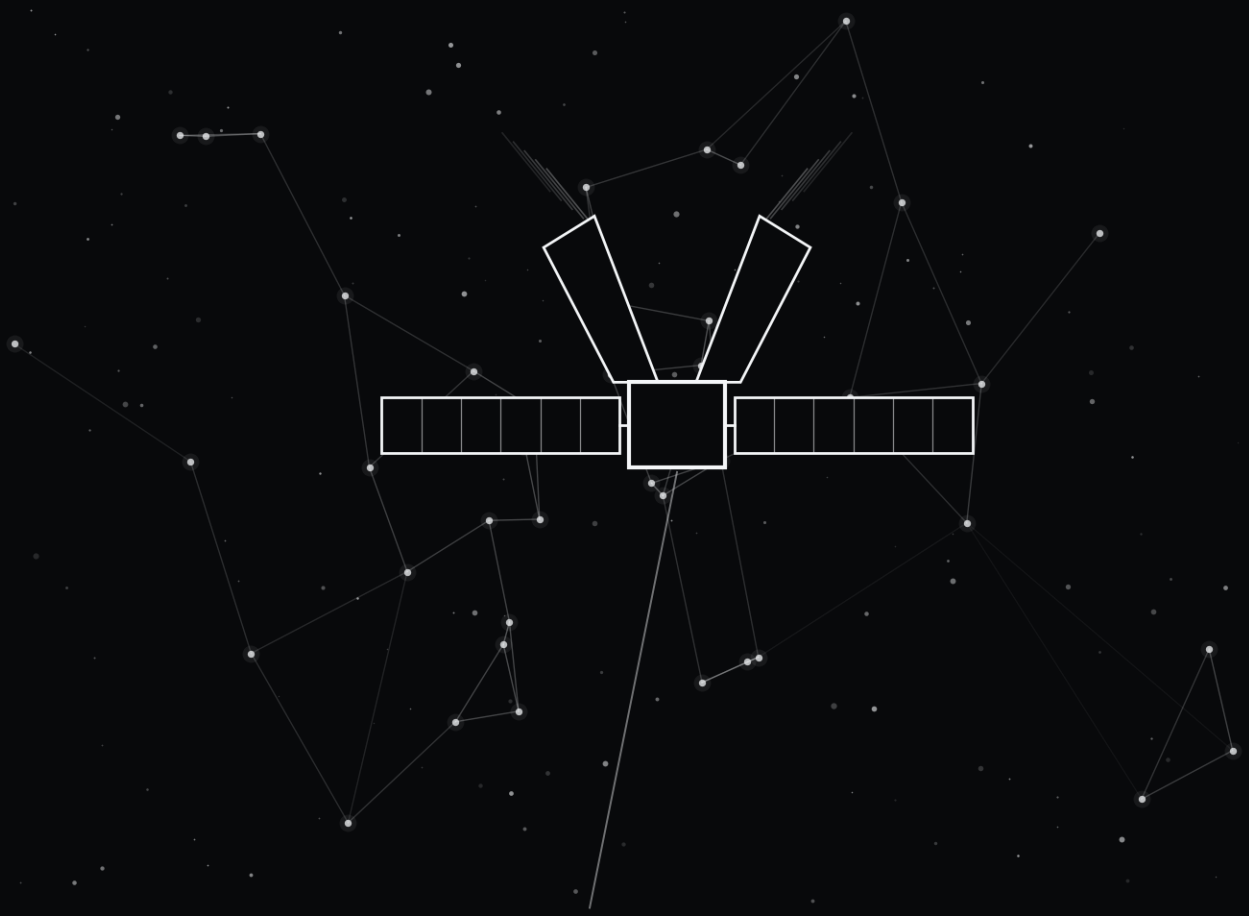
THE DATA CENTERS IN LEO

A GUIDE INTO ORBITAL INTELLIGENCE



IBRAHIM QUABBOUA

STARCAP

THE DATA CENTERS IN LEO

A guide into orbital intelligence

Ibrahim Quabboua

StarCap

First edition, revised after red-team review — 27 July 2026

CONTENTS

Foreword

Foreword

Disclosure, before you read anything else

A note on method

PART I — WHY ORBIT, WHY NOW

Chapter 1 — The wall Earth hit

Chapter 2 — What SpaceX actually changed

Chapter 3 — The inversion

PART II — THE PHYSICS YOU CANNOT ARGUE WITH

Chapter 4 — Power

Chapter 5 — Heat

Chapter 6 — Radiation

Chapter 7 — Bandwidth

Where Part II leaves us

PART III — ANATOMY OF A COMPUTE SATELLITE

Chapter 8 — Walking the machine

PART IV — THE SECTOR SKELETON

Chapter 9 — Launch and cargo

Chapter 10 — Power

Chapter 11 — Thermal

PART IV — THE SECTOR SKELETON (continued)

Chapter 12 — Radiation-tolerant compute

Chapter 13 — The compute platform

Chapter 14 — Communications and the ground segment

PART IV — THE SECTOR SKELETON (concluded)

Chapter 15 — In-space assembly and servicing

Chapter 16 — Debris, traffic and insurance

Chapter 17 — Outside the United States

PART V — DOES IT PENCIL

Chapter 18 — The cost model, and who would buy the output

Chapter 19 — What Earth gets back

PART VI — THE HORIZON

Chapter 20 — 2030, 2035, 2040

Chapter 21 — What would delay this, and what would prove it wrong

Chapter 22 — How to use this book

Closing

APPENDICES

Appendix A — Constants, formulas and conventions

Appendix B — Glossary

Appendix C — Sources

Appendix D — Disclosure

Appendix E — The companies, layer by layer

Foreword

[PLACEHOLDER — TO BE WRITTEN BY AN INVITED CONTRIBUTOR]

This page is reserved for a foreword by an investor active in the space sector.

Suggested brief to send with the invitation: eight hundred to twelve hundred words on why a book about one narrow vertical is more useful to an investor than another survey of the space economy, and on what the reader should do with a set of falsifiers rather than a set of predictions. The contributor should be given the full manuscript and complete freedom, including freedom to disagree with its conclusions in print — a foreword that argues with Chapter 21 would be worth more than one that endorses it.

Nothing may be printed on this page until a named person has written and approved it.

Foreword

> **[PLACEHOLDER — TO BE WRITTEN BY AN INVITED CONTRIBUTOR]**
>
> This page is reserved for a foreword by an engineer or academic working on spacecraft thermal, power or radiation systems.
>
> *Suggested brief:* six hundred to a thousand words on whether the physics in Parts II and III is stated fairly, where a practising engineer would push back, and what the reader should be most sceptical of. A technical foreword that names two or three places where the book is optimistic will do more for its credibility than any endorsement.
>
> *Nothing may be printed on this page until a named person has written and approved it.*

Disclosure, before you read anything else

The author is the founder of StarCap, a fund that holds or intends to hold positions in private companies across most of the layers described in Part IV. **Several companies discussed favourably in this book are held. This should be assumed to bias the book, and the bias is structural rather than occasional.**

Three specific places to be suspicious:

- The layers the fund is concentrated in are described as “structural.” A reader is entitled to ask whether the structural argument came first.
- The layer with no investable company (thermal, Chapter 11) is described as “the finding.” That is a convenient way to describe an absence.
- The layer that cannot be bought at venture scale (radiation-tolerant compute, Chapter 12) is described as “solved by others,” which is also the reason not to own it.

I believe each of those judgements is correct on its merits, and each sector chapter states its reasoning so you can check. But the correlation between what the analysis praises and what the author owns is real, and hiding it in an appendix would have been the wrong choice. The full disclosure is in Appendix D.

The strongest correction available to a reader is Chapter 21, which is the case against everything argued here, and Chapter 18, which reaches a conclusion notably worse for the author’s own positions than earlier drafts did.

A note on method

This book was written in the open in the sense that matters: every number in it is either cited to a public source or derived in front of the reader from constants listed in Appendix A. Where a derivation contradicted something stated earlier in the book, the contradiction was printed and the earlier figure corrected rather than quietly revised — Chapters 2 and 8 are the clearest example.

This edition was revised after the manuscript was deliberately attacked — a red-team pass in which the strongest available objections were written out and then answered in the text rather than in a rebuttal. Six substantive things changed as a result: a demand section was added because the book had none; behind-the-meter gas generation replaced the grid queue as the benchmark competitor; a cost of capital was introduced, which raised costs more than any engineering variable; availability, fleet attrition and performance fade entered the model; the manufacturing base case was moved to the less flattering figure; and revenue is now stress-tested alongside cost. The conclusions got narrower and, in one important way, clearer: the book now separates **today’s cost** from **tomorrow’s curve**, computes both, and reaches a date rather than a verdict. Where a revision contradicted an earlier chapter, the earlier chapter was corrected and the correction left visible.

Four numbers in this book are load-bearing and inadequately sourced: what a large spacecraft costs per kilogram to manufacture, what an orbital solar array costs per watt, the true capacity of the space insurance market, and the per-orbit debris flux at the reference altitude. Each is flagged where it appears. A reader who resolves any one of them will know something the author does not, and I would like to hear about it.

The figures were generated from the equations and assumptions stated in the text, not adapted from other publications. Where a figure is a schematic rather than a model output, it says so on the figure.

Company-level facts are current as of **27 July 2026** and will not stay current. Appendix E is written to age as slowly as possible: it contains judgements about structure and no valuations, round sizes or dated financials, all of which move quarterly and none of which belong in a book.

PART I — WHY ORBIT, WHY NOW

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part I of VI.

There is a version of this book that opens with a rendering: a silver slab of solar panels the size of a football field, wings of radiators glinting, a data centre in orbit above a blue Earth. That version is a pitch deck.

This one opens with a number instead. In 2023, a rack of AI accelerators drew about forty kilowatts. In 2026, the rack shipping to hyperscalers draws somewhere between one hundred and ninety and two hundred and thirty. The architecture named for 2027 is specified at six hundred, and the roadmap past it says one megawatt.[1]

A megawatt. Into a cabinet the size of a wardrobe.

Everything else in this book follows from asking what happens next to a civilisation that wants a great deal of that, and cannot get it fast enough. If you understand only the first three chapters, you will still understand the argument. The remaining hundred pages are the engineering of it.

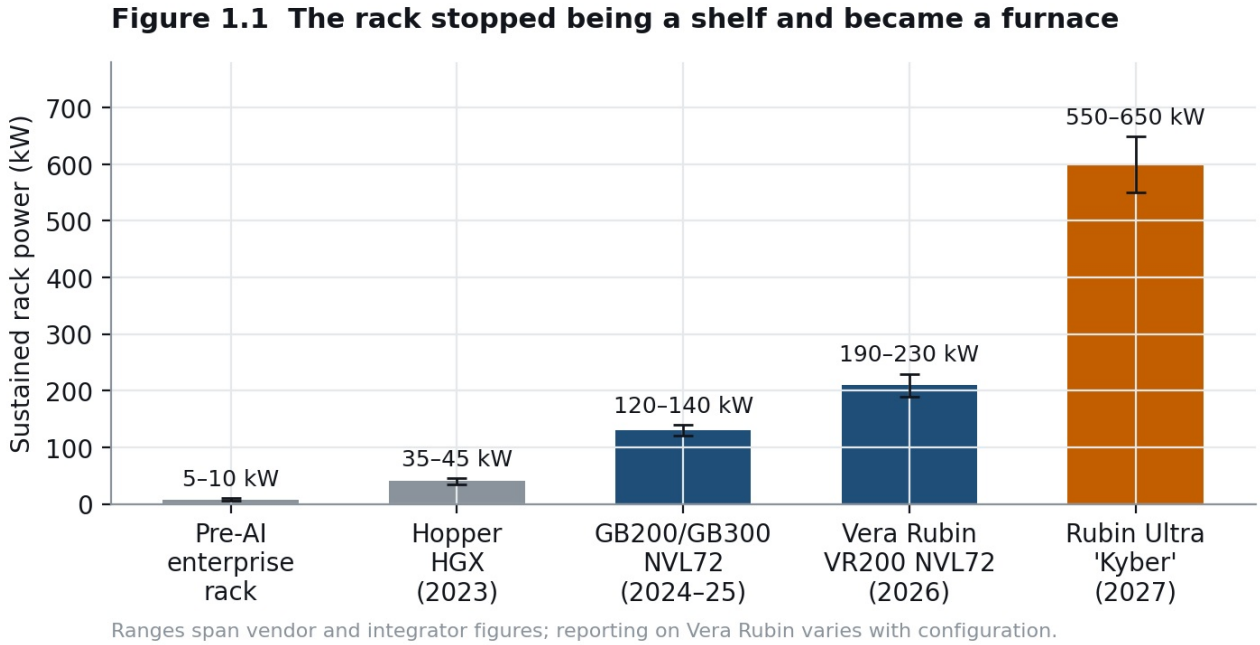
Chapter 1 — The wall Earth hit

What Nvidia did

For thirty years, the unit of computing was the chip. Nvidia's quiet structural achievement of the last three years was to move that unit up a level: the rack became the computer. Seventy-two accelerators, wired together by a switch fabric fast enough that software treats them as one enormous processor, cooled by liquid because air stopped working.

That change was a performance decision. Its consequences were an energy decision.

Figure 1.1 — *The rack stopped being a shelf and became a furnace*



Read the figure as a slope, not as five numbers. A pre-AI enterprise rack drew five to ten kilowatts, which is a large domestic oven. The Hopper generation reached roughly forty. Blackwell's NVL72 landed at one hundred and twenty to one hundred and forty, and reporting on the 2026 Vera Rubin generation puts a fully configured rack between one hundred and ninety and two hundred and thirty kilowatts depending on how it is specified.[1] The next rack architecture is quoted at six hundred, with megawatt-class densities named beyond it.

Two things break at that slope, and both matter later in this book.

The first is cooling. Air carries roughly three thousand times less heat per unit volume than water, which is why liquid cooling stopped being an exotic option and became the only option somewhere between forty and one hundred and twenty kilowatts per rack.[2] An enormous amount of the world's existing data centre floorspace cannot host this hardware at all without being rebuilt.

The second is that essentially all of that power leaves the building as heat. This is not an engineering failure; it is thermodynamics doing its job. A data centre is best understood as a machine that converts electricity into two products: tokens, and warm air. The tokens are the ones you sell.

> **PRIMER — Why "it all becomes heat" is exactly true** Electrical energy entering a computer does work: it charges and discharges transistor gates, drives current down copper, spins fans. None of that work is stored. There is no reservoir of potential energy left in a server at the end of the day. By conservation of energy, whatever went in as watts must come out, and with the trivial exception of the light leaving the fibre optics and the small amount of radio energy escaping, it all comes out as heat. A 200 kW rack is a 200 kW heater that happens to think. Hold onto this: in Chapter 5 it becomes the single hardest constraint in orbit.

What Google did

If Nvidia raised the numerator, Google spent fifteen years shrinking the denominator.

The industry measures this with PUE — power usage effectiveness, the ratio of total facility power to the power that actually reaches the computers. A PUE of 2.0, common in the 2000s, means every watt of computing carried a second watt of chillers, pumps, transformers and lighting. Google drove its fleet toward roughly 1.1, and in doing so taught the entire industry how: hot-aisle containment, custom power delivery, machine-learning control of the cooling plant, and siting facilities where the power and the climate are favourable rather than where the real estate is convenient.

Google also did something less discussed. It designed its own accelerator, the TPU, and it designed it at rack scale before rack scale was fashionable. That matters to us in Chapter 12, because a company that designs its own silicon can also decide to design silicon that survives orbit — and it has.

But notice what fifteen years of efficiency work implies for the future. **If your PUE is already 1.1, perfecting your cooling entirely buys you nine percent.** The overhead has been squeezed out. What remains is the load itself, and the load is growing at the slope in Figure 1.1. Efficiency is no longer where the answer is.

The wall

So the industry turned to the grid, and found a queue.

The specifics are worth carrying with you, because the whole orbital argument rests on them being true and durable:

- Projects reaching commercial operation in the PJM territory in 2025 had spent an average of about eight years in the interconnection queue.[3]
- New high-capacity connections in the dense hubs — Northern Virginia, Phoenix, Dallas — are quoted at four to seven years, and EPRI's 2026 analysis warns that a data centre relying only on the grid could face up to ten years to energisation in some regions.[4]
- Large power transformers, which every new substation needs, run two to four years of lead time.[5]
- EPRI's same analysis puts US data centres at nine to seventeen percent of national electricity consumption by 2030.[4]

Notice the shape of this problem. It is not that electricity is expensive. In much of the United States, industrial power is cheap by world standards. It is that electricity is **not available at a specific place on a specific date**, and no amount of money compresses a transformer factory's order book or a transmission permitting process into the eighteen months an AI roadmap runs on.

This is the sentence to keep:

> **The binding constraint on artificial intelligence is no longer silicon, and it is no longer the price of power. It is the queue for power.**

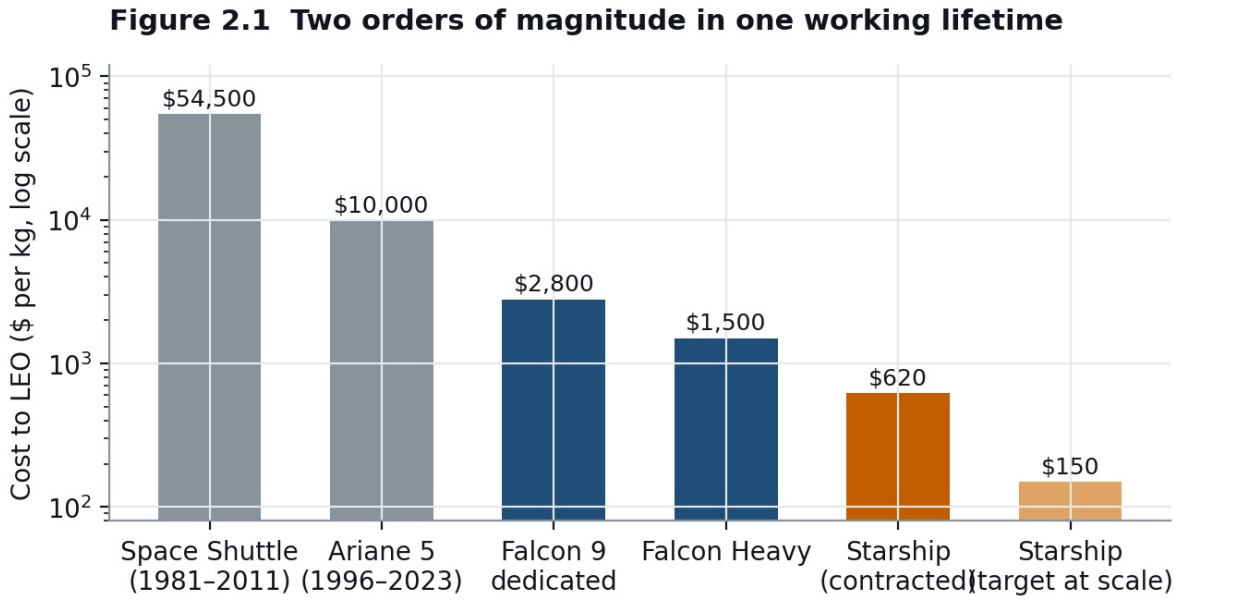
An engineer reading that should immediately ask the right follow-up question, which is not “how do we build more grid?” — many capable people are on that — but rather: *is there anywhere that a gigawatt of primary energy is already sitting, unqueued, unpermitted, and unowned?*

There is exactly one such place, and we have been able to reach it since 1957.

Chapter 2 — What SpaceX actually changed

Putting computers in space is not a new idea. Nothing in this book is a new idea; that is rather the point, and each sector chapter in Part IV opens by naming the person who had it first. What changed is not the concept. What changed is a price.

Figure 2.1 — *Two orders of magnitude in one working lifetime*



Starship contracted figure is inferred: a $90M contract disclosed in Voyager's 10-K against a 100–150 t station. The target column is company guidance, not a price anyone has paid.

The Space Shuttle delivered payload to low Earth orbit at something on the order of fifty thousand dollars per kilogram. A dedicated Falcon 9 today lists around two thousand seven hundred to three thousand. And in March 2026, for the first time, a Starship price entered the public record — not through a press release but through an SEC filing. Voyager Technologies disclosed a $90 million contract for a single launch, and Voyager’s payload is the Starlab station, which is designed to go up whole rather than be assembled in pieces.[6]

Divide the disclosed contract by a station mass in the hundred-to-one-hundred-fifty-tonne range and you get something near six hundred dollars per kilogram.[7]

Be careful with that number, and be careful with it in public. It is an **inference**, not a published price: the filing gives a contract value, and the mass comes from the station’s design, and nobody at either company has confirmed the division. It is also a price for a single very large dedicated payload, which is the most favourable case there is. And Starship has not yet flown that mission. Treat $600/kg as the best evidence currently available rather than as an achieved fact, and treat the company’s own long-run guidance of $100–200/kg as what it is: a target.

Why reuse changes the shape of the curve, not just its level

A first-principles reader should not accept “reuse makes it cheaper” without asking how much cheaper, and why there is a floor.

The cost of a launch decomposes into four things: propellant, hardware, operations, and the amortisation of the development programme. Propellant is startlingly cheap — the liquid oxygen and methane in a large vehicle costs on the order of a million dollars, which is a rounding error against a $90 million contract. Hardware is expensive and, in an expendable rocket, is destroyed on every flight. Operations and refurbishment sit in between.

So the arithmetic of reuse is simply this: if the hardware costs H to build and flies N times, its contribution per flight is H/N plus refurbishment. Going from $N = 1$ to $N = 10$ does not make launch ten percent cheaper; it removes ninety percent of the largest line item. Going from $N = 10$ to $N = 100$ does much less, because by then propellant, range operations and refurbishment dominate. That is why the curve in Figure 2.1 falls steeply and then flattens, and it is why $100/kg is difficult in a way that $600/kg is not.

The ratio that decides everything

Here is the calculation that tells you whether any of this is serious. It takes four inputs and you should change all of them as you read.

> **WORKED EXAMPLE — What fraction of an orbital data centre is launch?**
>
> Take a reference node of **100 kW** of compute — call it half a modern rack, plus the spacecraft around it. Assume for now a total wet mass of **3,000 kg**. *(This is a deliberately optimistic placeholder. Chapter 8 builds the vehicle part by part and arrives at 4,710 kg, which makes the ratios below roughly 50% worse. The optimistic version is shown here because it is the version the sector quotes, and it is instructive to see how far it moves once you do the work.)* The hardware inside — accelerators, storage, networking — costs roughly **$3–4M** at 2026 prices.
>
> - At Falcon 9 pricing (~$2,800/kg): launch = **$8.4M**, or roughly **2.4×** the value of the payload it carries.
> - At the contracted Starship figure (~$600/kg): launch = **$1.8M**, or roughly **half** the payload value.
> - At the guided target (~$150/kg): launch = **$450k**, or roughly **12%**.
>
> Nothing about the physics changed between those three lines. Only the price of the ride changed — and it moved the launch bill from *the dominant cost of the system* to *a line item smaller than the GPUs*.

That is the entire reason this book exists in 2026 and could not have been written in 2016. The physics of orbital compute was as sound then as it is now. It was simply unaffordable, and unaffordable ideas look identical to wrong ideas until the price moves.

One more consequence, and it is the one most people miss. When launch is the dominant cost, every engineering decision optimises for **mass**. When launch becomes cheap, mass stops being the master constraint and the design is free to get heavy where heaviness helps — thicker shielding, bigger radiators, deployable structures with margin. Much of Part IV is about companies who have understood that the design rules changed and are building for the new regime rather than the old one.

Chapter 3 — The inversion

Now put the two halves together, because the argument of this book is a single comparison.

Figure 3.1 — *The inversion*

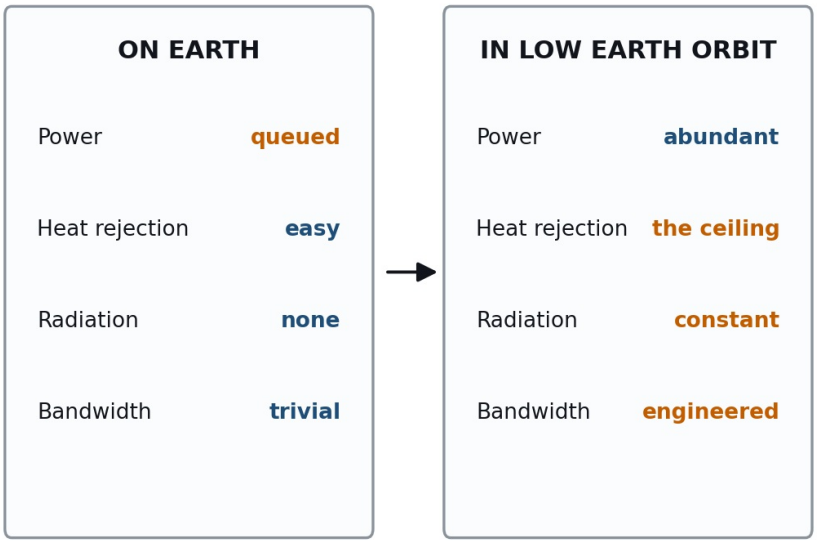


Figure 3.1 Moving compute to orbit does not remove constraints. It swaps a scheduling problem for a physics problem.

On Earth, energy is scarce in the sense that matters — scarce in *time*, rationed by a queue — while getting rid of heat is a solved commodity problem. You buy a chiller. You dig a cooling loop. Worst case, you evaporate water.

In low Earth orbit, both halves flip. Sunlight arrives at about 1,361 watts per square metre, unfiltered by atmosphere, on nobody's schedule and with nobody's permission. In the right orbit a spacecraft sees it almost continuously. Nothing in that sentence involves a transformer lead time.

And heat becomes the hardest problem you have. Not because there is a lot of it — there is exactly as much as on Earth, because the physics of Chapter 1's primer does not care where the rack is — but because in a vacuum there is no air to blow over a fin and no water to boil off. There is one and only one exit: **infrared radiation into the dark sky**. Radiators are the entire cooling system, they are large, and they set the mass and the deployment complexity of the spacecraft. Chapter 5 does this arithmetic honestly, and it is the chapter I would most like a sceptic to attack.

Radiation, meanwhile, goes from *not a consideration* to *a permanent design constraint*, and bandwidth goes from *free* to *a budget you must engineer*. These four — **power, heat, radiation, bandwidth** — are the spine of the book. Every sector in Part IV is a company attacking one of them.

So it should be said plainly, at the start, that going to orbit does not remove constraints. It **substitutes** them. You trade a scheduling problem for a physics problem: an interconnection queue for a radiator area, a transformer order book for a launch manifest. That trade is only a good one if the physics problem is smaller than the scheduling problem, and whether it is depends on numbers we will build up over the next hundred pages.

The argument that survives even if orbit is expensive

There is a weak version of the orbital case and a strong one, and most writing on this subject makes the weak one.

The weak version is a **cost** argument: orbital compute will be cheaper per GPU-hour than terrestrial compute. Over a two-year view this is difficult to defend, because it leans on a Starship price nobody has yet paid and on hardware lifetimes nobody has yet demonstrated.

The strong version is an **availability** argument, and it is already true.

> **WORKED EXAMPLE — What a queue costs per megawatt**
>
> Take one megawatt of modern accelerators. At about 130 kW per rack, that is roughly eight racks, or on the order of 550 to 600 GPUs once you account for the facility overhead of a low-PUE site. At prevailing cloud rates in the range of $3–4 per GPU-hour, and assuming high utilisation:
>
> 575 GPUs × $3.50/hr × 8,760 hr ≈ **$17.6M per megawatt-year** of gross revenue.
>
> Run the same calculation at $2/hr and you get about $10M; at $6/hr, about $30M. Call the honest range **$10–30M per megawatt-year**.
>
> Now apply the queue. A five-year wait for energisation is not a five-year delay in spending — it is five years of that revenue never earned, on the order of **$50–150M of foregone gross revenue per megawatt**, against a launch bill measured in single-digit millions per hundred kilowatts.

(Note for the manager: the v2 fund thesis uses ~$70M per megawatt-year for this figure. Built up from rack density and public GPU rental rates I cannot reproduce it — the number lands three to four times lower. The book uses the built-up range. The thesis document should be reconciled to it before either is published, because a sceptical reader will run exactly this calculation.)

Even at the bottom of that range, the conclusion holds and it is the most important sentence in Part I:

> **An orbital data centre does not have to be cheaper than a terrestrial one. It has to exist sooner.**

Two warnings about that sentence, both discharged in Part V. The first is that “the alternative” is not only a grid connection — it is a gas turbine on site, and Chapter 18 prices that competitor properly. The second is that orbit is not instant either: from decision to first operating megawatt is three to four and a half years. The availability argument survives both, in a narrower and more precise form than this page implies: orbit wins today for buyers who pay a premium for sovereignty, resilience or proximity to a sensor, and it wins for everyone else once manufacturing has come down its learning curve — which Chapter 18 dates to the early 2030s.

That is a much weaker claim than the one usually made for this sector, and weaker claims are the ones that survive contact with engineering. It is also why the rest of this book spends its time on radiators, radiation dose and optical links rather than on renderings. If the availability argument is the real one, then the question is not whether orbital compute is desirable. It is whether it is *buildable*, part by part, at the scale the argument requires.

That question has nine answers, one per subsystem, and they are what Part II and Part III are for.

Source notes, Part I

1. Rack power figures: Nvidia GB200/GB300 NVL72 at 120–140 kW and VR200 NVL72 at roughly 190–230 kW, with the Rubin Ultra “Kyber” architecture specified near 600 kW — SemiAnalysis, *Vera Rubin — Extreme Co-Design* (Feb 2026); ModulEdge, *Nvidia Vera Rubin: Data Center Impact* (Jun 2026). Reported figures differ by configuration; some sources quote 120.8 kW peak for a baseline VR NVL72. Ranges are used throughout rather than point values. **To verify against Nvidia’s own published specifications before publication.**
2. Air versus water volumetric heat capacity, and the practical density at which air cooling fails — The Strange Review, *The Liquid Revolution* (Mar 2026).
3. PJM interconnection queue duration — Data Center Knowledge, *AI Data Center Boom Rewires US Power Supply Chain* (May 2026), citing PJM data and filings.
4. Interconnection wait times of 4–7 years in major hubs, up to 10 years to energisation in some regions, and the 9–17% of national consumption projection — EPRI, *Powering Intelligence 2026*, as summarised in Sector Signals (Jun 2026). **Cite EPRI directly in the final text.**
5. Large power transformer lead times of 2–4 years — iMasons / industry procurement reporting, summarised in Construction Owners, *How Long It Actually Takes to Power a Data Center in 2026* (Jun 2026).
6. Voyager Technologies Form 10-K, filed 10 March 2026, disclosing a $90M launch contract; the Starlab station is manifested on Starship. Reported by Jeff Foust and subsequently by The Motley Fool (Mar 2026). **Cite the 10-K page directly.**
7. The ~$600/kg figure is an inference from the disclosed contract value divided by the station’s design mass range, not a disclosed price. Independent commentary reaches ~$612/kg using a 147 t assumption — Brownstone

Research (Apr 2026).

Figures 1.1, 2.1 and 3.1 generated for this volume; underlying values as cited above.

PART II — THE PHYSICS YOU CANNOT ARGUE WITH

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part II of VI.

Part I ended on a substitution: orbit trades a scheduling problem for a physics problem. This part prices the physics problem.

Four constraints govern everything that follows — **power, heat, radiation, bandwidth**. Every company in Part IV is attacking one of them. Every failure mode in Part VI is one of them winning. If you read this part with a pencil, you will be able to sanity-check any claim made about this sector for the next decade, including the ones in this book.

A note on method. Each chapter builds its numbers from constants rather than quoting a conclusion, and every number is given with the assumptions that produced it. Where the answer depends heavily on an assumption, the chapter says so and shows the sensitivity. Where I have used somebody else's published figure, it is cited. Where a number is a schematic illustration rather than a model output, it is labelled as one. This is the standard I would want applied to me, and it is the standard by which the sector's promotional material mostly fails.

Chapter 4 — Power

The constant

Every argument for compute in orbit begins with one number: **1,361 watts per square metre**. That is the solar irradiance at Earth's distance from the Sun, above the atmosphere, and it is the closest thing this industry has to a free lunch.

It is not free on the ground, because the atmosphere takes a share, clouds take another, night takes half, the seasons take more, and the panel spends most of the day pointing somewhere other than directly at the Sun. A good terrestrial site delivers perhaps 2,000 to 2,300 kWh per square metre per year onto a tilted panel. A spacecraft in the right orbit sees the full constant, almost all the time, pointed almost perfectly.

> **PRIMER — Why space cells are not the cells on your roof** Terrestrial panels are silicon, roughly 20–22% efficient, and chosen because they are cheap per watt. Spacecraft use multi-junction cells — typically three stacked semiconductor layers, each tuned to a different slice of the spectrum — at roughly 30–32% efficiency. They cost far more per watt and would be economically absurd on a roof. In orbit, where the cost that matters is the mass and area you had to launch, paying ten times more for fifty percent more output per square metre is obviously correct. This is a preview of a pattern that recurs throughout the book: **orbital engineering optimises different variables, so it reaches different answers, and the answers look wrong until you notice the variables changed.**

The comparison, built up

Take a 30% multi-junction cell. Facing the Sun, it produces 1,361 × 0.30 ≈ **408 watts per square metre**.

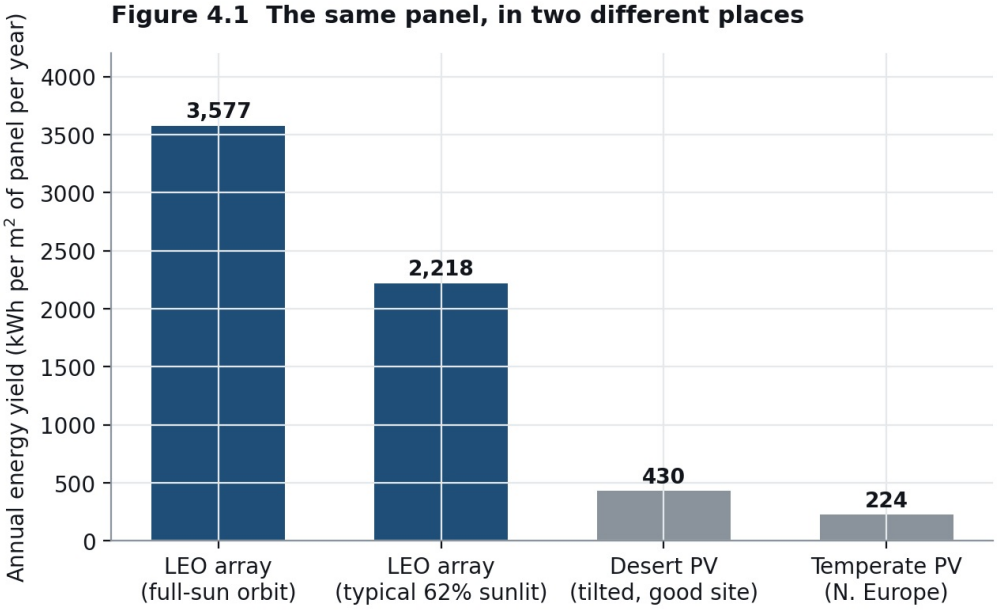
In a dawn–dusk sun-synchronous orbit — an orbit whose plane stays roughly perpendicular to the Sun line, so the spacecraft rides the day–night terminator — the array can be in sunlight essentially continuously for much of the year. Run that for a full year:

408 W/m² × 8,760 h = **3,577 kWh per square metre per year.**

A conventional LEO orbit, sunlit perhaps 62% of each revolution, yields around 2,200. A tilted panel at an excellent desert site, at 22% module efficiency and a realistic 0.85 performance ratio, yields about 430. A northern European site yields about 224.

Figure 4.1 — *The same panel, in two different places*

Figure 4.1 The same panel, in two different places



Orbit: 1,361 W/m² at 30% triple-junction efficiency. Terrestrial: plane-of-array irradiance x 22% module x 0.85 performance ratio. Per panel area, not per land area — terrestrial arrays also need row spacing.

So the honest ratio is roughly **eight to one** against a very good terrestrial site, per square metre of panel. That is the figure Google's own Project Suncatcher work arrives at independently, describing solar panels in the right orbit as up to eight times more productive than on Earth.[8]

Two adjustments, one in each direction. In favour of orbit: this compares panel area to panel area, and a terrestrial solar farm needs land between the rows, so the ratio per *acre* is larger still. Against orbit: the terrestrial panel is bolted to the ground and the orbital one had to be launched, deployed, pointed, and kept in an orbit that decays.

What 100 kilowatts actually looks like

Divide through. A 100 kW node needs 100,000 / 408 ≈ **245 square metres** of array — call it 280 m² once you allow for pointing losses, cell degradation, and the power lost between the array and the chips. That is a square roughly 17 metres on a side, or two wings each the size of a tennis court, folded into a launch fairing.

Mass follows from **specific power**, the watts per kilogram a given array technology delivers. Rigid panels sit around 40–100 W/kg; modern flexible roll-out arrays reach roughly 150 W/kg at beginning of life. At 150 W/kg, 100 kW of generation is about 670 kg of array. At a conservative 100 W/kg it is a tonne.

Scale that to a megawatt and the array alone is seven to ten tonnes. Hold that number; Chapter 5 adds a bigger one to it.

Eclipse, and the freedom nobody talks about

If the orbit is not full-sun, the spacecraft passes into Earth's shadow for roughly 35 minutes of each 90-minute revolution. Conventional spacecraft carry batteries. For a data centre, that is expensive: 100 kW through a 35-minute eclipse is about 58 kWh of storage per orbit, which at 200 Wh/kg is nearly 300 kg of battery per 100 kW, before depth-of-discharge margin and before the batteries' own thermal management.

But a data centre has an option a weather satellite does not. **It can simply stop computing.**

Terrestrial data centres are built for continuous availability because their customers demand it and their capital is idle when they are off. An orbital node running batch workloads — training runs, overnight rendering, scientific simulation, anything checkpointable — can be a *load-following machine*: full power in sunlight, idle in eclipse, no batteries at all beyond housekeeping. That is a 35% haircut on utilisation in exchange for deleting a mass and a failure mode.

Whether that trade is good depends entirely on the workload, and it is one of the sharper questions to ask any company in this sector: *do you carry batteries, and if so, which customer forced you to?* Choosing the full-sun orbit instead is the third answer, and it is the one Google's design takes — a dawn–dusk sun-synchronous orbit around 650 km.[8]

Where power breaks

Four things, and they are where the startups in Chapter 10 live.

1. **Deployment.** A 280 m² array is a mechanism, and mechanisms are the classic single point of failure on spacecraft. Nothing about the physics fails here; the hinge does.
2. **High voltage in plasma.** Moving a megawatt around a spacecraft at the tens of volts traditional buses use would require absurd conductor mass, so you want hundreds of volts. But LEO is not empty — it is a tenuous plasma, and high-voltage surfaces in plasma arc. This is a genuinely under-discussed constraint on scaling orbital power, and it is why high-voltage power management is a differentiator rather than a commodity.
3. **Degradation.** Arrays lose output to radiation, typically a low single-digit percentage per year. Over a five-year life that is a real haircut on the revenue model and it belongs in any honest unit-economics sheet.
4. **Drag.** A large array is a large sail. At 500–650 km the atmosphere is thin but not absent, and a high area-to-mass spacecraft decays measurably. Station-keeping propellant is a consumable, and consumables set service life.

Chapter 5 — Heat

This is the chapter the thesis lives or dies on, and it is the one I would most like a hostile engineer to attack.

Restating the problem exactly

From the primer in Chapter 1: essentially all electrical power entering a computer leaves it as heat. A 100 kW node is a 100 kW heater. A 1 MW node is a 1 MW heater, which is roughly the thermal output of two hundred domestic ovens running flat out, inside a structure the size of a bus, in a vacuum.

There are three ways to move heat: **conduction**, **convection**, and **radiation**. In orbit the second one is gone. There is no air to blow across a fin, no water to evaporate, no ground to sink into. Conduction still works *inside* the spacecraft — that is how you get heat from the chip to the panel — but at the boundary with the universe there is exactly one mechanism, and it is radiation.

The equation, and the one thing it tells you

A surface at temperature T radiates according to the Stefan–Boltzmann law. Net of what it absorbs from its surroundings:

$$q = \varepsilon \sigma (T^4 - T_{sink}^4)$$

where q is watts per square metre, ε is emissivity (how close the surface is to a perfect radiator — good radiator coatings reach 0.85 or better), and $\sigma = 5.67 \times 10^{-8}$ W/m²K⁴.

Everything about orbital thermal design is contained in that exponent. **Heat rejection scales as the fourth power of temperature.** Run your radiator ten percent hotter in absolute terms and you reject about 46% more heat from the same panel. That single fact should reorganise how you evaluate this sector.

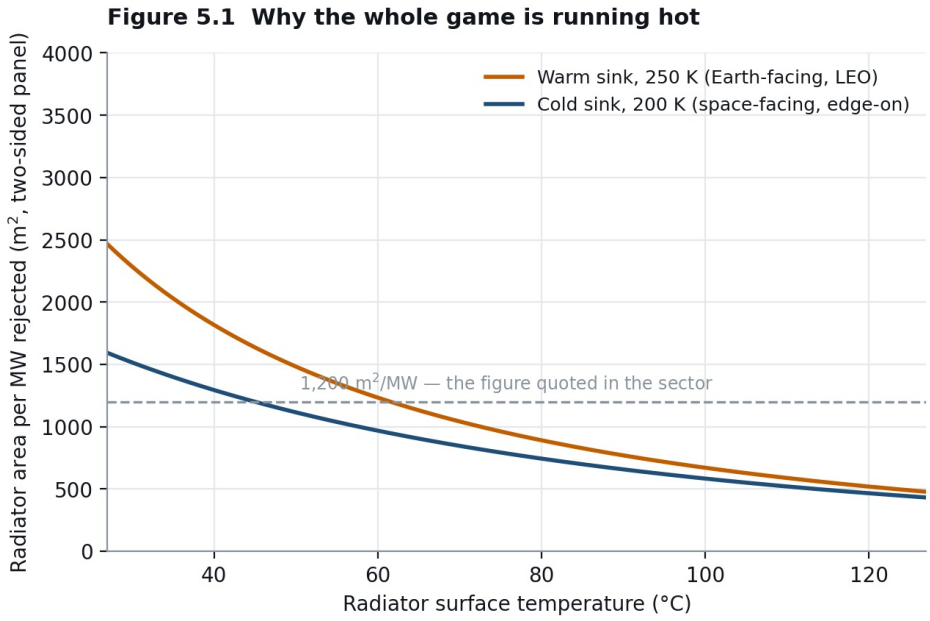
> **PRIMER — The sink is not absolute zero** Deep space is 2.7 K, and it is tempting to put that into the equation, which makes the sink term vanish and the numbers look wonderful. Do not. A radiator in low Earth orbit stares at a planet: Earth emits roughly 240 W/m² of infrared and reflects around 30% of the sunlight hitting it. A panel facing the Earth is being warmed by it. Depending on orientation, orbit and coatings, the *effective* sink temperature for a LEO radiator is somewhere in the range of roughly 200 to 270 K, not 2.7. Spacecraft designers work hard on geometry — radiators mounted edge-on to the Earth, shaded from the Sun — precisely to buy back those degrees.

The number

Take ε = 0.85 and run the equation across plausible radiator temperatures and two sink assumptions, for a two-sided panel radiating from both faces:

Radiator surface temp	Net flux, 250 K sink	Area per MW	Net flux, 200 K sink	Area per MW
40 °C (313 K)	274 W/m²	1,820 m²	386 W/m²	1,300 m²
50 °C (323 K)	336 W/m²	1,490 m²	448 W/m²	1,120 m²
70 °C (343 K)	479 W/m²	1,040 m²	590 W/m²	850 m²
100 °C (373 K)	745 W/m²	670 m²	856 W/m²	580 m²

Figure 5.1 — *Why the whole game is running hot*



q = $\varepsilon\sigma(T^4 - T_{sink}^4)$ with ε = 0.85. Area falls as roughly the fourth power of temperature: 40 °C to 100 °C cuts the radiator by about two thirds.

The figure most often quoted in this sector is **roughly 1,200 square metres per megawatt**. You can now see exactly what that figure assumes: a two-sided panel at around 40–50 °C with a reasonably cold sink. It is a fair number, not a promotional one — but it is not a law of nature either. It is the answer at one point on a curve, and a company that can run its radiators at 70 °C instead of 45 °C needs *half* the panel.

This is the single most useful lens in the sector, so it is worth stating as a rule:

> **Every degree you can run hotter is area you do not have to launch. The thermal problem is therefore not really a thermal problem — it is a question of how hot you are willing to let the silicon get.**

Terrestrial data centres run chips cool because cooling is cheap and hot silicon fails sooner. In orbit, cooling is the dominant mass in the vehicle. The economically correct answer is almost certainly to run hotter than any terrestrial operator would tolerate, accept a shorter hardware life, and replace more often. Nobody in this sector talks in these terms yet. I expect the winners will.

The half of the problem the equation does not cover

Radiator area is the *easy* half. The hard half is getting a megawatt of heat from a few hundred square centimetres of silicon out to a thousand square metres of panel.

The chain is a series of temperature drops, and every drop costs you area. Junction to cold plate. Cold plate to coolant. Coolant pumped through a loop. Loop into the radiator root. Root out along the fin, which is never isothermal. If the junction runs at 90 °C and you lose 30 degrees down the chain, you are radiating at 60, not 90, and Figure 5.1 charges you for it.

The available technologies, in ascending order of capability and descending order of flight heritage:

- **Heat pipes and loop heat pipes.** Passive, extremely reliable, decades of orbital heritage — and limited to

hundreds of watts to low kilowatts per line. They do not scale to megawatts.

- **Pumped single-phase loops.** Mechanical pumps circulating a fluid. Flight-proven at the tens-of-kilowatts scale on the ISS. Needs pumping power and introduces moving parts.
- **Pumped two-phase loops.** The coolant boils, which carries far more heat per kilogram of flow and holds temperature nearly constant. This is the right answer at megawatt scale, and it is the least flight-proven. Managing two-phase flow in microgravity, where vapour does not obligingly rise, is a genuinely hard problem.

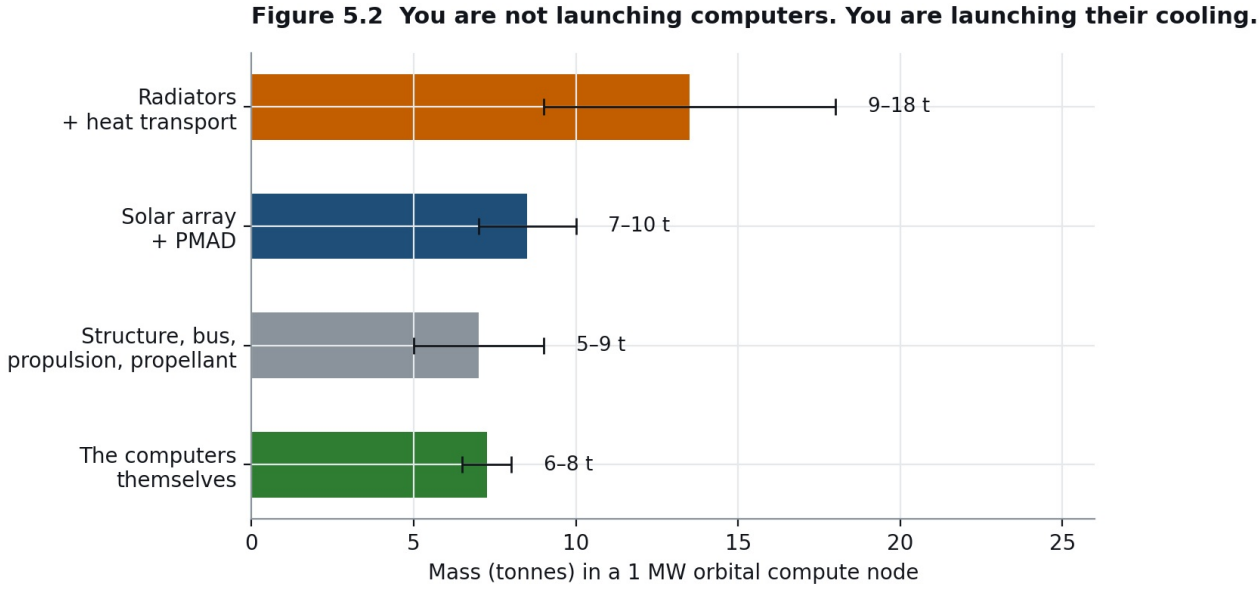
This is the honest statement of where the sector is: **between roughly ten watts and five hundred kilowatts, thermal management is solved with hardware that has flown for decades. Above a megawatt, it is not solved, and the companies claiming gigawatts have not built the intermediate step.**

Mass, and the sentence that follows from it

Deployable spacecraft radiators run roughly 5–12 kg per square metre including structure, plumbing and fluid. At 1,200 m² and the middle of that range, a megawatt of rejection is on the order of **9 to 18 tonnes**.

Now assemble the whole vehicle.

Figure 5.2 — *You are not launching computers. You are launching their cooling.*



Build-up in Chapter 8. Total 27–45 t per MW; the payload that earns revenue is roughly a fifth of it.

A megawatt-class orbital compute node lands somewhere around **27 to 45 tonnes**, of which the computers — the only part that earns revenue — are roughly six to eight. **The revenue-generating payload is about a fifth of what you launch.** Everything else is the apparatus required to feed it and cool it.

That is not a refutation. Terrestrial data centres are also mostly not-computers: substations, chillers, generators, concrete. It is, however, the number that should be in your head whenever someone multiplies a launch price by a rack weight and declares the economics obvious. At $600 per kilogram, a megawatt node costs **$16–27 million to launch**, against perhaps $15–20 million of computers inside it. Launch is not a rounding error at megawatt scale. It is comparable to the payload.

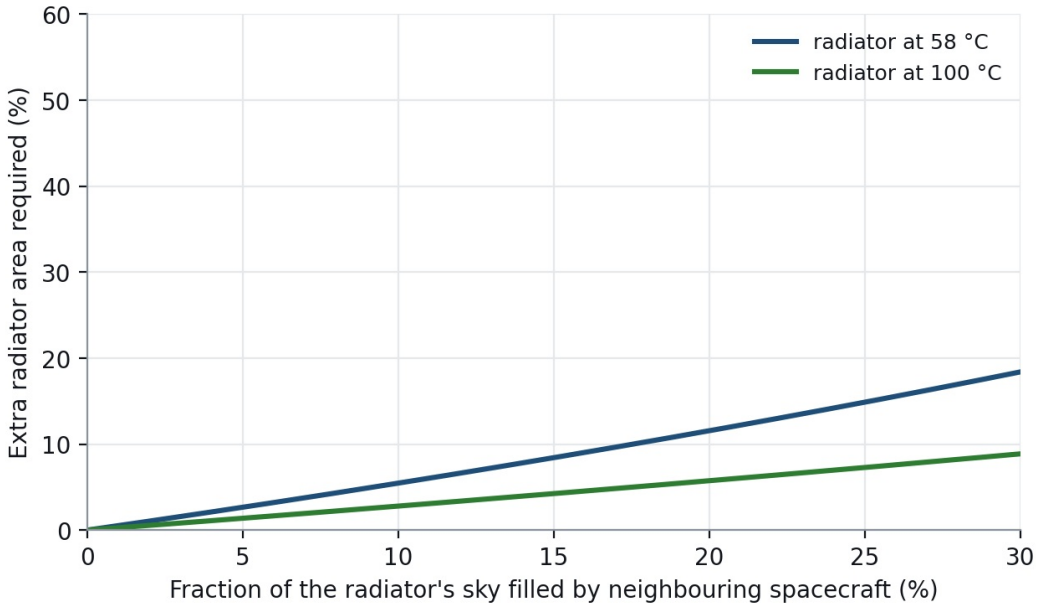
The complication a cluster introduces

One more effect belongs here, because Chapter 13’s preferred architecture quietly undermines Chapter 5’s arithmetic.

Everything above assumes a radiator with a clean view of cold sky. A formation-flying cluster does not have one. If your neighbours sit a few hundred metres away, they occupy part of your radiator’s field of view, and they are not cold sky — they are spacecraft at roughly three hundred kelvin. The effective sink is then a weighted mix: the cold background over most of the hemisphere, and warm hardware over the rest.

Figure 5.3 — *A cluster radiates partly into itself*

Figure 5.3 A cluster radiates partly into itself



Neighbouring spacecraft at ~300 K replace cold sky in the radiator's field of view. Cooler radiators suffer more, because they have less margin over the sink. Formation spacing is therefore a thermal decision, not only a control one.

The penalty is modest at realistic spacings and it is not zero. With a tenth of the radiator's sky filled by neighbours, a panel running at 58 °C needs roughly **five to six percent more area**; at a fifth of the sky, more than ten percent. Cooler radiators suffer proportionally more, because they have less temperature margin over the sink to begin with — the same fourth-power logic, working against you.

Two consequences worth carrying forward. **Formation spacing is a thermal decision, not only a control decision**, and a cluster design that optimises station-keeping propellant without a view-factor analysis has optimised the wrong variable. And every radiator area figure in this book, including the ones in Chapter 8's build-up, is a single-vehicle number that should be inflated by several percent for any clustered architecture.

Where heat breaks

- **Two-phase transport at megawatt scale** — the unretired technical risk of the entire sector.
- **Deployment.** A thousand square metres of radiator has to fold into a fairing and unfold without a single jam. This is the same mechanism risk as the solar array, doubled.
- **Micrometeoroid and debris puncture.** A radiator is a large, thin, fluid-filled target with the largest surface area on the vehicle. Puncture is not hypothetical over five years; it is an actuarial certainty at some rate, which is why panel isolation and loop redundancy are design requirements and not refinements. This is also the chapter that quietly determines what Chapter 16's insurers will charge.
- **Degradation of optical properties.** Emissivity is a coating property, and coatings erode under atomic oxygen and UV. A radiator that loses emissivity loses capacity on exactly the fourth-power curve above.

Chapter 6 — Radiation

What is actually out there

Three distinct populations, with three different consequences:

1. **Trapped particles.** Protons and electrons captured by Earth's magnetic field into the Van Allen belts. In low orbit you mostly miss them, except over the **South Atlantic Anomaly**, where the inner belt dips low, and near the poles, which is precisely where sun-synchronous orbits go. A polar or sun-synchronous data centre takes more dose than an equatorial one. That is a real cost of the full-sun orbit Chapter 4 recommended.
2. **Galactic cosmic rays.** Very high energy nuclei from outside the solar system. Low flux, extremely penetrating, effectively impossible to shield with any mass you would launch.

3. **Solar particle events.** Sporadic, occasionally enormous, and the reason a spacecraft needs a safe mode rather than merely an average-case design.

These do two categorically different things to electronics.

Total ionising dose is cumulative and gradual: charge builds up in insulating layers, thresholds shift, leakage rises, and eventually the part is out of spec. It is a wear-out mechanism, measured in rad(Si), and it defines a lifetime.

Single event effects are instantaneous: one particle deposits enough charge in the wrong place to flip a bit (an upset), glitch a signal (a transient), or — worst case — trigger a parasitic structure that draws destructive current until power is removed (**latch-up**). Upsets are a reliability nuisance. Latch-up destroys hardware.

What Google’s test actually showed, and why it matters so much

Until recently, the honest answer to “can a modern AI accelerator survive orbit?” was that nobody had published a test. In late 2025 Google published one, and it is the most important single data point in this sector.

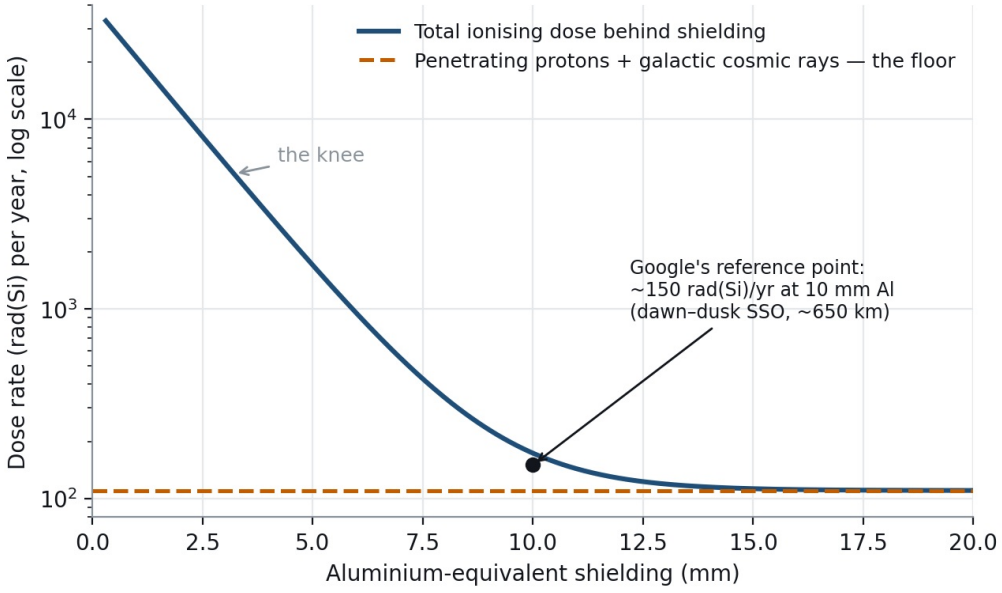
Google tested a Trillium v6e TPU and its host server in a 67 MeV proton beam. Behind roughly 10 mm aluminium-equivalent shielding, they put the expected dose in their target sun-synchronous LEO at about **150 rad(Si) per year — a five-year mission dose near 750 rad(Si)**. The high-bandwidth memory was the most sensitive subsystem and began showing irregularities only after about **2 krad(Si)**, close to three times the five-year dose. No hard failures attributable to total dose appeared up to the maximum tested **15 krad(Si)**.[9]

Read that carefully, because it is easy to over-read. It does not say frontier silicon is immune to space. It says that for one accelerator, in one orbit, behind realistic shielding, **cumulative dose is not the binding constraint over a five-year life.** Single event effects, long-duration behaviour, and thermal cycling are not settled by a beam test.

But it moves the argument. The old assumption was that space compute required bespoke radiation-hardened silicon, which runs generations behind commercial parts and costs five to ten times more. If commercial accelerators survive with shielding and software discipline, the cost premium collapses toward something like one-and-a-half to two times — and the difference between those two worlds is the difference between this sector being arithmetic and being fantasy.

Figure 6.1 — *Shielding has a knee, and then it stops paying*

Figure 6.1 Shielding has a knee, and then it stops paying



Curve shape is illustrative and anchored to Google's published figure; absolute dose depends on altitude, inclination and solar cycle and must be modelled per orbit. The point is the shape: the first few millimetres do nearly all the work.

The shielding calculation, and why you shield the box

Shielding is mass, and mass is area × thickness × density. Aluminium is 2.7 g/cm³, so 10 mm of it over a one-square-metre surface is 27 kg.

That is the whole reason the correct architecture shields **the electronics enclosure, not the spacecraft**. Wrap a compute box of, say, 12 m² of surface in 10 mm equivalent and you have spent roughly 320 kg — real, but affordable inside a 30-tonne vehicle. Wrap the whole vehicle, radiators included, and you have spent your entire mass budget protecting panels that do not care.

And note the shape of Figure 6.1. The first few millimetres do nearly all the work, because they stop the soft trapped electrons and lower-energy protons that dominate the flux. After the knee, you are adding mass against penetrating protons and cosmic rays that mostly ignore it — and at high thickness, incoming particles produce secondary showers inside the shield, so the returns can go worse than flat. **There is an optimum, it is a few millimetres to a centimetre, and past it you are launching lead for nothing.**

The three answers that are not shielding

- **Orbit selection.** Altitude and inclination change dose by more than any plausible shielding decision. This is a free variable and it is coupled to Chapter 4's full-sun orbit preference and Chapter 5's sink temperature. These trades are not separable, which is why spacecraft design is iterative and why "we'll just put it in SSO" is not a design.
- **Software.** Error-correcting memory, checkpointing, redundant execution, watchdogs that reset a hung device. Hyperscale software already assumes hardware fails constantly; that culture is an asset here, and it is why the compute layer may prove more tractable than the mechanical layers.
- **Current-limited power.** Latch-up is survivable if the supply detects the current spike and cycles the device before it cooks. This is a power-electronics design decision, and it is one of the few places where a supplier can differentiate on something other than price.

The reframe that matters for an investor

Radiation is not a wall. **It is a depreciation schedule.**

> **WORKED EXAMPLE — What a shorter hardware life actually costs**
>
> Take a 200 kW rack costing about $3.5M. Straight-line it over five years and the hardware costs $700k per year. If radiation and thermal cycling cut useful life to three years, it costs $1.17M per year — a difference of $470k.
>
> Now compare that with what the rack earns. At the $10–30M per megawatt-year range built in Part I, 200 kW earns roughly $2–6M a year gross.
>
> The life reduction costs somewhere around **8–20% of gross revenue**. Painful. Not disqualifying.

That is the right way to hold radiation risk in your head. It does not decide whether orbital compute works. It decides the margin, and therefore the price you should be willing to pay for a company that has not yet demonstrated multi-year hardware life on orbit.

Where radiation breaks

Long-duration behaviour of dense HBM stacks; the absence of any five-year on-orbit dataset for frontier accelerators; extreme solar particle events; and thermal cycling, which is technically Chapter 5's problem but kills hardware on the same timescale and is much less discussed.

Chapter 7 — Bandwidth

The link budget, in one paragraph

Radio power spreads. A transmitter radiating into a beam of angular width θ spreads its energy over an area that grows as the square of distance, so received power falls as $1/d^2$. You claw it back with antenna gain — focusing the beam — and gain rises as you shrink the wavelength or grow the aperture. That is the entire argument for optical links: at infrared wavelengths, a modest telescope produces a beam divergence measured in microradians rather than degrees, so almost all the transmitted photons arrive at the receiver instead of warming the sky. Optical links also sidestep spectrum licensing entirely, which is not a physics advantage but is a very real commercial one.

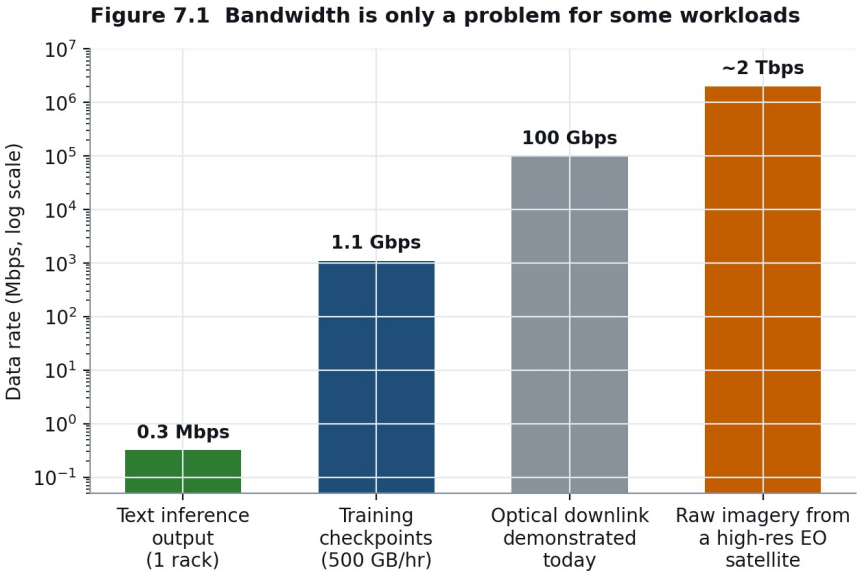
The price is pointing. Microradian beams require the two spacecraft to know where each other are and to hold aim through vibration and thermal distortion. Optical terminals are, in practice, precision mechanisms with a laser

attached, and that is why the terminal suppliers in Chapter 14 are a real business rather than a commodity.

How much data does an orbital data centre actually move?

This is the question that gets skipped, and the answer is far more encouraging than intuition suggests — for some workloads, and far worse for others.

Figure 7.1 — *Bandwidth is only a problem for some workloads*



Selling tokens from orbit is bandwidth-trivial. Shipping raw sensor data down is not — which is the argument for putting the compute next to the sensor.

Work it through:

- **Selling inference.** A rack serving, say, 10,000 output tokens per second at roughly four bytes per token generates about 40 kB/s — **0.32 megabits per second**. A thousand racks would need 320 Mbps. Text inference is, in bandwidth terms, nothing. You could downlink a substantial orbital inference business over a link that would embarrass a home broadband connection.
- **Training checkpoints.** A 500 GB checkpoint every hour is about 1.1 Gbps sustained. Real, manageable, and mostly avoidable by checkpointing to storage in orbit rather than to the ground.
- **Getting training data up.** Here is the wall. Ten petabytes of training corpus over a 100 Gbps link takes **about nine days of continuous transmission**, assuming you have contact with a ground station the whole time, which you do not. Training on Earth-resident data in orbit is a data-logistics problem long before it is a compute problem.
- **Data that is already up there.** A high-resolution imaging satellite generates raw data far faster than it can ever downlink it, which is why so much of it is discarded or heavily compressed on board. This is the strongest near-term case in the sector: **put the compute next to the sensor** and downlink conclusions rather than pixels.

So the ranking of natural first workloads falls straight out of the arithmetic: process data that is already in orbit; serve inference where the output is small; run batch training on resident datasets; and only last, and reluctantly, ship large terrestrial corpora upward.

Latency, which is not the problem people think

A spacecraft at 550 km is 1.8 milliseconds away at the speed of light straight overhead, and perhaps 3–7 ms at a realistic slant range. Add ground segment and backhaul and an orbital node sits, in latency terms, roughly where a data centre in a neighbouring city sits. That is fine for essentially all inference, fine for training, and not fine for latency-critical trading or real-time control loops.

The real availability problem is not distance but **geometry**: any given spacecraft is above any given ground station for only a few minutes per pass. Continuous service therefore requires either many ground stations, or inter-satellite links that relay traffic across the constellation to whichever node currently has a ground contact. The second option is what turns a fleet of computers into *one* computer, and it is why the relay layer is strategically more important than its revenue currently suggests.

Where bandwidth breaks

Cloud cover, which blocks optical ground links and forces geographically diverse station networks; pointing and acquisition under thermal distortion; the capital cost of a ground segment nobody counts in the spacecraft budget; and the regulatory question of who is allowed to downlink what, where — which Chapter 17 takes up when we leave the United States.

Where Part II leaves us

Constraint	Status	Who has to solve it
Power	Physics strongly favourable — roughly 8× per square metre. Engineering is deployment, high-voltage handling and degradation.	Array and power-management suppliers (Ch 10)
Heat	Solved below ~500 kW with decades of heritage. **Unsolved at megawatt scale.** The binding constraint.	Thermal specialists (Ch 11) — the thinnest layer in the sector
Radiation	Materially de-risked for dose by Google's testing. Open on long-duration and single-event behaviour. A depreciation schedule, not a wall.	Compute and power-electronics designers (Ch 12)
Bandwidth	Not a constraint for the natural first workloads. A hard constraint for uploading terrestrial data.	Optical terminals and relay operators (Ch 14)

One of those four rows is materially worse than the others, and it is the one with the fewest investable companies attacking it. That asymmetry is the most actionable observation in this book, and Part III now takes it apart component by component, before Part IV names the people trying.

Source notes, Part II

8. Solar yield in orbit versus Earth, the ~8× figure, and the dawn–dusk sun-synchronous reference orbit at approximately 650 km — Google Research, *Exploring a space-based, scalable AI infrastructure system design* (Project Suncatcher), and the accompanying paper, *Towards a future space-based, highly scalable AI infrastructure system design* (arXiv:2511.19468). The independent build-up in this chapter reaches 8.3×, which is a satisfying agreement given the two calculations were done from different starting points.
9. TPU radiation testing: 67 MeV proton beam; estimated ~150 rad(Si)/year and ~750 rad(Si) five-year shielded mission dose in the target sun-synchronous LEO behind ~10 mm Al equivalent; HBM irregularities from ~2 krad(Si); no hard TID-attributable failures to 15 krad(Si) — same sources as note 8.

Constants used throughout: solar constant 1,361 W/m²; Stefan–Boltzmann $\sigma = 5.670 \times 10^{-8}$ W/m²K⁴; aluminium density 2.70 g/cm³. Multi-junction cell efficiency, radiator emissivity, radiator areal mass, array specific power and coolant loop performance are stated as ranges reflecting current flight hardware; each is a variable the reader should change.

Open items for verification before publication: radiator areal mass (5–12 kg/m²) and array specific power (~150 W/kg) are quoted from general spacecraft-engineering practice and should be sourced to specific flight hardware datasheets. The 200–270 K effective sink range should be replaced with a properly cited thermal-environment reference for the specific reference orbit used in Chapter 8.

Figures 4.1, 5.1, 5.2, 6.1 and 7.1 generated for this volume from the equations and assumptions stated in the text.

PART III — ANATOMY OF A COMPUTE SATELLITE

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part III of VI.

Part II established four constraints. This part builds one machine that satisfies all four at once, part by part, and prices it.

The reason to do this before meeting any companies is simple: **you cannot evaluate a subsystem supplier until you know what the system needs from it.** A thermal company sounds impressive in isolation. Held against a mass budget it either closes a gap that matters or it does not.

So Chapter 8 defines a reference design and then takes it apart. Every number is a build-up, every assumption is named, and where the build-up disagrees with something stated earlier in this book, the disagreement is shown rather than quietly smoothed over.

Chapter 8 — Walking the machine

The reference node

Two vehicles carry the rest of the book:

Reference Node A — 100 kW. One modest compute payload on a single spacecraft. This is roughly the scale the sector is actually flying towards in the late 2020s, and everything in it exists in flight-proven form today.

Reference Node B — 1 MW. Ten times the compute, and the scale at which the sector's stated ambitions begin. Not simply Node A ten times over, because some things scale better than linearly and one thing — heat transport — does not scale at all with current hardware.

Both are assumed to fly in a **dawn–dusk sun-synchronous orbit at roughly 650 km**, which is the orbit Google's Project Suncatcher design selects and, for reasons that will become clear in the propulsion section, very close to the correct answer.

One convention, stated once and used throughout: a **modern AI rack is taken as 200 kW, 1.4 tonnes, and $3.5M**. Reported figures range from about 130 kW to 230 kW per rack depending on generation and configuration, so this is a choice rather than a fact, and it is among the assumptions most worth changing when you run your own version of these numbers.

Figure 8.1 — *Reference Node A, subsystem by subsystem*

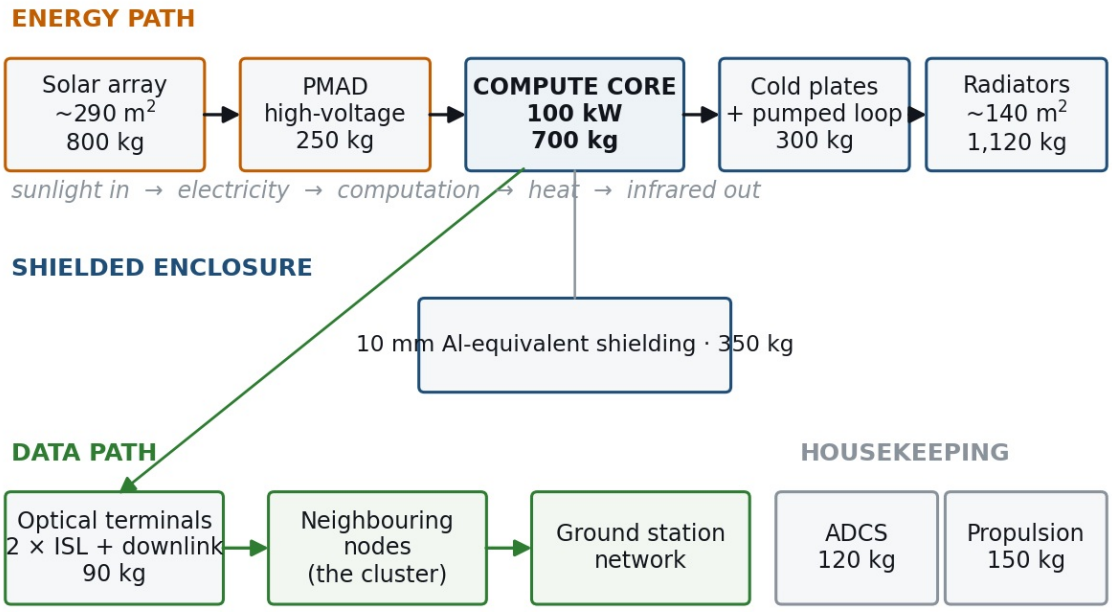


Figure 8.1 Reference Node A — 100 kW of compute, 4.7 tonnes of spacecraft. Masses built up in Table 8.1.

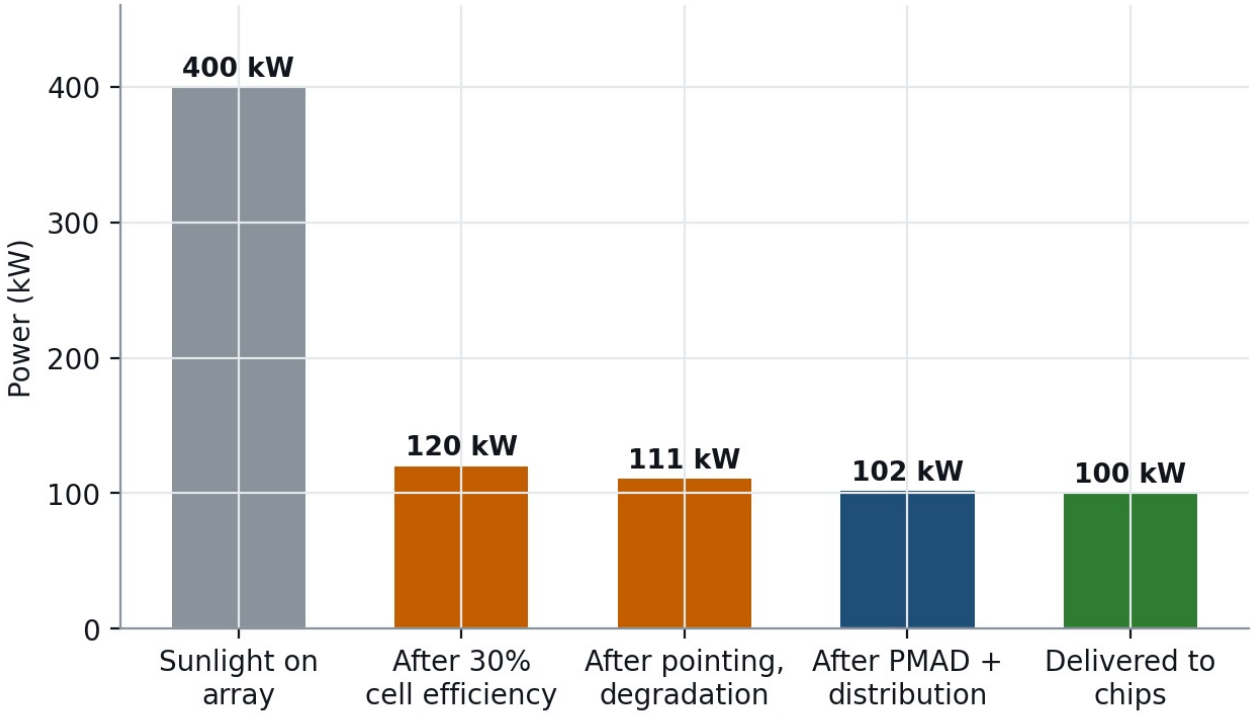
> **PRIMER — What a "satellite bus" actually means** Spacecraft are described as a **payload** and a **bus**. The payload is the thing the mission exists for — a camera, a radio, or in our case a rack of accelerators. The bus is everything that keeps the payload alive and pointed: structure, power, thermal control, attitude determination and control, propulsion, and command and data handling. The commercial significance is that buses have become products. You can buy one, as you would buy a server chassis. That is what companies like Apex and Loft Orbital sell, and it is why an orbital compute startup does not have to be a spacecraft manufacturer — which in turn is why this sector could form as quickly as it did.

1. The power chain, sized backwards

The first mistake an outsider makes is sizing the solar array to the solar array. You size it to the load, and then you walk every loss backwards.

Figure 8.2 — *Sizing backwards from the chips*

Figure 8.2 Sizing backwards from the chips



You do not size the array to the array. You size it to the load, then walk every loss backwards. Delivering 100 kW to silicon needs ~290 m^2 of array intercepting ~400 kW of sunlight.

To deliver 100 kW to silicon you need roughly 102 kW off the power bus, about 111 kW from the array after distribution and conversion losses, and about 120 kW at beginning of life once you have reserved margin for five years of radiation degradation and imperfect sun pointing. At 408 W/m² of cell output, that is **roughly 290 square metres of array** intercepting about 400 kW of raw sunlight.

Three engineering decisions hide inside that paragraph, and each is a company in Part IV.

Voltage. Moving 120 kW around a spacecraft at the 28 or 100 volts traditional buses use means enormous currents and therefore enormous conductor mass. You want to go to several hundred volts. But low Earth orbit is a tenuous plasma, and high-voltage surfaces in plasma can arc — a failure mode that has damaged real spacecraft. High-voltage power management for orbit is genuinely hard, which is why it is a differentiator rather than a commodity, and it is the specific thing that makes a high-power bus a product rather than a bigger version of a small one.

Deployment. Two wings of roughly 145 m² each must survive launch folded and then unfold once, perfectly, with no opportunity for repair. Mechanisms are the classic spacecraft single point of failure. The physics never fails here. The hinge does.

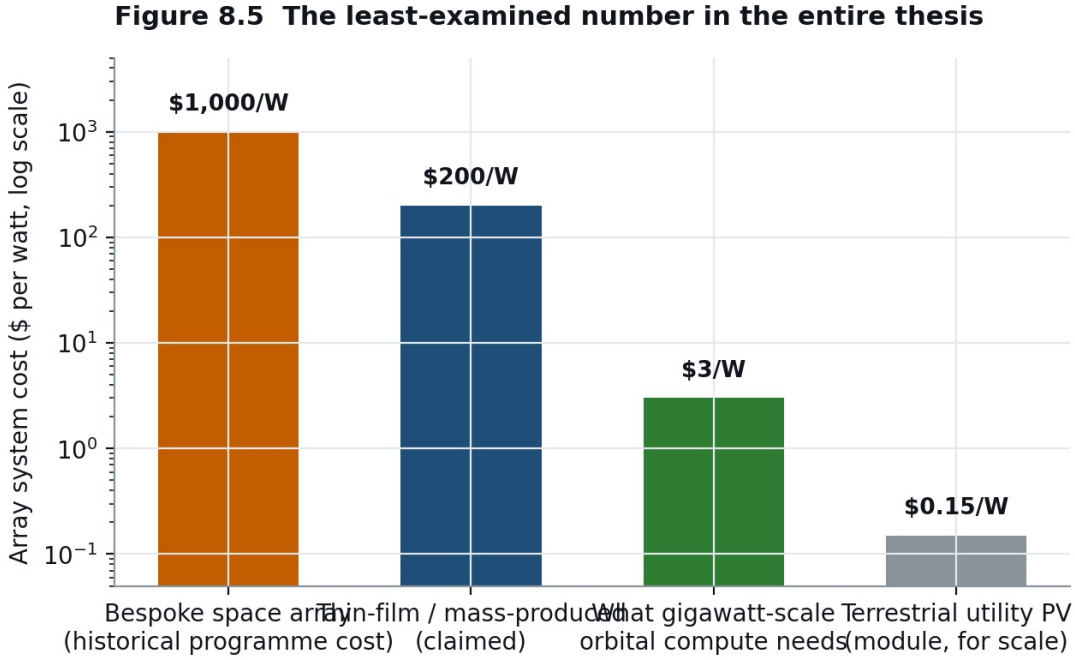
Cell chemistry — and a live argument. Chapter 4 asserted that multi-junction cells at 30% are obviously correct for orbit, because area and mass are what you pay for. That reasoning is sound at Node A's scale. It may be wrong at gigawatt scale, and the sector is now openly split: in February 2026 Rocket Lab introduced **silicon** solar arrays aimed specifically at space-based data centres, arguing that low cost per watt at industrial manufacturing scale is what gigawatt-class orbital power actually requires.[10]

Both positions can be right. Silicon is roughly two-thirds the efficiency, so it needs about 50% more area and carries more mass per watt — a real penalty when launch is expensive and the spacecraft is small. But it is manufacturable by the square kilometre, and when you need gigawatts rather than kilowatts, manufacturability beats elegance. **Which cell chemistry wins is a proxy for whether this industry is building spacecraft or building infrastructure**, and it is worth asking every company in the power layer which side of that they are on.

2. The number nobody in this sector wants to build up

Which brings us to the cost of the array, and to the single weakest link in every published orbital-compute economic model I have examined, including the ones I find persuasive.

Figure 8.5 — *The least-examined number in the entire thesis*



Three orders of magnitude separate what space arrays have historically cost from what gigawatt-scale orbital compute assumes they will cost. No published contract closes this gap. Treat any business plan that skips it with suspicion.

Bespoke space solar array *systems* — cells, blanket, structure, deployment mechanisms, qualification — have historically been quoted at figures on the order of **$1,000 per watt**, with thin-film approaches claiming a path below $200/W.[11] Terrestrial utility-scale modules are around fifteen cents.

At $1,000/W, the 120 kW array on Node A costs $120 million, which is forty times the cost of the computers it powers and roughly forty times the entire launch bill. The thesis dies instantly.

At the $3/W that gigawatt-scale orbital compute plans implicitly assume, that same array costs $360,000, and nobody thinks about it again.

Three orders of magnitude separate those two worlds, and no publicly disclosed contract closes the gap. That does not mean the gap cannot close — the historical figure reflects one-off programmes with heavy qualification burdens and no production volume, which is exactly the cost structure that mass manufacture destroys. Starlink already demonstrates that spacecraft components fall by orders of magnitude when built in the thousands. But it does mean this:

> **When you assess a company in this sector, find out what it assumes power hardware costs per watt, and whether that number comes from a quote or from a slide. This is the assumption most likely to be doing silent work in a business plan, and the one least likely to be challenged in a room full of people discussing radiators.**

It is also why the power layer is strategically more important than its share of the spacecraft mass suggests, and why the fund thesis's concentration there is defensible on structure rather than on enthusiasm.

3. The compute payload, and the ways vacuum changes it

Now the part that earns the money. A 100 kW payload at our convention is half a rack: roughly **700 kg and $1.75M** of accelerators, memory, storage and switching.

Four things change when you put it in vacuum.

No air means no air cooling anywhere. This is more disruptive than it sounds. Terrestrial servers use liquid cooling for the hot components and air for everything else — voltage regulators, memory, drives, switch ASICs. In vacuum there is no "everything else" path. Every component that dissipates meaningful power needs a conduction path to a cold plate, which means a substantially redesigned board and chassis, not a standard rack bolted to a

satellite. The vehicles being flown today are the beginning of that redesign, and it is one of the more underappreciated engineering programmes in the sector.

No convection means hot spots are permanent. On Earth, a badly cooled corner of a board is rescued by moving air. In vacuum, heat goes where the metal takes it, and nowhere else.

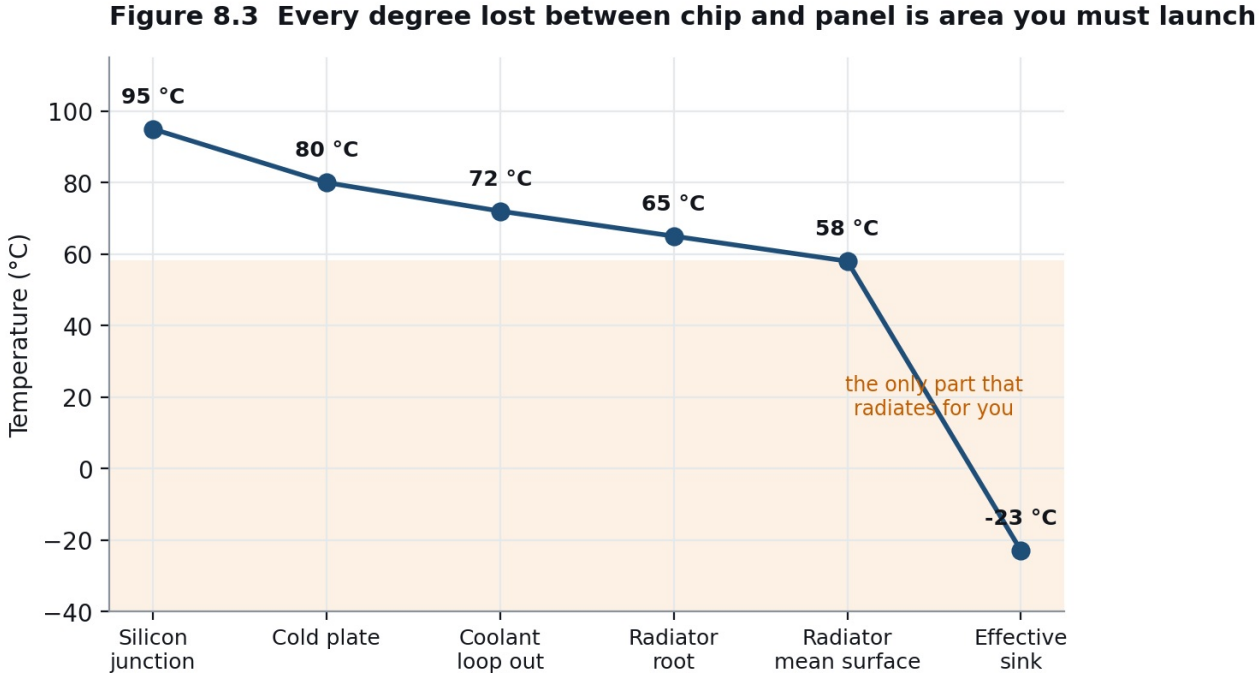
Outgassing and materials. Plastics, adhesives and lubricants release volatiles in vacuum, which then condense on the coldest, most sensitive surfaces available — typically your optics and your radiators. Every material in the payload has to be qualified for this, which quietly rules out large parts of the commercial supply chain.

Shielding is packaging. Chapter 6 established that you shield the box, not the ship. For Node A, roughly 10 mm aluminium-equivalent around a compact enclosure adds about **350 kg**. That figure scales with the *surface area* of the enclosure, not its contents, which is a strong argument for packing compute densely — the same square-cube logic that makes big things cheaper per unit than small ones.

4. The thermal chain, in detail

Chapter 5 sized the radiator. Here is everything between the chip and it.

Figure 8.3 — *Every degree lost between chip and panel is area you must launch*



Illustrative budget. A 37 °C drop from junction to radiator surface costs roughly 40% more radiator area than a design that loses only 15 °C — see Figure 5.1.

The junction runs at perhaps 95 °C. The cold plate touching it sits at 80. The coolant leaving the payload is at 72. By the time heat reaches the radiator root you are at 65, and the mean radiating surface — because a fin is never isothermal, and the tip is always cooler than the root — is nearer 58.

That 37-degree cascade is not waste in a moral sense; it is the price of moving heat. But look back at Figure 5.1 and price it. Radiating at 58 °C instead of 95 °C costs you roughly **40% more panel area** — which on Node A is about 40 square metres and 320 kilograms you launched purely to pay for temperature drops.

This is why the thermal layer is a real engineering business and not a plumbing subcontract. A company that shaves fifteen degrees out of that chain has removed a quarter of the radiator from the vehicle.

For Node A, the numbers are comfortable: 100 kW at around 58 °C surface temperature needs roughly **140 m² of two-sided radiator**, about 1,120 kg of panel and another 300 kg of pumps, fluid and plumbing. All of that is within the envelope of flight-proven pumped single-phase hardware.

For Node B, they are not comfortable. A megawatt needs roughly **1,400 m²** and pushes you into pumped two-phase transport, which is where flight heritage runs out. **The step from Node A to Node B is not a scaling exercise. It**

is a technology programme, and any company that describes it as the former is telling you something about itself.

5. Talking: optical terminals and the ground

Node A carries two or three optical terminals — roughly 90 kg all in — for inter-satellite links and, where the architecture allows, direct optical downlink. RF remains as a low-rate backup for command and telemetry, because you never want your only path to the spacecraft to depend on a mechanism holding a microradian aim.

Chapter 7 established that a compute node's own traffic is modest. What is not modest is the ground segment: a network of optical or Ka-band stations, sited for geographic diversity against cloud cover, is a capital programme in its own right, and it appears in almost no spacecraft cost estimate you will read. When someone quotes you dollars per GPU-hour in orbit, ask whether the ground stations are in it.

6. Pointing: three requirements that fight each other

Attitude control on a compute satellite is more interesting than on almost any other spacecraft, because it has three simultaneous pointing demands that are geometrically incompatible:

1. The **solar array** wants to face the Sun.
2. The **radiator** wants to face deep space and specifically to avoid facing the Sun or the Earth, since both warm it and Chapter 5 showed how expensive warm is.
3. The **optical terminals** want to face other spacecraft and, periodically, the ground.

You cannot satisfy all three with one rigid body, which is why real designs use gimballed arrays, gimballed terminals, and a body attitude chosen to keep the radiators edge-on to the Earth. That is three more mechanism sets, and mechanisms are where spacecraft fail.

The dawn–dusk sun-synchronous orbit is popular precisely because it partially resolves this fight: the Sun stays roughly perpendicular to the orbit plane, so a fixed geometry can keep arrays sunward and radiators edge-on to Earth without constant slewing. The choice of orbit is not a detail. **It is the decision that makes the rest of the spacecraft possible**, and it is a good early question for any company: *what orbit, and why that one?*

Hardware is modest — reaction wheels, star trackers, magnetorquers to dump accumulated momentum — around **120 kg** for Node A.

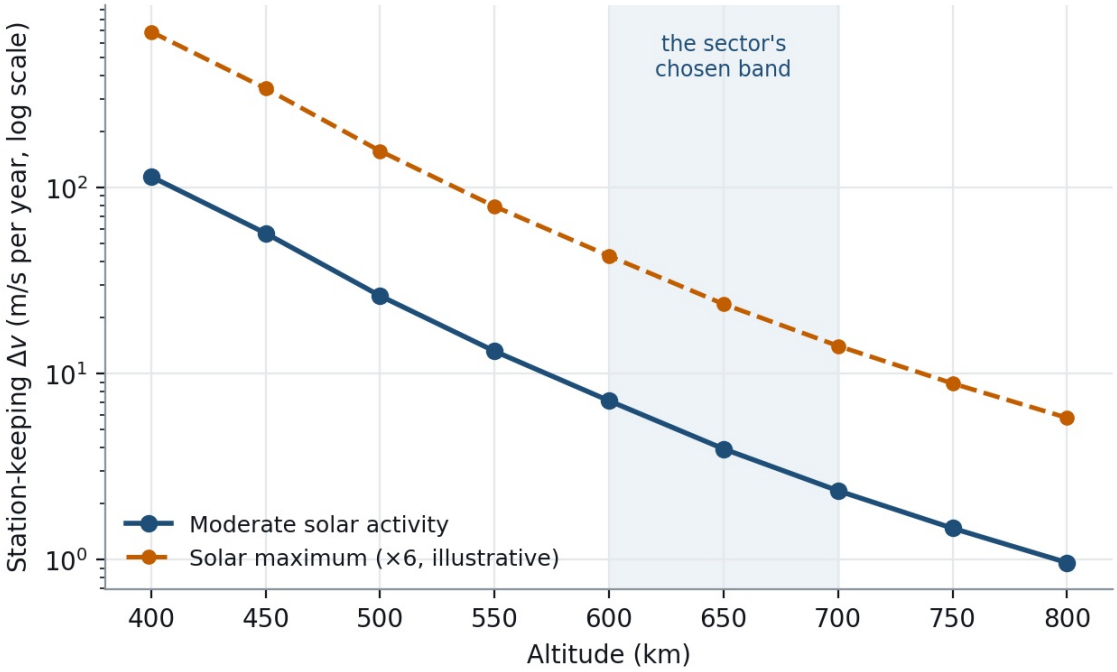
7. Staying up: drag and station-keeping

A compute node is mostly sail. Roughly 290 m² of array and 140 m² of radiator on a 4.7-tonne vehicle gives an area-to-mass ratio far worse than a conventional satellite, and low Earth orbit is not empty.

Work the drag calculation and the result is genuinely reassuring — with a sharp cliff at the bottom.

Figure 8.4 — *Why everyone is choosing 600–700 km*

Figure 8.4 Why everyone is choosing 600–700 km



Area-to-mass ratio 0.02 m^2/kg, drag coefficient 2.2. Atmospheric density varies by roughly an order of magnitude over the solar cycle, which is why the band matters more than the point value.

Altitude	Station-keeping Δv per year	Five-year propellant, Node A
400 km	~114 m/s	~180 kg
500 km	~26 m/s	~42 kg
600 km	~7 m/s	~11 kg
650 km	~4 m/s	~6 kg
800 km	~1 m/s	~2 kg

Assumes area-to-mass 0.02 m²/kg, drag coefficient 2.2, electric propulsion at 1,600 s specific impulse. Atmospheric density varies by roughly an order of magnitude over the eleven-year solar cycle; multiply accordingly for solar maximum.

At 650 km, keeping a compute node in place for five years costs a few kilograms of propellant. Drag is a non-issue. At 400 km it costs nearly 200 kg and a much larger thruster, and at solar maximum considerably more than that.

So the altitude choice is triple-constrained and lands in a narrow band. Below about 500 km, drag punishes you. Above about 700 km, orbital debris lifetime and radiation dose both rise, and you lose the natural decay that keeps the orbit self-cleaning. In between sits the band everyone is choosing, for reasons that are entirely physical rather than fashionable. Allow **150 kg** of thruster, tank and propellant on Node A.

8. Fitting inside the rocket

Node A stows into a conventional fairing without heroics. Node B does not stow into anything smaller than Starship, and even there the constraint is not mass but **folding**.

A megawatt node carries roughly 2,900 m² of array and 1,400 m² of radiator. As flat sheets a few millimetres thick, that is only tens of cubic metres — volumetrically trivial against a fairing measured in hundreds of cubic metres. The problem is that all of it must be creased, restrained through launch loads, and then released in sequence without a single jam, and the mechanism count grows with the number of folds.

This is the strongest engineering argument for **modularity**: rather than one 1 MW spacecraft with a thousand hinges, fly ten 100 kW nodes flying in formation and link them optically. Google’s design takes exactly this route, envisaging clusters of many satellites holding formation at separations of a few hundred metres.[8] It is also the

argument that eventually pulls in-space assembly and servicing from science fiction into the supply chain, which is Chapter 15.

9. Dying properly, and the problem nobody has solved

Two end-of-life issues, and the second is a business-model problem disguised as an engineering one.

Disposal. From 650 km, natural orbital decay takes far longer than the five-year post-mission disposal window that US regulation now expects, so a compute node must actively deorbit. Lowering perigee enough to guarantee re-entry costs on the order of 100 m/s — perfectly affordable, but it must be budgeted at the start and it must still work after five years of radiation. A dead 5-tonne spacecraft with 290 m² of sail area at 650 km is precisely the debris problem Chapter 16 is about.

Refresh. Here is the awkward one. Spacecraft are designed for lifetimes of five to fifteen years. AI accelerators are economically obsolete in roughly three. A terrestrial data centre solves this by rolling in new racks; the building outlives twenty generations of hardware.

An orbital data centre cannot do that today. It either accepts that it is flying depreciating silicon for its full mission life — which quietly destroys the return, since the revenue per rack collapses as newer hardware arrives — or it develops the ability to replace payloads on orbit, which is servicing, which does not commercially exist at this scale.

This, and not radiation or thermal, may turn out to be the sector's real structural weakness. It is the question I would put to every operator: *what is your hardware refresh strategy, and what does your model assume about revenue per rack in year four?* I have not yet seen a satisfying answer, and Chapter 15 is where the companies who could provide one live.

The build-up

Table 8.1 — Reference Node A: 100 kW of compute

Subsystem	Mass	Notes
Compute payload	700 kg	Half a rack at the 200 kW / 1.4 t convention
Enclosure and shielding	350 kg	~10 mm Al-equivalent around a compact box
Solar array	800 kg	120 kW BOL at ~150 W/kg
Power management and harness	250 kg	High-voltage PMAD, distribution
Battery	80 kg	Housekeeping only — full-sun orbit assumed
Radiators	1,120 kg	~140 m² two-sided at ~8 kg/m²
Heat transport	300 kg	Pumps, fluid, plumbing, cold plates
Structure and mechanisms	750 kg	Primary structure, deployment hardware
Attitude control	120 kg	Wheels, star trackers, magnetorquers
Communications	90 kg	Optical terminals plus RF backup
Propulsion and propellant	150 kg	Electric, five years at 650 km, plus deorbit
Total	**≈ 4,710 kg**	

Now the economics, and an admission.

Chapter 2 used 3,000 kg as a placeholder for this vehicle. The build-up says 4,710. That matters, because the placeholder made the launch bill look like half the value of the computers, and the real number does not:

	Falcon 9 (~$2,800/kg)	Starship contracted (~$600/kg)	Starship target (~$150/kg)
Launch cost, Node A	$13.2M	$2.8M	$0.7M
Versus $1.75M of computers	**7.5×**	**1.6×**	**0.4×**

At Node B the ratio improves — roughly 37 tonnes carrying $17.5M of computers, so about $22M of launch against

the payload, or **1.3×** — because structure and avionics scale sublinearly while compute scales linearly. Scale helps. It does not rescue.

The honest summary is that at contracted Starship pricing, **launch is roughly the same size as the hardware bill**, and the sector's economics only become comfortable at the target price nobody has yet paid. That is a materially more cautious conclusion than Chapter 2's placeholder implied, and it is why the correction is printed here rather than fixed silently.

Where each part breaks — and who is trying

Subsystem	The thing that breaks	Where the companies are
Solar array	Deployment mechanisms; cost per watt; high-voltage arcing in plasma	Ch 10 — power
Power management	Arcing, mass of distribution at high current	Ch 10 — power
Compute payload	Vacuum-compatible packaging; long-duration silicon behaviour	Ch 12 — radiation-tolerant compute
Heat transport	Two-phase flow at megawatt scale — **unsolved**	Ch 11 — thermal
Radiators	Deployment; micrometeoroid puncture; coating degradation	Ch 11 — thermal
Optical terminals	Microradian pointing; ground-segment capital	Ch 14 — comms
Attitude control	Three incompatible pointing demands; mechanism count	Ch 13 — platforms
Propulsion	Modest, provided you stay above 500 km	Ch 9 / Ch 15 — mobility
Disposal and refresh	**No commercial answer to hardware refresh on orbit**	Ch 15 — servicing

Two rows in that table have no adequate supplier: megawatt-scale heat transport, and on-orbit hardware refresh. Everything else has at least one credible company attacking it.

That asymmetry — one badly-served constraint and one unsolved business model, in a sector where everything else is buyable — is the map. Part IV walks it.

Source notes, Part III

0. Rocket Lab's February 2026 introduction of silicon solar arrays aimed at gigawatt-scale space-based data centres, positioned on cost per watt at manufacturing scale rather than efficiency — Rocket Lab Corporation press release, 26 February 2026.
1. Space solar array system costs: a figure on the order of $1,000/W for conventional systems with maximum specific power around 106 W/kg at beginning of life, against thin-film claims below $200/W — Small Satellite Conference paper, *Self-Deploying Thin-Film PV Solar Array Structure*. **This source is old and the figure is almost certainly high for a modern mass-produced array. It is used here to establish the size of the gap, not the current price. Finding a defensible 2026 figure is the single highest-value piece of remaining research for this book, and quite possibly for the fund.**
2. (continued from Part II) Cluster architecture with satellites holding formation at separations of a few hundred metres in dawn–dusk sun-synchronous orbit near 650 km — Google Research, Project Suncatcher.

Reference-node assumptions requiring sourcing before publication: rack power and mass convention (200 kW, 1.4 t, $3.5M); array specific power (150 W/kg); radiator areal mass (8 kg/m²); atmospheric density profile used in Figure 8.4. Each is stated in the text where used so a reader can substitute their own.

Figures 8.1 to 8.5 generated for this volume from the assumptions stated above.

PART IV — THE SECTOR SKELETON

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part IV of VI — chapters 9 to 11.

Parts I to III built the argument and the machine. This part takes the machine apart into layers and asks, for each one, the same five questions:

1. **Who thought of this first**, and why it stayed on paper.
2. **What the physics says** — the constraint, from first principles.
3. **What actually breaks** in hardware.
4. **Who is attacking it**, and where each bet is fragile.
5. **What to watch** — the specific event that tells you the layer is working.

A note on how companies appear here. Named companies date faster than anything else in a book, so the *judgement* in each chapter is about the layer, and the companies are evidence for it. Where I hold a position, it is disclosed in the appendix. Where a fact comes from a company rather than an independent source, it is labelled as a claim rather than a result. And where I am relying on my own fund's research rather than a public document, I say so, because you should discount it accordingly.

Chapter 9 — Launch and cargo

The old idea

In 1903, a provincial Russian schoolteacher named Konstantin Tsiolkovsky published the equation that still governs every launch:

$\Delta v = I_{sp} \cdot g \cdot \ln(m_0 / m_f)$

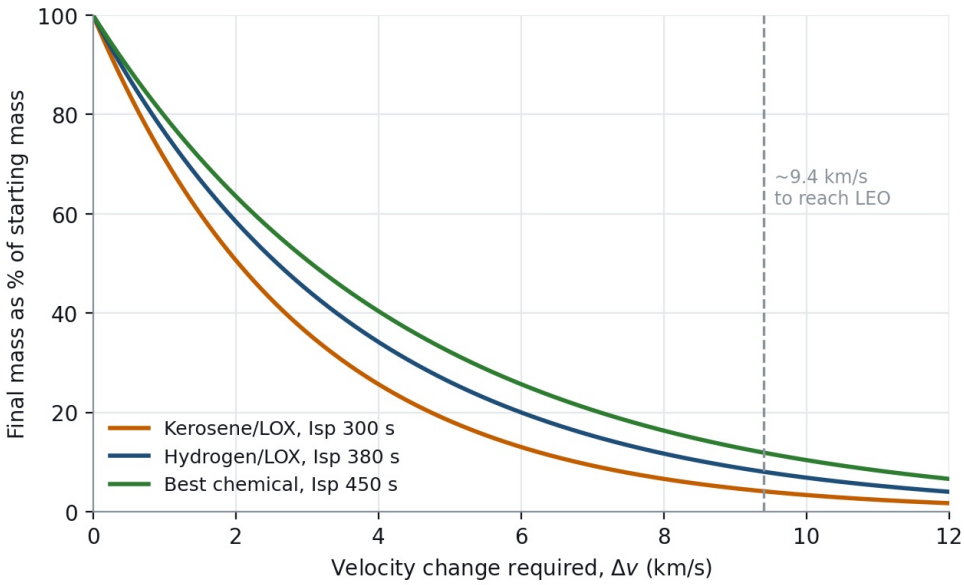
The velocity you can gain equals your engine's efficiency multiplied by the natural logarithm of how much mass you started with divided by how much you finished with. Verne had already fired a crew at the Moon from a cannon in 1865, and Goddard would fly the first liquid-fuelled rocket in 1926. But Tsiolkovsky's line is the one that mattered, because it converted a dream into an accounting problem with a very unpleasant term in it.

The physics

The unpleasant term is the logarithm. To gain velocity linearly, mass must grow exponentially.

Figure 9.1 — *The tyranny Tsiolkovsky described in 1903*

Figure 9.1 The tyranny Tsiolkovsky described in 1903



$\Delta v = I_{sp}\, g \ln(m_0/m_f)$. Reaching orbit leaves roughly 4–8% of liftoff mass, and that residue is structure plus payload, not payload alone. The curve is exponential, which is why a century of engineering moved the price by reuse rather than by propulsion.

Reaching low Earth orbit takes roughly 9.4 km/s once you include drag and gravity losses. Put that into the equation with the best chemical propulsion available and you finish with something like four to eight percent of your liftoff mass — and that residue is not payload. It is tanks, engines, structure, avionics *and* payload. Payload alone is typically two to four percent.

Here is the important consequence, and it is why this chapter is short. **Nobody is going to improve this.** Chemical propulsion is within a few percent of its theoretical energy limits, and has been for decades. There is no propulsion breakthrough coming that changes the cost of getting to orbit.

Which means the entire hundredfold cost reduction described in Chapter 2 came from somewhere else: not from making the rocket better, but from **not throwing it away**. Reuse does not beat the rocket equation. It amortises it. That distinction is worth holding, because it tells you where further gains can and cannot come from — cadence, refurbishment cost and manufacturing volume, not physics.

What breaks

Three things, none of them thermodynamic.

Cadence. A reusable vehicle only pays back if it flies often. The economics of the entire orbital compute sector are downstream of how many times per year a heavy vehicle can be turned around, and that is an industrial and regulatory problem — pads, range availability, environmental review — as much as a technical one.

Fairing volume, not mass. Chapter 8 showed that a megawatt node is volumetrically modest but mechanically complex. As launch mass gets cheap, the binding constraint moves to how much *folded structure* you can pack and reliably unfold. Vehicle diameter starts to matter more than vehicle payload.

The last mile. A rocket drops you where its trajectory ends, not where your constellation needs to be. Chapter 8's thermal and radiation analysis pushed us into a narrow altitude band and a specific orbit plane; getting a heavy node into precisely that plane, and later moving or disposing of it, is a separate vehicle and a separate business.

Who is attacking it

SpaceX is the layer, in a way that has no parallel elsewhere in this book. It sets the price everyone else models against, it has disclosed the only Starship contract price in the public record, and it is simultaneously a competitor in orbital compute — having filed for a very large constellation of compute satellites and merged with an AI company. Any honest treatment of this sector has to describe SpaceX as both the enabling infrastructure and the largest single competitive risk to everyone using it.

Rocket Lab and **Stoke Space** matter as the second sources. A sector whose economics rest on one provider's price list is a sector with a single point of commercial failure, and the value of a credible second heavy vehicle is greater

than its market share suggests.

Impulse Space is the last-mile answer: orbital transfer vehicles that take a payload from where the rocket left it to where it needs to be, in hours rather than months of natural drift. Founded by SpaceX's first employee, revenue-generating, with flight heritage and a large backlog.[12]

The investment stance, stated plainly

I do not think launch is where the returns are, and I do not hold it. Not because it will do badly — it may do very well — but because the largest player is about to be, or already is, a publicly discoverable asset, and the second tier is public too. Whatever excess return existed in launch has been competed into the open.

The last mile is different. It is a smaller market, it is not price-discovered, and it becomes structurally more valuable exactly as the thermal and radiation constraints of Chapters 5 and 6 push compute constellations into specific, awkward orbits. **Value in this layer has moved from getting to space to getting somewhere specific in space.**

What to watch

- Starship flying paying customers at or near the contracted price, rather than the target price.
- Turnaround time between heavy flights — the number that determines whether $600/kg becomes $200/kg.
- Whether a second provider reaches heavy-lift reuse. Until one does, every model in this book has a single-supplier risk sitting under it.

Chapter 10 — Power

The old idea

This is the layer with the deepest intellectual pedigree, and it is worth spending a paragraph on because it reframes what the sector is doing.

In 1964 the Soviet astronomer **Nikolai Kardashev** proposed classifying civilisations not by their politics or their technology but by the energy they command: a Type I civilisation uses the energy available on its planet, a Type II the output of its star. The scale was proposed as a tool for searching for extraterrestrial intelligence, but its underlying claim is the one that matters here — **that the ceiling on what a civilisation can do is set by the power it can gather, and that the next available increment is not on the planet.**

Four years later, in 1968, the engineer **Peter Glaser** published the solar power satellite concept: collect sunlight in orbit where it is uninterrupted, and send the energy down. Gerard **O'Neill** built an entire architecture on this premise in *The High Frontier* in 1976. Asimov had already written the fictional version in 1941.

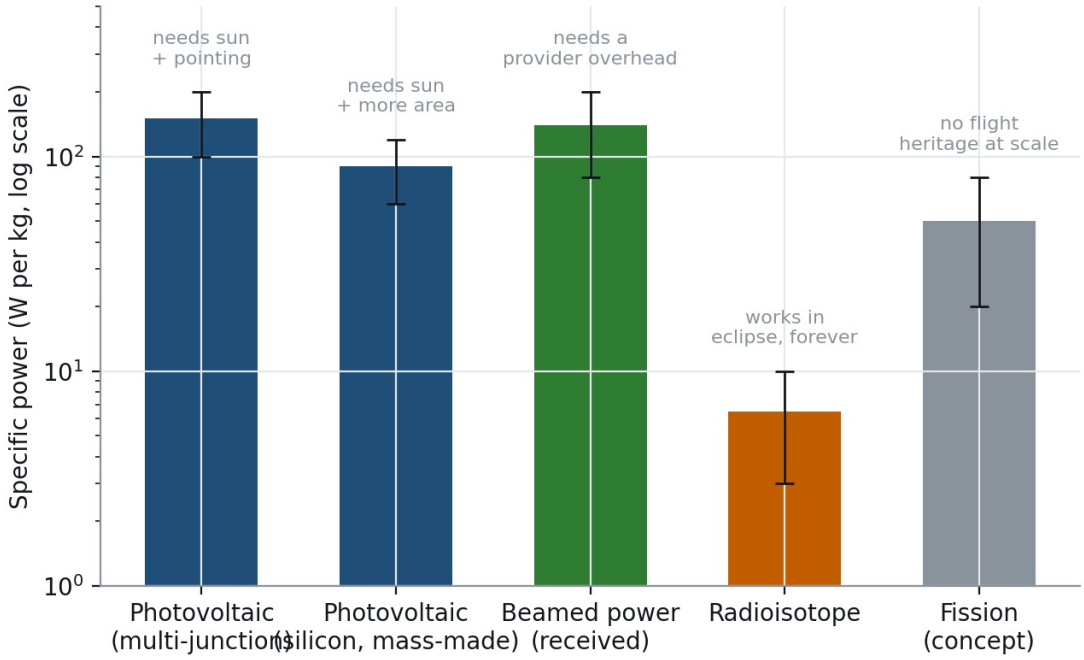
None of it happened, for one reason. At Space Shuttle prices, putting a kilogram in orbit cost roughly fifty thousand dollars, so the energy hardware could never repay its ride. **The idea was never wrong. It was priced out.** Chapter 2 is the story of that price changing, and this chapter is the first place where the change bites.

The physics

Four ways to have a watt in orbit, and they differ on axes that matter more than efficiency.

Figure 10.1 — *Four ways to get a watt in orbit, and what each one costs you*

Figure 10.1 Four ways to get a watt in orbit, and what each one costs you



Ranges are order-of-magnitude for current or near-term hardware at the spacecraft level. Specific power is only one axis: cost per watt, eclipse behaviour and deployment risk decide most real architectures.

Photovoltaic is the default: high specific power, no fuel, but useless in eclipse and dependent on pointing and deployment. Within it sits the multi-junction versus silicon argument from Chapter 8 — efficiency per square metre against cost per watt at manufacturing scale.

Beamed power is the interesting one, and it deserves more attention than it gets. The insight is not that beaming sunlight is more efficient — it is not, since you pay conversion losses twice. The insight is **who owns the receiver**. If you concentrate and beam sunlight onto a client spacecraft's *existing* solar array, the customer needs no new hardware at all: no retrofit, no custom receiver, no integration programme. That converts a hard sale into an easy one, and it turns a power company into shared infrastructure serving every operator rather than a supplier picking a winner.

Radioisotope power is the opposite trade: a few watts per kilogram, which is terrible, in exchange for output that does not care about eclipse, sun angle, distance from the Sun, or dust. It will never power a megawatt data centre. It is the correct answer for anything that must survive the dark.

Fission is the only technology that could plausibly deliver megawatts without square kilometres of array — and the thermal analysis of Chapter 5 applies to it doubly, since a reactor's waste heat must also be radiated. There is no flight heritage at the relevant scale.

What breaks

Everything in Chapter 8's power section, plus the number that chapter refused to let pass: **cost per watt**. Recall Figure 8.5 — historical bespoke space arrays on the order of $1,000/W against an implicit assumption near $3/W in gigawatt-scale plans. That gap is this layer's central commercial question, and any company here that cannot answer it precisely is selling a story.

Beyond it: deployment mechanisms, high-voltage arcing in LEO plasma, radiation degradation of cells, and — for beamed power — pointing accuracy and conversion efficiency at range.

Who is attacking it

K2 Space is the clearest structural bet in the layer, because it does not sell power as an accessory to a satellite; the satellite *is* the power system. A large bus designed from the start around high power, high-voltage avionics and a high-power thruster, sized for what heavy lift now permits rather than for what a Falcon-class fairing used to permit. It has raised at a multi-billion valuation against substantial signed contracts and has won a programme of

record.[12]

Rocket Lab entered this layer in February 2026 with silicon arrays explicitly aimed at gigawatt-scale space data centres, betting on manufacturability over efficiency.[10]

Star Catcher is the beamed-power bet described above — concentrating and beaming solar energy onto clients' existing arrays, with signed power purchase agreements ahead of first flight and multiple orbital-compute companies among its named customers.[12] Its founder previously built and sold an in-space manufacturing company, which is the repeat-founder profile that historically predicts more than any other single signal.

Zeno Power is the radioisotope bet, and its moat is unusual: not physics — radioisotope power is sixty years old — but a **fuel supply chain**. Recycled strontium-90 and americium-241 with contracted industrial partners is not something a competitor replicates by raising money faster.[12]

Aetherflux sits across two layers, generating power and consuming it in its own compute payload. Strong founder, top-tier syndicate, and — at the time of writing — nothing yet demonstrated in orbit, which is the correct reason to hold it smaller than its narrative would justify.

The investment stance

Power is where I am most concentrated, and the reason is structural rather than enthusiastic.

Every operator in this sector needs watts, and none of them agree on whose compute wins. A supplier that sells to all of them does not require me to pick the winner in the layer above — which matters enormously over a five-year hold through a period when the operator layer will consolidate. Chapter 8 also showed that power hardware carries the single largest unresolved cost in the whole thesis, and unresolved costs are where pricing power lives.

The caution: this layer's valuations are the most sensitive to the $/W question. If mass-produced arrays arrive at $3/W, the power layer becomes a commodity and the margin moves elsewhere. If they arrive at $100/W, the sector's gigawatt plans do not happen and the layer is small. Both outcomes are bad for a naive long position, which is why the specific companies here are chosen for moats — a supply chain, a shared-infrastructure position, a bus architecture — rather than for exposure to the theme.

What to watch

- **The first published dollars-per-watt for a mass-produced orbital array.** This single number moves more of the model than anything else in this book.
- A successful spacecraft-to-spacecraft power beaming demonstration on orbit. It has never been done commercially; doing it creates a new category.
- Whether high-power buses win programmes of record against high-rate manufacturers. That contest tells you whether power or production volume is the durable differentiator.

Chapter 11 — Thermal

The old idea

In 1960, **Freeman Dyson** published a two-page paper in *Science* proposing that an advanced civilisation would eventually capture the full output of its star, and that we could search for such civilisations by looking for the infrared they radiate.

Read that again, because it is the most under-appreciated sentence in the history of this subject. **Dyson's insight was not the sphere. It was the waste heat.** He understood immediately that any structure gathering enormous energy must dump an almost equally enormous amount of heat, and that the dumping — not the gathering — would be the visible signature. The first person to think rigorously about civilisation-scale computing around a star concluded that you would find it by its radiators.

Science fiction then spent sixty years forgetting this. Almost no spacecraft in film has radiators. They have engines, weapons and windows, and no way at all to get rid of the heat those imply. The omission is so consistent that it functions as a useful test: **when someone shows you a rendering of an orbital data centre, look for the radiators, and if they are not obviously the largest thing in the picture, the picture is fiction.**

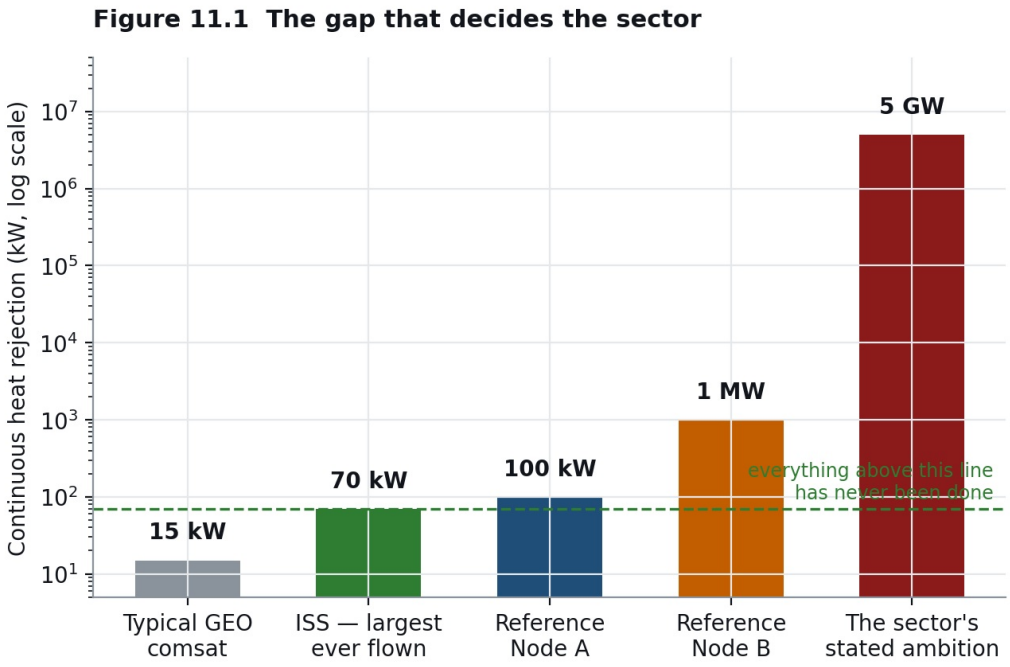
The physics

Chapter 5 did the arithmetic; this chapter needs only its two conclusions. Rejection scales as the fourth power of temperature, so running hotter is the dominant lever. And radiator area is not really the hard part — **getting a megawatt of heat from the silicon to the panel is.**

There is a third lever, largely unexploited, that follows directly from the fourth-power law: **actively pumping heat uphill.** A heat pump consumes electrical power to raise the rejection temperature above the source temperature. It sounds wasteful — you are adding heat to get rid of heat — but because area falls as T^4, raising the radiator from 27 °C to around 100 °C can cut radiator area by roughly sixty percent and system mass by roughly a third even after accounting for the pump's own power and its own waste heat.[13] In a vehicle where radiators are the largest mass item, that is an enormous trade, and it is available today in principle. That it is not yet a product tells you how immature this layer is.

What breaks — and the number that frames the whole sector

Figure 11.1 — *The gap that decides the sector*



The International Space Station rejects 70 kW through two pumped ammonia loops and six radiator assemblies — the largest active thermal control system ever flown. A single reference compute node exceeds it. The stated ambition exceeds it 70,000-fold.

The International Space Station rejects **70 kilowatts**, through two independent pumped ammonia loops and six deployed radiator assemblies.[14] That is the largest active thermal control system ever flown by anyone.

Reference Node A — a single modest compute satellite, half a rack of accelerators — needs 100 kW. **The entry-level vehicle in this sector exceeds the largest thermal system in the history of spaceflight.** The megawatt node needs fourteen times the ISS. The gigawatt-scale ambitions in the sector's public filings need tens of thousands of times.

One more comparison, because it cuts both ways. The ISS rejects its 70 kW from roughly 420 square metres of panel — about 166 watts per square metre. Chapter 5's numbers give a data centre radiator at 58 °C roughly twice that per square metre. The difference is simply temperature: the ISS runs its loops near room temperature because it contains humans. **A machine that contains no humans can run hot, and that is the sector's single largest structural advantage over every thermal system previously flown.** The physics is on the sector's side here. The hardware is not there yet.

What specifically is missing:

- **Two-phase pumped transport at megawatt scale.** Heat pipes top out in the low kilowatts. Single-phase pumped loops are proven to tens of kilowatts. Two-phase is the right answer above a megawatt and has no flight heritage at that scale, and managing boiling flow in microgravity — where vapour does not rise — is genuinely

hard.

- **Deployment of a thousand square metres** of thin panel, reliably, once.
- **Micrometeoroid and debris puncture** of the largest, thinnest, fluid-filled surface on the vehicle. The ISS radiators were designed with debris protection precisely because this is a certainty over years, not a possibility.
- **Coating degradation**, which attacks emissivity, which sits inside the fourth-power law.

Who is attacking it

Here the chapter becomes unusual, because the honest answer is: **almost nobody you can invest in.**

Spacebilt has the best technology fit I have found — thermal tiles rejecting heat to deep space, already supplying orbital data-centre nodes for a commercial station operator. It has no visible venture round, no accessible cap table, and no mark. It is a supplier and an acquisition target, not an investment.

Sophia Space is the closest thing to an investable pure-play. Its TILE architecture makes thermal integration the product rather than a subsystem — modules combining solar generation on one face with radiative cooling on the other, racked together into compute. It is seed-stage with first customer deliveries targeted for 2028.[12] The strategic ambiguity is worth naming: the same tile can be sold *to* orbital data centre operators or assembled *into* a competing one, and only the first configuration is a supplier business.

Celeroton and companies like it — oil-free high-speed turbomachinery for cryocoolers and heat pumps — are exactly the technology the heat-pump argument above requires. They are European industrial suppliers with no venture round to buy into.

The investment stance, and why the emptiness is the finding

The layer that most determines whether this sector works has the fewest investable companies in it. That is not an oversight on my part; it is a structural observation, and it admits three readings.

The optimistic reading: this is a gap, and gaps get filled. A well-founded thermal company with a megawatt-class two-phase loop would have every operator in the sector as a customer.

The realistic reading: thermal is so central to the vehicle that operators will build it in-house rather than buy it. Cooling is not a component you bolt on; it shapes the whole spacecraft. If that is right, the value in this layer accrues to the operators, and the correct exposure is through them.

The pessimistic reading: the layer is empty because the problem is harder than the sector admits, and the people who understand it best are not starting companies to solve it.

I hold the middle reading, and I act on it by reserving capital against the layer conditionally rather than committing to it — funded only against a demonstrated supply agreement with a named operator, because a thermal company that is not selling to operators is competing with them.

What to watch

- **The first commercial deployable radiator and direct-to-chip liquid loop operating on orbit at rack scale.** This is the nearest genuine test of the architecture and it is imminent.
- Anyone announcing a pumped **two-phase** loop qualified above a few hundred kilowatts. That announcement, from anyone, reprices the whole sector.
- Anyone shipping a **space-rated heat pump**. It is the highest-leverage unbuilt product in this book.
- Conversely: an operator quietly raising its published radiator area per megawatt is telling you the ΔT chain in Figure 8.3 is worse in practice than in analysis.

Chapters 12 to 17 — radiation-tolerant compute, the compute platform itself, communications and ground segment, in-space assembly and servicing, debris and insurance, and innovation outside the United States — follow in the next instalment.

Source notes, Part IV (chapters 9–11)

2. Company-level facts in these chapters — funding rounds, valuations, contract values, customer names, flight

history and staffing — are drawn from the StarCap Fund I Thesis v2 research file (27 July 2026) and its underlying sources. **These are the author's own fund research and should be independently verified against primary sources before publication.** Valuations and round terms in this sector move quarterly; any figure printed in a book will be stale, which is why the chapters above are written to be read for structure rather than for marks.

3. The heat-pump trade — raising rejection temperature by active pumping to cut radiator area by roughly 60% and system mass by roughly a third — is worked through in the spacecraft adsorption thermal storage patent literature, which computes it for a 10 kW system rejecting at 375 K against a 300 K baseline. The arithmetic is general and follows directly from Chapter 5's fourth-power relationship.
4. ISS External Active Thermal Control System: two independent single-phase ammonia loops of about 35 kW each, six deployed radiator assemblies of eight panels at 3.33 × 2.64 m, 70 kW total — NASA ISS ATCS overview and the *Journal of Spacecraft and Rockets* radiator performance literature. Radiator area and the resulting ~166 W/m² are computed from those panel dimensions in this volume.
5. (continued from Part III) Rocket Lab silicon solar arrays for space-based data centres, 26 February 2026.

Historical sources to cite properly before publication: Tsiolkovsky (1903); Kardashev's civilisation classification (1964); Glaser's solar power satellite concept (1968); O'Neill, *The High Frontier* (1976); Dyson, *Search for Artificial Stellar Sources of Infrared Radiation*, Science (1960). These are quoted from general knowledge here and each needs a direct citation.

Figures 9.1, 10.1 and 11.1 generated for this volume.

PART IV — THE SECTOR SKELETON (continued)

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part IV of VI — chapters 12 to 14.

Chapter 12 — Radiation-tolerant compute

The old idea

In 1956 John von Neumann gave a series of lectures with a title that reads like a description of this entire chapter: *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components*. His question was how to build a computer you can trust out of parts you cannot. His answer — redundancy, voting, and accepting error as a design input rather than a defect — is the intellectual ancestor of everything that follows.

Richard Hamming had already published error-correcting codes in 1950. Triple modular redundancy, where three processors compute the same thing and majority-vote the answer, became the standard spacecraft technique. The Apollo Guidance Computer flew with 2,048 words of erasable memory and a design philosophy of graceful degradation.

The relevant point for us is that **the fault-tolerance problem was solved conceptually before the transistor was widely deployed.** What changed is not the theory. It is the economics of which parts you apply it to.

The physics

Chapter 6 laid out the two damage modes. Here is the part that decides architecture.

Historically, shrinking transistors made parts *less* sensitive to accumulated dose — thinner gate oxides trap less charge. That was a gift to the industry: each new process node was, roughly, more radiation-tolerant than the last. Google’s own analysis notes that this favourable trend does not continue below the 5 nm node.[15]

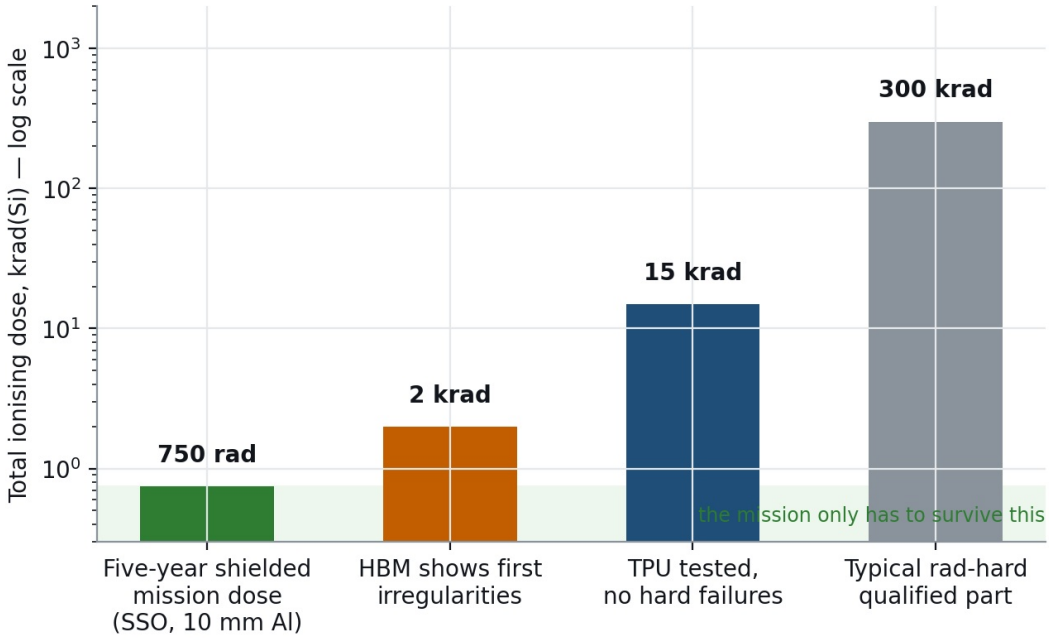
Meanwhile the opposite happens for single events. Smaller, lower-voltage transistors hold less charge per stored bit, so it takes less energy from a passing particle to flip one. Density compounds it: more bits per square centimetre means more targets.

So modern silicon is, crudely, **better at surviving the slow damage and worse at surviving the fast one** — which is exactly the profile that favours the hyperscale answer. Accumulated dose you cannot fix in software. Bit flips you can.

What the numbers actually say

Figure 12.1 — *Commercial silicon has more margin than the sector assumed*

Figure 12.1 Commercial silicon has more margin than the sector assumed



Mission dose, HBM sensitivity threshold and maximum tested dose from Google's published Trillium TPU proton-beam results. Rad-hard qualification level shown for scale; those parts run generations behind commercial silicon in performance.

Put the published figures on one axis and the argument becomes visual. The five-year shielded mission dose in the reference orbit is about 0.75 krad. High-bandwidth memory — the most sensitive subsystem — showed its first irregularities near 2 krad. The tested part reached 15 krad with no hard failures attributable to dose.[9]

The mission needs to clear the leftmost bar. The evidence says commercial silicon clears it with roughly a factor of three of margin on its weakest component, and considerably more on the processor itself.

That is a smaller margin than it looks, and it is worth being precise about why. Three times is not much when the dose estimate itself depends on solar cycle phase, exact orbit, and how much shielding survives the mass budget review. It is also a *dose* margin, and dose was never the thing that was going to kill you suddenly.

What breaks

Latch-up, which destroys hardware rather than corrupting it, and is a power-electronics problem: detect the current excursion, cycle the device, restart. Every operator in this sector is quietly building this capability and almost none of them talk about it.

Memory, which is both the most sensitive component and the largest by area in a modern accelerator. If there is a hardware surprise in this sector over the next five years, my expectation is that it comes from HBM stacks rather than from logic.

The absence of data. Nobody has multi-year on-orbit statistics for frontier accelerators. Beam tests simulate a dose; they do not simulate five years of thermal cycling, vibration-settled solder, coolant contact and unattended operation.

Who is attacking it — and why I do not own any of them

Google is the only organisation designing its own silicon with orbit in mind, and it has published the testing to prove it, alongside a design for satellite clusters and a first launch of prototypes with an Earth-imaging partner.[8] **Nvidia** has flown commercial accelerators in orbit and now offers a radiation-tolerant line built on commercial silicon rather than bespoke rad-hard parts. **Ramon.Space** and **Microchip** occupy the traditional space-processor niche, the latter supplying processors to commercial station operators.

Now the uncomfortable conclusion, which is the reason this chapter exists in a book about investing.

This layer has been solved by companies you cannot buy at venture scale, and that is a good outcome, not a missed one. The radiation problem is being retired by two of the largest semiconductor organisations on Earth, funded by AI revenues that dwarf this sector, and they are publishing the results. A private radiation-hardened

silicon company competing with that is a poor investment regardless of how good its engineering is.

A guidebook that only tells you where to invest is a sales document. **This is a layer to understand and not to own.** Its progress is the strongest de-risking event the sector has had — Chapter 6's reframe holds, that radiation is a depreciation schedule rather than a wall — and you capture that progress by owning the operators and suppliers whose economics improve because of it.

What to watch

- The first published **on-orbit** single-event rates for a frontier accelerator, as opposed to beam-test results.
- Any evidence on HBM behaviour over years rather than hours.
- Whether the sub-5 nm dose trend Google flagged turns into a practical ceiling on which generations of silicon can fly.

Chapter 13 — The compute platform

The old idea

In October 1945, in *Wireless World*, a young Arthur C. Clarke pointed out that three stations in geostationary orbit could cover the whole planet, and worked out the orbital mechanics to prove it. He did not patent it. The paper is the founding document of the idea that **infrastructure — not exploration, not science, but plumbing — is a reason to go to space.**

Everything in this chapter is a descendant of that observation. The satellite is not the mission. It is the building.

The physics — three architectures, and the trade between them

There are only three ways to arrange compute in orbit, and each is a different answer to Chapter 8's mechanism problem.

The monolith. One large spacecraft, one enormous radiator, one very large array. Fewest interfaces, best mass efficiency, worst deployment risk, and a single failure takes everything.

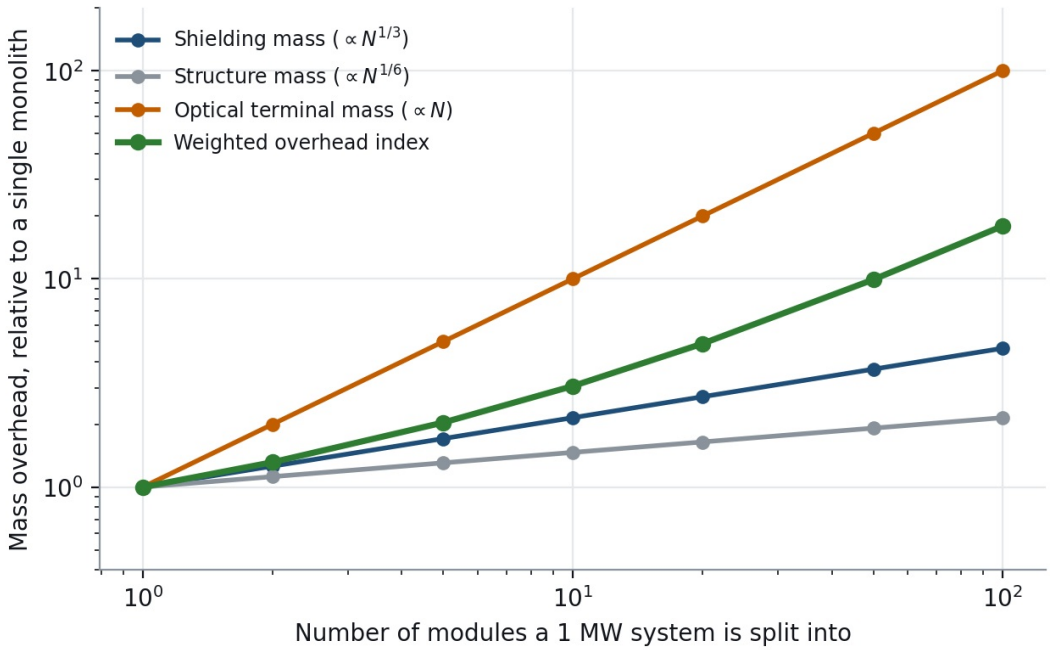
The modular cluster. Many small identical units flying in formation, linked optically, behaving as one machine. This is Google's approach — clusters of many satellites holding formation at separations of a few hundred metres.[8] It is also, in a different form, the tile architectures.

The hosted payload. Put the compute on somebody else's platform — a commercial station, or a standardised bus flying other customers' payloads. Lowest capital, lowest control, fastest to orbit.

The trade between the first two is quantifiable, and I have not seen it done in public, so here it is.

Figure 13.1 — *Modularity is bought, not free*

Figure 13.1 Modularity is bought, not free



Radiator area is indifferent to how you split the system; shielding, structure and inter-module links are not. What modularity buys back — deployment simplicity, launch granularity, graceful failure — does not appear on this chart.

Radiator area is indifferent to how you split the system — a megawatt needs its square metres whether it is one machine or a hundred. But three things are not indifferent:

- **Shielding** scales with enclosure *surface area*, and surface grows as the two-thirds power of volume. Split one megawatt into ten modules and total shielded surface rises by about **2.2×**; into a hundred, by about 4.6×. This is simply the square-cube law, and it is the strongest physical argument for packing compute densely.
- **Optical terminals** scale roughly linearly with module count, because every module must talk to its neighbours.
- **Structure** scales mildly.

There is a fourth, which Chapter 5's self-heating section quantifies: a cluster's members occupy part of each other's radiator sky, so every module needs a few percent more panel than the same power would need flying alone. **Formation spacing is a thermal decision before it is a control decision.**

So the monolith is mass-optimal and the cluster is not. What the cluster buys in return does not appear on that chart: simpler deployment per unit, the ability to launch incrementally as capital allows, graceful degradation when one unit dies, and — critically for Chapter 8's refresh problem — the possibility of replacing a fraction of the fleet each year with current silicon rather than flying one ageing monolith for a decade.

Read that last point carefully, because it may be decisive. The modularity penalty is a one-time mass cost of roughly two. The refresh problem is a recurring revenue cost that compounds. If I had to bet on which architecture wins, I would bet on the one that can be upgraded, and pay the two.

What breaks

Formation-keeping propellant and control for clusters. Deployment for monoliths. For hosted payloads, dependence on a platform operator whose priorities are not yours. And for all three, the refresh problem from Chapter 8, which nobody has solved.

Who is attacking it

Starcloud is the flight-heritage leader: it flew a commercial accelerator in orbit in late 2025 and trained a model there, and its second vehicle carries a newer accelerator with roughly ten times the compute and a hundred times the power generation, using direct-to-chip liquid cooling into deployable radiators, with booked workloads from named cloud customers.[12] That combination — flown hardware, real customers, and the first serious test of the thermal architecture at rack scale — makes it the single most informative company in the sector regardless of whether you own it.

SpaceX with xAI is the scale competitor: a filing for a very large compute constellation, captive launch, and its own AI workloads to fill it. Chapter 9 already made the point that the sector's enabler is also its largest competitive risk. The strength test is whether an independent operator is distinguished enough that SpaceX cannot erase it by announcement.

Google is the research leader, with published physics and prototype satellites planned.

Axiom Space is the hosted-platform route, with data-centre nodes already on orbit and a business that is cross-subsidised by existing revenue — which in a sector full of pre-revenue narratives is a genuinely different risk profile.

Sophia Space is the tile bet, and **Lonestar** occupies the lunar storage niche. **Madari** and similar sovereign-backed entrants matter more than their size suggests, for reasons Chapter 17 takes up. And there is a tail of very early entrants with enormous constellation filings and no hardware, which is where the sector's promotional energy concentrates.

The investment stance

Ambition is not a differentiator in this sector. Flight heritage is. A hundred-thousand-satellite filing costs a lawyer's fee. Getting one rack to work in vacuum for a year costs everything the company has.

I hold this layer, deliberately, including companies that compete directly with SpaceX — but only where something has flown, and never at a size where being wrong about which architecture wins is fatal. The reason to own it despite the competitive risk is that this is where the asymmetric outcomes are: if orbital compute works at all, the operators capture the revenue, and an acquisition by a larger player is a good outcome rather than a capped one.

What to watch

- **A second-generation compute satellite running paying workloads on orbit.** This is the nearest thing the sector has to a verdict, and it is imminent.
- Anyone publishing an actual price for orbital GPU-hours. Until someone does, every economic model in this sector — including mine — is theory.
- Whether cluster architectures demonstrate formation-keeping at the separations their designs assume.

Chapter 14 — Communications and the ground segment

The old idea

In 1880 Alexander Graham Bell transmitted speech on a beam of sunlight, using a vibrating mirror and a selenium receiver. He called the photophone his greatest invention — greater, he said, than the telephone. It was useless at the time, because the atmosphere and the technology were not ready, and it took a century for the laser and the optical fibre to prove him right.

Optical communication in space is the same idea with the atmosphere removed, which is the one place it was always going to work best.

The physics

A beam spreads at roughly the wavelength divided by the aperture. That single relationship explains why this layer went optical.

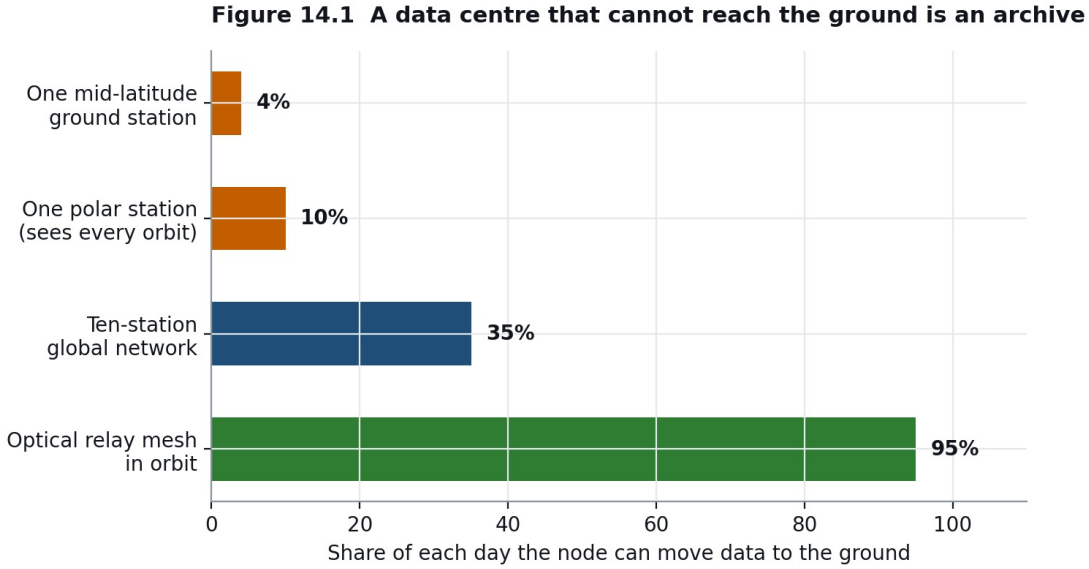
> **WORKED EXAMPLE — Why lasers and not radio**
>
> **Optical:** 1,550 nm wavelength through a 10 cm telescope gives a divergence of about **15 microradians**. At 1,000 km range, the beam is about **15 metres** across.
>
> **Radio:** Ka-band at 30 GHz — a 1 cm wavelength — through a 1 metre dish gives about **10 milliradians**. At the same 1,000 km, the beam is about **10 kilometres** across.
>
> The optical beam concentrates the same transmitted power into an area roughly **400,000 times smaller**. That is the entire argument, and no amount of clever modulation closes it.

The price is that you must hit a 15-metre spot from 1,000 kilometres away, from a vibrating spacecraft, while both

ends are moving at 7.5 km/s. Optical terminals are precision mechanisms with lasers attached, and that is why they are a product rather than a component.

The constraint nobody budgets for

Figure 14.1 — *A data centre that cannot reach the ground is an archive*



Illustrative geometry for a ~650 km sun-synchronous node: roughly ten minutes of contact per pass, a handful of usable passes per station per day. The relay layer's value is not bandwidth. It is duty cycle.

Chapter 7 showed that an orbital compute node's traffic is modest in bandwidth terms. The problem is not how fast you can move data. It is **how often you are allowed to.**

A node at 650 km passes over a given mid-latitude ground station for roughly ten minutes at a time, a handful of times a day. That is on the order of four percent of the day. A polar station sees a sun-synchronous satellite on nearly every orbit, which helps considerably. A ten-station global network gets you to something like a third. Only an **inter-satellite relay mesh** — where your traffic hops across the constellation to whichever node currently has a ground contact — gets you to continuous service.

So the relay layer's value is not bandwidth. It is **duty cycle**, and duty cycle is the difference between selling a cloud service and operating an archive. That is a twenty-fold multiplier on the usefulness of the same hardware, and it is why this layer is strategically more important than its revenue currently suggests.

The other half of it is capital nobody counts. A geographically diverse ground network — sited against cloud cover, licensed, staffed, backhauled — is a real infrastructure programme. **When someone quotes you dollars per GPU-hour in orbit, ask whether the ground segment is in the number.** It usually is not.

Who is attacking it

Kepler Communications operates a deployed optical relay constellation in low Earth orbit, running as an IP mesh, with compute and storage on the relay satellites themselves and a roadmap to substantially higher link rates. Commercial orbital data-centre nodes have already ridden and used it.[12] It is simultaneously the backhaul for orbital compute and, because its relays carry accelerators, a distributed compute platform of its own.

Skyloom and the other terminal suppliers sell the mechanism. **Starlink** operates by far the largest laser mesh in existence but does not sell open relay to third parties — and that, rather than any technical risk, is the live commercial question in this layer.

The investment stance

This is the cleanest picks-and-shovels position in the sector: a deployed network, real switching costs, revenue from customers who are not dependent on orbital compute succeeding, and a service every operator needs regardless of which operator wins.

The risk is specific and worth stating rather than glossing: **if SpaceX opens its laser mesh as a commercial service, pricing in this layer compresses immediately.** It would not eliminate the incumbents — defence traffic

will not route through a commercial mesh, and sovereign customers will pay for alternatives — but it would change the margin structure permanently. I hold the layer with that risk explicitly priced rather than assumed away.

What to watch

- Any signal that Starlink laser links become purchasable by third parties.
- Operational 100 Gbps-class optical relay, not demonstrated but in revenue service.
- Whether orbital compute operators build their own ground networks or buy contact time. That decision tells you where this layer's margin ends up.

Chapters 15 to 17 — in-space assembly and servicing, debris and insurance, and innovation outside the United States — complete Part IV in the next instalment.

Source notes, Part IV (chapters 12–14)

5. The observation that the historically favourable trend in radiation sensitivity with shrinking process nodes does not continue beyond 5 nm — Google Research, *Towards a future space-based, highly scalable AI infrastructure system design* (arXiv:2511.19468).

9, 8. (continued from Part II) Trillium TPU proton-beam results, mission dose figures, cluster architecture and reference orbit — Google Research, Project Suncatcher.

2. (continued) Company-level facts — flight history, funding, contracts, customers — from the StarCap Fund I Thesis v2 research file, 27 July 2026. **Author's own fund research; verify against primary sources before publication.**

Historical sources to cite properly before publication: von Neumann, *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components* (1956); Hamming (1950); Clarke, *Extra-Terrestrial Relays*, Wireless World (October 1945); Bell's photophone (1880). Quoted from general knowledge here; each needs a direct citation.

Beam divergence figures in Chapter 14 computed as $\theta \approx \lambda/D$ and stated in the text. The modularity scaling in Figure 13.1 follows from the square-cube law as derived in the chapter; the weighting of the composite index is the author's and should be treated as illustrative rather than as a design tool.

Figures 12.1, 13.1 and 14.1 generated for this volume.

PART IV — THE SECTOR SKELETON (concluded)

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part IV of VI — chapters 15 to 17.

Chapter 15 — In-space assembly and servicing

The old idea

In 1929 an Austro-Hungarian army officer writing under the name Hermann Noordung published *The Problem of Space Travel*, containing the first detailed engineering description of a habitable rotating space station — a structure obviously too large to launch whole, and therefore obviously something you would **assemble in orbit**. Von Braun popularised the same vision in *Collier's* in 1952. O'Neill industrialised it in 1976.

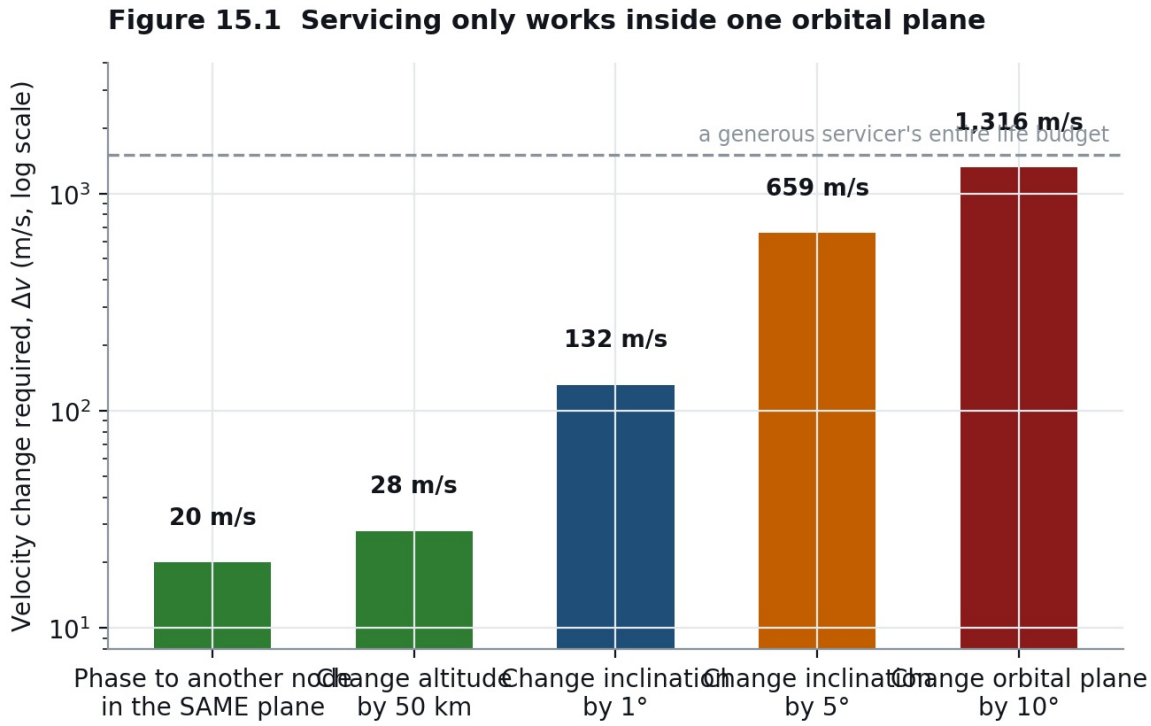
Assembly in orbit is therefore among the oldest ideas in spaceflight, and unlike most of the ideas in this book it has actually been done. Skylab was saved in 1973 by astronauts deploying a sunshade over a station that had lost its thermal shield. The Hubble Space Telescope was repaired in 1993 and serviced four more times. The International Space Station was built from dozens of separately launched modules.

So the question is not *can* we assemble and service things in orbit. **It is whether it can be done without astronauts, and at a price that a commercial operator would pay.** Every servicing success so far has been a national programme costing billions.

The physics, and the constraint that decides everything

Getting from one satellite to another is not like driving to another building. Orbits are defined by six parameters, and changing some of them is nearly free while changing others is ruinous.

Figure 15.1 — *Servicing only works inside one orbital plane*



Plane changes cost $\Delta v = 2v\sin(\Delta i/2)$ — at orbital velocity, a single degree costs more than raising your altitude by two hundred kilometres. A constellation that wants to be serviced must be designed co-planar from the first launch.

Moving to a different position *within the same orbital plane* costs almost nothing: you change altitude slightly, drift ahead or behind, and circularise. Tens of metres per second, plus patience.

Changing the **plane** costs Δv = 2v·sin(Δi/2). At an orbital velocity of 7.5 km/s, one single degree costs about **132 m/s** — more than raising your altitude by two hundred kilometres. Ten degrees costs over 1,300 m/s, which is most of a servicer's entire propellant life for one trip.

From that one relationship comes the most actionable design rule in this chapter:

> **A constellation that intends to be serviced must be co-planar from its first launch. Serviceability is not a capability you add later; it is an orbital architecture decision made before anything flies.**

There is one loophole worth knowing. Earth's equatorial bulge causes orbital planes to precess, and the rate depends on altitude and inclination. A servicer parked at a slightly different altitude drifts in plane angle for free, at perhaps a degree every few days. **Time substitutes for propellant.** That works beautifully for a patient logistics vehicle and not at all for a repair call.

Why this chapter matters more than its market size suggests

Chapter 8 ended on an unsolved problem: spacecraft last five to fifteen years, AI accelerators are economically obsolete in about three, and an orbital operator currently has no way to roll in new racks.

This is the chapter where that gets solved, or does not.

> **WORKED EXAMPLE — What a payload swap is worth**
>
> Reference Node A carries a 700 kg compute payload worth about $1.75M new. The rest of the vehicle — array, radiators, structure, propulsion, roughly 4,000 kg of it — has no reason to be obsolete after three years.
>
> If you can replace the payload alone, you launch 700 kg instead of 4,710 kg. At $600/kg that is **$420k instead of $2.8M** — a saving of about $2.4M per node per refresh cycle, before counting the revenue difference between current and three-year-old silicon.
>
> A servicer that could visit ten co-planar nodes on one mission is therefore addressing something on the order of **$24M of avoided cost per cycle** for a single modest constellation. That is a real business, and it does not exist.

The counter-argument is Chapter 13's: if modules are cheap enough, you do not service them, you replace them whole and deorbit the old ones. Which answer wins depends on the ratio between the payload's value and the rest of the vehicle's value — and on Figure 8.5's unresolved question of what the power hardware actually costs. **If arrays and radiators turn out to be cheap, disposable modules win and this chapter is a footnote. If they are expensive, servicing is the sector's most valuable unbuilt business.**

What breaks

Autonomous rendezvous and proximity operations without a cooperative target. Docking interfaces, which do not exist as an industry standard for this application. Thruster plume impingement on the delicate, enormous surfaces of Chapter 8. Fluid transfer for two-phase thermal loops, which is materially harder than propellant transfer. And the fact that a servicer approaching a $50M asset can destroy it.

Who is attacking it

Impulse Space is the mobility layer described in Chapter 9, and mobility is the prerequisite for logistics. **Outpost** is working on Earth return, which is the other half of the loop — bringing hardware down rather than only up. **Starfish Space**, **Astroscale** and the established primes' servicing vehicles have demonstrated approach and docking with uncooperative targets. Europe's robotic assembly demonstrator programme is flying its first mission in this window.[16]

The investment stance

This layer has disappointed investors for fifteen years, and I hold it lightly and specifically.

The pattern has been consistent: technically impressive demonstrations, and a customer base that turns out to prefer buying a new satellite. What is different now — and the reason to keep watching rather than dismiss it — is that the customer has changed. Previous servicing markets were about extending the life of a satellite whose *whole value* was ageing. Orbital compute is the first application where a small, valuable, rapidly-obsoleting component sits inside a large, slowly-obsoleting vehicle. That is precisely the structure that makes servicing economic.

So: not a bet on servicing as a category. A bet on mobility, with servicing as the option attached.

What to watch

- **The first commercial payload swap on orbit**, in any application. It would reprice this layer immediately.
- An orbital compute operator publishing a docking interface standard — which would tell you it has decided serviceability is worth designing for.
- Whether the compute constellations now being filed are co-planar. Figure 15.1 means that single architectural choice determines whether servicing is even possible for them.

Chapter 16 — Debris, traffic and insurance

The old idea

In 1978 a NASA scientist named **Donald Kessler** published a paper describing what happens when the density of objects in orbit passes a threshold: collisions generate fragments, fragments cause further collisions, and the process becomes self-sustaining. The Kessler syndrome is now the single most cited constraint on the long-term use of low Earth orbit, and it was described before the first Shuttle flight.

The physics

Orbital velocity is about 7.5 km/s, and two objects can meet at up to twice that. Kinetic energy goes as the square of velocity, which produces numbers that do not match intuition at all.

> **WORKED EXAMPLE — What a fleck of paint does at orbital speed**
>
> A **1 millimetre** aluminium sphere weighs about 1.4 milligrams. At a 10 km/s closing speed it carries about **70 joules** — enough to punch through thin panel and open a fluid line.
>
> A **1 centimetre** fragment weighs about 1.4 grams and carries about **70 kilojoules** — several times the muzzle energy of a heavy rifle round, delivered to a spacecraft with no armour.
>
> A **10 centimetre** object carries about **70 megajoules**. Nothing survives that, and the collision itself creates thousands of new fragments.

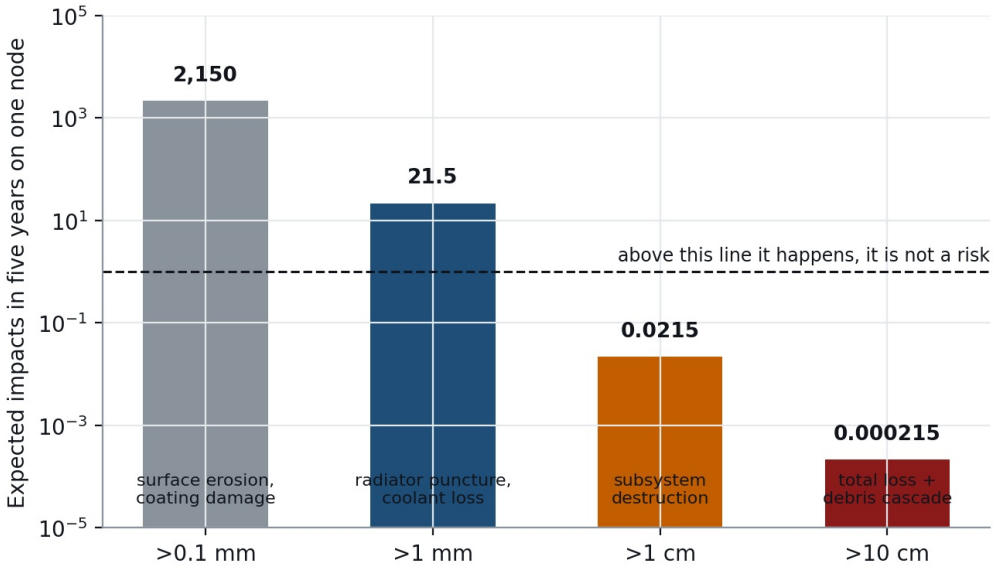
Why compute nodes are different, and worse

Here is the point this chapter exists to make, and I have not seen it made elsewhere.

An orbital data centre is, per tonne, the largest target ever flown. Chapter 8’s reference node carries roughly 290 m² of solar array and 140 m² of radiator — over 400 square metres of cross-section on a 4.7 tonne vehicle. A conventional satellite of similar mass presents a few square metres. **The collision cross-section is up to a hundred times higher for the same mass in orbit.**

Figure 16.1 — *A compute node is the largest target ever flown*

Figure 16.1 A compute node is the largest target ever flown



430 m^2 of array and radiator over five years, against order-of-magnitude debris and meteoroid flux for this altitude band. A 1 cm aluminium fragment at 10 km/s carries about 70 kJ — several times a heavy rifle round. Flux figures need modelling per orbit.

Run the flux against that area over five years and the picture separates cleanly into three regimes:

- **Sub-millimetre impacts are constant.** Thousands of them. They erode coatings, which attacks emissivity, which sits inside Chapter 5's fourth-power law. This is a slow degradation of thermal capacity, not an event.
- **Millimetre-class impacts are certain** — on the order of twenty over five years for one node. Most hit array or structure and do little. Some hit a radiator, and a radiator is a thin, fluid-filled surface. **This is why loop isolation and panel redundancy are requirements rather than refinements**, and it is the specific reason the ISS radiators were designed with debris protection built in.
- **Centimetre-class impacts are a low-probability, total-loss event** — a couple of percent per node over five years on these numbers. For a single satellite that is an acceptable risk. For a constellation of a thousand nodes it means you should expect to lose roughly twenty of them to debris, and you must have designed for it.

The constraint nobody is modelling: insurance capacity

The space insurance market is small. Global annual premiums are on the order of half a billion dollars — **less than the cost of a single large terrestrial data centre campus** — and the market has had loss-making years recently.[17]

Now consider what this sector proposes to insure. A gigawatt-class orbital constellation is tens of billions of dollars of hardware, in an environment where the loss modelling is immature, the failure modes are correlated across the fleet, and the largest single asset class has no actuarial history whatsoever.

The binding constraint on scaling this sector may turn out to be underwriting capacity rather than physics. Debt financing requires insurance. Insurance requires a loss model. A loss model requires history. The sector has none, which means early constellations will either be equity-financed at high cost of capital, self-insured by balance sheets large enough to absorb the loss, or state-backed. All three of those favour very large incumbents over startups, and that is a competitive dynamic worth understanding before it becomes obvious.

Who is attacking it

Space situational awareness providers — **LeoLabs**, **Slingshot** and the government catalogues — sell the tracking that makes conjunction avoidance possible. Specialist brokers and underwriters price the risk. Debris removal companies exist and remain in search of a customer willing to pay.

The investment stance

I do not hold this layer, and I want to be clear that this is a sizing judgement rather than a quality judgement. Tracking and insurance are real businesses; they are simply too small, at current scale, to move a concentrated

fund, and debris removal has no proven buyer.

But this chapter belongs in a guidebook because **it is a cost line and a gating risk in every other chapter.** Every operator's model needs a debris allowance, a conjunction-avoidance propellant budget, and an insurance premium — and most published models contain none of the three.

What to watch

- **The first insurance policy written on an orbital data centre, and its rate.** That number is the market's honest estimate of the risk, and it is more informative than any technical paper.
- Any conjunction event involving a very large-area spacecraft. The first one will reprice the entire sector's risk assumptions in a week.
- Whether regulators begin treating cross-sectional area, rather than satellite count, as the quantity to limit. That would change constellation design more than any other plausible rule.

Chapter 17 — Outside the United States

Nearly everything in this book so far has been American. That is a fair reflection of where the private capital and the launch capacity sit, and it is also a distortion, because the country with the most **flown** orbital compute hardware is not the United States.

China

China treats orbital computing as infrastructure policy rather than as a venture opportunity, and it started earlier.

In May 2025 a Long March 2D placed twelve satellites in orbit as the first batch of the **Three-Body Computing Constellation**, led by Zhejiang Lab with the Chengdu company ADA Space — described as the world's first dedicated orbital computing constellation, with a combined capability of 5 peta-operations per second and 30 terabytes of onboard storage across the twelve.[18] The stated plan reaches 100 satellites by 2027, within a wider "Star-Compute" programme targeting 2,800 satellites and a combined 1,000 POPS.[19]

It is not the only one. A 300-satellite Tiansuan constellation runs out of Beijing University of Posts and Telecommunications; a further roughly 1,000-satellite network, Xingshu, was announced in 2026 by Fudan University and a Shanghai partner; Beijing established a dedicated space-computing innovation centre in mid-2026 whose priority areas include heat-tolerant, radiation-tolerant chips designed specifically for orbit.[20]

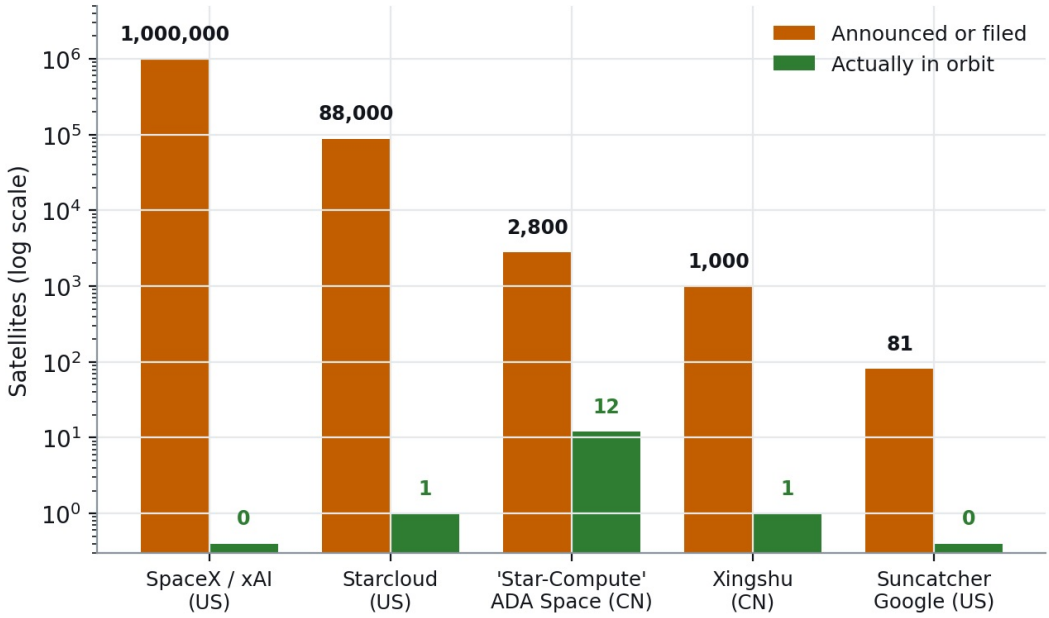
Two things in the Chinese approach are worth taking seriously on the engineering merits.

The first is the **stated rationale**, which is not energy at all. Chinese researchers frame the problem as downlink: that a very large share of the data generated in orbit is never transmitted or processed because bandwidth to the ground is insufficient.[18] That is precisely Chapter 7's strongest near-term argument — put the compute next to the sensor — and it is notable that the programme with the most hardware in orbit is the one pursuing the workload the physics most clearly supports.

The second is that Chinese teams are converging on the **same COTS-plus-software answer** as everyone else: commercial chips rather than slow rad-hard parts, with software-defined fault tolerance and ground-based digital twins doing the reliability work.[20] When independent programmes under different constraints arrive at the same architecture, that architecture is probably right.

Figure 17.1 — *Filings are cheap. Hardware is not.*

Figure 17.1 Filings are cheap. Hardware is not.



Dedicated compute satellites only. Zero is drawn at the axis floor for visibility. The country with the most announced capacity and the country with the most flown hardware are not currently the same country.

I hold no Chinese exposure, by mandate, and I would note that this is a decision about jurisdiction and not about capability. The relevant question for a Western investor is not whether to buy these programmes but **what they do to price**. A state-backed constellation that does not need to earn a commercial return can set the price of orbital inference in the markets it serves, in the same way state-backed manufacturing has repeatedly done in terrestrial industries.

Europe

Europe's contribution has been a serious feasibility study rather than hardware. **ASCEND** — coordinated by Thales Alenia Space with a consortium including ArianeGroup, Airbus, DLR, HPE and Orange — was funded by the European Commission under Horizon Europe and concluded in 2024 that space-based data centres are technically, economically and environmentally feasible, framed around EU net-zero goals and data sovereignty. Its stated ambition is one gigawatt deployed before 2050, against a projected 23 GW terrestrial data centre market by 2030, with modular infrastructure assembled in orbit by robotic systems.[21]

Read that carefully and Europe's position becomes clear. **The target date is 2050.** The American programmes are talking about 2027. Europe has produced the most rigorous public feasibility analysis in the field and paired it with a timeline a generation behind the people building hardware.

That is not merely a criticism. Europe's realistic roles are as a **customer** — sovereign, regulated, willing to pay for data jurisdiction — as a **supplier** of optical terminals, robotics and thermal engineering, and as a **regulator** whose rules on data location will shape where orbital capacity can legally serve. All three matter. None of them is a constellation.

Russia, and everyone else

Russia is not a factor in this sector. Its launch industry has contracted sharply and there is no visible orbital compute programme.

The Gulf states are more interesting than their space heritage suggests, because they combine three things this sector needs: sovereign capital with a long horizon, cheap energy and a strategic interest in owning digital infrastructure rather than renting it. Gulf sovereign funds already appear on the cap tables of Western station and compute companies, and regional manufacturing joint ventures are being built. Japan and India appear as component suppliers and investors rather than as programme leaders.

The stance, and the wrinkle worth thinking about

The sector will not remain a single market. Data jurisdiction is the reason: a data centre in orbit is under the jurisdiction of the state that licensed the satellite, not of the territory it flies over, which makes orbital compute an unusually clean instrument of digital sovereignty. That argues for **several regional constellations rather than one global cloud**, which is good for suppliers who sell to everyone and harder for operators hoping for winner-take-all economics.

And a genuine legal oddity that I have not seen resolved anywhere: advanced AI accelerators are subject to export controls in most jurisdictions. **What happens when the accelerator is launched into orbit and then sold, leased, or served to a customer in a restricted country?** Where the chip physically is, who controls it, and which regime applies are not settled questions. They will be, and the answer will shape this industry more than several of the engineering problems in this book.

What to watch

- Whether the Chinese constellation reaches its stated 2027 scale, and whether it publishes performance.
- Any European decision to fund ASCEND beyond study phase, which would pull its timeline forward by a decade.
- The first export-control ruling on accelerators in orbit.

Part IV is complete. Part V builds the full economic model — cost per GPU-hour, satellites per gigawatt, and what Earth gets back — and Part VI sets the horizon and the falsifiers.

Source notes, Part IV (chapters 15–17)

6. European Robotic Orbital Support Services in-orbit demonstrator, led by Thales Alenia Space, first mission scheduled in the 2026 window — reported alongside the ASCEND study results.
7. Space insurance market scale is stated here as an order-of-magnitude figure from general industry reporting. **This needs a properly sourced 2026 premium and capacity figure before publication; the argument depends on the comparison being right.**
8. First batch of twelve satellites of the Three-Body Computing Constellation launched 14 May 2025 on a Long March 2D from Jiuquan, led by ADA Space and Zhejiang Lab; combined 5 POPS and 30 TB onboard storage; described as the first dedicated orbital computing constellation — SpaceNews, 14 May 2025. Downlink-loss rationale and the 100-satellite-by-2027 plan — Xinhua, April 2026.
9. The wider Star-Compute programme targeting 2,800 satellites and 1,000 POPS — SpaceNews and Xinhua, as above.
0. Tiansuan (BUPT), the Xingshu network announced by Fudan University and a Shanghai partner in 2026, the Beijing space-computing innovation centre, and the software-defined fault-tolerance approach with ground digital twins — Chinese technology press, 2026. **Verify each independently; reporting on Chinese programmes varies in reliability and the constellation numbers quoted are announcements, not deployments.**
1. ASCEND (Advanced Space Cloud for European Net zero emission and Data sovereignty), coordinated by Thales Alenia Space under Horizon Europe, results announced 27 June 2024; consortium, feasibility conclusion, 1 GW before 2050 ambition and the 23 GW 2030 market estimate — Thales Alenia Space press release and contemporaneous reporting.

Historical sources to cite properly before publication: Noordung, *Das Problem der Befahrung des Weltraums* (1929); von Braun in *Collier's* (1952); Kessler and Cour-Palais, *Collision Frequency of Artificial Satellites: The Creation of a Debris Belt*, JGR (1978).

Debris flux figures in Figure 16.1 are order-of-magnitude values for this altitude band and **must be replaced with per-orbit ORDEM or MASTER model output before publication.** Kinetic energies computed from aluminium spheres at 10 km/s closing velocity. Plane-change costs computed as $2v \cdot \sin(\Delta i/2)$ at $v = 7.55$ km/s.

Figures 15.1, 16.1 and 17.1 generated for this volume.

PART V — DOES IT PENCIL

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part V of VI.

Everything so far has been physics and engineering. This part turns it into money, and it does so in the open: every input is stated, every step is shown, and the model is simple enough that you can rebuild it in a spreadsheet in an afternoon and disagree with me precisely.

That is deliberate. The most common failure in this sector's economic writing is a model whose conclusion is determined by an assumption buried three layers down. The second most common is a model that stress-tests its costs and holds its revenue constant.

This part avoids both, and in doing so it does something the sector rarely does: it separates **today's number from tomorrow's curve.** Those are different questions, and conflating them is why the debate about orbital compute is so bad. Today's number is a snapshot of a first-of-a-kind industry building single-digit units at venture cost of capital, and it is unflattering — as the first number always is. The curve is what happens when a manufactured product goes down a learning rate, and it is the only thing that has ever mattered in an infrastructure industry.

Both are computed here. The snapshot first, because a projection you cannot anchor is a fantasy. Then the curve, because a snapshot you cannot project is a photograph of a rocket on the pad.

Chapter 18 — The cost model, and who would buy the output

First: who buys this?

Cost models are the easy part. The question that decides this sector is who signs the contract, and it is almost never asked in print.

Start by being clear about who does *not*. **A hyperscaler will not put production inference on a satellite** in the foreseeable future. Enterprise procurement requires an availability commitment, a support path, a second source and a settled legal jurisdiction; Chapter 8 established that an orbital node cannot be physically touched after launch, and Chapter 17 established that the jurisdiction question is genuinely unresolved. The named cloud "booked workloads" reported by operators today should be read as what they almost certainly are — funded pilots, valuable as validation, not evidence of a procurement channel.

So who is left? Three buyers, and they share a property that turns out to matter more than anything in the cost model: **all three pay a premium, and none of them are buying commodity compute.**

1. Data that is already in orbit. Earth-observation operators are downlink-limited, not compute-limited. Chinese researchers building orbital compute frame their entire rationale this way, and Chapter 7's arithmetic supports it: raw sensor data is enormous, conclusions are tiny. For this buyer the alternative is not a cheaper terrestrial GPU — it is *discarding the data*. That is not a price-sensitive comparison, and it is available today at kilowatt scale rather than megawatt scale.

2. Defence and intelligence. Latency to the sensor, resilience against attacks on terrestrial infrastructure, and sovereign control of the processing chain. Government customers buy attributes rather than throughput, and have historically paid multiples of commercial rates for both. It is not an accident that the private companies in this sector with real revenue have defence-weighted revenue.

3. Sovereign data jurisdiction. A satellite falls under the jurisdiction of the state that licensed it, not of the territory it flies over. For a state wanting compute outside another power's legal reach, that is a feature with no terrestrial substitute, and Europe's public programme is explicitly framed around it.

The consequence reorganises the rest of this chapter:

> **Orbital compute does not need to beat the commodity price of a GPU-hour. It needs to beat it for buyers who are not buying a commodity. Any model that benchmarks orbital cost against spot cloud rental is measuring against the wrong customer — and it is the harshest possible test, which is why this chapter uses it anyway.**

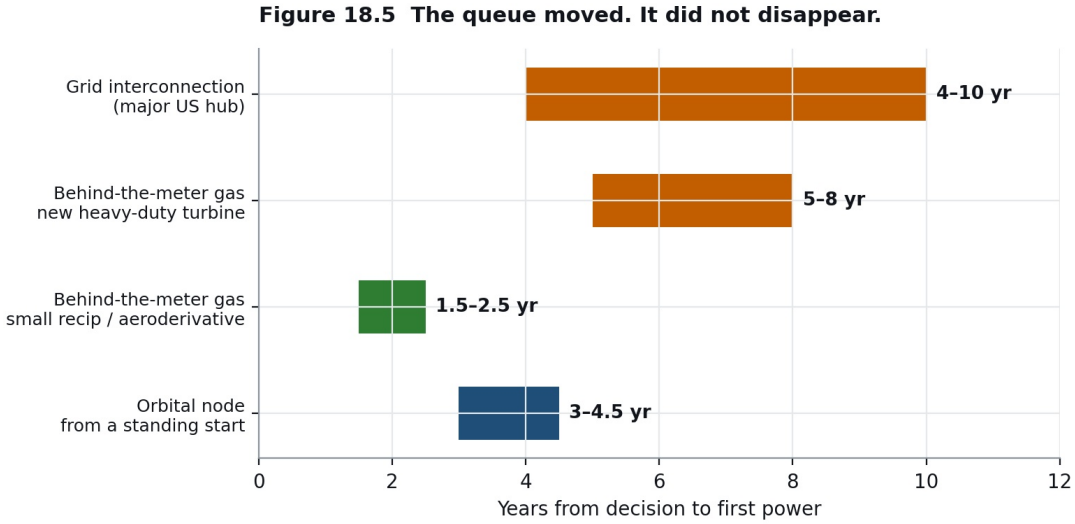
Second: what is the real competitor?

Part I framed the alternative as “wait four to ten years for a grid connection.” That is the wrong counterfactual, and it is the strongest attack on this book’s argument.

The real alternative is **behind-the-meter generation**: put a gas turbine on site and skip the utility entirely. It is already the dominant workaround, and if it is fast, the availability argument collapses and nothing else here saves it.

So how fast is it?

Figure 18.5 — *The queue moved. It did not disappear.*



Heavy-duty gas turbine order books at the three global manufacturers now stretch to 2029–2030, with quoted lead times of five years and guidance to plan for seven to eight. Small units are fast but cap out well below hyperscale.

Three companies in the world build heavy-duty gas turbines at scale. All three are sold out on a horizon no other industrial market can absorb: order backlogs stretching into 2029 and 2030, lead times that have gone from two or three years in 2022 to roughly five today, with manufacturers now advising customers to plan on seven-to-eight-year timelines, and steep price inflation on delivery slots.[22]

That is the answer, and it is more interesting than a simple defence. **The bottleneck did not vanish when the industry routed around the utility. It relocated from the interconnection queue to the turbine factory** — which is the same structural story this book opened with: compute is limited by industrial capacity, not by physics and not by money.

Two honest qualifications. Smaller reciprocating and aeroderivative units have much shorter lead times, on the order of eighteen months to two years, and are genuinely fast — but they are less efficient, dirtier, permit-constrained on air quality in exactly the metros where demand concentrates, and they do not reach hyperscale. And an operator who ordered turbine slots two years ago has power today while an orbital operator does not. **Behind-the-meter gas is a real and partial escape at tens of megawatts, and it meets the same wall at hundreds.**

It also sets a cost benchmark that is not soft: a behind-the-meter site costs roughly the same per megawatt-year as a grid-connected one, because cheaper fuel offsets the generation capital.

The inputs, all of them, on one page

Input	Value used	Source
Reference vehicle	1 MW compute, 37 tonnes	Ch 8
Compute hardware	6.8 t, $17.5M	Ch 8 convention
Spacecraft mass, everything else	30.2 t	Ch 8, Table 8.1
Spacecraft build cost	**$2,000/kg base**; $1,000 stretch; $20,000 bespoke	the open question
Launch price	$150/kg base; $600/kg bespoke case	Ch 2
Compute hardware life	5 years base (3–7 tested)	Ch 6
Spacecraft service life	7 years base (5–10 tested)	Ch 8
Cost of capital	**12% base (8–20% tested)**	new — see below
Availability	**92%**	new — see below
Fleet attrition	**2% of hull per year**	Ch 16
Performance fade	**1.5% per year**	Ch 16
Operations	$2M per MW-year, including ground segment	estimate
Insurance	3% of hull per year	Ch 16
Revenue	$9M / $18M / $30M per MW-year	three worlds

Four lines are new relative to how this sector usually models itself, and each exists because leaving it out flatters the answer:

Cost of capital. Earlier drafts used straight-line amortisation and called themselves “a cash cost model, not a valuation.” That was a dodge. A long-dated, capital-intensive, high-risk asset must carry a capital charge, and the honest instrument is a capital recovery factor — the annual payment that returns capital at a required rate over the asset’s life. At 12% over seven years that is 22% of capital per year, against the 14% straight-line implies. **This single correction adds more cost than launch does.**

Availability. Terrestrial data centres sell 99.9%. An orbital node cannot be touched: one jammed deployment, one unrecovered latch-up, one radiator puncture and it is a total loss. 92% is a judgement, unproven in either direction, and it appears in the tornado so you can move it.

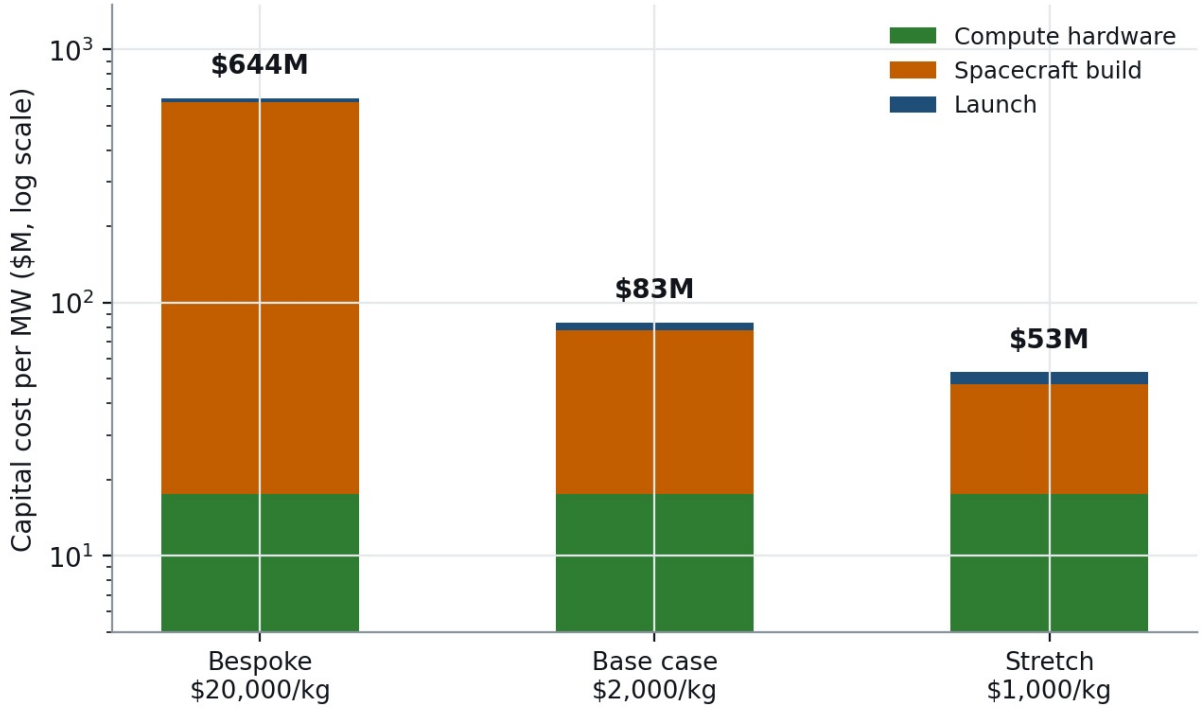
Fleet attrition and performance fade. Chapter 16 established that millimetre impacts are certain and that coating erosion attacks emissivity, which sits inside a fourth-power law. A model assuming constant thermal performance for seven years contradicts its own Chapter 16. Attrition replaces 2% of hull value annually; fade derates output 1.5% a year.

Availability and fade together give an **effective output factor of 0.879** — a node sells about seven eighths of its nameplate GPU-hours. Every unit cost below is per *effective* GPU-hour.

The three worlds

Figure 18.1 — *Launch is the smallest bar on this chart*

Figure 18.1 Launch is the smallest bar on this chart



The only thing that changes materially between the columns is what it costs to manufacture the spacecraft.

	Compute	Spacecraft build	Launch	Capex/MW	Cost per effective GPU-hour
Bespoke ($20,000/kg)	$17.5M	$604M	$22.2M	$644M	**$38.62**
Base ($2,000/kg)	$17.5M	$60.4M	$5.5M	$83M	**$5.56**
Stretch ($1,000/kg)	$17.5M	$30.2M	$5.5M	$53M	**$3.72**
Terrestrial, same capital treatment					*$1.86*
Terrestrial with behind-the-meter gas					*$1.92*
Prevailing commodity market price					*~$3.50*

Read that slowly, because it is the honest snapshot and it is not yet a business.

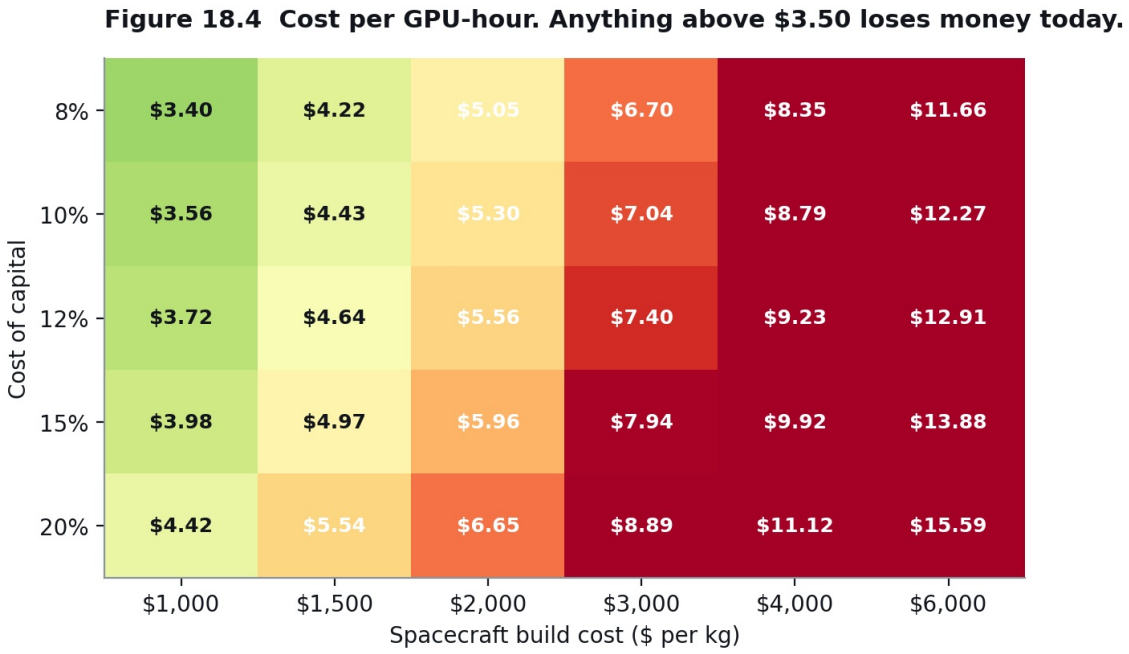
At today's manufacturing costs and today's cost of capital, orbital compute does not clear the commodity price. The base case loses money. The stretch case, which assumes manufacturing economics nobody has yet demonstrated for a 37-tonne deployable vehicle, lands at $3.72 against a $3.50 market — close, and not there.

Now hold that beside two other snapshots. Utility-scale solar in 2010 cost roughly four times the wholesale power price it was competing with. Lithium cells in 2010 cost about seven times what the automotive industry said it needed. Falcon 9 in 2012 was more expensive per kilogram than the incumbent it was built to undercut. **In each case the snapshot was correct and the conclusion drawn from it was wrong**, because the snapshot measured a first-of-a-kind unit and the thing that mattered was the slope.

The right question is therefore not *does this clear today* — it plainly does not — but **at what point on the manufacturing curve does it clear, and can the industry get there.** That is a calculable question, and the next

section calculates it.

Figure 18.4 — *Cost per GPU-hour across manufacturing cost and cost of capital*



Green cells clear the current market price. The thesis needs cheap manufacturing AND cheap capital — neither alone is enough, and the sector currently has neither.

The matrix is the honest summary of this book's economics. Green cells clear today's commodity price; there are few of them, and they all sit in the bottom-left corner where a company manufactures at a thousand dollars a kilogram and borrows like a utility. **The thesis needs cheap manufacturing and cheap capital simultaneously, and the sector currently has neither.**

Which brings demand back with force, and explains the sequence the industry will actually follow. **The premium buyers come first because they clear earlier.** Against the stretch case at $3.72, a buyer paying 30% over commodity for sovereignty, latency to the sensor or resilience makes the business work today. The commodity market clears later, when the curve has done its work.

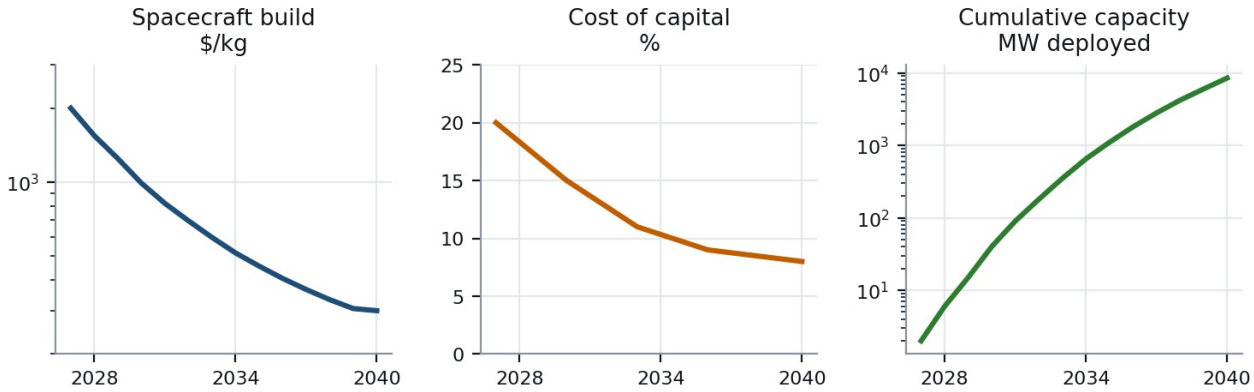
> **The verdict this arithmetic supports: orbital compute is a premium-buyer business now and a commodity-compute business later, and the interesting question is the size of "later." A pitch promising cheaper AI in orbit this year is wrong. A pitch promising it by the early 2030s is arguing about a learning rate, which is a real argument with real evidence on both sides.**

When does the math work?

Three things move together over the next decade, and none of them requires a breakthrough. Each is an extrapolation of a process already running.

Figure 18.7 — *The three curves underneath the crossover*

Figure 18.7 The three curves underneath the crossover



Manufacturing learning drives the first, maturity of the asset class drives the second, and the second depends on the third. Each is a projection with a stated assumption, not a forecast — change any of them and Figure 18.6 moves.

One: manufacturing goes down a learning curve. Wright's law — unit cost falls by a fixed percentage for every doubling of cumulative production — is among the most durable empirical regularities in industrial economics, and it holds across aircraft, semiconductors, solar modules, batteries and launch vehicles. Observed learning rates cluster between 10% and 20% per doubling. The relevant question for this sector is not whether it applies but how many doublings it gets, and that depends on deployed capacity.

Take a deployment path that reaches a gigawatt in the mid-2030s — aggressive, but slower than the sector's own filings. From two megawatts of cumulative capacity to a thousand is about nine doublings. **At a 15% learning rate, that takes spacecraft manufacturing from $2,000/kg to roughly $450/kg.** Not by inventing anything: by building the ninth hundred of something rather than the first ten.

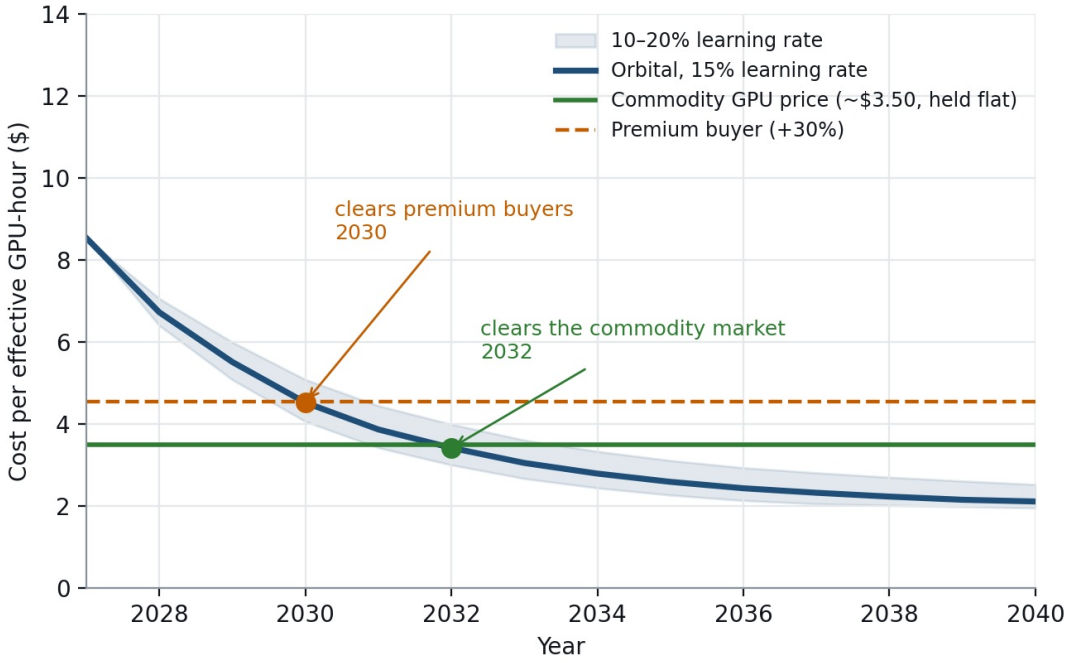
Two: the cost of capital falls as the asset class matures. This is the driver nobody models and it is worth as much as the manufacturing curve. Today an orbital compute constellation is financed at venture rates because it is unproven, uninsured and unrepeated. That is exactly what solar farms and wind projects looked like in the 2000s, and both are now project-financed at single-digit rates against contracted offtake. The path is not mysterious: flight heritage, then contracted revenue, then insurance, then debt. **Going from 20% to 9% cost of capital cuts the annual cost of the same hardware by roughly a third.**

Three: availability improves with flight heritage, as it has for every spacecraft class ever fielded, and every point of availability is revenue against a fixed cost base.

Put the three together against a flat revenue assumption — no help from rising GPU prices, no help from improving performance per watt, both of which would accelerate it:

Figure 18.6 — *When the math starts working*

Figure 18.6 When the math starts working



Wright's law applied to spacecraft manufacturing against the deployment curve in Figure 18.7, with cost of capital falling from venture rates to infrastructure rates and availability improving with flight heritage. Revenue is held flat, which is conservative.

Year	Build cost	Cost of capital	Cost per GPU-hour
2027	$2,000/kg	20%	$8.53
2030	~$990/kg	15%	$4.52
2032	~$700/kg	~12%	**$3.41**
2035	~$460/kg	10%	**$2.59**
2040	~$300/kg	8%	**$2.11**

Orbital compute clears the premium market around 2030 and the commodity market around 2032, on a 15% learning rate. At 20% it happens a year earlier; at a pessimistic 10% it slips to 2034. The band in Figure 18.6 is that whole range, and the honest headline is: **somewhere between 2031 and 2034, on assumptions that require nothing new to be invented.**

Three things worth saying about that result.

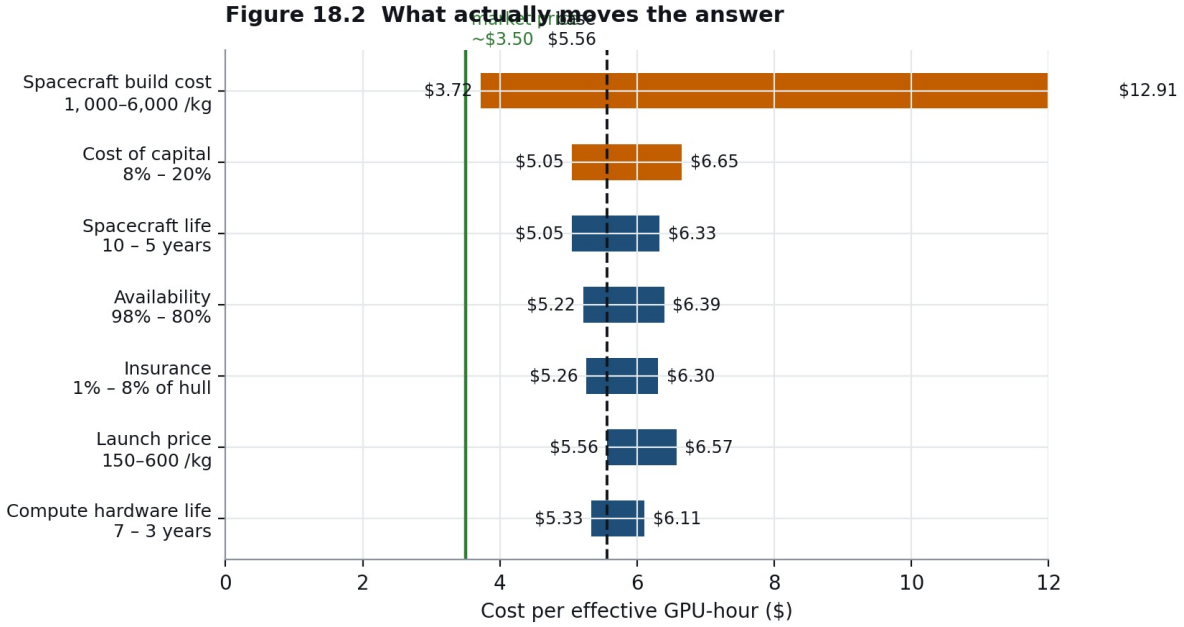
It is **conservative on revenue.** Holding the price of a GPU-hour flat for fourteen years while performance per watt improves every generation is almost certainly wrong in the sector's favour: orbital costs scale with watts — radiators, arrays, mass — while revenue scales with useful output. Every doubling of tokens per watt is a doubling of revenue against a fixed thermal and structural cost base. That single effect could pull the crossover in by years, and it is deliberately excluded here.

It is **self-referential in a way you should watch.** The manufacturing curve needs deployment, and deployment needs the manufacturing curve. Industries do escape that loop — subsidised early demand, government anchor customers, a vertically integrated player who eats the early losses — but they do not escape it automatically. **This is exactly what the premium buyers are for.** Defence, sovereign and sensor-adjacent customers pay enough to fund the first nine doublings. They are not a consolation prize for a business that cannot reach the commodity market; they are the mechanism by which it gets there.

And it explains **why the investable moment is now rather than in 2032.** By the time the curve crosses, the companies that rode it down are priced accordingly. The window for an investor is the period when the physics is settled, the engineering is legible, and the arithmetic has not yet caught up — which is a fair description of the present.

What actually moves the answer

Figure 18.2 — *What actually moves the answer*



One variable at a time from the base case. Everything left of the green line makes money at today's prices; everything right of it does not. Launch price is the second least important input in the model.

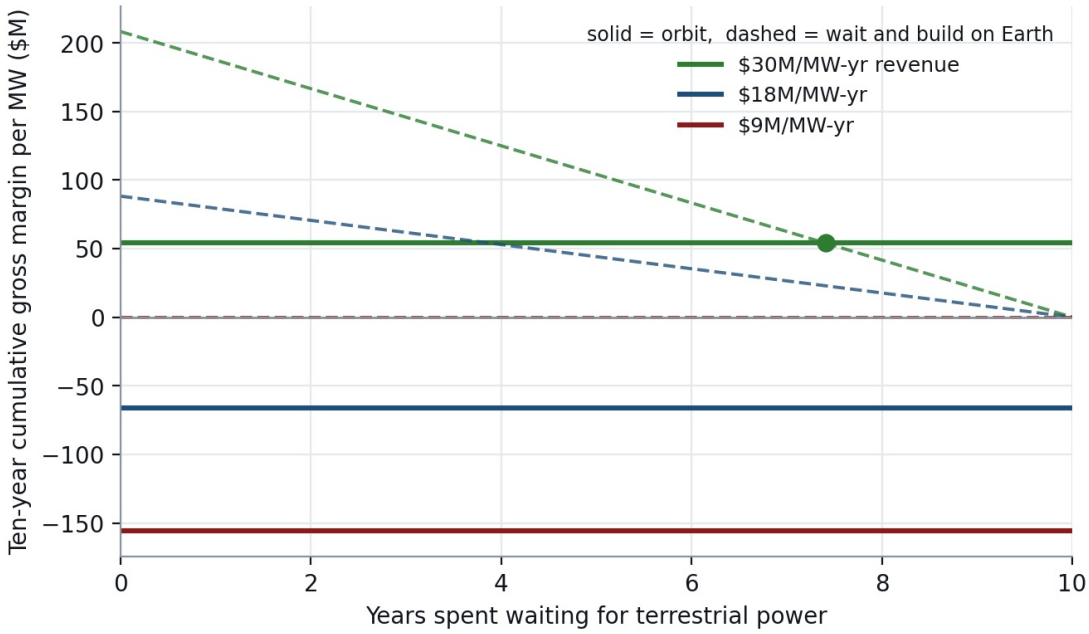
One variable at a time from the base case:

1. **Spacecraft build cost** — dominates everything.
2. **Cost of capital** — second, and absent from every published model in this sector I have read. A company that finances like infrastructure has a structural advantage over one financing like a startup, independent of its engineering.
3. **Spacecraft service life** — a ten-year vehicle nearly halves the capital charge.
4. **Compute hardware life** — Chapter 6’s depreciation argument and Chapter 15’s servicing argument arriving together.
5. **Availability** — larger than launch price, and completely unmeasured.
6. **Launch price** — sixth. The number this sector discusses most is the second least important input in its own economics.
7. **Insurance** — smallest today, most likely to move if Chapter 16’s capacity problem bites.

The queue argument, priced across three revenue worlds

Figure 18.3 — *The argument, honestly, across three revenue worlds*

Figure 18.3 The argument, honestly, across three revenue worlds



At base-case costs the orbital line only clears the terrestrial one in the high-revenue world, and then only past a long wait. This is a much narrower result than the sector's usual claim, and it is the one the arithmetic supports.

Earlier drafts held revenue fixed at $18M per MW-year while stress-testing seven cost inputs. That was the second dodge, and it mattered: $18M derives from spot GPU rental rates set during a capital-expenditure boom, and those rates are likelier to fall than the cost inputs are.

Three worlds, then. At **$9M** — rates halving as supply catches up — orbital compute is unviable at any manufacturing cost in this model, and so is a good deal of the terrestrial industry. At **$18M**, base-case orbital loses money and only the stretch case survives, with a long break-even. At **$30M** — premium buyers, sovereign contracts, defence rates — orbital clears terrestrial past about **seven and a half years** of waiting.

A narrower window than Part I implied — **at today's costs.** Run the same three worlds against the 2032 cost line from Figure 18.6 and the picture inverts: the $18M world clears comfortably, and the $9M world becomes the marginal case rather than the impossible one. The break-even queue length falls with the cost curve, which is the point of the cost curve.

It also explains the shape of the actual sector today: the companies with revenue sell to governments, not to clouds. That is not a limitation of the thesis. It is the first rung.

How long does orbit itself take?

The availability argument claims capacity is needed *now*. Orbit should be held to the same clock.

> **WORKED EXAMPLE — Time to first orbital megawatt**
>
> Design and qualification of a new megawatt-class vehicle: 24–36 months, assuming the thermal architecture works first time. Manufacturing and integration: 6–12 months. Launch slot and commissioning: 3–6 months.
>
> **Three to four and a half years from decision to first operating megawatt** — for an operator that already has a flown predecessor. For a new entrant, longer.

Set against Figure 18.5 that is sober but not bad. Orbit already beats a grid connection in a congested hub and beats a new heavy-duty turbine slot. It is slower than a small reciprocating unit, and it is not the instant capacity the marketing implies.

The direction of travel matters more than the current position. Every one of those timelines is fixed by physical infrastructure except the orbital one, which is fixed by a production line that does not exist yet. **Grid queues lengthen as demand grows; turbine backlogs lengthen as demand grows; spacecraft build times shorten as production scales.** Time-to-power is the one competitive dimension where orbit's trend line points the right way

while everyone else's points the wrong way.

How big is a gigawatt, physically?

> **WORKED EXAMPLE — One gigawatt of orbital compute**
>
> - **1,000 reference nodes** of 1 MW each.
> - **37,000 tonnes** to orbit — roughly **370 launches** at 100 tonnes of usable payload, close to four years of a hundred-flight-a-year fleet doing nothing else.
> - **2.9 km² of solar array** and **1.4 km² of radiator**: about three square kilometres of thin deployed hardware.
> - Roughly **5,000 racks** of accelerators.
> - At stretch-case costs, about **$53 billion of capital** — and at 12%, an annual capital charge near **$13 billion**.
>
> None of it is impossible. All of it is far larger than the word "constellation" implies.

Chapter 19 — What Earth gets back

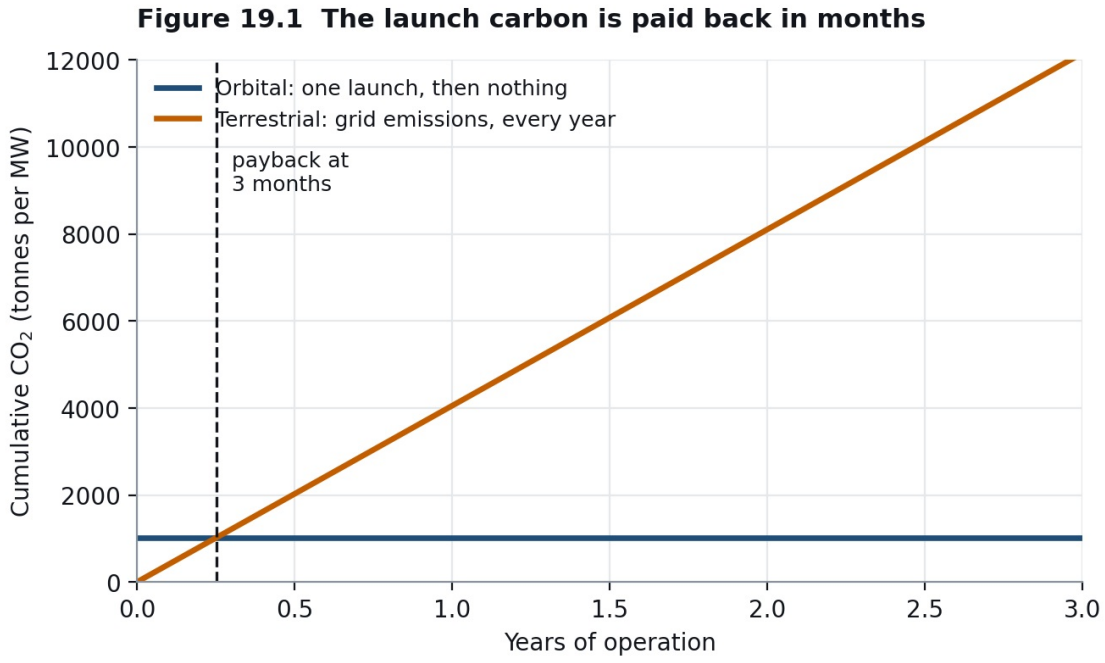
There is a second case, made loudly by the European study and quietly by everyone else, that orbital compute is better for the planet. This chapter prices it: encouraging in ratio, small in absolute terms.

Energy and carbon

A terrestrial megawatt of compute draws about 1.25 MW from the grid once cooling and conversion are included — roughly 11 GWh a year. At a US average grid intensity near 370 gCO_2/kWh, that is about **4,050 tonnes of CO_2 per megawatt-year**. An orbital megawatt draws none of it.

Against that sits the launch. A methane-fuelled heavy vehicle burns roughly a thousand tonnes of fuel, producing on the order of 2,750 tonnes of CO_2. A 37-tonne node claiming its share of a hundred-tonne launch carries about **1,020 tonnes** of embodied launch emission.

Figure 19.1 — *The launch carbon is paid back in months*



Grid intensity 370 gCO_2/kWh, PUE 1.25, methane combustion at 2.75 t CO_2 per tonne of fuel, 37 t node as a share of a 100 t launch. Excludes stratospheric soot and water-vapour effects, which are poorly quantified and not favourable.

Payback is **about three months**. After that, each year of operation is roughly four thousand tonnes that did not happen.

Two caveats belong in the same breath. This counts only combustion CO_2; rockets also inject soot and water vapour directly into the stratosphere, where radiative effects are poorly quantified and not favourable — at hundreds of launches per gigawatt that is a research gap, not a footnote. And against a data centre powered by dedicated new renewables, the advantage narrows substantially.

Water and land

A terrestrial megawatt cooled evaporatively can consume on the order of **twenty million litres of water a year**, though modern closed-loop facilities have cut this by an order of magnitude. An orbital megawatt consumes none. Powering 1.25 MW continuously from terrestrial solar needs several hectares before the data centre is built; orbit uses none — but it consumes orbital volume and cross-sectional area, which Chapter 16 established is scarce and unpriced. The environmental case tends to skip that.

The dividends that are not about energy

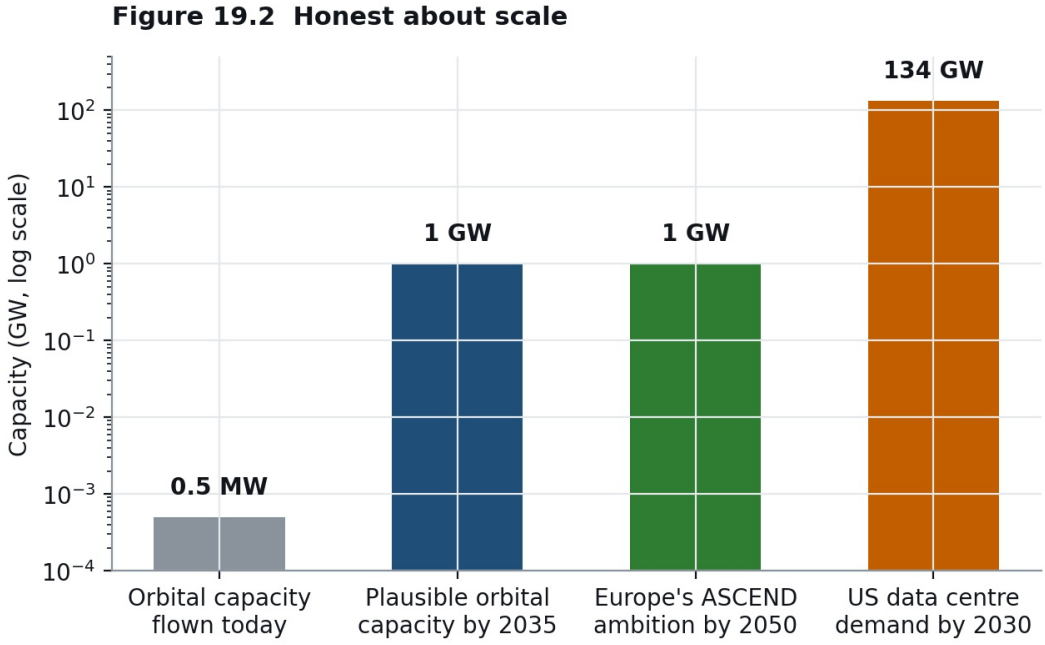
The sensor dividend is the most certain public benefit in this book and needs no gigawatt to materialise: compute next to the sensor turns Earth observation from a system that photographs and forgets into one that watches and reports.

Resilience cuts both ways: compute detached from a national grid survives grid failure, and an orbital fleet is exposed to a severe solar storm in a correlated way a distributed terrestrial fleet is not.

Spillover may be the largest effect of all. Megawatt-scale two-phase cooling, high-voltage distribution, radiation-tolerant commercial silicon and fully autonomous operation are all things terrestrial data centres will want as rack densities pass six hundred kilowatts. Chapter 1's slope does not stop.

Now the honest part

Figure 19.2 — *Honest about scale*



Even a successful decade puts orbital compute at well under one percent of demand. This is a relief valve for capacity that cannot be built in time — not a climate solution, and any book claiming otherwise is selling something.

US data centre demand alone heads toward roughly 134 GW by 2030. Europe's most ambitious public programme targets one gigawatt in orbit by 2050. A wildly successful decade might put a gigawatt up by the mid-2030s — well under one percent of demand.

Orbital compute is not a climate solution and cannot become one at any plausible scale in the next fifteen years. It is a relief valve for capacity that cannot be built in time, and its environmental advantages are a bonus rather than a reason for it to exist. Any deck leading with the carbon argument is selling something, and the giveaway is that the carbon argument is the only one in this sector that requires no engineering to be true.

The reason to build in orbit today is a premium buyer and a long queue. The reason to build in orbit in 2032 is that it will be the cheaper way to add a gigawatt. Everything in this chapter is a dividend on a decision made for a different reason.

Part VI sets the horizon and then makes the case against everything argued here.

Source notes, Part V

2. Gas turbine supply: three manufacturers of heavy-duty units at scale, backlogs extending into 2029–2030, lead times reported to have risen from two or three years in 2022 to roughly five today, guidance to plan on seven-to-eight-year timelines, and steep inflation in delivery-slot pricing — GE Vernova investor disclosures and Q1 2026 results, Siemens Energy backlog reporting, and industry analysis, 2025–2026. **The price-inflation figure in particular needs a primary source before publication.**

All figures in Chapter 18 are computed in this volume from the inputs tabled at the head of the chapter. The model applies a capital recovery factor at the stated cost of capital rather than straight-line amortisation, charges fleet attrition annually against hull value, and expresses unit costs per *effective* GPU-hour after availability and performance fade. A reader who prefers straight-line accounting and full availability will get numbers roughly 40% lower; **the gap between those two accounting treatments is larger than any engineering variable in this book, which is itself the point.**

Spacecraft build cost per kilogram remains the dominant and least-sourced input. The base case was moved from $1,000/kg to $2,000/kg in this revision because the Starlink comparison — 800 kg of one repeated simple design at thousands of units — is a weak analogy for 37 tonnes of deployable radiator and high-voltage distribution at perhaps dozens of units. A model should not take its base case from its most optimistic analogy.

Terrestrial baseline: $12M/MW facility capex over 15 years, industrial power at $60/MWh, $1M operations. The behind-the-meter case adds generation capital at $1,500/kW over 15 years against fuel at $40/MWh. Both need sourcing to a current market survey. Prevailing commodity GPU rental near $3.50 per hour is a volatile, discount-heavy public cloud figure and is the single number most likely to be wrong in either direction.

Chapter 19: grid intensity 370 gCO_2/kWh; methane combustion at 2.75 t CO_2 per tonne of fuel; launch propellant load and payload share order-of-magnitude. **Stratospheric soot and water-vapour forcing at high launch cadence remains an acknowledged and unquantified gap.**

3. The projection in Figures 18.6 and 18.7 applies Wright’s law at learning rates of 10%, 15% and 20% per doubling of cumulative deployed capacity, against an assumed deployment path reaching roughly one gigawatt in the mid-2030s. Learning rates in that band are observed across aircraft, semiconductors, photovoltaics, batteries and launch vehicles; **the specific rate for large deployable spacecraft is unobserved and is the projection’s central assumption.** The cost-of-capital path from 20% to 8% follows the financing maturation of utility-scale renewables and is an analogy, not a forecast. Revenue is deliberately held flat across the whole projection, which is conservative. **Every input is stated so the reader can rebuild the curve with their own numbers; anyone who does should expect a different date, and the range of defensible dates is roughly 2031 to 2034.**

Figures 18.1 to 18.7, 19.1 and 19.2 generated for this volume.

PART VI — THE HORIZON

The Data Centers in LEO — Ibrahim Quabboua, StarCap. Draft, Part VI of VI.

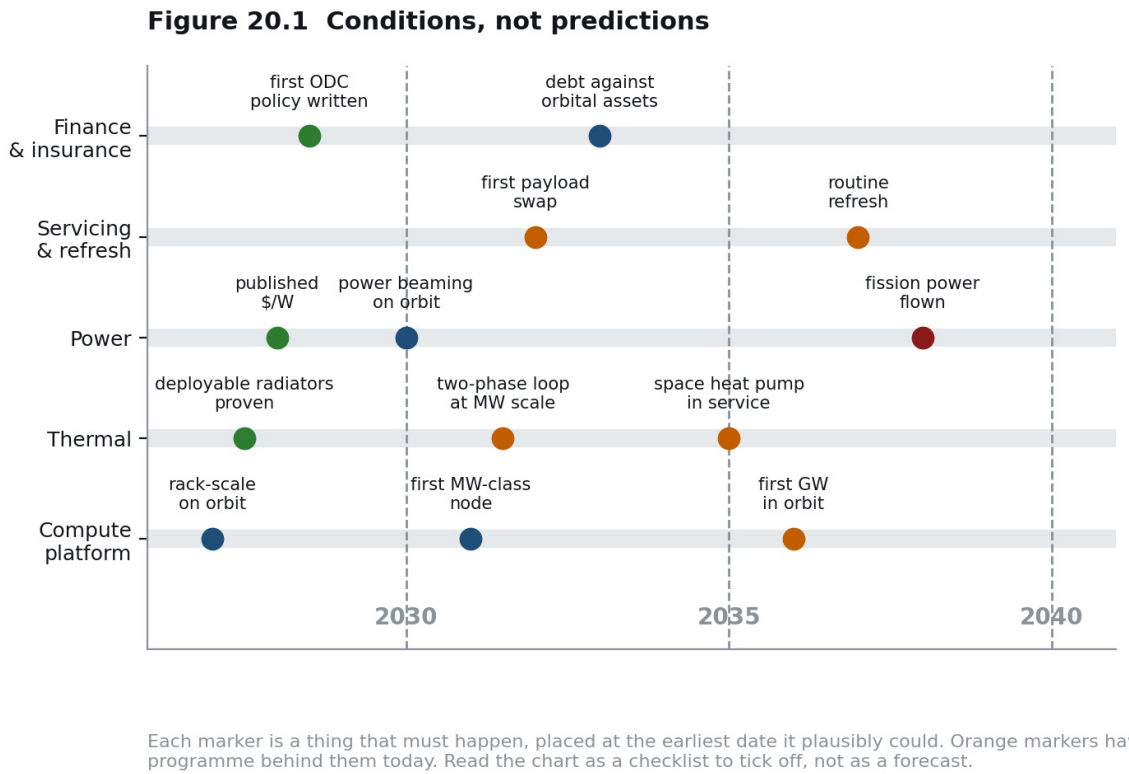
Chapter 20 — 2030, 2035, 2040

How to read a horizon

Predictions in this sector age badly and publicly. So this chapter does not make them. It sets out **conditions** — things that must happen for the sector to reach a given scale by a given date — and leaves the forecasting to you.

The distinction matters. “There will be a gigawatt in orbit by 2036” is a claim I cannot support. “There cannot be a gigawatt in orbit by 2036 unless a two-phase thermal loop is qualified at megawatt scale, a spacecraft factory reaches constellation economics, and a heavy vehicle sustains a launch cadence measured in hundreds per year” is a claim I can support completely, and it is far more useful, because each clause is something you can watch.

Figure 20.1 — *Conditions, not predictions*



Each marker is a thing that must happen, placed at the earliest date it plausibly could. Orange markers have no credible programme behind them today. Read the chart as a checklist to tick off, not as a forecast.

2030 — the proving decade

By 2030 the sector should have answered whether the machine works at all. What that requires:

Condition	Why it matters
Rack-scale compute operating on orbit with paying customers	Proves the vacuum-packaging and thermal architecture of Chapter 8
Deployable radiators and a direct-to-chip loop demonstrated in flight	Chapter 5’s easy regime, confirmed rather than modelled
A published price for orbital compute	Until this exists, every model including mine is theory
A published dollars-per-watt for a mass-produced orbital array	Chapter 8’s largest unresolved number
The first insurance policy written on a compute node	Chapter 16’s financing gate opens or does not
First megawatt-class node in flight or in build	The step from Chapter 8’s Node A to Node B

My expectation for 2030 is **tens of megawatts in orbit, not hundreds** — a handful of operators, most revenue from sovereign, defence and sensor-adjacent buyers rather than the commodity cloud, and the megawatt-scale thermal question closing.

That sounds modest and it is the most important rung on the ladder. Chapter 18's crossover requires roughly nine doublings of manufacturing experience, and tens of megawatts delivered to customers who pay a premium is precisely how the first four of them get funded. **A sector at 40 MW in 2030 is not a sector that failed to reach a gigawatt. It is a sector on schedule to.**

2035 — the manufacturing decade

Everything after 2030 is a factory question. Part V showed why: at bespoke spacecraft costs the business is absurd, and at constellation-scale manufacturing it works. Nothing between now and 2035 matters as much as which of those worlds arrives.

Condition	Why it matters
A spacecraft production line at $1,000–2,000/kg for a 30-tonne class vehicle	The single decisive variable in Part V
Two-phase thermal transport qualified above a megawatt	The unretired technical risk
A demonstrated payload swap or module replacement on orbit	Chapter 15's answer to the refresh problem
A second heavy-lift provider at sustained cadence	Removes the sector's single-supplier risk
Debt financed against orbital compute assets	The capital structure the sector needs to scale
Regional constellations under separate jurisdictions	Chapter 17's sovereignty logic playing out

If all six land, a gigawatt in orbit becomes a 2036–2038 proposition. If the first one does not, none of the others rescue it, and the sector stabilises at a few hundred megawatts serving specialised workloads.

2040 — the infrastructure decade

Two futures, and they are genuinely far apart.

If the manufacturing question resolved well, orbital compute by 2040 is boring infrastructure: several gigawatts across multiple sovereign constellations, routine on-orbit refresh, in-space assembly of structures too large to launch, possibly the first fission power in orbit, and a price for orbital GPU-hours that a procurement officer compares against terrestrial without excitement. The interesting companies by then are the ones nobody writes about — the terminal supplier, the radiator manufacturer, the servicing operator.

If it did not, orbital compute in 2040 is a real but modest industry doing exactly one thing extremely well: processing sensor data next to the sensor. Every Earth-observation constellation carries meaningful compute; nobody trains frontier models in orbit; the gigawatt filings are a historical curiosity of the mid-2020s. That outcome is not a failure. It is a useful industry that was oversold, which is the most common ending for a technology of this kind.

I do not know which one happens. I hold the position I hold because the first outcome is worth many multiples and the second is not worth zero.

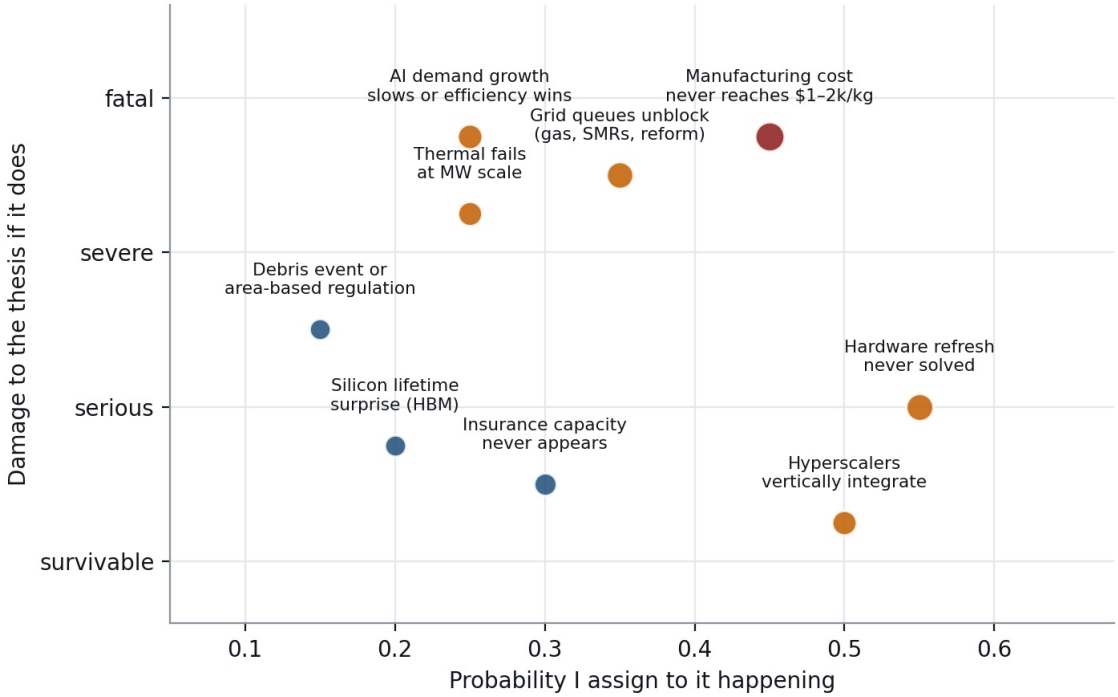
Chapter 21 — What would delay this, and what would prove it wrong

A guidebook that cannot argue against itself is a brochure. This chapter is the strongest case I can make against everything in this book, made as well as someone who disagrees with me would make it.

One framing note before the list. Chapter 18 showed that this sector's economics are a **date**, not a verdict: the arithmetic crosses somewhere between 2031 and 2034 on assumptions that require nothing new to be invented. Most of what follows therefore moves the date rather than cancelling the industry. Two items genuinely cancel it, and they are marked. The rest are schedule risk — which is the normal condition of an infrastructure business, and which is priced by *when* you buy rather than *whether*.

Figure 21.1 — *How this thesis dies*

Figure 21.1 How this thesis dies



Probabilities are the author's judgement over a ten-year horizon, not model output. They are printed so that a reader can disagree with a specific number rather than with a mood.

The two that would genuinely cancel it, and the ones that move the date

1. Manufacturing cost never falls (cancels). Part V's model is unambiguous: at $20,000 per kilogram of spacecraft, orbital compute costs $23 per GPU-hour against a $3.50 market. The entire thesis rests on a 37-tonne power-and-cooling spacecraft being manufacturable at Starlink-like economics — and Starlink satellites are around 800 kg of relatively simple hardware, not thirty tonnes of deployable radiator and high-voltage power distribution. **The analogy that carries the whole model may not hold**, and this is the most probable route to the sector failing quietly rather than dramatically.

2. The queue unblocks everywhere at once (cancels). Chapter 18 shows the orbital case only clears terrestrial past roughly seven and a half years of waiting, and then only in the premium-revenue world. That is a narrow margin against a moving target. Behind-the-meter gas is already the standard workaround, and although Chapter 18 shows the turbine order book has simply become the new queue, that is a supply constraint and supply constraints resolve. Small modular reactors are being financed specifically for data centres; fuel cells are shipping; ERCOT has restructured its interconnection review; political pressure to accelerate permitting for AI infrastructure is intense everywhere. **If time-to-power on Earth compresses to two or three years by any route, the availability argument — the strong form of the case, the one that survives when the cost case fails — evaporates entirely.** Nothing else in this book saves it.

2b. Capital stays expensive (moves the date). Figure 18.4 shows the thesis needs cheap manufacturing *and* infrastructure-grade capital together, and Figure 18.7 assumes the second arrives as the asset class matures — flight heritage, then contracted revenue, then insurance, then debt, exactly the path renewables walked. If that path stalls, the crossover slips by years. This is a financing problem masquerading as an engineering one, and it will not be solved in a cleanroom.

3. GPU prices fall (moves the date). Every cost figure in Chapter 18 is measured against a commodity rental rate near $3.50 an hour, set during a capital-expenditure boom. Halve it and orbital compute is unviable at any manufacturing cost in the model — as is a large share of the terrestrial industry. The premium buyers of Chapter 18 are the hedge against this, and they are a smaller market than the sector's addressable-market slides assume.

4. Demand growth slows (moves the date — and efficiency may help rather than hurt). The whole edifice assumes compute demand outruns power supply indefinitely. Two things could break that: a slowdown in AI capital expenditure, or continued rapid improvement in performance per watt. Note that Nvidia claims roughly ten times the inference throughput per watt from one generation to the next. **If efficiency gains outpace demand growth**

for even a few years, the power crunch that justifies going to orbit could resolve on Earth.

It is worth noting that this one cuts both ways, and Chapter 18 excluded the favourable half. Orbital costs scale with **watts** — radiators, arrays, mass — while revenue scales with **useful output**. Every doubling of tokens per watt doubles revenue against a fixed thermal and structural cost base, which improves orbital unit economics faster than terrestrial ones, where the power constraint was the thing being relieved. Efficiency is a risk to the *rationale* and a gift to the *arithmetic*, and which dominates is genuinely open.

The three that would be serious but survivable

5. Hardware refresh is never solved. The most *likely* of the failure modes, and the most under-analysed. If an operator must fly the same accelerators for a full spacecraft life, revenue per rack collapses as newer silicon arrives on Earth, and the returns in Part V's model — which assume three-to-five-year hardware amortisation — are simply wrong. This does not kill the sector; it caps it, and it shifts value to whoever solves servicing.

6. Thermal does not close at megawatt scale. Chapter 11 is honest that two-phase transport above a megawatt has no flight heritage. If it proves intractable, orbital compute stops at the hundred-kilowatt scale and becomes the edge-processing business of the 2040 downside case.

7. Hyperscalers vertically integrate. If SpaceX, Google and the large cloud providers build the compute layer themselves at a scale that makes independents irrelevant, the operator layer loses its value — though the power, thermal and relay layers keep selling into a larger market. This costs an investor the high-performance sleeve and leaves the picks-and-shovels intact, which is precisely why the portfolio in Part IV is weighted the way it is.

The ones I think are overrated as risks

Radiation was retired as a first-order risk by the testing described in Chapter 12; it is now a depreciation schedule. **Launch cost** is, per Figure 18.2, the fourth most important input and is already better than the model needs. **Latency** is a non-issue for the natural workloads. I could be wrong about all three, but each would need new evidence rather than new argument.

The falsifiers — what would change my mind, specifically

I would substantially **reduce** conviction if:

- a published spacecraft build cost for a large compute vehicle came in above $5,000/kg;
- **time-to-power on Earth** — by grid, turbine or reactor — fell below three years in two major US markets;
- **commodity GPU rental rates halved** without a corresponding premium market emerging;
- an operator disclosed a cost of capital above 18% for its constellation build;
- a second-generation compute satellite failed thermally on orbit;
- performance-per-watt improvements ran ahead of data centre demand growth for two consecutive years.

I would substantially **increase** conviction if:

- an operator published a price for orbital GPU-hours **at a premium to commodity** and sold capacity forward to a named sovereign or defence buyer;
- an operator raised **debt** against an orbital compute asset, which would prove the capital-cost problem is solvable;
- a spacecraft build cost near $1,000/kg was demonstrated for a vehicle above ten tonnes;
- a two-phase loop was qualified above 500 kW;
- a payload swap was demonstrated on orbit in any application;
- an insurer wrote a policy on a compute node at a rate below five percent.

Those are twelve specific, observable events. Write them down. Any of them arriving should move your view further than any amount of additional argument, including mine.

Chapter 22 — How to use this book

The point of a guidebook is that it survives contact with a company you have never heard of. So here is the whole book compressed into questions you can ask in a first meeting.

1. **What orbit, and why that one?** Altitude and plane determine drag, dose, sink temperature and serviceability simultaneously. A company that has not thought hard about this has not designed a spacecraft. *(Chapters 4, 6, 8, 15)*
2. **What temperature does your radiator run at?** Rejection scales as T^4. This one number tells you more about their mass budget than their whole deck. *(Chapter 5)*
3. **What does your power hardware cost per watt, and is that a quote or a slide?** The least-examined number in the sector. *(Chapter 8)*
4. **What does the spacecraft cost per kilogram to build, at what production volume?** The variable that decides the business. *(Chapter 18)*
5. **How do you move heat from the chip to the panel above 500 kW?** If the answer is “pumped two-phase,” ask what has flown. *(Chapters 5, 11)*
6. **What is your hardware refresh strategy, and what does your model assume about revenue per rack in year four?** *(Chapters 8, 15)*
7. **Is your constellation co-planar?** If not, it can never be serviced economically. *(Chapter 15)*
8. **Is the ground segment in your cost per GPU-hour?** It usually is not. *(Chapters 7, 14)*
9. **What is your cross-sectional area, and what have you budgeted for debris and insurance?** *(Chapter 16)*
10. **Who is the premium buyer, and what premium do they pay?** At commodity GPU prices this business does not clear. If the answer is “hyperscalers,” they have not read their own model. *(Chapter 18)*
11. **What cost of capital does your model assume, and can you actually raise at it?** Second most important input in the sector and almost never stated. *(Chapter 18)*
12. **What could SpaceX announce next quarter that would erase you?** If there is no good answer, there is no moat. *(Chapters 9, 13)*

And one question to ask yourself rather than them: **which of the eight falsifiers in Chapter 21 has moved since I last looked?**

Closing

The argument of this book has been narrower than the sector’s usual claims, and I want to restate it plainly at the end.

Artificial intelligence is no longer limited by chips or by the price of electricity. It is limited by the queue for power — by transformers, permits and interconnection studies measured in years. Orbit is the one place where a very large amount of energy is already available and unqueued, and since 2024 the price of getting there has fallen far enough to make using it arguable.

Going to orbit does not remove constraints. It substitutes a scheduling problem for a physics problem: an interconnection queue for a radiator, a transformer order book for a launch manifest. That trade is good if, and only if, the physics problem is smaller than the scheduling problem. Worked through honestly in Part V, the answer today is: not yet for a commodity buyer, already yes for a premium one — and the gap between those two closes on a curve, not on a hope.

That curve is the argument of this book. At today’s costs, orbital compute is a first-of-a-kind industry building single-digit units at venture rates, and it looks expensive, exactly as solar looked in 2010 and reusable launch looked in 2012. Apply the learning rates that every manufactured hardware industry has followed, add the financing maturation that every infrastructure asset class has followed, and the arithmetic crosses the commodity price somewhere between **2031 and 2034** — with nothing new invented, no breakthrough required, and revenue held deliberately flat throughout.

Premium buyers fund the descent. Defence, sovereign and sensor-adjacent customers pay enough today to buy the first nine doublings of manufacturing experience, and those doublings are what deliver the commodity market later. That sequence — pay a premium early, industrialise, cross over — is not a workaround. It is how every infrastructure industry in history has been built.

But the ideas underneath it are old, and they are not going away. Kardashev’s observation that a civilisation is bounded by the energy it can gather; Glaser’s satellite; O’Neill’s high frontier; Dyson’s insight that the waste heat is

the visible signature of thinking at scale. None of those people were wrong. They were early, by a factor of about fifty thousand dollars per kilogram.

What changed is a price, and prices keep moving. What follows from it is engineering — and the engineering, part by part, is what this book has been about.

If the curve holds, the 2030s are the decade in which computation stops being something a civilisation does on its planet’s surface and starts being something it does in its planet’s sunlight. That is a large claim, and this book has tried to earn it the only way a claim like that can be earned: by writing down every number, showing where it came from, and marking clearly the four that could still prove it wrong.

APPENDICES

Appendix A — Constants, formulas and conventions

Quantity	Value used
Solar constant at 1 AU	1,361 W/m²
Stefan–Boltzmann constant σ	5.670×10^{-8} W/m²K⁴
Standard gravity g	9.80665 m/s²
Earth gravitational parameter μ	3.986×10^{14} m³/s²
Orbital velocity at 650 km	≈ 7.53 km/s
Aluminium density	2,700 kg/m³
Multi-junction cell efficiency	30%
Radiator emissivity ε	0.85
Effective LEO sink temperature	200–270 K
Radiator areal mass	5–12 kg/m²
Array specific power	~150 W/kg
Rack convention	200 kW, 1.4 t, $3.5M

Formulas used throughout

- Rocket equation: $\Delta v = I_{sp} \cdot g \cdot \ln(m_0/m_f)$
- Radiative heat rejection: $q = \varepsilon \sigma (T^4 - T_{sink}^4)$
- Plane change cost: $\Delta v = 2 v \sin(\Delta i/2)$
- Drag deceleration: $a = \tfrac{1}{2} \rho v^2 C_d (A/m)$
- Beam divergence: $\theta \approx \lambda / D$

Appendix B — Glossary

ADCS — attitude determination and control system; how a spacecraft knows and changes where it is pointing. **COTS** — commercial off-the-shelf; ordinary terrestrial hardware flown without space-specific redesign. **Dawn–dusk SSO** — a sun-synchronous orbit riding the day/night terminator, giving near-continuous sunlight. **GCR** — galactic cosmic rays; high-energy particles from outside the solar system, effectively unshieldable. **HBM** — high-bandwidth memory; the memory stacks beside an AI accelerator, and its most radiation-sensitive part. **ISL** — inter-satellite link; a connection between spacecraft, now usually optical. **ISAM** — in-space assembly and manufacturing (and servicing). **Kessler syndrome** — self-sustaining collisional cascade of orbital debris. **Latch-up** — a destructive parasitic conduction path triggered by a single particle strike. **PMAD** — power management and distribution. **PUE** — power usage effectiveness; total facility power divided by IT power. **SAA** — South Atlantic Anomaly; where the inner radiation belt dips closest to Earth. **SEU / SEE** — single-event upset / effect; instantaneous fault caused by one particle. **Specific power** — watts generated per kilogram of hardware. **TID** — total ionising dose; cumulative radiation damage, measured in rad(Si). **Two-phase loop** — a cooling loop in which the coolant boils, carrying far more heat per unit of flow.

Appendix C — Sources

Full source notes appear at the end of each part. Primary sources relied on most heavily:

- Google Research, *Exploring a space-based, scalable AI infrastructure system design* and *Towards a future space-based, highly scalable AI infrastructure system design* (arXiv:2511.19468) — solar yield, reference orbit, cluster architecture, and the TPU radiation testing that underpins Chapters 6 and 12.
- Voyager Technologies Form 10-K, 10 March 2026 — the only publicly disclosed Starship contract price.

- NASA ISS Active Thermal Control System documentation and the *Journal of Spacecraft and Rockets* radiator literature — the 70 kW benchmark in Chapter 11.
- EPRI, *Powering Intelligence 2026*, and PJM interconnection data — the queue figures in Chapter 1.
- Thales Alenia Space, ASCEND study results, June 2024 — Chapter 17.
- SpaceNews and Xinhua reporting on the Three-Body Computing Constellation — Chapter 17.
- StarCap Fund I Thesis v2, 27 July 2026 — company-level research, disclosed as the author's own.

Verification status. This is a working draft. Items flagged in the source notes as requiring verification — most importantly spacecraft build cost per kilogram, array cost per watt, space insurance market scale, and per-orbit debris flux — must be sourced to primary documents before publication. The structure of the arguments does not depend on them; several of the conclusions do.

Appendix D — Disclosure

The author is the founder of StarCap and holds, or intends to hold, positions in private companies across the layers described in Part IV, including in the compute platform, power, mobility and network layers. Several companies discussed favourably in this book are held. Several companies discussed unfavourably are not.

This book is not investment advice, does not constitute an offer of any security, and is written for readers who intend to reach their own conclusions. Where the author's fund research has been used in place of an independent source, this is stated in the source notes. Valuations, funding rounds and contract values in this sector change quarterly and every such figure printed here should be assumed out of date.

Appendix E — The companies, layer by layer

How to read this appendix. It contains no valuations, no round sizes and no dated financials. That is deliberate: those numbers move every quarter and would make the book wrong within a year of printing. What follows is structural judgement — what each company does, why it might matter to orbital compute, where it is fragile, and the single event that would most change the assessment. Those things age in years rather than months.

Everything here is the author's opinion as of **27 July 2026**, informed by his own fund's research, and should be read alongside the disclosure at the front of this book. Companies are grouped by the chapter that covers their layer.

One organising idea. Chapter 18 dates the moment orbital compute clears the commodity market to the early 2030s, and identifies the mechanism that gets it there: premium buyers funding the first nine doublings of manufacturing experience. **The right question to ask of every company below is therefore not "is this profitable today" but "is this company positioned to be on the curve when it crosses, and can it survive the descent."** Companies with premium revenue now and manufacturing leverage later are the ones that compound through it.

Launch and mobility *(Chapter 9)*

SpaceX. Sets the price everyone models against, holds the only publicly disclosed heavy-lift contract price, and — having filed for a very large compute constellation and merged with an AI company — competes directly with the operators who depend on it. There is no clean way to hold a view on this sector without holding a view on SpaceX. *Fragile:* nothing structurally; the risk it presents is to everyone else. *Watch:* whether Starship flies paying customers at contracted rather than target prices, and turnaround time between heavy flights.

Rocket Lab. The credible second heavy provider, and since early 2026 also a power-layer entrant with silicon arrays aimed explicitly at orbital data centres. Its strategic significance exceeds its market share: a sector priced off one supplier has a single point of commercial failure. *Fragile:* competing in two layers against much larger incumbents in each. *Watch:* whether the silicon-array bet wins design slots against multi-junction suppliers — that contest decides Chapter 8's open architectural question.

Impulse Space. Last-mile orbital mobility, founded by SpaceX's first employee, revenue-generating with flight heritage. Chapters 5 and 6 push compute constellations into specific awkward orbits, and something has to put them there. *Fragile:* the most mature private name in the theme is also the one where expectations are highest. *Watch:* whether compute constellations start buying orbital transfer as a line item rather than accepting the orbit the rocket gives them.

Power *(Chapter 10)*

K2 Space. The satellite *is* the power system — a large bus designed around high power, high-voltage avionics and a high-power thruster, sized for what heavy lift now permits. The cleanest structural expression of the view that power, not production rate, is the durable differentiator in buses. *Fragile:* a competing supplier on the same government programme is direct evidence the layer is contestable. *Watch:* on-orbit confirmation of high-voltage bus performance at rated power.

Star Catcher. Concentrates and beams solar energy onto client spacecraft's *existing* arrays — no retrofit, no custom receiver. Shared infrastructure that sells to every operator rather than requiring you to pick one, with contracted customers ahead of first flight and a repeat founder who previously built and sold an in-space manufacturing company. *Fragile:* small team against a capital-intensive constellation; the hard version of the physics has never been done. *Watch:* a spacecraft-to-spacecraft beaming demonstration on orbit — it would create a category.

Zeno Power. Radioisotope power. Terrible specific power, irrelevant for megawatts, and exactly right for anything that must survive the dark. The moat is not the physics — which is sixty years old — but a recycled-isotope fuel supply chain that venture money cannot compress. *Fragile:* revenue is defence and civil space, not orbital compute; it is a dual-use holding, not a pure-play. *Watch:* whether the fuel supply agreements convert into repeat unit orders.

Aetherflux. Generates power and consumes it in its own compute payload — the only name owning both ends. Strongest founder record in the group and, at the time of writing, nothing demonstrated in orbit. That combination is precisely why it should be held smaller than its narrative justifies. *Fragile:* weakest ratio of expectation to evidence in the sector. *Watch:* the first power-beaming demonstration from low Earth orbit.

Thermal *(Chapter 11)*

Sophia Space. The closest thing to an investable pure-play in the layer that most determines whether the sector works. Its tile architecture makes thermal integration the product rather than a subsystem. *Fragile:* early stage, and strategically ambiguous — the same tile can be sold *to* operators or assembled *into* a competing data centre, and only the first configuration is a supplier business. *Watch:* a signed supply agreement with a named operator. Absent that, it is competing with its customers.

Spacebilt. Thermal tiles already supplying commercial orbital data-centre nodes — the best technology fit in the layer and not investable at venture scale. Track it as an acquisition target rather than a position. *Watch:* who buys it.

Celeroton and the turbomachinery suppliers. Oil-free high-speed compressors for cryocoolers and heat pumps: the exact hardware Chapter 11's heat-pump argument requires. European industrial suppliers, no venture round. *Watch:* anyone packaging this into a space-rated heat pump — the highest-leverage unbuilt product in this book.

Compute platform *(Chapter 13)*

Starcloud. Flight heritage leader: flew a commercial accelerator, trained a model in orbit, and its second vehicle is the first serious test of direct-to-chip liquid cooling into deployable radiators at rack scale with named cloud customers. The single most informative company in the sector whether or not you own it. *Fragile:* constellation scale requires capital far beyond what has been raised, and reported strategic interest from the dominant launch provider is both an exit path and a dependency. *Watch:* paying workloads running on orbit, and any published price for orbital compute.

Google (Project Suncatcher). The only organisation designing its own silicon with orbit in mind, and the only one publishing its physics. Not investable at this scale; its progress is nonetheless the strongest de-risking event the sector has had. *Watch:* the prototype satellites, and any published on-orbit single-event data.

Axiom Space. The hosted-platform route, with data-centre nodes already on orbit and a compute business cross-subsidised by existing revenue — a genuinely different risk profile in a sector full of pre-revenue narratives. *Fragile:* leadership churn and a binary government station selection sit underneath it. *Watch:* whether third parties choose to host compute on a station rather than fly their own free-flyer.

Loft Orbital. Flies other people's payloads on standardised shared satellites and lets customers run inference on hardware already in orbit — the lowest-friction route to orbital compute for a customer who does not want to own a satellite. A sovereign manufacturing joint venture gives it a distribution channel the US-centric names lack. *Fragile:* a shared-platform business is a margin business. *Watch:* how much of its bookings become compute rather than sensing.

Apex Space. Productised buses built ahead of demand, running toward high annual production. Excellent, and held small precisely because rate manufacturing is a real moat against incumbents and no moat at all against the next well-funded productiser. Its demand is proliferated national-security constellations, not orbital compute. *Watch:* whether a compute operator standardises on a bought bus rather than building its own.

Network *(Chapter 14)*

Kepler Communications. A deployed optical relay constellation running as an IP mesh, with compute and storage on the relay satellites themselves. Simultaneously the backhaul for orbital compute and a distributed compute platform of its own — and Chapter 14 showed the relay layer's value is duty cycle, a twentyfold multiplier on the usefulness of identical hardware. *Fragile:* if the dominant operator opens its laser mesh as a commercial service, pricing compresses immediately. *Watch:* exactly that.

Skyloom and the terminal suppliers. Microradian pointing mechanisms with lasers attached. Unglamorous, necessary, and sold to everyone. *Watch:* whether terminal supply becomes a bottleneck as constellation counts rise.

Outside the roster

ADA Space, Zhejiang Lab and the Chinese programmes *(Chapter 17)* are excluded by mandate rather than by capability, and they currently have more dedicated compute hardware in orbit than anyone else. The relevant question for a Western investor is not whether to buy them but what a state-backed constellation that does not need a commercial return does to price.

Thales Alenia Space and the ASCEND consortium *(Chapter 17)* have produced the most rigorous public feasibility analysis in the field, attached to a 2050 timeline. Europe's realistic roles here are customer, component supplier and regulator.

The two vacancies

Two positions in this appendix do not exist and should:

1. **A megawatt-class thermal transport company** with qualified two-phase hardware, selling to operators rather than competing with them.
2. **An on-orbit hardware refresh business** — servicing, module replacement, or a modular architecture sold as a refresh service.

Chapters 11 and 15 argue that these are the sector's two unserved constraints. If either company is founded and funded well, it should be assumed to be the most valuable new entrant in this appendix.

End of book.