FRONTIER RESEARCH

ENERGY & COMPUTE · DECEMBER 2024

The Reactor and the Rack

Why the AI power constraint and the small-modular trade are not the same trade.

Leonardo Fabbri
Founder & Chief Executive · AI Lab Monaco

Artificial Intelligence Laboratory Monaco

Between 20 September and 3 December 2024, four of the largest buyers of electricity on earth committed to nuclear power. Constellation agreed to restart Three Mile Island Unit 1 for Microsoft. Google contracted with Kairos Power. Amazon anchored a funding round at X-energy and signed with two utilities. Meta went to the market asking for one to four gigawatts. The tape read this as the moment small modular reactors became the answer to the AI power problem. Read the agreements themselves and a narrower picture emerges: almost every megawatt with a delivery date inside this decade comes from a reactor that already exists, and almost every small-modular commitment is an option on the 2030s. The AI capital cycle and the reactor construction cycle are roughly a decade out of phase. That gap is the subject of this note.

The load is real. The forecast is not precise.

Three credible houses published estimates of United States data-centre electricity demand in 2024. EPRI put data centres at between 4.6% and 9.1% of national generation by 2030.[1] Goldman Sachs put them near 8% by the same date.[2] Lawrence Berkeley National Laboratory, publishing this month, measured 4.4% in 2023 and projected a range of 6.7% to 12% by 2028.[3] The direction of travel is not in dispute. The magnitude is: the credible band spans a factor of roughly two and a half. A range that wide is not a forecast. It is a statement that the denominator is unknown.

What was actually signed.

Of the capacity contracted in the autumn, the megawatts that arrive before 2030 come from large light-water reactors that already exist — Three Mile Island Unit 1 at 835 MW, targeted for 2028;[4] Palisades, targeted for 2025 on a federal loan that closed in September.[5] The small-modular content is different in kind. Google's agreement with Kairos targets a first reactor by 2030 and 500 MW only by 2035.[6] Amazon's Washington State project begins at four modules.[7] Meta's request for proposals asks for capacity that has not been designed, sited or licensed.[8] These are real commitments. They are not near-term electrons.

Why the phase difference matters.

The compute cycle is being spent now. What clears the market between 2025 and 2029 is gas turbines, uprates and restarts of the existing fleet, demand-side flexibility, and above all position in the interconnection queue. The marginal megawatt for this cycle is not a new reactor. A portfolio expressing "AI needs power" through small-modular equities is therefore expressing a view about 2032 while being marked against the news flow of 2027.

Cost, fuel, permission.

Three constraints bind, in that order. Modularity is a cost hypothesis, and it has never been tested at commercial scale in the West: the last Western build, Vogtle 3 and 4, delivered at roughly two and a half times its original estimate and about seven years late.[9,10] Fuel is a physical constraint — the high-assay low-enriched uranium that most advanced designs require is produced in the United States at a first-phase contracted rate of 900 kilograms a year,[11] against a Department of Energy estimate of more than 40 tonnes a year needed by 2030.[12] Permission is the newest of the three, and the least priced.

The order nobody read.

On 1 November 2024 the Federal Energy Regulatory Commission rejected the amended interconnection agreement that would have allowed Talen Energy to serve up to 960 MW directly to an Amazon data centre at Susquehanna.[13] The Commission did not prohibit co-location; it declined to accept that agreement on that record and left the general question open. But co-location — putting the load beside the reactor, behind the meter — is what makes the nuclear-and-data-centre pairing economically distinctive. Until the structure settles, the value of a reactor to a hyperscaler turns on a tariff question, not an engineering one.

The purest exposure is not listed.

The listed small-modular names are pre-revenue and have not built a reactor. The names with revenue — fuel cyclers, forgers, turbine islands — are diversified businesses for which small modular reactors are a rounding error this decade. The developers with the most credible near-term fleets are private. Position sizing should reflect that, not the narrative.

Leonardo Fabbri
Founder & Chief Executive, AI Lab Monaco

Monaco · December 2024
Frontier Research

CONTENTS

Eight sections. The first three establish what is being asked of the grid, what was actually contracted this autumn, and what has actually been built anywhere in the world. The middle three take the three constraints that decide whether small modular reactors arrive in time — cost, physical supply, and regulatory permission. The last two deal with implementation and with what would prove us wrong.

01 **The load** 05
What the compute build actually asks of the grid — and how wide the honest range still is

02 **Ninety days in autumn** 07
What the hyperscalers signed between September and December, and when the electrons actually arrive

03 **What has actually been built** 09
The world's operating small-modular fleet, in full, and what the empirical record does and does not license

04 **The cost hypothesis** 11
Vogtle, the Carbon Free Power Project, and the number modularity has to beat to matter

05 **Three physical constraints** 13
Enriched fuel, licensing throughput, and a supply chain that was allowed to atrophy

06 **Behind the meter** 15
The order of 1 November, and why the structure of the trade is not yet settled

07 **The vehicle problem** 17
What is listed, what is private, and what the difference implies for position sizing

08 **Risks and falsification** 19
The named monitorables that would confirm or break this argument, and the windows they fall in

— **References · Methodology, scope and disclosures** 21
Forty numbered sources, palette convention, and the standing statement on the nature of this publication

EXHIBITS

01 The compute cycle and the reactor cycle are out of phase
02 Three houses, three answers, one direction
03 What makes this load different from the last one
04 The autumn contracts, sorted by first electron
05 Most of the contracted capacity is not new, and most of the new capacity is not near
06 The operating small-modular fleet is two units, neither of them Western
07 Vogtle is the benchmark modularity has to beat
08 The first Western small-modular project died on price, not on physics
09 Enrichment capacity is roughly two orders of magnitude short
10 Behind the meter or in front of it: what the November order left open
11 The purest exposure is not listed
12 Placeholder — total return of the listed complex against the index
13 What would confirm this argument, and what would break it

COVER DATE
December 2024
Date of the underlying research

EXTENT
23 pages · 13 exhibits · 40 sources
One exhibit published as a reserved placeholder

AUTHOR
Leonardo Fabbri
Founder & Chief Executive, AI Lab Monaco

Five conclusions, each stated with the confidence we actually hold it at. Where a judgment depends on a range rather than a point, we give the range. Where it depends on something that has not happened yet, Section 08 names the observable that would settle it.

1 **The load is real; the forecast is not precise, and should not be traded as if it were.**
The three most-cited 2024 estimates of the United States data-centre electricity share put the late-decade number somewhere between 4.6% and 12%. We present the spread rather than choosing within it, and note that the last time this forecast was attempted the extrapolation from workload to energy proved wrong by an order of magnitude in the same direction. High conviction on direction; low conviction on magnitude.

2 **What the hyperscalers bought in the autumn was mostly not small modular reactors.**
The contracted megawatts with a delivery date before 2030 are restarts and re-contracting of existing large light-water plant. The small-modular content consists of development agreements, equity and a request for proposals — commitments whose first electrons are 2030 at the earliest.

3 **The two cycles are roughly a decade out of phase, and nothing announced this year closes the gap.**
The compute capital cycle is being spent between 2025 and 2029. The earliest credible Western small-modular electron is 2030, and fleet-scale delivery is mid-decade after that. Nothing announced in 2024 shortens that. High conviction on the structural claim; the timing claim is the one to monitor, and Section 08 names the two observations that would overturn it.

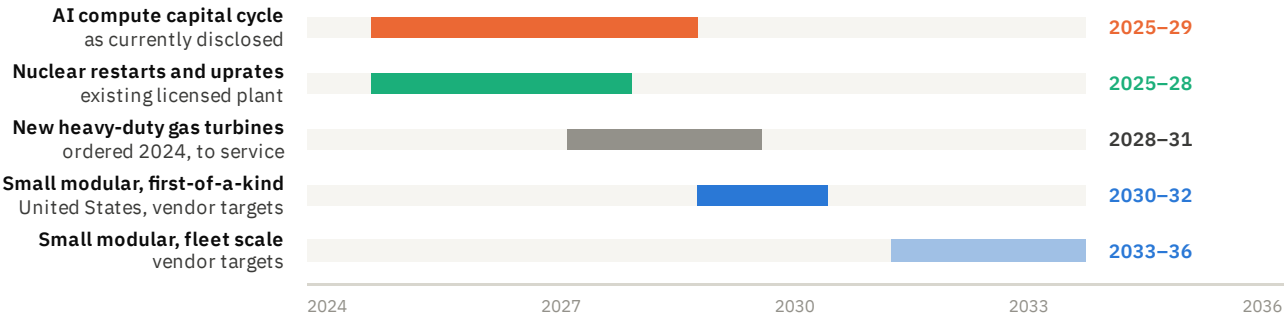
4 **Cost, fuel and permission bind before technology does — and permission is the least priced of the three.**
Modularity remains a cost hypothesis untested in the West. Domestic high-assay enrichment is roughly two orders of magnitude below the 2030 requirement. And the co-location structure the whole pairing depends on was, on 1 November, rejected in its first serious test.

5 **The purest exposure is private, so the listed complex is a poor instrument for the view.**
Listed pure-plays are pre-revenue with no reactor built; listed revenue-generating names are diversified businesses for which small modular reactors are immaterial this decade. The security that has actually captured the theme so far is the existing reactor, not the future one. This is a sizing constraint rather than an argument against the thesis, but it is a binding one.

EXHIBIT 01

The compute cycle and the reactor cycle are out of phase

Indicative delivery windows for the principal sources of firm supply, against the window in which the current artificial-intelligence capital cycle is being spent.



Source: AI Lab Monaco, from company announcements and public filings cited at references 4–8 and 14–19. **These windows are our characterisation of vendor and operator targets, not published point estimates**, and vendor targets for first-of-a-kind nuclear plant have historically slipped. The gas-turbine window reflects reported 2024 order-to-service lead times for heavy-duty units and is the least firmly sourced row in this exhibit.

01

DEMAND

The load

Everyone agrees data centres will consume materially more electricity by the end of the decade. Nobody agrees by how much, and the disagreement is wider than most capacity plans assume.

Start with what is measured rather than forecast. Lawrence Berkeley National Laboratory, publishing this month, puts United States data-centre consumption at roughly 176 terawatt-hours in 2023 — about 4.4% of national electricity use, against roughly 1.9% in 2018.[3] That is a doubling of share in five years, and it happened before generative-model training became a material line item. The base case for further growth does not require heroic assumptions.

The forecasts are another matter. EPRI, in May, published four scenarios putting data centres between 4.6% and 9.1% of national generation by 2030.[1] Goldman Sachs, in April, projected 160% growth in data-centre power demand and a share near 8% by the same date.[2] Berkeley's own projection runs to 6.7%–12% by 2028.[3] Taken together the credible band spans a factor of roughly two and a half on the same variable in the same decade.

Before accepting any of them, note what happened last time this forecast was made. Between 2010 and 2018 global data-centre compute grew by something over five hundred per cent while data-centre electricity use grew by about six per cent, because efficiency gains at the server, the rack and the facility absorbed almost all of it.[32] Forecasts made in 2010 that extrapolated workload growth into energy growth were wrong by an order of magnitude. Whether that decoupling continues is the central open question inside every range on this page, and it is the strongest available argument for the bottom of the band.

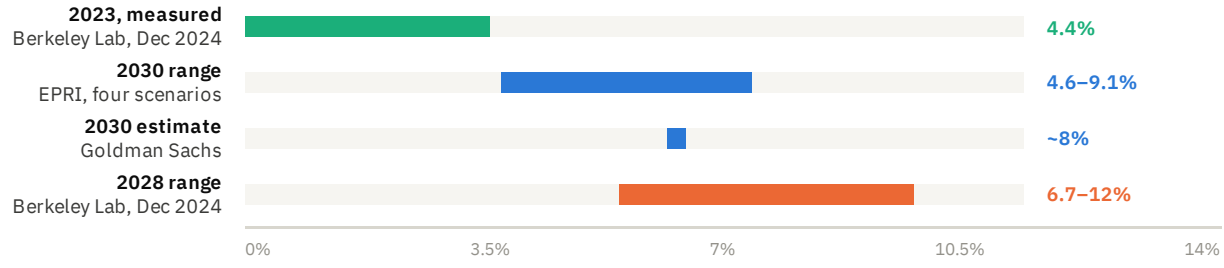
That spread is not a failure of the analysts. It reflects genuine uncertainty about accelerator shipment volumes, utilisation rates, the pace of efficiency gains at both the chip and the facility level, and how much inference workload migrates to the edge. But it does mean something practical: a capacity plan built against the top of the range and financed against the bottom of it is the standard way utilities and developers get into trouble. We carry the range forward rather than a point, and we say so wherever it matters.

There is a second distinction that gets lost in the headline percentages. These are estimates of *energy* — terawatt-hours over a year. The problem a hyperscaler is actually solving is *capacity*: firm megawatts available at three in the morning in February. A data centre is not a large factory. It is a large factory that never stops. Its load factor is close to unity, and it degrades badly under interruption, because the assets inside it depreciate on a three-to-five-year clock whether or not they are running.

EXHIBIT 02

Three houses, three answers, one direction

Data centres as a share of United States electricity consumption. Measured 2023 value, and the published forecast ranges, on a common axis.



Source: EPRI, *Powering Intelligence*, May 2024;[1] Goldman Sachs Research, April 2024;[2] Lawrence Berkeley National Laboratory, *2024 United States Data Center Energy Usage Report*, December 2024.[3] Target years differ — Berkeley's upper range is a 2028 figure and is therefore not directly comparable with the 2030 rows. Bars are positioned on a common percentage axis; the Goldman row is a point estimate rendered with a nominal width for legibility.

The continuity of the duty cycle is the whole argument for firm generation, and it is a good argument. Intermittent capacity can supply the energy; it cannot by itself supply the availability, and the firming — storage, gas, or overbuild plus curtailment — carries a cost that belongs in the comparison. Where a buyer has also committed publicly to carbon-free power on an hourly rather than annual matching basis, the set of technologies that clears is small. Nuclear is in it. That is not a narrative; it is arithmetic.

The commitments themselves are worth reading precisely, because they are stricter than they are usually described. Annual matching — buying enough renewable output over a year to equal consumption — can be satisfied with solar and wind almost anywhere. Hourly matching, which Google adopted as a 2030 target in 2020 and which several peers have since echoed, cannot.[34] It requires generation that is there in the specific hour the load draws. That is a procurement constraint with the force of a physical one, and it is self-imposed, which makes it more binding rather than less.

Density compounds the siting problem. Vendor materials for the highest-density 2024 rack designs describe cabinet-level draws on the order of 120 kW, against roughly 10–15 kW for a conventional enterprise rack.[20] The consequence is not only more power but more power per square metre, which means the load cannot be thinned out across cheap land to match whatever the local grid can deliver. It wants to be concentrated, and it wants to be concentrated where fibre, water and a permitting regime already exist.

Those places are increasingly the places where the transmission system cannot deliver on the timeline requested. That is the origin of the co-location idea: if the queue is the binding constraint, put the load next to the generator and step around the queue entirely. It is an elegant answer, and as Section 06 sets out, it is the part of the structure that turned out to be least settled.

None of this yet says anything about small modular reactors. It says that the load is large, continuous, dense, concentrated, and in a hurry — and that a buyer with those characteristics will pay a premium for firm carbon-free capacity sited beside it. Whether that premium can be captured by a technology that has not yet been built in the West is the question the rest of this note addresses.

EXHIBIT 03

What makes this load different from the last one

Three characteristics that distinguish data-centre load growth from the industrial load growth utilities are organised around.

~100%

Load factor

Effectively continuous demand. Unlike industrial load, it does not shape to the day, the shift or the season, so it cannot be served unfirmed.

~120 kW

Draw per cabinet, top of 2024 range

Against 10–15 kW for a conventional enterprise rack. The load is denser as well as larger, which concentrates rather than distributes the siting problem.

4.4 → 12%

US share, 2023 to 2028 upper bound

Berkeley Lab's December range. The upper bound is close to double the lower, on a five-year horizon, from a measured base.

Source: Lawrence Berkeley National Laboratory, December 2024;[3] vendor specifications for 2024 high-density rack architectures.[20] The load-factor figure is a characterisation of typical hyperscale operating practice rather than a published point estimate, and varies materially by facility and workload mix.

WHAT WE TAKE FORWARD

The case for firm, dispatchable, carbon-free capacity sited beside the load is sound, and it does not depend on resolving the forecast spread. What depends on the spread is how much capacity, and by when — so we carry the range rather than a point. Throughout the rest of this note we keep the energy question and the capacity question separate, because the technologies that answer them are not the same technologies.

02

THE CONTRACTS

Ninety days in autumn

Four buyers, six agreements, and a set of delivery dates that do not say what the headlines said they said.

The sequence is worth setting out in order, because the order is part of what made it feel like a regime change. On 20 September Constellation announced it would restart Three Mile Island Unit 1 under a twenty-year power purchase agreement with Microsoft.[4] On 30 September the Department of Energy closed a $1.52bn loan to support the restart of Palisades.[5] On 14 October Google signed with Kairos Power.[6] On 16 October Amazon announced three things at once: an anchor investment in X-energy's Series C-1, a project with Energy Northwest in Washington State, and an agreement to explore development at Dominion's North Anna site.[7] On 3 December Meta issued a request for proposals for one to four gigawatts of new nuclear capacity.[8]

Read as a set, these are not the same kind of instrument. A twenty-year power purchase agreement against an existing licensed reactor is a firm commercial contract with a known asset and a datable restart. A master plant development agreement is a framework for building something that does not yet exist at a site not yet chosen. An equity cheque is an option. A request for proposals is an invitation to bid. Collapsing all four into "the hyperscalers are buying nuclear" loses the only variable that matters for the current capital cycle, which is when the electrons show up.

It is also worth being clear about what did not happen. No hyperscaler placed a reactor order. None signed a construction contract. None took delivery risk on a plant. The largest capital commitment disclosed in the sequence was an equity cheque into a development company, and the largest capacity number was an invitation to bid. This is not a criticism — it is the rational way to buy an option on a technology whose cost you cannot yet observe — but it is a different thing from procurement, and the distinction disappeared almost entirely in how the sequence was reported.

EXHIBIT 04

The autumn contracts, sorted by first electron

Every publicly announced hyperscaler nuclear transaction of 2024, with the instrument type and the earliest date at which the capacity is targeted to deliver.

DATE	BUYER	COUNTERPARTY & ASSET	CAPACITY	INSTRUMENT	FIRST ELECTRON
4 Mar	Amazon	**Talen Energy** Cumulus campus, Susquehanna, PA	960 MW	Asset purchase & behind-the-meter supply	● Operating
20 Sep	Microsoft	**Constellation** Crane Clean Energy Center (TMI-1), PA	835 MW	20-year power purchase agreement	● 2028 target
30 Sep	—	**Holtec / US DOE** Palisades restart, MI · $1.52bn loan	800 MW	Federal loan; no hyperscaler offtake disclosed	● 2025 target
14 Oct	Google	**Kairos Power** Fleet agreement; sites not announced	500 MW	Master plant development agreement	● 2030 first unit
16 Oct	Amazon	**X-energy / Energy Northwest** Richland, WA · four modules, to twelve	320 MW to 960 MW	Development agreement plus Series C-1 anchor	● Early 2030s
16 Oct	Amazon	**Dominion Energy** North Anna, VA · technology not selected	~300 MW	Agreement to explore development	● Not established
3 Dec	Meta	**Open to developers** Request for proposals; responses due Feb 2025	1–4 GW	Request for proposals	● Not established

Source: company announcements and US Department of Energy releases, March–December 2024.[4–8,13] The 960 MW at Susquehanna is subject to the amended interconnection service agreement rejected by FERC on 1 November 2024 (Section 06); the underlying reactors operate, but the contracted supply arrangement does not currently stand. Capacity figures are as announced and are gross where disclosed.

Sorted this way, the picture is unambiguous. The capacity with a delivery date inside this decade is existing large light-water plant being restarted or re-contracted. The small-modular content — Kairos, X-energy — targets the early 2030s at the earliest. The largest single number in the table, Meta's one-to-four gigawatts, attaches to no technology, no site and no licence.

The strongest argument against reading this as we do is worth stating at full force, because we think it is partly right. Nuclear's historic failure mode in the West has not been engineering. It has been the allocation of financing risk. Projects are capital-intensive, long-dated, and exposed to construction overrun in markets where no counterparty will accept that exposure at a price the developer can afford. The Carbon Free Power Project, discussed in Section 04, died precisely there.

What appeared this autumn is the counterparty the sector has lacked for forty years: a buyer that is investment-grade, price-insensitive relative to its own gross margins, contractually committed to carbon-free power, and in a hurry. That materially changes the cost of capital available to a developer, and it changes it now rather than in 2030. Anyone who says the autumn agreements do not matter has not understood what they fix.

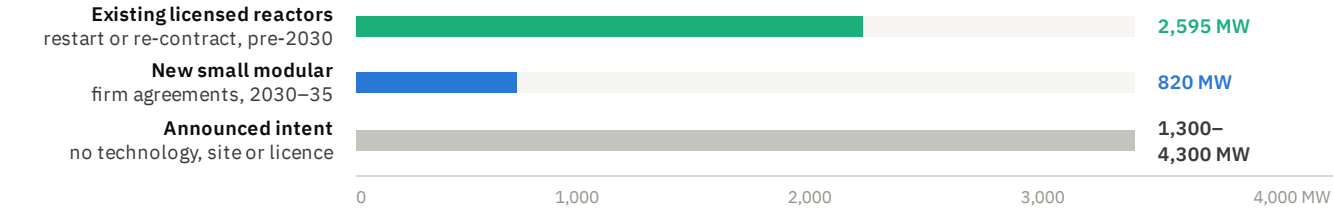
The Amazon structure deserves particular attention because it is the most integrated of the set. An equity position in the developer, a utility partner holding the licence, a site with an existing nuclear workforce, and — through TRISO-X — a fuel fabrication route inside the same corporate perimeter.[29,39] Vertical integration is the standard answer to a supply chain that does not yet exist, and it is the answer a buyer reaches for when it does not believe the market will deliver the inputs in time. That is a considered position and it should be read as one.

But there is a limit to what an offtake agreement can do, and it is exactly the limit that bears on this note. A creditworthy offtaker compresses the *financing* risk. It does not compress the *schedule*. Licensing, first-of-a-kind engineering, long-lead forgings, fuel qualification and commissioning consume calendar time that no amount of balance sheet shortens. The AI capital cycle is constrained by the schedule, not by the financing. The two problems have different solutions and the autumn agreements solve the other one.

EXHIBIT 05

Most of the contracted capacity is not new, and most of the new capacity is not near

Announced 2024 capacity, grouped by asset status and target delivery window.



Source: AI Lab Monaco, aggregated from Exhibit 04. Existing-reactor total comprises Susquehanna (960 MW, subject to the rejected interconnection agreement), Three Mile Island Unit 1 (835 MW) and Palisades (800 MW). New small-modular total comprises the Kairos fleet agreement (500 MW by 2035) and the initial X-energy deployment (320 MW). Announced intent comprises the Meta request for proposals (1–4 GW) and the Dominion exploratory agreement (~300 MW). **Grouping is ours.**

WHAT WE TAKE FORWARD

The autumn agreements matter more for what they do to nuclear's cost of capital than for what they do to the 2027 supply stack. A creditworthy, price-insensitive, carbon-motivated offtaker is the thing the sector has lacked for four decades, and it now exists. That is a genuine structural change and we do not minimise it. It is not the same thing as capacity arriving in time for this compute cycle, and a position should not conflate them.

03

THE RECORD

What has actually been built

There are two operating small modular reactors in the world. Neither is in a Western country, and neither was built to a disclosed commercial cost target.

The first analytical error in this sector is treating "small modular reactor" as a technology. It is a size class, and it contains at least two populations with very different risk. The first is miniaturised light-water plant — NuScale, the BWRX-300, Rolls-Royce's design — which are scaled-down versions of the physics that runs the existing global fleet. The second is Generation IV: high-temperature gas-cooled, sodium fast, molten salt, fluoride salt. These are not miniaturisations of anything that has operated commercially in the West. Kairos, X-energy, TerraPower and Oklo are all in the second population.

The terminology also misleads. The IAEA's threshold for "small" is 300 megawatts electric, which makes a small modular reactor roughly the size of a mid-scale gas plant rather than something that fits in a shipping container.[14] And "small" is not the operative word: the economic claim rests entirely on "modular", meaning factory-built, repeated, transportable units. A small reactor built bespoke on site captures none of the thesis and inherits all of the per-kilowatt cost penalty of losing scale economies. Size is a consequence of the strategy, not the strategy.

The IAEA catalogues more than seventy designs under development.[14] The overwhelming majority will never be built. That fact alone should discipline how much weight any single developer's timeline is given, and it argues for looking at the layers of the stack that are indifferent to which design wins — fuel, enrichment, forgings — rather than picking among developers. Section 05 explains why that layer has a problem of its own.

EXHIBIT 06

The operating small-modular fleet is two units, neither of them Western

Operating, under construction, and in licensing. Status as at December 2024.

UNIT	COUNTRY	TECHNOLOGY	CAPACITY	STATUS
Akademik Lomonosov Pevek · 2 × KLT-40S	Russia	Floating pressurised water	70 MWe	**Operating** Commercial operation May 2020
HTR-PM, Shidaowan Shandong · 2 modules, 1 turbine	China	High-temperature gas-cooled	210 MWe	**Operating** Commercial operation Dec 2023
Linglong One (ACP100) Hainan	China	Integral pressurised water	125 MWe	Under construction First concrete July 2021
CAREM-25 Atucha	Argentina	Integral pressurised water	32 MWe	Construction stalled Under way since 2014
BWRX-300, Darlington Ontario Power Generation	Canada	Boiling water	300 MWe	Site preparation Construction licence pending
Natrium, Kemmerer TerraPower	United States	Sodium fast with storage	345 MWe	Non-nuclear works Permit application filed Mar 2024
NuScale US460 6 × 77 MWe module	United States	Integral pressurised water	462 MWe	Design under review 50 MWe variant certified Jan 2023
Aurora, Idaho Oklo	United States	Sodium fast microreactor	~15 MWe	Pre-application Prior application denied Jan 2022

Source: IAEA, *Advances in Small Modular Reactor Technology Developments*, 2024;[14] World Nuclear Association;[15] company and regulator filings.[16–19,21,22] Selection is ours and is not exhaustive: it covers the units that are operating, in construction, or material to the Western investment case. Capacity is gross electrical output as published.

The two operating units deserve more respect than they usually get in Western commentary. HTR-PM is a serious engineering achievement: a pebble-bed high-temperature gas-cooled reactor in commercial operation, demonstrating passive safety characteristics at commercial scale, which is exactly what the Generation IV case has been promising for thirty years.[16] Akademik Lomonosov demonstrates that marine-derived pressurised water modules can be factory-built, floated to site and operated.[17] Both prove feasibility.

Neither proves economics. Neither was procured under competitive commercial discipline; neither publishes an auditable capital cost per kilowatt that a Western developer could underwrite against. The honest reading is that the physics and the manufacturing are demonstrated, and the cost is not observed.

It is not an accident that both are where they are. Both were built inside systems with state-directed capital, no commercial hurdle rate applied to the project, a vertically integrated domestic supply chain, and a regulator inside the same state apparatus as the builder. Those are the conditions under which a first-of-a-kind nuclear plant gets finished. None of them is available to a Western developer, and a Western programme must therefore substitute something else — which, on the evidence of Section 02, is the hyperscaler balance sheet. Whether that substitution works is the open question of the decade for this sector.

What the record does not license

It is tempting to run the argument further than the evidence allows. The absence of a Western small modular reactor does not demonstrate that one cannot be built on budget. It demonstrates that the cost distribution is *unobserved*. An unobserved distribution is not the same as a bad one. But it is an unpriced one, and in a market where the narrative is strong and the comparables are absent, unpriced risk is more often underpriced than over. That asymmetry, rather than any technical pessimism, is what drives our caution.

The second thing the record does not license is aggregation. A light-water module and a sodium fast reactor face different licensing paths, different fuel supply chains, different qualification programmes and different failure modes. NuScale's design has been certified by the Nuclear Regulatory Commission; Oklo's first application was denied on completeness grounds in January 2022.[22] Those two facts describe genuinely different positions in the regulatory queue, and a basket that holds both as "SMR exposure" is not expressing a view — it is declining to have one.

A third point, easily missed, concerns which population the hyperscalers actually chose. Google went to Kairos, Amazon to X-energy: both Generation IV, both without a commercially operating Western precedent, both requiring fuel that the West does not yet produce at scale. The buyers with the most sophisticated technical diligence functions in the world selected the higher-technical-risk population. That is either a signal that the light-water route's cost problem is considered unsolvable, or a signal that these are cheap options rather than considered procurement decisions. We do not think the evidence yet distinguishes between those two readings, and we would not pretend otherwise.

WHAT WE TAKE FORWARD

Feasibility is established; economics are not. We treat "small modular reactor" as a category containing at least two distinct risk profiles — miniaturised light-water plant with a certified design and a known fuel cycle, and Generation IV designs with no commercially operating Western precedent and an immature fuel supply — and we decline to hold them as a single exposure. Where a claim in this note depends on which population is meant, we say which.

04

COST

The cost hypothesis

Modularity is an argument about learning rates. The West's two most recent data points argue the other way, and one of them was a small modular reactor.

State the modularity case at its strongest, because it is a good case. Nuclear's cost problem in the West is a construction problem, not a physics problem. Bespoke, site-built, first-of-a-kind megaprojects have terrible learning curves: every unit is a prototype, craft labour is assembled and dispersed, and rework compounds. Factory fabrication of repeated identical units moves the work into an environment where learning accumulates — shipyard economics rather than construction economics. If that transfer can be made, the learning rate should look like manufacturing, and manufacturing learning rates are the reason solar modules cost what they now cost.

The difficulty is that the argument requires volume, and volume requires the cost the volume is supposed to produce. A first-of-a-kind unit carries the full burden of engineering, licensing and tooling amortisation. Nobody orders the fleet that justifies the factory until somebody demonstrates the cost that only the factory can deliver. This circularity is the real obstacle, and it is commercial rather than technical.

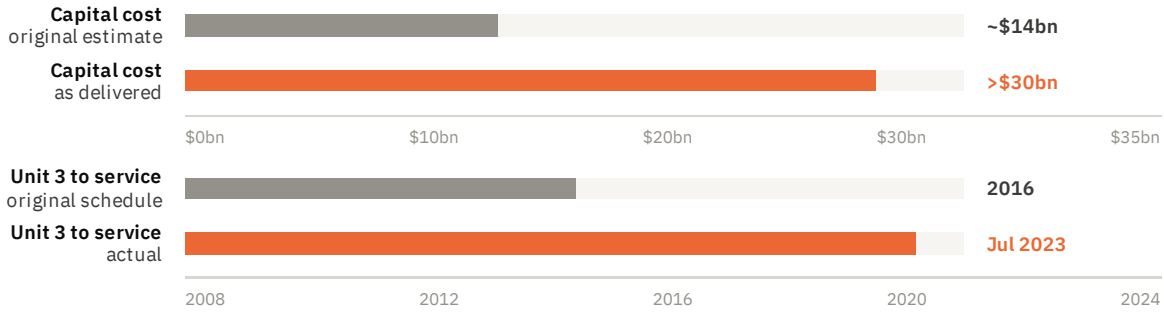
Europe supplies two more observations and they run the same way. Olkiluoto 3 in Finland was ordered in 2005 for delivery in 2009 and entered regular production in April 2023.[38] Flamanville 3 in France began construction in 2007 against a 2012 target; it loaded fuel in May 2024 and reached first criticality in September, at a cost several times the original estimate.[37] Three Western first-of-a-kind reactor projects, three overruns between two and four times, and slippage measured in years rather than months. That is a pattern, not a run of bad contractors.

The modular reply is that these were precisely the wrong kind of project — large, bespoke, site-assembled, each one a prototype — and that not doing this is the entire proposal. The reply is coherent, and we think it is the right diagnosis. It is also, as at December 2024, wholly prospective: no Western programme has yet produced the second identical unit that would begin to demonstrate it.

EXHIBIT 07

Vogtle is the benchmark modularity has to beat

Vogtle Units 3 and 4, the most recent Western nuclear construction: original estimate against delivered outcome, on cost and on schedule.



Source: Georgia Power and Southern Company disclosures;[9,10] Unit 4 entered commercial operation in April 2024. Delivered cost is stated as a floor: reported totals vary with the treatment of the Westinghouse bankruptcy, owner costs and financing, and we have not attempted to reconcile them. Vogtle is a large AP1000 build rather than a small modular reactor; it is offered as the relevant Western construction precedent, not as a like-for-like comparator.

Vogtle is the general precedent. The specific one is better, because it was an actual small modular reactor project, and because it failed in a way that is diagnostic.

The Carbon Free Power Project would have built a NuScale plant in Idaho for a consortium of Utah municipal utilities. It had what the sector says it needs: a certified design — the first small modular design certified by the Nuclear Regulatory Commission[23] — federal cost-share, a site, and a willing buyer. In 2021 the target power price was $58 per megawatt hour. In January 2023 it was revised to $89, an increase of just over half.[24] In November 2023 the project was terminated because subscription had not reached the level needed to proceed.[25]

Note what did not kill it. Not physics, not licensing, not public opposition. The design was certified. The project died because the price moved and the buyers had the right to walk. That is the cleanest available test of the commercial hypothesis in the West, and the hypothesis failed it.

What this does not license

One cancelled project does not establish a cost floor, and the conditions were unusually hostile: an inflation shock in exactly the input categories that dominate nuclear construction, a rapid rise in the discount rate applied to a thirty-year asset, and — decisively — an offtaker consortium of small municipal utilities with limited balance sheets and a contractual exit. Change the counterparty and the outcome may change with it.

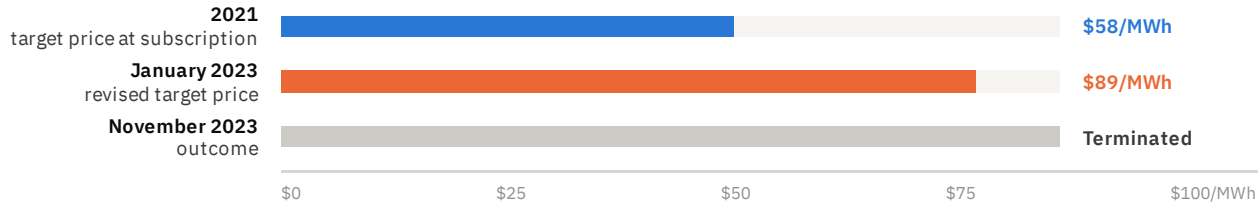
Which is precisely what Section 02 described happening this autumn. The Carbon Free Power Project's failure mode was counterparty fragility, and the counterparty problem is the one the hyperscaler agreements address. This is the strongest argument in favour of the sector available in December 2024, and we hold it seriously. It does not, however, touch first-of-a-kind engineering cost — the number that moved from $58 to $89 — and that number is set by the supply chain, the labour market and the regulator, none of which have a hyperscaler counterparty.

One further asymmetry deserves recording. A cost overrun on a conventional plant is absorbed by a regulated utility and, ultimately, a rate base. A cost overrun on a plant contracted to a technology company under a fixed-price agreement lands on the developer, who in every current case is a venture-funded company without the balance sheet to absorb it. Until the first fleet contract is signed we will not know which party has agreed to hold that risk, and the identity of that party is more informative about the project's likelihood of completion than any engineering milestone.

EXHIBIT 08

The first Western small-modular project died on price, not on physics

Carbon Free Power Project: target power price, and the outcome.



Source: NuScale Power and Utah Associated Municipal Power Systems disclosures, 2021–2023.[24,25] Target prices are the levelised figures published for subscription purposes and are net of the Department of Energy cost-share award; they are not directly comparable with market power prices.

WHAT WE TAKE FORWARD

Modularity is a hypothesis about learning rates that needs volume to test, and volume needs the demonstrated cost that only volume can produce. The autumn offtake agreements attack the demand side of that circularity, and that is real progress. They do nothing about first-of-a-kind engineering cost — which is the number that moved from $58 to $89 and closed the last Western attempt.

05

CONSTRAINTS

Three physical constraints

Enriched fuel, licensing throughput, and a supply base that was allowed to atrophy. None of the three is solved by capital, and the first of them lands hardest on exactly the designs the hyperscalers chose.

Fuel. Most Generation IV designs require high-assay low-enriched uranium — HALEU — enriched above the roughly 5% ceiling of the existing commercial fuel cycle and below the 20% weapons-relevant threshold. Kairos, X-energy, TerraPower and Oklo all need it. Until recently the only commercial supplier at scale was Russia. The Prohibiting Russian Uranium Imports Act, signed on 13 May 2024, bans Russian enriched uranium imports with waivers available only through 2027.[27]

Domestic supply is at demonstration scale. Centrus completed the first phase of its Department of Energy contract at Piketon, Ohio with a 20-kilogram delivery, and the second phase is contracted at a rate of 900 kilograms a year.[11] The Department has estimated that more than 40 tonnes a year could be required by 2030.[12] The gap is roughly two orders of magnitude, and it is not a gap a purchase order closes: enrichment is capital equipment with multi-year lead times and its own licensing pathway.

The timing of the waiver regime deserves attention on its own. Russian enriched uranium may be imported under waiver only through 2027.[27] That is precisely the window in which the first advanced-reactor cores need to be fabricated, qualified and loaded. The constraint therefore tightens rather than eases across exactly the period the schedules in Exhibit 04 depend on, and it does so by operation of statute rather than by market outcome.

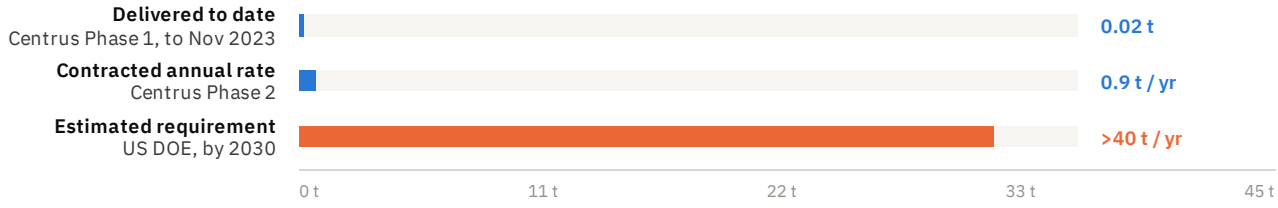
The nuance matters more than the headline. The light-water designs — the BWRX-300, NuScale, Rolls-Royce — run on conventional low-enriched uranium and are not exposed to this constraint. The fuel bottleneck therefore binds the Generation IV population and spares the population with the certified design. Google and Amazon both selected from the constrained side.

Enrichment is also only half of it. Enriched uranium has to be fabricated into fuel, and the fuel forms these designs use are themselves novel: TRISO particles for the gas-cooled reactors, metallic and salt-based forms elsewhere. Fabrication capacity is being built — X-energy's TRISO-X facility at Oak Ridge is the furthest advanced[39] — but a fuel form must also be qualified by the regulator, which is a multi-year programme of irradiation testing that runs on its own clock and cannot be accelerated with money. A reactor cannot start without a qualified first core, and the first core is on the critical path of every schedule in Exhibit 04.

EXHIBIT 09

Enrichment capacity is roughly two orders of magnitude short

United States high-assay low-enriched uranium: delivered to date, contracted annual rate, and the agency estimate of annual requirement by 2030. Common linear scale.



Source: Centrus Energy disclosures, 2023–2024;[11] US Department of Energy HALEU Availability Program materials.[12] The first two bars are rendered at a minimum visible width and are not to scale against the third; their true widths on this axis would be 0.04% and 2%. The 2030 requirement is an agency projection covering the whole advanced-reactor programme, not a firm order book, and should be read as an indication of order of magnitude.

Licensing throughput. The Nuclear Regulatory Commission has never licensed a commercial Generation IV power reactor. The ADVANCE Act, signed on 9 July 2024, reduces application fees for advanced reactor applicants and directs the Commission toward faster review — a genuine and constructive change.[26] But the binding variable is not fee policy; it is review capacity against novel physics. Each first-of-a-kind design consumes staff time disproportionate to its size, because the review must construct the regulatory precedent as it goes rather than apply one.

The observable to watch is TerraPower's construction permit application, filed in March 2024 — the first for a commercial advanced reactor in the United States.[19] Its duration will set the market's expectation for everything behind it. The failure mode is visible in Oklo's January 2022 experience: the application was not refused on safety grounds but denied for insufficient information,[22] which is the outcome that costs years rather than the outcome that ends a programme.

There is a structural fix in progress. The Commission is developing a risk-informed, technology-inclusive licensing framework — Part 53 — intended to license advanced reactors on their actual risk profile rather than by analogy to light-water rules written for a different machine.[35] If it works it is the single most important regulatory development in the sector. It is also not yet finalised, and the designs now in the queue are being reviewed under the old framework while the new one is written.

Supply chain and labour. Three decades with almost no Western orders hollowed out the base that builds reactors: heavy forgings, code-qualified components, nuclear-grade welders, and the quality systems that certify them. A significant share of the Vogtle overrun is attributable to rebuilding that capability inside a live project.[9]

Labour is the part of this least amenable to capital. Nuclear construction at scale requires thousands of code-qualified craft workers — welders, pipefitters, electricians certified to nuclear quality standards — and Vogtle at peak employed craft labour in the thousands on a single site.[9] Those certifications take years to obtain and lapse without work. A programme that intends to build a fleet needs a training pipeline running well before the first order, and no such pipeline is currently running at fleet scale in the West.

This is the constraint on which modularity has the best case. Smaller components mean more qualified suppliers, shorter forgings, transportable assemblies and less site craft labour. If any part of the modular thesis is going to work first, it is this part. The difficulty is the same circularity as Section 04: the supply base deepens with volume, and volume waits on the cost that a deeper supply base would deliver.

What the constraints do not license

None of the three is permanent, and it would be wrong to read them as a ceiling. Enrichment capacity is being built; the licensing precedent is being constructed now; supply bases recover when order books appear. The claim here is narrower and it is about time. Each of these constraints resolves on a multi-year clock, they resolve partly in sequence rather than in parallel, and none of them resolves inside the window in which the current compute capital cycle is being committed.

WHAT WE TAKE FORWARD

Fuel is the hard constraint and it binds the Generation IV designs specifically — which is to say, the ones Google and Amazon selected. Licensing is the uncertain constraint, and TerraPower's permit review is the observable that will resolve it. Supply chain is the constraint modularity is most likely to ease, and the one that eases last, because it needs the order book that the other two delay.

06

REGULATION

Behind the meter

The structure that makes reactor adjacency valuable was rejected in its first serious test, six weeks ago, and the market has largely carried on as though it were settled.

Ask why a data centre would want to sit beside a reactor rather than simply buy its output. Four reasons, in descending order of importance. It bypasses the interconnection queue, which in the constrained regions is measured in years. It avoids transmission and distribution charges on the largest cost line a very large customer faces. It permits a directly negotiated price rather than a market-clearing one. And it supports the cleanest possible version of an hourly-matched carbon-free claim, because the electrons are demonstrably the reactor's.

The first of those reasons is doing most of the work. At the end of 2023 roughly 2,600 gigawatts of generation and storage capacity sat in United States interconnection queues, and the typical project completing the process had spent about five years in it.[40] For a buyer whose competitive position depends on standing up capacity within eighteen months, a five-year queue is not a cost — it is a disqualification. Co-location is the only structure that removes it outright.

In March, Amazon bought Talen Energy's Cumulus data centre campus adjacent to the Susquehanna nuclear station for $650m, with an arrangement to take up to 960 MW behind the meter.[28] Because that changes how the plant relates to the transmission system, it required an amended interconnection service agreement, filed with the Federal Energy Regulatory Commission. Transmission-owning utilities objected, on the argument that a very large co-located load continues to rely on the grid for backup and reliability services while avoiding the cost allocation that funds them — a cost shift onto other ratepayers.

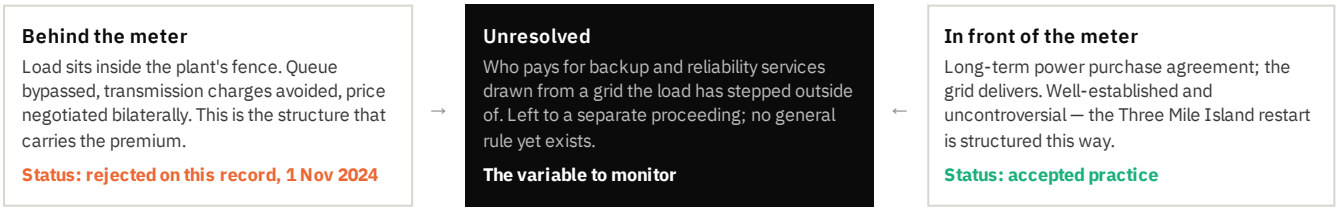
The objection is not frivolous, and it is worth stating in its strongest form. A co-located load does not truly leave the grid. It relies on the transmission system for backup when the reactor is down for refuelling or an unplanned outage, and it relies on system-wide reliability services it no longer contributes to funding. If the largest and most creditworthy customers can contract around network cost allocation, the fixed costs of the network fall on a smaller and poorer remaining base. That is a distributional argument, and distributional arguments tend to win in regulatory proceedings.

On 1 November the Commission rejected the amended agreement.[13] It is important to be precise about what that decision was and was not. It was not a prohibition on co-location. The Commission found that the filing parties had not justified the specific arrangement on the record before it, and left the general questions to be worked out elsewhere. It is a procedural setback, not a doctrinal one.

EXHIBIT 10

Behind the meter or in front of it: what the November order left open

The two structures available for pairing a large load with a nuclear plant, and their status as at December 2024.



Source: FERC order of 1 November 2024 in the amended interconnection service agreement proceeding concerning the Susquehanna station;[13] Constellation disclosures.[4] Characterisation of the two structures and their commercial consequences is ours. The order was decided by a divided Commission and its reasoning is specific to the record presented.

Why we think this is the most under-priced item in the sector at the end of 2024: the autumn announcements were received as commercial news, and this order is structural news. If behind-the-meter service cannot be arranged cleanly, the fallback is a front-of-meter power purchase agreement — the structure Constellation and Microsoft used at Three Mile Island. That is a perfectly sound instrument, and it delivers carbon-free energy under a long-term contract. What it does not do is bypass the queue or avoid the transmission cost allocation. It removes precisely the advantages that made physical adjacency worth paying for.

Put differently: strip out co-location and a reactor becomes a supplier of firm carbon-free energy under contract, which is valuable, rather than a way of obtaining interconnected capacity years earlier than the queue allows, which is the thing the compute cycle actually needs. The premium and the urgency both attach to the second description.

The scarcity is already visible in prices that have nothing to do with reactors. PJM's capacity auction for the 2025/26 delivery year, held in July 2024, cleared at $269.92 per megawatt-day against $28.92 for the previous year — an increase of roughly nine times in the largest wholesale market in the United States.[33] A capacity auction prices the option to be available, which is exactly the attribute this load requires and exactly what an intermittent resource cannot sell. The market has already repriced firmness. What it has not settled is who may contract for it directly, and on what terms.

The political economy runs the same way. A framework that visibly shifts network costs from the largest and most profitable customers onto residential ratepayers in the states hosting the build-out is not a stable equilibrium, whatever its efficiency merits. The likely landing zone is therefore neither prohibition nor unrestricted bypass, but a defined tariff for co-located load — which would resolve the uncertainty while removing part of the economic advantage that made co-location attractive in the first place.

There is a second-order point worth recording. The largest single capacity number announced by any hyperscaler in 2024 — the 960 MW at Susquehanna — currently rests on an agreement that does not stand. Any aggregate of "hyperscaler nuclear capacity" that includes it is counting something contingent, and most published aggregates do.

What this does not license

We are not forecasting that co-location will be prohibited, and we do not think that is the likely outcome. The economic logic is strong, the political pressure runs toward accommodation, and the Commission expressly declined to decide the general question. A subsequent order that establishes a workable framework — most plausibly one that assigns co-located load a defined share of transmission and reliability costs — would convert this from a headwind into a tailwind, and would do so quickly. Our claim is only that the structure is currently unresolved while being widely treated as settled, and that the asymmetry between those two states is not in the price.

WHAT WE TAKE FORWARD

Co-location is what makes a reactor worth more to a hyperscaler than the energy it produces, and co-location is not yet a settled structure. This is a genuinely two-sided variable with a plausible near-term resolution in either direction, which is unusual in this sector and is why we rank it first among the monitorables in Section 08. It is also the item most likely to move before any reactor is built.

07

IMPLEMENTATION

The vehicle problem

The listed complex is either pre-revenue with no reactor built, or diversified to the point where small modular reactors are immaterial. The developers the hyperscalers actually chose are private.

A structural view is only as good as the instrument available to express it, and this sector has an unusually poor one. The pure-play listed names have no revenue from reactors because they have not built any; their equity is an option on a development programme. The names that do generate revenue derive it overwhelmingly from other things — naval propulsion, enrichment services under legacy contracts, uranium mining, gas turbines — for which small modular reactors are a rounding error this decade.

Between those two groups sits the interesting case. Centrus is the sole domestic producer of high-assay low-enriched uranium, which Section 05 identified as the hardest physical constraint in the chain. That is a genuinely scarce position. It is also a small revenue line inside a business whose economics still turn on legacy separative-work contracts and on a Russian import waiver regime that cuts in both directions.

This shape is not unique to nuclear, and it is worth naming as a general phenomenon. When a theme arrives faster than the securities that express it, capital reaches for the nearest listed proxy, and the proxy re-rates on the theme rather than on its own cash flows. The re-rating is not irrational — it is what the available instrument set permits — but it decouples the security from the claim it is standing in for, and the decoupling is invisible until the claim is tested.

EXHIBIT 11

The purest exposure is not listed

Where small-modular exposure sits in public and private markets, and how much of each business it currently represents.

NAME	LISTING	WHAT THE BUSINESS IS	SMR REVENUE TODAY
NuScale Power	NYSE : SMR	Certified light-water design; Romanian project in engineering	None — pre-revenue
Oklo	NYSE : OKLO	Sodium fast microreactor developer; pre-application at the NRC	None — pre-revenue
Nano Nuclear Energy	Nasdaq : NNE	Early-stage microreactor developer	None — pre-revenue
Centrus Energy	NYSE : LEU	Enrichment; sole domestic HALEU producer	Small, contracted
BWX Technologies	NYSE : BWXT	Naval reactors, medical isotopes, nuclear components	Immaterial
GE Vernova	NYSE : GEV	Gas, wind, electrification; owner of the BWRX-300	Immaterial
Cameco	NYSE : CCJ	Uranium mining; minority interest in Westinghouse	Immaterial
Constellation · Vistra · Talen	CEG · VST · TLN	Existing merchant nuclear and thermal fleets	Not applicable
X-energy · TerraPower Kairos Power · Holtec	Private	The four developers hyperscalers and utilities actually contracted with	Not accessible

Source: company disclosures and public filings as at December 2024. Materiality assessments are ours and are qualitative. Securities are named to illustrate the structure of the opportunity set; **nothing here is a recommendation, and this note does not carry ratings or price targets** — see the disclosures on page 22.

The last row of Exhibit 11 is the finding. Google chose Kairos. Amazon chose X-energy. The Wyoming project is TerraPower's. The Palisades restart is Holtec's. All four are private, and none of them can be bought. The public market's small-modular pure-plays are, in each case, companies that the hyperscalers did not select.

There is a second, subtler point about what the listed market actually re-rated during 2024. The clearest beneficiaries of the AI power theme among liquid securities were not developers of future reactors but owners of existing ones. That revaluation ran on a different mechanism: higher forward power prices, scarcity value in firm dispatchable capacity, and a newly demonstrated willingness by very large buyers to pay above-market for carbon-free output from plant that already exists. That is the near-term AI power trade, and it is a different trade from the one this note is about.

Two further asymmetries are worth naming. First, the listed pure-plays are pre-revenue developers whose equity behaves like a long-dated option: convex to news, sensitive to the discount rate, and diluted by the equity issuance that funding a development programme requires. Second, the layer with genuine scarcity — enrichment and fuel fabrication — is largely either inside a diversified issuer or private, so the bottleneck this note identifies as the hardest is also among the hardest to own directly. Where a bottleneck cannot be bought, it usually gets expressed through whatever adjacent security trades, and that substitution is where the mispricing tends to accumulate.

We would put the implication plainly. Conviction on a structural claim does not license conviction on an instrument. Where the purest expression of a view is unavailable, the correct response is a smaller position held over a longer horizon — not a worse instrument held at the same size because it is the one that trades.

One practical consequence follows for anyone building exposure here. The layer of the stack most likely to be repriced first is not the developer and not the reactor — it is the constraint. Enrichment, fuel fabrication, heavy forgings and the qualified craft base all bind before any reactor is delivered, they bind regardless of which design wins, and they bind on a timetable set by physics rather than by narrative. That is a more defensible place to express the view than a bet on which of seventy designs gets built.

EXHIBIT 12

Placeholder — total return of the listed complex against the index

Reserved for a total-return comparison of the listed nuclear and small-modular cohort against a broad equity benchmark, year to date 2024.

EXHIBIT RESERVED — DATA NOT SOURCED

This exhibit is intentionally unpopulated. The series convention is to mark an exhibit as reserved rather than to publish figures we have not sourced from a pricing vendor.

To populate: total return, 31 December 2023 to the cover date, for (i) an equal-weighted basket of the listed names in Exhibit 11, (ii) the merchant nuclear operators shown separately, and (iii) the S&P 500 as reference. Separating (i) from (ii) is the point of the exhibit — it is the evidence that would test the decomposition claimed on page 20.

Source: to be populated from a pricing vendor prior to circulation. We note in Section 08 that the 2024 price action in this complex runs against the argument of this note, and that we are unable to decompose it fully with public data.

WHAT WE TAKE FORWARD

The instrument problem is not an argument against the structural view; it is a constraint on how the view can responsibly be held. The purest assets are private, the liquid proxies are either optionless or diluted, and the security that has actually captured the AI power theme so far is the existing reactor rather than the future one.

08

FALSIFICATION

Risks and falsification

Separating conviction on the structural claim from conviction on timing, and naming the observables that would settle either one.

We hold two claims at different levels of confidence, and they should be tested separately. **The structural claim** — that firm, carbon-free, dispatchable capacity sited near concentrated computational load is scarce and will command a premium — we hold with high confidence. It follows from the load characteristics in Section 01 and does not depend on any view about reactor vendors.

The timing claim — that small modular reactors will not supply material capacity to the compute cycle now being financed — we hold with moderate confidence, and it is the claim to monitor. It rests on construction and licensing durations that have historically been underestimated, which is a reason to trust it, and on the assumption that nothing changes the slope, which is a reason not to.

The distinction matters for how a position is held. A structural claim held with high confidence justifies a durable allocation that is not marked against quarterly news. A timing claim held with moderate confidence justifies a monitored position with pre-committed revision triggers. Conflating the two produces the characteristic failure of thematic investing: a correct structural view abandoned at the point of maximum drawdown because it was sized as though the timing were also known.

What follows is the list of triggers. We have tried to make each one a public, dated, binary-enough event rather than a direction of travel, because a monitorable that cannot embarrass us is not a monitorable.

EXHIBIT 13

What would confirm this argument, and what would break it

Named observables, the windows in which they fall, and the direction each would push the argument of this note.

MONITORABLE	WINDOW	WHAT IT WOULD TELL US	DIRECTION
FERC establishes a workable co-location framework	2025	Restores the structural premium on reactor adjacency; the fastest-moving variable in the note	Breaks §06
A hyperscaler signs a fixed-price EPC with completion guarantee	2025–26	Attaches a number to the cost hypothesis and names who bears the overrun	Breaks §04
NRC issues the Natrium construction permit	2025–26	First US advanced-reactor construction permit; sets the review-duration precedent	Breaks §05
Darlington discloses an auditable cost per kilowatt at final investment decision	2025	The first Western small-modular capital cost that can be underwritten against	Either
Contracted US HALEU output rises above roughly 5 tonnes a year	2026–27	Fuel constraint eases materially for the Generation IV population	Breaks §05
A light-water small modular project reaches final investment decision	2025–26	The certified-design population acquires a real order book	Breaks §04
Any first-of-a-kind target slips beyond 2032	Any	The historical slippage pattern reasserts and the phase gap widens	Confirms
A second small modular project is cancelled on price	Any	Two observations rather than one; the cost hypothesis loses its benefit of the doubt	Confirms
Data-centre load growth resolves toward the bottom of the range	2025–27	The premium on firm capacity compresses and urgency drains from the whole thesis	Confirms

Source: AI Lab Monaco. Windows are our estimates of when each observable is likely to become visible, not published schedules. "Breaks" indicates that the observation would require us to revise the named section; two such observations inside eighteen months would require revising the timing claim itself.

EVIDENCE CHECK · WHAT RUNS AGAINST THIS NOTE

The listed nuclear and small-modular complex has re-rated materially during 2024 — the year in which we argue the timing is wrong. That is the single largest piece of evidence against the argument set out here, and we record it as such rather than in a footnote. Our reading is that the great majority of that move is attributable to the revaluation of *existing* reactors and merchant generation, on the mechanism described on page 18, rather than to expectations of small-modular delivery. But we cannot fully separate the two effects with public data. Exhibit 12 is reserved for the decomposition that would test it. If that decomposition comes back the other way — if the pre-revenue developers, not the existing-fleet owners, carried the move on fundamentals rather than on flows — then the market is pricing something we have not identified, and this note is the thing that needs revising.

Three further ways we could be wrong

We may be underweighting what the offtaker changes. Section 04 describes a circularity: no fleet order without a demonstrated cost, no demonstrated cost without a fleet order. A hyperscaler with a trillion-dollar balance sheet and a strategic rather than financial motive could break that in a single transaction by accepting completion risk that no utility would accept. Nothing in the physics prevents it. If it happens, our timing claim degrades quickly, and it would most likely happen without warning.

We may be wrong about the appropriate discounting horizon. It is not irrational for equity to price 2035 cash flows in 2024 — that is what growth equity does, and the criticism that a market is "early" is the least reliable criticism in the profession. Our objection is narrower than it may read: it is to the conflation of a near-term power trade with a long-dated construction trade under one label, not to the existence of the second trade.

Schedules can compress when governments decide they should. Our base case extends the historical relationship between novel nuclear licensing and calendar time. That relationship has been broken before, in both directions, when a programme acquired national-security framing. There are visible reasons why advanced nuclear might acquire that framing in the next administration. We do not forecast it, and we would not be surprised by it.

The category may resolve rather than persist. We have insisted on separating light-water from Generation IV designs. It is possible that within eighteen months the market does that separation for us — that one population attracts the orders and the other does not — in which case the aggregate statements in this note become less useful than the disaggregated ones, and Exhibit 06 rather than Exhibit 05 becomes the operative table.

The one thing we would not revise

Nothing in the foreseeable evidence would change the judgment in Section 01 that the published range for data-centre load is too wide to trade as a point estimate. If anything, we expect that range to widen before it narrows, as efficiency gains and inference-workload migration become measurable. Positions sized against the top of the range should be understood as bets on the range as well as on the technology.

WHAT WE TAKE FORWARD

High conviction on the structural claim; moderate conviction on timing, monitored against nine named observables. Two of those observables — a workable co-location framework and a fixed-price construction contract — would between them require us to rewrite most of this note. Neither is unlikely. That is the honest shape of the position, and it argues for a small, patient exposure rather than a confident one.

REFERENCES

Every quantitative claim in this note is footnoted to a numbered source below. All sources predate the cover date; nothing published after 31 December 2024 has been used.

1 Electric Power Research Institute. *Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption.* May 2024.

2 Goldman Sachs Research. *Generational Growth: AI, Data Centers and the Coming US Power Demand Surge.* April 2024.

3 Shehabi, A. et al. *2024 United States Data Center Energy Usage Report.* Lawrence Berkeley National Laboratory, December 2024.

4 Constellation Energy. "Constellation to launch Crane Clean Energy Center, restoring jobs and carbon-free power to the grid." Press release, 20 September 2024.

5 US Department of Energy, Loan Programs Office. Loan guarantee to Holtec Palisades for restart of the Palisades Nuclear Plant. 30 September 2024.

6 Google. "New nuclear clean energy agreement with Kairos Power." 14 October 2024.

7 Amazon. "Amazon signs agreements for innovative nuclear energy projects." 16 October 2024.

8 Meta Platforms. Request for proposals for new nuclear generation capacity in the United States. 3 December 2024.

9 Georgia Power / Southern Company. Vogtle Unit 3 commercial operation announcement and associated cost disclosures. July 2023.

10 Southern Company. Vogtle Unit 4 commercial operation announcement. 29 April 2024.

11 Centrus Energy Corp. HALEU Operation, Piketon, Ohio: Phase 1 completion and Phase 2 contract disclosures. 2023–2024.

12 US Department of Energy. HALEU Availability Program: programme materials and requirement estimates. 2024.

13 Federal Energy Regulatory Commission. Order rejecting amended interconnection service agreement concerning the Susquehanna Steam Electric Station. 1 November 2024.

14 International Atomic Energy Agency. *Advances in Small Modular Reactor Technology Developments.* ARIS supplement, 2024 edition.

15 World Nuclear Association. "Small Nuclear Power Reactors." Information paper, updated 2024.

16 China National Nuclear Corporation; IAEA Power Reactor Information System. HTR-PM, Shidaowan: entry into commercial operation. December 2023.

17 Rosatom; IAEA Power Reactor Information System. Akademik Lomonosov floating power unit: commercial operation. May 2020.

18 Ontario Power Generation. Darlington New Nuclear Project: programme updates and regulatory status. 2024.

19 TerraPower. Natrium demonstration project, Kemmerer, Wyoming: construction permit application (March 2024) and commencement of non-nuclear construction (June 2024).

20 NVIDIA Corporation. Rack-scale architecture specifications for 2024 accelerated computing platforms.

21 China National Nuclear Corporation. Linglong One (ACP100) demonstration project, Hainan: construction status. 2021–2024.

22 US Nuclear Regulatory Commission. Denial of the Oklo Power combined licence application for the Aurora facility. 6 January 2022.

23 US Nuclear Regulatory Commission. NuScale small modular reactor design certification rule. *Federal Register*, January 2023.

24 NuScale Power and Utah Associated Municipal Power Systems. Revised Carbon Free Power Project target price disclosure. January 2023.

25 NuScale Power and Utah Associated Municipal Power Systems. Joint announcement of the termination of the Carbon Free Power Project. 8 November 2023.

26 Accelerating Deployment of Versatile, Advanced Nuclear for Clean Energy (ADVANCE) Act of 2024. Public Law 118-67, enacted 9 July 2024.

27 Prohibiting Russian Uranium Imports Act. Public Law 118-62, enacted 13 May 2024.

28 Talen Energy Corporation. Sale of the Cumulus data centre campus at Susquehanna to Amazon Web Services. 4 March 2024.

29 X-energy Reactor Company. Series C-1 financing round announcement. October 2024.

30 Oklo Inc. Completion of business combination with AltC Acquisition Corp and listing on the New York Stock Exchange. May 2024.

31 Declaration to Triple Nuclear Energy, COP28, Dubai. 2 December 2023.

32 Masanet, E., Shehabi, A., Lei, N., Smith, S. and Koomey, J. "Recalibrating global data center energy-use estimates." *Science*, vol. 367, 2020.

33 PJM Interconnection. 2025/2026 Base Residual Auction results. 30 July 2024.

34 Google. *24/7 Carbon-Free Energy by 2030*: commitment and methodology papers. 2020–2024.

35 US Nuclear Regulatory Commission. Part 53 risk-informed, technology-inclusive regulatory framework rulemaking, pursuant to the Nuclear Energy Innovation and Modernization Act, Public Law 115-439. 2019–2024.

36 Tennessee Valley Authority. Clinch River Nuclear Site early site permit (2019) and BWRX-300 programme authorisation (2024).

37 Électricité de France. Flamanville 3 EPR: fuel loading, May 2024; first criticality, September 2024.

38 Teollisuuden Voima Oyj. Olkiluoto 3 EPR: entry into regular electricity production, April 2023.

39 X-energy Reactor Company / TRISO-X. TRISO-X Fuel Fabrication Facility, Oak Ridge, Tennessee: licence application and construction status. 2022–2024.

40 Rand, J. et al. *Queued Up: Characteristics of Power Plants Seeking Transmission Interconnection.* Lawrence Berkeley National Laboratory, April 2024.

ON SOURCE SELECTION

We have preferred primary documents throughout — the company announcement rather than the report of it, the regulator's order rather than the summary, the statute rather than the commentary. Where an estimate could only be obtained from a research house we have named the house and, where two houses disagree materially, presented both rather than selecting between them. Three of the sources above are projections produced by parties with an interest in the outcome they project: references 12, 19 and 29. They are used for order of magnitude and are identified as such at the point of use.

SCOPE AND METHOD

This note assembles publicly available primary sources — company announcements, regulatory filings and orders, statute, and agency and laboratory publications — and reasons from them. It contains no proprietary data, no channel checks and no management access. Where a figure is our aggregation, characterisation or estimate rather than a published point value, the exhibit source note says so explicitly.

PUBLICATION HISTORY

Cover-dated December 2024, which is the date of the underlying research. Every source cited predates 31 December 2024, and the analysis has been constructed so that no information postdating the cover date enters it. Where an event's outcome was still open at the cover date — most importantly the co-location questions discussed in Section 06 — the note says that it was open rather than resolving it.

RESERVED EXHIBITS

Exhibit 12 is published as a labelled placeholder. The house convention is to mark an exhibit reserved rather than to publish figures that have not been sourced from a pricing vendor. It is identified in the contents list on page 3, and Section 08 states plainly that the evidence it would carry runs against the argument of this note.

PALETTE CONVENTION

The series binds four data colours to fixed meanings. This note's variables are physical and financial rather than model-architectural, so the bindings are extended rather than reassigned, preserving each slot's function: amber-orange for the compute layer and its demand; blue for the structurally-favoured-but-not-yet-dominant option, here firm new nuclear; green for the installed base, here existing reactors; grey for neutral or unestablished. Every bar carries a direct value label, which remains a requirement rather than a stylistic choice.

THIS IS NOT INVESTMENT RESEARCH

AI Lab Monaco is an independent research laboratory. It is not a broker-dealer, an investment adviser or a regulated financial institution, and this publication is not investment research within the meaning of any applicable regulation. It carries no ratings, no price targets, no earnings estimates and no recommendation to buy, sell or hold any security. Securities and issuers are named solely to illustrate the structure of the opportunity set discussed in Section 07; naming an issuer is not an assessment of its merits. Nothing here constitutes financial, legal, tax or accounting advice, and nothing here takes account of the objectives, financial situation or needs of any recipient. Readers should obtain independent professional advice before acting on anything in this document.

USE OF QUANTITIES

Capacity is stated as gross electrical output in megawatts unless otherwise noted, and as announced by the counterparty rather than adjusted by us. Currency is United States dollars in nominal terms of the year cited; we have not deflated historical construction costs to a common year, which understates the real severity of the older overruns discussed in Section 04. Dates are the date of public announcement, which for regulatory matters may postdate the decision itself.

FORWARD-LOOKING STATEMENTS

This note discusses construction schedules, licensing outcomes, fuel supply and regulatory decisions that had not occurred at the cover date. Statements about them are forward-looking and subject to substantial uncertainty. Nuclear construction schedules in particular have historically been unreliable, which is itself an argument made in the note. Past performance is not a guide to future performance. No representation or warranty, express or implied, is made as to the accuracy or completeness of information obtained from third-party sources.

VIEWS AND INTERESTS

Views are those of the author as at December 2024 and are subject to change without notice. AI Lab Monaco does not undertake to update this note. The author and the laboratory may hold interests in securities or private companies mentioned. No issuer named in this note has commissioned, reviewed, approved or paid for it, and no compensation has been received from any issuer in connection with its preparation.

INDEPENDENCE AND CORRECTIONS

No issuer, developer, utility or regulator named in this note was contacted in its preparation, and none has seen it in draft. Errors of fact will be corrected in a revised edition with the correction stated on this page rather than made silently, and the cover date will be retained with the revision noted separately.

TERMINOLOGY

"Small modular reactor" follows the IAEA convention of a reactor with electrical output up to 300 megawatts designed for factory fabrication of repeated units. We distinguish throughout between miniaturised light-water designs and Generation IV designs, because they carry materially different licensing, fuel and supply-chain risk and are frequently discussed as though they did not. "Firm capacity" means capacity available on demand irrespective of weather or time of day.

KNOWN LIMITATIONS

Three, stated plainly. We have not modelled levelised cost for any design, because the first-of-a-kind inputs are not observable and a model built on vendor targets would restate those targets with false precision. We have not decomposed 2024 equity performance in the sector, which is the reserved Exhibit 12, and Section 08 records that this is the principal gap in our evidence. And we have relied on public disclosure by parties with an interest in the timelines they publish; where a schedule is a vendor target we have labelled it as one.

DISTRIBUTION AND CITATION

Distributed to subscribers and clients of AI Lab Monaco. Not for distribution to any person or in any jurisdiction where such distribution would be contrary to law. May be quoted with attribution. Suggested form: AI Lab Monaco (2024), *The Reactor and the Rack: small modular reactors and the AI load curve*, Frontier Research, December 2024.

The Reactor and the Rack · Frontier Research · December 2024
Set in IBM Plex Sans and IBM Plex Serif. Cover artwork generated for this note.
© 2024 LF Artificial Intelligence Laboratory Monaco. All rights reserved.

Firm power is a schedule problem before it is a technology problem.



Frontier Research — AI Lab Monaco's note series on the technologies, constraints and capital flows shaping artificial intelligence. Notes are identified by title and date, and are written to be read independently.

LF Artificial Intelligence Laboratory Monaco
2 Avenue Princesse Grace · MC 98000 Monaco
info@artificialintelligencemonaco.com
artificialintelligencemonaco.com