1337
ADVISORS

The Evolution of Compute Markets and Future Compute Rights

A Market Structure Analysis of How Competition for Intelligence May Transform the Procurement, Allocation, and Reallocation of Future Compute Capacity

Mike Ossowski
1337 Advisors

Contents

Executive Summary

The competition for intelligence is changing how compute is acquired.

For most of the cloud era, compute was infrastructure: acquired when needed, consumed, released. The procurement horizon was short. Future access carried little economic weight.

Artificial intelligence changes that. As intelligence becomes a primary driver of competitive advantage, the resources required to produce it become strategically valuable. Compute sits at the center. Organizations are no longer competing only for customers, talent, and capital. They are competing for the ability to generate, deploy, and scale intelligence, and that competition increasingly plays out in advance.

The result is a structural shift in compute procurement. Organizations are reserving capacity months and years before consumption. Hyperscalers are committing hundreds of billions to GPU infrastructure before workloads are fully defined. Neoclouds emerged to serve organizations that could not access hyperscaler capacity at all. These are not demand signals. They are access signals: evidence that future availability itself has become valuable.

Once procurement and consumption separate by time, two things become unavoidable: uncertainty and misallocation.

Every future commitment is a forecast. Some organizations will reserve too much. Others too little. Model efficiency will improve. Workloads will change. Strategic priorities will shift. The gap between what was reserved and what is actually needed creates a problem that procurement mechanisms alone cannot solve. It is not a supply problem but an allocation problem. Scarcity and surplus can coexist. The market may contain unmet demand and underutilized capacity at the same time.

The challenge of reallocation is moving future access toward the participants who value it most. It is the market function that remains almost entirely underdeveloped.

This paper examines the economic forces creating demand for that function. It does not propose a specific market structure. It argues that as competition for intelligence intensifies, future access to compute will require mechanisms for discovery, valuation, and reallocation that extend well beyond bilateral procurement.

The central thesis: as compute becomes a strategic resource defined by future access rather than present consumption, compute markets will evolve toward infrastructure designed to establish, allocate, and reallocate claims on future capacity. Those claims, which this paper calls future compute rights, are not the premise of this analysis. They are its conclusion.

Part I · The Competition for Intelligence

Artificial intelligence is transforming intelligence into a scalable economic resource. Compute sits at the center of that transition.

---

1. Compute Has Become a Strategic Economic Input

The competition for intelligence is reshaping what it means to control compute.

For most of the cloud era, compute was a technology cost. Organizations consumed capacity when they needed it, released it when they didn't, and optimized spend against utilization. That framework made sense when compute functioned as a support layer: infrastructure behind products, not a primary input to competitive output.

Artificial intelligence changes the role of compute because it changes the role of intelligence itself.

Historically, intelligence inside an organization was embedded in people: analysts, engineers, lawyers, operators, executives. Scaling intelligence required hiring, training, and retaining additional talent. Intelligence scaled at the speed of people.

Artificial intelligence introduces a different architecture. Intelligence can be generated, deployed, and scaled through systems, embedded in software, distributed across workflows, and applied continuously. Once intelligence becomes producible at scale, the inputs required to produce it become strategically important.

Those inputs are specific:

- **Data:** intelligence must be trained, grounded, evaluated, and refined
- **Talent:** systems must be designed, deployed, monitored, and improved
- **Energy:** intelligence-producing infrastructure consumes significant power
- **Semiconductors:** they determine the performance characteristics of the computational layer

- **Networks and storage:** intelligence systems require movement, persistence, and access to information

Compute sits at the center of this system. It is the resource through which data, algorithms, energy, and hardware convert into usable output. A model architecture without compute remains unrealized. A dataset without compute remains underutilized. An AI strategy without compute remains aspirational.

This transition is already visible in market behavior. Technology companies are committing hundreds of billions to GPU infrastructure, datacenter capacity, and power agreements, not against fully defined workloads but against anticipated demand. Neoclouds like CoreWeave built businesses specifically because hyperscaler capacity was unavailable to organizations that needed it. GPU brokers and spot marketplaces emerged to serve demand that couldn't find a path through existing channels. These are market responses to a resource becoming strategically scarce.

Compute is a particularly strong candidate for market evolution because it possesses characteristics most strategic resources do not. It can be measured, allocated, contracted, reserved, and priced. It can be differentiated by hardware type, geography, networking, service level, and time horizon. It is specific enough to be meaningfully differentiated and comparable enough to be allocated at scale.

Talent cannot be standardized this way. Data resists cross-context comparison. Energy already has mature market structures in most regions. Compute sits at the intersection of technical specificity and economic comparability, and that combination creates the conditions for market structure to emerge.

Markets do not develop sophisticated infrastructure around resources that are simple, abundant, and easy to procure. They develop it when resources become scarce, strategically significant, forward-looking, and imperfectly allocated. Commodity markets, energy futures, freight derivatives, bandwidth exchanges: each emerged when bilateral procurement became insufficient to coordinate access to a resource that mattered enough.

Compute is entering that territory. The signal is not simply that demand is growing. The signal is that compute is increasingly being acquired for future use. Not because it is needed today, but because access tomorrow cannot be assumed.

That separation of procurement from consumption is the foundational shift. Once an organization reserves capacity months or years before it will be consumed, it is no longer simply buying compute. It is securing a claim on future access. And once claims on future access become economically meaningful, a different set of questions emerges:

- How should future capacity be allocated?
- How should reservations be valued?
- What happens when actual demand diverges from the forecast embedded in a commitment?
- Can excess capacity be reassigned?
- Who has the right to use what was reserved?

These are not technical questions. They are market structure questions, and compute markets are not yet equipped to answer them well.

Part II · The Emergence of Future Compute Markets

Future access is becoming increasingly valuable. As organizations commit further into the future, uncertainty becomes a defining market characteristic.

2. The Emergence of Future Compute Commitments

> KEY INSIGHT
>
> Organizations do not reserve compute years in advance because they want to. They reserve compute because they increasingly believe they cannot afford not to.

Most discussions frame compute reservations as procurement decisions. They are better understood as forecasts.

Every reservation contains an implicit prediction: about future workloads, future demand, future model architectures, future competitive conditions. The contractual structure varies across reserved instances, dedicated clusters, and long-term capacity agreements, but the economic substance is consistent. Beneath every future commitment sits an assumption about a future state of the world.

The reason organizations make these commitments is not confidence. It is the absence of a viable alternative.

When access to advanced compute cannot be assumed, the cost of being capacity-constrained may substantially exceed the cost of over-reservation. Delayed model training delays product timelines. Constrained inference capacity limits deployment scale. Falling behind on infrastructure creates competitive gaps that are difficult to close. In that environment, reserving capacity before demand is certain is not reckless. It is rational. The expected cost of insufficient capacity exceeds the expected cost of too much.

This calculus is already driving observable behavior. Organizations with no immediate training requirement are signing multi-year GPU commitments. Enterprises are reserving dedicated clusters against workloads that exist only in roadmaps. The scale of announced capital commitments makes the point directly.

EXHIBIT 1: SELECTED ANNOUNCED CAPITAL COMMITMENTS IN AI INFRASTRUCTURE

Organization	Location	Commitment	Notes
Amazon	Indiana	$15B	Data center expansion, 2026
Meta	Indiana	$10B	Large-scale GPU cluster
Microsoft	Wisconsin	$13B+	Multi-campus buildout
Meta	Tennessee	$10B	1,500 acres, 13 buildings
Amazon	Ohio	$10B+	Continued regional expansion
Microsoft	Quebec	$7.5B CAD	600 MW pipeline
OpenAI / SoftBank (Stargate)	Texas	$500B program	1.2 GW operational, 9 GW by 2029

Source: 1337 Advisors Platform · platform.1337advisors.com/infrastructure

> **Observation:** *These are forward commitments. The workloads that will consume this infrastructure do not yet fully exist. Organizations are not procuring against known demand. They are reserving access against anticipated scarcity, and doing so at a scale that makes the cost of being wrong very high.*

The procurement decision is no longer about satisfying known demand. It is about avoiding constraint when demand materializes.

Reservations are not evidence of confidence in forecasts. They are evidence that organizations no longer trust they can acquire capacity on demand.

The challenge is that every forecast is wrong. Technological progress does not follow roadmaps. Model efficiency improves in ways that alter infrastructure requirements. Funding environments change. Strategic priorities shift. The further a commitment extends into the future, the wider the gap between the assumptions embedded in it and what actually occurs.

This is not a planning failure. It is the unavoidable consequence of planning.

Consider the two organizations this dynamic reliably produces.

The first reserved insufficient capacity. Demand materialized faster than expected. New opportunities emerged. Infrastructure is now the constraint on growth.

The second reserved too much. Workloads failed to develop. Model efficiency improved. The organization controls capacity it no longer needs.

These look like opposite problems. Economically, they are the same problem from opposite sides of the market. One has compute. One needs compute. Both arrived there because the future diverged from the forecast that justified the original decision.

The market has become increasingly effective at helping organizations make forward commitments. It has invested almost nothing in helping them manage the consequences when those commitments prove wrong. Procurement solves for acquisition. It does not solve for the gap between what was acquired and what is actually needed. As commitments grow longer and larger, that gap becomes an increasingly important source of market friction.

---

3. Demand Risk and Forecast Uncertainty

> KEY INSIGHT
>
> Scarcity encourages earlier commitments. Earlier commitments increase uncertainty.

The separation of procurement from consumption carries a structural paradox.

Scarcity encourages earlier commitments. Earlier commitments increase uncertainty.

As advanced compute becomes strategically important, organizations become increasingly concerned about failing to secure capacity when needed. That concern pushes procurement decisions further into the future. But moving decisions further into the future increases the probability that original assumptions become incomplete or obsolete before consumption occurs. The market attempts to reduce access risk by committing earlier. In doing so, it increases forecast risk.

Time is the mechanism. A capacity decision made thirty days before use can be anchored to a workload that exists, a customer that is known, a requirement that is reasonably clear. A capacity decision made three years before use is a strategic commitment made under conditions that may change entirely: model architectures shift, hardware generations alter relative economics, products get cancelled, competitive dynamics change.

The longer the gap between procurement and consumption, the wider the range of possible divergence between assumption and outcome. That divergence is compounded by the fact that there is no single price for compute. The same GPU, across different providers, trades across a range wide enough to make long-range cost forecasting nearly impossible.

EXHIBIT 2: ON-DEMAND PRICE SPREAD BY GPU ACROSS TRACKED PROVIDERS

GPU	Architecture	Floor Price	Ceiling Price	Spread	Providers
H100 SXM 80GB	Hopper	$2.04/hr	$12.29/hr	502%	13
H200 SXM 141GB	Hopper	$2.29/hr	$10.84/hr	374%	11
A100 SXM4 80GB	Ampere	$1.36/hr	$5.12/hr	277%	11
B200 SXM 180GB	Blackwell	$3.95/hr	$14.00/hr	254%	7
AMD MI300X	CDNA3	$1.99/hr	$6.00/hr	202%	5

Source: 1337 Advisors Platform · platform.1337advisors.com/signals. Data as of June 2026.

> **Observation:** *A 502% spread on the H100 SXM means compute has no single price. An organization forecasting infrastructure costs three years out cannot anchor to a number. This dispersion reflects real differences in provider trust, reliability, operational overhead, and geographic cost structure. It also means that any future reallocation mechanism must first solve a valuation problem: when excess reserved capacity needs to move, what is it actually worth?*

Over-reservation and under-reservation carry asymmetric costs, and the asymmetry matters. Over-reservation creates inefficiency: excess spend and underutilized capacity. Under-reservation creates strategic constraint: missed deployment windows, delayed model development, and lost competitive position. When the downside of insufficient capacity exceeds the downside of excess capacity, rational organizations bias toward over-reservation.

That behavior should not be read as waste. In supply-constrained markets, excess reservation is a form of insurance. The organization is not predicting it will consume every unit it reserves. It is saying the cost of being short at the wrong moment is too high to tolerate. A utilization lens reads this as inefficiency. A strategic access lens reads it as risk management.

At the individual firm level, demand risk looks like a planning problem. At the market level, it becomes an allocation problem. When many organizations make forward commitments under uncertainty, the market produces structural mismatches. Some participants control more future capacity than they need. Others face shortfalls. The mechanisms to connect them are largely absent.

The question is not whether forecasts will be accurate. They will not be. The question is whether the market can respond efficiently when forecasts are wrong.

---

4. Scarcity, Optionality, and Reservation Value

> KEY INSIGHT
>
> A decision can look inefficient in hindsight and still have been rational at the time it was made.

Demand risk explains why future commitments are uncertain. Scarcity explains why organizations make them anyway.

If advanced compute were abundant and easily substitutable, reservations would have limited value beyond eventual consumption. Organizations could wait for demand to materialize and acquire capacity when needed. Scarcity changes that calculation. When capacity is constrained, future access itself acquires value. It protects not only because it enables consumption but because it guards against a future state in which access may be unavailable, more expensive, or controlled by competitors.

A reservation functions as an option: the right, not the obligation, to convert future access into actual consumption. Utilization alone does not capture that value.

A reservation that goes partially unused may still have been economically rational. It may have protected a product roadmap from an infrastructure shortage. It may have prevented a competitor from securing the same capacity. It may have given an organization time to make better decisions as uncertainty resolved. These are real economic benefits that no utilization metric can see.

This reframes how excess reserved capacity should be interpreted. It is not always evidence of poor planning. In supply-constrained markets, it is often the visible cost of preserving strategic optionality: the price of being able to act when compute became most valuable.

To understand what an organization is actually paying for when it reserves compute, it helps to decompose the cost directly.

EXHIBIT 3: H100 SXM ON-DEMAND COST DECOMPOSITION BY PROVIDER

Provider	Location	Power $/GPU·hr	Hardware $/GPU·hr	Rental $/GPU·hr	Residual
Hyperstack	Manchester, UK	$0.136	$0.951	$2.04	40%
Nebius	Mäntsälä, Finland	$0.039	$0.951	$2.15	47%
RunPod	US (multi-region)	$0.078	$0.951	$3.29	64%
DigitalOcean	New York, NY	$0.157	$0.951	$3.39	63%
Crusoe	Calhan, CO	$0.034	$0.951	$3.90	71%
Lambda Labs	Austin, TX	$0.070	$0.951	$3.99	71%
CoreWeave	Saddle Brook, NJ	$0.078	$0.951	$6.16	81%
OCI	Ashburn, VA	$0.073	$0.951	$10.75	89%
GCP	Council Bluffs, IA	$0.046	$0.951	$11.53	90%
AWS	Ashburn, VA	$0.060	$0.951	$12.29	91%
Azure	Ashburn, VA	$0.060	$0.951	$12.29	91%

Hardware amortization: $25,000 over 3 years = $0.951/hr. PUE: 1.35. Average electricity share of on-demand price: 1.9%.

Source: 1337 Advisors Platform · platform.1337advisors.com/power

> **Observation:** *Electricity accounts for less than 2% of the on-demand price of an H100 SXM. Hardware depreciation is consistent across all providers at $0.951/hr. What varies is the residual: what you pay above the actual cost of the underlying asset. At Hyperstack, residual is 40%. At AWS and Azure, it reaches 91%. You are not paying for electricity or even primarily for the GPU. You are paying for trust, compliance infrastructure, operational resilience, and accumulated enterprise relationships. That residual is a market structure premium, and it has nothing to do with GPU performance.*

The power procurement market offers the clearest structural analog for how early positioning creates durable advantage. Microsoft's twenty-year nuclear agreement with Constellation around Three Mile Island and AWS's 1,920-megawatt commitment at Susquehanna were both made against uncertain future demand. Neither organization knew precisely how much power it would need or when. Both understood that failing to secure long-term power access in a constrained market carried greater risk than over-committing. The consequence of those decisions is now visible in the divergence of cost curves across power procurement strategies.

EXHIBIT 4: POWER COST FORWARD CURVES BY PROCUREMENT STRATEGY (H100 SXM)

Period	Henry Hub Spot	Renewable PPA	New Nuclear PPA	Market Rate
Q2 2026	$3.80/MMBtu	$0.0385/GPU·hr	$0.0805/GPU·hr	$0.0629/GPU·hr
Q3 2026	$3.95/MMBtu	$0.0385/GPU·hr	$0.0817/GPU·hr	$0.0639/GPU·hr
Q4 2026	$4.60/MMBtu	$0.0385/GPU·hr	$0.0829/GPU·hr	$0.0682/GPU·hr
Q2 2027	$4.10/MMBtu	$0.0385/GPU·hr	$0.0853/GPU·hr	$0.0649/GPU·hr
Q4 2027	$5.00/MMBtu	$0.0385/GPU·hr	$0.0879/GPU·hr	$0.0709/GPU·hr
Q4 2028	$5.40/MMBtu	$0.0385/GPU·hr	$0.0931/GPU·hr	$0.0735/GPU·hr

Nuclear PPA: early contracts (2023-2025) at ~$0.075/kWh. New entrants (2026): $0.100/kWh baseline, +6% annual scarcity repricing.

Source: 1337 Advisors Platform · platform.1337advisors.com/power

> **Observation:** *Three providers buying power today face three completely different cost structures that will diverge further over time. The renewable PPA holder faces a flat line regardless of commodity markets. The market rate buyer tracks Henry Hub with full commodity exposure. The new nuclear entrant faces rising costs as viable nuclear capacity becomes increasingly scarce — early contracts set an anchor that new entrants cannot replicate. A provider that signed a renewable PPA in 2021 holds a structural cost advantage that compounds annually over any provider entering the market today. Power procurement is not an energy decision. It is a long-duration competitive positioning decision with consequences that extend across the entire productive life of the infrastructure.*

Compute reservations are developing identical dynamics. The organizations that secured dedicated GPU clusters when capacity was accessible operate under fundamentally different economic conditions than those attempting to procure that same capacity today. The reservation is no longer just a cost. It is a position.

As reservations acquire optionality value independent of utilization, a structural tension emerges. Capacity committed to one participant may not reflect its highest-value use as conditions change. The reservation solved an access problem when it was made. It may create an allocation problem over time.

Part III · Allocation and Market Structure

Forecast divergence creates allocation challenges. The question increasingly becomes how future access is coordinated once conditions change.

5. Allocation Challenges and Forecast Divergence

> KEY INSIGHT
>
> Scarcity and surplus can coexist simultaneously.

Forecast errors do not remain theoretical. They create allocation problems.

Once organizations reserve capacity at scale under uncertainty, the market reliably produces a structural condition: some participants control more future capacity than they need, others need more than they control. These are not separate problems. They are the same problem viewed from opposite sides of the market.

What makes this particularly consequential is that both conditions can exist simultaneously. Scarcity and surplus coexist not because aggregate supply is insufficient, but because access to that supply is locked into earlier assumptions that no longer reflect actual demand. The system contains unmet demand and underutilized capacity at the same time.

This is an allocation problem, not a supply problem.

Supply expansion, meaning more datacenters, more GPUs, and more power, is necessary and will continue. But it does not resolve misallocation. A market built only around supply growth can still fail to move capacity toward the participants who value it most if the mechanisms for reallocation are absent. Scarcity and misallocation are separate pathologies requiring separate solutions.

The rigidity of long-term commitments is the mechanism that creates this friction. A buyer who reserved too much has no efficient way to surface that excess to a buyer who needs more. A provider whose customer's requirements have changed has limited ability to redirect capacity without renegotiation. The commitment that solved an access problem when it was signed may become an allocation problem as the market evolves around it.

This tension between contractual certainty and economic efficiency is inherent to forward markets. Long-term commitments create value: visibility for providers, supply assurance for buyers. They also create rigidity. The more capacity locked into forward agreements, the more important it becomes to understand whether those agreements can adapt as reality changes.

The resolution in mature resource markets has been the gradual development of reallocation infrastructure: mechanisms that allow participants to discover who holds excess, value that access, and transfer it efficiently without collapsing the certainty that made the original commitment worthwhile. Compute markets have not yet built these mechanisms.

The result is a market that is active from a procurement perspective and increasingly inefficient from an allocation perspective. Organizations continue sourcing capacity. Providers continue signing agreements. But future access, once committed, remains largely frozen in place regardless of how conditions change.

As compute becomes more strategically important and reservation horizons extend, the cost of that inefficiency grows. The question is no longer whether enough compute can be built. It is whether future access to compute can move toward the participants who value it most. On that question, the market has barely started.

---

6. Procurement, Allocation, and Reallocation

> KEY INSIGHT
>
> Procurement addresses scarcity. Reallocation addresses misallocation.

Compute markets have invested heavily in procurement. They have invested almost nothing in reallocation. That imbalance matters more as forward commitments grow.

Procurement, allocation, and reallocation are distinct market functions that solve different problems. Procurement determines how a participant acquires access to capacity. Allocation determines who initially controls capacity. Reallocation determines how access moves when the original allocation no longer reflects actual demand. Every market that handles a strategically important resource eventually needs to perform all three functions well.

The historical emphasis on procurement was reasonable. Acquiring capacity was the first-order problem. The neocloud sector, GPU marketplaces, and broker networks that emerged over the last several years represent significant progress on that problem. Participants can now discover, evaluate, and secure compute more efficiently than was possible five years ago.

The same cannot be said for reallocation.

The information required to reallocate future access is fundamentally harder to assemble than the information required to procure it. During procurement, a buyer evaluates providers, hardware configurations, pricing structures, and service levels against a known requirement. During reallocation, the market must determine who controls excess future access, what characteristics that capacity possesses, whether it can be reassigned, and which participants would value it more. This assessment plays out across a fragmented, largely opaque landscape of bilateral commitments.

The result: a participant with excess future capacity may not know who needs it. A participant short on future capacity may not know who holds excess. Capacity exists within the system but cannot be efficiently discovered, valued, or redirected. Mismatches that could be resolved remain unresolved, not because the capacity doesn't exist, but because the market lacks the infrastructure to move it.

Early signals of this evolution are already visible. ICE and Ornn's launch of GPU compute futures represents the first attempt to introduce price discovery and transferability into compute markets. Broker networks and spot markets are early, incomplete forms of reallocation infrastructure. These developments are market responses to growing friction: evidence that participants are beginning to feel the cost of inadequate reallocation mechanisms.

What the market does not yet have is infrastructure designed specifically to handle the reallocation of future compute access: mechanisms for discovering who controls excess, valuing that access, and transferring it efficiently without undermining the provider relationships and contractual certainty that make long-term commitments possible in the first place.

Procurement addresses scarcity. Reallocation addresses misallocation. The second problem is growing faster than the market's ability to solve it.

Part IV · The Evolution of Future Compute Rights

The final section examines whether future access itself may become economically significant enough to drive new market structures.

7. The Evolution of Future Compute Rights

> **KEY INSIGHT**
>
> Future compute rights are not the starting point of the analysis. They offer a smart solution to the problems addressed in it.

The preceding sections trace a sequence.

Competition for intelligence increases the strategic importance of compute. Strategic importance increases the value of future access. Future access requires long-dated commitments. Long-dated commitments create forecast uncertainty. Forecast uncertainty creates allocation mismatches. Allocation mismatches create demand for reallocation. Reallocation requires infrastructure the market has not yet built.

Taken individually, each step looks like a procurement or planning problem. Viewed as a sequence, they describe something more significant: compute markets may be approaching a phase in which future access itself becomes economically important enough to require a new kind of market infrastructure.

That infrastructure is what this paper means by future compute rights.

The term should not be understood as a technology concept or a product specification. It should be understood as a market response to an economic problem: the need to establish, allocate, value, and transfer claims on future compute capacity when that future access has become strategically valuable.

A reservation already begins to resemble such a claim. It gives one participant control over future access that another participant may later value more. It possesses optionality value independent of eventual consumption. It can become misallocated. What it lacks is the infrastructure required to move efficiently when the original allocation no longer reflects actual demand. Future compute rights would close that gap.

The power market again offers a useful structural analog. Long-term power purchase agreements, which were commitments made against uncertain future demand, created the same misallocation pressure that compute reservations are beginning to create. The resolution was not simply more generation capacity. It was the development of secondary markets, transfer mechanisms, and financial instruments that allowed power access to move toward higher-value uses without collapsing the certainty that made long-term contracts worthwhile. Compute is not electricity. But the economic pressure is structurally identical, and the market evolution it produces may follow recognizable patterns.

The geographic dimension of that pressure is worth examining directly. Power costs across US regions alone vary by a factor of seven, and the infrastructure decisions required to access lower-cost power are effectively permanent.

EXHIBIT 5: US REGIONAL POWER PRICING BY GRID OPERATOR

Region	Grid Operator	Price Range ($/kWh)	Characteristics
Nordic Region	—	$0.03 – $0.05	Hydro-dominant, best-in-class cost
Central	SPP	$0.04 – $0.07	Wind-rich, consistently low cost
Texas	ERCOT	$0.04 – $0.08	Isolated grid, no capacity market
Midwest	MISO	$0.05 – $0.08	Large renewable corridor
Mid-Atlantic	PJM	$0.08 – $0.12	Largest ISO by load, 13 states plus DC
New York	NYISO	$0.12 – $0.20	Constrained, urban density
New England	ISO-NE	$0.14 – $0.22	High-cost, limited new capacity
California	CAISO	$0.18 – $0.28	Most expensive, regulatory complexity

US interconnection queue in constrained markets: 5-7 years.

Source: 1337 Advisors Platform · platform.1337advisors.com/infrastructure

> **Observation:** *A 7x difference in power cost exists within the United States alone. For a hyperscaler campus consuming 300-500 MW, that difference is worth hundreds of millions of dollars annually. But interconnection queues in constrained markets run 5-7 years, and long-term power agreements lock in economics for decades. An organization cannot simply move to the cheaper grid. The geographic decisions being made today about where to build and which grid to connect to will determine cost structure for the entire productive life of the infrastructure. Silicon is global. Power infrastructure is permanent. The providers that understood this earliest hold structural advantages that cannot be replicated at any price available in today's market.*

The questions that point toward future compute rights are already being asked:

- Who controls future access to advanced GPU capacity, and how is that access documented?
- How should excess reserved capacity be valued relative to current market alternatives?
- What happens when an organization's compute requirements change materially before a commitment matures?
- How can access move between participants without undermining the provider relationships that make long-term commitments viable?

These questions do not yet have market structure answers. But they are precisely the questions that precede market structure development in every resource market that has reached this stage. They were asked about power before PPA secondary markets developed. They were asked about freight before forward rate derivatives developed. They were asked about bandwidth before capacity trading emerged.

Compute appears to be entering this stage now.

The conclusion of WP-01 is not that future compute rights will emerge in a specific form. It is that the economic conditions required for them to emerge are already present, and the market friction created by their absence is growing. The industry's attention is currently focused on supply: how many GPUs can be deployed, how much power can be secured, how quickly datacenters can be built. Those questions remain important. But supply alone will not resolve the market structure challenges created by long-dated commitments, forecast uncertainty, and misallocation.

The more important question is how future access to compute gets coordinated once that access becomes strategically important. The economic conditions that produce that question are already present. The market infrastructure that answers it does not yet exist. The gap between those two facts is where future compute rights sit. WP-02 begins to lay out their structure.

The Next Phase of Analysis

WP-01 has established the diagnosis: why compute is becoming strategically important, why future access is becoming valuable, why forecast divergence is unavoidable, and why efficient reallocation may become an increasingly critical market function.

WP-02, *The Lifecycle of a Future Compute Right*, turns from diagnosis to design. It will examine how a future compute right might be created, defined, transferred, and redeemed, covering the attributes required to describe future access, the role of providers in establishing claims on capacity, the conditions under which those claims might move between participants, and the mechanisms required to convert future access into actual compute consumption.

WP-01 asks why future compute rights may emerge. WP-02 explores how they might work.

Research Roadmap

Paper	Title	Core Question
WP-01	The Evolution of Compute Markets and Future Compute Rights	Why future compute rights may emerge
WP-02	The Lifecycle of a Future Compute Right	How future compute rights might work
WP-03	Economics and Market Structure of Future Compute Rights	What market structure creates the most efficient allocation and reallocation of future access

WP-02 focus areas: Creation · Metadata · Ownership · Transferability · Redemption

WP-03 focus areas: Incentives · Liquidity · Pricing · Market Design · Risk Management