AI, Energy and the Price of Intelligence

If AI demand booms and power becomes the binding constraint, the frontier slows and compute gets scarce. The intuitive conclusion is that the price of intelligence rises. The evidence points elsewhere: the price of a fixed level of intelligence keeps collapsing, the frontier premium persists, and the risk that matters is a margin collapse, not a demand collapse.

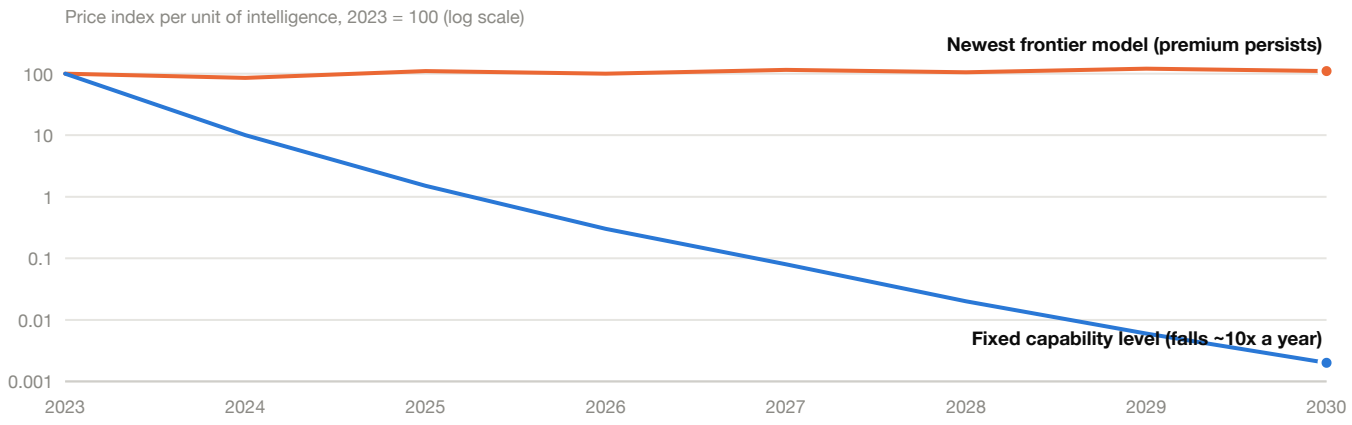
THE BULL CASE

- **Demand is elastic.** Every fall in price has been absorbed by new uses. Agents multiply tokens per task by 10 to 100x. Jevons has held every year so far.
- **Efficiency is independent of the grid.** Chips, distillation, quantization and sparse models push intelligence per watt up 2 to 3x a year whether or not anyone builds a power plant.
- **Inference escapes the constraint.** Serving is distributable: behind the meter, following cheap power across time zones, or offshore. Only frontier training needs a gigawatt in one place.
- **Supply is arriving.** Gen IV fission permits, hyperscaler nuclear deals, and solar plus storage as the cheapest new generation in most of the world.
- **Scarcity protects incumbents.** Whoever owns power, land and interconnect holds a real moat while the frontier premium lasts.

THE BEAR CASE

- **Energy caps US frontier growth.** Interconnect queues, turbine backlogs and permitting kill exactly the gigawatt sites frontier training needs. Models improve more slowly than chip supply allows.
- **The frontier moves.** China adds more generation each year than the entire US data center fleet draws. The strategic risk is geographic, not price.
- **Plateau plus open models strips the premium.** If capability stalls for 18 months while open weights close the gap, frontier capex is stranded against commodity pricing. This is fiber in 2000: usage kept compounding, the builders went bankrupt anyway.
- **ROI disappointment comes first.** The likelier trigger is buyers stopping because agents did not deliver value, not because models stopped improving.

Two curves, not one: the price of intelligence depends on which intelligence



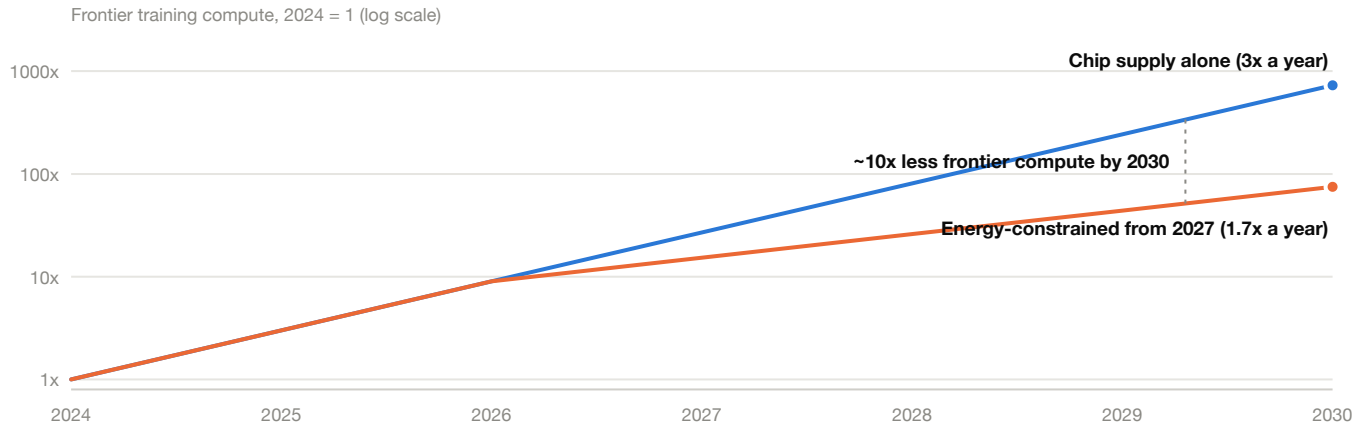
Illustrative scenario, not a forecast series. Blue tracks the pattern since 2023: the price to serve a given capability level falls roughly tenfold a year as smaller, cheaper and open models match it. Orange is the newest frontier model, which holds a premium because its supply is bounded by the largest training runs. An energy constraint tightens orange and leaves blue alone.

Where the constraint bites, and where it does not

Energy does not make intelligence expensive. It caps how fast the top of the capability curve can move, and it decides where that work happens. Four links in the intuitive chain are weaker than they look.

- **Scarcity shows up as rationing, not price.** Labs compete on price and ration by other means: rate limits, usage caps, enterprise-first allocation, shifting compute between training and inference. The symptom of an energy-constrained industry is a late model and tighter limits, not a higher API bill.
- **Open models do not free up watts.** An open-weight token costs roughly the same joules as a comparable closed one. Under a power constraint they compete for the same scarce energy. They collapse pricing power, not compute demand.
- **Training bears the constraint; inference does not.** The gigawatt campus is the bottleneck. Serving can go wherever power is cheap.
- **Demand does not asymptote even if quality does.** “Good enough” changes who gets paid, not how many tokens get run.

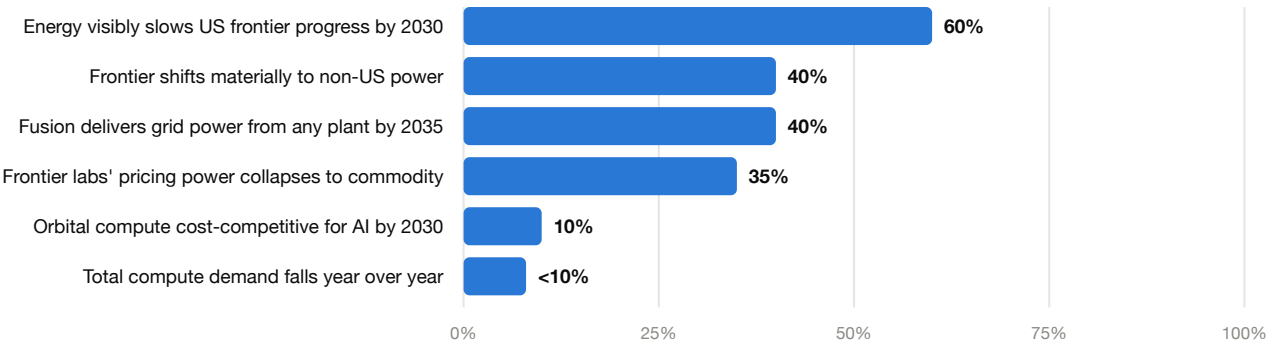
What an energy cap does to the frontier



Illustrative compounding. If US power additions cap frontier compute growth near 1.7x a year from 2027, the frontier in 2030 has roughly a tenth of the compute it otherwise would. Models still improve on a slower curve, and the gap is an incentive to train wherever power is unconstrained.

> **The collapse to watch is a margin collapse, not a demand collapse. Fiber got built, usage kept compounding, and the price per bit still destroyed the builders.**

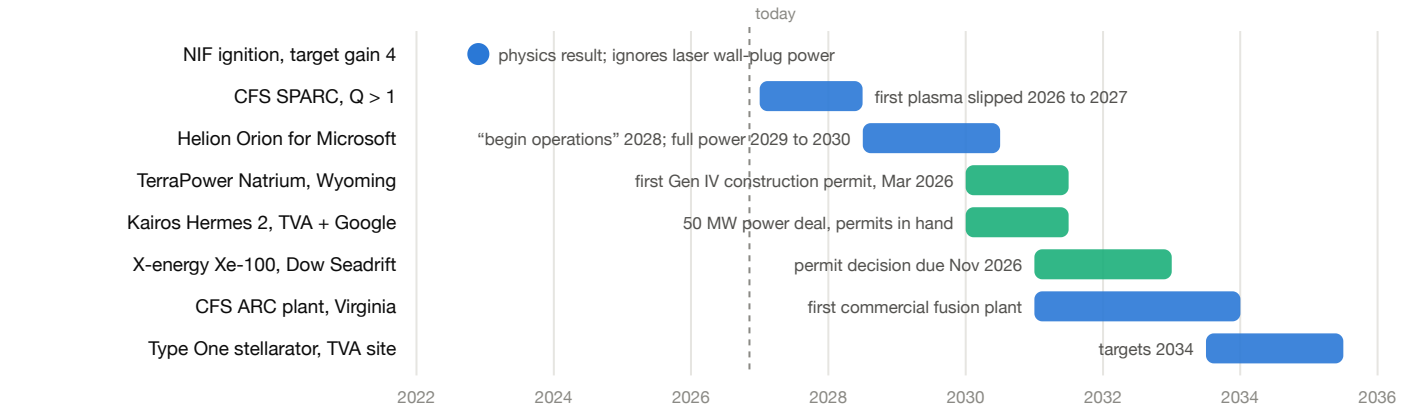
Scenario odds through 2030, with two energy markers



Subjective probabilities from this analysis. Fusion and orbital compute are the two supply-side stories most often offered as the way out, and neither arrives inside the window where the constraint binds.

The supply side: what actually arrives, and when

Fusion is about a year from its first commercially relevant net-energy demonstration and a decade or more from meaningful grid power. Fission is the real near-term breakthrough, moved by permits and hyperscaler money rather than new physics. Solar plus batteries already happened. None of it changes the 2028 to 2030 window.



Bars show the plausible range for first sustained output; dates from company statements and NRC filings.

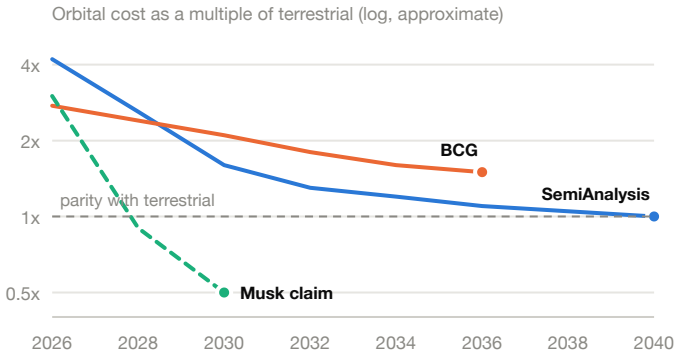
FUSION ODDS

COMPANY	GRID BY ~2035	EVER MATTERS
Commonwealth Fusion Systems	30%	60%
Helion	20%	40%
TAE	10%	25%
Pacific Fusion	10%	30%
Type One Energy	10%	30%
Stellarator others (Proxima)	5%	30%
Zap Energy	5%	20%
General Fusion	<5%	10%

CFS rests on the well-understood tokamak physics behind ITER, made small by high-field magnets. Helion has the best economics if it works and has never published a Q value. AI shortens design loops by perhaps 20 to 30% but does not fix magnets, materials or tritium.

MUSK'S SPACE DATA CENTERS

His argument is not that Earth energy fails. It is that terrestrial deployment is capped by grids and politics, orbital solar yields about five times a ground panel, and Starship collapses launch cost. The binding constraint in orbit is cooling, not power: radiation is the only way to shed heat, and radiator mass can exceed the compute itself. The economics are a pure launch-cost bet, and Starship has not yet flown a payload to orbit.



Approximate readings of published models. SemiAnalysis: more than 4x terrestrial cost in 2026, parity near 2040. BCG: a persistent premium near 1.5x a decade out. The story is strongest as a reason a million Starship launches become a necessity, and it was told in an IPO filing.

> **If watts are the scarce asset, rent flows to whoever owns power and interconnect. That is the world Musk is positioning for, whether or not the orbital math closes.**

Sources. CFS SPARC and Virginia plant pages; Axios (Sept 2026) on Helion's Orion timeline; Fusion Industry Association 2026 report; LLNL on NIF's 8.6 MJ shot; Live Science on Wendelstein 7-X; X-energy 2026 8-K and 10-Q; SMR Intel NRC tracker; SemiAnalysis, "The Case for Space Datacenters"; BCG, "Space-Based Data Centers: Cost and Outlook to 2040"; TechCrunch (Feb, May 2026) and NPR (Apr 2026) on SpaceX orbital compute; WEF on orbital cooling; DeepMind and EPFL tokamak control. Probabilities and scenario curves are the author's own.