal cloud data center on Earth, the answer as of today, April 4, 2026, is mostly no.

…However, I think that is the wrong question. A space-based computing platform should not try to win on ordinary cost per rack, normal latency, or quick upgrade cycles. It should try to win on operations per joule, long-duration throughput, and the ability to realize a physical computing substrate that is extremely hard or impossible to operate efficiently on Earth.

8.2 Rapid Single Flux Quantum Logic: A Brief Primer

…Rapid Single Flux Quantum, or RSFQ, logic is a superconducting digital technology in which information is carried by single quanta of magnetic flux moving through Josephson junction circuits (Likharev and Semenov, 1991). Unlike CMOS, which uses voltage states, RSFQ uses superconducting switching events.

…A characteristic switching energy scale is:

$$E_{switch} = \Phi_{B_0} I_c \approx 2 \times 10^{-19}\ J\ per\ junction$$

Here, Φ_{B_0} is the magnetic flux quantum and I_c is the junction critical current. This energy scale is roughly 10^5 times smaller than the switching energy of modern CMOS logic (Holmes et al., 2013). However, RSFQ requires cryogenic operation below the superconducting transition temperature of niobium, and in practice systems are operated near 4 K. In other words, RSFQ is attractive because each switching event uses extremely little energy compared with ordinary CMOS. However, that advantage only exists if the system is kept extremely cold so the hardware can remain superconducting. This means RSFQ is not just a better computer; it is a technology that only becomes practical in a cryogenic environment, such as that provided by cold dark space.

8.3 The L2 Cryogenic Advantage

…The Sun-Earth L2 point is about 1.5 million km from Earth on the anti-solar side and is already used for major space observatories such as The James Webb Space Telescope, JWST (Gardner et al., 2006). With a multilayer sunshield, a spacecraft in this threshold can passively reach about:

$$T_{shield} \approx 40\ K$$

…The cosmic microwave background (left over radiation from the early universe) also provides a natural far-field sink (distant thermal background which absorbs outgoing heat radiation from the focused object) at:

$$T_{CMB} = 2.725\ K$$

…If I start from a passive 40 K environment and cool RSFQ processors down to 4 K, the ideal Carnot refrigeration penalty is:

$$COP^{-1} = \frac{T_{hot} - T_{cold}}{T_{cold}} = \frac{(40\ -\ 4)}{4} = 9\ W$$

On Earth, cooling from 300 K down to 4 K gives:

$$COP^{-1}(terrestrial, 300\ K\ \rightarrow\ 4\ K) = \frac{300\ -\ 4}{4} = 74\ W$$

…So, in an idealized comparison, the space system needs only about 9 W of refrigeration work per watt of processor dissipation, while the terrestrial cryogenic system needs about 74 W. That is about an 8 times advantage in refrigeration burden before I even include the intrinsic switching-energy advantage of RSFQ over CMOS.

8.4 The Energy Economic Argument

…To show why this matters, I can compare the energy cost of a very large computation. Consider 10^{24} floating-point operations, which is a reasonable order-of-magnitude estimate for a major AI training run in the 2030 to 2035 period.

…For terrestrial CMOS with energy per operation of about 10^{-14} J:

$$E_{CMOS} = 10^{24} \times 10^{-14} = 10^{10}\ J$$

…That is about 277,778 MWh, which at \$0.10/kWh corresponds to about \$27.8 billion in electricity alone.

…For RSFQ with energy per operation of about $10^{-19}\ J$:

$$E_{RSFQ} = 10^{24} \times 10^{-19} = 10^{5}\ J$$

…That is only about 0.028 MWh before refrigeration overhead. Even if I multiply by the L2 refrigeration penalty, the effective energy remains only about 0.28 MWh. At that point, the dominant cost shifts away from electricity and toward hardware and launch. So, the argument for an L2 RSFQ facility is not that it is a slightly better data center. The argument is that it may operate in an entirely different energy regime.

8.5 Suitable Workloads and the Latency Exemption

…The round-trip delay to L2 is on the order of 10 seconds, so interactive applications are eliminated immediately. However, there are important workloads that do not depend on interactivity:

- Large-scale AI training runs those last days to months, where a delay of seconds is negligible compared with the full job duration.
- Genomic analysis and protein folding, which are heavily compute-intensive and can be handled through batched input and output.
- Climate and Earth system modeling, where simulations run for very long times and communication latency is not the limiting factor.
- Cryptographic workloads such as zero-knowledge proof generation and large post-quantum computations.
- Archival and cold storage systems, where retrieval delays of seconds are already acceptable in some terrestrial systems.

…For these workloads, the important metric is not latency but operations per joule per dollar over the facility lifetime. On that metric, an L2 solar-powered RSFQ facility may have no direct terrestrial equivalent. So, my conjecture is not that I should move a normal data center into space. It is that I should build a superconducting computing platform in space that is physically difficult to realize efficiently on Earth.

9. Discussion and Conclusion

…The physics of space-based data centers shows a sharp tradeoff. On one side, space offers really thermal and power advantages: passive radiative cooling, strong continuous solar input, and access to very low environmental temperatures. On the other side, ordinary hardware in space faces radiation damage, nontrivial shielding requirements, strict latency limits, high launch cost, and the inability to refresh hardware easily, risking it becoming obsolete prior to hardware expiration.

…Because of that, I do not think space-based computing should be treated as a near-term replacement for our current terrestrial cloud infrastructure. For ordinary interactive services, Earth is still the better place to compute. However, I do think space may become a strong complement for a narrower set of workloads that are energy-intensive, long-running, and insensitive to delay. Not to mention the clear advantages if we are computing on say, the Moon or Mars.

…The L2 RSFQ conjecture is speculative, but it is grounded in real physical arguments. James Webb Space Telescope has already shown that passive temperatures near 40 K are achievable at L2. RSFQ logic has an enormous switching-energy advantage over CMOS. Carnot analysis shows that the refrigeration penalty is far smaller when starting from 40 K than from 300 K. When I combine these points, the result is a plausible path toward a computing platform that could differ from terrestrial systems by orders of magnitude in energy efficiency.

…There are still serious open questions. These include the radiation tolerance of RSFQ circuits at L2, the mass and reliability of the cryocoolers, long-duration servicing strategy, and whether the economic crossover with terrestrial systems will really occur at future launch prices. Even so, none of those questions obviously closes the door.

…So, my final conclusion is that ordinary space data centers are still limited, but a specialized cryogenic superconducting compute platform in deep space may represent a genuinely new class of energy-efficient infrastructure.

Works Cited

Alena, Richard L. *Mission Radiation Environment Modeling and Analysis: Avionics Trade Study for GCD Rad-Neuro Project*. NASA/TM-20220011775, National Aeronautics and Space Administration, 2022.

Advanced Research Projects Agency–Energy. *Cooling Operations Optimized for Leaps in Energy, Reliability, and Carbon Hyperefficiency for Information Processing Systems (COOLERCHIPS).* U.S. Department of Energy, 22 Sept. 2022.

Baumann, Robert C. "Radiation-Induced Soft Errors in Advanced Semiconductor Technologies." IEEE Transactions on Device and Materials Reliability, vol. 5, no. 3, 2005, pp. 305-316.

Bernstein, K., et al. "Device and Architecture Outlook for Beyond CMOS Switches." Proceedings of the IEEE, vol. 98, no. 12, 2010, pp. 2169-2184.

Bett, Andreas W., et al. "III-V Solar Cells under Monochromatic Illumination." Proceedings of the 33rd IEEE Photovoltaic Specialists Conference, 2009, pp. 1-5.

Fixsen, D. J. "The Temperature of the Cosmic Microwave Background." The Astrophysical Journal, vol. 707, no. 2, 2009, pp. 916-920.

Gardner, Jonathan P., et al. "The James Webb Space Telescope." Space Science Reviews, vol. 123, no. 4, 2006, pp. 485-606.

Gorski, Alex, and Owen Comstock. “Residential Electric Bills in Hawaii and Connecticut Are Twice Those in New Mexico, Utah.” *Today in Energy*, U.S. Energy Information Administration, 12 May 2025.

Holmes, D. S., A. L. Ripber, S. A. Cybart, and R. C. Dynes. "Energy Scale of a Josephson Junction." Physical Review Letters, vol. 111, 2013, p. 167001.

International Energy Agency. “Executive Summary.” *Energy and AI*, 10 Apr. 2025.

International Energy Agency. Electricity 2024: Analysis and Forecast to 2026. IEA, 2024.

Kopp, G., and J. L. Lean. "A New, Lower Value of Total Solar Irradiance: Evidence and Climate Significance." Geophysical Research Letters, vol. 38, no. 1, 2011, pp. 1-7.

Kosinski, Robert J. "A Literature Review on Reaction Time." Clemson University, 2008.

Landauer, Rolf. "Irreversibility and Heat Generation in the Computing Process." IBM Journal of Research and Development, vol. 5, no. 3, 1961, pp. 183-191.

Landis, Geoffrey A., et al. "Space Solar Power: A Fresh Look." AIAA SPACE 2011 Conference & Exposition, 2011, pp. 1-8.

Likharev, Konstantin K., and Vasili K. Semenov. "RSFQ Logic/Memory Family: A New Josephson-Junction Technology for Sub-Terahertz-Clock-Frequency Digital Systems." IEEE Transactions on Applied Superconductivity, vol. 1, no. 1, 1991, pp. 3-28.

Shea, M. A., and D. F. Smart. "A Summary of Major Solar Proton Events." Solar Physics, vol. 127, no. 2, 1990, pp. 297-320.

SpaceX. "Falcon 9." SpaceX, 2023.

The Green Grid. "PUE: A Comprehensive Examination of the Metric." The Green Grid, 2012.

Uptime Institute. Global Data Center Survey Results 2022. Uptime Institute, 2022.

Van Allen, James A. "The Geomagnetically Trapped Corpuscular Radiation." Journal of Geophysical Research, vol. 64, no. 11, 1959, pp. 1683-1689.

Wrbanek, John D., and Susan E. Wrbanek. *Space Radiation and Impact on Instrumentation Technologies*. NASA/TP-2020-220002, National Aeronautics and Space Administration, 2020.