The Basis of Intelligence: Building a Derivatives Market for AI Compute

Monday, August 31, 2026

INTRODUCTION

Twice this year, the heads of two very different derivatives exchanges reached for the same metaphor. Introducing Kalshi's compute forward curves in July, CEO Tarek Mansour declared: "Compute is the new oil. Like every commodity before it, it needs a real derivatives market."[1] Two months earlier, CME Group Chairman and CEO Terry Duffy had struck the same note when announcing CME's planned compute futures, calling compute "the new oil of the 21st century."[2] The metaphor captures something real about the market's trajectory. Oil became the organizing commodity of the twentieth-century industrial economy; compute may play a comparable role in the AI economy. As explored later in this article, electricity is the closer structural analog for market design, but the underlying signal is the same: AI compute infrastructure is beginning to look like a commodity market in formation.

Both executives are talking about something more than just massive arrays of GPUs and other high-performing chip sets. The unit at issue is the time slice; that is, access to a defined hardware accelerator (a GPU or other specialized chip designed for the intensive parallel computations that AI workloads require), in a defined location, with a defined service level. That unit is already being quoted, reserved, financed and prepared for trading. Notably, the market is being financialized before it has been fully standardized.

Sponsors of exchange-traded funds ("ETFs") have filed for funds that would hold compute futures; CME Group and Intercontinental Exchange ("ICE") have each announced cash-settled GPU compute futures; bilateral platforms are matching forward transactions in physical capacity; and perpetual products are already trading in offshore markets.[3]

---

[1] Kalshi, *Kalshi Announces Compute Forward Curves, Becoming the Exchange for the AI Economy* (July 14, 2026), https://news.kalshi.com/p/compute-forward-curves.

[2] Press Release, CME Group, CME Group and Silicon Data Partner to Launch First Compute Futures (May 12, 2026), https://www.cmegroup.com/media-room/press-releases/2026/5/12/cme_group_and_silicondatapartnertolaunchfirstcomputefutures.html.

[3] For the purpose of this Alert, "compute" is defined as computing power, which is the processing power primarily used by the large language models in the AI economy.

Yet as of the date of this publication, no GPU or other hardware accelerator compute futures contract has received final approval from the Commodity Futures Trading Commission (the "CFTC") or begun trading on a U.S. exchange. That said, the CFTC recently issued a Request for Comment ("RFC") on the listing of compute derivatives contracts, addressing the statutory Core Principles for designated contract markets, the challenges in developing a compute derivatives market, and posing questions on issues related to compute cash markets (*i.e.*, payment for a present service), market oversight and manipulation, customer protection, and perpetual compute futures.[4] It is now apparent to all that units of compute capacity will become or support some type of financial product. But what market infrastructure will compute capacity need to become a credible one?

In this article, we examine:

- The products now being announced or filed, such as compute-futures ETFs, cash-settled futures from CME Group and ICE, platform-based forwards, perpetuals and exchange-for-physical ("EFP") structures — and how the product layers fit together;
- Why a GPU-hour fits comfortably within the definition of a "commodity" in the Commodity Exchange Act (the "CEA");
- What the electricity markets can teach us about hubs, nodes and basis risk, and where the analogy breaks down; and
- The problems the market must solve before it scales — benchmark integrity, price transparency, supplier concentration and hardware depreciation — and the documentation opportunities each of them creates.

1. THE MARKET IS BEING BUILT IN LAYERS

1.1. ETFs ARE ARRIVING BEFORE FUTURES

ETF sponsors are already positioning for investor demand for compute exposure. ProShares has filed for the ProShares AI Computing Power ETF, together with leveraged and inverse versions;[5] Roundhill has filed for the Roundhill Compute ETF (GPUX);[6] Leverage Shares has filed unlevered and 2x compute funds;[7] Volatility Shares has filed unlevered, 2x, -1x and -2x AI Computing Power ETFs;[8] and a filing for the brilliantly named OK Computer

---

[4] Press Release, CFTC, *CFTC Requests Comments on the Listing of Compute Derivatives Contracts*, Release No. 9286-26 (Aug. 19, 2026), https://www.cftc.gov/PressRoom/PressReleases/9286-26; Request for Comment on the Listing of Compute Derivatives Contracts, 17 C.F.R. pts. 1, 38, RIN 3038-AF77 (Aug. 19, 2026).

[5] ProShares Trust, Registration Statement (Form N-1A) (2026), https://www.sec.gov/Archives/edgar/data/1174610/000117461026000177/f45283d1.htm.

[6] Roundhill Compute ETF Prospectus, https://www.roundhillinvestments.com/assets/pdfs/gpux_prospectus.pdf.

[7] Themes ETF Trust, Post-Effective Amendment (Form 485APOS) (2026), https://www.sec.gov/Archives/edgar/data/1976322/000182912626005829/themesetftrust_485apos.htm.

[8] Volatility Shares Trust filing, https://www.otcmarkets.com/filing/html?id=19495796&guid=eLF-kenypr1Jr3h; *see also* VettaFi, *AI ETFs: The Next Wave Emerges* (June 4, 2026), https://etfdb.com/thematic-investing-content-hub/next-wave-ai-etfs/.

Power ETF rounds out the field.[9] The funds generally contemplate investing in compute futures tied to benchmark rental prices for the GPUs used in both AI model training and generating user-requested inference. Because those futures have not been approved or listed, the filings say more about expected investor demand than about a functioning underlying market — a sequencing that commentators were quick to note.[10]

That sequencing creates a familiar ETF problem. A futures-based fund must roll expiring contracts. The roll becomes predictable when the fund's schedule and assets are visible. Other traders can trade ahead of the fund, particularly when open interest is thin and liquidity is shallow. With compute units, that risk is magnified because a newly listed futures contract may have few natural counterparties, wide spreads, and a cash index that is still forming. The leveraged and inverse filings raise the stakes further, since products that reset daily concentrate on trading around the market close and amplify the footprint of any rebalancing flows. A perpetual structure that can trade 24/7/365 and which forgoes an expiry date can eliminate roll risk, but, as discussed below, it does not eliminate the risk that the perpetual's reference benchmark may be unreliable if the underlying spot index is built from thin and opaque price inputs. It may consequently fail to produce a trustworthy settlement price, creating vulnerability to manipulation and distortion.

1.2. FUTURES AND FORWARDS TARGET THE SAME NEED DIFFERENTLY

Exchange-traded futures would add standardized price discovery and central clearing to a market that is currently fragmented across hyperscalers, neo-cloud providers, colocation operators, and private platforms. The market is moving quickly: on May 12, 2026, CME Group announced plans to launch two cash-settled compute futures in partnership with Silicon Data, whose indices are described as the first daily benchmarks for on-demand GPU rental rates.[11] Then, on August 11, CME announced that these two compute futures will be launched on October 5, 2026, subject to regulatory approval.[12] The Silicon Data index family currently covers H100, H200, A100, B200 and MI300X rental prices,[13] and Silicon Data states that its inputs are drawn from neo-cloud providers, hyperscalers, colocation markets and private rental platforms and are standardized for term and scale.[14]

A week later, on May 19, 2026, ICE announced its own suite of U.S. dollar-denominated, cash-settled GPU compute futures based on Ornn's Compute Price Index ("OCPI"), which Ornn describes as the first compute index built exclusively from executed transactions.[15] OCPI's current coverage includes NVIDIA H100, H200, B200 and

---

[9] *See OK Computer Power ETF Files Third Application for Compute Futures That Don't Exist Yet*, Crypto Briefing (May 21, 2026), https://cryptobriefing.com/ok-computer-power-etf-compute-futures/.

[10] *See The Race to Offer Compute Futures to Masses Has Already Started*, Bloomberg (May 20, 2026), https://www.bloomberg.com/news/articles/2026-05-20/the-race-to-offer-compute-futures-to-masses-has-already-started.

[11] Press Release, CME Group, *supra* note 2.

[12] Press Release, CME Group, *CME Group and Silicon Data to Launch Compute Futures on October 5 to Unlock New Way to Hedge AI Risks* (Aug. 11, 2026), https://www.cmegroup.com/media-room/press-releases/2026/8/11/cme_group_and_silicondatatolaunchcomputefuturesonoctober5tounloc.html.

[13] Silicon Data, *GPU Rental Price Indices*, https://www.silicondata.com/products/silicon-index.

[14] *See, e.g.*, Silicon Data, *H100 Rental Price Index*, https://www.silicondata.com/products/silicon-index/h100.

[15] Press Release, Intercontinental Exchange, *ICE and Ornn to Launch GPU Compute Futures Contracts* (May 19, 2026), https://ir.theice.com/press/news-details/2026/ICE-and-Ornn-to-Launch-GPU-Compute-Futures-Contracts/default.aspx.

RTX 5090 accelerator models. Both contracts remain under exchange review presumably to resolve questions of whether they can be shown not to be readily susceptible to manipulation and, therefore, eligible for self-certification.[16] Risk of manipulation is discussed further in Part V. The methodological distinction between the two index families (one providing standardized multi-source observations, the other verified transaction prints) also impacts this risk.

Forward markets are closer to today's commercial reality. Platforms such as Compute Exchange match buyers and sellers of future GPU capacity, after which the parties negotiate and execute the commercial terms directly.[17] A forward contract can customize the exact GPU model referenced as well as the cluster size, region, interconnect, start date and service level — features that a standardized futures contract cannot fully capture.

Kalshi has taken a third route. On July 14, 2026, it introduced GPU compute forward curves for NVIDIA B200, H200 and A100 capacity, derived from prices in its own weekly and monthly compute contracts rather than from spot rental transactions.[18] If those curves deepen, they could become inputs for options, structured products and bilateral hedges. A downward-sloping curve would benefit a buyer seeking to lock in future capacity at a known price, while allowing a seller to trade current spot economics for revenue certainty. But an implied curve is not the same as a liquid cash market. Its usefulness will depend on volume, transparency, and the relationship between the exchange's contracts and actual compute transactions.

1.3. PERPETUALS AND EFPs ADDRESS DIFFERENT PROBLEMS

On May 28, 2026, Architect Financial Technologies announced that it had acquired a CFTC-designated contract market ("DCM") and plans to launch the American Innovation Exchange, offering futures and options on compute costs tied to multiple GPU vendors and models, subject to regulatory review.[19] Its "ComputeConnect" product would facilitate exchanges of futures for physical capacity.[20] Meanwhile, its AX perpetual futures exchange — regulated by the Bermuda Monetary Authority — already offers GPU-hour perpetual futures to non-U.S. participants.[21]

With an EFP,[22] on the other hand, a buyer could purchase H200 futures six months ahead and, as the delivery window approaches, exchange the futures position for a capacity commitment with the required cluster size, location, interconnect and service window. The physical leg would trade at a premium or discount to the futures

---

[16] *Id.*; *see also* Ornn, *Price Index Methodology*, https://data.ornn.com/docs/price-index.

[17] Compute Exchange, *Forward Contracts*, https://compute.exchange/forward-contracts.

[18] Kalshi, *supra* note 1; *see also Kalshi Ramps Up Effort to Build Markets for AI Computing Power*, Bloomberg (July 14, 2026), https://www.bloomberg.com/news/articles/2026-07-14/kalshi-ramps-up-effort-to-build-markets-for-ai-computing-power.

[19] Press Release, Architect Financial Technologies, *Architect Financial Technologies to Launch U.S. Futures Exchange for Compute and AI Commodities Following DCM Acquisition* (May 28, 2026), https://www.prnewswire.com/news-releases/architect-financial-technologies-to-launch-us-futures-exchange-for-compute-and-ai-commodities-following-dcm-acquisition-302784564.html; Architect, *AI Exchange*, https://architect.co/ai-exchange/.

[20] Architect, *ComputeConnect*, https://architect.co/computeconnect/.

[21] Architect, *AX*, https://architect.co/ax/.

[22] An EFP is a privately negotiated transaction in which two parties simultaneously exchange a futures position for an equivalent or closely related position in the underlying cash or physical market. One party buys the futures and sells the related physical position, while the other party sells the futures and buys the physical position.

price depending on reliability, flexibility and other attributes. The futures leg supplies standardization, the physical leg provides specificity, and the negotiated difference is the basis.

Perpetuals take a different route. By concentrating liquidity in a single non-expiring instrument, they avoid the liquidity fragmentation and predictable roll activity that can accompany a strip of dated contracts. A perpetual's reference price and funding mechanism is critical. On May 29, 2026, the CFTC approved KalshiEX's bitcoin perpetual futures contract under Regulation 40.3 and simultaneously issued a policy statement explaining that perpetual contracts referencing asset classes other than some digital assets will be evaluated on their own terms.[23] The CFTC's bitcoin determination emphasized bitcoin's deep, active and continuous spot market. In June 2026, the CFTC separately sought comment on extending standard futures contracts to 24/7 trading and on perpetual contracts referencing physically delivered or storable energy commodities, later extending the comment period to August 26, 2026.[24] Those initiatives treat 24/7 trading and perpetual structure as distinct design questions. As discussed further below in Part V, compute units do not yet have deep, continuous and broadly distributed price formation comparable to that of, for example, bitcoin. This will be a hurdle for achievement of a futures product, including a perpetuals product.

1.4. REGULATORY ISSUES: CORE PRINCIPLE 3 AND THE PATH TO LISTING

A DCM may list only contracts that are not readily susceptible to manipulation under CEA Core Principle 3.[25] For a cash-settled compute future, the analysis turns largely on the settlement index and the exchange's ability to monitor the reference market.[26] Part 38, which sets forth the implementing rules of the CEA Core Principles, requires the DCM to monitor the pricing of the index to which the contract will be settled, assess whether the methodology remains appropriate and maintain access — directly or through information-sharing arrangements — to trader positions and transactions in the reference market.[27]

For a novel compute contract — and especially for any compute perpetual — case-specific engagement and CFTC review under Regulation 40.3, rather than self-certification, is the more realistic path.[28]

---

[23] Press Release, CFTC, *CFTC Approves BTCPERP Contract Submitted by KalshiEX, LLC* (May 29, 2026), https://www.cftc.gov/PressRoom/PressReleases/9240-26; CFTC, *Policy Statement Concerning the Listing of Perpetual Contracts*, 91 Fed. Reg. (June 3, 2026), https://www.federalregister.gov/documents/2026/06/03/2026-11020/policy-statement-concerning-the-listing-of-perpetual-contracts. For a fuller discussion, see CahillNXT, *True Perpetuals Come Onshore* (June 2, 2026), https://www.cahill.com/publications/client-alerts/2026-06-02-true-perpetuals-come-onshore.

[24] Press Release, CFTC, *Request for Comment on the Extension of Standard Futures Contracts to 24/7 Trading and on Perpetual Contracts Referencing Physically Delivered or Storable Energy Commodities* (June 22, 2026), https://www.cftc.gov/PressRoom/PressReleases/9259-26; Press Release, CFTC, *CFTC Extends Public Comment Period* (July 23, 2026), https://www.cftc.gov/PressRoom/PressReleases/9271-26.

[25] 7 U.S.C. § 7(d)(3); 17 C.F.R. §§ 38.200–38.201; 17 C.F.R. pt. 38, app. C.

[26] As noted by the CFTC in its RFC, the majority of compute derivatives initially would be cash-settled because of the potential infrastructure-related challenges associated with delivering the underlying commodity. *See* RFC, p.9.

[27] 17 C.F.R. §§ 38.251–38.254.

[28] On July 9, 2026, the Commission stayed CME's self-certified extension of crude oil futures to 24/7 trading under Regulation 40.2(c), with the Chairman pointedly criticizing the exchange's decision to press ahead while a related comment period remained open, and the Commission encouraging exchanges to consult with staff before

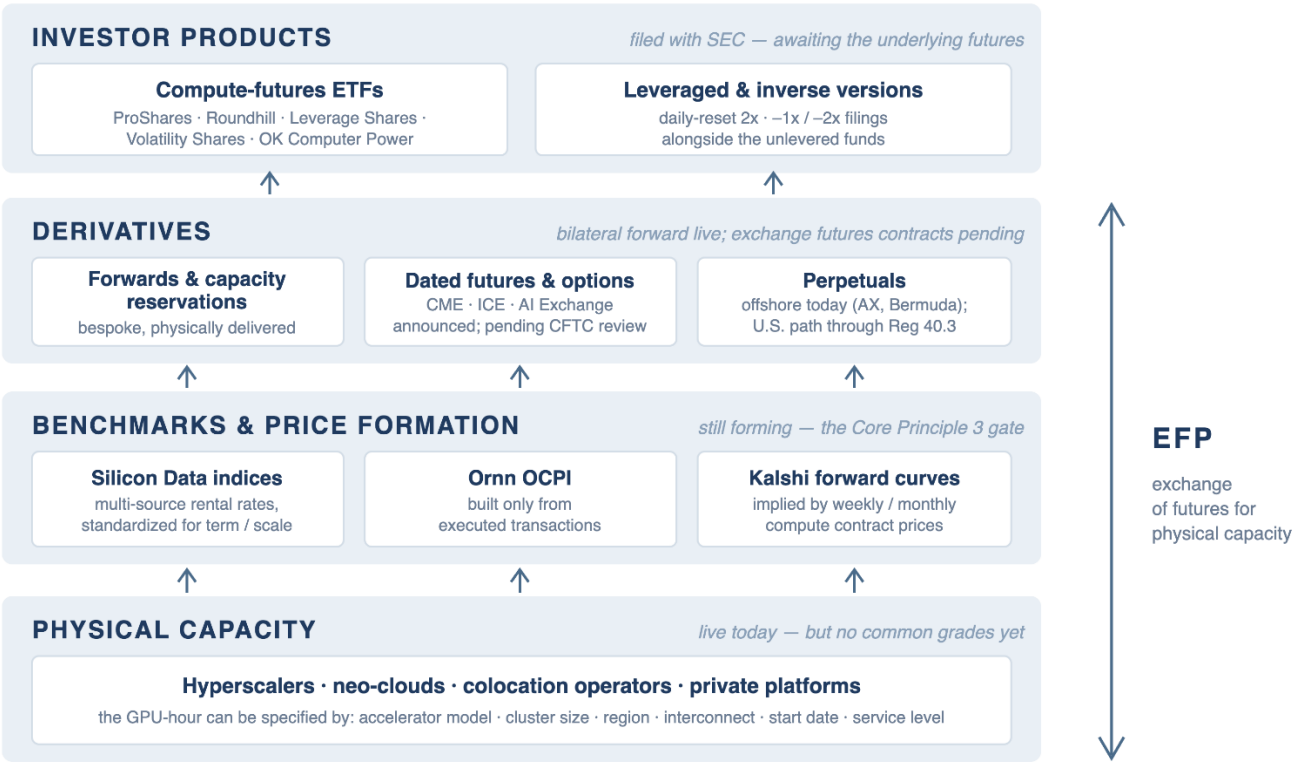


Figure 1 *— The layered structure of the emerging compute market. Each layer depends on the one beneath it, and the EFP mechanism ties exchange-traded contracts back to physical capacity. As of the date of this Alert, the upper layers are the furthest advanced on paper, while the two layers beneath them remain furthest from standardization.*

2. WHY COMPUTE CAN BECOME A TRADEABLE COMMODITY

2.1. GPU-HOURS FIT THE CEA'S DEFINITION OF "COMMODITY"

A GPU-hour is the exclusive use of one specified graphics processing unit for one hour. It is the core unit around which the emerging compute derivatives market is being built. CME Group's announced contracts reference the rental price of a specific GPU model (NVIDIA H100 or B200), with each contract sized at 730 GPU-hours, approximating one continuously rented GPU for an average month (365 days multiplied by 24 hours, divided by 12 months). These are financially settled rental-price contracts; no physical GPU time changes hands. ICE has announced contracts on two different indices: Ornn's transaction-based OCPI, which is separated by GPU model, and NATIVX's COIL Index, which adjusts quoted compute prices to account for distortion from different regional electricity prices.[29]

---

proposing novel contracts. Press Release, CFTC, *CFTC to Stay Self-Certified Contract on 24/7 Trading for Crude Oil Futures* (July 9, 2026), https://www.cftc.gov/PressRoom/PressReleases/9265-26.

[29] Press Release, Intercontinental Exchange, ICE and NATIVX to Launch Energy-Normalized Compute Futures Contracts (July 1, 2026), https://www.businesswire.com/news/home/20260701837320/en/ICE-and-NATIVX-to-Launch-Energy-Normalized-Compute-Futures-Contracts.

An important limitation is that the GPU-hour is hardware specific rather than a universal measure of computational work. An H100-hour, a B200-hour and an RTX 5090-hour each deliver different throughput depending on precision, workload, memory, interconnect and software stack. That is why both CME and Ornn maintain separate indexes by GPU model rather than converting everything into a single performance metric such as floating-point operations.

The CEA's definition of "commodity" is deliberately broad. It reaches not only enumerated agricultural products but also "all other goods and articles" and "all services, rights, and interests" in which contracts for future delivery are or become the subject of dealing.[30] A right to access a defined quantity of GPU capacity for a defined period can fit within this framework whether it is described commercially as a service, a capacity reservation, or an interest in machine time. GPU-hours are therefore a viable underlying for a futures contract, subject to the CEA's other requirements. The CFTC's RFC reinforced this view, recognizing that while compute may not yet exhibit the fungibility, standardization and sufficient liquidity of commodities that typically underlie a derivatives market, these novel issues are "informed by historical antecedents" in "similar commodities," and the Commission cited the example of electricity.[31]

2.2. THE COMMERCIAL CASE FOR A FORWARD PRICE

AI labs and enterprises face exposure to the cost of running training and inference workloads on particular clusters of GPUs and other accelerator hardware. Hyperscalers and neo-clouds face the opposite side of that exposure. They have large, fixed investments in GPUs, power, cooling, networking and data-center capacity, with revenue that varies as utilization changes. Both sides need better visibility into future prices, and both need tools to hedge costs, revenue and capital commitments.

A credible compute derivatives market could provide four things the cash market does not yet provide consistently: (1) forward price discovery; (2) a hedge for compute costs and capacity revenue; (3) a planning signal for data center and fleet investment; and (4) a reference price for financing, collateral valuation and GPU-backed lending.[32] The economic case is straightforward even if the market design lags.

2.3. FORWARDS CAN DELIVER A COMMERCIAL HEDGE (IF THEY ARE TRULY PHYSICAL)

Forward contracts fit the current GPU procurement model because they can be tailored to a bilateral relationship. A buyer can request quotes from several counterparties, choose to buy or sell future capacity, and use a broker or marketplace to find the other side.

---

[30] 7 U.S.C. § 1a(9); 17 C.F.R. § 1.3; *see also* Gary E. Kalbaugh, *Why Regulate Commodities?*, 57 Suffolk U. L. Rev. 43, 44 (2024).

[31] *See* RFC, p.12 & n.45 (citing *Proposal To Exempt Certain Transactions Involving Not-for-Profit Electric Utilities*, 77 FR 164 (Aug. 23, 2012)) ("Unlike many physical commodities, electric energy is not capable of being purchased in large commercial quantities ahead of time, delivered, and stored for later consumption or use. That is, electric energy must be used or consumed on an as-needed basis.").

[32] This is echoed by the CFTC in its RFC, noting that "[a] compute futures market may therefore provide a means for managing and assuming price risks, discovering prices, or disseminating pricing information as to general trends in AI adoption. Futures markets aggregate the disparate views of individual market participants and are designed to reveal prices that allows such market participants to efficiently hedge risks or speculate (in this case, on, for example, the future scale of AI demand). In short, compute futures contracts would allow financial market 'to aggregate and reveal information about the future of the AI economy'." *See* RFC, p.11.

The CEA excludes from "future delivery" a sale of a cash commodity for deferred shipment or delivery, and excludes from the definition of "swap" a sale of a nonfinancial commodity for deferred shipment or delivery when the transaction is intended to be physically settled.[33] Those exclusions can provide a route for bona fide commercial forwards, but they are not a blanket exemption for any contract labeled a "forward." The analysis turns on the facts such as the parties' ability and intent to make or take delivery, the terms of delivery, whether the arrangement is leveraged or margined, and whether the transaction is genuinely commercial rather than a disguised financial instrument.[34]

With compute, "delivery" should mean provisioned, exclusive and verifiable access to conforming accelerator capacity for the contract window, not merely a promise to use best efforts to provide that delivery. Existing market materials describe delivery as reserved access to a defined volume of compute beginning on an agreed date.[35] The documentation opportunity is to convert that commercial concept into objective grades, conformance tests, uptime and service-level remedies, substitution rights and benchmark fallbacks.

Forwards nevertheless have limits. They remain bespoke and therefore difficult to net, offset or use as a portfolio hedge. Delivery obligations must be real; the parties must have the ability to make and take delivery; and any subsequent book-out ordinarily will be a new, separately negotiated transaction rather than an embedded right of offset. Those frictions are precisely why standardized futures, swaps and EFPs are needed alongside the forward market. Forwards alone cannot solve every risk.

A mature market also needs a common and standardized grading system. A contract for "Grade A H200, U.S. East" should identify the specific hardware (by SKU), region of installation, interconnect methodology, performance standard, service level tier, and verification method that make one provider's capacity comparable to another's. Until those attributes are standardized, the market will have price discovery without full fungibility.

3. ELECTRICITY IS A WORTHY ANALOGY — UP TO A POINT

3.1. HUBS, NODES AND CLUSTER-SPECIFIC EXPOSURES

Electricity shows how a heterogeneous time- and location-dependent product can become a tradeable commodity complex. Compute is similar, although it does not yet have electricity's institutional infrastructure:

- GPU families — H100, H200, B200 and similar models — play the role of benchmark generation types.
- Cloud providers, neo-clouds, marketplaces and colocation operators supply the physical and commercial layer.
- A GPU deployment in a particular data center functions much like an electricity generation node. In electricity markets, nodes represent specific generation or delivery points connected to a grid through transmission infrastructure, and prices at each node reflect local supply/demand conditions, congestion, and capacity constraints. Similarly, a GPU cluster in a specific data center is tied to a fixed physical location with finite network interconnects, power, and cooling capacity and

[33] 7 U.S.C. §§ 1a(27), 1a(47)(B)(ii).

[34] 7 U.S.C. § 2(c)(2)(D). Leveraged or margined products offered to retail participants present a separate analysis under the CEA's retail commodity transaction provisions.

[35] *See, e.g.*, Compute Exchange, *supra* note 17; Forward Compute, *Forward Contracts*, https://www.forwardcompute.ai/products/forward-contracts.

the price of compute at that location will reflect utilization, provider, hardware configuration, and infrastructure constraints unique to that site.

- A broad index — an H200 benchmark built from multiple providers, for example — functions more like a power hub than a single vendor's quote. It can be useful for hedging general market exposure without matching every detail of a particular deployment.

3.2. BASIS RISK IS THE NEXT PRODUCT OPPORTUNITY

Electricity hedgers use hub futures even when their physical exposure sits at a specific node. The difference is considered "basis" risk and the electricity markets have developed nodal-to-hub basis swaps to manage it. Compute will likely develop similar architecture with different drivers. A cluster's price may reflect utilization, hardware configuration, network performance, power and cooling costs, location, reliability, and the commercial terms of the reservation. A broad compute index will hedge market movements without perfectly hedging the price of the specific cluster a customer actually needs.

That gap creates room for an over-the-counter compute basis market. A dealer or specialized desk could exchange a cluster-specific price for an index price (plus or minus an agreed spread) letting a standardized future hedge broad exposure while a basis swap addresses the residual deployment risk. EFPs offer a complementary solution. Rather than settling the residual in cash, the parties use the futures price as the anchor and negotiate the premium or discount for the actual capacity. If EFPs become liquid and transparent, they could remove basis from being purely a private negotiation and render it a quoted market variable.

4. WHAT THE MARKET MUST SOLVE BEFORE IT SCALES

4.1. INTRANSPARENCY AND CONCENTRATION WILL WIDEN BASIS

Wholesale electricity prices benefit from a public market architecture. For example, rates are filed, and system operators publish nodal prices under a transparency framework administered by the Federal Energy Regulatory Commission.[36] Compute has no comparable disclosure mandate or regulatory oversight body. Moreover, compute markets are inherently more global than electricity: while electricity markets are bounded by the physical reach of transmission grids and regulated by national or regional authorities, compute capacity can be purchased from providers anywhere in the world and delivered over the internet, making price formation diffuse and cross-border surveillance more difficult. That opacity will widen the difference between a market-wide index and any one provider's price. It also means that a benchmark may reflect the reporting choices of a small number of large suppliers rather than the full economics of the market. More data on volume, transaction frequency and contributor concentration will be better; alongside clear contribution standards, auditability, publication of methodology and a process for handling missing, stale or outlier observations.

4.2. BENCHMARK INTEGRITY IS THE GATING ISSUE

A third-party index does not shift responsibility away from the exchange. The DCM must conduct due diligence, monitor the index and its methodology, and maintain access — directly or through information-sharing arrangements — to relevant information from the reference market.[37] As noted in the RFC, Appendix C to Part 38

[36] *See* 16 U.S.C. §§ 824d, 824t.

[37] CEA Core Principles 3&4.

requires appropriate consideration to be given to the commercial acceptability, public availability, and timeliness of the price series that is used to calculate the cash settlement price.[38]

A benchmark built from thin, voluntary or non-public inputs can be moved by a party that controls capacity, reports prices or holds a position in the derivative. The CFTC has made the same observation, asking what protections would prevent a provider from manipulating a cash settlement index by adjusting a posted rate, directing capacity onto or away from a venue whose transactions the methodology treats as input data, or executing or declining to execute transactions during the observation window.[39] The mitigants are contribution standards, transaction verification, weighting rules, stale-price and outlier treatment, governance, audit rights, fallback methodologies and information-sharing arrangements.

Concentration makes the problem harder. If a small group of hyperscalers and neo-clouds controls most available capacity, the market may be unable to demonstrate broad deliverable supply or broad price formation. A provider that can influence both the physical market and the benchmark should be addressed through methodology, disclosure, surveillance, position controls and governance. This can reduce the potential for manipulation of the cash settlement price, and enhance the reliability of that price as an indicator of cash market values.[40]

4.3. HARDWARE DEPRECIATION CHANGES THE CONTRACT

Electricity does not depreciate: a megawatt-hour in 2030 is the same product as one in 2025. Compute is different. An H100-hour in 2030 may describe older, less valuable hardware — or a product with no meaningful rental market at all. Every index and forward curve must therefore address when a hardware generation is retired, how a successor model is treated, and what happens when the reference series ceases to be published. No amount of financial engineering removes that physical and technological risk.

Hardware depreciation thus creates inherent tenor limits. If the curve is expected to decline as new accelerators arrive and supply expands, a ten-year compute future may be economically artificial. Long-dated hedges may instead require bilateral contracts with negotiated fallbacks, capacity commitments or structured exposure across hardware generations. A robust and sustainable definition of “GPU hour” must be developed. Perpetual contracts face an acute version of this problem. Because a perpetual has no natural expiry, hardware obsolescence can degrade the underlying spot index over time as rental volume declines and index observations thin, eventually threatening the funding mechanism and settlement integrity. Properly designed perpetual contracts should therefore include a sunset trigger tied to minimum index liquidity, with a predeclared final cash settlement hierarchy, so that the instrument can be wound down in an orderly fashion before the reference market becomes unreliable.

Technological turnover also fragments liquidity. A benchmark tied to one GPU model can become less relevant as workloads migrate, forcing the market to list overlapping contracts and spreads across generations. The result may be thinner liquidity precisely where users want specificity. Perpetuals and inter-model spreads may help, but only if the underlying data and reference mechanics remain robust.

Market participants will need benchmark disruption events, index provider replacement and acquisition provisions, fallbacks for hardware becoming outdated or obsolete, successor model adjustments, service level adjusted floating amounts, substitution and outage mechanics, *force majeure* provisions tailored to cluster failures, and clear rules for

---

[38] *See* RFC, n.38.

[39] *See* RFC, at pp.15-16.

[40] *See* RFC, p.10.

valuation and termination. Sophisticated legal input will be required to help define the market's grades, delivery standards and fallback economics.

4.4. SWAPS ARE READY NOW

Swaps can and are being traded in the compute market today. Their versatility — together with a long history of robust contractual fallbacks for disrupted or inaccurate price sources in established commodity markets — makes them a natural vehicle for compute risk management.

In May 2026, FalconX announced what it described as the first OTC swap referencing the forward price of compute.[41] FalconX acted as dealer, with Superstate founder Robert Leshner as counterparty. The transaction referenced the Ornn's OCPI for NVIDIA H100 GPUs. FalconX did not disclose the trade's notional amount, maturity, fixed price or settlement calculation.

Economically, one party pays a fixed price for a specified quantity of GPU-hours while the other pays the benchmark price, with the difference settled in cash. A cloud provider or AI lab could pay fixed price to hedge rising rental costs; a GPU operator could receive fixed price to protect against falling rates. These are customized OTC contracts, rather than standardized, exchange-traded products; they introduce dealer credit exposure and make price discovery difficult. The benchmark itself is a material source of risk given fragmented H100 rental markets. The FalconX transaction is best understood as early proof of market infrastructure rather than evidence of a mature swap market, but it demonstrates that the contractual mechanics of commodity swaps can be adapted for compute.

5. CONCLUSION AND NEXT STEPS

AI compute is rapidly moving toward financialization for the same reasons that prior derivatives markets have developed: the underlying commercial problem of buyers needing to manage future capacity costs, sellers needing to manage utilization and capital intensity, and investors wanting exposure to the growth of the AI infrastructure stack. But the market will only truly mature when a "GPU hour" has a credible and widely accepted definition, transparent benchmarks, enforceable delivery or settlement mechanics, and enough liquidity to support hedging without creating a new layer of basis risk.

The immediate opportunities for market participants are many and include building the benchmarks, the basis market, the EFP structure, the collateral and valuation framework, the second-order products (ETFs etc.) and the documentation that makes all of those pieces interoperable. Right now, the ambition driving the development of compute derivatives markets is outpacing the practical reality of implementation, but the direction of travel is clear: once price formation and governance catch up with the hardware, GPU hours can become a new financial building block for AI infrastructure and a new trading product for many different types of traders and investors.

The CahillNXT team is prepared to assist clients in evaluating compute-linked products and index methodologies, structuring physical forwards and capacity agreements designed to remain within the CEA's forward exclusions, engaging with the CFTC on novel contract designs, structuring and documenting compute swaps and related products, and developing the grading, fallback and benchmark-disruption documentation this market will require. We welcome the opportunity to discuss how these issues apply to your specific circumstances.

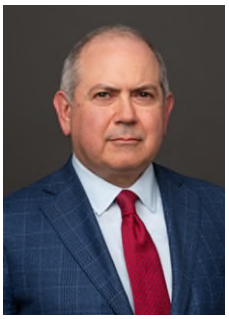
---

[41] Press Release, FalconX, FalconX Executes First OTC Compute Forward Trade (May 27, 2026), https://www.falconx.io/newsroom/falconx-executes-first-otc-compute-forward-trade.

CONTACT

If you have any questions about the issues addressed in this Alert, or if you would like a copy of any of the materials mentioned in it, please do not hesitate to call or email authors Gary Kalbaugh (partner) at 212.701.3505 or gkalbaugh@cahill.com, Lewis Rinaudo Cohen (partner) at 212.701.3758 or lrcohen@cahill.com, or Chloe Chan (law clerk) at 212.701.3058 or cchan@cahill.com; or email publicationscommittee@cahill.com.

Gary Kalbaugh
Partner
212.701.3505
gkalbaugh@cahill.com



Lewis Rinaudo Cohen
Partner
212.701.3758
lrcohen@cahill.com

Chloe Chan
Law Clerk
212.701.3058
cchan@cahill.com