Computation DePIN

November 2025

Created by:

Solus Partners & KOLs

CONTENT

EXECUTIVE SUMMARY

DEPIN MARKET CAP	YTD REVENUE LEADER	SLA/ATTESTATION GAP
$14.6B ↑ +0.2%	**$147M** Aethir ARR	**$5-10B**
Aggregate market capitalization across all DePIN protocols	*Q3 2025: $39.86M | +5x YoY growth*	*Opportunity blocked by compliance gaps*

		Argentum AI			
15.8	14.4	14.4	13.65	12.1	11.35
Leader	Watch	Watch (Early)	Watch	Watch	Risky

- DePIN compute remains **price-competitive for batchable AI workloads** (fine-tunes, offline inference, rendering)
- **Public, network-wide SLA/latency dashboards are still limited**; enterprises should contract for audited uptime, job receipts and, where available, TEE/attestation
- **Fees and activity are concentrated in a few networks** (Aethir, io.net, Helium/other DePIN verticals); revenue growth is visible but uneven QoQ
- **Institutional adoption depend on** verifiable SLAs, predictable billing, regional pinning/data-residency, and export-control posture
- Argentum AI enters as an early, compliance-aware compute marketplace focused on institutional security, re-using second-gen GPUs and multi-region diversification; monitor for public SLA dashboards and audited benchmarks

INTRODUCTION

This report reviews the DePIN compute segment (AI and data-center workloads) as of 05 Nov 2025

The focus is operational: supply sources, orchestration and verification, SLA posture, and enterprise controls. We summarize network activity and fees, highlight where telemetry is reliable, and surface gaps that limit institutional adoption

In this report:

- mapped the Computation DePIN landscape
- defined workloads and SLAs, mapped segments (GPU marketplaces, de-cloud)
- publicly available data (pricing, usage claims) are compiled into a weighted comparative scorecard
- assessed 6 providers on six pillars (0–20 scale): Performance/SLA, Cost, Supply Quality, Demand Quality, Compliance, Token Sustainability.

Methodology & scope notes

- Reference window. All facts are current as of 05 Nov 2025. 2024–2025 sources; older items are tagged with their year
- Evidence priority. Public docs, dashboards, audits, credible research; numbers are cross-checked with ≥2 sources
- Metrics tracked. Usage/fee revenue, active devices and capacity, SLA/latency evidence, attestation/TEE coverage, pricing examples, procurement artifacts (DPAs/SCCs, regional pinning, export-control posture).
- Exclusions. No token price/FDV analysis. Market-cap figures appear only when tied to infrastructure context and not to draw investment conclusions.

MARKET CONTEXT

AI compute demand is compounding; advanced GPUs remain supply-constrained.

DePIN exploits cost and location arbitrage , but enterprise buyers require auditable SLAs and datahandling assurances

DEPIN MARKET CAP

$14.6B ↑ +0.2%

Aggregate market capitalization across all DePIN protocols

TOTAL PROTOCOLS TRACKED

423 (102 Compute related)

Live: 50+ | Emerging: 100+ | Watch: 273

YTD REVENUE LEADER

$147M Aethir ARR

Q3 2025: $39.86M | +5x YoY growth

ENTERPRISE CLIENTS

150+ (compute only)

Aethir: 435K GPU containers | 90+ countries

ADDRESSABLE MARKET 2027

$80-150B

Batch training, inference, rendering, agents

SLA/ATTESTATION GAP

$5-10B

Opportunity blocked by compliance gaps

KEY FINDINGS

1. **Many decentralized providers show lower list prices than big clouds**. In the context of the workload type, best nearterm fit might be batch inference, scheduled fine-tunes, burst training

Aethir's pricing calculator showed an Nvidia H100 instance at ~$1.25/hr (mid-2025) - up to 90% cheaper than cloud on a per-hour basis

Akash's public marketplace had H200 GPU rentals around $1.40/hr (substantially below AWS rates).

io.net and Nosana similarly claim 60-70% cost savings

2. **Realized savings depend on workload tolerance for variability**. Clients only see cost benefits if their workloads can tolerate variable throughput or availability. Where strict low latency (p95 < 100–150 ms) or high uptime is required, these networks may not yet be suitable without careful engineering

- In the terms of reliability of the technology for today's date public, network-wide SLA dashboards (uptime, p95) and hardware attestation coverage are inconsistent. Enterprises bear more execution and audit risk
- Compliance: enterprise adoption depends on clear controls for data residency, export controls (advanced GPUs), and PII rules (GDPR/CCPA)

To date, few compute DePIN projects have built-in features for geo-fencing or legal contracts with node operators. The **EU AI Act** (phasing in 2025–2027) will likely require model transparency and risk controls even for infrastructure providers.

3. **Supply Quality.** The supply side ranges from enterprise-grade data center GPUs to consumer-grade cards contributed by hobbyists.

Aethir's pricing calculator showed an Nvidia H100 instance at ~$1.25/hr (mid-2025) - up to 90% cheaper than cloud on a per-hour basis

Some projects are instituting quality tiers - e.g. ***Akash*** *is working on attestation via Trusted Execution (TEE/TPM) and GPU vendor attestation (e.g. NVidia's NVML-based attestation) so that clients can request "verified" nodes. Others like* ***Nosana*** *offer "Confidential Jobs" where data is handled privately, but this stops short of proving hardware integrity*

SECTOR SNAPSHOT

DePIN is a relative young sector besides some of the first-movers like Helium, Akash and Arweave. More significant monetization started in the last year:

- **YoY fees did ~5x**, with continued growth **driven by Aethir and io.net**
- Aethir offering GPU-compute has a large share here, though these fees are based on buy-backs
- The **World Economic Forum** projects the valuation of the DePIN sector to **$3.5T by 2028 (~90x vs. 2025)**, indicating that the growth is going to continue

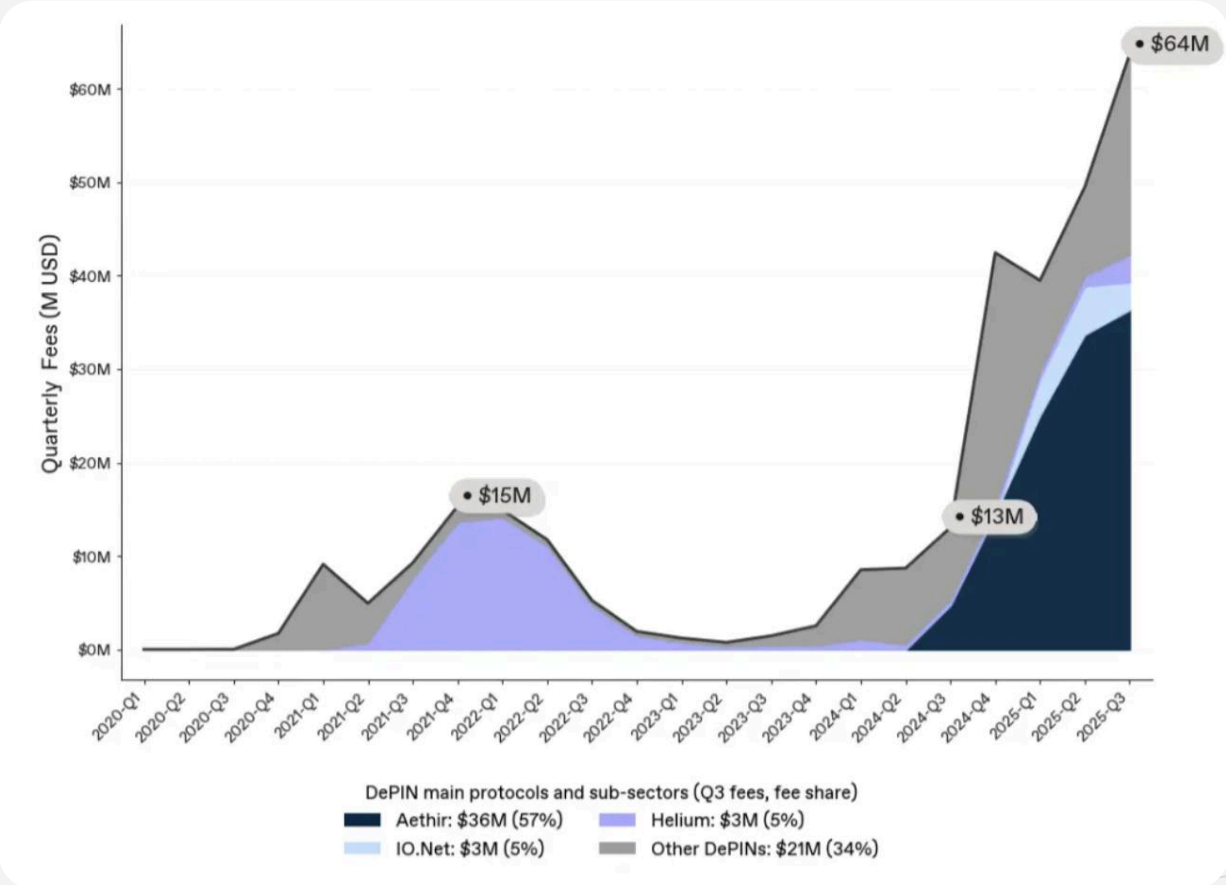


Figure 1. On-chain fees for DePIN, 2020–Q3 2025

Fees accelerated into 2025 with visible contributions from Aethir and io.net, while Helium and other DePINs remain smaller in absolute terms. The growth path is volatile; sequential slowdowns appear around mid-2024 followed by a stronger 2025 ramp

SECTOR SNAPSHOT

Figure 2. DePIN Market Cap / Volume / Projects / Devices

Project count and device growth outpaced market cap at times, implying **supply expansion ahead of sustained demand**. Spikes in volume tend to precede device plateaus, consistent with episodic incentives or narrative waves.

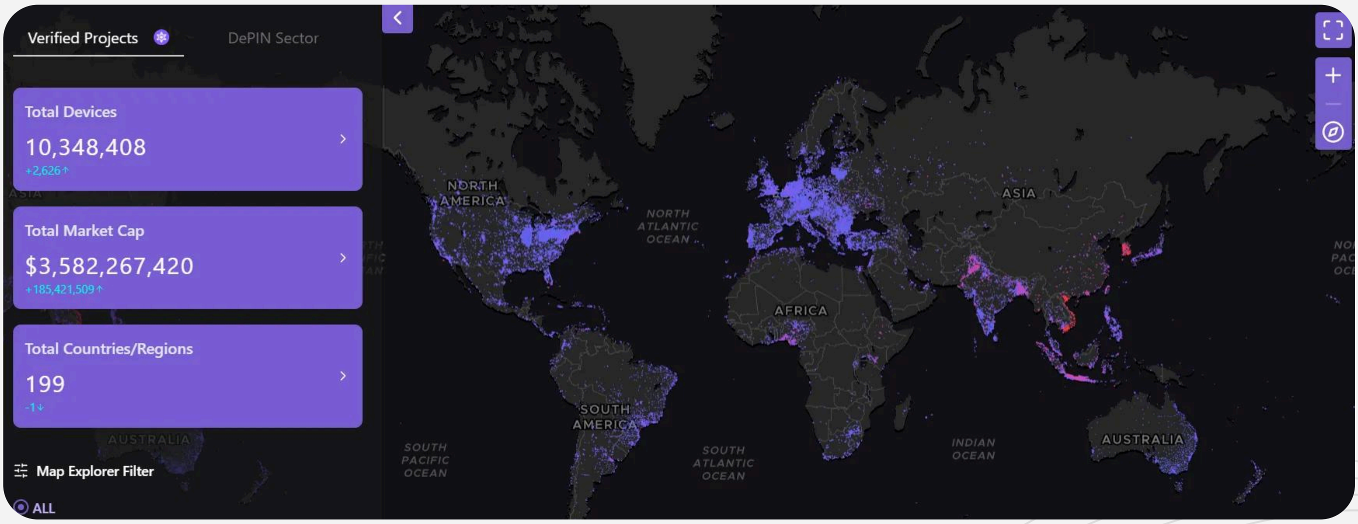


Figure 3. Global distribution of verified DePIN devices

Node density clusters in North America, Europe, and parts of East Asia. Regional gaps persist across Africa and parts of Latin America, which affects latency/SLA guarantees for global portfolios. For institutions, region pinning and redundancy across jurisdictions remain required controls.

MARKET SIZE (TAM / SAM / SOM)

Category	Definition	Market size (USD) by 2030	Source
TAM	AI-related datacenter/ compute spend	**5.2T** (AI-ready DCs) within **6.7T** total datacenter capex cumulatively through 2030	McKinsey
SAM	GPU/AI compute delivered as cloud services	GPUaaS : **12.26B** AI-as-a-Service: **91.2B**	GPUaaS AIaaS
SOM	Near-term DePIN share	**20-30B**	GPUaaS DePINscan

Key growth drivers for Computation DePiN

- **Increasing AI workloads** that are forcing a step-up in data-center investment, as McKinsey sees $6.7T cumulative DC capex by 2030, with $5.2T for AI-ready facilities

  > *Reuters reported ongoing Nvidia GPU shortages in China and parts of Asia in 2025: even major cloud providers cannot fully satisfy regional demand*

- **Uneven access to high-end GPUs** creates pricing and allocation frictions that decentralized markets can arbitrage

  > *Nvidia H100 80GB on AWS can cost $10-$12 per hour, versus ~$1-2/hour on some decentralized networks (though with caveats)*

- **Regulatory changings.** The EU AI Act's phased obligations are driving compliance-aware infrastructure choices, while the IEA expects data-centers electricity demand to more than double by 2030, pushing operators to tap cheaper power geographies and alternative supply
- **Technological Maturation.** The software stack for distributed computing is improving. Containerization and orchestration tools (Kubernetes, etc.) make it easier to distribute workloads across independent providers. The rise of modular AI workloads can align well with decentralized networks

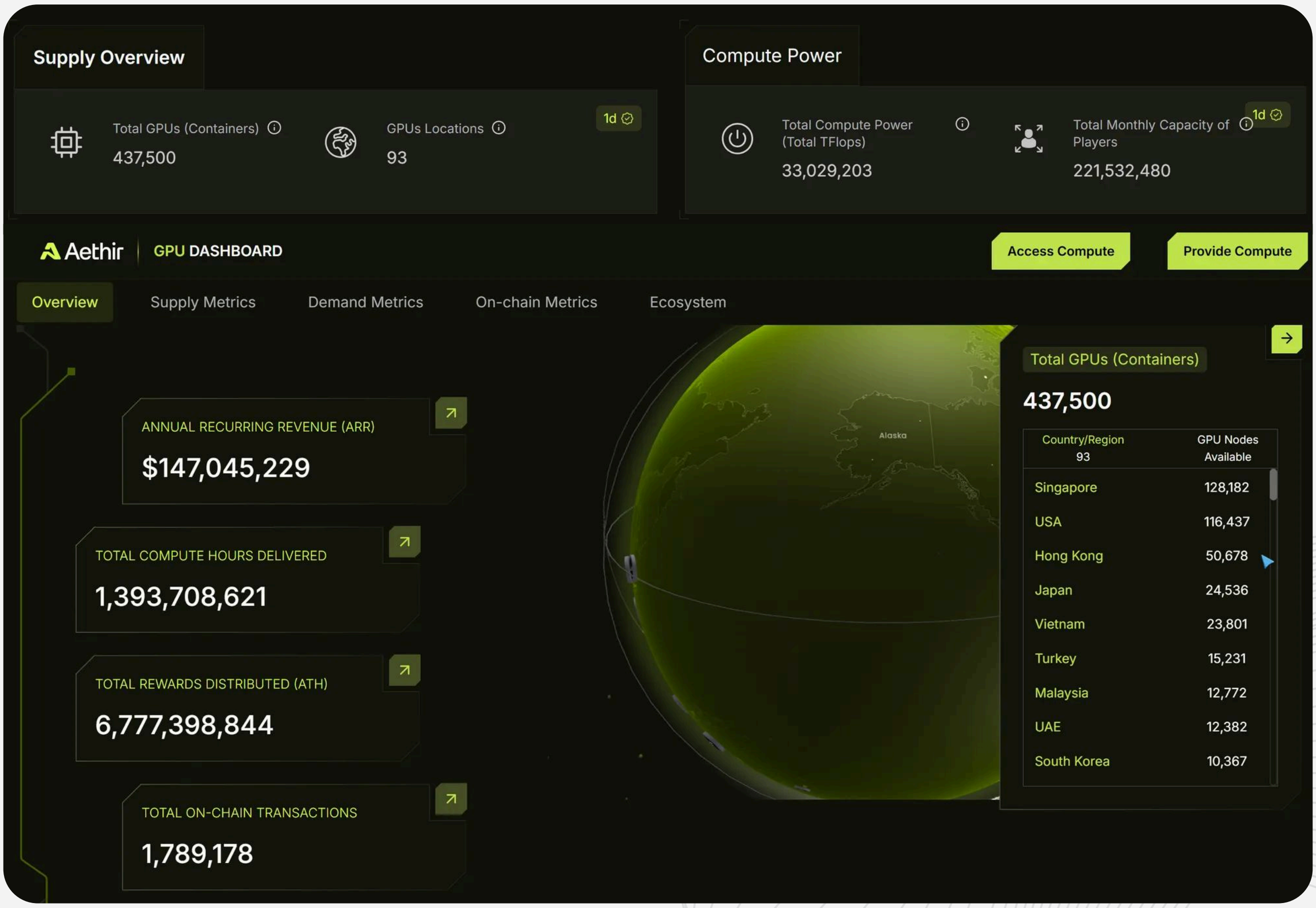
PROJECTS

AETHIR

Decentralized Enterprise GPU Cloud

Aethir is an enterprise-focused distributed GPU cloud, aiming to provide globally distributed, low-latency compute infrastructure for AI, gaming, and metaverse applications

It operates a **network of high-end GPU nodes (NVIDIA GB200, H100, H200, etc.) across 90+ countries** via partnerships with data centers and telecom edges. Aethir's platform allows resource providers to contribute hardware and users to deploy containerized workloads on-demand through a portal



AETHIR Decentralized Enterprise GPU Cloud

> *“We're past the point of asking whether decentralized infrastructure can work at scale — it already does”*

The real question now is how quickly enterprises will recognize that the centralized cloud model they've relied on is fundamentally hitting a ceiling. **GPU scarcity isn't a temporary supply chain issue; it's structural.** Costs keep climbing, and when you're competing for the same pool of H100s as every other AI company, you're not just paying a premium — you're accepting that your infrastructure roadmap is dictated by someone else's capacity planning.

What we're seeing with DePIN isn't just about cost arbitrage, though the 80–90% savings certainly get attention. It's about **rebuilding the economics of compute from the ground up.**

At Aethir:

- We've distributed **over 435,000 enterprise-grade GPUs**
- Across **93 countries**
- While maintaining **99.99% uptime**

That's not a pilot program — that's production infrastructure serving real enterprise workloads. The companies working with us aren't taking a bet on decentralization as a concept; they're solving immediate problems around:

- Access to compute
- Cost predictability
- Geographic distribution

— issues that centralized providers simply can't address at scale.

> *“The narrative around trust and reliability in DePIN needs to catch up with reality”*

We're not asking enterprises to compromise on performance or take on execution risk. We're offering:

- Verifiable **SLAs**
- Transparent **on-chain data**
- A **network architecture** that's proven itself under real-world conditions

The infrastructure revolution isn't coming — it's already here, and the gap between early adopters and everyone else is widening fast

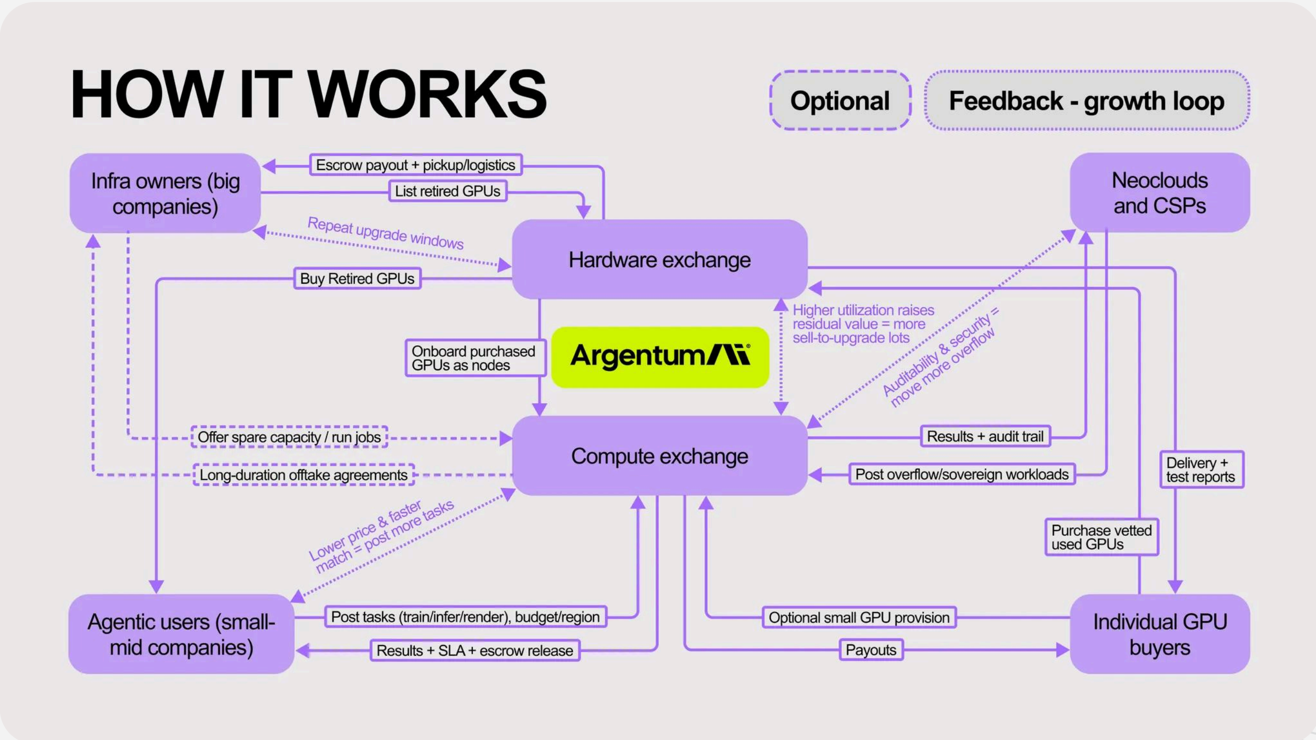
Argentum AI

GPU & HW Marketplace

Decentralized compute & hardware marketplace with dynamic auctions for enterprises to bid on and secure access to satisfy compute needs worldwide

Unique Features

- Enterprise **access to data centers worldwide that meets security and compliance needs** while scaling globally
- Lets small teams use top-tier compute without large upfront capex
- **Extend second and third generation GPUs** by building long-term demand across applications and need statements



Frigg

ArgentumAI aiming at the right hole second life for enterprise gpus = big economic opportunity and the missing compliance + transparency layer for depin = exactly needed

how much they actually deliver is pure execution i bucket ArgentumAI as "thesis very strong want to see metrics over the next 1-2 years"

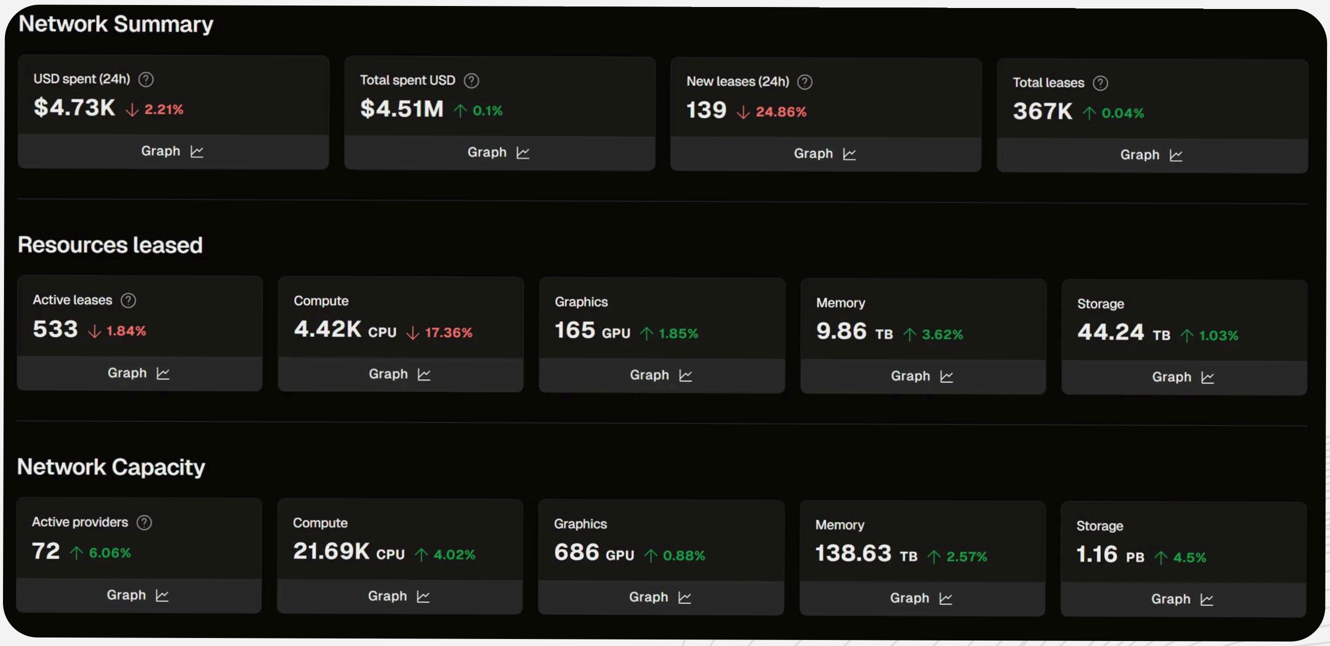
AKASH NETWORK

General Compute

Akash is a decentralized cloud compute marketplace built on the Cosmos blockchain. It allows anyone to rent out computing resources via a permissionless, on-chain bidding system

Users (tenants) submit deployment requests (containerized applications defining CPU, GPU, memory needs, etc.), which providers bid to fulfill. The winning bid is recorded as a lease smart contract, and the workload runs on the provider's infrastructure in a container

Unique Features: reverse auction mechanism for pricing, Kubernetes/Docker integration, plans for trusted execution support (SGX/TPM attestation)

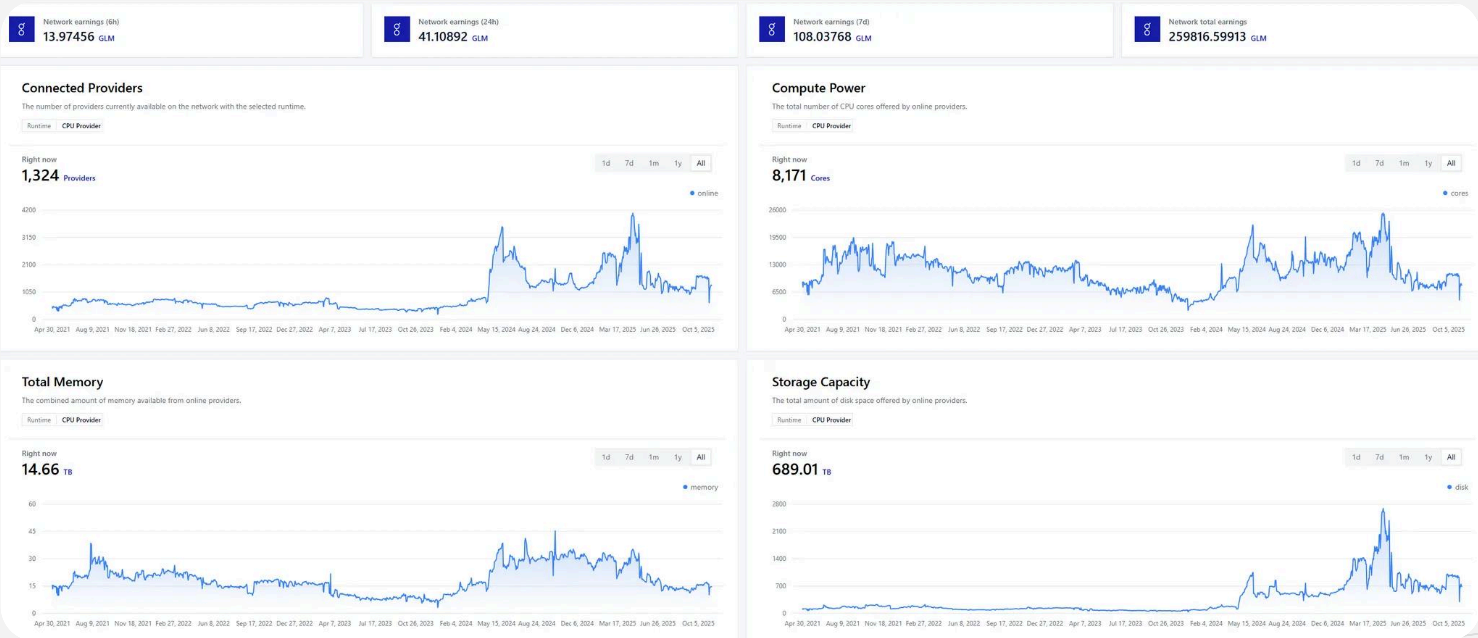


GOLEM NETWORK

P2P Compute

Golem is one of the earliest decentralized compute marketplaces (originating in 2016–2017). It enables users to buy and sell CPU/GPU time in a peer-to-peer network. Golem's vision is often summarized as an Airbnb for computers, leveraging the world's unused processing power for tasks like rendering, scientific computing, machine learning, and more.

Unique Features: task verification via checksums for deterministic tasks, redundant computing for non-deterministic tasks, reputation-based matching, application-specific approach with templates



IO.NET

AI/ML GPU Cloud

io.net is a decentralized GPU network designed for scalable, cost-efficient AI computing. Originally an internal solution for quant finance computing, it pivoted to address the global GPU shortage by **aggregating underutilized compute power from independent data centers, miners, and consumer GPUs.** io.net's platform functions as a cloud service where developers can deploy machine learning workloads across a distributed pool of GPUs via a unified API.

Unique Features: abstracts underlying nodes into single virtual cluster, advanced orchestration with distributed training support, built-in fault tolerance, custom data loading and synchronization tooling

NOSANA

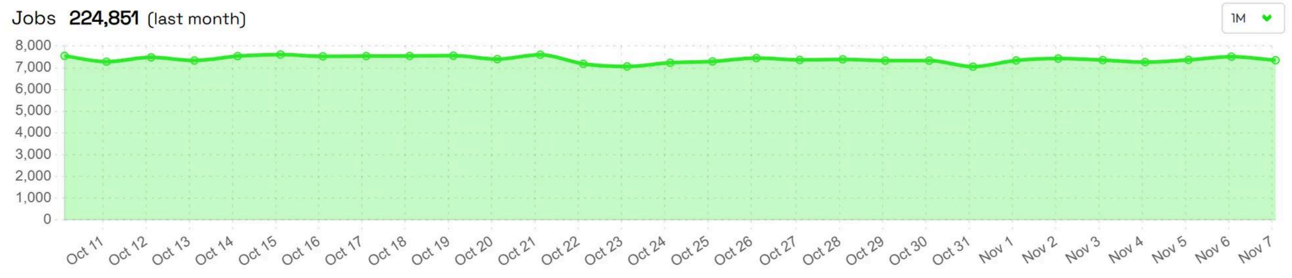
AI Inference GPU Network

Nosana is a decentralized GPU compute network focusing on making AI inference affordable and accessible. It is built on Solana and targets businesses and developers who need GPU power for AI tasks, particularly model inference and certain CI/CD workloads. Nosana allows anyone with spare GPU capacity to become a host and list their resources on an open marketplace.

Unique Features: easy DevOps integration for CI/CD pipelines, Confidential Jobs with encrypted data transfer, leverages idle gaming PCs and non-datacenter GPUs, low entry barrier

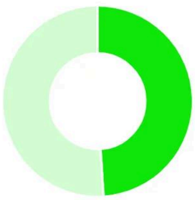
Statistics

Completed jobs	Running jobs	Queued jobs	Total job hours	Online GPU Hosts
2,695,209	805	4	2,631,928	1,608



GPUs

803 / 1638
hosts available



PREMIUM | COMMUNITY | OTHERS

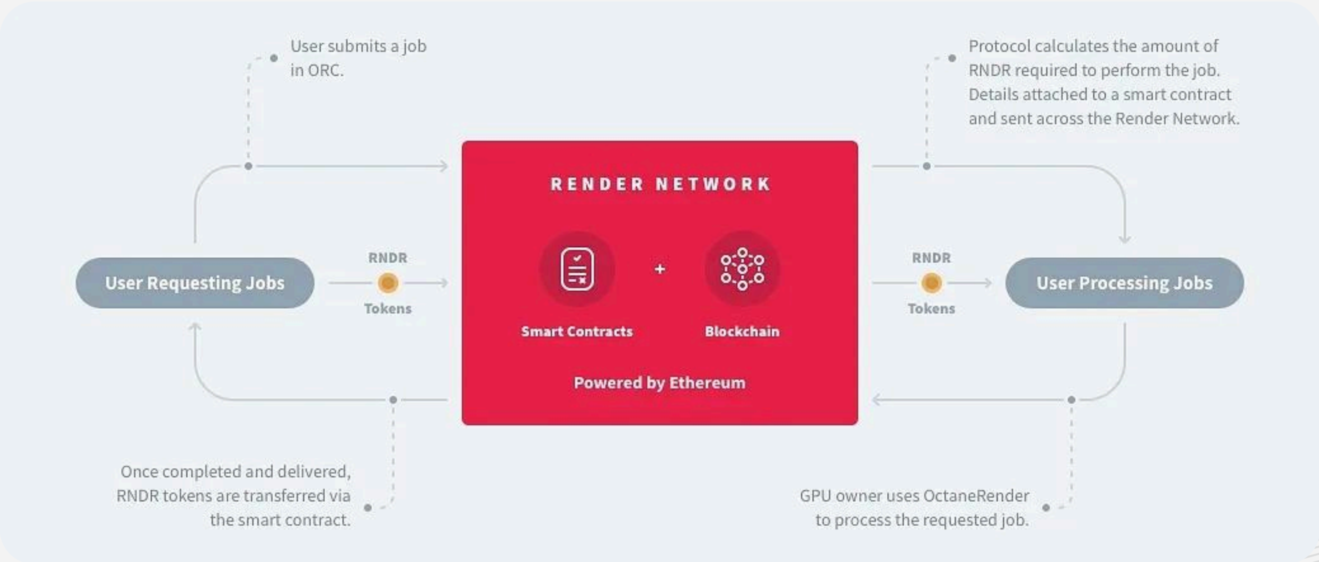
GPU	Price	Availability
NVIDIA 3060	$0.048/h	19 / 31 hosts
NVIDIA 4060	$0.064/h	12 / 17 hosts
NVIDIA 3070	$0.080/h	23 / 39 hosts
NVIDIA 3080	$0.096/h	46 / 68 hosts
NVIDIA 4070	$0.096/h	20 / 32 hosts
NVIDIA 4000/A4000	$0.128/h	3 / 6 hosts
NVIDIA 4080	$0.160/h	8 / 14 hosts
NVIDIA 3090	$0.192/h	110 / 140 hosts
NVIDIA 5070	$0.240/h	66 / 71 hosts

RENDER

Decentralized Rendering

The Render Network is a decentralized GPU rendering platform that connects artists and content creators with providers of GPU power to render digital content. Founded in 2017 by OTOY, it enables creators to submit rendering jobs via OctaneRender and pay in RNDR tokens, while node operators receive RNDR for completing work.

Unique Features: OctaneBench scoring determines node work allocation, parallel task distribution across nodes, token burning with each job, professional VFX integration (Marvel movies)



ARCHITECTURE

Compute DePIN architecture elements

- **a marketplace or scheduling layer (often on-chain) where jobs are posted and matched to nodes**

on-chain reverse auction: a client posts a deployment request (with desired resources), providers bid, and the lowest bid within a time wins, resulting in an on-chain lease contract

has a similar on-chain marketplace but with more of a focus on enterprise integration, and introduces "checker nodes" that monitor performance off-chain (details proprietary)

aggregates supply off-chain but settles usage fees on-chain (Solana), with its own scheduling software handling the matching

- **an off-chain execution layer (nodes running the computations in containers or VMs)**

leverages Docker containers orchestrated by the provider's Kubernetes cluster – the provider essentially runs a Kubernetes node that accepts the Akash deployment if they win the bid

introduced its own sandbox execution environment (was VM-based in early versions, now container-based in Golem Unlimited and New Golem) to run arbitrary code from requesters in a secure manner across providers

have their own container orchestration tailored to GPU tasks (possibly with support for frameworks like PyTorch, etc., out of the box).

- **a verification or settlement mechanism to confirm results and release payment**

has leases that define a rate (e.g. X AKT per block) and the user funds a wallet that the provider draws from per block of execution. If funds run out, the container stops

requires users to lock RNDR tokens which are released upon job completion to the node operators

(centralized but crypto-friendly marketplace) dynamically prices per second and allows payment via traditional means – decentralized counterparts may move in that direction as well for ease of use

SUPPORTED GPUS & HOSTING

Platform	Supported GPUs	Hosting	Sources
Akash	Documented support for NVIDIA H100, H200, A100	Independent providers (data centers/hosts), Kubernetes containers	*Akash Network*
Aethir	NVIDIA H100, H200, GB200	200+ locations / 93–95 countries (global DC + edge footprint)	*Aethir H1 2025 Report*
io.net	Broad NVIDIA lineup incl. V100/A100/H100 and consumer RTX	Community-contributed DC + edge nodes aggregated into a single GPU pool	*Io.net guides*
Nosana	Multiple NVIDIA markets (presumably RTX 30-series, 40-series, etc.)	Hosts list GPUs on an open marketplace for AI inference/ compute	*Nosana documentation*
Vast.ai*	Wide range of NVIDIA (inc. consumer-grade)	40+ certified data centers via its marketplace. (centralized marketplace)	*Vast's site*
Render	GPUs from RTX 3050 to RTX 5090	Globally distributed node operators (artists/studios/ partners)	*RNKB FAQ*

**(Vast.ai is a centralized service but included for context as a major GPU marketplace)*

Aethir and Akash focus on data-center grade GPUs, while io.net and Nosana accept more heterogeneous hardware (including miners' GPUs or gaming rigs). Render started in the creative community so its supply is broad (workstation GPUs)

BENCHMARKS AND SLA ANALYSIS

Provider	Region (example)	p95 Latency	Cost / 1M tokens (LLM)	Uptime
Fireworks AI	us-east	Fireworks listed among top throughput options. LLM: Groq is 0.26s avg for Llama-3.370B (3rd-party test). [source]	Llama-3 70B listed at $0.90/M tokens	99.97 - 99.99%
Together AI	us-west	Together AI appears among low-latency Llama-3.3-70B providers (no p95 published)	Model-specific $/M shown in pricing (varies by model/version) Pricing	100% recent uptime
Deep Infra	us-east	N/A (no public p95)	Llama-3.3-70B-Instruct-Turbo $0.13/M input / $0.38/M output (current), with earlier reports citing $0.23/$0.40 for Llama-3.x-70B. [source]	N/A
Akash	Varies by provider	N/A (depends on provider)	Akash sells GPU hours, not pertoken inference. (e.g., H100 SXM5 $1.18-$2.53/hr, H200 $1.95- $3.35/hr)	Marketplace lists hourly GPU

Annexe:

- **Latency p95:** vendors rarely publish p95, though I used third-party evaluations where available (e.g., Vellum's Llama-3.3-70B test for relative latency/throughput across providers)
- **Uptime and success rate:** for uptime were used official status pages (Fireworks, Together); for decentralized networks, rely on the job logs and, where available, case studies

ECONOMICS

GPU-hour price (USD)

Name	Type	Price ($)	Source
Google Vertex AI A3	NVIDIA H100	**11.27/GPU-hr** (regiondependent)	*Cloud Google, section Accelerators*
Lambda	NVIDIA H100-class	**2.49/hr**	*Lambda.ai*
Vast.ai	NVIDIA H100 PCIE	**1.2/hr**	*Pricing Vast.ai*
Akash	NVIDIA H100 80GB	Average **1.08/hr**	*Akash pricing*
Aethir	NVIDIA H100	**1.25/hr**	*Aethir pricing*
io.net	NVIDIA H100 PCIe	**1.7/hr** (Container as a Service)	*Io.net*

These **numbers show 5-10× cheaper rates than top cloud providers**. It should be noted that clouds offer other services (managed storage, networking, support) and often the decentralized prices don't include data egress or ancillary costs (which can be higher if transferring large training sets, for example). Nonetheless, the raw compute cost advantage is clear

Convert to tokens:

$$\text{Cost}_{(token)} = \frac{USD/hr}{Token\ USD\ Price}$$

For example, at $1.25/hr (Aethir) and token price $0.50, cost = 2.5 tokens/hr.

Frigg

on raw compute price DePIN can really hit up to 5-10x advantage

once you include full ops costs in some enterprise scenarios that shrinks to ~2-3x

still in AI even 2-3x savings is huge

ECONOMICS

Utilization, Gross Margin

Gross Margin

$$GM = \frac{P_{(hr)} - \left(PUE * kW * p_{(kWh)}\right) - marketplace\ fee - var.OPEX/hr}{P_{(hr)}}$$

Key inputs:

- PUE (Power Usage Effectiveness): Uptime Institute 2024 survey reports average ~1.54; best-in-class (e.g., Facebook Prineville) near 1.15.
- GPU power draw (H100): ~700 W typical.
- Electricity tariffs (Aug 2025): US commercial ~13-14¢/kWh (EIA); EU non-household ~€19.0/100 kWh (Eurostat).
- Marketplace fees: 5-10% typical across DePIN platforms

Example calculation (US, Akash H100 at $1.08/hr, PUE 1.4, 13¢/kWh):

$$\text{Energy cost/hr} = 1.4 \times 0.7\,\text{kW} \times \$0.13 = \$0.127$$

$$\text{Marketplace fee} = 0.08 \times \$1.08 = \$0.086$$

$$\text{Var. OpEx} = \$0.05\ (\text{estimated})$$

$$GM = \frac{\$1.08 - \$0.127 - \$0.086 - \$0.05}{\$1.08} \approx 69\%$$

This margin supports sustainable operations if utilization remains >50% and no hardware depreciation is incurred. In practice, utilization variance (demand spikes, off-hours) is the dominant swing factor for unit profitability

PUE plays the role of the KPI of the given data center, as it represents the ratio of the total consumed energy by the data center to the energy spent directly for the IT infrastructure

Therefore, **a lower PUE (i.e., a more efficient data center) reduces energy/hr in the formula and increases gross margin**

ECONOMICS

Payback, Fee Coverage Ratio

Payback period (months)

$$Payback = \frac{CapEx/GPU}{(P_{(hr)} - VC/hr) * u * 720}$$

Where:

- CapEx/GPU: H100 acquisition cost (~$15k–$20k used)
- VC - variable costs per hour
- u – utilization (realized busy time ÷ available time for a GPU). DePIN GPU supply is elastic (adding nodes is fast compared with hyperscalers' long build cycles), but demand comes in spikes, so realized utilization is the swing factor for unit returns.
- 720 - hours per 30-day month

At $15k CapEx, $1.08 revenue, $0.26 variable cost, 60% utilization:

$$\text{Payback} = \frac{\$15000}{(\$1.08 - \$0.26) \times 0.6 \times 720} \approx 52 \text{ months}$$

Fee Coverage Ratio (protocol sustainability)

$$\text{Fee Coverage Ratio} = \frac{Protocol\ fees\ (same\ period)}{Token\ emissions\ (same\ period)}$$

Akash proxy: Messari reports lease revenue of $1.0m (Q1-25) and $0.82m (Q2-25) as activity consolidated into higher-value deployments [source]

In DePIN, utilization drives return more than power line items at current price levels

Providers that win longer-lived, batchable workloads and operate at reasonable PUE with transparent verification/SLAs are positioned to sustain margins and compress payback even when headline $/GPU-hr is competitive [source]

TOKENOMICS

What to measure

- **On-chain transaction volume (compute)** - value of lease settlements / job payments on the protocol over period *t*
- **Active payers (AP)** - count of unique tenant addresses that opened/paid at least one compute lease in period *t*
- **Fees vs emissions (FCR)** - Treasury runway (months):

$$TR = \frac{Treasury\ (USD\ or\ token)}{Average\ quarterly\ grants\ +\ OpEx} * 3$$

Provider	On-chain demand proxy	Fees	Emissions	FCR	Treasury / runway
Akash	New & active leases (stats.akash), lease revenue (Messari)	AKT-denom. fee revenue (Messari)	Inflation 4-8%	2.9%	Community pool 7.17m
Render	Burned tokens/job spend (BME); Render Credits conversions	Fees effectively > burn(<= 95%)	Emissions per RENDER	11%	Treasury not primary; burn is the sustainability signal
io.net	On-chain: rewards distribution Off-chain: buyer invoices/contracts	Limited public fee capture	Disinflation to 800M cap: 95% to GPU/5% to CPU rewards	16.1%	Treasury/ runway not public
Aethir	$20.5M Q1 2025; $127M ARR	Transaction based	~259% initial, tapering	17.7%	$344M DATC: Predictive Oncology (NASDAQ: POAI)

Aethir offers the most complete on-chain dataset today for a token-economics analysis with their dashboard [source], Render's burns are a direct proxy for paid usage under BME, and io.net's known emissions + reward split, so we sustainability using emissions until full fee data is available

The $140B Compute Arbitrage

AI compute costs are skyrocketing. Fine-tuning a model on AWS can cost $11.27/hour for an H100 GPU, while Aethir offers the same job for $1.25/hour—saving a single team $8K/month. Scaled across AI's $80–150B market by 2027, this is the biggest infrastructure arbitrage since the move to cloud.

Enterprises hesitate due to perceived risks: decentralized networks, crypto tokens, and regulatory concerns. DePIN is overhyped yet underutilized, lacking institutional infrastructure

Proven Use Cases:

- Aethir: Regional gaming compute, 65% cheaper and faster than AWS
- Akash: AI lab model evaluation, 59% cost reduction
- Render: VFX studios, 60–70% cheaper rendering
- io.net: Startup access to 40x H100 GPUs in 48 hours, 70% cheaper
- Nosana: DeFi CI/CD, 80% cost savings

ArgentumAI Institutional DePIN

- Focuses on idle enterprise GPUs (V100s, P100s, RTX A-series) as revenue assets
- Compliance-first: DAO LLC, KYC/AML, regional pinning, SLA escrow
- Targets CIOs, labs, and regulated enterprises, not just crypto-native devs

Why It Matters:

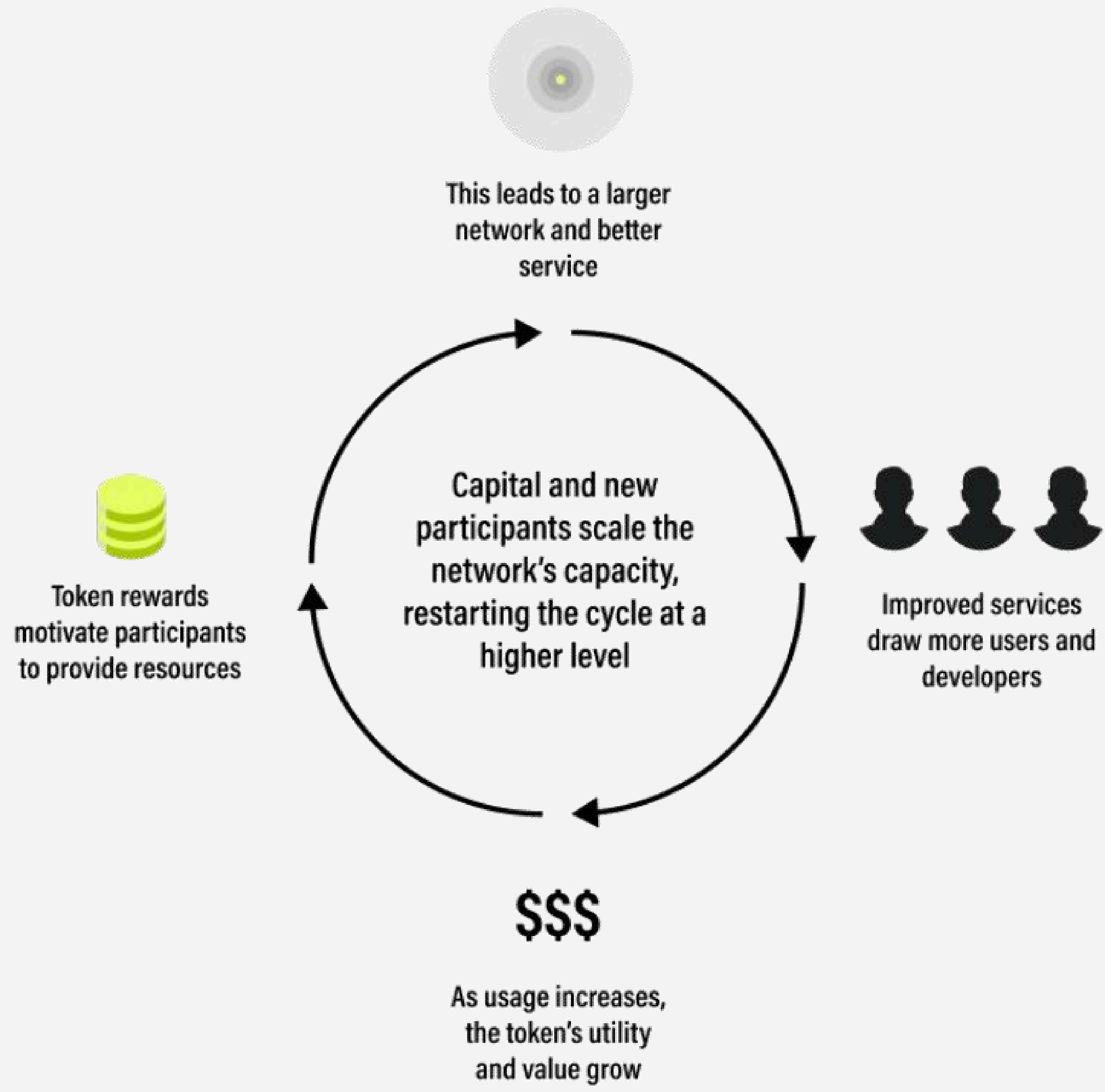
- Compliance-first networks can unlock institutional spend and idle hardware.
- DePIN adoption is already happening at scale, but execution risk remains: liquidity, competitive response, and enterprise onboarding.

DePIN is no longer theoretical—enterprise-grade providers like ArgentumAI prove the model works for regulated workloads, offering massive cost arbitrage while bridging compliance gaps.

THE DEPIN FLYWHEEL

The DePIN Flywheel is a positive feedback loop that aligns incentives with sustained network growth:

- Token rewards attract infrastructure providers
- Supply-side growth improves services
- Improved services attract builders and users
- Increased usage leads to value growth and improved utility



GOVERNANCE AND REGULATION

Geographical placement of nodes

Routing workloads across borders can trigger cross-border transfer rules (GDPR Chapter V), sanctions/KYC screening (OFAC), and contract duties (GDPR Art. 28; CCPA service-provider/contractor terms). Controllers must ensure processors (node operators) follow documented locations and transfer mechanism

U.S. export controls on advanced AI chips

The **U.S. Bureau of Industry and Security (BIS) controls advanced computing ICs (ECCN 3A090) and certain systems (4A090)** first issued Oct-2022 and tightened in Oct-2023 via interim final rules

These laws restrict exports to China/Macau and impose end-user controls. [Federal Register]

BIS issued additional clarifications and rulemakings in 2024 and 2025 expanding the FDP rule scope: they updated due-diligence expectations and trade analyses (i.e. introducing restrictions tied to AI model weights in certain contexts)

DePIN teams brokering GPUs must screen destinations, counterparties, and cluster composition accordingly

GDPR (EU)

GDPR Chapter V requires an adequacy decision (e.g., EU–US Data Privacy Framework) **or appropriate safeguards** (e.g., Standard Contractual Clauses) before personal data leaves the EEA.

Article 28 (Chapter IV) mandates a DPA with processors/sub-processors (node operators), with documented processing instructions, location, and security measures (records of processing and “data protection by design” also apply) [GDRP Chapter IV].

What this means for DePIN: enforce region pinning at the scheduler, use SCCs (or DPF-certified U.S.

recipients), and maintain a registry of approved jurisdictions/providers for each workload class.

COMPLIANCE READINESS MATRIX

Status Legend

● ✓ Live ● → Roadmap ● ✗ N/A ● ? Unknown

PROTOCOL	EU AI ACT	GDPR/DPA	REGIONAL PINNING	EXPORT CONTROLS (BIS)	SLA DASHBOARDS	TEE/ATTESTATION
Aethir	● → Roadmap	● → Limited	● ✓ Live	● ? Unknown	● ✓ Live	● → Roadmap
Akash	● → Roadmap	● → Roadmap	● → Limited	● ✗ N/A	● → Roadmap	● → Roadmap
io.net	● ✗ N/A	● → Limited	● → Limited	● ✗ N/A	● → Limited	● ✗ N/A
Render	● ✗ N/A	● → Limited	● ✗ N/A	● ✗ N/A	● → Limited	● ✗ N/A
Nosana	● ✗ N/A	● ✗ N/A	● ✗ N/A	● ✗ N/A	● ✗ N/A	● ✗ N/A
ArgentumAI	● → Roadmap	● → Roadmap	● → Roadmap	● → Roadmap	● → Roadmap	● → Roadmap

EU AI Act: Compliance with European Union regulations on AI system transparency and accountability

GDPR/DPA: Data protection agreements and privacy compliance for EU citizen data

Regional Pinning: Ability to restrict data/compute to specific geographic regions

Export Controls (BIS): US Bureau of Industry and Security compliance for GPU/technology export

SLA Dashboards: Published service level agreements with monitoring and reporting

TEE/Attestation: Hardware-level security with cryptographic proof of execution environment

OVERALL COMPARISON MATRIX

Protocol	Focus	Access	Second Life HW	Orchestration/ Verification	SLA/Attestation Evidence	Pricing/ Allocation	Notes
	Decentralized GPU cloud (AI/gaming)	Enterprise-grade GPUs	–	Checker nodes; SLO guide	No public network-wide GPU-TEE attestation	Marketplace rentals	*High enterprise focus; telemetry transparency improving*
	Decentralized cloud (general compute)	Global SMEs	–	On-chain leases; roadmap to NVTrust/TDX/ SEV-SNP	Roadmap published; provider-level evidence varies	Provider bid/ deploy	*Transparent pricing; QoS varies by provider*
	GPU DePIN clusters(Solana)	B2C+B2B crypto-native	–	Central scheduler; emissions split known	No public network-wide SLA dashboard	On-chain cluster tasks	*Large pool; heterogeneous performance*
	GPU rendering → AI	B2C creators	–	BME burns; OctaneBench tiers	No public GPU-TEE docs for AI	Per-minute usage	*Strong in rendering; AI expanding*
	AI inference	Global SMEs	–	Confidential Jobs (private transfer/ logging)	No hardware TEE docs	Per data processed	*Early B2B stage*
	General compute (CPU-lean)	B2C creators	–	Task receipts / marketplace	N/A	N/A	*Legacy network with active ecosystem*
Argentum AI	Marketplace for GPUcompute & hardware	Global B2B+B2C+Institutons	Yes	Planned contracts + benchmarking	Yes (TEE/ ZK,SLA+escrow)	Dynamic auctions +SLAS	*High institutional focus; early; watch for SLA dashboards & audits*

COMPARATIVE SCORING MODEL

In order to create a framework for comparing DePIN networks across economic and operational dimensions we took 6 main criteria.

Each of them is rated between 0 and 20, multiplied by a weight, and rolled up into a composite score (0-20)

Criteria	Type	Weigh	Source
Performance & SLA	p95 latency, uptime	25%	0 = unreliable (<90 % success); 12 = 99 % uptime; 20 = ≥ 99.9 % uptime + p95 latency ≤ 150 ms
Cost Advantage	Relative $/GPU-hr vs AWS at same SLA	20%	0 = more expensive; 12 = 10–20 % cheaper; 20 = more than 35 % cheaper with stable SLA
Supply	NVIDIA H100/A100 share, PUE of data centers	10%	0 = mostly consumer GPUs; 12 = mixed; 20 = enterprise-grade GPUs with attested nodes
Demand	Real usage vs incentives	10%	0 = mostly airdrops; 12 = some B2B pilots; 20 = more than 50 % revenue from nonincentivized clients
Compliance	Data residency, export controls, legal structure	20%	0 = non-compliant; 12 = partial policies; 20 = full compliance and legal entity clarity
Token Sustainability	Fee coverage ratio, emissions schedule, treasury runway	15%	0 = fees < 10 % of emissions; 12=20-40%; 20 = > 80 % fee coverage + deflationary model

Total Weight = 100%

Fee Coverage Ratio (protocol sustainability)

$$Total\ Score = \sum_{i=1}^{6} (S_i * W_i)$$

Range	Category	Interpretation
15.5-20	Leader	High SLA, strong token sustainability
12-15.5	Watch	Promising, but requires improvement in either compliance, or demand
<12	Risky	Weak SLA, instable

COMPARATIVE SCORING

Protocol	SLA	Cost	Supply	Demand	Law	Ops	Total	Rank	Source
	16	18	18	14	13	16	**15.8**	**Leader**	Pricing calculator shows H100 example at $1.25/hr and “up to 90%” savings (demo calc)
	14	17	14	16	12	14	**14.4**	**Watch**	Official claim of 30k+ GPUs and up to 70% lower cost blog notes $20M annualized onchain revenue
ArgentumAI	14	16	16	12	14	14	**14.4**	**Watch (Early)**	Deck, whitepaper, roadmap, traction to date
	12	16	13	14	14	13	**13.65**	**Watch**	Public GPU pricing shows H200 at $1.40/hr vs hyperscalers (large discount); attestation/TEE work formalized via roadmap
	12	12	15	14	10	12	**12.1**	**Watch**	Network is primarily GPU rendering today. Pricing model tied to OctaneBench tiers
	10	15	11	11	10	11	**11.35**	**Risky**	Focus on AI inference with “up to 6× cheaper” claim docs show Confidential Jobs (private transfer/logging) but not hardware TEEs

**ArgentumAI* score reflects early-stage status projected based on the current potential and execution; execution risk is higher given limited live pilots (as of Nov 2025)*

NOTES TO THE TABLE

PRICING DISCLOSURE

Aethir *H100 demo ≈ $1.25/hr;* ***Akash*** *H200 ≈ $1.40/hr vs hyperscalers).* ***io.net*** *and* ***Nosana*** *cite %-savings without full rate cards*

SUPPLY MIX

Aethir *focuses on enterprise-grade GPUs (GB200/H200/H100/B200).* ***io.net*** *aggregates heterogeneous sources, including gaming/miner rigs.* ***Render****'s installed base is rendering-centric*

ARGENTUMAI (WATCH/EARLY)

Early-stage; active institutional outreach. Focus: reuse of second-gen/idle GPUs, multi-region diversification, security. SLA dashboard pending; no public rate card

SLA VISIBILITY

Aethir *provides an SLO/Service Guide.* ***Akash, Render, io.net****, and* ***Nosana*** *lack network-wide audited SLA dashboards; until public p95 latency/uptime is posted, scores stay mid or below*

DEMAND SIGNALS

io.net *reports ≈ $20M annualized on-chain revenue and 30k+ GPUs.* ***Aethir l****ists recent case studies.* ***Render*** *has strong creator demand.* ***Akash*** *shows steady adoption via pricing/availability pages*

DATA QUALITY & SCORING

Where only marketing claims exist and audited telemetry is absent (public p95 latency/uptime, TEE-attested logs), scoring is conservative.

REKTNOMIST

> ***I think we're watching the early formation of a completely new layer of cloud infrastructure***

You can see how networks like **Aethir** and **io.net** are starting to unlock real efficiency, running H100 GPUs at a fraction of what hyperscalers charge, but at the same time, I believe the price advantage isn't the full story

> ***What really determines whether this sector grows beyond early crypto users is trust***

Most of these networks still can't show verifiable SLAs or audited uptime, and that's a big deal if you're an enterprise dealing with compliance-heavy workloads. In my view, **until we get consistent attestation data, transparent job receipts, and region pinning built into the orchestration layer, the institutional market will hesitate**. I think the opportunity is massive though, especially as regulations like the EU AI Act and US export controls force infrastructure providers to rethink where and how they process data

The funny thing is that **what started as a *"cheap compute"* experiment is now turning into a kind of *global GPU clearinghouse***. Second-generation cards, idle enterprise hardware, even miner GPUs are all finding new life in these decentralized markets.

And from what I see, this repurposing loop is what gives DePIN real staying power. I'd argue we're still in the reliability and compliance phase, not the adoption phase yet, but if the networks can close that verification gap, I genuinely think DePIN will become a core part of the global compute economy in the next few years

MARKET POTENTIAL & GAPS

DEMAND DRIVERS

Workloads best suited to DePIN today are batchable: fine-tuning, offline inference, rendering, and synthetic data generation. Agent pipelines and evaluation jobs with relaxed tail-latency also fit. Unit economics hinge on achieved utilization (GPU busy hours) and the realized $/GPU-hr after egress and storage.

PERFORMANCE CONSTRAINTS

Wide-area bandwidth and multi-cluster orchestration limit distributed training scale; underutilization rises when cross-region links are slow. Schedulers need topology awareness, retry logic, and job pre-emption to keep GPUs saturated.

COMPLIANCE & PROCUREMENT

Institutions require DPAs/SCCs/UK IDTA where PII is processed, export-control screening for advanced GPUs, and regional pinning with auditable provider registries. Billing must be predictable (reserved + spot) with invoices and receipts.

VERIFICATION & SLAs

Required controls include audited uptime/latency dashboards, TEEs (SEV-SNP/TDX/NVTrust) with attestation evidence, and reproducible benchmarks for specific GPU tiers. Job receipts and immutable logs should be standard.

SUPPLY

Large pools of idle or second-generation GPUs exist in DCs and enterprise labs. Repurposing economics are favorable where power, PUE, and cooling are adequate; consumer rigs display higher failure variance. Failure domains should be isolated at scheduler level.

INSTITUTIONAL CHECKLIST

What enterprises still require: (1) verifiable SLAs and attestation coverage, (2) predictable billing and enterprise contracts, (3) regional compliance toolkits and provider registries, (4) security reviews and pen-test summaries with remediation timelines.

ArgentumAI's Potential

ArgentumAI appeared in the info space as a strong early-stage differentiated player addressing current DePIN gaps: *if it successfully implements compliance features and showcases reliable performance with older institutional grade hardware, it could attract users who need cheaper compute but held back by lack of trust in other networks*

Given it's early, the next milestones would be launching a public dashboard of performance (to build credibility) and onboarding a mass of institutional GPU providers across regions

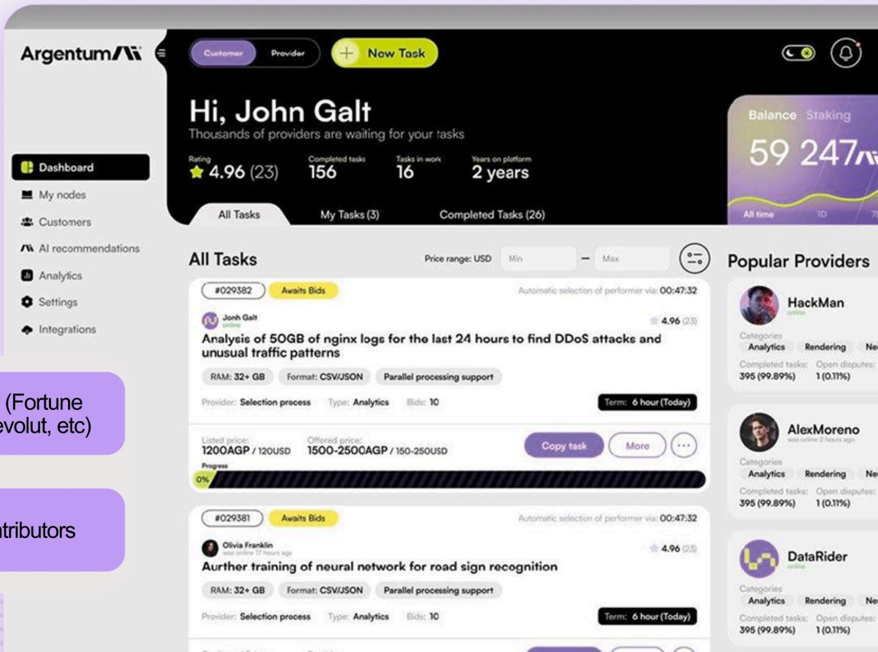
Strengths:

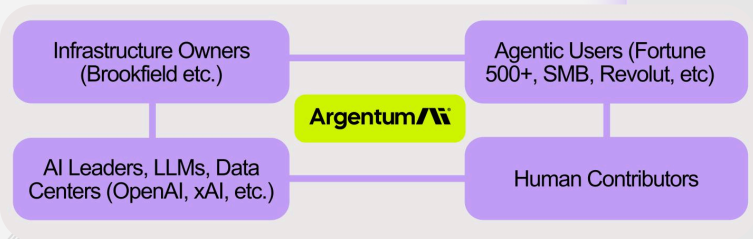
- AI-augmented marketplace design (real-time benchmarking) addresses limitations around static pricing
- Second-life GPU focus taps underserved institutional hardware monetization channel
- Compliance roadmap (DAO LLC, KYC/AML, data residency controls) aligns with institutional requirements faster than peers
- Argentum has a strong team, advisors, and network needed to launch and execute an institutional-level project amid active Web3 adoption

Risks:

- Execution complexity: Multi-layer marketplace (hardware + compute) requires operational sophistication competitors have not yet mastered.
- Competitive response: If incumbents (Aethir, io.net) improve SLA transparency and compliance faster, ArgentumAI's differentiation narrows.

If those indicators are positive, ArgentumAI can become a worthy player that complements existing networks by adding a layer of trust and inclusivity of hardware that others leave idle





CONCLUSIONS

		Argentum AI			
15.8	14.4	14.4	13.65	12.1	11.35
Leader	Watch	Watch (Early)	Watch	Watch	Risky

- Growth is real but concentrated; fee ramps are visible in 2025 yet uneven across protocols; YoY fees did ~5x
- Best near-term fit: batchable AI jobs; interactive low-latency workloads still favor centralized clouds
- Networks with transparent leases/usage (e.g., Akash; Render burns as proxy) are easier to diligence
- Attestation and SLA dashboards are the primary adoption bottlenecks for regulated buyers
- Regional pinning and export-control screening must be integral to orchestration and contracts.
- ArgentumAI: early, but targets two gaps - reuse of second-gen GPUs and security/region diversification for institutions. Monitor for public SLA dashboards, attestation coverage, and audited benchmarks

Frigg

DePIN compute can move from side role to a core infra role in the AI boom but not just by being cheap it needs auditability and compliance

AUTHORS

Nick Smohorzhevskyi

Chief Investment Officer (CIO) of Solus Group & Solus Partners

Maksym Kovalenko

Analyst & Research at Solus Group & Solus Partners

Valeriia Charkovska,

Investment analyst intern at Solus Group

Aethir ArgentumAI

Frigg

KOL, Mapping DeFi, zk contributor

ALEX

KOL. Research

SALAZAR.ETH

KOL. Defi. Contributor

REKTNOMIST

KOL. Research

Want to collaborate?

Contact us

AUTHORS

CryptoGoos
KOL

RUKIZCUKIZ
KOL

DILLION_EMPIRE
KOL

KONG TRADING
KOL

ZENBIRB
KOL

LISA FLORENTINA
KOL

ELITE
KOL.

MR BRASS
KOL

Leshka.eth
KOL

CRYPTO GIDEON
KOL

Want to collaborate?

Contact us

DARC by solus group

SOLUS Group - an alliance of Web 3.0 degen-companies: experts in marketing, fundraising, and consulting for tier 1 Web 3.0 businesses. Tame the bull with Solus.

SOLUS PARTNERS — IR advisory company for Web 3.0

DARC by solus group — Degen Associate Research Club by Solus Partners

SOLUS AGENCY — Web 3.0 Marketing & Consulting Marketing Agency

SOLUS.d&a — Drops & Ambassador programs degen strategists

MERGEWAVE capital — M&A and IR advisory company for IT/Fintech

@NICK_SOLUS